TECHNISCHE UNIVERSITÄT WIEN

HABILITATION

Eigenvalue asymptotics of finite difference and differential operators

zur Erlangung der Lehrbefugnis (venia docendi)
an der Fakultät für Mathematik und Geoinformation
der Technischen Universität Wien

vorgelegt von

Dipl.-Ing. Dr.techn. Benjamin Eichinger

Wien, Juni 2024

Contents

Chapter 1

Introduction

This habilitation thesis consists of 5 papers, which can be grouped into 2 categories. The first category deals with approximation problems for rational functions and the second one is concerned with with eigenvalue asymptotics of truncations of continuum Schrödinger operators.

The papers related to approximation problems for rational functions are *1) Orthogonal rational functions with real poles, root asymptotics, and GMP matrices*; *2) Asymptotics of Chebyshev rational functions with respect to subsets of the real line* and *3) Finite-Gap CMV Matrices: Periodic Coordinates and a Magic Formula* deal with L^2- and L^∞-approximation problems for rational functions. In the L^2-setting, we restrict our attention to measures that are supported on the real line or the unit circle, so that we can also study associated self-adjoint and unitary operators, respectively. These results are presented in Section 1.1.3, 1.1.4 and 1.1.5. Section 1.1.1 is of preliminary nature and aims to provide an overview on existing results and should motivate the results presented in this thesis.

Eigenvalue asymptotics of truncations of continuum Schrödinger operators are discussed in Section 1.2. It summarizes the results from the papers *4) Stahl–Totik regularity for continuum Schrödinger operators* and *5) Asymptotics for Christoffel functions associated to continuum Schrödinger operators*. The latter paper builds on the former one, as well as the preprint *An approach to universality using Weyl m-functions*, which is not part of this habilitation thesis but for completeness reviewed in Section 1.2.2.

The papers in this habilitation thesis arose in collaboration with Jacob S. Christiansen, Milivoje Lukić, Tom Vandenboom and Giorgio Young. I am deeply grateful to all of them for inspiring discussions and collaborations.

1.1 Extremal rational functions

This section is concerned with extremal rational functions. We start with some terminology. We will mainly be interested in associated self-adjoint operators. Thus, poles and spectral sets will be subsets of $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$.

For $\mathbf{c} \in \overline{\mathbb{R}}$ we denote

$$r(z, \mathbf{c}) = \begin{cases} \frac{1}{\mathbf{c}-z}, & \mathbf{c} \neq \infty, \\ z, & \mathbf{c} = \infty. \end{cases}$$

We fix a compact proper subset $\mathsf{E} \subset \overline{\mathbb{R}}$ containing infinitely many points, and a sequence $\mathbf{C} = (\mathbf{c}_k)_{k=1}^\infty$ with $\mathbf{c}_k \in \overline{\mathbb{R}} \setminus \mathsf{E}$. The sequence $\mathbf{C}$ can have repetitions, which are used to designate multiplicity: we consider the spaces of rational functions $\mathcal{L}_n$ defined by

$$\mathcal{L}_n = \left\{ \frac{P(z)}{R_n(z)} : \; P \in \mathcal{P}_n \right\}, \tag{1.1}$$

where $\mathcal{P}_n$ denotes the set of polynomials of degree at most n and

$$R_n(z) = \prod_{\substack{1\le k\le n \\ \mathbf{c}_k\neq\infty}} (z - \mathbf{c}_k). \tag{1.2}$$

Of course, the spaces $\mathcal{L}_n$ could also be defined iteratively by

$$\mathcal{L}_n = \operatorname{span}\left\{ r(z, \mathbf{c}_n)^{d_n} \right\} \oplus \mathcal{L}_{n-1}, \quad \mathcal{L}_0 = \{1\},$$

where d_n denotes the multiplicity of the pole $\mathbf{c}_n$ up to that point,

$$d_n = \sum_{\substack{1\le k\le n \\ \mathbf{c}_k=\mathbf{c}_n}} 1.$$

The case E being a compact subset of $\mathbb{R}$ and $\mathbf{c}_k \equiv \infty$ leads to polynomials, i.e., $\mathcal{L}_n = \mathcal{P}_n$ for all n. Let $\|\cdot\|_{\mathsf{E}}$ denote the sup-norm on E and for a probability measure μ with $\operatorname{supp}\mu \subset \overline{\mathbb{R}}$, we denote by $\|\cdot\|_{L^2(\mu)}$ the corresponding $L^2(\mathbb{R}, \mu)$-norm. In this case we assume that $\mathbf{c}_k \in \overline{\mathbb{R}} \setminus \operatorname{supp}\mu$. In the context of this section we will be concerned with 4 different extremal problems:

Problem 1 (Orthogonal extremal problem)**.**

$$\kappa_n(\mathbf{c}_n) := \sup\{\operatorname{Re}\lambda \mid \exists h \in \mathcal{L}_{n-1}, \; \|\lambda r(\cdot, \mathbf{c}_n)^{d_n} + h\|_{L^2(\mu)} \le 1\}. \tag{1.3}$$

By strict convexity of the L^2-norm, the extremal problem has a unique maximizer, which will be denoted by τ_n. By the nature of the Gram–Schmidt process, it can be seen that τ_n corresponds to the orthonormal rational function for the increasing sequence of spaces $\mathcal{L}_n$. In particular, for $\mathbf{c}_k \equiv \infty$, we obtain the orthonormal polynomials associated to μ.

Problem 2 (Christoffel extremal problem)**.** *For* $x_* \in \overline{\mathbb{R}} \setminus \{\mathbf{c}_k : 1 \le k \le n\}$,

$$\kappa_n(x_*) := \sup\{\operatorname{Re} f(x_*) \mid f \in \mathcal{L}_n, \; \|f\|_{L^2(\mu)} \le 1\}.$$

We adopt the name Christoffel extremal problem, since for the case $\mathcal{L}_n = \mathcal{P}_n$, the reciprocal of $\kappa_n(x_*)$ is typically called the Christoffel-function.

The corresponding problems for the sup-norm are:

Problem 3 (Chebyshev extremal problem)**.**

$$m_n(\mathbf{c}_n) := \sup\{\operatorname{Re}\lambda \mid \exists h \in \mathcal{L}_{n-1}, \; \|\lambda r(\cdot, \mathbf{c}_n)^{d_n} + h\|_{\mathsf{E}} \le 1\}.$$

Problem 4 (Residual extremal problem)**.** *For* $x_* \in \overline{\mathbb{R}} \setminus (\mathsf{E} \cup \{\mathbf{c}_k : 1 \le k \le n\})$,

$$m_n(x_*) := \sup\{\operatorname{Re} f(x_*) \mid f \in \mathcal{L}_n, \; \|f\|_{\mathsf{E}} \le 1\}.$$

In the literature, the above stated problems are most intensively studied in the polynomial case $\mathcal{L}_n = \mathcal{P}_n$. One goal of the papers [19, 20] is to provide a unified perspective on the subject.

There is a vast literature on orthogonal and Chebyshev polynomials. Thus, the following is a very partial list of references: For a book reference on orthogonal rational functions, we refer to [4]. For Chebyshev polynomials associated to subsets of $\mathbb{R}$ we recommend the survey [45], the recent papers [7, 8] and the recent survey article [9]. For Problem 4 in the polynomial setting see [10] and also [16, 21].

1.1.1 Extremal polynomials, periodic coordinates and a magic formula

In this section, we recall certain aspects of the theory of orthogonal polynomials. We call a measure μ trivial, if it is a finite sum of dirac masses and non-trivial otherwise. Let μ be a non-trivial probability measure with compact support in $\mathbb{C}$ and denote $\mathsf{E} = \operatorname{supp}\mu$ its topological support. Orthonormal polynomials can be obtained from (1.3) by setting $\mathcal{L}_n = \mathcal{P}_n$. In this case we get

$$\kappa_n(\infty) = \sup\{\operatorname{Re}\lambda \mid P(z) = \lambda z^n + Q(z), Q \in \mathcal{P}_{n-1}, \|P\|_{L^2(\mu)} \leq 1\}.$$

The unique maximizer is the orthonormal polynomial of degree n associated to μ and will be denoted by $p_n(z)$.

Jacobi matrices and Favard's theorem

We will now restrict our attention to the case that $\operatorname{supp}\mu$ is a compact subset of $\mathbb{R}$. Then, there exist $a_n > 0, b_n \in \mathbb{R}$ such that the orthonormal polynomials satisfy

$$\begin{aligned} zp_n(z) &= a_{n+1}p_{n+1}(z) + b_np_n(z) + a_{n-1}p_{n-1}(z), \quad n \geq 1, \\ zp_0(z) &= a_1p_1(z) + b_0p_0(z). \end{aligned} \tag{1.4}$$

Let $M_z = M_{z,\mu}$ denote the multiplication operator in $L^2(\mu)$. Then (1.4) says that the matrix representation of M_z with respect to the orthonormal basis $(p_n)_{n=0}^\infty$ is of the form

$$J = J(((a_{n+1}, b_n))_{n=0}^\infty) = \begin{pmatrix} b_0 & a_1 & 0 & 0 & 0 \\ a_1 & b_1 & a_2 & 0 & 0 \\ 0 & a_2 & b_2 & a_3 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{pmatrix}. \tag{1.5}$$

The coefficients $((a_{n+1}, b_n))_{n=0}^\infty$ are called *Jacobi parameters*. Since $\operatorname{supp}\mu$ is compact, the operator M_z is bounded along with the Jacobi parameter. Thus, if (1.5) is considered as the matrix representation of J with respect to the standard basis, $(e_n)_{n=0}^\infty$, of $\ell^2(\mathbb{N}_0)$, then J defines a bounded operator on $\ell^2(\mathbb{N}_0)$ called *Jacobi matrix*.

Recall that a *Herglotz function* is an analytic map form $\mathbb{C}_+ := \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$ into $\mathbb{C}_+ \cup \mathbb{R}$. We will denote the set of all Herglotz functions by $\mathcal{N}_0$. Herglotz functions admit an integral representation. That is, for any $f \in \mathcal{N}_0$ there exists unique $a \in \mathbb{R}$, $b \geq 0$ and a positive measure ν with $\int \frac{d\nu(\xi)}{1+\xi^2} < \infty$ such that

$$f(z) = a + bz + \int \left(\frac{1}{\xi - z} - \frac{\xi}{1+\xi^2}\right) d\nu(\xi). \tag{1.6}$$

The measure ν can be recovered from f by Stieltjes inversion formula

$$\lim_{\varepsilon \to 0} \frac{1}{\pi} \int_c^d f(t + i\varepsilon)dt = \frac{1}{2}\left(\nu((c,d)) + \nu([c,d])\right), \tag{1.7}$$

for $c < d$.

Above we have described the map $\mu \mapsto J$. Let us now describe the inverse map. Let $a_n > 0, b_n \in \mathbb{R}$ be bounded sequences and J be defined by (1.5). Due to the boundedness of the coefficients, J defines a self-adjoint operator on $\ell^2(\mathbb{N}_0)$. Define the *Weyl m-function* by

$$m(z) = \langle (J - z)^{-1} e_0, e_0 \rangle. \tag{1.8}$$

The function m is a Herglotz function. Let μ be the measure in the integral representation (1.6) of m. Stated differently, μ is the spectral measure of J corresponding to the cyclic vector e_0. These two maps are inverses of each other. This result is standard and can be found e.g. [32, Theorem 10.9]

Theorem 1.1 (Favard's theorem)**.** *Let μ be a non-trivial probability measure with compact support in $\mathbb{R}$, $(p_n)_{n=0}^\infty$ the associated orthonormal polynomials and J be defined by (1.4) and (1.5). If $a_n > 0, b_n \in \mathbb{R}$ are bounded sequences and $J = J(((a_{n+1}, b_n))_{n=0}^\infty)$, let μ be the measure in the integral representation of m defined by (1.8). Then the map $\mu \mapsto J$ and $J \mapsto \mu$ are mutually inverses.*

Transfer matrices and coefficient stripping

Writing (1.4) into matrix form, we get

$$\begin{pmatrix} p_{n+1}(z) \\ -a_{n+1}p_n(z) \end{pmatrix} = \mathfrak{a}(z, a_{n+1}, b_n) \begin{pmatrix} p_n(z) \\ -a_n p_{n-1}(z) \end{pmatrix}$$

where

$$\mathfrak{a}(z, a, b) = \begin{pmatrix} \frac{z-b}{a} & \frac{1}{a} \\ -a & 0 \end{pmatrix}.$$

In particular, defining the *transfer matrix* $\mathfrak{A}_n(z)$ by

$$\mathfrak{A}_n(z) := \mathfrak{a}(z, a_n, b_{n-1}) \cdots \mathfrak{a}(z, a_1, b_0),$$

we see that

$$\begin{pmatrix} p_n(z) \\ -a_n p_{n-1}(z) \end{pmatrix} = \mathfrak{A}_n(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The second column of $\mathfrak{A}_n(z)$ is given in terms of the orthonormal polynomials of the second kind

$$q_n(z) := \int \frac{p_n(\xi) - p_n(z)}{\xi - z} d\mu(\xi), \tag{1.9}$$

with $a_0 q_{-1}(z) = -1$ by convention. By means of (1.4) and (1.9) one gets

$$\begin{pmatrix} q_n(z) \\ -a_n q_{n-1}(z) \end{pmatrix} = \mathfrak{A}_n(z) \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

that is

$$\mathfrak{A}_n(z) = \begin{pmatrix} p_n(z) & q_n(z) \\ -a_n p_{n-1}(z) & -a_n q_{n-1}(z) \end{pmatrix}. \tag{1.10}$$

The transfer matrix also plays an important role in order to describe shifts of Weyl m-functions. Let $J^{(n)}$ denote the Jacobi matrix which is obtained from J by stripping the first n rows and columns. To be precise, denote by S_+ the right-shift operator on $\ell^2(\mathbb{N}_0)$, acting on the standard basis by $S_+ e_n = e_{n+1}$. Then

$$J^{(n)} = (S_+^*)^n J (S_+)^n.$$

If $m^{(n)}$ denotes the corresponding Weyl m-function, then it can be seen that

$$m^{(1)}(z) = \frac{-1}{z - b_0 + a_1^2 m(z)}, \tag{1.11}$$

see e.g. [44, Theorem 3.2.4.]. It will be convenient to write this in projective coordinates of the Riemann sphere $\overline{\mathbb{C}}$. For $u, v \in \mathbb{C}^2$, we write $u \sim v$, if there exists $\lambda \in \mathbb{C}$ such that $u = \lambda v$. In projective coordinates, we can rewrite (1.11) as

$$\begin{pmatrix} m(z) \\ 1 \end{pmatrix} \sim \mathfrak{a}(z, a_1, b_0) \begin{pmatrix} m^{(1)}(z) \\ 1 \end{pmatrix}.$$

Iterating this, yields

$$\begin{pmatrix} m(z) \\ 1 \end{pmatrix} \sim \mathfrak{A}_n(z) \begin{pmatrix} m^{(n)}(z) \\ 1 \end{pmatrix}. \tag{1.12}$$

Direct and inverse spectral theory of periodic Jacobi matrices

If J is periodic, the relation (1.12) becomes particularly relevant. We will assume in this section that there exists $q > 0$ such that for all n

$$a_{n+q} = a_n \quad \text{and} \quad b_{n+q} = b_n.$$

As will become clear below, it is more natural in this context to extend J by periodicity to an operator on $\ell^2(\mathbb{Z})$. That is, let J be an operator acting on $\ell^2(\mathbb{Z})$, whose matrix representation in the standard basis is

$$J = \left(\begin{array}{ccc|ccc} \ddots & \ddots & & & & \\ \ddots & b_{-2} & a_{-1} & & & \\ & a_{-1} & b_{-1} & a_0 & & \\ \hline & & a_0 & b_0 & a_1 & \\ & & & a_1 & b_1 & a_2 \\ & & & & a_2 & \ddots & \ddots \\ & & & & & \ddots \end{array}\right).$$

We will use the term *half-line* for operators acting on $\ell^2(\mathbb{N}_0)$ and *full-line* for operators acting on $\ell^2(\mathbb{Z})$. Let $\ell^2_+ = \ell^2(\mathbb{N}_0), \ell^2_- = \ell^2(\mathbb{Z}_{<0})$ and $P_\pm$ denote the orthogonal projections from $\ell^2(\mathbb{Z})$ onto $\ell^2_\pm$ and $J_\pm = P_\pm J P_\pm$ and let e_n, $n \in \mathbb{Z}$, denote the standard basis of ℓ^2. Moreover, define the finite range operator $F : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ by

$$F(v) = \langle v, e_0 \rangle e_{-1} + \langle v, e_{-1} \rangle e_0.$$

With this notation we have

$$J = J_+ \oplus J_- + a_0 F. \tag{1.13}$$

Let $m_\pm$ denote the Weyl m-functions associated to $J_\pm$. Combining (1.12) with periodicity, we get:

Theorem 1.2. *Let J be a q-periodic full-line Jacobi matrix and $m_\pm$ be defined as above. Then m_+ is a solution of*

$$\mathfrak{A}_q(z)_{21} m(z)^2 + (\mathfrak{A}_q(z)_{22} - \mathfrak{A}_q(z)_{11}) m(z) - \mathfrak{A}_q(z)_{12} = 0, \tag{1.14}$$

where $\mathfrak{A}_q(z)_{ij}$ denote the entries of the transfer matrix $\mathfrak{A}_q(z)$. Moreover, the second solution is given by $(a_0^2 m_-(z))^{-1}$.

Proof. We will only show that m_+ is a solution of (1.14). That $(a_0^2 m_-(z))^{-1}$ is the second solution of (1.14) can be found for instance in [44, Theorem 5.2.2.]

Periodicity implies that $J_+^{(q)} = J_+$ and $m_+^{(q)} = m_+$. Thus, (1.12) yields

$$m_+(z) = \frac{\mathfrak{A}_q(z)_{11} m_+(z) + \mathfrak{A}_q(z)_{12}}{\mathfrak{A}_q(z)_{21} m_+(z) + \mathfrak{A}_q(z)_{22}},$$

which is (1.14). □

Periodic discriminant

Define

$$\Delta(z) := \operatorname{tr} \mathfrak{A}_q(z) = p_q(z) - a_0 q_{q-1}(z). \tag{1.15}$$

Clearly, Δ is a polynomial of degree q and is called the (polynomial) *discriminant*. In view of (1.19) and (1.20) below, it is clear that studying Δ is essential in order to describe the spectrum of J_+ and J. We collect the crucial properties of Δ in the following theorem. These properties are for instance proved in [32, Section 10.10.].

Theorem 1.3. *Let Δ be defined by (1.15). Then it holds that*

(i) For all $z \in \mathbb{C}$ $\Delta(\overline{z}) = \overline{\Delta(z)}$;

(ii) All zeros of Δ are real and simple;

(iii) $\Delta'(c) = 0$ implies $c \in \mathbb{R}$ and $|\Delta(c)| \geq 2$;

In particular, for

$$\mathsf{E} := \Delta^{-1}([-2,2]) = \{z \in \mathbb{C} \mid \Delta(z) \in [-2,2]\} \tag{1.16}$$

it holds that

$$\mathsf{E} \subset \mathbb{R}. \tag{1.17}$$

We will see below that $\sigma_{\text{ess}}(J_+) = \sigma(J) = \mathsf{E}$. In general, E is the union of g disjoint intervals, where $0 \leq g \leq p-1$. We will write E in the form

$$\mathsf{E} = [\mathbf{b}_0, \mathbf{a}_0] \setminus \bigcup_{j=1}^{g} (\mathbf{a}_j, \mathbf{b}_j), \tag{1.18}$$

where $\mathbf{b}_0 < \mathbf{a}_1 < \mathbf{b}_1 < \cdots < \mathbf{b}_g < \mathbf{a}_0$. The sets $(\mathbf{a}_j, \mathbf{b}_j)$ are called gaps of E. If for some critical point c we have $|\Delta(c)| = 2$, we refer to it as a closed gap. Subsets of E on which Δ is monotonic are called bands of E. To be precise, let $A = \Delta^{-1}((-2,2))$. If (c,d) is some connected component of A, then $[c,d]$ is called a band of the spectrum. Note that two bands join, if and only if the gap is closed. We say that *all gaps are open*, if for every critical point c it holds that $|\Delta(c)| > 2$. Note that this corresponds exactly to the case that $g = p-1$.

By Theorem 1.3, all solutions of $\Delta(z)^2 = 4$ belong to $[\mathbf{b}_0, \mathbf{a}_0]$ and we can thus choose an analytic branch of $\sqrt{\Delta(z)^2 - 4}$ on $\mathbb{C} \setminus (-\infty, \mathbf{a}_0]$, so that $\sqrt{\Delta(z)^2 - 4} > 0$ for $z > \mathbf{a}_0$. With this choice of the square root, using that $\det \mathfrak{A}_p(z) = 1$, Theorem 1.2 and (1.10), one obtains

$$m_+(z) = -\frac{\beta(z) - \sqrt{\Delta(z)^2 - 4}}{2a_q p_{q-1}(z)}, \quad \beta(z) = p_q(z) + a_q q_{q-1}(z) \tag{1.19}$$

and the second solution of the quadratic equation yields

$$-\frac{1}{a_0^2 m_-(z)} = \frac{\beta(z) + \sqrt{\Delta(z)^2 - 4}}{2a_q p_{q-1}(z)}. \tag{1.20}$$

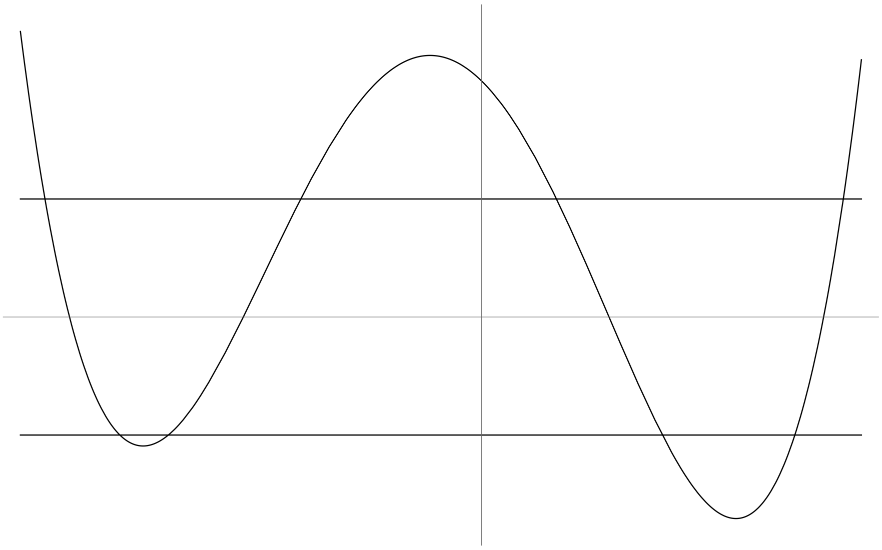


Figure 1.1: Discriminant of a 4-periodic Jacobi matrix

Dirichlet divisors and isospectral torus

In the following, we will find the spectrum of J and J_+ and characterize the spectral data that allows to solve the inverse problem, i.e., to uniquely recover J and J_+. Let us define

$$g_{n,m}(z) = \langle (J-z)^{-1} e_m, e_n \rangle.$$

For a full-line Jacobi matrix J on ℓ^2, the vectors $\{e_{-1}, e_0\}$ are cyclic for J, and we are thus led to consider the matrix-valued M-function

$$M(z) = \begin{pmatrix} g_{-1,-1}(z) & g_{-1,0}(z) \\ g_{0,-1}(z) & g_{0,0}(z) \end{pmatrix}.$$

A calculation using the second resolvent identity and (1.13) shows that

$$M(z) = \begin{pmatrix} m_-(z)^{-1} & a_0 \\ a_0 & m_+(z)^{-1} \end{pmatrix}^{-1}. \tag{1.21}$$

In particular,

$$-\frac{1}{g_{-1,-1}(z)} = -\frac{1}{m_-(z)} + a_0^2 m_+(z), \quad -\frac{1}{g_{0,0}(z)} = -\frac{1}{m_+(z)} + a_0^2 m_-(z) \tag{1.22}$$

and

$$g_{0,0}(z) = \frac{m_+(z)}{m_-(z)} g_{-1,-1}(z). \tag{1.23}$$

Combining (1.19), (1.20) and (1.22) we get

$$g_{-1,-1}(z) = -\frac{p_{q-1}(z)}{a_0\sqrt{\Delta^2-4}}, \quad g_{0,0}(z) = -\frac{q_q(z)}{\sqrt{\Delta^2-4}}. \tag{1.24}$$

We say that a full-line Jacobi matrix J is *reflectionless* on a set $A \subset \mathbb{R}$, if for a.e. $\xi \in A$

$$m_+(\xi + i0) = \frac{1}{\overline{a_0^2 m_-(\xi + i0)}}, \tag{1.25}$$

where $f(\xi+i0)=\lim_{y\downarrow 0} f(\xi+iy)$ denotes the a.e. existent normal boundary values of a Herglotz function. It follows from the explicit formulae (1.19) and (1.20) that a periodic Jacobi matrix J is reflectionless on E given by (1.16).

Lemma 1.4. *Let J be a q-periodic full-line Jacobi matrix, $\mathsf{E}=\Delta^{-1}([-2,2])$. Then J is reflectionless on E.*

Zeros of p_{q-1} are called *Dirichlet eigenvalues*, since they correspond to eigenvalues of the truncation of J_+ onto $\ell^2(\{0,\cdots,q-2\})$. In principle, due to (1.19) and (1.20), poles of m_+ and $(a_0^2m_-)^{-1}$ can occur at Dirichlet eigenvalues. However, it turns out that Dirichlet eigenvalues that belong to E do not lead to poles of these functions and if a Dirichlet eigenvalue belongs to $\mathbb{R}\setminus\mathsf{E}$, then it is either a pole of m_+ or a pole of $(a_0^2m_-)^{-1}$ and not of both:

Lemma 1.5. *Assume that $p_{q-1}(x)=0$ or $q_q(x)=0$, then $|\Delta(x)|\geq 2$. Moreover, if $A=\Delta^{-1}((-2,2))$, then there is exactly one Dirichlet eigenvalue in each connected component of $[\mathbf{b}_0,\mathbf{a}_0]\backslash A$. If x is a Dirichlet eigenvalue such that $x\in(\mathbf{a}_j,\mathbf{b}_j)$, $1\leq j\leq g$, then x is either a pole of m_+ or a pole of $(a_0^2m_-)^{-1}$.*

Proof. The first statement is a direct consequence of $\det\mathfrak{A}_q=1$ and the AM-GM inequality. A standard Herglotz function argument shows that the zeros of Δ and p_{q-1} strictly interlace, which implies the second statement. Finally, if m_+ and $(m_-)^{-1}$ would have a pole, (1.23) and (1.24) would imply that $g_{0,0}$ has a pole, contradicting (1.24). □

This motivates the following definition. Let x_j denote the Dirichlet eigenvalue in $[\mathbf{a}_j,\mathbf{b}_j]$, $1\leq j\leq g$[1]. Then we set

$$\varepsilon_j=\begin{cases}-1, & \text{if } x_j\in(\mathbf{a}_j,\mathbf{b}_j) \text{ and } x_j \text{ is a pole of } (a_0^2m_-)^{-1}\\ 1, & \text{otherwise.}\end{cases}$$

There is a certain ambiguity, since we set $\varepsilon_j=1$, if $x_j\in\{\mathbf{a}_j,\mathbf{b}_j\}$. This is somewhat arbitrary and is only for definiteness, as will become clear below. The collection of all Dirichlet eigenvalues together with the marker ε_j is called *Dirichlet divisor*

$$D=((x_j,\varepsilon_j))_{j=1}^g.$$

Together with E, it allows us to uniquely recover, J,J_+,J_-. We collect the full spectral characterization in the following theorem; see [32, Theorem 10.77.,Theorem 10.82.]

Theorem 1.6. *Let J be a q-periodic full-line Jacobi matrix, $\mathsf{E}=\Delta^{-1}([-2,2])$ and D the associated Dirichlet divisor. Then the following holds:*

(i) J has purely absolutely continuous spectrum of multiplicity 2 and is given by $\sigma(J)=\mathsf{E}$,

(ii) J_+ has simple spectrum and it holds

a) The essential spectrum is given by $\sigma_{\mathrm{ess}}(J_+)=\mathsf{E}$,

b) The discrete spectrum is given by

$$\sigma_d(J_+)=\{x_j\mid x_j \textit{ is Dirichlet eigenvalue in } (\mathbf{a}_j,\mathbf{b}_j) \textit{ and } \varepsilon_j=1\}.$$

[1] Note that we do not consider Dirichlet eigenvalues that belong to closed gaps.

(iii) J can be uniquely recovered from E and D via the formulae

$$a_0^2 m_+(z) = \frac{1}{2}\left(-\frac{1}{g_{-1,-1}(z)} - (z+\alpha) + \sum_{j=1}^{g} \frac{\varepsilon_j \sigma_j}{x_j - z}\right) \tag{1.26}$$
$$-\frac{1}{m_-(z)} = \frac{1}{2}\left(-\frac{1}{g_{-1,-1}(z)} + (z+\alpha) + \sum_{j=1}^{g} \frac{\varepsilon_j \sigma_j}{x_j - z}\right)$$

where

$$\sigma_k = \frac{\sqrt{\prod_{j=0}^{g}(x_k - \mathbf{a}_j)(x_k - \mathbf{b}_j)}}{\prod_{k\neq j}(x_k - x_j)}, \quad \textit{and,} \quad \alpha = -\frac{1}{2}\left(\mathbf{a}_0 + \mathbf{b}_0 + \sum_{j=1}^{g}(\mathbf{a}_j + \mathbf{b}_j - 2x_j)\right).$$

Remark 1.7. (i) By (1.24), $g_{-1,-1}$ is determined by D and E. Moreover, since $m_+(z) = -\frac{1}{z} + o(1)$, for $z \to \infty$, we can recover a_0 from (1.26). Thus, we get $m_\pm$, which by Favard's theorem allows to uniquely recover $J_\pm$ and a_0. By (1.13) this uniquely allows to recover J. Thus, if E is fixed, the map $J \mapsto D$ is injective.

(ii) Although we have argued with the explicit formulae in terms of orthonormal polynomials, which was possible due to the periodicity of J, it can be shown that the crucial properties that are required to prove Theorem 1.6 is (1.21) and the fact that J is reflectionless on E. This will become important below, where we will associate Jacobi matrices to spectral sets that are not necessarily spectral sets of periodic Jacobi matrices.

Let us define the set of all Dirichlet divisors by

$$\mathcal{D}(\mathsf{E}) = \{((x_j, \varepsilon_j))_{j=1}^{g} \mid x_j \in [\mathbf{a}_j, \mathbf{b}_j], \varepsilon_j \in \{-1, 1\}, 1 \le j \le g\}/\sim, \tag{1.27}$$

with the identification $(x_j, 1) \sim (x_j, -1)$, if $x_j \in \{\mathbf{a}_j, \mathbf{b}_j\}$ and equip it with the product topology of circles. We have described the map from periodic Jacobi matrices with spectrum E into $\mathcal{D}(\mathsf{E})$ and showed that this map is injective. For the given E, we define the *isospectral torus* of Jacobi matrices by

$$\mathcal{T}_{\mathsf{E}}^{J} := \{J \mid J \text{ is full-line periodic Jacobi matrix with } \sigma(J) = \mathsf{E}\}. \tag{1.28}$$

By construction of the set E, we see that $\mathcal{T}_{\mathsf{E}}^{J}$ is not empty. Moreover, we have seen that the map periodic Jacobi matrix to its Dirichlet divisor is an injective map from $\mathcal{T}_{\mathsf{E}}^{J}$ into $\mathcal{D}(\mathsf{E})$. But much more is true:

Theorem 1.8. *The map $J \mapsto D$ sets up a bijection between $\mathcal{T}_{\mathsf{E}}^{J}$ and $\mathcal{D}(\mathsf{E})$. If we equip $\mathcal{T}_{\mathsf{E}}^{J}$ with the strong operator topology and $\mathcal{D}(\mathsf{E})$ with the product topology of circles, then this map is a homeomorphism.*

The proof can be found e.g. in [44, Section 5.13.] and in a much more general setting in [46]. This shows that $\mathcal{T}_{\mathsf{E}}^{J}$ is homeomorphic to $\mathbb{T}^g$, which justifies the name isospectral torus.

Example 1.9. If $\mathsf{E} = [-2, 2]$, then $\mathcal{T}_{\mathsf{E}}^{J}$ consists of the single "free" Jacobi matrix J_0, whose Jacobi coefficients are given by $a_n \equiv 1$ and $b_n \equiv 0$.

Note that $J_0 = S + S^{-1}$, where S denotes the right shift operator on ℓ^2, defined by $Se_n = e_{n+1}$. The following characterization of $\mathcal{T}_{\mathsf{E}}^{J}$ bears the name "magic formula":

Theorem 1.10. *Let J be a full-line Jacobi matrix and let E be the spectrum of a q-periodic Jacobi matrix with discriminant Δ. Then the following holds*

$$J \in \mathcal{T}_J(E) \iff \Delta(J) = S^q + S^{-q}. \tag{1.29}$$

Note that $S^q + S^{-q}$ can be viewed as a matrix version of J_0, where the 1 and 0 coefficients are substituted by the $q \times q$ identity and zero matrix. This point of view of considering $\Delta(J)$ as a block Jacobi matrix with matrix coefficients will be crucial below.

This description is only possible if we a priori assume that E is the spectrum of a periodic Jacobi matrix. In this case we have seen that there exists a polynomial discriminant. However, this gives a strong restrictions on E. In particular, a generic union of intervals will not be the spectrum of a periodic Jacobi matrix as the following theorem shows. Let $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$ and $G_{\mathsf{E}}(z,\infty)$ denote the potential theoretic Green function of Ω with pole at ∞. That is, the unique harmonic function in $\Omega \setminus \{\infty\}$ which has vanishing boundary values on E and satisfies $G_{\mathsf{E}}(z,\infty) = \log|z| + O(1)$ as $z \to \infty$. Moreover, let $\omega_{\mathsf{E}}(A,z)$ denote the harmonic measure of Ω, the set $A \subset \mathsf{E}$ and the point $z \in \Omega$. For later reference, we formulate the following theorem for sets E, which are arbitrary union of g intervals written in the form (1.18). Let us denote $\mathsf{E}_k = \mathsf{E} \cap [\mathbf{b}_0, \mathbf{a}_k]$. Then we have

Theorem 1.11. *Let $\mathsf{E} \subset \mathbb{R}$ with no isolated points and assume that there exists a polynomial, P, of degree q such that $P^{-1}([-2,2]) = \mathsf{E}$. Then*

$$B(z) = \frac{P(z) - \sqrt{P(z)^2 - 4}}{2},$$

is a function on Ω. Moreover, we have

$$\log|B(z)| = -qG_{\mathsf{E}}(z,\infty)$$

and for $1 \le k \le g$, $q\omega_{\mathsf{E}}(\mathsf{E}_k,\infty) \in \mathbb{N}$. This holds if and only if E is the spectrum of a q-periodic Jacobi matrix.

Proof. Recall that the Joukowsky map $f(\zeta) = \zeta + \frac{1}{\zeta}$ maps $\mathbb{D}$ conformally onto $\overline{\mathbb{C}} \setminus [-2,2]$. Let $g(u) = \frac{u - \sqrt{u^2-4}}{2}$, denote its inverse. Since by assumption P maps Ω into $\overline{\mathbb{C}} \setminus [-2,2]$, $B(z) = g(P(z))$ is a well defined function in Ω. By construction $|B| \le 1$ in Ω, and $|B(z)| \to 1$ as $z \to \mathsf{E}$. By the asymptotic behavior for G_{E} at ∞, we see that

$$\log|B(z)| + qG_{\mathsf{E}}(z,\infty)$$

is harmonic in Ω and vanishes on the boundary, and thus by the maximum principle vanishes identically. Computing the change of the argument of $G_{\mathsf{E}}(z,\infty)$ when it circles around E_k and the fact that B is a single valued function, implies the statement about the harmonic measure. If E is the spectrum of a q-periodic Jacobi matrix, then we can choose $P = \Delta$. We refer to [44, Theorem 5.5.25.] for the other direction of the last statement. □

The Damanik–Killip–Simon approach and the Yuditskii discriminant

We start this section with a classical result of Szegő. Let μ be a non-trivial probability measure supported on $\partial\mathbb{D} = \{z \in \mathbb{C} \mid |z| = 1\}$ and let Φ_n denote the monic orthogonal polynomials in $L^2(\mu)$. These polynomials obey the recursive relation

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

with uniquely determined $\alpha_n \in \mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$, where $\Phi_n^*(z) = z^n\overline{\Phi_n(1/\bar{z})}$. In this context, the coefficients α_n are called Verblunsky coefficients. Let f_μ denote the density of μ with respect to the normalized Lebesgue measure on $\partial\mathbb{D}$. Then the following statement holds:

$$\sum_{n=0}^{\infty} |\alpha_n|^2 < \infty \iff \int \log(f_\mu(\theta))\frac{\mathrm{d}\theta}{2\pi} > -\infty. \tag{1.30}$$

Szegő's theorem represents what Simon calls a gem of spectral theory [44, Section 1.4]: a one-to-one correspondence between a class of measures and a class of coefficients. This theorem can also be interpreted in terms of certain unitary operators, called CMV matrices as will be discussed later in this introduction. Remarkably, this theorem has an analog for self-adjoint operators, that was proved by Killip and Simon [26].

Theorem 1.12. *Let μ be a compactly supported real probability measure with half-line Jacobi matrix J_+ and Jacobi parameters $((a_{n+1}, b_n))_{n=0}^{\infty}$ and write $\mathrm{d}\mu = f_\mu(\xi)d\xi + d\mu_s$. Then*

$$\sum_{n=0}^{\infty} |a_{n+1} - 1|^2 + |b_n|^2 < \infty \tag{1.31}$$

if and only if

(i) $\sigma_{\mathrm{ess}}(J_+) = [-2, 2]$,

(ii) $\displaystyle\int_{-2}^{2} \log(f_\mu(\xi))\sqrt{\mathrm{dist}(\xi, \mathbb{R} \setminus [-2, 2])}\mathrm{d}\xi > -\infty$,

(iii) The eigenvalues $\xi_n \notin [-2, 2]$ obey $\displaystyle\sum_{n=0}^{\infty} \sqrt{\mathrm{dist}(\xi_n, [-2, 2])}^3 < \infty$.

This has been generalized by Damanik, Killip and Simon to arbitrary spectral sets of periodic Jacobi matrices, with the additional assumption that all gaps are open. In order to formulate the result, we need a notion of convergence to the isospectral torus, which will substitute (1.31). Recall that $\mathcal{T}_J([-2, 2]) = \{J_0\}$ and note that (1.31) means that $J - J_0$ is a Hilbert-Schmidt operator. The generalization of Theorem 1.12 is formulated in terms of half-line operators. The isospectral torus of half-line Jacobi matrices is defined similarly to (1.28):

$$\mathcal{T}_{\mathsf{E}}^{J_+} = \{J_+ \mid J_+ \text{ is half-line periodic Jacobi matrix with } \sigma_{\mathrm{ess}}(J_+) = \mathsf{E}\}.$$

On bounded half-line Jacobi matrices J_+, we consider the metric

$$d(J_+, \tilde{J}_+) = \sum_{k=0}^{\infty} e^{-k}(|a_{k+1} - \tilde{a}_{k+1}| + |b_k - \tilde{b}_k|). \tag{1.32}$$

On norm-bounded sets of Jacobi matrices, convergence in this metric corresponds to strong operator convergence. However, instead of distance to a fixed Jacobi matrix $\tilde{J}_+$, we will consider the distance to $\mathcal{T}_{\mathsf{E}}^{J_+}$,

$$d(J_+, \mathcal{T}_{\mathsf{E}}^{J_+}) = \inf_{\tilde{J}_+ \in \mathcal{T}_{\mathsf{E}}^{J_+}} d(J_+, \tilde{J}_+) = \min_{\tilde{J}_+ \in \mathcal{T}_{\mathsf{E}}^{J_+}} d(J_+, \tilde{J}_+).$$

Then, the generalization of Theorem 1.12 for periodic Jacobi matrices is the following:

Theorem 1.13. *Let* E *be the spectrum of a periodic Jacobi matrix such that all gaps are open. Let* μ *be a compactly supported real probability measure with half-line Jacobi matrix* J_+ *and denote* $\mathrm{d}\mu = f_\mu(\xi)d\xi + d\mu_s$. *Then*

$$\sum_{n=0}^{\infty} d((S_+^*)^n J_+ (S_+)^n, \mathcal{T}_{\mathsf{E}}^{J_+}) < \infty \tag{1.33}$$

if and only if

(i) $\sigma_{\text{ess}}(J_+) = \mathsf{E}$,

(ii) $\displaystyle\int_{-2}^{2} \log(f_\mu(\xi))\sqrt{\operatorname{dist}(\xi, \mathbb{R}\setminus\mathsf{E})}\mathrm{d}\xi > -\infty$,

(iii) The eigenvalues $\xi_n \notin \mathsf{E}$ *obey* $\displaystyle\sum_{n=0}^{\infty} \sqrt{\operatorname{dist}(\xi_n, \mathsf{E})}^3 < \infty$.

Besides its own interest, the main reason for mentioning these results is the approach that allowed to extend Theorem 1.12 to Theorem 1.13. The proof of Theorem 1.13 essentially consists of four steps. Let us briefly recall that a block Jacobi matrix is of the form

$$\mathbf{J} = \begin{pmatrix} \mathfrak{w}_0 & \mathfrak{v}_0 & & & \\ \mathfrak{v}_0^* & \mathfrak{w}_1 & \mathfrak{v}_1 & & \\ & \mathfrak{v}_1^* & \mathfrak{w}_2 & \mathfrak{v}_2 & \\ & & \mathfrak{v}_2^* & \ddots & \ddots \\ & & & \ddots & \end{pmatrix} \tag{1.34}$$

where $\mathfrak{v}_j$ and $\mathfrak{w}_j$ are $d \times d$ matrices, $\mathfrak{w}_j = \mathfrak{w}_j^*$, and $\det \mathfrak{v}_j \neq 0$ for each j. Type 3 block Jacobi matrices have each $\mathfrak{v}_j$ lower triangular and positive on the diagonal. If $\mathbf{J}$ is a $d \times d$-block Jacobi matrix, then it has a $d \times d$-spectral measure $d\Xi$. Note that if J_+ is a Jacobi matrix, then for a periodic discriminant of degree q, $\Delta(J_+)$ is a $q \times q$ block Jacobi matrix.

Let us now outline the proof of Theorem 1.13. Recall that Δ is the discriminant for the spectrum E of a periodic Jacobi matrix.

1. Prove a version of Theorem 1.12 for block Jacobi matrices. The analogous coefficient condition (1.31) in this theorem is $\mathbf{J} - (S_+^q + (S_+^*)^q)$ is a Hilbert Schmidt operator,
2. Apply this block version to the block Jacobi matrix $\Delta(J_+)$,
3. Translate the condition on the matrix spectral measure of $\Delta(J_+)$ into a condition on the spectral measure of J_+,
4. Translate the condition $\Delta(J_+) - (S_+^q + (S_+^*)^q)$ is a Hilbert–Schmidt operator into a condition on the Jacobi coefficients of J_+.

In all these steps, the technically most challenging part is step 4, since the coefficients of J_+ appear in $\Delta(J_+) - (S_+^q + (S_+^*)^q)$ in a very complicated way. It is important to mention that this approach relies in a crucial way on existence of the function Δ, which, due to the magic formula (1.29), maps any element of the isospectral torus to the block Jacobi matrix version of J_0. This provides a serious obstacle to extend the Damanik–Killip–Simon approach to other spectral sets, such as arbitrary union of intervals.

The isospectral torus of finite-gap sets and the Yuditskii discriminant

In this section we will discuss spectral sets, E, that are arbitrary finite union of intervals. We will refer to them in the following as finite-gap sets. In particular, we do not assume that there exists a periodic Jacobi matrix with the given spectral set. Still we write E in the form (1.18). Clearly, the definition of $\mathcal{D}(\mathsf{E})$ (1.27) carries over to this setting without change. It is not so immediately clear, how to define the isospectral torus $\mathcal{T}_{\mathsf{E}}^J$ of Jacobi matrices in this setting. As already indicated above, the property of being reflectionless is the key to extend the notion of isospectral torus to arbitrary finite-gap sets.

$$\mathcal{T}_{\mathsf{E}}^J = \{J \mid J \text{ is full-line reflectionless Jacobi matrix on } \mathsf{E} \text{ and } \sigma(J) = \mathsf{E}\}. \tag{1.35}$$

It can then be shown, that $\mathcal{T}_{\mathsf{E}}^J$ can equivalently be given by

$$\mathcal{T}_{\mathsf{E}}^J = \{J \mid J \text{ is full-line almost periodic Jacobi matrix with } \sigma(J) = \sigma_{\text{ac}}(J) = \mathsf{E}\}.$$

Using definition (1.35), the spectral characterization of Theorem 1.6 for elements of $\mathcal{T}_{\mathsf{E}}^J$ still holds without change. It also still holds that $\mathcal{T}_{\mathsf{E}}^J$ and $\mathcal{D}(\mathsf{E})$ are homeomorphic, so that $\mathcal{T}_{\mathsf{E}}^J$ can be identified with a g-dimensional torus, where g is the number of gaps of E. However, it was seen in Theorem 1.11 that existence of a polynomial P with $P^{-1}([-2,2]) = \mathsf{E}$ imposes a strong restriction on the harmonic measure of E_k, and thus generically does not hold for arbitrary finite-gap sets.

In [53], Yuditskii overcame this issue by introducing a rational discriminant in terms of the Ahlfors function at ∞. Let $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$. The Ahlfors function is the unique analytic function, $w : \Omega \to \mathbb{D}$, with $w(\infty) = 0$ which maximizes $\mathrm{Re}(zw(z))|_{z=\infty}$. The Ahlfors function at ∞ will be denoted by w_∞. Moreover, let $G_{\mathsf{E}}(z, z_0)$ denote the Green function of the domain Ω with pole at z_0. We collect properties of the Ahlfors function and the Yuditksii discriminant in the following theorem:

Theorem 1.14. *Let E be a finite-gap set written as in (1.18) and w_∞ the Ahlfors function for $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$ at ∞. Then for $1 \le j \le g$ there exists $\mathbf{c}_j \in (\mathbf{a}_j, \mathbf{b}_j)$ such that $w_\infty(z) = 0$ if and only if $z \in \{\infty, \mathbf{c}_1, \cdots, \mathbf{c}_g\}$. It holds that*

$$-\log |w_\infty(z)| = G_{\mathsf{E}}(z, \infty) + \sum_{j=1}^{g} G_{\mathsf{E}}(z, \mathbf{c}_j).$$

If

$$\Delta_{\mathsf{E}}(z) := w_\infty(z) + \frac{1}{w_\infty(z)},$$

then

$$\Delta_{\mathsf{E}}(z) = \lambda_{g+1} z + d + \sum_{j=1}^{g} \frac{\lambda_j}{\mathbf{c}_j - z}, \quad \lambda_j > 0 \tag{1.36}$$

and

$$\Delta_{\mathsf{E}}^{-1}([-2,2]) = \mathsf{E} \tag{1.37}$$

Moreover, Δ_{E} is the unique rational function in $\mathcal{N}_0$ with $\Delta_{\mathsf{E}}(\infty) = \infty$ such that (1.37) holds.

If J is a Jacobi matrix and $\mathbf{c} \in \mathbb{R} \setminus \sigma(J)$, then $(\mathbf{c} - J)^{-1}$ is an infinite band matrix. This suggests that in order to apply the Damanik–Killip–Simon approach with the rational Yuditskii discriminant it is required to use operators that are suitable for the rational function Δ_{E}. This is achieved by using orthogonal rational functions instead of orthogonal polynomials.

Fix a vector $\mathbf{C}_{g+1} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ with no repetitions such that $\mathbf{c}_{k_\infty} = \infty$ for some $1 \le k_\infty \le g+1$ and let $\mathbf{C} = (\mathbf{c}_k)_{k=1}^{\infty}$ be the sequence obtained from $\mathbf{C}_{g+1}$ by periodic repetition, i.e., $\mathbf{c}_k = \mathbf{c}_j$ if and only if $k = j \mod (g+1)$. Let τ_n be the sequence of orthogonal rational functions as in Problem 1.3 associated to $\mathbf{C}$. A half-line GMP matrix [53] is the matrix representation of the multiplication operator $M_{z,\mu}$ in the basis $(\tau_n)_{n=0}^{\infty}$; its matrix elements are

$$A_{mn} = \int \xi \tau_n(\xi) \overline{\tau_m(\xi)} \, \mathrm{d}\mu(\xi). \tag{1.38}$$

The condition that $\mathbf{c}_{k_\infty} = \infty$ for some k_∞ guarantees that $A_{mn} = 0$ for $|m-n| > g+1$, so these matrix elements generate a bounded operator A on $\ell^2(\mathbb{N}_0)$ such that $A_{mn} = \langle A e_n, e_m \rangle$. We say that $A \in \mathbb{A}(\mathbf{C}_{g+1})$.

GMP matrices have the property that some of their resolvents are also GMP matrices; namely, for any $k \neq k_\infty$, $(\mathbf{c}_k - A)^{-1} \in \mathbb{A}(f(\mathbf{C}_{g+1}))$ where f is the Möbius transform $f : z \mapsto (\mathbf{c}_k - z)^{-1}$ and $f(\mathbf{C}_{g+1}) = (f(\mathbf{c}_1), \dots, f(\mathbf{c}_g))$. Note that the special case $g = 0$, $\mathbf{C}_1 = (\infty)$ gives precisely a Jacobi matrix.

Similar to Jacobi matrices, GMP matrices have a very special structure that allows to give a definition of a full-line GMP matrix associated to a fixed finite sequence of poles $\mathbf{C}_{g+1}$ as above:

Definition 1.15. Fix a finite sequence $\mathbf{C}_{g+1} = (\mathbf{c}_1, \dots, \mathbf{c}_g, \infty)$[2] and let X^- denote the upper triangular part of a matrix X (excluding the diagonal), and X^+ the lower triangular part (including the diagonal). Then we say that a full-line operator A is GMP structured, and denote it by $A \in \mathbb{A}$, if it is a $(g+1)$-block Jacobi matrix

$$A = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ & A_{-1}^* & B_{-1} & A_0 & & \\ & & A_0^* & B_0 & A_1 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix} \tag{1.39}$$

such that

$$A_j = \delta_g \vec{p}_j^{\,*}, \quad B_j = (\vec{q}_j \vec{p}_j^{\,*})^- + (\vec{p}_j \vec{q}_j^{\,*})^+ + \hat{\mathbf{C}}, \quad \vec{p}_j, \vec{q}_j \in \mathbb{R}^{g+1},$$

and

$$\hat{\mathbf{C}} = \begin{pmatrix} \mathbf{c}_1 & & & \\ & \ddots & & \\ & & \mathbf{c}_g & \\ & & & 0 \end{pmatrix}, \quad \vec{p}_j = \begin{pmatrix} p_0^{(j)} \\ \vdots \\ p_g^{(j)} \end{pmatrix}, \quad \vec{q}_j = \begin{pmatrix} q_0^{(j)} \\ \vdots \\ q_g^{(j)} \end{pmatrix}, \quad p_g^{(j)} > 0.$$

We then say an operator $A \in \mathbb{A}$ is a full-line GMP matrix, if the resolvents $(\mathbf{c}_\ell - A)^{-1}$ exist for all $1 \le \ell \le g$ and $S^{-\ell}(\mathbf{c}_\ell - A)^{-1} S^\ell \in \mathbb{A}$. In this case we write $A \in \mathrm{GMP}(\mathbf{C}_{g+1})$. Again we call the generating coefficients $(\vec{p}_j, \vec{q}_j)_{j \in \mathbb{Z}}$ the GMP coefficients of A.

[2] The splitting into $(g+1) \times (g+1)$ blocks is at the position of k_∞. Therefore, to simplify the definition we assume that $k_\infty = g+1$.

Let us note some special properties of GMP matrices. $A \in \mathrm{GMP}(\mathbf{C}_{g+1})$ is a $2g+3$ diagonal matrix. However, on the most outer diagonal only one entry in each block is non-vanishing, namely

$$p_g^{(j)} = \langle Ae_{2g+1+j(g+1)}, e_{g+j(g+1)}\rangle > 0. \tag{1.40}$$

However, in contrast to Jacobi matrices, where taking higher powers of J increases the number of non-trivial diagonals, by forming resolvents $(\mathbf{c}_\ell - A)^{-1}$, GMP matrices stay in the same class. Only the position of the non-vanishing entry at the most outer diagonal is shifted.

For $\mathbf{C}_{g+1}$ and E fixed, we define the isospectral torus of full-line GMP matrices by

$$\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{g+1}) = \{A \in \mathrm{GMP}(\mathbf{C}_{g+1}) \mid A \text{ is g+1 periodic and } \sigma(A) = \mathsf{E}\}. \tag{1.41}$$

For an arbitrary choice of poles, $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{g+1})$ may be empty. However, the Ahlfors function provides a choice of poles so that $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{g+1})$ is always non-empty. If $\mathbf{c}_k$, for $1 \le k \le g+1$, denote the zeros of the Ahlfors function associated to the set E as described in Theorem 1.14, then we use the notation $\mathbf{C}_{\mathsf{E}} = (\mathbf{c}_1 \cdots, \mathbf{c}_g, \infty)$. In this case, $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}})$ is up to a certain identification in bijective correspondence to the isospectral torus of almost periodic Jacobi matrices $\mathcal{T}_{\mathsf{E}}^J$ as defined in (1.35). Moreover, the Yuditskii discriminant allows to characterize $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}})$ in terms of a magic formula:

$$A \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}) \iff \Delta_{\mathsf{E}}(A) = S^{g+1} + S^{-(g+1)}.$$

With this in hand, the Damanik–Killip–Simon approach is applicable for GMP matrices and is successfully applied in [53] to prove Theorem 1.13 for spectral sets E being an arbitrary union of intervals. Again it is proved following the Damanik–Killip–Simon approach:

1. A block Jacobi version of Theorem 1.12 exists already from Damanik–Killip–Simon [13],
2. For a given measure μ associate a Jacobi matrix J and a GMP matrix A and apply the block version of Theorem 1.12 to the block Jacobi matrix $\Delta_{\mathsf{E}}(A)$,
3. Translating the measure condition is the same as above,
4. Translate the condition $\Delta_{\mathsf{E}}(A) - (S^{g+1} + S^{-(g+1)})$ is a Hilbert-Schmidt operator back to conditions for J.

Again, step 4 is the technically most challenging part and is resolved by introducing the Jacobi flow, which is a certain flow on GMP matrices corresponding to the shift action $S^{-1}JS$ on the level of Jacobi matrices.

We decided to provide a detailed account on this method for two reasons. First of all, we believe that it is a strong method in the spectral theory of Jacobi matrices. Secondly, we successfully applied it in order to settle a conjecture of Simon on regular Jacobi matrices and the Cesàro–Nevai class, which is part of the habilitation thesis and will be presented below.

1.1.2 Stahl–Totik regularity and a conjecture of Simon

In this section, we recall the application of potential theory to the asymptotic analysis of extremal polynomials. Standard references on potential theory are e.g. [1, 37]. Potential theory in $\mathbb{C}$ is concerned with the study of subharmonic potentials

$$\Phi_\rho(z) = \int \log|z - x| \mathrm{d}\rho(x),$$

for some compactly supported probability measure ρ. They enter naturally in studying growth of polynomials as the degree $n \to \infty$, due to the simple observation that if $P_n(z) = \prod_{j=1}^n (z - z_j(n))$, then

$$\frac{1}{n} \log |P_n(z)| = \Phi_{\rho_n}(z),$$

where

$$\rho_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{z(j)}, \tag{1.42}$$

denotes the normalized zero counting measure. We will assume that $K \subset \mathbb{C}$ is compact and of positive logarithmic capacity, Cap_K, and define Ω to be the unbounded connected component of $\overline{\mathbb{C}} \setminus K$. Therefore, there exists a Green function, $G_K(z, \infty)$, with pole at ∞. It can be represented in terms of an logarithmic potential by

$$G_K(z, \infty) = \int_K \log |z - x| d\rho_K(x) + \gamma_K, \tag{1.43}$$

where $\gamma_K = -\log(\mathrm{Cap}_K)$ denotes the Robin constant and ρ_K the equilibrium measure of the set K.

Let μ be a non-trivial, real probability measure with compact support and denote by E its essential support (the support with isolated points removed)

$$\mathsf{E} = \mathrm{ess\,supp}(\mu).$$

Let $((a_{n+1}, b_n))_{n=0}^{\infty}$ denote the associated Jacobi parameters. Then

$$\limsup_{n\to\infty} \left(\prod_{\ell=1}^{n} a_\ell \right)^{1/n} \leq \mathrm{Cap}_{\mathsf{E}}\,. \tag{1.44}$$

The measure μ is called *Stahl-Totik regular*, if

$$\lim_{n\to\infty} \left(\prod_{\ell=1}^{n} a_\ell \right)^{1/n} = \mathrm{Cap}_{\mathsf{E}}\,.$$

Stahl–Totik regularity is equivalent to root asymptotics of orthonormal polynomials:

Theorem 1.16. *Let μ be a non-trivial, real probability measure with compact support, $\mathsf{E} = \mathrm{ess\,supp}(\mu)$ and $(p_n)_{n=0}^{\infty}$ the associated orthonormal polynomials. Then μ is Stahl-Totik regular if and only if*

$$\lim_{n\to\infty} |p_n(z)|^{1/n} = G_{\mathsf{E}}(z, \infty),$$

uniformly on compact subsets of $\mathbb{C} \setminus [\min \mathsf{E}, \max \mathsf{E}]$.

Let J_+ be a half-line Jacobi matrix with Jacobi parameters $((a_{n+1}, b_n))$. J_+ belongs to the Nevai class of $\mathsf{E} = [-2, 2]$, if $a_n \to 1$, $b_n \to 0$. It holds that $\mathrm{Cap}_{[-2,2]} = 1$. Therefore, since by Weyl's theorem for elements of the Nevai class it holds that $\sigma_{\mathrm{ess}}(J_+) = [-2, 2]$, the Nevai condition implies regularity. The converse is false. However, in [43] it is shown that if $\mathrm{ess\,supp}(\mu) = [-2, 2]$ and μ is regular, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (|a_n - 1| + |b_n|) = 0. \tag{1.45}$$

Jacobi matrices satisfying (1.45) are said to belong to the Cesáro-Nevai class of $[-2, 2]$; cf. [25, 41, 43]. Let E be a finite-gap set and $\mathcal{T}_{\mathsf{E}}^{J_+}$ denote the isospectral torus of half-line Jacobi matrices. A half-line Jacobi matrix J_+ is said to belong to the Cesáro–Nevai class of E if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} d((S_+^*)^n J_+ (S_+)^n, \mathcal{T}_{\mathsf{E}}^{J_+}) = 0.$$

Applying the Damanik–Killip–Simon approach, Simon extended it to spectra of periodic Jacobi matrices with all gaps open.

Theorem 1.17. *Let E be the spectrum of a periodic Jacobi matrix with all gaps open. If μ is Stahl–Totik regular on E, then the associated Jacobi matrix belongs to the Cesáro–Nevai class of E.*

Motivated by this, Simon conjectured that Theorem 1.17 extends to arbitrary finite-gap sets.

Conjecture 1. *Let E be a finite-gap set. If μ is Stahl–Totik regular on E, then the associated Jacobi matrix belongs to the Cesáro–Nevai class of E.*

In [19], we developed a theory of Stahl–Totik regularity for orthogonal rational functions and settled the Simon conjecture by applying the Damanik–Killip–Simon approach with the Yuditskii discriminant.

1.1.3 Orthogonal rational functions with real poles, root asymptotics, and GMP matrices

The theory of Stahl–Totik regularity as presented in Section 1.1.2 assigns a special role to the interior point ∞ of the domain $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$, as it is formulated in terms of orthonormal polynomials and the associated potential theoretic quantities Green function at ∞, the capacity and the equilibrium measure. In particular, due to this special role of ∞, it is a non-trivial question, whether the push-forward of a Stahl–Totik regular measure with respect to a Möbius transform that does not preserve ∞ is again Stahl–Totik regular. Motivated by Conjecture 1 we were seeking in [19] for a more conformal invariant theory of Stahl–Totik regularity.

As a direct consequence of our construction, we obtain that the property of being Stahl–Totik regular is conformally invariant in the following sense: The set of Möbius transformations which preserve $\overline{\mathbb{R}}$ is the semidirect group product $\mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}$, whose normal subgroup $\mathrm{PSL}(2,\mathbb{R})$ corresponds to the orientation preserving case. Denote by $f_*\mu$ the pushforward of μ, defined by $(f_*\mu)(A) = \mu(f^{-1}(A))$ for Borel sets A.

Theorem 1.18. *Let $f \in \mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}$. If μ is a Stahl–Totik regular measure on $\mathbb{R}$ and $\infty \notin \mathrm{supp}(f_*\mu)$, then the pushforward measure $f_*\mu$ is also Stahl–Totik regular.*

Fix a non trivial probability measures μ with arbitrary $\operatorname{supp}\mu \subset \overline{\mathbb{R}}$ and a finite sequence with no repetitions, $\mathbf{C}_{g+1} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ with $\mathbf{c}_k \in \overline{\mathbb{R}} \setminus \operatorname{supp}\mu$ for all k. Let $(\tau_n)_{n=0}^{\infty}$ denote the sequence of orthogonal rational functions obtained as maximizers of Problem 1 for the sequence of poles obtained from $\mathbf{C}_{g+1}$ by periodic extension. Let

$$\mathsf{E} = \operatorname{ess\,supp}\mu$$

and assume for simplicity that E is not polar. Denote by $G_{\mathsf{E}}(z,z_0)$ the Green function in $\Omega=\overline{\mathbb{C}}\setminus\mathsf{E}$ with pole at z_0 and define

$$\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}_{g+1})=\frac{1}{g+1}\sum_{k=1}^{g+1}G_{\mathsf{E}}(z,\mathbf{c}_k). \tag{1.46}$$

The growth rate of τ_n can be bounded from below in terms of $\mathcal{G}_{\mathsf{E}}$:

Theorem 1.19. *For all $z\in\mathbb{C}\setminus\mathbb{R}$,*

$$\liminf_{n\to\infty}|\tau_n(z)|^{1/n}\geq e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}_{g+1})}.$$

Note that for orthonormal polynomials $(a_1\dots a_n)^{-1}$ appearing in (1.44) is the leading coefficient of p_n. Recall from Problem 1 that

$$\kappa_n=\sup\{\operatorname{Re}\lambda\mid\exists h\in\mathcal{L}_{n-1},\ \|\lambda r(\cdot,\mathbf{c}_n)^{d_n}+h\|_{L^2(\mu)}\leq 1\}.$$

Since $\tau_n=\kappa_n r(\cdot,\mathbf{c}_n)^{d_n}+h$ for $h\in\mathcal{L}_{n-1}$, κ_n is explicitly characterized as a kind of leading coefficient for τ_n with respect to the pole at $\mathbf{c}_k$ where $n=j(g+1)+k$, $1\leq k\leq g+1$. Below, we will also relate the constants κ_n to off-diagonal coefficients of associated GMP matrices leading to a generalization of (1.44) for orthogonal rational functions.

It is a basic property of the Green function that the limits

$$\gamma_{\mathsf{E}}^k=\begin{cases}\lim_{z\to\mathbf{c}_k}(G_{\mathsf{E}}(z,\mathbf{c}_k)+\log|z-\mathbf{c}_k|), & \mathbf{c}_k\neq\infty\\ \lim_{z\to\mathbf{c}_k}(G_{\mathsf{E}}(z,\mathbf{c}_k)-\log|z|), & \mathbf{c}_k=\infty\end{cases}$$

exist. Note that, if $\mathbf{c}_k=\infty$, γ_{E}^k is precisely the Robin constant for the set E. We further define constants λ_k by

$$\log\lambda_k=\gamma_{\mathsf{E}}^k+\sum_{\substack{1\leq\ell\leq g+1\\ \ell\neq k}}G_{\mathsf{E}}(\mathbf{c}_k,\mathbf{c}_\ell).$$

Theorem 1.20. *For all $1\leq k\leq g+1$, for the subsequence $n(j)=j(g+1)+k$,*

$$\liminf_{j\to\infty}\kappa_{n(j)}^{1/n(j)}\geq\lambda_k^{1/(g+1)}.$$

Again Stahl–Totik regularity is now characterized through equality:

Theorem 1.21. *The following are equivalent:*

(i) For some $1\leq k\leq g+1$, for the subsequence $n(j)=j(g+1)+k$,

$$\lim_{j\to\infty}\kappa_{n(j)}^{1/n(j)}=\lambda_k^{1/(g+1)},$$

(ii) For all $1\leq k\leq g+1$, for the subsequence $n(j)=j(g+1)+k$,

$$\lim_{j\to\infty}\kappa_{n(j)}^{1/n(j)}=\lambda_k^{1/(g+1)},$$

(iii)

$$\lim_{n\to\infty}\left(\prod_{\ell=1}^{g+1}\kappa_{n+\ell}\right)^{1/n}=\left(\prod_{k=1}^{g+1}\lambda_k\right)^{1/(g+1)},$$

(iv) *For q.e.* $z \in \mathsf{E}$*, we have* $\limsup_{n\to\infty} |\tau_n(z)|^{1/n} \le 1$*,*

(v) *For some* $z \in \mathbb{C}_+$*,* $\limsup_{n\to\infty} |\tau_n(z)|^{1/n} \le e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}_{g+1})}$*,*

(vi) *For all* $z \in \mathbb{C}$*,* $\limsup_{n\to\infty} |\tau_n(z)|^{1/n} \le e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}_{g+1})}$*,*

(vii) *Uniformly on compact subsets of* $\mathbb{C}\setminus\mathbb{R}$*,* $\lim_{n\to\infty} |\tau_n(z)|^{1/n} = e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}_{g+1})}$*.*

Definition 1.22. The measure μ is $\mathbf{C}_{g+1}$-regular if it obeys one (and therefore all) of the assumptions of Theorem 1.21.

For the case $\mathbf{C}_1 = (\infty)$ this is exactly Stahl–Totik regularity for orthogonal polynomials. However, for multiple poles we point out that there is a phenomena, which has no analog for orthogonal polynomials. Namely, control of the growth rate of $\kappa_{n(j)}$ along a single subsequence allows to control the full sequence. This will be crucial below in order to characterize Stahl–Totik regularity by only using entries of a GMP matrix A and not its resolvents.

Moreover, we show that the regular behavior described by Theorem 1.21 is independent of the set of poles $\mathbf{C}_{g+1}$:

Theorem 1.23. *Let* $g, \tilde{g} \in \mathbb{N}_0$ *and* $\mathbf{C}_{g+1}, \tilde{\mathbf{C}}_{\tilde{g}+1}$ *be two finite sequences of elements from* $\overline{\mathbb{R}}\setminus\operatorname{supp}\mu$*, not necessarily of the same length. Then* μ *is* $\mathbf{C}_{g+1}$*-regular if and only if it is* $\tilde{\mathbf{C}}_{\tilde{g}+1}$*-regular.*

Corollary 1.24. *Let* $\operatorname{supp}\mu \subset \mathbb{R}$*. Let* $\mathbf{C}_{g+1}$ *be a finite sequence of elements from* $\overline{\mathbb{R}}\setminus\operatorname{supp}\mu$*. Then* μ *is* $\mathbf{C}_{g+1}$*-regular if and only if it is Stahl–Totik regular.*

Theorem 1.23 shows that Stahl–Totik regularity is indeed a property of the measure and not of the particular chosen sequence of rational functions. Thus, Theorem 1.21 should not be seen as describing equivalent conditions for a new class of measures, but rather a new set of regular behaviors for the familiar class of Stahl–Totik regular measures.

In the case of orthonormal polynomials p_n, it is known that p_n has at most one zero in any connected component of $\overline{\mathbb{R}}\setminus\operatorname{supp}\mu$ and no zero in the connected component of $\overline{\mathbb{R}}\setminus\operatorname{supp}\mu$ containing ∞. For orthogonal rational functions the situation is similar. Write

$$\tau_n(z) = \frac{P(z)}{R_n(z)}, \tag{1.47}$$

where R_n is defined as in (1.2). Similar as for polynomials, one can show that P has at most one simple zero in each connected component of $\overline{\mathbb{R}}\setminus\operatorname{supp}\mu$ and no zero in the connected component which contains $\mathbf{c}_k$. However, it can happen that $P(\mathbf{c}_j) = 0$, for $j \neq k$. If this happens, we call $\mathbf{c}_j$ a "generalized zero" of τ_n. In this case, due to cancellation, the degree of τ_n is decreased. However, since this can happen at most for all $\mathbf{c}_j \neq \mathbf{c}_k$ and zeros are simple, we always have

$$n - g \le \deg\tau_n \le n. \tag{1.48}$$

Considering the case $\mathbf{C}_2 = (\infty, 0)$ and a measure which is symmetric with respect to $z \mapsto -z$ shows that cancellation can indeed happen.

Next we turn towards describing the limit distribution of the normalized zero counting measure. We define

$$\nu_n = \frac{1}{n}\sum_{w:\tau_n(w)=0}\delta_w.$$

Although we normalize by n, ν_n may not be a probability measure: however, due to (1.48), $1 - g/n \leq \nu_n(\overline{\mathbb{R}}) \leq 1$. Therefore, normalizing by $\deg \tau_n$ instead of by n would not affect the limits as $n \to \infty$.

In the regular case, limits of ν_n are described in terms of the harmonic measure. Let $\omega_{\mathsf{E}}(A, z)$, denote the harmonic measure of the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$ the set $A \subset \mathsf{E}$ and $z \in \overline{\mathbb{C}} \setminus \mathsf{E}$. We define the probability measure on E

$$\rho_{\mathsf{E},\mathbf{C}_{g+1}} = \frac{1}{g+1} \sum_{j=1}^{g+1} \omega_{\mathsf{E}}(\cdot, \mathbf{c}_j). \tag{1.49}$$

Theorem 1.25. *The following holds:*

(a) If μ is $\mathbf{C}_{g+1}$ regular, then $\lim_{n\to\infty} \nu_n = \rho_{\mathsf{E},\mathbf{C}_{g+1}}$ in the w^*-topology of $C(\overline{\mathbb{R}})^*$.

(b) If $\lim_{n\to\infty} \nu_n = \rho_{\mathsf{E},\mathbf{C}_{g+1}}$ in the w^*-topology of $C(\overline{\mathbb{R}})^*$, then μ is $\mathbf{C}_{g+1}$ regular or there exists a polar set $X \subset \mathsf{E}$ such that $\mu(\overline{\mathbb{R}} \setminus X) = 0$.

We turn to applications of $\mathbf{C}_{g+1}$-regularity for self-adjoint operators. Fix a sequence $\mathbf{C}_{g+1} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ such that $\mathbf{c}_{k_\infty} = \infty$ for some $1 \leq k_\infty \leq g+1$ and let $A \in \mathbb{A}(\mathbf{C}_{g+1})$ be a half-line GMP matrix with matrix elements (1.38). Just as regularity for Jacobi matrices can be formulated in terms of its off-diagonal entries, we will show that regularity of a GMP matrix can be characterized in terms of its entries in the outermost non-trivial diagonal. Half-line GMP matrices have the same block structure as full-line GMP matrices, except, the splitting of the blocks may be shifted due to the position of $\mathbf{c}_{k_\infty}$ in $\mathbf{C}_{g+1}$. In particular, like in (1.40), the positive entry on outermost diagonal is

$$\beta_j = \langle A e_{(j+1)(g+1)+k_\infty}, e_{j(g+1)+k_\infty} \rangle$$

Theorem 1.26. *Fix a probability measure μ with $\operatorname{supp} \mu \subset \mathbb{R}$ and a sequence $\mathbf{C}_{g+1} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ with $\mathbf{c}_{k_\infty} = \infty$. Then*

$$\limsup_{j\to\infty} \left(\prod_{\ell=1}^{j} \beta_\ell \right)^{1/j} \leq \lambda_{k_\infty}^{-1}. \tag{1.50}$$

Moreover, the measure μ is Stahl–Totik regular if and only if

$$\lim_{j\to\infty} \left(\prod_{\ell=1}^{j} \beta_\ell \right)^{1/j} = \lambda_{k_\infty}^{-1}. \tag{1.51}$$

The statement follows by relating the coefficients β_j to the constants $\kappa_{k_\infty + j(g+1)}$. In the formulation of Theorem 1.26 we see the importance of Theorem 1.21 (i). If we would need convergence along all subsequence, then instead of (1.50) we would also need to consider entries of the resolvents $(\mathbf{c}_j - A)^{-1}$ for all finite $\mathbf{c}_j \in \mathbf{C}_{g+1}$. At the same time, regularity also implies a convergence result similar to (1.50) for the outermost nonvanishing entries of the resolvents, and thus in fact provides $g+1$ criteria for regularity.

With the notion of regularity developed for orthogonal rational functions we are now able to combine the Damanik–Killip–Simon approach with the Yuditskii discriminant to prove Conjecture 1:

Theorem 1.27. *Let $\mathsf{E} \subset \mathbb{R}$ be a finite-gap set and $\mathcal{T}_{\mathsf{E}}^{J_+}$ the associated isospectral torus of Jacobi matrices. If J_+ is a regular half-line Jacobi matrix with $\sigma_{\text{ess}}(J_+) = \mathsf{E}$, then J_+ belongs to the Cesáro–Nevai of E, i.e.,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} d((S_+^*)^m J_+ S_+^m, \mathcal{T}_{\mathsf{E}}^{J_+}) = 0. \tag{1.52}$$

As already mentioned in Section 1.1.2, Theorem 1.27 was proved by Simon for spectra of periodic Jacobi matrices with all gaps open. It was then proved with completely different methods by Krüger [27] under the additional assumption $\inf_n a_n > 0$. While this is a common assumption in the ergodic literature, regular Jacobi matrices do not always satisfy it: [41, Example 1.4] can easily be modified to give a regular Jacobi matrix with spectrum $[-2, 2]$ and $\inf a_n = 0$. Thus, Theorem 1.27 first settles Simon's Conjecture 1 in full generality.

1.1.4 Asymptotics of Chebyshev rational functions with respect to subsets of the real line

This section will be concerned with the sup-norm problems stated in the beginning of this introduction. Fix a compact proper subset $\mathsf{E} \subset \overline{\mathbb{R}}$ containing infinitely many points. Connected components of $\overline{\mathbb{R}} \setminus \mathsf{E}$ are called gaps of E. We fix a sequence of poles $\mathbf{C} = (\mathbf{c}_k)_{k=1}^{\infty}$ with $\mathbf{c}_k \in \overline{\mathbb{R}} \setminus \mathsf{E}$ and consider the associated Chebyshev problem

$$m_n(\mathbf{c}_n) := \sup\{\operatorname{Re} \lambda \mid \exists h \in \mathcal{L}_{n-1}, \|\lambda r(\cdot, \mathbf{c}_n)^{d_n} + h\|_{\mathsf{E}} \le 1\} \tag{1.53}$$

and for $x_* \in \overline{\mathbb{R}} \setminus (\mathsf{E} \cup \{\mathbf{c}_k : 1 \le k \le n\})$, the residual problem

$$m_n(x_*) := \sup\{\operatorname{Re} f(x_*) \mid f \in \mathcal{L}_n, \ \|f\|_{\mathsf{E}} \le 1\}, \tag{1.54}$$

where $\mathcal{L}_n$ was defined in (1.1).

In contrast to Section 1.1.3, we do not assume that the sequence is periodic. Let us point out that this assumption was natural and necessary in Section 1.1.3, since the periodic occurrences of ∞ yielded a finite band structure of the associated GMP matrices.

We will start with results describing the solutions for fixed n. It is well known that Chebyshev polynomials of degree n associated to a compact subset of $\mathbb{R}$ can be characterized by the alternation theorem. This characterization yields an explicit formula that allows to compute asymptotics as the degree tends to ∞; cf. [7, 8]. We will prove a unified version of the alternation theorem for the Chebyshev and the residual problem for rational functions. This will then allow us to prove so-called Szegő-Widom asymptotics for the extremal functions.

It will be convenient to use the notion of divisors on the Riemann sphere $\overline{\mathbb{C}}$. Divisors on $\overline{\mathbb{C}}$ are elements of the free Abelian group over $\overline{\mathbb{C}}$. They can be realized as functions $D : \overline{\mathbb{C}} \to \mathbb{Z}$ or as formal sums. Namely, if $D \in \mathbb{Z}^{\overline{\mathbb{C}}}$ then alternatively it is common to identify D with the formal sum

$$\sum_{z\in\overline{\mathbb{C}}} n_z z, \quad \text{where } n_z = D(z).$$

Note that $n_z \neq 0$ only for finitely many $z \in \overline{\mathbb{C}}$. We will mainly use the functional realization. The degree of D is the integer $\deg D = \sum_z D(z)$, and the divisor D is integral if $D(z) \ge 0$ for all z. We also write $D_1 \le D_2$, if $D_2 - D_1$ is integral and denote by $\operatorname{supp} D = \{z \in \overline{\mathbb{C}} : D(z) \neq 0\}$ the support of D. In particular, for a meromorphic nonconstant function $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, we denote its polar divisor by $(f)_\infty$; the polar divisor assigns to each pole the multiplicity of that pole, and takes zero values elsewhere. Similarly, for $w \in \mathbb{C}$, we define $(f)_w = (1/(f-w))_\infty$. The value $\deg(f)_w$ is independent of w and corresponds to the degree of f. We also follow the convention to set $(f)_w = 0$, if f is a constant. For a given sequence $\mathbf{C}$, we define the special divisor:

$$D_n^\infty(\mathbf{c}) = \#\{k : \ \mathbf{c}_k = \mathbf{c}, \ 1 \leq k \leq n\}. \tag{1.55}$$

By definition, we have $\deg D_n^\infty = n$. Any integral divisor, D, of degree n generates an $n+1$ dimensional vector space by

$$\mathcal{L}(D) = \{f : \overline{\mathbb{C}} \to \overline{\mathbb{C}} \mid f \text{ is meromorphic and } (f)_\infty \leq D\},$$

The space $\mathcal{L}_n$ from (1.1) can now equivalently be introduced by

$$\mathcal{L}_n = \mathcal{L}(D_n^\infty).$$

This allows us to present (1.53) and (1.54) in a unified way:

Problem 5. *For a real integral divisor D_n^∞ with $\deg D_n^\infty = n$ containing only points in $\overline{\mathbb{R}} \setminus \mathsf{E}$, and a point $x_* \in \overline{\mathbb{R}} \setminus \mathsf{E}$, denote $d_n = D_n^\infty(x_*)$ and $\mathcal{L}_n = \mathcal{L}(D_n^\infty)$ and find*

$$m_n(x_*) := \sup\{\operatorname{Re} \lim_{x \to x_*} \frac{F_n(x)}{r(x,x_*)^{d_n}} \mid F_n \in \mathcal{L}_n, \ \|F_n\|_{\mathsf{E}} \leq 1\}. \tag{1.56}$$

The Chebyshev problem corresponds to $d_n > 0$ and the residual problem corresponds to $d_n = 0$. Throughout this section, we work in the general setting of Problem 5. For fixed n, we will call the gap containing x_* the extremal gap (for n).

The Chebyshev alternation theorem compresses in a very elegant way several crucial properties of Chebyshev polynomials associated to subsets of $\mathbb{R}$. Let $\mathsf{E} \subset \mathbb{R}$ be compact. The Chebyshev alternation theorem states that the Chebyshev polynomial, T_n, of degree n is uniquely characterized by the property, that there are $x_1 < \cdots < x_{n+1}$, $x_i \in \mathsf{E}$, such that

$$T_n(x_j) = (-1)^{n+1-j}.$$

This characterization can be understood in the following way: T_n has n simple real zeros all belonging to $[\min \mathsf{E}, \max \mathsf{E}]$. Between consecutive zeros, there is at least one point belonging to E on which $|T_n|$ attains the sup-norm of T_n, leading to $n-1$ points of alternation. In addition, there is one at each gap edge of the extremal gap (in this case the one containing ∞) which sums up to $n+1$ points of alternation. In particular x_1 and x_{n+1} will always be counted, because of the natural order of $\mathbb{R}$. To account for that in the general setting, we use the following notation:

Definition 1.28. For a sequence $(t_j)_{j=0}^m$ in $\overline{\mathbb{R}}$ with $m \geq 2$, we say that the sequence is cyclically ordered if it has no repetitions and there exists $f \in \mathrm{PSL}(2,\mathbb{R})$ such that $f(t_0) = \infty$ and $f(t_1) < f(t_2) < \cdots < f(t_m)$. We will also use cyclic interval notation: for distinct $a, b \in \overline{\mathbb{R}}$, we denote

$$(a,b) = \{c \mid (a,c,b) \text{ is cyclically ordered}\}, \qquad [a,b] = \{a,b\} \cup (a,b).$$

This gives a well-defined cyclic order, since $\mathrm{PSL}(2,\mathbb{R})$ transformations preserve orientation on $\overline{\mathbb{R}}$.

In the setting of rational functions the counting for the alternation theorem is essentially more delicate, and the relative ordering of the poles and alternation points play an important role. The reason for this is that if between two zeros there is a gap with a pole $\mathbf{c}_j$, then the sign at the next gap edge depends on the parity of the pole. This makes it necessary to define the following *sign function*:

$$S_n(x) = \sum_{\substack{1\le k\le n \\ \mathbf{c}_k\neq x_*}} \chi_{[x_*,\mathbf{c}_k)}(x) = \sum_{\mathbf{c}\in\overline{\mathbb{R}}\setminus\{x_*\}} D_n^\infty(\mathbf{c})\chi_{[x_*,\mathbf{c})}(x).$$

Recall that a function F is called real if for all $z\in\mathbb{C}$, $\overline{F(\overline{z})} = F(z)$.

Definition 1.29. For a real function $F\in\mathcal{L}_n$ with $\|F\|_{\mathsf{E}}\le 1$, a set of distinct points $x_1,\dots,x_m\in\mathsf{E}$ such that the sequence $(x_*,x_1,\dots,x_m)$ is cyclically ordered and satisfies the following alternation property

$$F(x_j) = (-1)^{m-j-S_n(x_j)} \tag{1.57}$$

for all $j=1,\dots,m$ is called an alternation set. We say that F has a maximal alternation set if $m=n+1$.

It should be noted that the notion of alternation set depends on the function F, the class $\mathcal{L}_n$, the set E, and the reference point x_*. We note that in what follows, whenever we refer to extremal functions, we mean this in the sense of Problem 5.

Theorem 1.30 (Alternation theorem)**.** *A real function $F\in\mathcal{L}_n$ with $\|F\|_{\mathsf{E}}\le 1$ is an extremal function if and only if it has a maximal alternation set.*

These results generalize standard results from the polynomial case: in the Chebyshev polynomial case, $S_n(x)\equiv 0$, and in the residual polynomial case, S_n has one jump which may or may not affect the alternation criterion, depending on degree. The case of Chebyshev rational functions was also previously formulated in [31]. In all the real extremal problems, previously considered in the literature, the extremizer is seen to be nonconstant. However, in the setting of residual rational functions, the extremizer can be a constant function, and the alternation theorem lets us characterize when this happens:

Theorem 1.31. *The extremal function F_n is constant if and only if the divisor D_n^∞ is of the form (1.55) for points $\mathbf{c}_1,\dots,\mathbf{c}_n$ such that the points $x_*,\mathbf{c}_1,\mathbf{c}_2,\dots,\mathbf{c}_n$ are in $n+1$ distinct gaps of E.*

In particular, for the Chebyshev problem, $x_*=\mathbf{c}_n$ so F_n is always nonconstant. Theorem 1.31 can be understood in the following way. Let's write again

$$F_n = \frac{P}{R_n},$$

where R_n is defined by (1.2). As for orthogonal rational functions, P has at most one simple zero in each gap of E and no zero in the extremal gap. In particular, P can have a zero at $\mathbf{c}_k$, which is then called a "generalized zero". Theorem 1.31 describes the case that P has n generalized zeros[3]. This is made again precise in the language of divisors. We call a point x a "generalized zero" of F_n if either $(F_n)_0(x)>0$ or if

$$D_n^\infty(x) - (F_n)_\infty(x) > 0.$$

[3] Where $\deg P = n-1$ is viewed as a generalized zero at ∞

Thus, this notion includes both actual zeros of F_n and places where there is a reduction in the order of the pole compared to the maximal allowed order. These generalized zeros are precisely counted by the divisor

$$D_n^0 := (F_n)_0 + D_n^\infty - (F_n)_\infty.$$

We turn to root asymptotics of F_n. From $\deg(F_n)_0 = \deg(F_n)_\infty$ it follows that

$$\deg D_n^0 = \deg D_n^\infty = n$$

so we can define the normalized pole counting measure

$$\mu_n := \frac{1}{n}\sum_{\mathbf{c}} D_n^\infty(\mathbf{c})\delta_{\mathbf{c}}$$

and normalized generalized zero counting measure

$$\nu_n := \frac{1}{n}\sum_{\mathbf{c}} D_n^0(\mathbf{c})\delta_{\mathbf{c}}.$$

The behavior of $\log|F_n|$ is governed by the zero and pole distributions. This corresponds to two Riesz representations, with $\log|F_n|$ superharmonic (respectively, subharmonic) away from the set of zeros (respectively, poles). The limiting pole distribution μ directly determines the root asymptotics of the functions F_n and the limiting zero distribution.

We assume that E is not a polar set and denote by $G_{\mathsf{E}}(z,w)$ the Green function and by $\omega_{\mathsf{E}}(dx,z)$ harmonic measure for this domain.

Theorem 1.32 (Root asymptotics)**.** *Assume that* E *is not a polar set,* $\overline{\bigcup_{n\geq 1}\operatorname{supp} D_n^\infty} \cap \mathsf{E} = \emptyset$*,* $\lim \mu_n = \mu$ *in the* w^**-topology of* $C(\overline{\mathbb{R}})^*$ *and* $(x_n^*)_{n=0}^\infty$ *be a sequence in* $\overline{\mathbb{R}} \setminus \mathsf{E}$ *not accumulating on* E*. Then uniformly on compact subsets of* $\mathbb{C} \setminus \mathbb{R}$*,*

$$\lim_{n\to\infty} \frac{1}{n}\log|F_n(z)| = \int G_{\mathsf{E}}(z,x)d\mu(x). \tag{1.58}$$

Moreover,

$$\lim \nu_n = \int \omega_{\mathsf{E}}(dz,x)d\mu(x).$$

in the w^**-topology of* $C(\overline{\mathbb{R}})^*$*.*

For the special case of a periodic sequence $\mathbf{c}_1, \cdots, \mathbf{c}_p$, we have $\mu = \frac{1}{p}\sum_{j=1}^p \delta_{\mathbf{c}_j}$ and (1.58) resembles (1.49).

Recall that in the setting of orthogonal rational functions we proved in Theorem 1.19 a universal lower bound for the extremal functions and Stahl–Totik regularity is equivalent to existence of the same upper bound leading to root asymptotics of the extremal functions. As Theorem 1.32 shows, the situation is different in the L^∞-setting. We will again show a universal lower bound. However, due to explicit representations of F_n in terms of the so-called n-extension $\mathsf{E}_n = F_n^{-1}([-1,1])$, we obtain a Bernstein-Walsh type upper bound for F_n in terms of the Green functions $G_{\mathsf{E}}(z,\mathbf{c})$. Combining this with the lower bound proves root asymptotics for F_n.

We also prove so-called Szegő–Widom asymptotics for F_n. For the sake of this introduction, we only describe heuristically what this type of notion means and refer to the original article for the precise results. If $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$ is simply connected, let $B_{\mathsf{E}}(\cdot,\mathbf{c})$ denote the Riemann map, which maps Ω conformally onto $\mathbb{D}$ normalized so that $B_{\mathsf{E}}(\mathbf{c},\mathbf{c}) = 0$ and $\partial_z B_{\mathsf{E}}(z,\mathbf{c})|_{z=\mathbf{c}} > 0$. It can be seen that

$$B_{\mathsf{E}}(z,\mathbf{c}) = e^{-G_{\mathsf{E}}(z,\mathbf{c}) - i\widetilde{G_{\mathsf{E}}(z,\mathbf{c})}}, \tag{1.59}$$

where $\widetilde{G_{\mathsf{E}}(z,\mathbf{c})}$ denotes the harmonic conjugate of $G_{\mathsf{E}}(z,\mathbf{c})$.

Considering

$$F_n\left(\prod_{\mathbf{c}} D_n^{\infty}(\mathbf{c})B_{\mathsf{E}}(z,\mathbf{c})\right), \tag{1.60}$$

yields an analytic, in fact normal family and one can describe convergent subsequence in the topology of local uniform convergence in Ω. If Ω is not simply connected, $B_{\mathsf{E}}(z,\mathbf{c})$ can still be defined locally by (1.59) and then using the monodromy theorem extended to a global multivalued analytic function in Ω. The corresponding subsequential limits of (1.60) are then described in terms of extremal solutions of a corresponding $H^{\infty}(\Omega)$ problem of so-called character automorphic functions. This notion goes back to the seminal paper of Widom [51] and is thus referred to Szegő-Widom asymptotics. In [20] we prove Szegő–Widom asymptotics for F_n associated for domains $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$ which are regular Parreau–Widom domains, so that the Direct Cauchy theorem holds in Ω. These domains are defined in [20, Section 4], but let us mention here that it also includes infinitely connected domains. This generalizes the results of [7] to Chebyshev and residual rational functions.

1.1.5 Finite-Gap CMV Matrices: Periodic Coordinates and a Magic Formula

It was mentioned in Section 1.1.1 that GMP matrices provide periodic coordinates for generically almost periodic Jacobi matrices associated to finite-gap sets. This was then used for instance in the proof of Theorem 1.27 to extend results from the spectral set $[-1,1]$ to spectral sets which are arbitrary finite-gap sets. The unitary analog of Jacobi matrices are called CMV matrices. Again, the isospectral torus of CMV matrices associated to a finite union of arcs of the unit circle generically consists of almost periodic operators. We will again use the terminology finite-gap sets. The goal of this section is to find periodic coordinates and prove a magic formula for CMV matrices belonging to the isospectral torus of finite-gap sets.

We start with the motivation for and definition of CMV matrices. Let μ be a non-trivial probability measure supported on $\partial\mathbb{D} = \{z \in \mathbb{C} \mid |z| = 1\}$ and ϕ_n the associated orthonormal polynomials. If M_z denotes the multiplication operator in $L^2(\mu)$, then $M_z^* = M_{1/z}$. We still have that $\langle M_z\phi_j, \phi_k\rangle = 0$, if $k > j+1$, but for $k \le j+1$, the corresponding entries may be different from zero. In particular, the corresponding matrix representation is not of finite band structure. To overcome this, Cantero, Moral and Velázquez (CMV) [5] suggested to instead orthogonalize the sequence $(1, z^{-1}, z, z^{-2}, z^2, \cdots)$. The multiplication operator in this basis is a five diagonal matrix of a special structure that will be defined below and the corresponding operators on $\ell^2(\mathbb{N}_0)$ are called CMV-matrices; cf [38, Chapter 4]. In the language of divisors, this means to consider the orthogonal rational functions associated to periodic repetition of $\mathbf{C}_2 = (0,\infty)$. Denoting the reflection with respect to the unit circle by $\hat{z} = 1/\overline{z}$, $z \in \overline{\mathbb{C}}$, then $\mathbf{C}_2 = (z_0, \hat{z}_0)$, with $z_0 = 0$. This perspective will be important in the following. We recall the structure of a half-line CMV matrices. Fix a sequence $(a_k)_{k=0}^{\infty} \in \mathbb{D}^{\mathbb{N}_0}$, set $\rho_k = \sqrt{1-a_k^2}$ and define

$$C_+ = \begin{pmatrix} \overline{a_0} & \overline{a_1}\rho_0 & \rho_1\rho_0 & 0 & & \\ \rho_0 & -\overline{a_1}a_0 & -\rho_1 a_0 & 0 & & \\ 0 & \overline{a_2}\rho_1 & -\overline{a_2}a_1 & \overline{a_3}\rho_2 & \rho_3\rho_2 & 0 \\ 0 & \rho_2\rho_1 & -\rho_2 a_1 & -\overline{a_3}a_2 & -\rho_3 a_2 & 0 \\ & 0 & 0 & \overline{a_4}\rho_3 & -\overline{a_4}a_3 & \\ & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

The structure in the first two rows is slightly different due to an "initial condition". Starting from the third row, the structure repeats two-periodically and can thus be used to also define full-line CMV matrices acting on $\ell^2(\mathbb{Z})$. Thus, the matrix representation of a full-line CMV matrix is of the form

$$C = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ \rho_{2l-1}\overline{a_{2l}} & -\overline{a_{2l}}a_{2l-1} & \overline{a_{2l+1}}\rho_{2l} & \rho_{2l}\rho_{2l+1} & & \\ \rho_{2l}\rho_{2l-1} & -\rho_{2l}a_{2l-1} & -\overline{a_{2l+1}}a_{2l} & -\rho_{2l+1}a_{2l} & & \\ & & \rho_{2l+1}\overline{a_{2l+2}} & -\overline{a_{2l+2}}a_{2l+1} & \overline{a_{2l+3}}\rho_{2l+2} & \rho_{2l+2}\rho_{2l+3} \\ & & \rho_{2l+2}\rho_{2l+1} & -\rho_{2l+2}a_{2l+1} & -\overline{a_{2l+3}}a_{2l+2} & -\rho_{2l+3}a_{2l+2} \\ & & & \ddots & \ddots & \ddots \end{pmatrix}.$$

We say a CMV matrix is (almost) periodic, if this property holds for the generating sequence $(a_k)_{k\in I}$, where $I = \mathbb{N}_0$ or $I = \mathbb{Z}$. Let $\mathsf{E} \subset \partial\mathbb{D}$ be a finite union of circular arcs. The isospectral torus of CMV matrices it defined by

$$\mathcal{T}_{\mathsf{E}}^{\mathrm{CMV}} = \{C \mid C \text{ is an almost periodic full-line CMV matrix}, \sigma(C) = \sigma_{\mathrm{ac}}(C) = \mathsf{E}\}.$$

For a definition comparable to (1.35) see [6, eq. (1.34)]. As for Jacobi matrices, it is a property of the set E, whether $\mathcal{T}_{\mathsf{E}}^{\mathrm{CMV}}$ consists of periodic or merely almost periodic CMV matrices. One goal of this section is to find a class of operators that allows to give a parametrization of $\mathcal{T}_{\mathsf{E}}^{\mathrm{CMV}}$ by periodic operators.

Recall that CMV matrices arise as the matrix representation for the multiplication operator in $L^2(\mu)$ with respect to orthogonal rational functions for periodic repetitions of poles $\mathbf{C}_2 = (z_0, \hat{z}_0)$, with $z_0 = 0$. The following generalizations seems natural. Fix a non-trivial probability measure on $\partial\mathbb{D}$ and $(z_j)_{j=1}^{\infty} \in \mathbb{D}$, with $|z_j| \leq K < 1$. Velázquez showed that the structure of the multiplication operator $M_{z,\mu}$ in $L^2(\mu)$ with respect to the orthogonal rational functions for the sequence of poles $\mathbf{C} = (z_1, \hat{z}_1, z_2, \hat{z}_2, \dots)$ is related to CMV matrices via an operator Möbius transform; specifically, denoting by $D_+ := \mathrm{diag}_{\mathbb{N}_0}(0, z_1, z_1, z_2, z_2, \dots)$, he showed that $M_{z,\mu}$ is unitarily equivalent to

$$b_{-D_+}(C_+) := \eta_{D_+}(1 + C_+ D_+^*)^{-1}(D_+ + C_+)\eta_{D_+}^{-1}, \quad \eta_{D_+} = \sqrt{1 - D_+ D_+^*} \tag{1.61}$$

for some half-line CMV matrix C_+; cf. [50, Theorem 5.1]. This theorem suggests we should study this new class of operator Möbius transforms of CMV matrices more closely. However, as for GMP matrices, periodic occurrences of $z_0 = 0$ will be important for a finite band structure and thus we will restrict our attention to periodic sequences of poles:

Definition 1.33 (MCMV matrices)**.** Fix $n \geq 1$, let $(z_k)_{k=0}^{n-1} \in \mathbb{D}^n$ with $z_0 = 0$, $\mathbf{C}_{2n} = (z_0, \hat{z}_0, z_1, \hat{z}_1, \dots, \hat{z}_{n-1})$, $(a_k)_{k\in\mathbb{Z}} \in \mathbb{D}^{\mathbb{Z}}$, and $\vartheta \in \mathbb{R}/2\pi\mathbb{Z}$. Denote by D_0 the $2n$-periodic diagonal matrix

$$D_0 := D_0(\vec{z}) = \mathrm{diag}_{\mathbb{Z}}(\dots, z_{n-2}, z_{n-1}, z_{n-1}, z_0 \mid z_0, z_1, z_1, z_2, \dots), \tag{1.62}$$

let $\Lambda_k(\vartheta)$ be the $2n \times 2n$ diagonal matrix $\Lambda_k(\vartheta) := \operatorname{diag}_{2n}(e^{ik\vartheta}, e^{-ik\vartheta}, \dots, e^{ik\vartheta}, e^{-ik\vartheta})$, and define

$$\Lambda(\vartheta) := \bigoplus_{k\in\mathbb{Z}} \Lambda_k(\vartheta),$$

where $\Lambda_k(\vartheta)$ acts on $\operatorname{span}\{e_k, e_{k+1}\}$ in $\ell^2(\mathbb{Z})$. With this notation, the (whole-line) MCMV matrix for $\vec{z}$, (a_k), and ϑ is defined by

$$A := A((a_k), \vartheta; \mathbf{C}_{2n}) = \Lambda(\vartheta)^* b_{\text{-}D_0}(C)\Lambda(\vartheta), \tag{1.63}$$

where $C = C((a_k))$ is the CMV matrix associated to $(a_k)_{k\in\mathbb{Z}}$.

The vertical bar in (1.62) denotes the splitting of $\ell^2(\mathbb{Z}) = \ell_- \oplus \ell^2_+$. The operator Möbius transformation $C \mapsto b_{\text{-}D_0}(C)$, defined on operators acting on $\ell^2(\mathbb{Z})$ is defined exactly as in (1.61) with D_0 replaced by D_+. For finite vectors $(z_k)_{k=0}^{n-1}$ we will henceforth always assume that $z_0 = 0$. Finally, the meaning of the conjugation by $\Lambda(\vartheta)$ will be clarified below; in short, this diagonal matrix enables us to change from periodicity up to a rotational phase to pure periodicity.

Like CMV matrices, an MCMV matrix A is again band-structured. If we split A into $2n\times 2n$ blocks $\mathbf{A_{ij}}$, then $\mathbf{A_{ij}} = \mathbf{0}$ if $|\mathbf{i}-\mathbf{j}| > 1$. Moreover, the off-diagonal blocks are of the form $\mathbf{A_{i,i-1}} = \mathbf{v^i}\delta_{2n-1}^\intercal$ and $\mathbf{A_{i,i+1}} = \mathbf{u^i}\delta_0^\intercal$ for some explicit vectors $\mathbf{u^i}, \mathbf{v^i} \in \mathbb{C}^{2n}$, i.e., A is of the form

$$A = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ \mathbf{0} & \mathbf{0}\ \ \mathbf{v^i} & \mathbf{A_{ii}} & \mathbf{u^i}\ \ \mathbf{0} & \mathbf{0} \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

Furthermore, since operator Möbius transforms preserve unitarity, MCMV matrices are likewise unitary operators. Thus MCMV matrices can be viewed as being "block-CMV". This special structure does not hold for arbitrary operator Möbius transforms of CMV matrices; it follows in our case from D_0 having periodically repeated zero entries.

As in Section 1.1.1 we start with discussing the spectrum of periodic MCMV matrices. For a fixed collection of poles $\mathbf{C}_{2n}$ as in Definition 1.33, we define the collection of all MCMV matrices by

$$\mathbb{A}(\mathbf{C}_{2n}) := \{A((a_k)), \vartheta; \mathbf{C}_{2n}) \mid (a_k) \in \mathbb{D}^{\mathbb{Z}},\ \vartheta \in \mathbb{R}/2\pi\mathbb{Z}\} \tag{1.64}$$

and give special consideration to the subset $\mathbb{A}_{\text{per}}(\mathbf{C}_{2n}) \subset \mathbb{A}(\mathbf{C}_{2n})$ of periodic operators, i.e.

$$\mathbb{A}_{\text{per}}(\mathbf{C}_{2n}) := \{A \in \mathbb{A}(\mathbf{C}_{2n}) \mid S^{2n}A = AS^{2n}\}, \tag{1.65}$$

where, as before, S is the right shift operator.

The spectrum of a periodic MCMV matrix is described by the trace of the corresponding monodromy matrix. Given $w \in \mathbb{D}$ we define the Blaschke factor

$$b_w(z) = \frac{z - w}{1 - \overline{w}z} \tag{1.66}$$

and for $a \in \mathbb{D}$ the $\mathbb{SU}(1,1)$ matrix

$$\mathcal{U}(a) := \frac{1}{\rho}\begin{pmatrix} 1 & a \\ \overline{a} & 1 \end{pmatrix}, \quad \rho = \sqrt{1 - |a|^2}.$$

For a fixed $A = A((a_k)), \vartheta; \mathbf{C}_{2n}) \in \mathbb{A}_{\mathrm{per}}(\mathbf{C}_{2n})$, define the monodromy matrix T_A by

$$T_A(z) := \mathcal{U}(a_0)\begin{pmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{pmatrix}\mathcal{U}(a_1)\begin{pmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{pmatrix}\mathcal{U}(a_2)\begin{pmatrix} b_{z_2}(z) & 0 \\ 0 & 1 \end{pmatrix}\cdots$$
$$\cdots\mathcal{U}(a_{2n-3})\begin{pmatrix} b_{z_{n-1}}(z) & 0 \\ 0 & 1 \end{pmatrix}\mathcal{U}(a_{2n-2})\begin{pmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{pmatrix}\mathcal{U}(a_{2n-1})\begin{pmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{pmatrix}. \tag{1.67}$$

Letting $B(z) = z\prod_{j=1}^{n-1} b_{z_j}(z)$, we introduce the following:

Definition 1.34 (Discriminant of an MCMV matrix)**.** Let $A = A((a_k), \vartheta; \mathbf{C}_{2n}) \in \mathbb{A}_{\mathrm{per}}(\mathbf{C}_{2n})$ be a periodic MCMV matrix. The discriminant of A is the rational function defined by

$$\Delta_A(z) := \frac{1}{B(z)}\operatorname{tr}(T_A(z)).$$

In [6, Lemma 4.2] we show that

(i) Δ_A is real, i.e., $\overline{\Delta_A(\hat{z})} = \Delta_A(z)$,

(ii) zeros of Δ_A lie on $\partial\mathbb{D}$ and are simple,

(iii) for all critical points $c \in \partial\mathbb{D}$ we have $|\Delta_A(c)| \geq 2$,

(iv) $\sigma(A) = \Delta_A^{-1}([-2, 2])$.

Resembling the crucial properties of the polynomial discriminant. In particular, this implies that $\sigma(A)$ is a finite-gap set.

Let $\mathsf{E} \subset \partial\mathbb{D}$ be a finite-gap set, $\mathbf{C}_{2n} = (z_0, \hat{z}_0, z_1, \hat{z}_1, \ldots, z_{n-1}, \hat{z}_{n-1})$, $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$ and consider as in (1.41)

$$\mathcal{T}_E(\mathbf{C}_{2n}, \lambda_*) = \{A \in \mathbb{A}_{\mathrm{per}}(\mathbf{C}_{2n}) \mid \sigma(A) = E, \Delta_A(\lambda_*) > 0\}.$$

The extra condition $\Delta_A(\lambda_*) > 0$ may seem surprising at the first sight compared to (1.28) or (1.41). This is explained, since in both cases the discriminant is normalized at ∞; specifically the polynomial discriminant has positive leading coefficient and thus a fixed behavior at ∞. Likewise the discriminant of periodic GMP matrices always has behavior $\lambda_{g+1}z$ for $z \to \infty$ with $\lambda_{g+1} > 0$. In the case of orthogonal rational functions associated to subsets of $\partial\mathbb{D}$ there is no special gap which could be used as a natural normalization. This explains the appearance of this extra condition.

The discriminant can be used to check whether a periodic MCMV matrix is an element of $\mathcal{T}_E(\mathbf{C}_{2n}, \lambda_*)$ or not, providing an algebraic description of the isospectral torus:

Theorem 1.35. *Fix $\mathbf{C}_{2n}$, $A_0 \in \mathbb{A}_{\mathrm{per}}(\mathbf{C}_{2n})$, $\lambda_* \in \partial\mathbb{D} \setminus \sigma(A_0)$ with $\Delta_{A_0}(\lambda_*) > 0$. Then*

$$\mathcal{T}_{\sigma(A_0)}(\mathbf{C}_{2n}, \lambda_*) = \{A \in \mathbb{A}_{\mathrm{per}}(\mathbf{C}_{2n}) : \Delta_A = \Delta_{A_0}\}. \tag{1.68}$$

In the above theorem, we assumed a priori that $\mathsf{E} = \sigma(A_0)$ is the spectrum of a periodic MCMV matrix. However, for an arbitrary finite-gap set E and an arbitrary choice of poles $\mathbf{C}_{2n}$, $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{2n}, \lambda_*)$ may be empty. This can be seen for example by the fact that not every finite-gap set is the spectrum of a periodic CMV matrix. However, we will show, that for a certain choice $\mathbf{C}_{\mathsf{E}}$ it is always non-empty, in fact it is in bijective correspondence with a $g+1$ dimensional torus, where $g+1$ is the number of gaps, respectively bands of E. This will be achieved by studying the Ahlfors function for the points 0 and ∞, respectively.

Consider a finite-gap set E having $g+1$ gaps. Recall that for any point z_0 in the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$, there exists an Ahlfors function w_{z_0} which maximizes the modulus of the derivative at z_0 (or, in the case $z_0 = \infty$, maximizes $\lim_{z\to\infty} |zw_\infty(z)|$) among all analytic functions on $\overline{\mathbb{C}} \setminus \mathsf{E}$ with modulus bounded by 1. This extremal property defines w_{z_0} uniquely up to a unimodular multiplier and, moreover, $w_{z_0}(z_0) = 0$. In the right normalization, these Ahlfors functions for $\mathsf{E} \subset \partial\mathbb{D}$ have the symmetry property

$$w_{\hat{z}_0}(\hat{z}) = \overline{w_{z_0}(z)}; \tag{1.69}$$

in particular, the zeros of w_∞ can be obtained by reflecting the zeros of w_0 with respect to $\partial\mathbb{D}$. In terms of these functions, we can define a special function, which we call the generalized discriminant, related to the set E:

Definition 1.36 (Generalized discriminant)**.** For a finite union of non-degenerate closed circular arcs $\mathsf{E} \subset \partial\mathbb{D}$, the generalized discriminant is defined by

$$\Delta_{\mathsf{E}} := \frac{1}{w_0 w_\infty} + w_0 w_\infty.$$

By (1.69), we see that Δ_{E} is real-valued on $\partial\mathbb{D}$; since $|w_{z_0}(z)| = 1$ for $z \in \mathsf{E}$ in the sense of nontangential limits and $|w_{z_0}(z)| < 1$ for $z \in \overline{\mathbb{C}} \setminus \mathsf{E}$, it follows that

$$\mathsf{E} = \Delta_{\mathsf{E}}^{-1}([-2, 2]).$$

The function Δ_{E} has $2(g+1)$ poles, half of which lie inside the unit disk. Moreover, there is exactly one critical point (i.e., a zero of Δ_{E}') in each band of E and in each gap of E. While Δ_{E} maps all critical points in bands to -2, the critical points in gaps have Δ_{E}-value strictly greater than 2. As will be crucial for our analysis, we define $\mathbf{C}_{\mathsf{E}}$ to be some fixed ordering of the poles of Δ_{E} inside $\mathbb{D}$, i.e.

$$\mathbf{C}_{\mathsf{E}} := (z_0 = 0, \hat{z}_0, z_1, \hat{z}_1, \dots, \hat{z}_g) \in \mathbb{D}^{g+1}, \quad z_k \in \mathbb{D} \text{ a pole of } \Delta_{\mathsf{E}}. \tag{1.70}$$

The following theorem shows that for the particular choice $\mathbf{C}_{\mathsf{E}}$, $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}, \lambda_*)$ is in bijective correspondence to the isospectral torus of almost periodic CMV matrices $\mathcal{T}_{\mathsf{E}}^{\mathrm{CMV}}$:

Theorem 1.37 (Periodic coordinates for finite-gap CMV matrices)**.** *Let $\mathsf{E} \subset \partial\mathbb{D}$ be a disjoint union of $g+1$ non-degenerate closed circular arcs, and let $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$. For $\mathbf{C}_{\mathsf{E}}$ as in (1.70) depending only on E, there is a unitary bijection between $\mathcal{T}_{\mathsf{E}}^{\mathrm{CMV}}$ and $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}, \lambda_*)$; i.e.*

$$\mathcal{T}_{\mathsf{E}}^{\mathrm{CMV}} \simeq \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}, \lambda_*).$$

In particular, for an almost-periodic CMV matrix $\mathcal{C}$ with absolutely continuous spectrum E, there exists an associated CMV matrix $C = C((a_k))$ with phase-periodic coefficients

$$a_{k+2(g+1)} = e^{-2i\vartheta} a_k, \quad k \in \mathbb{Z} \tag{1.71}$$

such that $\mathcal{C}$ is unitarily equivalent to the periodic MCMV matrix $A((a_k), \vartheta; \mathbf{C}_{\mathsf{E}}) \in \mathbb{A}_{\mathrm{per}}(\mathbf{C}_{\mathsf{E}})$. Moreover, the spectral measures of the half-line restrictions $\mathcal{C}_+$ and A_+ and the cyclic vector e_0 coincide.

Remark 1.38. (i) As a consequence of (1.71), the operator $b_{\text{-}D_0}(C)$ is periodic up to a phase. By conjugating it with $\Lambda(\vartheta)$ – and this is the main purpose of introducing such a diagonal matrix – we get that $A((a_k), \vartheta; \mathbf{C}_{\mathsf{E}})$ becomes periodic in the standard sense. This is particularly important in view of Theorem 1.39 below, since by Naiman's lemma [36] an operator satisfying the right-hand side of (1.72) is necessarily periodic.

(ii) Notice that (1.71) is invariant under the substitution $\vartheta \mapsto \vartheta + \pi$. Consequently, there are in fact two MCMV matrices, $A((a_k), \vartheta; \mathbf{C}_{\mathsf{E}})$ and $A((a_k), \vartheta + \pi; \mathbf{C}_{\mathsf{E}})$, in $\mathbb{A}_{\text{per}}(\mathbf{C}_{\mathsf{E}})$ with spectrum E having unitary equivalent half-line restrictions to $\mathcal{C}_+$. However, notice that this same map sends $e^{i\vartheta}$ to $-e^{i\vartheta}$, which, by (1.67), changes the sign of the discriminant $\Delta_A \mapsto -\Delta_A$. For this reason, the normalization $\Delta_A(\lambda_*) > 0$ fixes uniquely one such MCMV matrix.

Using the generalized discriminant, we can characterize $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}, \lambda_*)$ by a magic formula such as in Theorem 1.10.

Theorem 1.39 (Magic Formula for MCMV matrices)**.** *Let $\mathsf{E} \subset \partial\mathbb{D}$ be a disjoint union of $g+1$ non-degenerate closed circular arcs, let $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$, and let $\mathbf{C}_{\mathsf{E}}$ be as in (1.70). Then, for any $A \in \mathbb{A}(\mathbf{C}_{\mathsf{E}})$,*

$$A \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}, \lambda_*) \iff \Delta_{\mathsf{E}}(A) = S^{2(g+1)} + S^{-2(g+1)}, \tag{1.72}$$

and in this case $\Delta_A = \Delta_{\mathsf{E}}$.

In this section, we have concentrated mainly on the operator theoretic perspective. As CMV matrices have close connections to Schur and Caratheodory functions, these results can also be interpreted in a function theoretic approach. At this place we will only review this connection, explain the consequences informally and refer to [6, Section 1.2] for the precise statements.

A Schur function is an analytic function $f : \mathbb{D} \to \overline{\mathbb{D}}$. Provided that f is not a finite Blaschke product, i.e. a finite product of Blaschke factors (1.66), the Schur algorithm

$$f_0(z) = f(z),$$
$$z f_{k+1}(z) = \frac{f_k(z) - \mathbf{a}_k}{1 - \overline{\mathbf{a}_k} f_k(z)}, \quad \mathbf{a}_k = f_k(0)$$

determines an infinite sequence of parameters $(\mathbf{a}_k) \in \mathbb{D}^{\mathbb{N}_0}$, also known as Schur parameters; conversely, any sequence $(\mathbf{a}_k) \in \mathbb{D}^{\mathbb{N}_0}$ determines a function $f \in \mathcal{S}$ by an associated continued fraction expansion (see, e.g., [38]).

A Caratheodory function is an analytic function on $\mathbb{D}$, with $\operatorname{Re} F \geq 0$ on $\mathbb{D}$ and $F(0) = 1$. The association

$$F(z) = \frac{1 + zf(z)}{1 - zf(z)}, \quad f(z) = \frac{1}{z} \frac{F(z) - 1}{F(z) + 1},$$

sets up a one-to-one correspondence between Schur and Caratheodory functions. Every Caratheodory function admits an integral representation. That is, there exists a unique probability measure μ on $\partial\mathbb{D}$, such that

$$F(z) = \int \frac{\zeta + z}{\zeta - z} d\mu(\zeta).$$

Let $\mathcal{C}_+ = \mathcal{C}_+((a_k))$ be a CMV matrix generated by the coefficients $(a_k)_{k \in \mathbb{N}_0}$. Then e_0 forms a cyclic vector for $\mathcal{C}_+$ and the spectral measure of $\mathcal{C}_+$ and the vector e_0 is the measure in the integral representation of the Caratheodory function

$$F(z) = \left\langle (\mathcal{C}_+ - z)^{-1} (\mathcal{C}_+ + z) \delta_0, \delta_0 \right\rangle.$$

If f is the Schur function associated to F, then by Geronimus' theorem [38, Chapter 3] $(a_k)_{k\in\mathbb{N}_0}$ are exactly the Schur parameters of f.

The context of Theorem 1.37 is that the almost periodic CMV matrix $\mathcal{C}_+$ and the periodic MCMV matrix A_+ are associated to the same Caratheodory function F and Schur function f. The Schur algorithm always uses zero as interpolation point. As an immediate corollary from the direct spectral theory of periodic MCMV matrices developed in [6, Section 4] one obtains that if one uses the zeros of the Ahlfors function as interpolation points, then f solves a periodic interpolation problem, see [6, Theorem 1.6]. Moreover, F is a solution of a quadratic equation of the form (1.14) where the coefficients are given in terms of the monodromy matrix of A_+, see [6, Corollary 1.7]. This resolved a conjecture of Simon [39, Conjecture 11.9.6]

1.2 Eigenvalue asymptotic for continuum Schrödinger operators

In this section, we describe several results about the local and global asymptotic behavior of eigenvalues of truncations of continuum Schrödinger operators to finite intervals as the length of the interval tends to ∞. Let $V \in L^1_{\mathrm{loc}}([0,\infty))$ be real valued and consider the associated Schrödinger operator, corresponding formally to

$$H_V = -\partial_x^2 + V. \tag{1.73}$$

We will in addition assume that V is such that the operator is limit point at $+\infty$. Thus, placing a Dirichlet condition at 0 yields a self-adjoint realization of H_V. That is, we will consider the domain of the operator

$$D(H_V) = \{f \in L^2((0,\infty)) \mid f, f' \in \mathrm{AC}_{\mathrm{loc}}([0,\infty)), -f'' + Vf \in L^2((0,\infty)), f(0) = 0\}, \tag{1.74}$$

where $\mathrm{AC}_{\mathrm{loc}}([0,\infty))$ denotes the set of functions that are absolutely continuous on bounded intervals. Let H_V^L denote the truncation of H_V to a finite interval $[0, L]$ with domain

$$D(H_V^L) = \{f \in L^2((0,L)) \mid f, f' \in \mathrm{AC}([0,L]), -f''+Vf \in L^2((0,L)), f(0) = f(L) = 0\}.$$

As will be explained below, H_V^L has only discrete spectrum corresponding to the zeros of fundamental solutions. For any $z \in \mathbb{C}$ the Dirichlet solution, $u(x,z)$, is the solution of the initial value problem

$$-\partial_x^2 u(x,z) + V(x)u(x,z) = zu(x,z), \quad u(0,z) = 0, \partial_x u(0,z) = 1. \tag{1.75}$$

Clearly ξ is an eigenvalue of H_V^L if and only if $u(L,\xi) = 0$. For fixed L, $u(L,\cdot)$ is an entire function that has infinitely many zeros, which are necessarily real and accumulate only at $+\infty$. The zeros of $u(L,\cdot)$ will be denoted by ξ_j^L, $j \in \mathbb{N}$. In this section, we will be concerned with asymptotic properties of ξ_j^L as $L \to \infty$. First we will study their global behavior, that is we will study limits of the zero counting measure

$$\rho_L = \frac{1}{L}\sum_{j=0}^{\infty} \delta_{\xi_j^L}. \tag{1.76}$$

In contrast to the normalized zero counting measure for orthogonal polynomials (1.42), this is not a finite measure. In Section 1.2.1, we introduce the notion of Stahl–Totik regularity for continuum Schrödinger operators and show that it implies that ρ_L has a unique limit that can be characterized in terms of the spectrum of the half-line operator H_V. Section 1.2.2 and 1.2.3 will then be concerned with the local asymptotic behavior of ξ_j^L around a point $\xi_0 \in \mathbb{R}$. In particular, we will be interested in the case when consecutive zeros around ξ_0 are asymptotically equally spaced.

1.2.1 Stahl–Totik regularity for continuum Schrödinger operators

In this section, we study Schrödinger operators as in (1.73) with domain (1.74), with focus on studying limits of the zero counting measure (1.76). Throughout this section, we will assume that V is uniformly locally integrable, i.e.,

$$\sup_{x\geq 0}\int_x^{x+1} |V(t)|\mathrm{d}t < \infty. \tag{1.77}$$

In particular, this implies that H_V is limit point at ∞.

As discussed in Section 1.1.2, Stahl–Totik regularity was initially studied in the setting of orthogonal polynomials with compactly supported measures of orthogonality and is formulated in terms of the potential theoretic Green function with pole at infinity, the logarithmic capacity and the equilibrium measure. All these notions cease to exist, when ∞ becomes a boundary point of the domain. In the setting of continuum Schrödinger operators, the boundary of the domain will be the closure of $\mathsf{E} = \sigma_{\text{ess}}(H_V)$ with respect to the Riemann sphere $\overline{\mathbb{C}}$ and thus, since H_V is unbounded, will contain ∞ as a boundary point. Let us mention that this is different from the setting of Section 1.1.3. Although for orthogonal rational functions we also allow unbounded sets E, but the set of poles are chosen from $\overline{\mathbb{R}} \setminus \mathsf{E}$ and are thus still interior points of the domain. In this section, we consider unbounded sets and still use the boundary point ∞ as "interpolation point".

Martin functions will serve as the counterpart of Green functions in the spectral theory of unbounded operators. Martin functions first appeared in [34], for a book treatment we refer to [2]. For Denjoy domains, i.e. $\mathsf{E} \subset \mathbb{R}$, the Martin function related to ∞ can be defined as a positive harmonic function on Ω which is bounded on each bounded subset of Ω and which vanishes q.e. on E. If E is semibounded, the cone of such functions is one-dimensional. Hence, one additional normalization is required, which will be specified later. Interestingly, whereas the Green function always has a $\log|z|$ behavior at its pole, the behavior of the Martin function can be more varied. We will show that the Martin function with pole at ∞ associated to resolvent domains of Schrödinger operators obeying (1.77) have maximal possible growth. Let E be semibounded and M be a Martin function. The Akhiezer–Levin condition in this setting is

$$\lim_{z \to -\infty} \frac{M(z)}{\sqrt{-z}} > 0 \tag{1.78}$$

(by general principles, the limit exists with a value in $[0, \infty)$). This condition was initially proposed by Akhiezer and Levin without a square root for subsets of $\mathbb{R}$ that are unbounded in both directions. In this setting, the condition can be understood as follows: if the set is sufficiently dense at ∞ such that, in potential-theoretical sense, locally at ∞ divides $\mathbb{C}$ into two half planes, then there exist two linearly independent Martin functions, or only one in the case of sparser sets. We refer to [23] for a comprehensive discussion on Martin functions in Denjoy domains.

For Akhiezer–Levin sets, we will normalize the Martin function so that the limit in (1.78) is equal to 1 and denote it by M_E. The following theorem shows that spectra of continuum Schrödinger operators satisfying (1.77) are always Akhiezer-Levin sets. Moreover, the higher asymptotic expansion that we obtained for M_E is crucial to define a notion of capacity in this setting.

Theorem 1.40. *For any potential V obeying (1.77) and $\mathsf{E} = \sigma_{\text{ess}}(H_V)$, the domain $\Omega = \mathbb{C} \setminus \mathsf{E}$ is Greenian, ∞ is a Dirichlet-regular point for Ω, Ω obeys the Akhiezer–Levin condition, and there exists $a_\mathsf{E} \in \mathbb{R}$ such that the Martin function has the asymptotic behavior*

$$M_\mathsf{E}(z) = \operatorname{Re}\left(\sqrt{-z} + \frac{a_\mathsf{E}}{2\sqrt{-z}}\right) + o\left(\frac{1}{\sqrt{|z|}}\right), \tag{1.79}$$

as $z \to \infty$, $\arg z \in [\delta, 2\pi - \delta]$, for any $\delta > 0$.

Each conclusion is strictly stronger than the other. Specifically, the presence of the second term in expansion (1.79) is not inherent to Akhiezer-Levin sets, but rather a characteristic associated with spectra of Schrödinger operators. It's crucial to note that spectra of Schrödinger operators with bounded potentials can exhibit extreme thinness, potentially possessing zero Hausdorff dimension [12] and zero lower box counting dimension [11]. In contrast, our result can be understood as a universal "thickness" result in the perspective of Martin functions.

The constant a_{E} serves as a substitute for the Robin constant and enters the following universal inequality:

Theorem 1.41. *If V is a potential obeying (1.77) and $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$, then*

$$a_{\mathsf{E}} \leq \liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t. \tag{1.80}$$

The inequality (1.80) is proved by comparing asymptotic expansions of M_{E} and of the Dirichlet solution $u(x,z)$ as defined in (1.75) for $z \to \infty$. In fact, the inequality (1.80) holds not just asymptotically but pointwise, leading to an universal lower bound on the growth rate of eigensolutions:

Theorem 1.42. *If V is a potential obeying (1.77) and $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$, then*

$$M_{\mathsf{E}}(z) \leq \liminf_{x\to\infty} \frac{1}{x}\log|u(x,z)|, \qquad \forall z \in \mathbb{C}\setminus[\min \mathsf{E}, \infty). \tag{1.81}$$

It is necessary to exclude $[\min \mathsf{E}, \infty)$ since for $z \in (\min \mathsf{E}, \infty)$, by Sturm oscillation theory [40], $u(\cdot, z)$ can have infinitely many zeros.

Stahl–Totik regularity is characterized through equality in (1.80).

Definition 1.43. The potential V is regular if

$$a_{\mathsf{E}} = \lim_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t. \tag{1.82}$$

We have already observed in the context of the inequalities (1.80) and (1.81), that an inequality of constants can actually be lifted to an inequality of functions. The same holds in the case of equality, leading to root asymptotics of Dirichlet solutions:

Theorem 1.44. *If V is a potential obeying (1.77) and $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$, the following are equivalent:*

(i) *V is regular,*

(ii) *For q.e. $z \in \mathsf{E}$, $\limsup_{x\to\infty} \frac{1}{x}\log|u(x,z)| \leq 0$,*

(iii) *For all $z \in \mathbb{C}$, $\limsup_{x\to\infty} \frac{1}{x}\log|u(x,z)| \leq M_{\mathsf{E}}(z)$,*

(iv) $\lim_{x\to\infty} \frac{1}{x}\log|u(x,z)| = M_{\mathsf{E}}(z)$ *uniformly on compact subsets of* $\mathbb{C}\setminus[\min \mathsf{E}, \infty)$.

Simon conjectured in [43] that a theory of Stahl–Totik regularity should exist also for continuum Schrödinger operators, that is similar to the one for Jacobi matrices presented in Section 1.1.2. Theorem 1.44 proves a theory of Stahl–Totik regularity in much higher generality as previously conjectured in [43]. See the introduction of [17] for more details on the subject.

Let us return to the initial question of describing limits of the zero counting measure ρ_L. We will first define a candidate for the limit measure. The Martin function can be extended to a subharmonic function to all of $\mathbb{C}$ and as such has a Riesz measure defined by

$$\rho_{\mathsf{E}} = \frac{1}{\pi}\Delta M_{\mathsf{E}}. \tag{1.83}$$

We call ρ_{E} the Martin measure for E. Let ρ_L denote the zero counting measure as defined in (1.76). Note that ρ_L is the Riesz measure of $\frac{1}{L}\log|u(L,\cdot)|$. Stahl–Totik regularity implies convergence of ρ_L to ρ_{E}:

Theorem 1.45. *Assume that V is regular. Then $\rho_L \to \rho_{\mathsf{E}}$ as $L \to \infty$ in the w^*-topology of $C_c(\mathbb{R})^*$.*

To formulate a converse of this theorem, let us recall the construction of a maximal spectral measure in terms of the Weyl m-function. Since H_V is limit point at ∞, for $z \in \mathbb{C} \setminus \mathbb{R}$, there exists (up to a scalar multiple unique) $\psi(x,z)$ satisfying

$$-\partial_x^2\psi(x,z) + V(x)\psi(x,z) = z\psi(x,z)$$

and $\psi(\cdot,z) \in L^2((0,\infty))$, which is called the Weyl solution at ∞. Again for $z \in \mathbb{C} \setminus \mathbb{R}$, the Weyl m-function is defined by

$$m(z) = \frac{\partial_x\psi(0,z)}{\psi(0,z)}. \tag{1.84}$$

The function m is a Herglotz functions and the measure μ in the integral representation (1.6) represents a maximal spectral measure for H_V. If μ is not supported on a polar set, then also the converse of Theorem 1.45 holds:

Theorem 1.46. *Assume that V obeys (1.77) and let μ be a maximal spectral measure for H_V. Suppose that ρ_L converges to ρ_{E} as $L \to \infty$ in the w^*-topology of $C_c(\mathbb{R})^*$. Then, either V is regular, or there exists a polar Borel set X such that $\mu(\mathbb{R} \setminus X) = 0$.*

We conclude with a sufficient condition for Stahl–Totik regularity in terms of the spectral measure.

Theorem 1.47. *Assume that V obeys (1.77), let $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$ and assume that E is a regular Parreau-Widom set. Denote by $\mathrm{d}\mu = f_\mu \mathrm{d}\xi + \mathrm{d}\mu_s$ the Radon–Nikodym decomposition of μ with respect to Lebesgue measure. If $f_\mu(\xi) > 0$ for Lebesgue-a.e. $\xi \in \mathsf{E}$, then V is regular.*

We have presented the results in a deterministic setting. For ergodic families of Schrödinger operators, the considered quantities can be recognized as ergodic notions such as Lyapunov exponent or density of states measure. The above mentioned results have particularly relevant applications in the ergodic setting. For the precise statements, we refer to the introduction of [17].

1.2.2 Bulk universality limits

As seen in the previous section, Stahl–Totik regularity is intimately related to convergence of the zero counting measure and thus to the global asymptotics of the eigenvalues of H_V^L. In contrast to this, this section will be concerned with the local behavior of the eigenvalues at some point $\xi_0 \in \mathbb{R}$.

The results presented in this section are from the paper [18][4], which find most applications for orthogonal polynomials. However, due to their general formulation in terms of canonical systems, they equally apply to continuum Schrödinger operators. For the sake of this introduction, we focus on this perspective. We will present the

[4]This paper is not part of the habilitation thesis.

results for potentials $V \in L^1_{\mathrm{loc}}([0,\infty))$ which are in the limit point case at $+\infty$ and set an arbitrary self-adjoint boundary condition at 0

$$\cos\beta f(0) + \sin\beta f'(0) = 0, \quad \beta \in [0,\pi). \tag{1.85}$$

Thus, the domain of the operator $H_{V,\beta}$ is

$$D(H_{V,\beta}) = \{f \in L^2((0,\infty)) \mid f, f' \in \mathrm{AC}_{\mathrm{loc}}([0,\infty)), -f'' + Vf \in L^2((0,\infty)), (1.85) \text{ holds}\}. \tag{1.86}$$

The corresponding Weyl m-function is

$$m_\beta(z) = \frac{\cos\beta\partial_x\psi(0,z) - \sin\beta\psi(0,z)}{\sin\beta\partial_x\psi(0,z) + \cos\beta\psi(0,z)}, \tag{1.87}$$

where $\psi(x,z)$ denotes again the Weyl solution at $+\infty$. Let μ_β denote the measure in the integral representation of the Herglotz function m_β. Let $u(x,z) = u_\beta(x,z)$, be the eigensolution satisfying

$$-\partial_x^2 u(x,z) + V(x)u(x,z) = zu(x,z), \quad u(0,z) = -\sin\beta, \partial_x u(0,z) = \cos\beta.$$

We define the reproducing kernel

$$K_L(z,w) = \int_0^L u(s,z)\overline{u(s,w)}ds. \tag{1.88}$$

It is indeed the reproducing kernel for the space of functions

$$\mathcal{S}_L := \left\{\int_0^L f(t)u(t,\cdot)dt \mid f \in L^2((0,L))\right\}$$

equipped with the $L^2(\mathbb{R},\mu_\beta)$ scalar product. We say that μ_β admits bulk universality limits at $\xi \in \mathbb{R}$, if

$$\lim_{L\to\infty} \frac{K_L\left(\xi + \frac{z}{\tau(L)}, \xi + \frac{w}{\tau(L)}\right)}{K_L(\xi,\xi)} = \frac{\sin(\pi(z-\overline{w}))}{\pi(z-\overline{w})}, \tag{1.89}$$

for an appropriate function $\tau(L) \to \infty$ as $L \to \infty$ and $z, w \in \mathbb{C}$. The motivation for this type of rescaling limits stems from random matrix theory, where they are used to describe the local eigenvalue statistics of unitary ensembles [14]. This goes back to the seminal work of Wigner [52], where he proposed to use local eigenvalue statistics of random matrices as a model for local statistical behavior of resonances in scattering theory.

There is a long history of results, finding the optimal conditions on the spectral measure to prove existence of universality limits (1.89), see introduction of [18] or the survey paper [30]. In the paper [18], for the first time bulk universality was proved with a pointwise purely local condition on the spectral measure, resolving a conjecture of Avila–Last–Simon [3] and Lubinsky [30]. The result in the setting of continuum Schrödinger operators is the following:

Theorem 1.48 ([18, Theorem 1.10])**.** *Fix $\beta \in [0,\pi)$ and consider the operator $H_{V,\beta}$ with domain (1.86) and m-function m_β as in (1.87). Let $J \subset \mathbb{R}$ and assume that for some $0 < \alpha < \pi/2$,*

$$f_\beta(\xi) := \frac{1}{\pi} \lim_{\substack{z\to\xi \\ \alpha\le\arg(z-\xi)\le\pi-\alpha}} \operatorname{Im} m_\beta(z) \tag{1.90}$$

converges uniformly in $\xi \in J$ *and* $0 < f_\mu(\xi) < \infty$. *Then*

$$\lim_{L\to\infty} \frac{K_L\left(\xi + \frac{z}{f(\xi)K_L(\xi,\xi)}, \xi + \frac{w}{f(\xi)K_L(\xi,\xi)}\right)}{K_L(\xi,\xi)} = \frac{\sin(\pi(z-\overline{w}))}{\pi(z-\overline{w})} \tag{1.91}$$

uniformly on compact regions of $(\xi, z, w) \in J \times \mathbb{C} \times \mathbb{C}$.

Since boundary behavior of Herglotz functions are not affected by a change of the measure away from a neighborhood of ξ, condition (1.90) is a local condition. For Lebesgue a.e. $\xi \in \mathbb{R}$, the limit (1.90) exists and it recovers the density of the absolutely continuous part of μ_β, i.e.,

$$d\mu_\beta(\xi) = f_\beta(\xi)d\xi + d\mu_s(\xi)$$

where μ_s is a singular measure with respect to Lebesgue measure. A point $\xi \in \mathbb{R}$ is called a Lebesgue point of μ_β, if it is a Lebesgue point of f_β and

$$\lim_{\varepsilon\to 0} \frac{\mu_s((\xi-\varepsilon, \xi+\varepsilon))}{2\varepsilon} = 0.$$

Pointwise, the limit (1.90) exists at every Lebesgue point of μ_β.

The connection to eigenvalue spacing is via the Freud–Levin theorem. For $L > 0$ and $\xi \in \mathbb{R}$, we denote by $\xi_j^L(\xi)$, for $j \in \mathbb{Z}$, the zeros of $u(L,\cdot)$ counted from ξ, i.e.,

$$\cdots < \xi_{-2}^L(\xi) < \xi_{-1}^L(\xi) < \xi \leq \xi_0^L(\xi) < \xi_1^L(\xi) < \dots \tag{1.92}$$

with no zeros of $u(L,\cdot)$ between ξ_j^L and ξ_{j+1}^L.

Theorem 1.49. *With the assumptions of Theorem 1.48 we have that for every* $\xi \in J$ *and* $j \in \mathbb{Z}$

$$\lim_{L\to\infty} f_\beta(\xi)K_L(\xi,\xi)\left(\xi_{j+1}^L(\xi) - \xi_j^L(\xi)\right) = 1. \tag{1.93}$$

The statement (1.93) is a consequence of (1.91). This connection was first proved for orthogonal polynomials by Levin and Lubinsky, rediscovering an idea of Freud [22, 29]. The adaptation to continuum Schrödinger operators was done in [15] for the case $\beta = \frac{\pi}{2}$, but can be directly translated to arbitrary β.

The equal eigenvalue spacing (1.93) with rate $K_L(\xi,\xi)$ raises the question about the asymptotic behavior of $K_L(\xi,\xi)$ as $L \to \infty$. This will be the context of the next section.

1.2.3 Asymptotics for Christoffel functions associated to continuum Schrödinger operators

In this section, we consider potentials satisfying (1.77) and consider the associated half-line Schrödinger operators $H_{V,\pi/2}$ with Neumann boundary conditions. However, the same proofs would also apply for Dirichlet boundary conditions. Let $K_L(z,w)$ be the corresponding reproducing kernel (1.88). Note that $u_{\pi/2}$ is the classical Neumann solution which we will henceforth denote by v, i.e.,

$$-\partial_x^2 v_(x,z) + V(x)v(x,z) = zv(x,z), \quad v(0,z) = 1, \partial_x v(0,z) = 0.$$

Then K_L is given by

$$K_L(z,w) = \int_0^L v(t,z)\overline{v(t,w)}\mathrm{d}t.$$

Note that in this case $\mathcal{S}_L$ can be given by

$$\mathcal{S}_L = \left\{ \int_0^L \cos(t\sqrt{z}) f(t) \mathrm{d}t \mid \ f \in L^2((0,L)) \right\}.$$

Motivated by (1.91) and (1.93), for $\xi \in \mathbb{R}$, we are interested in asymptotics of $K_L(\xi,\xi)$ as $L \to \infty$. For $z \in \mathbb{C}$, the function

$$\lambda_L(z) = \frac{1}{K_L(z,z)}$$

is called the *Christoffel function*. It quantifies the growth rate of eigensolutions with respect to L, which is known to be an important quantity in spectral theory. It is used for instance in subordinacy theory [24] or in [28] to describe the absolutely continuous spectrum of H_V. In contrast to universality limits (1.91), it is known that asymptotics of $\lambda_L(\xi)$, for $\xi \in \mathbb{R}$, depend on some global properties of the spectral measure, see e.g. the discussion proceeding Theorem 1.2 in [49]. It turns out that a sufficient global condition in order to prove asymptotics of $\lambda_L(\xi)$ is Stahl–Totik regularity as presented in Section 1.2.1. Recall that exemplified by Theorem 1.47, this is a fairly weak assumption on the spectral measure.

We need to set up some notation in order to formulate the main statement of this section. Let

$$m(z) = m_{\pi/2}(z) = -\frac{\psi(0,z)}{\partial_x \psi(0,z)}$$

and $\mu = \mu_{\pi/2}$ the corresponding spectral measure and write its decomposition $\mathrm{d}\mu = f_\mu \mathrm{d}\xi + \mathrm{d}\mu_s$ with respect to the Lebesgue measure. Let $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$ and ρ_E the corresponding Martin measure (1.83) and write $\mathrm{d}\rho_\mathsf{E} = f_\mathsf{E} \mathrm{d}\xi + \mathrm{d}\rho_{\mathsf{E},s}$ its Lebesgue decomposition. The result for Stahl–Totik regular potentials is the following:

Theorem 1.50. *Let V be a Stahl–Totik regular potential such that $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$ is Dirichlet regular and μ the corresponding spectral measure. Let $I \subset \mathrm{Int}(\mathsf{E})$ be a closed interval such that μ is absolutely continuous in a neighborhood of J and its density f_μ is positive and continuous at every point of J. Then we have*

$$\lim_{L\to\infty} L\lambda_L(\xi) = \frac{f_\mu(\xi)}{f_\mathsf{E}(\xi)}, \tag{1.94}$$

uniformly for $\xi \in J$.

Initially, Christoffel functions were studied in the setting of orthogonal polynomials. In this context, the reproducing kernel K_L is replaced by the Christoffel-Darboux kernel, defined analogously to (1.88), but using the orthogonal polynomial of degree n instead of the Neumann solution at L. For compactly supported measures, a typical result would be

$$\lim_{n\to\infty} n\lambda_n(\xi) = \frac{f_\mu(\xi)}{f_\mathsf{E}(\xi)}, \tag{1.95}$$

where $\lambda_n(\xi)$ is the Christoffel function associated to the orthonormal polynomials and $f_\mathsf{E}(\xi)$ denotes the density of the equilibrium measure. A fundamental result of Máté–Nevai–Totik [35] establishes (1.95) for the case $\mathsf{E} = [-2,2]$. More precisely, it is shown that (1.95) holds provided that μ is Stahl–Totik regular on $[-2,2]$, $f_\mu(\xi) > 0$, $\log f_\mu$ is integrable in a neighborhood of ξ, and ξ is a Lebesgue point of both the measure μ and the Szegő function associated to f_μ. This has been extended by Totik to arbitrary compact sets [47].

Our approach is inspired by a method used by Simon and we obtain a full analog for continuum Schrödinger operators of all results in [42]. Let us mention that the assumptions in [42] or in Theorem 1.50 are stronger than the ones in [35]. However, the conclusion is also stronger, since uniformity in (1.94) requires continuity of f_μ, see also [48]. It is an interesting question if our method could also be used to prove (1.94) under Lebesgue point and local Szegő conditions as used by Máté–Nevai–Totik.

Limits of Christoffel functions for continuum Schrödinger operators were first studied by Maltsev in [33]. At that time the notion of regularity for continuum Schrödinger operators was not available and Maltsev proved (1.94) for potentials $V = \mathring{V} + \tilde{V}$, where $\mathring{V}$ is a periodic continuous potential, $\tilde{V}$ is so that $\sigma_{\text{ess}}(V) = \sigma_{\text{ess}}(\mathring{V})$ and $\tilde{V}$ is Césaro decaying, i.e.,

$$\lim_{L\to\infty} \frac{1}{L} \int_0^L |\tilde{V}(x)| dx = 0. \tag{1.96}$$

Thus, our result generalize [33] in several directions. First of all, if E is the spectrum of a continuum Schrödinger operator, then generically there is no periodic potential so that the essential spectrum of the associated operator is E. Moreover, even if E is the spectrum of a periodic Schrödinger operator, a regular potential does not necessarily satisfy (1.96). A counterexample can be found even in the simplest case $\mathsf{E} = [0, \infty)$ with $\mathring{V} \equiv 0$. It is shown in [17, Example 1.13] that the potential defined piecewise by $V(x) = (-1)^{\lfloor 2n(x-n) \rfloor}$ on $x \in [n-1, n)$ for integer n is regular with $\sigma_{\text{ess}}(L_V) = [0, \infty)$, but (1.96) does not hold. On the other hand, since periodic potentials are regular, it follows from (1.82) and (1.96) that the potentials considered in [33] are Stahl–Totik regular.

We can now combine Theorem 1.50 with Theorem 1.48 and Theorem 1.49. For $\xi \in \mathbb{R}$, let $\xi_j^L(\xi)$ denote the zeros of $\partial_L v(L, \cdot)$ as in (1.92).

Theorem 1.51. *With the assumptions of Theorem 1.50 we have*

$$\lim_{L\to\infty} \frac{K_L\left(\xi + \frac{z}{L}, \xi + \frac{w}{L}\right)}{K_L(\xi, \xi)} = \frac{\sin(\pi f_{\mathsf{E}}(\xi)(z - \overline{w}))}{\pi f_{\mathsf{E}}(\xi)(z - \overline{w})},$$

and

$$\lim_{L\to\infty} L f_{\mathsf{E}}(\xi) \left(\xi_{j+1}^L(\xi) - \xi_j^L(\xi) \right) = 1 \tag{1.97}$$

uniformly for $\xi \in I$.

We conclude with a summary of Section 1.2. In Section 1.2.1 we presented a theory of Stahl Totik regularity for continuum Schrödinger operators and showed that for regular potentials we can describe the global asymptotics of the eigenvalues of H_V^L as $L \to \infty$. In Section 1.2.2 we showed that universality limits and equal eigenvalue spacing with scale $\lambda_L(\xi)$ hold under a weak purely local condition on the Weyl m-function. Finally, in Section 1.2.3 we showed that Stahl–Totik regularity is a suitable global assumption on the potential that allows for the computation of asymptotics $\lambda_L(\xi)$ to obtain universality and equal eigenvalue spacing with explicit scale L.

Bibliography

[1] D. H. Armitage and S. J. Gardiner, *Classical potential theory*, Springer Monographs in Mathematics, Springer Verlag, London, 2001.

[2] ———, *Classical potential theory*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2001.

[3] A. Avila, Y. Last, and B. Simon, *Bulk universality and clock spacing of zeros for ergodic Jacobi matrices with absolutely continuous spectrum*, Anal. PDE **3** (2010), no. 1, 81–108.

[4] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njå stad, *Orthogonal rational functions*, Cambridge Monographs on Applied and Computational Mathematics, vol. 5, Cambridge University Press, Cambridge, 1999.

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

[6] J. S. Christiansen, B. Eichinger, and T. VandenBoom, *Finite-gap CMV matrices: periodic coordinates and a magic formula*, Int. Math. Res. Not. IMRN (2021), no. 18, 14016–14085.

[7] J. S. Christiansen, B. Simon, P. Yuditskii, and M. Zinchenko, *Asymptotics of Chebyshev polynomials, II: DCT subsets of* $\mathbb{R}$, Duke Math. J. **168** (2019), no. 2, 325–349.

[8] J. S. Christiansen, B. Simon, and M. Zinchenko, *Asymptotics of Chebyshev polynomials, I: subsets of* $\mathbb{R}$, Invent. Math. **208** (2017), no. 1, 217–245.

[9] ———, *Widom factors and Szego-Widom asymptotics, a review*, Toeplitz operators and random matrices—in memory of Harold Widom, Oper. Theory Adv. Appl., vol. 289, Birkhäuser/Springer, Cham, [2022] ©2022, pp. 301–319.

[10] ———, *Asymptotics of Chebyshev polynomials, V. residual polynomials*, Ramanujan J. **61** (2023), no. 1, 251–278.

[11] D. Damanik, J. Fillman, and A. Gorodetski, *Multidimensional almost-periodic Schrödinger operators with Cantor spectrum*, Ann. Henri Poincaré **20** (2019), no. 4, 1393–1402.

[12] D. Damanik, J. Fillman, and M. Lukic, *Limit-periodic continuum Schrödinger operators with zero measure Cantor spectrum*, J. Spectr. Theory **7** (2017), no. 4, 1101–1118.

[13] D. Damanik, R. Killip, and B. Simon, *Perturbations of orthogonal polynomials with periodic recursion coefficients*, Ann. of Math. (2) **171** (2010), no. 3, 1931–2010.

[14] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics, vol. 3, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.

[15] B. Eichinger, *Asymptotics for Christoffel functions associated to continuum Schrödinger operators*, J. Anal. Math., 1–35.

[16] ———, *Szego-Widom asymptotics of Chebyshev polynomials on circular arcs*, J. Approx. Theory **217** (2017), 15–25.

[17] B. Eichinger and M. Lukić, *Stahl–Totik regularity for continuum Schrödinger operators*, arXiv:2001.00875 (2020).

[18] B. Eichinger, M. Lukić, and B. Simanek, *An approach to universality using Weyl m-functions*, arXiv:2108.01629 (2021).

[19] B. Eichinger, M. Lukić, and G. Young, *Orthogonal rational functions with real poles, root asymptotics, and GMP matrices*, Trans. Amer. Math. Soc. Ser. B **10** (2023), 1–47.

[20] B. Eichinger, M. Lukić, and G. Young, *Asymptotics of Chebyshev Rational Functions with Respect to Subsets of the Real Line*, Constr. Approx. **59** (2024), no. 3, 541–581.

[21] B. Eichinger and P. Yuditskii, *The Ahlfors problem for polynomials*, Mat. Sb. **209** (2018), no. 3, 34–66.

[22] G. Freud, *Orthogonale Polynome*, Lehrbücher und Monographien aus dem Gebiete der Exakten Wissenschaften, Mathematische Reihe, Band 33, Birkhäuser Verlag, Basel-Stuttgart, 1969.

[23] S. J. Gardiner and T. Sjödin, *Potential theory in Denjoy domains*, Analysis and mathematical physics, Trends Math., Birkhäuser, Basel, 2009, pp. 143–166.

[24] D. J. Gilbert and D. B. Pearson, *On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operators*, J. Math. Anal. Appl. **128** (1987), no. 1, 30–56.

[25] L. Golinskii and S. Khrushchev, *Cesàro asymptotics for orthogonal polynomials on the unit circle and classes of measures*, J. Approx. Theory **115** (2002), no. 2, 187–237.

[26] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. **158** (2003), no. 1, 253–321.

[27] H. Krüger, *Probabilistic averages of Jacobi operators*, Comm. Math. Phys. **295** (2010), no. 3, 853–875.

[28] Y. Last and B. Simon, *Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*, Invent. Math. **135** (1999), no. 2, 329–367.

[29] E. Levin and D. S. Lubinsky, *Applications of universality limits to zeros and reproducing kernels of orthogonal polynomials*, J. Approx. Theory **150** (2008), no. 1, 69–95.

[30] D. S. Lubinsky, *An update on local universality limits for correlation functions generated by unitary ensembles*, SIGMA Symmetry Integrability Geom. Methods Appl. **12** (2016), Paper No. 078, 36.

[31] A. L. Lukashov, *On Chebyshev-Markov rational functions over several intervals*, J. Approx. Theory **95** (1998), no. 3, 233–352.

[32] M. Lukić, *A first course in spectral theory*, Graduate Studies in Mathematics, vol. 226, American Mathematical Society, Providence, RI, [2022] ©2022.

[33] A. Maltsev, *Universality limits of a reproducing kernel for a half-line Schrödinger operator and clock behavior of eigenvalues*, Comm. Math. Phys. **298** (2010), no. 2, 461–484.

[34] R. S. Martin, *Minimal positive harmonic functions*, Trans. Amer. Math. Soc. **49** (1941), 137–172.

[35] A. Mate, P. Nevai, and V. Totik, *Szego's extremum problem on the unit circle*, Ann. of Math. (2) **134** (1991), no. 2, 433–453.

[36] P. B. Naiman, *On the theory of periodic and limit-periodic Jacobian matrices*, Dokl. Akad. Nauk SSSR **143** (1962), 277–279.

[37] T. Ransford, *Potential theory in the complex plane*, London Mathematical Society Student Texts, vol. 28, Cambridge University Press, Cambridge, 1995.

[38] B. Simon, *Orthogonal polynomials on the unit circle. Part 2*, American Mathematical Society Colloquium Publications, vol. 54, American Mathematical Society, Providence, RI, 2005, Spectral theory.

[39] ______, *Orthogonal polynomials on the unit circle. Part 2*, American Mathematical Society Colloquium Publications, vol. 54, American Mathematical Society, Providence, RI, 2005, Spectral theory.

[40] ______, *Sturm oscillation and comparison theorems*, Sturm-Liouville theory, Birkhäuser, Basel, 2005, pp. 29–43.

[41] ______, *Equilibrium measures and capacities in spectral theory*, Inverse Probl. Imaging **1** (2007), no. 4, 713–772.

[42] ______, *Two extensions of Lubinsky's universality theorem*, J. Anal. Math. **105** (2008), 345–362.

[43] ______, *Regularity and the Cesàro-Nevai class*, J. Approx. Theory **156** (2009), no. 2, 142–153.

[44] ______, *Szegő's theorem and its descendants*, M. B. Porter Lectures, Princeton University Press, Princeton, NJ, 2011, Spectral theory for L^2 perturbations of orthogonal polynomials.

[45] M. Sodin and P. Yuditskii, *Functions that deviate least from zero on closed subsets of the real axis*, Algebra i Analiz **4** (1992), no. 2, 1–61.

[46] ______, *Almost periodic Jacobi matrices with homogeneous spectrum, infinite-dimensional Jacobi inversion, and Hardy spaces of character-automorphic functions*, J. Geom. Anal. **7** (1997), no. 3, 387–435.

[47] V. Totik, *Asymptotics for Christoffel functions for general measures on the real line*, J. Anal. Math. **81** (2000), 283–303.

[48] ______, *Universality and fine zero spacing on general sets*, Ark. Mat. **47** (2009), no. 2, 361–391.

[49] ______, *Asymptotics of Christoffel functions on arcs and curves*, Adv. Math. **252** (2014), 114–149.

[50] Luis Velázquez, *Spectral methods for orthogonal rational functions*, J. Funct. Anal. **254** (2008), no. 4, 954–986.

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

[52] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. (2) **62** (1955), 548–564.

[53] P. Yuditskii, *Killip-Simon problem and Jacobi flow on GMP matrices*, Adv. Math. **323** (2018), 811–865.

Chapter 2

Papers

2.1 Orthogonal rational functions with real poles, root asymptotitics, and GMP matrices, B. Eichinger, G. Young, M. Lukić, **Trans. Amer. Math. Soc. 10**, (2023), 1–47

2.2 Asymptotics of Chebyshev rational functions with respect to subsets of the real line, B. Eichinger, G. Young, M. Lukić, **Constr. Approx 59**, (2024), 541-581

2.3 Finite-gap CMV matrices: periodic coordinates and a magic formula, B. Eichinger, J. S. Christiansen, T. VandenBoom, **Int. Math. Res. Not.**, (2021), 14016–14085

2.4 Stahl–Totik regularity for continuum Schrödinger operators, B. Eichinger, M. Lukić, to appear in **Anal. PDE**

2.5 Asymptotics for Christoffel functions associated to continuum Schrödinger operators, B. Eichinger, **J. Anal. Math.**, 1–35

TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY, SERIES B
Volume 10, Pages 1–47 (January 17, 2023)
https://doi.org/10.1090/btran/117

ORTHOGONAL RATIONAL FUNCTIONS WITH REAL POLES, ROOT ASYMPTOTICS, AND GMP MATRICES

BENJAMIN EICHINGER, MILIVOJE LUKIĆ, AND GIORGIO YOUNG

Abstract. There is a vast theory of the asymptotic behavior of orthogonal polynomials with respect to a measure on $\mathbb{R}$ and its applications to Jacobi matrices. That theory has an obvious affine invariance and a very special role for ∞. We extend aspects of this theory in the setting of rational functions with poles on $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, obtaining a formulation which allows multiple poles and proving an invariance with respect to $\overline{\mathbb{R}}$-preserving Möbius transformations. We obtain a characterization of Stahl–Totik regularity of a GMP matrix in terms of its matrix elements; as an application, we give a proof of a conjecture of Simon – a Cesàro–Nevai property of regular Jacobi matrices on finite gap sets.

1. Introduction

There is a vast theory of orthogonal polynomials with respect to measures on $\mathbb{C}$ and their root asymptotics, exemplified by the Ullman–Stahl–Totik theory of regularity. Let μ be a compactly supported probability measure and $\{p_n\}_{n=0}^\infty$ the corresponding orthonormal polynomials, obtained by the Gram–Schmidt process from $\{z^n\}_{n=0}^\infty$ in $L^2(d\mu)$. Then

$$\liminf_{n\to\infty} |p_n(z)|^{1/n} \geq e^{G_{\mathsf{E}}(z,\infty)} \tag{1.1}$$

for z outside the convex hull of $\operatorname{supp}\mu$, where E is the essential support of μ and G_{E} denotes the potential theoretic Green function for the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$; if that domain is not Greenian, one takes $G_{\mathsf{E}} = +\infty$ instead. For measures compactly supported in $\mathbb{R}$, this theory can be interpreted in terms of self-adjoint operators. In particular, for any bounded half-line Jacobi matrix

$$J = \begin{pmatrix} b_1 & a_1 & & \\ a_1 & b_2 & a_2 & \\ & a_2 & \ddots & \ddots \\ & & \ddots & \end{pmatrix}$$

with $a_\ell > 0$, $b_\ell \in \mathbb{R}$,

$$\limsup_{n\to\infty} \left(\prod_{\ell=1}^n a_\ell \right)^{1/n} \leq \operatorname{Cap} \sigma_{\mathrm{ess}}(J), \tag{1.2}$$

Received by the editors October 30, 2020, and, in revised form, December 1, 2020, and February 16, 2022.

2020 *Mathematics Subject Classification.* Primary 47B36, 42C05.

The first author was supported by Austrian Science Fund FWF, project no: J 4138-N32. The second author was supported in part by NSF grant DMS–1700179. The third author was supported in part by NSF grant DMS–1745670.

where Cap denotes logarithmic capacity. For both of these universal inequalities, the case of equality (and existence of limit) is called Stahl–Totik regularity [27]; the theory originated with the case $\mathsf{E} = [-2, 2]$, first studied by Ullman [30].

We extend aspects of this theory to the setting of rational functions with poles in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. One motivation for this is the search for a more conformally invariant theory. Statements such as (1.1), (1.2) rescale in obvious ways with respect to affine transformations (automorphisms of $\mathbb{C}$) which preserve $\mathbb{R}$, so it is obvious that an affine pushforward of a Stahl–Totik regular measure is Stahl–Totik regular. However, the point ∞ has a very special role throughout the theory: for a Möbius transformation f which does not preserve ∞, $p_n \circ f$ are rational functions with a pole at $f^{-1}(\infty)$, and $f(J)$ as defined by the functional calculus is not a finite band matrix. Thus, it is a nontrivial question whether a Möbius pushforward of a Stahl–Totik regular measure is Stahl–Totik regular.

The set of Möbius transformations which preserve $\overline{\mathbb{R}}$ is the semidirect group product $\mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}$, whose normal subgroup $\mathrm{PSL}(2,\mathbb{R})$ corresponds to the orientation preserving case. Denote by $f_*\mu$ the pushforward of μ, defined by $(f_*\mu)(A) = \mu(f^{-1}(A))$ for Borel sets A. As an example of our techniques, we obtain the following:

Theorem 1.1. *Let $f \in \mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}$. If μ is a Stahl–Totik regular measure on $\mathbb{R}$ and $\infty \notin \mathrm{supp}(f_*\mu)$, then the pushforward measure $f_*\mu$ is also Stahl–Totik regular.*

However, we will mostly work in the more general setting when multiple poles on $\overline{\mathbb{R}}$ are allowed, which arises naturally in the spectral theory of self-adjoint operators. Denote by $T_{f,d\mu}$ the multiplication operator by f in $L^2(d\mu)$. The matrix representation for $T_{x,d\mu(x)}$ in the basis of orthogonal polynomials is a Jacobi matrix, and through this classical connection, the theory of orthogonal polynomials is inextricably linked to the spectral theory of Jacobi matrices. In this matrix representation, resolvents $T_{(\mathbf{c}-x)^{-1},d\mu(x)}$ are not finite-diagonal matrices. However, in a basis of orthogonal rational functions with poles at $\mathbf{c}_1, \dots, \mathbf{c}_g, \infty$, the multiplication operators $T_{(\mathbf{c}_1-x)^{-1},d\mu(x)}, \dots, T_{(\mathbf{c}_g-x)^{-1},d\mu(x)}, T_{x,d\mu(x)}$ all have precisely $2g+1$ nontrivial diagonals. The corresponding matrix representations are called GMP matrices; they were introduced by Yuditskii [32].

Self-adjoint operators and their matrix representations are an important part of this work, so we choose to present the theory in a more self-contained way, using self-adjoint operators from the ground up; this has similarities with [22]. Some proofs could be shortened by using orthogonal polynomials with respect to varying weights [27, Chapter 3], but some facts rely on the precise structure obtained by the periodically repeating sequence of poles.

We should also compare this to the case of CMV matrices: for a measure supported on the unit circle, Stahl–Totik regularity is still defined in terms of orthogonal polynomials, but the CMV basis [4, 22] is given in terms of positive and negative powers of z, i.e., orthonormal rational functions with poles at ∞ and 0. The symmetries in that setting lead to explicit formulas for the CMV basis in terms of the orthogonal polynomials; it is then a matter of calculation to relate the exponential growth rate of the CMV basis to that of the orthogonal polynomials, and to interpret regularity in terms of the CMV basis. In our setting, there is no such symmetry and no formula for orthonormal rational functions in terms of orthonormal polynomials.

In order to state our results in a conformally invariant way, we will use the following notations and conventions throughout the paper. The measure μ will be a probability measure on $\overline{\mathbb{R}}$. We denote by $\operatorname{supp}\mu$ its support in $\overline{\mathbb{R}}$, and we consider its essential support (the support with isolated points removed), denoted

$$\mathsf{E} = \operatorname{ess\,supp}\mu.$$

We will always assume that μ is nontrivial; equivalently, $\mathsf{E} \neq \emptyset$.

Fix a finite sequence with no repetitions, $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ with $\mathbf{c}_k \in \overline{\mathbb{R}} \setminus \operatorname{supp}\mu$ for all k. Consider the sequence $\{r_n\}_{n=0}^{\infty}$ where $r_0 = 1$ and for $n = j(g+1) + k$, $1 \le k \le g+1$,

$$r_n(z) = \begin{cases} \frac{1}{(\mathbf{c}_k - z)^{j+1}} & \mathbf{c}_k \in \mathbb{R}, \\ z^{j+1} & \mathbf{c}_k = \infty. \end{cases} \tag{1.3}$$

Applying the Gram–Schmidt process to this sequence in $L^2(d\mu)$ gives the sequence of orthonormal rational functions $\{\tau_n\}_{n=0}^{\infty}$ whose behavior we will study. We note that the special case $\operatorname{supp}\mu \subset \mathbb{R}$, $g = 0$, $\mathbf{C} = (\infty)$ corresponds to the standard construction of orthonormal polynomials associated to the measure μ (note that, since we denote by $\operatorname{supp}\mu$ the support in $\overline{\mathbb{R}}$, the statement $\operatorname{supp}\mu \subset \mathbb{R}$ implies that μ is compactly supported in $\mathbb{R}$), and our first results are an extension of the same techniques.

The first result is a universal lower bound on the growth of $\{\tau_n\}_{n=0}^{\infty}$ in terms of a potential theoretic quantity. If E is not a polar set, we use the (potential theoretic) Green function for the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$, denoted G_{E}, and we define

$$\mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) = \begin{cases} \frac{1}{g+1} \sum_{k=1}^{g+1} G_{\mathsf{E}}(z, \mathbf{c}_k) & \mathsf{E} \text{ is not polar}, \\ +\infty & \mathsf{E} \text{ is polar}. \end{cases} \tag{1.4}$$

Theorem 1.2. *For all $z \in \overline{\mathbb{C}} \setminus \mathbb{R}$,*

$$\liminf_{n\to\infty} |\tau_n(z)|^{1/n} \ge e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}.$$

This is a good place to point out that our current setup is not related to the recent paper [13], in which the behavior was compared to a Martin function at a boundary point of the domain. Here, the behavior is compared to a combination of Green functions (1.4), all the poles are in the interior of the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$, and the difficulty comes instead from the multiple poles.

Another universal inequality for orthonormal polynomials comes from comparing their leading coefficients to the capacity of E. In our setting, the analog of the leading coefficient must be considered in a pole-dependent way. Denote

$$\mathcal{L}_n = \operatorname{span}\{r_\ell \mid 0 \le \ell \le n\}.$$

By the nature of the Gram–Schmidt process, there is a $\kappa_n > 0$ such that

$$\tau_n - \kappa_n r_n \in \mathcal{L}_{n-1}.$$

The Gram–Schmidt process can be reformulated as the $L^2(d\mu)$-extremal problem

$$\kappa_n = \max\left\{\operatorname{Re}\kappa : f = \kappa r_n + h, h \in \mathcal{L}_{n-1}, \|f\|_{L^2(d\mu)} \le 1\right\}. \tag{1.5}$$

By strict convexity of the L^2-norm, these L^2-extremal problems have unique extremizers given by $f = \tau_n$, and κ_n is explicitly characterized as a kind of leading coefficient for τ_n with respect to the pole at $\mathbf{c}_k$ where $n = j(g+1)+k$, $1 \le k \le g+1$.

Below, we will also relate the constants κ_n to off-diagonal coefficients of certain matrix representations.

The growth of the leading coefficients κ_n will be studied along sequences $n = j(g+1)+k$ for a fixed k, and bounded by quantities related to the pole $\mathbf{c}_k$. If E is not a polar set, it is a basic property of the Green function that the limits

$$\gamma_{\mathsf{E}}^k = \begin{cases} \lim_{z\to\mathbf{c}_k}(G_{\mathsf{E}}(z,\mathbf{c}_k)+\log|z-\mathbf{c}_k|), & \mathbf{c}_k \neq \infty, \\ \lim_{z\to\mathbf{c}_k}(G_{\mathsf{E}}(z,\mathbf{c}_k)-\log|z|), & \mathbf{c}_k = \infty \end{cases}$$

exist. Note that if $\mathbf{c}_k = \infty$, γ_{E}^k is precisely the Robin constant for the set E. We further define constants λ_k by

$$\log \lambda_k = \begin{cases} \gamma_{\mathsf{E}}^k + \sum_{\substack{1\le \ell\le g+1 \\ \ell\neq k}} G_{\mathsf{E}}(\mathbf{c}_k,\mathbf{c}_\ell) & \mathsf{E} \text{ is not polar,} \\ +\infty & \mathsf{E} \text{ is polar.} \end{cases} \tag{1.6}$$

Theorem 1.3. *For all* $1\le k\le g+1$, *for the subsequence* $n(j) = j(g+1)+k$,

$$\liminf_{j\to\infty} \kappa_{n(j)}^{1/n(j)} \ge \lambda_k^{1/(g+1)}. \tag{1.7}$$

Theorem 1.4. *The following are equivalent:*

(i) *For some* $1\le k\le g+1$, *for the subsequence* $n(j)=j(g+1)+k$,
$$\lim_{j\to\infty} \kappa_{n(j)}^{1/n(j)} = \lambda_k^{1/(g+1)};$$
(ii) *For all* $1\le k\le g+1$, *for the subsequence* $n(j)=j(g+1)+k$,
$$\lim_{j\to\infty} \kappa_{n(j)}^{1/n(j)} = \lambda_k^{1/(g+1)};$$
(iii)
$$\lim_{n\to\infty}\left(\prod_{\ell=1}^{g+1}\kappa_{n+\ell}\right)^{1/n} = \left(\prod_{k=1}^{g+1}\lambda_k\right)^{1/(g+1)};$$
(iv) *For q.e.* $z\in\mathsf{E}$, *we have* $\limsup_{n\to\infty}|\tau_n(z)|^{1/n}\le 1$;
(v) *For some* $z\in\mathbb{C}_+$, $\limsup_{n\to\infty}|\tau_n(z)|^{1/n}\le e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}$;
(vi) *For all* $z\in\mathbb{C}$, $\limsup_{n\to\infty}|\tau_n(z)|^{1/n}\le e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}$;
(vii) *Uniformly on compact subsets of* $\mathbb{C}\setminus\mathbb{R}$, $\lim_{n\to\infty}|\tau_n(z)|^{1/n} = e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}$.

Definition 1.5. The measure μ is $\mathbf{C}$-regular if it obeys one (and therefore all) of the assumptions of Theorem 1.4.

In this terminology, Stahl–Totik regularity is precisely (∞)-regularity, i.e., $\mathbf{C}$-regularity for the special case $\operatorname{supp}\mu\subset\mathbb{R}$, $g=0$, $\mathbf{C}=(\infty)$. Theorems 1.2, 1.3, 1.4 are closely motivated by foundational results for Stahl–Totik regularity. A new phenomenon appears through the periodicity with which poles are taken in (1.3) and the resulting subsequences $n(j)=j(g+1)+k$: since κ_n is a normalization constant for τ_n, it is notable that control of κ_n along a single subsequence $n(j)=j(g+1)+k$ in Theorem 1.4(i) provides control over the entire sequence. This phenomenon doesn't have an exact analog for orthogonal polynomials, where $g=0$. We will also see below that this is essential in order to characterize the regularity of a GMP matrix using only the entries of the matrix itself and not its resolvents.

Moreover, we show that the regular behavior described by Theorem 1.4 is independent of the set of poles $\mathbf{C}$:

Theorem 1.6. *Let $\mathbf{C}_1, \mathbf{C}_2$ be two finite sequences of elements from $\overline{\mathbb{R}} \setminus \operatorname{supp}\mu$, not necessarily of the same length. Then μ is $\mathbf{C}_1$-regular if and only if it is $\mathbf{C}_2$-regular.*

Corollary 1.7. *Let $\operatorname{supp}\mu \subset \mathbb{R}$. Let $\mathbf{C}$ be a finite sequence of elements from $\overline{\mathbb{R}} \setminus \operatorname{supp}\mu$. Then μ is $\mathbf{C}$-regular if and only if it is Stahl–Totik regular.*

Thus, Theorem 1.4 should not be seen as describing equivalent conditions for a new class of measures, but rather a new set of regular behaviors for the familiar class of Stahl–Totik regular measures.

We consistently work with poles on $\overline{\mathbb{R}}$ since our main interest is tied to self-adjoint problems. Some of our results are in a sense complementary to the setting of [27, Section 6.1], where poles are allowed in the complement of the convex hull of $\operatorname{supp}\mu$, and the behavior of orthogonal rational functions is considered with respect to a Stahl–Totik regular measure. Due to this, it is natural to expect that these results hold more generally, for measures on $\mathbb{C}$ and general collections of poles and Möbius transformations. Moreover, in our setup the poles are repeated exactly periodically, but we expect this can be generalized to a sequence of poles which has a limiting average distribution. Related questions for orthogonal rational functions were also studied by [3, 10].

As noted in [27, Section 6.1], poles in the gaps of $\operatorname{supp}\mu$ can cause interpolation defects in the problem of interpolation by rational functions. In our work, these interpolation defects show up as possible reductions in the order of the poles. For example, consider $\mathbf{C} = (\infty, 0)$. Then, by construction, τ_{2j+1} is allowed a pole at 0 of order at most j. However, if μ is symmetric with respect to $z \mapsto -z$, the functions τ_n will have an even/odd symmetry. Since τ_{2j+1} contains a nontrivial multiple of z^{j+1}, it follows that $\tau_{2j+1}(z) = (-1)^{j+1}\tau_{2j+1}(-z)$. By this symmetry, the actual order of the pole at 0 is $j + 1 - k$ for some even k, so it cannot be equal to j (it will follow from our results that in this case, the order of the pole is $j - 1$). The same effect can be seen for the pole at ∞ for $\mathbf{C} = (0, \infty)$. In the polynomial case, this does not occur: p_n always has a pole at ∞ of order exactly n.

We will consider at once the distribution of zeros of τ_n and the possible reductions in the order of the poles. We will prove that all zeros of τ_n are real and simple, and that $n - g \le \deg \tau_n \le n$. We define the normalized zero counting measure

$$\nu_n = \frac{1}{n} \sum_{w : \tau_n(w) = 0} \delta_w.$$

Although we normalize by n, ν_n may not be a probability measure: however $1 - g/n \le \nu_n(\overline{\mathbb{R}}) \le 1$. Therefore, normalizing by $\deg \tau_n$ instead of by n would not affect the limits as $n \to \infty$.

We will now describe the weak limit behavior of the measures ν_n as $n \to \infty$. To avoid pathological cases, we assume that E is not polar; in that case, denoting by $\omega_{\mathsf{E}}(dx, w)$ the harmonic measure for the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$ at the point w, we define the probability measure on E,

$$\rho_{\mathsf{E},\mathbf{C}} = \frac{1}{g+1} \sum_{j=1}^{g+1} \omega_{\mathsf{E}}(\mathrm{d}x, \mathbf{c}_j).$$

The results below describe weak limits of measures in the topology dual to $C(\overline{\mathbb{R}})$.

Theorem 1.8. *Let μ be a probability measure on $\overline{\mathbb{R}}$. Assume that E is not a polar set.*

(a) *If* μ *is* $\mathbf{C}$ *regular, then* $\text{w-lim}_{n\to\infty} \nu_n = \rho_{\mathsf{E},\mathbf{C}}$.
(b) *If* $\text{w-lim}_{n\to\infty} \nu_n = \rho_{\mathsf{E},\mathbf{C}}$, *then* μ *is* $\mathbf{C}$ *regular or there exists a polar set* $X \subset \mathsf{E}$ *such that* $\mu(\overline{\mathbb{R}} \setminus X) = 0$.

We now turn to matrix representations of self-adjoint operators. Fix a sequence $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ such that $\mathbf{c}_{k_\infty} = \infty$ for some $1 \le k_\infty \le g+1$. A half-line GMP matrix [32] is the matrix representation for multiplication by x in the basis $\{\tau_n\}_{n=0}^\infty$ for this sequence $\mathbf{C}$; its matrix elements are

$$A_{mn} = \int \overline{\tau_m(x)} x \tau_n(x)\, d\mu(x).$$

The condition that $\mathbf{c}_{k_\infty} = \infty$ for some k_∞ guarantees that $A_{mn} = 0$ for $|m-n| > g+1$, so these matrix elements generate a bounded operator A on $\ell^2(\mathbb{N}_0)$ such that $A_{mn} = \langle e_m, A e_n \rangle$, where $(e_n)_{n=0}^\infty$ denotes the standard basis of $\ell^2(\mathbb{N}_0)$. We say that $A \in \mathbb{A}(\mathbf{C})$.

GMP matrices have the property that some of their resolvents are also GMP matrices; namely, for any $k \neq k_\infty$, $(\mathbf{c}_k - A)^{-1} \in \mathbb{A}(f(\mathbf{C}))$ where f is the Möbius transform $f : z \mapsto (\mathbf{c}_k - z)^{-1}$ and $f(\mathbf{C}) = (f(\mathbf{c}_1), \dots, f(\mathbf{c}_{g+1}))$.

Note that the special case $g = 0$, $\mathbf{C} = (\infty)$ gives precisely a Jacobi matrix. A Jacobi matrix is said to be regular if it is obtained by this construction from a regular measure; analogously, we will call a GMP matrix regular if it is obtained from a regular measure. Just as regularity of a Jacobi matrix can be characterized in terms of its off-diagonal entries, we will show that regularity of a GMP matrix can be characterized in terms of its entries in the outermost nontrivial diagonal. We will also obtain a GMP matrix analog of the inequality (1.2).

The GMP matrix has an additional block matrix structure; in particular, for a GMP matrix with $\mathbf{c}_{k_\infty} = \infty$, on the outermost nonzero diagonal $m = n - g - 1$, the only nonzero terms appear for $n = j(g+1) + k_\infty$, and those are strictly positive. Thus, we denote

$$\beta_j = \langle e_{j(g+1)+k_\infty}, A e_{(j+1)(g+1)+k_\infty} \rangle. \tag{1.8}$$

Theorem 1.9. *Fix a probability measure* μ *with* $\operatorname{supp} \mu \subset \mathbb{R}$ *and a sequence* $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_{g+1})$ *with* $\mathbf{c}_k = \infty$. *Then*

$$\limsup_{j\to\infty} \left(\prod_{\ell=1}^{j} \beta_\ell \right)^{1/j} \le \lambda_{k_\infty}^{-1}. \tag{1.9}$$

Moreover, the measure μ *is Stahl–Totik regular if and only if*

$$\lim_{j\to\infty} \left(\prod_{\ell=1}^{j} \beta_\ell \right)^{1/j} = \lambda_{k_\infty}^{-1}. \tag{1.10}$$

The proof will use a relation between the sequence $\{\beta_j\}_{j=1}^\infty$ and the constants $\{\kappa_{j(g+1)+k_\infty}\}_{j=1}^\infty$. In particular, the characterization of regularity in Theorem 1.9 is made possible by the characterization of regularity in terms of the subsequence $\{\kappa_{j(g+1)+k}\}_{j=1}^\infty$ for any single k. Theorem 1.9 also corroborates the perspective that regularity of the measure is the fundamental notion which manifests itself equally well in many different matrix representations.

Since the resolvents $(\mathbf{c}_k - A)^{-1}$ are also GMP matrices and their measures are pushforwards of the original measure, they are also regular GMP matrices; in this sense, Theorem 1.9 provides $g+1$ criteria for regularity, one corresponding to each subsequence $n(j) = j(g+1)+k$, $1 \le k \le g+1$.

As an application of this theory, we show that it provides a proof of a theorem for Jacobi matrices originally conjectured by Simon [23]. Let $\mathsf{E} \subset \mathbb{R}$ be a compact finite gap set,

$$\mathsf{E} = [\mathbf{b}_0, \mathbf{a}_0] \setminus \bigcup_{k=1}^{g} (\mathbf{a}_k, \mathbf{b}_k), \tag{1.11}$$

and denote by $\mathcal{T}_{\mathsf{E}}^{+}$ the set of almost periodic half-line Jacobi matrices with $\sigma_{\mathrm{ess}}(J) = \sigma_{\mathrm{ac}}(J) = \mathsf{E}$ [5, 14]. Through algebro-geometric techniques and the reflectionless property, this class of Jacobi matrices has been widely studied for their spectral properties and quasiperiodicity (see also [26, 31] for more general spectral sets). They also provide natural reference points for perturbations, which is our current interest. On bounded half-line Jacobi matrices J, we consider the metric

$$d(J, \tilde{J}) = \sum_{k=1}^{\infty} e^{-k}(|a_k - \tilde{a}_k| + |b_k - \tilde{b}_k|). \tag{1.12}$$

On norm-bounded sets of Jacobi matrices, convergence in this metric corresponds to strong operator convergence. However, instead of distance to a fixed Jacobi matrix $\tilde{J}$, we will consider the distance to $\mathcal{T}_{\mathsf{E}}^{+}$,

$$d(J, \mathcal{T}_{\mathsf{E}}^{+}) = \inf_{\tilde{J} \in \mathcal{T}_{\mathsf{E}}^{+}} d(J, \tilde{J}) = \min_{\tilde{J} \in \mathcal{T}_{\mathsf{E}}^{+}} d(J, \tilde{J}).$$

Denote by S_+ the right shift operator on $\ell^2(\mathbb{N}_0)$, $S_+ e_n = e_{n+1}$. The condition $d((S_+^*)^m J S_+^m, \mathcal{T}_{\mathsf{E}}^{+}) \to 0$ as $m \to \infty$ is called the Nevai condition. For $\mathsf{E} = [-2, 2]$, this corresponds simply to the commonly considered condition $a_n \to 1$, $b_n \to 0$ as $n \to \infty$ [18]. In general, as a consequence of [21], the Nevai condition implies regularity. The converse is false; however:

Theorem 1.10. *If $\mathsf{E} \subset \mathbb{R}$ is a compact finite gap set and J is a regular Jacobi matrix with $\sigma_{\mathrm{ess}}(J) = \mathsf{E}$, then*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{m=1}^{N} d((S_+^*)^m J S_+^m, \mathcal{T}_{\mathsf{E}}^{+}) = 0. \tag{1.13}$$

The condition (1.13) is described as the Cesàro–Nevai condition; it was first studied by Golinskii–Khrushchev [15] in the OPUC setting with essential spectrum equal to $\partial\mathbb{D}$. Theorem 1.10 was conjectured by Simon [23] and proved in the special case when E is the spectrum of a periodic Jacobi matrix with all gaps open by using the periodic discriminant and techniques from Damanik–Killip–Simon [7] to reduce to a block Jacobi setting. It was then proved by Krüger [17] by very different methods under the additional assumption $\inf_n a_n > 0$. While this is a common assumption in the ergodic literature, regular Jacobi matrices do not always satisfy it: [22, Example 1.4] can easily be modified to give a regular Jacobi matrix with spectrum $[-2, 2]$ and $\inf a_n = 0$. We prove Theorem 1.10 in full generality by applying Simon's strategy and, instead of the periodic discriminant and techniques from [7], using the Ahlfors function, GMP matrices, and techniques of Yuditskii [32].

For the compact finite gap set $\mathsf{E} \subset \mathbb{R}$, among all analytic functions $\overline{\mathbb{C}} \setminus \mathsf{E} \to \mathbb{D}$ which vanish at ∞, the Ahlfors function Ψ takes the largest value of $\mathrm{Re}(z\Psi(z))|_{z=\infty}$.

The Ahlfors function has precisely one zero in each gap, denoted $\mathbf{c}_k \in (\mathbf{a}_k, \mathbf{b}_k)$ for $1 \le k \le g$, a zero at $\mathbf{c}_{g+1} = \infty$, and no other zeros; see also [25, Chapter 8]. In particular, for the finite gap set E, this generates a particularly natural sequence of poles $\mathbf{C}_{\mathsf{E}} = (\mathbf{c}_1, \dots, \mathbf{c}_g, \infty)$.

The Ahlfors function was used by Yuditskii [32] to define a discriminant for finite gap sets,

$$\Delta_{\mathsf{E}}(z) = \Psi(z) + \frac{1}{\Psi(z)}. \tag{1.14}$$

This function is not equal to the periodic discriminant, but it has some similar properties and it is available more generally (even when E is not a periodic spectrum). Namely, Δ_{E} extends to a meromorphic function on $\overline{\mathbb{C}}$ and $(\Delta_{\mathsf{E}})^{-1}([-2,2]) = \mathsf{E}$. It was introduced by Yuditskii to solve the Killip–Simon problem for finite gap essential spectra. In fact, the discriminant is a rational function of the form

$$\Delta_{\mathsf{E}}(z) = \lambda_{g+1} z + d + \sum_{k=1}^{g} \frac{\lambda_k}{\mathbf{c}_k - z} \tag{1.15}$$

for some $d \in \mathbb{R}$; in particular, we will explain that the constants $\lambda_j > 0$ in (1.15) match the general definition (1.6).

As a first glimpse of our proof of Theorem 1.10, we note that it uses the following chain of implications. Starting with a regular Jacobi matrix with essential spectrum E, by a change of one Jacobi coefficient, which does not affect regularity, we can assume that $\mathbf{c}_k \notin \operatorname{supp} \mu$ (Lemma 7.1). Under this assumption, regularity of the Jacobi matrix implies regularity of the corresponding GMP matrix A and the resolvents $(\mathbf{c}_k - A)^{-1}$, $k = 1, \dots, g$, which can be characterized in terms of their coefficients by Theorem 1.9. By properties of the Yuditskii discriminant, this further implies regularity of the block Jacobi matrix $\Delta_{\mathsf{E}}(A)$. Let us briefly recall that a block Jacobi matrix is of the form

$$\mathbf{J} = \begin{bmatrix} \mathfrak{w}_0 & \mathfrak{v}_0 & & & \\ \mathfrak{v}_0^* & \mathfrak{w}_1 & \mathfrak{v}_1 & & \\ & \mathfrak{v}_1^* & \mathfrak{w}_2 & \mathfrak{v}_2 & \\ & & \mathfrak{v}_2^* & \ddots & \ddots \\ & & & \ddots & \end{bmatrix}, \tag{1.16}$$

where $\mathfrak{v}_j$ and $\mathfrak{w}_j$ are $d \times d$ matrices, $\mathfrak{w}_j = \mathfrak{w}_j^*$, and $\det \mathfrak{v}_j \neq 0$ for each j. Type 3 block Jacobi matrices have each $\mathfrak{v}_j$ lower triangular and positive on the diagonal. An extension of regularity to block Jacobi matrices was developed by Damanik–Pushnitski–Simon [8]; in particular, $\mathbf{J}$ is regular for the set $[-2,2]$ if $\sigma_{\mathrm{ess}}(\mathbf{J}) = [-2,2]$ and

$$\lim_{n\to\infty} \left(\prod_{j=1}^{n} |\det \mathfrak{v}_j| \right)^{1/n} = 1. \tag{1.17}$$

This chain of arguments will result in Lemma 1.11:

Lemma 1.11. *Let J be a regular Jacobi matrix, $\mathsf{E} = \sigma_{\mathrm{ess}}(J)$ a finite gap set, and $\mathbf{C}_{\mathsf{E}}$ the corresponding sequence of zeros of the Ahlfors function. Assuming $\mathbf{c}_k \notin \sigma(J)$ for $1 \le k \le g$, denote by A the GMP matrix corresponding to J with respect to the* *sequence* $\mathbf{C}_{\mathsf{E}}$*. Then* $\Delta_{\mathsf{E}}(A)$ *is a regular type 3 block Jacobi matrix with essential spectrum* $[-2,2]$*.*

With Lemma 1.11, it will follow that $\mathbf{J} = \Delta_{\mathsf{E}}(A)$ obeys a Cesàro–Nevai condition. That Cesàro–Nevai condition will imply (1.13) by a modification of arguments of [32]. The strategy is clear: just as [32] uses a certain square-summability in terms of $\mathfrak{v}_j, \mathfrak{w}_j$ to prove finiteness of ℓ^2-norm of $\{d((S_+^*)^m J S_+^m, \mathcal{T}_{\mathsf{E}}^+)\}_{m=0}^\infty$, we will use Cesàro decay in terms of $\mathfrak{v}_j, \mathfrak{w}_j$ to conclude the Cesàro decay (1.13). This can be expected due to a certain locality in the dependence between the terms of the series considered; this idea first appeared in [23] in the setting of periodic spectra with all gaps open. However, some care is needed, since the locality is only approximate in some steps; this is already visible in (1.12). Also, substantial modifications are needed throughout the proof due to the possibility of $\liminf \|\mathfrak{v}_j\| = 0$ (this cannot happen in the Killip–Simon class), which locally breaks some of the estimates. The fix is that this can only happen along a sparse subsequence, but the combination of a bad sparse subsequence and approximate locality means that we cannot simply ignore a bad subsequence once from the start; we must maintain it throughout the proof. A related issue arises with the Cesàro version of a Killip–Simon type functional. We will describe the necessary modifications to the detailed analysis in [32].

The rest of the paper will not exactly follow the order given in this section. In Section 2, we describe the behavior of our problem with respect to Möbius transformations, and we describe the distribution of zeros of the rational function τ_n. In Section 3, we recall the structure of GMP matrices and relate their matrix coefficients to the quantities κ_n, and use this to provide a first statement about exponential growth of orthonormal rational functions on $\overline{\mathbb{C}} \setminus \overline{\mathbb{R}}$. In Section 4, we combine this with potential theoretic techniques to characterize limits of $\frac{1}{n}\log|\tau_n|$ as $n \to \infty$ and prove the universal lower bounds. In Section 5, we prove the results for $\mathbf{C}$-regularity and Stahl–Totik regularity. In Section 6 we describe a proof of Theorem 1.10.

2. Orthonormal rational functions and Möbius transformations

In Section 1, starting from the measure μ and sequence of poles $\mathbf{C}$, we defined a sequence $\{r_n\}_{n=0}^\infty$ and the orthonormal rational functions $\{\tau_n\}_{n=0}^\infty$. In the next statement, we will denote these by $r_n(z;\mathbf{C})$ and $\tau_n(z;\mu,\mathbf{C})$, in order to state precisely the invariance of the setup with respect to Möbius transformations.

Lemma 2.1. *If* f *is a Möbius transformation which preserves* $\overline{\mathbb{R}}$*, then*

$$\tau_n(z;\mu,\mathbf{C}) = \rho^n \tau_n(f(z); f_*\mu, f(\mathbf{C})), \tag{2.1}$$

where $f(\mathbf{C}) = (f(\mathbf{c}_1), \dots, f(\mathbf{c}_{g+1}))$ *and*

$$\rho = \begin{cases} +1 & f \in \mathrm{PSL}(2,\mathbb{R}), \\ -1 & f \in (\mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}) \setminus \mathrm{PSL}(2,\mathbb{R}). \end{cases}$$

Proof. Note that the sequence $\{r_n\}_{n=0}^\infty$ does not have this property: $r_n(z;\mathbf{C})$ is not equal to $\rho^n r_n(f(z); f(\mathbf{C}))$. However, if we denote

$$\mathcal{L}_n(\mathbf{C}) = \mathrm{span}\{r_\ell(\cdot;\mathbf{C}) \mid 0 \le \ell \le n\},$$

then it suffices to have

$$r_n(f(z); f(\mathbf{C})) - c_n \rho^n r_n(z; \mathbf{C}) \in \mathcal{L}_{n-1}(\mathbf{C}) \tag{2.2}$$

for some constants $c_n > 0$. If (2.2) holds, then applying the Gram–Schmidt process to the sequences $\{r_n(f(z); f(\mathbf{C}))\}_{n=0}^{\infty}$ and $\{r_n(z; \mathbf{C})\}_{n=0}^{\infty}$ will give the same sequence of orthonormal functions, up to the sign change ρ^n, which is precisely (2.1).

Note that if (2.1) holds for f_1, f_2, it holds for their composition, so it suffices to verify (2.2) for a set of generators of $\mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}$. In particular, (2.2) is checked by straightforward calculations for affine transformations and for the inversion $f(z) = -1/z$, which implies the general statement since affine maps and inversion generate $\mathrm{PSL}(2,\mathbb{R}) \rtimes \{\mathrm{id}, z \mapsto -z\}$. □

Let us emphasize what Lemma 2.1 does and what it doesn't do. Since the Möbius transformation acts on both the measure and the sequence of poles, Lemma 2.1 does not by itself prove Theorem 1.1. Lemma 2.1 can only say that if μ is Stahl–Totik regular, then $f_*\mu$ is $(f(\infty))$-regular, which is not sufficient unless f is affine. The proof of Theorem 1.1 will be more involved.

However, Lemma 2.1 provides a very useful conformal invariance for many of our proofs. This can be compared to choosing a convenient reference frame. Since potential theoretic notions such as Green functions are conformally invariant, our results will be invariant with respect to Möbius transformations. We will often use this invariance in the proofs to fix a convenient point at ∞.

Note that this will be possible even though some objects are not conformally invariant. Some of our results compare the sequences κ_n with the λ_k, and although those objects are not preserved under conformal transformations, both sequences are affected in a compatible way so that the inequalities and equalities are preserved. Explicitly, fix k and $n = j(g+1)+k$ and a Möbius transformation $f \in \mathrm{PSL}(2,\mathbb{R})$ (a reflection can be considered separately). Let us denote a local dilation factor $f'(\mathbf{c}_k) = \lim_{z\to\mathbf{c}_k} \frac{r_k(z,\mathbf{C})}{r_k(f(z),f(\mathbf{C}))}$. Then, we use Lemma 2.1 to compute

$$\kappa_{n(j)} = \lim_{z\to\mathbf{c}_k} \frac{\tau_n(z,\mu,\mathbf{C})}{r_n(z,\mathbf{C})} = \lim_{z\to\mathbf{c}_k} \frac{\tau_n(f(z),f_*\mu,f(\mathbf{C}))}{r_n(z,\mathbf{C})} = \frac{\tilde{\kappa}_{n(j)}}{f'(\mathbf{c}_k)^{j+1}},$$

where $\tilde{\kappa}_{n(j)}$ is the leading coefficient $\tau_n(z,f_*\mu,f(\mathbf{C})) - \tilde{\kappa}_n r_n(z,f(\mathbf{c}_k)) \in \mathcal{L}_{n-1}(f(\mathbf{C}))$. If E is nonpolar, the Green function is conformally invariant so we find by another computation

$$\begin{aligned}
\log \tilde{\lambda}_k &:= \lim_{w\to f(\mathbf{c}_k)} \big(G_{f(\mathsf{E})}(w,f(\mathbf{c}_k)) - \log|r_k(w,f(\mathbf{C}))|\big) + \sum_{\substack{1\le \ell\le k\\ \ell\ne k}} G_{f(\mathsf{E})}(f(\mathbf{c}_k),f(\mathbf{c}_\ell)) \\
&= \lim_{z\to\mathbf{c}_k} \left(G_{\mathsf{E}}(z,\mathbf{c}_k) - \log|r_k(z,\mathbf{C})| + \log\left|\frac{r_k(z,\mathbf{C})}{r_k(f(z),f(\mathbf{C}))}\right|\right) + \sum_{\substack{1\le \ell\le k\\ \ell\ne k}} G_{\mathsf{E}}(\mathbf{c}_k,\mathbf{c}_\ell) \\
&= \log\lambda_k + \log(f'(\mathbf{c}_k)),
\end{aligned}$$

where we have used that $f \in \mathrm{PSL}(2,\mathbb{R}) \implies f' > 0$ on $\overline{\mathbb{R}}$. Thus, $\tilde{\lambda}_k = f'(\mathbf{c}_k)\lambda_k$. If E is polar, then $f(\mathsf{E})$ is as well. From these calculations, it becomes elementary to verify that statements such as those in Theorems 1.3, 1.4 are conformally invariant.

Note that technical ingredients of the proof, such as polynomial factorizations, give a preferred role to ∞ so they break symmetry. For instance, we will often use the observation that the subspace $\mathcal{L}_n$ can be represented as

$$\mathcal{L}_n = \left\{ \frac{P}{R_n} \mid P \in \mathcal{P}_n \right\} \tag{2.3}$$

for some suitable polynomial R_n with factors which account for finite poles $\mathbf{c}_k \neq \infty$. We will use the representation (2.3) after placing a convenient point at ∞. This idea is already seen in the next proof.

Lemma 2.2. *All zeros of the rational function τ_n are simple and lie in $\overline{\mathbb{R}}$. Moreover, $n - g \leq \deg \tau_n \leq n$.*

Let $n = j(g+1) + k$, $1 \leq k \leq g+1$, and denote by I the connected component of $\mathbf{c}_k$ in $\overline{\mathbb{R}} \setminus \operatorname{supp} \mu$. Then τ_n has no zeros in I and at most one zero in any other connected component of $\overline{\mathbb{R}} \setminus \operatorname{supp} \mu$.

Proof. Fix $1 \leq k \leq g+1$ and without loss of generality, assume $\mathbf{c}_k = \infty$. Then, in the representations (2.3), we can notice that $R_{n-1} = R_n$. In particular, then $\tau_n \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$ implies the representation $\tau_{n(j)} = \frac{P_n}{R_n}$ for some polynomial P_n of degree n.

Recall that τ_n, $n = k + (j-1)(g+1)$ is the unique minimizer for the extremal problem (1.5). By complex conjugation symmetry, the minimizer is real. To proceed further, we study zeros of P_n by using Markov correction terms.

We say that a rational function M is an admissible Markov correction term if $M > 0$ a.e. on E and $M(z)P_n(z) \in \mathcal{P}_{n-1}$. In this case, using $\langle M\tau_n, \tau_n\rangle > 0$, we see that the function $g(\epsilon) = \|\tau_n - \epsilon M \tau_n\|^2$ obeys

$$g'(0) = -2\langle M\tau_n, \tau_n\rangle < 0.$$

Thus, for small enough $\epsilon > 0$, the function

$$\tilde{\tau}_n = \tau_n - \epsilon M \tau_n$$

obeys $\|\tilde{\tau}_n\|_{L^2(d\mu)} < \|\tau_n\|_{L^2(d\mu)}$. Since $\tilde{\tau}_n$ is of the form $\tilde{\tau}_n = \kappa_n z^{j+1} + h(z)$ for some $h(z) \in \mathcal{L}_{n-1}$ and in particular has the same leading coefficient as τ_n, the function $\tilde{\tau}_n / \|\tilde{\tau}_n\|_{L^2(d\mu)} \in \mathcal{L}_n$ contradicts extremality of τ_n. In other words, for the extremizer τ_n, there cannot be any admissible Markov correction terms.

Assume that P_n has a nonreal zero $w \in \overline{\mathbb{C}} \setminus \overline{\mathbb{R}}$. Then, since τ_n is real, $P_n(\overline{w}) = 0$, so the Markov correction term $M(z; w) = \frac{1}{(z-w)(z-\bar{w})}$ would be admissible, leading to contradiction.

Assume that P_n has two zeros x_1, x_2 in the same connected component of $\overline{\mathbb{R}} \setminus \operatorname{supp} \mu$; then, the Markov correction term

$$M(z; x_1, x_2) = \frac{1}{(z - x_1)(z - x_2)}$$

would be admissible, leading to contradiction.

There are no zeros of P_n in I. Otherwise, if $x \in I$ was a zero, the Markov term

$$M(z, x) = \begin{cases} \frac{1}{z-x}, & x < \inf \mathsf{E}, \\ \frac{1}{x-z}, & x > \sup \mathsf{E} \end{cases}$$

would be admissible.

Finally, all zeros of P_n are simple: otherwise, if $x_0 \in \mathbb{R}$ was a double zero, the Markov term

$$M(z, x_0) = \frac{1}{(z - x_0)^2}$$

would be admissible.

The properties of zeros of τ_n follow from those of P_n. There may be cancellations in the representation $\tau_n = \frac{P_n}{R_n}$, but since P_n has at most a simple zero at $\mathbf{c}_\ell$, the only possible cancellations are simple factors $(z-\mathbf{c}_\ell)$, $\ell \neq k$. Thus, $n-g \leq \deg \tau_n \leq n$. □

The use of Markov correction factors is standard in the Chebyshev polynomial literature and is applied here with a modification for the L^2-extremal problem (in the L^∞-setting, singularities in M are treated with a separate argument near the singularity, which would not work here).

Corollary 2.3. *The measures ν_n are a precompact family with respect to weak convergence on $C(\overline{\mathbb{R}})$. Any accumulation point $\nu = \lim_{\ell\to\infty} \nu_{n_\ell}$ is a probability measure and* $\operatorname{supp}\nu \subset \mathsf{E}$*.*

Proof. By Lemma 2.2, $\nu_n(\overline{\mathbb{R}}) \leq 1$, so precompactness follows by the Banach-Alaoglu theorem. If $\nu = \lim_{\ell\to\infty} \nu_{n_\ell}$, then since $1 - \frac{g}{n_\ell} \leq \nu_{n_\ell}(\overline{\mathbb{R}}) \leq 1$, $\nu(\overline{\mathbb{R}}) = 1$.

Let $(\mathbf{a}, \mathbf{b})$ be a connected component of $\overline{\mathbb{R}} \setminus \mathsf{E}$. Let us prove that $\nu((\mathbf{a}, \mathbf{b})) = 0$. By Möbius invariance, it suffices to assume that (a, b) is a bounded subset of $\mathbb{R}$.

Fix $r \in \mathbb{N}$. As $\operatorname{supp}\mu \setminus \mathsf{E}$ is a discrete set, we have

$$\#\{x \in \operatorname{supp}\mu : \mathbf{a} + 1/r < x < \mathbf{b} - 1/r\} = M < \infty.$$

So, by Lemma 2.2, $\nu_{n_\ell}((\mathbf{a}+1/r, \mathbf{b}-1/r)) \leq \frac{2M+1}{n_\ell}$ and by the Portmanteau theorem and sending $r \to \infty$, $\nu((\mathbf{a}, \mathbf{b})) = 0$ and $\operatorname{supp}\nu \subset \mathsf{E}$. □

3. GMP matrices and exponential growth of orthonormal rational functions

In this section, we consider orthonormal rational functions through the framework of GMP matrices. We begin by recalling the structure of GMP matrices [32]. The GMP matrix has a tridiagonal block matrix structure, with the beginnings of new blocks corresponding to occurrences of $\mathbf{c}_{k_\infty} = \infty$. Explicitly,

$$A = \begin{bmatrix} \tilde{B}_0 & \tilde{A}_0 & & & \\ \tilde{A}_0^* & B_1 & A_1 & & \\ & A_1^* & B_2 & A_2 & \\ & & A_2^* & \ddots & \ddots \\ & & & \ddots & \end{bmatrix},$$

where $\tilde{B}_0$ is a $k_\infty \times k_\infty$ matrix, $\tilde{A}_0$ is a $k_\infty \times (g+1)$ matrix. For $j \geq 0$, A_j, B_j are $(g+1)\times(g+1)$ matrices; while for $j \geq 1$ these appear in A unmodified in the above, $\tilde{A}_0$ and $\tilde{B}_0$ are projections of A_0 and B_0 respectively. More precisely, let X^- denote the upper triangular part of a matrix X (excluding the diagonal) and X^+ the lower triangular part (including the diagonal). Then, indexing the entries of A_j, B_j, $j \geq 0$ from 0 to g, we see they are of the form

$$A_j = \vec{p}_j \vec{\delta}_0^{\intercal}, \quad B_j = \hat{\mathbf{C}} + (\vec{q}_j \vec{p}_j^{\intercal})^+ + (\vec{p}_j \vec{q}_j^{\intercal})^-, \tag{3.1}$$

where $\vec{p}_j, \vec{q}_j \in \mathbb{R}^{g+1}$, with $(\vec{p}_j)_0 > 0$ and $\hat{\mathbf{C}} = \operatorname{diag}\{0, \mathbf{c}_{k_\infty+1}, \dots, \mathbf{c}_{g+1}, \mathbf{c}_1, \dots, \mathbf{c}_{k_\infty-1}\}$ (with the obvious modification if $k_\infty = 1$ or $k_\infty = g+1$) and $\vec{\delta}_0$ denotes the standard first basis vector of $\mathbb{R}^{g+1}$. $\tilde{A}_0$ and $\tilde{B}_0$ are projections of A_0 and B_0,

$$\tilde{A}_0 = \Pi A_0 \quad \tilde{B}_0 = \Pi B_0 \Pi^*$$

with Π the block matrix $\Pi := \left[0_{k_\infty \times (g+1-k_\infty)} | I_{k_\infty \times k_\infty}\right]$. We will refer to $\{\vec{p}_j, \vec{q}_j\}_{j\geq 1}$ as the GMP coefficients of A. While the precise structure will not be essential throughout the paper, we point out two things. First on the outermost diagonal of A in each block there is only one nonvanishing entry, given by $(\vec{p}_j)_0$, which is positive and which is at a different position depending on the position of ∞ in the sequence $\mathbf{C}$. And secondly, in general as a self-adjoint matrix B_j could depend on $(g+1)(g+2)/2$ parameters, but we see that in fact they only depend on $2(g+1)$. This is not that surprising due to their close relation to three-diagonal Jacobi matrices. A similar phenomenon also appears for their unitary analogs [6].

Remark 1. For later reference, we provide an alternative point of view on the block structure of A. The structure provided above is chosen so that c_{k_∞} is at the first diagonal position of the B-blocks. Recall also that to these blocks we attached a column $\vec{p}$ (with positive first entry $(\vec{p})_0 > 0$) to the right and a row $\vec{p}_j^{\intercal}$ at the bottom. If, instead of viewing this as a block matrix structure with blocks of size $(g+1)\times(g+1)$, we view this structure as overlapping blocks of size $(g+2)\times(g+2)$ which overlap at the positions of c_{k_∞}, then those would contain all nonvanishing entries of the GMP matrix (i.e., it would also include the vector $\vec{p}_j$). Moreover, the positive entries are exactly at the upper right and the lower left corner of the bigger block. Now placing the window of size $(g+1)\times(g+1)$ on the top of the bigger block corresponds to the structure presented above. We will encounter in Section 6 that in other settings it may be more natural to place the block at the lower corner, and in this case the B blocks will have structure similar to (3.1).

Now the various notations for the off-diagonal blocks A_j, the vectors $\vec{p}_j$ which determine them, and the coefficients β_j defined in (1.8) are related as

$$\beta_j = \langle e_{j(g+1)+k_\infty}, A e_{(j+1)(g+1)+k_\infty}\rangle = (A_j)_{00} = (\vec{p}_j)_0.$$

The coefficients β_j are a special case of the coefficients Λ_n defined for $n = j(g+1)+k$, $1 \leq k \leq g+1$ as

$$\Lambda_n = \begin{cases} \langle e_{j(g+1)+k}, (\mathbf{c}_k - A)^{-1} e_{(j+1)(g+1)+k}\rangle & k \neq k_\infty, \\ \langle e_{j(g+1)+k}, A e_{(j+1)(g+1)+k}\rangle & k = k_\infty. \end{cases} \tag{3.2}$$

Namely $\beta_j = \Lambda_{j(g+1)+k_\infty}$, and the coefficients $\Lambda_{j(g+1)+\ell}$ for $k \neq k_\infty$ instead occur as outermost diagonal coefficients for the GMP matrix $(\mathbf{c}_k - A)^{-1}$. In our later applications to the discriminant of A, both the coefficients of A and of its resolvents will appear, so we will work with Λ_n throughout.

Next, we connect the coefficients (3.2) to the solutions of the L^2-extremal problem (1.5).

Lemma 3.1. *For all n,*

$$\frac{\kappa_n}{\kappa_{n+g+1}} = \Lambda_n. \tag{3.3}$$

Proof. Let $n = j(g+1)+k$. By self-adjointness,

$$\Lambda_n = \langle e_n, r_k(A) e_{n+g+1}\rangle = \langle r_k \tau_n, \tau_{n+g+1}\rangle = \langle \kappa_n r_{n+g+1} + h, \tau_{n+g+1}\rangle$$

for some $h \in \mathcal{L}_{n+g}$. By orthogonality, $\langle \tau_{n+g+1}, h\rangle = 0$, so $\langle \tau_{n+g+1}, r_{n+g+1}\rangle = \frac{1}{\kappa_{n+g+1}}$ implies that

$$\Lambda_n = \langle \tau_{n+g+1}, \kappa_n r_{n+g+1} + h\rangle = \frac{\kappa_n}{\kappa_{n+g+1}}. \qquad \square$$

We now adapt to GMP matrices ideas from the theory of regularity for Jacobi matrices [22].

Lemma 3.2. *Let $A \in \mathbb{A}(\mathbf{C})$. For all $j \geq 1$, $\|\vec{p}_j\| \leq \|A\|$.*

Proof. Fix $j \geq 1$ and denote $n = j(g+1) + k_\infty$. For any $\ell = 0, \dots, g$,

$$(p_j)_\ell = \langle e_{n-g-1+\ell}, Ae_n\rangle = \int \tau_{n-g-1+\ell}(x) x \tau_n(x) d\mu(x).$$

Since the vectors $\tau_{n-g-1+\ell}$ are orthonormal, by the Bessel inequality,

$$\|\vec{p}_j\|^2 \leq \int |x\tau_n(x)|^2 d\mu(x) \leq \|A\|^2 \int |\tau_n(x)|^2 d\mu(x) = \|A\|^2$$

since $\|A\| = \sup_{x \in \operatorname{supp} \mu} |x|$. $\square$

Lemma 3.3. *For $z \in \mathbb{C} \setminus \mathbb{R}$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log |\tau_n(z)| > 0. \tag{3.4}$$

Proof. We adapt the proof of [22, Proposition 2.2]. It suffices to prove (3.4) along the subsequences $n(j) = j(g+1) + k$, $j \to \infty$, for $1 \leq k \leq g+1$. Moreover, due to $\overline{\mathbb{R}}$-preserving conformal invariance, it suffices to fix k and prove

$$\liminf_{j\to\infty} \frac{1}{n(j)} \log |\tau_{n(j)}(z)| > 0 \tag{3.5}$$

under the assumption that $\mathbf{c}_k = \infty$. This allows us to use the associated GMP matrix $A \in \mathbb{A}(\mathbf{C})$.

Note that for any m, since $\{\tau_\ell\}_{\ell=0}^\infty$ is an orthonormal basis of $L^2(d\mu)$,

$$\sum_\ell A_{m\ell}\tau_\ell(z) = \sum_\ell \langle z\tau_m(z), \tau_\ell(z)\rangle \tau_\ell(z) = z\tau_m(z).$$

This equality holds in $L^2(d\mu)$, but since all functions are rational, it also holds pointwise. Thus, if we fix $z \in \mathbb{C} \setminus \mathbb{R}$, the sequence $\vec{\varphi} = \{\tau_\ell(z)\}_{\ell=0}^\infty$ is a formal eigensolution for A at energy z, i.e. $(A - z)\vec{\varphi} = 0$ componentwise. Since A is represented as a block tridiagonal matrix, let us also write $\vec{\varphi}$ in a matching block form, as $\vec{\varphi}^\top = \begin{bmatrix} \vec{u}_0^\top & \vec{u}_1^\top & \vec{u}_2^\top & \dots \end{bmatrix}$ where

$$\vec{u}_0^\top = \begin{bmatrix} \tau_0(z) & \dots & \tau_{k-1}(z) \end{bmatrix}, \quad \vec{u}_j^\top = \begin{bmatrix} \tau_{n(j-1)-1}(z) & \dots & \tau_{n(j)-1}(z) \end{bmatrix}, \quad j \geq 1.$$

We also consider the projection of $\vec{\varphi}$ onto the first $j+1$ blocks,

$$\vec{\varphi}_j^\top = \begin{bmatrix} \vec{u}_0^\top & \dots & \vec{u}_j^\top & 0 & \dots \end{bmatrix},$$

and compute $(A - z)\vec{\varphi}_j$. By the block tridiagonal structure of A, for $m < n(j-1)$ we have $\langle e_m, (A - z)\vec{\varphi}_j\rangle = 0$. For $0 \leq \ell \leq g$, we have

$$\langle e_{n(j-1)+\ell}, (A - z)\vec{\varphi}\rangle - \langle e_{n(j-1)+\ell}, (A - z)\vec{\varphi}_j\rangle = (p_j)_\ell \tau_{n(j)}(z)$$

so that $\langle e_{n(j-1)+\ell}, (A - z)\vec{\varphi}_j\rangle = -(p_j)_\ell \tau_{n(j)}(z)$. Moreover,

$$\langle e_{n(j)}, (A - z)\vec{\varphi}_j\rangle = \langle e_{n(j)}, A\vec{\varphi}_j\rangle = (\vec{p}_j)^* u_j(z).$$

For $m > n(j)$, we again have $\langle e_m, (A-z)\vec{\varphi}_j\rangle = 0$. In conclusion, $(A-z)\vec{\varphi}_j$ has only two nontrivial blocks,

$$((A-z)\vec{\varphi}_j)^\top = \begin{bmatrix} 0 & \dots & 0 & -(\vec{p}_j \tau_{n(j)}(z))^\top & ((\vec{p}_j)^* u_j)^\top & 0 & \dots \end{bmatrix}.$$

In particular, we can compute

$$\langle \vec{\varphi}_j, (A-z)\vec{\varphi}_j\rangle = -\vec{u}_j^* \tau_{n(j)}(z)\vec{p}_j. \tag{3.6}$$

Since A is self-adjoint and $\vec{\varphi}_j \in \ell^2(\mathbb{N}_0)$, by a standard consequence of the spectral theorem [29, Lemma 2.7.],

$$|\operatorname{Im} z| \|\vec{\varphi}_j\|^2 \le |\langle \vec{\varphi}_j, (A-z)\vec{\varphi}_j\rangle|.$$

Using (3.6) and the Cauchy–Schwarz inequality gives

$$|\operatorname{Im} z| \sum_{m=0}^{j} \|\vec{u}_m\|^2 \le |\tau_{n(j)}(z)| \|\vec{p}_j\| \|\vec{u}_j\|.$$

By Lemma 3.2, with $C = |\operatorname{Im} z| / \|A\|$,

$$C \sum_{m=0}^{j} \|\vec{u}_m\|^2 \le |\tau_{n(j)}(z)| \|\vec{u}_j\|. \tag{3.7}$$

Applying the AM-GM inequality to the right-hand side of (3.7) gives

$$|\tau_{n(j)}(z)| \|\vec{u}_j(z)\| \le \frac{1}{2}\left(C \|\vec{u}_j(z)\|^2 + C^{-1} |\tau_{n(j)}(z)|^2 \right)$$

which together with (3.7) implies

$$|\tau_{n(j)}(z)|^2 \ge C^2 \sum_{m=0}^{j} \|\vec{u}_m\|^2. \tag{3.8}$$

Since $|\tau_{n(j)}(z)|^2 \le \|\vec{u}_{j+1}\|^2$, this implies that

$$\sum_{m=0}^{j+1} \|\vec{u}_m\|^2 \ge \left(1 + C^2\right) \sum_{m=0}^{j} \|\vec{u}_m\|^2.$$

Since $\|\vec{u}_0\| \ge |\tau_0(z)| = 1$, this implies by induction that

$$\sum_{m=0}^{j} \|\vec{u}_m\|^2 \ge \left(1 + C^2\right)^j.$$

Combining this with (3.8) gives a lower bound on $|\tau_{n(j)}(z)|$ which implies (3.5). □

The estimates in the previous proof also lead to the following:

Corollary 3.4. *For any $z \in \mathbb{C} \setminus \mathbb{R}$, the quantities*

$$\liminf_{j\to\infty} \frac{1}{j(g+1)+k} \log |\tau_{j(g+1)+k}(z)|, \qquad \limsup_{j\to\infty} \frac{1}{j(g+1)+k} \log |\tau_{j(g+1)+k}(z)|$$

are independent of $k \in \{1, \dots, g+1\}$.

Proof. Assume $j \geq 1$. For $k - g - 1 \leq \ell \leq k - 1$, the estimate (3.8) gives

$$|\tau_{j(g+1)+k}(z)|^2 \geq C^2 \|\vec{u}_j\|^2 \geq C^2 |\tau_{j(g+1)+\ell}(z)|^2$$

which implies

$$\liminf_{j\to\infty} \frac{1}{j(g+1)+k} \log|\tau_{j(g+1)+k}(z)| \geq \liminf_{j\to\infty} \frac{1}{j(g+1)+\ell} \log|\tau_{j(g+1)+\ell}(z)| \tag{3.9}$$

and

$$\limsup_{j\to\infty} \frac{1}{j(g+1)+k} \log|\tau_{j(g+1)+k}(z)| \geq \limsup_{j\to\infty} \frac{1}{j(g+1)+\ell} \log|\tau_{j(g+1)+\ell}(z)|. \tag{3.10}$$

Clearly, the right-hand sides don't change if ℓ is shifted by $g+1$, so (3.9), (3.10) hold for all $k, \ell \in \{1, \dots, g+1\}$ with $k \neq \ell$. By symmetry, since the roles of k, ℓ can be switched, we conclude that equality holds in (3.9), (3.10). □

4. Growth rates of orthonormal rational functions

In this section, we will combine the positivity (3.4) with potential theory techniques in order to study exponential growth rates of orthonormal rational functions. Our main conclusions will be conformally invariant, but our proofs will use potential theory arguments and objects such as the logarithmic potential of a finite measure ν,

$$\Phi_\nu(z) = \int \log|z - x| d\nu(x),$$

which is well defined when $\operatorname{supp} \nu$ does not contain ∞.

Theorem 4.1. *Fix $1 \leq k \leq g+1$ and denote by I the connected component of $\overline{\mathbb{R}} \setminus \operatorname{supp} \mu$ containing $\mathbf{c}_k$. Suppose there is a subsequence $n_\ell = j_\ell(g+1) + k$ such that $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$ and $\frac{1}{n_\ell} \log \kappa_{n_\ell} \to \alpha \in \mathbb{R} \cup \{-\infty, +\infty\}$ as $\ell \to \infty$. Then uniformly on compact subsets of $(\overline{\mathbb{C}} \setminus \overline{\mathbb{R}}) \cup (I \setminus \{\mathbf{c}_k\})$, we have*

$$h(z) := \lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}(z)|.$$

The function h is determined by ν and α; in particular, if $\mathbf{c}_k = \infty$,

$$h(z) = \alpha + \Phi_\nu(z) - \frac{1}{g+1} \sum_{\substack{m=1 \\ m\neq k}}^{g+1} \log|\mathbf{c}_m - z|. \tag{4.1}$$

Moreover,

(a) *$\alpha = -\infty$ is impossible;*
(b) *If $\alpha = +\infty$, the limit is $h = +\infty$;*
(c) *If $\alpha \in \mathbb{R}$, the limit h extends to a positive harmonic function on $\overline{\mathbb{C}} \setminus (\mathsf{E} \cup \{\mathbf{c}_1, \dots, \mathbf{c}_{g+1}\})$ such that*

$$h(z) = -\frac{1}{g+1} \log|\mathbf{c}_m - z| + O(1), \quad z \to \mathbf{c}_m \neq \infty, \tag{4.2}$$

$$h(z) = \frac{1}{g+1} \log|z| + O(1), \quad z \to \mathbf{c}_m = \infty. \tag{4.3}$$

Proof. By using $\overline{\mathbb{R}}$-preserving conformal invariance, we can assume without loss of generality that $\mathbf{c}_k = \infty$. We will use the representation (2.3) of the subspace $\mathcal{L}_n$. For $n = j(g+1)+k$, counting degrees of the poles leads to

$$\tau_n = \frac{P_n}{R_n}, \qquad R_n(z) = \prod_{m=1}^{k-1}(\mathbf{c}_m - z)\prod_{\substack{m=1\\ m\neq k}}^{g+1}(\mathbf{c}_m - z)^j,$$

with $\deg P_n = n$. This may not be the minimal representation of τ_n, but by the proof of Lemma 2.2, the only possible cancellations are simple factors $(\mathbf{c}_m - z)$ for each $m \neq k$, so we get the minimal representation $\tau_n(z) = P(z)/Q(z)$ with

$$P(z) = \kappa_n \prod_{w:\tau_n(w)=0}(z-w), \quad Q(z) = \prod_{\substack{m=1\\ m\neq k}}^{g+1}(\mathbf{c}_m - z)^{j+\delta_{m,j}},$$

where $|\delta_{m,j}| \leq 1$ for each j. All that matters is that $\delta_{m,j}/j \to 0$ as $j \to \infty$. It will be useful to turn this rational function representation into a kind of Riesz representation,

$$\log|\tau_n(z)| = \log\kappa_n + n\int \log|x-z|\,d\nu_n(x) - \sum_{\substack{1\leq m\leq g+1\\ m\neq k}}(j+\delta_{m,j})\log|\mathbf{c}_m - z|. \tag{4.4}$$

Since $\mathbf{c}_k = \infty$, note that $K = \overline{\mathbb{R}}\setminus I$ is a compact subset of $\mathbb{R}$. Denote $\Omega = \mathbb{C}\setminus K$. For any $z\in\Omega$, the map $x \mapsto \log|x-z|$ is continuous on K, so $\Phi_{\nu_{n_\ell}}(z) \to \Phi_\nu(z)$ as $\ell\to\infty$. In fact, convergence is uniform on compact subsets of Ω: since $\operatorname{supp}(\nu_{n_\ell}) \subset K$ and $\nu_{n_\ell}(K) \leq 1$ for all ℓ, the estimate

$$\log\left|\frac{x-z_1}{x-z_2}\right| \leq \log\left(1+\frac{|z_1-z_2|}{\operatorname{dist}(z_2,K)}\right) \leq \frac{|z_1-z_2|}{\operatorname{dist}(z_2,K)}, \qquad z_1,z_2\in\Omega,$$

implies uniform equicontinuity of the potentials Φ_{n_ℓ} on compact subsets of Ω, and the Arzelà–Ascoli theorem implies uniform convergence on compacts.

Note that (b) follows from (4.1). By Corollary 2.3, $\operatorname{supp}\nu \subset \mathsf{E}$ and $\Phi_\nu(z)$ is harmonic on $\mathbb{C}\setminus\mathsf{E}$, so the right hand side extends to a harmonic function on $\mathbb{C}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$ and we denote this extension also by h. By Lemma 3.3, h is positive on $\mathbb{C}_+\cup\mathbb{C}_-$, so $\alpha \neq -\infty$; moreover, by the mean value property, h is positive on $\mathbb{C}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$.

The remaining asymptotic properties follow from (4.1). Under the assumption $\mathbf{c}_k = \infty$, $\operatorname{supp}\nu$ is a compact subset of $\mathbb{R}$, and $\Phi_\nu(z) = \log|z| + O(1)$, $z\to\infty$. It then follows that $h(z) = \frac{1}{g+1}\log|z| + O(1)$ as $z\to\infty$. Of course, $h(z) = -\frac{1}{g+1}\log|z-\mathbf{c}_m| + O(1)$ near each $\mathbf{c}_m \neq \mathbf{c}_k$. □

Theorem 4.1 motivates interest in positive harmonic functions on $\overline{\mathbb{C}}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$. If E is polar, by Myrberg's theorem [2, Theorem 5.3.8], any such function is constant. If E is not polar, knowing the asymptotic behavior of h at the poles, positivity of h improves to the following lower bound on h. Lemma 4.2 reflects a standard minimality property of the Green function [11, Section VII.10].

Lemma 4.2. *Assume that E is a nonpolar closed subset of $\overline{\mathbb{R}}$. Let h be a positive superharmonic function on $\overline{\mathbb{C}}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$. Suppose* $h(z) + \frac{1}{g+1}\log|z-\mathbf{c}_k|$

has an existent limit at $\mathbf{c}_k$ *for each finite* $\mathbf{c}_k$*, and* $h(z) - \frac{1}{g+1}\log|z|$ *has an existent limit at* ∞ *if one of the* $\mathbf{c}_k = \infty$*. Then*

$$h(z) \geq \mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) \tag{4.5}$$

for $z \in \overline{\mathbb{C}} \setminus \mathsf{E}$*. For* $1 \leq k \leq g+1$*, define*

$$\alpha_k = \begin{cases} \lim_{z\to\mathbf{c}_k}(h(z) + \frac{1}{g+1}\log|z - \mathbf{c}_k|), & \mathbf{c}_k \neq \infty, \\ \lim_{z\to\infty}(h(z) - \frac{1}{g+1}\log|z|), & \mathbf{c}_k = \infty. \end{cases}$$

Then

$$\alpha_k \geq \frac{\log\lambda_k}{g+1}. \tag{4.6}$$

Proof. We will use a stronger, q.e. version of the maximum principle [20, Thm 3.6.9]. Define

$$\tilde{h}(z) := \mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) - h(z),$$

which is bounded at $\mathbf{c}_k$ for $1 \leq k \leq g+1$ and so extends to a subharmonic function on $\overline{\mathbb{C}} \setminus \mathsf{E}$. Since $\mathcal{G}_{\mathsf{E}}$ vanishes q.e. on E, we have for q.e. $t \in \mathsf{E}$,

$$\limsup_{z\to t} \tilde{h}(z) = -\liminf_{z\to t} h(z) \leq 0.$$

Now we show $\tilde{h}$ is bounded above on $\overline{\mathbb{C}} \setminus \mathsf{E}$. Let $\mathcal{U}$ be a union of small neighborhoods containing the points $\mathbf{c}_k$ in $\overline{\mathbb{C}} \setminus \mathsf{E}$. By the definition of the Green function, $\mathcal{G}_{\mathsf{E}}(z, \mathbf{C})$ defines a harmonic and bounded function on $\overline{\mathbb{C}} \setminus (\mathsf{E} \cup \mathcal{U})$. That is, there exists M such that for all $z \in \overline{\mathbb{C}} \setminus (\mathcal{U} \cup \mathsf{E})$ we have

$$\mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) \leq M.$$

Since $h \geq 0$, it follows on $\overline{\mathbb{C}} \setminus (\mathcal{U} \cup \mathsf{E})$ that

$$\tilde{h}(z) = \mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) - h(z) \leq \mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) \leq M.$$

On the other hand, by properties of the Green functions we have

$$\frac{\log\lambda_k}{g+1} = \begin{cases} \lim_{z\to\mathbf{c}_k}(\mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) + \frac{1}{g+1}\log|z - \mathbf{c}_k|), & \mathbf{c}_k \neq \infty, \\ \lim_{z\to\infty}(\mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) - \frac{1}{g+1}\log|z|), & \mathbf{c}_k = \infty. \end{cases}$$

Then, by assumption, for $1 \leq k \leq g+1$, $\tilde{h}(z) = \frac{\log\lambda_k}{g+1} - \alpha_k + o(1)$ as $z \to \mathbf{c}_k$ and, in particular, the difference is bounded in a small neighborhood of $\mathbf{c}_k$. Thus, $\tilde{h}$ is bounded above on $\overline{\mathbb{C}} \setminus \mathsf{E}$.

So, by the maximum principle $\tilde{h} \leq 0 \implies \mathcal{G}_{\mathsf{E}}(z, \mathbf{C}) \leq h(z)$ on $\overline{\mathbb{C}} \setminus \mathsf{E}$. Since $0 \geq \lim_{z\to\mathbf{c}_k} \tilde{h}(z) = \frac{\log\lambda_k}{g+1} - \alpha_k$, we have (4.6). □

Lemma 4.3. *Under the same assumptions as Lemma* 4.2*, the following are equivalent:*

(i) *Equality in* (4.6) *for all* k *with* $1 \leq k \leq g+1$
(ii) *Equality in* (4.6) *for a single* k *with* $1 \leq k \leq g+1$
(iii) *Equality holds in* (4.5)

Proof. (i) $\implies$ (ii) is trivial. Suppose then (ii); with the notation of Lemma 4.2, by assumption, $\tilde{h}(\mathbf{c}_k) = 0$ and $\tilde{h}$ achieves a global maximum. By the maximum principle for subharmonic functions [20, Theorem 2.3.1], $\tilde{h} \equiv 0$ on $\overline{\mathbb{C}} \setminus \mathsf{E}$. Finally, if (iii) holds, then evaluating $\tilde{h}(\mathbf{c}_k)$ for each $1 \leq k \leq g+1$ yields (i). □

We will now prove Theorems 1.2 and 1.3.

Proof of Theorem 1.2. Using conformal invariance, we take $\mathbf{c}_k = \infty$. Fix $z \in \overline{\mathbb{C}} \setminus \overline{\mathbb{R}}$ and select a sequence $(n_\ell)_{\ell=1}^\infty$ such that

$$\liminf_{n\to\infty} \frac{1}{n} \log|\tau_n(z)| = \lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}(z)|.$$

By precompactness of the (ν_n), we may pass to a further subsequence, which we denote again by $(n_\ell)_{\ell=1}^\infty$, so that $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$ and $\frac{1}{n_\ell}\log\kappa_{n_\ell} \to \alpha$ for some ν and α. Then for h as in Theorem 4.1,

$$\lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}(z)| = h(z),$$

on $\mathbb{C} \setminus \mathbb{R}$. If $\alpha = +\infty$, then there is nothing to show. Suppose $\alpha < \infty$. If E is not polar we apply (a) of Theorem 4.1 to find $\alpha \in \mathbb{R}$, and we may use (c) of the same theorem and Lemma 4.2 to conclude.

If instead E is polar, by Myrberg's theorem, h is constant on $\mathbb{C}\setminus(\mathsf{E}\cup\{\mathbf{c}_1, \dots, \mathbf{c}_{g+1}\})$. Computing the limit at $\mathbf{c}_k$ we see $h \equiv +\infty$. In particular, $\liminf_{n\to\infty} \frac{1}{n}\log|\tau_n(z)| = +\infty$ for $z \in \mathbb{C} \setminus \mathbb{R}$. □

Proof of Theorem 1.3. Fix $1 \le k \le g+1$ and assume again by conformal invariance that $\mathbf{c}_k = \infty$. Using precompactness of the measures (ν_n), we find a subsequence $n_\ell = j_\ell(g+1) + k$ with

$$\lim_{\ell\to\infty} \frac{1}{n_\ell} \log\kappa_{n_\ell} = \liminf_{j\to\infty} \frac{1}{n(j)} \log\kappa_{n(j)} =: \alpha$$

and $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$. If $\alpha = +\infty$, we are done. Suppose then $\alpha < \infty$, then we have by Theorem 4.1(a), $\alpha \in \mathbb{R}$. Furthermore, if E is nonpolar, by (c) and Lemma 4.2, $h(z) \ge \mathcal{G}_{\mathsf{E}}(z, \mathbf{C})$ on $\overline{\mathbb{C}} \setminus \mathsf{E}$. In particular, by the representation (4.1) we see that $\alpha = \lim_{z\to\infty}(h(z) - \frac{1}{g+1}\log|z|)$, and so (4.6) yields the desired inequality.

If instead E is polar, by Theorem 1.2, for each $z \in \mathbb{C} \setminus \mathbb{R}$,

$$h(z) = \lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}| \ge \liminf_{n\to\infty} \frac{1}{n} \log|\tau_n(z)| = +\infty,$$

and so by Theorem 4.1(b), $\alpha = +\infty$. □

5. Regularity

We will begin by proving a version of Theorem 1.4 for a fixed k.

Lemma 5.1. *Fix $k \in \{1, \dots, g+1\}$. Along the subsequence $n(j) = j(g+1) + k$, the following are equivalent:*

(i) $\lim_{j\to\infty} \kappa_{n(j)}^{1/n(j)} = \lambda_k^{1/(g+1)}$;
(ii) *For q.e. $z \in \mathsf{E}$, we have* $\limsup_{j\to\infty} |\tau_{n(j)}(z)|^{1/n(j)} \le 1$;
(iii) *For some $z \in \mathbb{C}_+$,* $\limsup_{j\to\infty} |\tau_{n(j)}(z)|^{1/n(j)} \le e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}$;
(iv) *For all $z \in \mathbb{C}$,* $\limsup_{j\to\infty} |\tau_{n(j)}(z)|^{1/n(j)} \le e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}$;
(v) *Uniformly on compact subsets of $\mathbb{C} \setminus \mathbb{R}$,* $\lim_{j\to\infty} |\tau_{n(j)}(z)|^{1/n(j)} = e^{\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})}$.

Proof. Using conformal invariance, we will assume throughout the proof that $\mathbf{c}_k = \infty$. First, suppose that E is polar. In this case (ii) is vacuous, and since $\mathcal{G}_{\mathsf{E}} \equiv +\infty$, (iii) and (iv) are trivially true. Since $\lambda_k = +\infty$, (i) follows from Theorem 1.3. As in the proof of Theorem 4.1, weak convergence of measures implies uniform on compacts convergence of their potentials. Thus, since ν_n are a precompact family, so are Φ_{ν_n}. Thus, the convergence $\lim_{j\to\infty} \frac{1}{n(j)} \log \kappa_{n(j)} = +\infty$ implies that $\lim_{j\to\infty} \frac{1}{n(j)} \log|\tau_{n(j)}(z)| = +\infty$ uniformly on compact subsets of $\mathbb{C}\setminus\mathbb{R}$, so (v) holds.

For the remainder of the proof, we will assume E is not polar. Moreover, we will repeatedly use the fact that if any subsequence of a sequence in a topological space has a further subsequence which converges to a limit, then the sequence itself converges to this limit. In particular, when concluding (v), we apply this fact in the Fréchet space of harmonic functions on $\mathbb{C}\setminus\mathbb{R}$ with the topology of uniform convergence on compact sets.

(iii) $\Longrightarrow$ (v): Given a subsequence of $n(j) = j(g+1)+k$, using precompactness of the measures ν_n, we pass to a further subsequence $n_\ell = j_\ell(g+1)+k$ with $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$ and $\lim_{\ell\to\infty} \frac{1}{n_\ell} \log \kappa_{n_\ell} =: \alpha$, with α real or infinite. By Theorem 4.1, uniformly on compact subsets of $\mathbb{C}\setminus\mathbb{R}$,

$$h(z) = \lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}(z)|$$

with h given by (4.1). Using the assumption, for some $z_0 \in \mathbb{C}_+$, we have

$$h(z_0) \le \limsup_{j\to\infty} \frac{1}{n(j)} \log|\tau_{n(j)}(z_0)| < \infty.$$

So, by Theorem 4.1, $\alpha \in \mathbb{R}$ and h has a harmonic extension to $\overline{\mathbb{C}}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$. Furthermore, by Lemma 4.2, $h \ge \mathcal{G}_\mathsf{E}$. By assumption, we have the opposite inequality at $z_0 \in \mathbb{C}_+$, and so, by the maximum principle for harmonic functions, $h = \mathcal{G}_\mathsf{E}$ on $\mathbb{C}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$, and in particular on $\mathbb{C}\setminus\mathbb{R}$. Thus, we have (v).

(v) $\Longrightarrow$ (iv): For $z \in \{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\}$, $\mathcal{G}_\mathsf{E}(z,\mathbf{C}) = +\infty$ and there is nothing to show. Fix $z \in \mathbb{C}\setminus\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\}$ and let $n_\ell = j_\ell(g+1)+k$ be a subsequence with $\lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}(z)| = \limsup_{j\to\infty} \frac{1}{n(j)} \log|\tau_{n(j)}(z)|$. By passing to a further subsequence, we may assume $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$, and $\lim_{\ell\to\infty} \frac{1}{n_\ell} \log \kappa_{n_\ell} =: \alpha$ where α is real or infinite. By the assumption, we have $h = \lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}| = \mathcal{G}_\mathsf{E}$ on $\mathbb{C}\setminus\mathbb{R}$. So, by (a) and (b), $\alpha \in \mathbb{R}$ and h extends to a harmonic function on $\overline{\mathbb{C}}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$. By the representation (4.1), we may extend h subharmonically to $\overline{\mathbb{C}}\setminus\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\}$. On this set, $\mathcal{G}_\mathsf{E}$ is also subharmonic, so, by the weak identity principle [20, Theorem 2.7.5], $h = \mathcal{G}_\mathsf{E}$ on $\overline{\mathbb{C}}\setminus\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\}$. Thus, by the principle of descent [27, A.III], we have

$$\lim_{\ell\to\infty} \frac{1}{n_\ell} \log|\tau_{n_\ell}(z)| \le h(z) = \mathcal{G}_\mathsf{E}(z,\mathbf{C}) \tag{5.1}$$

and (iv) follows.

(v) $\Longrightarrow$ (i): Given a subsequence of $n(j) = j(g+1)+k$, we use precompactness of the ν_n to pass to a further subsequence $n_\ell = j_\ell(g+1)+k$ with $\lim_{\ell\to\infty} \frac{1}{n_\ell} \log \kappa_{n_\ell} =: \alpha \in \mathbb{R}\cup\{-\infty,+\infty\}$ and $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$. Then in the notation of Theorem 4.1 and by assumption, for a $z \in \mathbb{C}\setminus\mathbb{R}$

$$\lim_{\ell\to\infty} \log|\tau_{n_\ell}(z)| = h(z) = \mathcal{G}_\mathsf{E}(z,\mathbf{C}).$$

So by Lemma 4.3, $\alpha = \frac{\log \lambda_k}{g+1}$. Thus, $\lambda_k^{1/(g+1)}$ is the only accumulation point of $\kappa_{n(j)}^{1/n(j)}$ in $\mathbb{R}\cup\{-\infty,+\infty\}$ and we have (i).

(i) $\Longrightarrow$ (v): As before, we fix a subsequence of $n(j) = j(g+1)+k$ and use precompactness to pass to a further subsequence $n_\ell = j_\ell(g+1)+k$ with $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$.

Then, by Theorem 4.1 and in the notation introduced there, uniformly on compact subsets of $\mathbb{C}\setminus\mathbb{R}$,

$$\lim_{\ell\to\infty}\frac{1}{n_\ell}\log|\tau_{n_\ell}(z)| = h(z),$$

where h is given by (4.1) with $\alpha = \frac{\log\lambda_k}{g+1}$. Thus, by Lemma 4.3(ii), $h(z) = \mathcal{G}_{\mathsf{E}}(z,\mathbf{C})$ on $\mathbb{C}\setminus\mathbb{R}$. Since the initial subsequence was arbitrary, we have (v).

(iv) $\implies$ (ii): Recalling that the Green function vanishes q.e. on E, the claim follows.

(ii) $\implies$ (v): Fixing a subsequence of $n(j)$, we again use precompactness to select a further subsequence $n_\ell = j_\ell(g+1)+k$ such that $\text{w-lim}_{\ell\to\infty}\nu_{n_\ell} = \nu$ and $\lim_{\ell\to\infty}\frac{1}{n_\ell}\log\kappa_{n_\ell} =: \alpha$, $\alpha\in\mathbb{R}\cup\{-\infty,+\infty\}$. By the upper envelope theorem, there is a polar set $X_1\subset\mathbb{C}$ such that on $\mathbb{C}\setminus X_1$, $\limsup_{\ell\to\infty}\Phi_{\nu_{n_\ell}} = \Phi_\nu$. Now, we let $X_2 := \{t\in\mathsf{E} : \limsup_{n\to\infty}\frac{1}{n}\log|\tau_n(t)| > 0\}$, which is polar by assumption, and $X_3 := \{z\in\mathbb{C} : \Phi_\infty(z) = -\infty\}$, which is polar by [20, Theorem 3.5.1]. Then, for a $t\in\mathsf{E}\setminus(X_1\cup X_2\cup X_3)$, we have

$$\alpha \le \limsup_{n\to\infty}\frac{1}{n}\log|\tau_n(t)| - \Phi_\nu(t) + \frac{1}{g+1}\sum_{\substack{m=1\\ m\neq k}}^{g+1}\log|\mathbf{c}_m - t| < \infty.$$

So $\alpha\in\mathbb{R}$ by Theorem 4.1(a). Thus, by (c) of the same theorem, uniformly on compact subsets of $\mathbb{C}\setminus\mathbb{R}$

$$h(z) = \lim_{\ell\to\infty}\frac{1}{n_\ell}\log|\tau_{n_\ell}(z)|$$

and h extends to a positive harmonic function on $\overline{\mathbb{C}}\setminus(\mathsf{E}\cup\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\})$ with logarithmic poles at each of the $\mathbf{c}_m$. So, $h-\mathcal{G}_{\mathsf{E}}$ extends to a harmonic function on $\overline{\mathbb{C}}\setminus\mathsf{E}$, and $h-\mathcal{G}_{\mathsf{E}}\ge 0$ there by Lemma 4.2. We now show that in fact $h=\mathcal{G}_{\mathsf{E}}$ using the stronger, q.e. maximum principle.

We use the equality in (4.1) to extend h to a subharmonic function on $\mathbb{C}\setminus\{\mathbf{c}_1,\dots,\mathbf{c}_{g+1}\}$. By the upper envelope theorem and the assumption again, for $t\in\mathsf{E}\setminus(X_1\cup X_2)$

$$h(t) = \limsup_{\ell\to\infty}\frac{1}{n_\ell}\log|\tau_{n_\ell}(t)| \le 0.$$

Then, for these t, since $\mathcal{G}_{\mathsf{E}}$ is positive, we have

$$\limsup_{\substack{z\to t\\ z\in\mathbb{C}\setminus\mathsf{E}}}(h(z)-\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})) \le \limsup_{\substack{z\to t\\ z\in\mathbb{C}\setminus\mathsf{E}}} h(z) \le h(t) \le 0$$

by upper semicontinuity. So, $\limsup_{\substack{z\to t\\ z\in\mathbb{C}\setminus\mathsf{E}}}(h(z)-\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})) \le 0$ for q.e. $t\in\mathsf{E}$.

Since h is upper semicontinuous on the compact set E, there is an M so that $\sup_{t\in\mathsf{E}} h(t)\le M$. As in the above, now for any $t\in\mathsf{E}$, we have

$$\limsup_{\substack{z\to t\\ z\in\mathbb{C}\setminus\mathsf{E}}}(h(z)-\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})) \le \limsup_{\substack{z\to t\\ z\in\mathbb{C}\setminus\mathsf{E}}} h(z) \le h(t) \le M.$$

So, there is a neighborhood $\mathcal{U}$ of E with $\sup_{z\in\mathcal{U}\cap(\overline{\mathbb{C}}\setminus\mathsf{E})}(h-\mathcal{G}_{\mathsf{E}}) \le M+1$. Since the difference is harmonic on $\overline{\mathbb{C}}\setminus\mathcal{U}$, we conclude that $\sup_{z\in\overline{\mathbb{C}}\setminus\mathsf{E}}(h(z)-\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}))<\infty$. Thus, by the maximum principle and the reverse inequality, $h=\mathcal{G}_{\mathsf{E}}$ on $\overline{\mathbb{C}}\setminus\mathsf{E}$. Since the first sequence was arbitrary, we have (v).

Since the implication (iv) $\implies$ (iii) is clear, we may conclude. □

We now put the subsequences together and use Corollary 3.4 to show that regular behavior occurs for one k if and only if it happens for all.

Proof of Theorem 1.4. Applying Lemma 5.1 for all k implies equivalence of conditions (ii), (iv), (v), (vi), (vii) from Theorem 1.4. By Corollary 3.4, for some $z \in \mathbb{C}_+$, the condition

$$\limsup_{j\to\infty} \frac{1}{j(g+1)+k} \log|\tau_{j(g+1)+k}(z)| \leq \mathcal{G}_{\mathsf{E}}(z, \mathbf{C})$$

holds for one value of k if and only if it holds for all. Due to Lemma 5.1, this immediately implies equivalence of conditions (i) and (iii) from Theorem 1.4. It remains to prove equivalence of (ii), (iii).

(ii) $\implies$ (iii): For $n \in \mathbb{N}$ and $1 \leq k \leq g+1$, denote by $N(n,k)$ the integer such that $n+1 \leq N(n,k) \leq n+g+1$ and $N(n,k)-k$ is divisible by $g+1$. Then $N(n,k)/n \to 1$ as $n \to \infty$ so (ii) implies $\lim_{n\to\infty} \kappa_{N(n,k)}^{1/n} = \lambda_k^{1/(g+1)}$. Taking the product over $k = 1, \dots, g+1$ gives (iii).

(iii) $\implies$ (ii): Similarly to the above, Theorem 1.3 shows that for all k,

$$\liminf_{n\to\infty} \kappa_{N(n,k)}^{1/n} \geq \lambda_k^{1/(g+1)}. \tag{5.2}$$

Thus, if (ii) was false, this would mean that for some $k = m$, $\limsup_{n\to\infty} \kappa_{N(n,m)}^{1/n} > \lambda_m^{1/(g+1)}$. Taking products over k, we would have

$$\begin{aligned}\limsup_{n\to\infty} \left(\prod_{k=1}^{g+1} \kappa_{N(n,k)}\right)^{1/n} &\geq \limsup_{n\to\infty} \kappa_{N(n,m)}^{1/n} \liminf_{n\to\infty} \left(\prod_{\substack{1\leq k\leq g+1\\ k\neq m}} \kappa_{N(n,k)}\right)^{1/n} \\ &> \left(\prod_{k=1}^{g+1} \lambda_k\right)^{1/(g+1)}\end{aligned}$$

(the last step again uses (5.2) for all $k \neq m$). This would contradict (iii), so the proof is complete. ☐

We now prove a seemingly special case of Corollary 1.7.

Proposition 5.2. *Assume that the sequence* $\mathbf{C}$ *contains* ∞. *Then* μ *is Stahl–Totik regular if and only if it is* $\mathbf{C}$*-regular.*

Proof. Let us assume that μ is $\mathbf{C}$-regular and let p_n denote the orthonormal polynomial with respect to μ. Fix $z \in \mathbb{C}$. Since ∞ is in $\mathbf{C}$, $p_n \in \mathcal{L}_{n(g+1)}$, so the orthonormal polynomials can be expressed on the basis of orthonormal rational functions as

$$p_n(z) = \sum_{m=0}^{n(g+1)} c_m \tau_m(z), \qquad \sum_{m=0}^{n(g+1)} |c_m|^2 = 1.$$

Thus, in particular, $|c_\ell| \leq 1$ and we get

$$|p_n(z)| \leq (1 + n(g+1)) \sup_{0\leq m\leq n(g+1)} |\tau_m(z)|. \tag{5.3}$$

By Theorem 1.4, for q.e. $z \in \mathsf{E}$,

$$\limsup_{\ell\to\infty} \frac{1}{\ell} \log|\tau_\ell(z)| \leq 0. \tag{5.4}$$

Thus, for q.e. $z \in \mathsf{E}$, (5.3) implies

$$\limsup_{n\to\infty} \frac{1}{n} \log |p_n(z)| \le 0. \tag{5.5}$$

Thus, μ is Stahl–Totik regular.

Conversely, assume that μ is Stahl–Totik regular. For $n = j(g+1)+k$, the polynomial R_n is a divisor of R_{g+1}^{j+1}, so we can write $\tau_n = \frac{P_n}{R_{g+1}^{j+1}}$ where $\deg P_n \le n+g$. For any $\epsilon > 0$ there exists a polynomial Q_ϵ such that $1-\epsilon \le Q_\epsilon R_{g+1} \le 1+\epsilon$ on E. Thus,

$$|\tau_n(z)| \le (1-\epsilon)^{-j-1} |P_n(z) Q_\epsilon^{j+1}(z)| \tag{5.6}$$

and $\|P_n Q_\epsilon^{j+1}\| \le (1+\epsilon)^{j+1}$ since τ_n is normalized. Since $P_n Q_\epsilon^{j+1}$ is a polynomial of degree at most $n+g+(j+1)\deg Q_\epsilon$, similarly to the above, representing it in the basis of polynomials shows

$$|P_n(z)Q_\epsilon^{j+1}(z)| \le (1+\epsilon)^{j+1}(n+g+1+(j+1)\deg Q_\epsilon) \sup_{0\le m\le n+g+(j+1)\deg Q_\epsilon} |p_n(z)|. \tag{5.7}$$

Since $n+g+(j+1)\deg Q_\epsilon = O(n)$ as $n\to\infty$, the supremum in (5.7) grows subexponentially whenever (5.5) holds. By (5.6), this implies

$$\limsup_{n\to\infty} \frac{1}{n} \log|\tau_n(z)| \le \frac{1}{g+1} \log\left(\frac{1+\epsilon}{1-\epsilon}\right).$$

Since $\epsilon > 0$ is arbitrary, we conclude that (5.5) implies (5.4), so (5.4) holds q.e. on E. ☐

From this seemingly special case, Theorem 1.6 and Corollary 1.7 follow easily:

Proof of Theorem 1.6. By applying a conformal transformation, the special case shows that μ is $\mathbf{C}_1$-regular if and only if it is $(\mathbf{c}_k)$-regular for any single $\mathbf{c}_k$ in $\mathbf{C}_1$. By applying this twice, we conclude that if $\mathbf{C}_1$, $\mathbf{C}_2$ have a common element, then μ is $\mathbf{C}_1$-regular if and only if it is $\mathbf{C}_2$-regular.

By applying that conclusion twice, we will finish the proof. Namely, for arbitrary $\mathbf{C}_1$, $\mathbf{C}_2$, choose a sequence $\mathbf{C}_3$ which has common elements with both $\mathbf{C}_1$ and $\mathbf{C}_2$. Then μ is $\mathbf{C}_1$-regular if and only if it is $\mathbf{C}_3$-regular if and only if it is $\mathbf{C}_2$-regular. ☐

Proof of Corollary 1.7. The result follows by taking $\mathbf{C}_2 = (\infty)$ in Theorem 1.6. ☐

Proof of Theorem 1.1. By Lemma 2.1, $f_*\mu$ is Stahl–Totik regular if and only if μ is $(f^{-1}(\infty))$-regular, and by Corollary 1.7, this is equivalent to Stahl–Totik regularity of μ. ☐

Proof of Theorem 1.8. (a) We note that by Corollary 1.7 we may use Theorem 1.4. Fix $1 \le k \le g+1$, and use conformal invariance to assume $\mathbf{c}_k = \infty$. Given a subsequence of $n(j) = j(g+1)+k$, we use precompactness to pass to a further subsequence $n_\ell = j_\ell(g+1)+k$ with $\text{w-lim}_{\ell\to\infty} \nu_{n_\ell} = \nu$. We write

$$\mathcal{G}_{\mathsf{E}}(z,\mathbf{C}) = \Phi_{\rho_{\mathsf{E},\mathbf{C}}}(z) + \frac{1}{g+1}\log\lambda_k - \frac{1}{g+1}\sum_{\substack{m=1\\ m\neq k}}^{g+1} \log|z-\mathbf{c}_m| \tag{5.8}$$

which we will use to show $\Phi_\nu = \Phi_{\rho_{\mathsf{E},\mathbf{C}}}$. By (ii), we may apply Theorem 4.1 with $\alpha = \frac{1}{g+1}\log\lambda_k$. Then, (vii) yields $h = \mathcal{G}_{\mathsf{E}}$ off the real line, and thus the equality between the representations (4.1) and (5.8) gives $\Phi_\nu(z) = \Phi_{\rho_{\mathsf{E},\mathbf{C}}}(z)$ on $\mathbb{C}\setminus\mathbb{R}$. By the weak identity principle, this equality extends to $\mathbb{C}$. Applying the distributional Laplacian to both sides gives $\nu = \rho_{\mathsf{E},\mathbf{C}}$. Thus, $\text{w-lim}_{n\to\infty}\nu_n = \rho_{\mathsf{E},\mathbf{C}}$.

(b) The main ingredient is a variant of Schnol's theorem; for any n, $\int |\tau_n|^2\, d\mu = 1$, so

$$\sum_{n=1}^{\infty} n^{-2}\int |\tau_n|^2\, d\mu < \infty.$$

By Tonelli's theorem, it follows that $\sum_{n=1}^{\infty} n^{-2}|\tau_n|^2 < \infty$ μ-a.e., so there exists a Borel set $B\subset\mathbb{C}$ with $\mu(\mathbb{C}\setminus B)=0$ such that

$$\limsup_{n\to\infty}\frac{1}{n}\log|\tau_n(z)| \le 0, \qquad \forall z\in B. \tag{5.9}$$

Suppose μ is not regular. Then, by Theorem 1.4(ii), there is a $1\le k\le g+1$ with

$$\limsup_{j\to\infty}\frac{1}{n(j)}\log\kappa_{n(j)} > \frac{1}{g+1}\log\lambda_k.$$

Using conformal invariance, we may assume $\mathbf{c}_k=\infty$, and we can pass to a subsequence $n_\ell = j_\ell(g+1)+k$ such that $\alpha := \lim_{\ell\to\infty}\frac{1}{n_\ell}\log\kappa_{n_\ell} > \frac{1}{g+1}\log\lambda_k$, where $\alpha\in\mathbb{R}\cup\{+\infty\}$ by Theorem 4.1(a). Since $\text{w-lim}_{n\to\infty}\nu_n = \rho_{\mathsf{E},\mathbf{C}}$, by comparing (4.1) and (5.8), we have for $z\in\mathbb{C}\setminus\mathbb{R}$,

$$\lim_{\ell\to\infty}\frac{1}{n_\ell}\log|\tau_{n_\ell}(z)| = \mathcal{G}_{\mathsf{E}}(z,\mathbf{C}) + d, \tag{5.10}$$

where $d = \alpha - \frac{\log\lambda_k}{g+1} > 0$. By the upper envelope theorem applied to the sequence $\{\nu_{n_\ell}\}_{\ell\in\mathbb{N}}$, there exists a polar set X such that (5.10) also holds for all $z\in\mathbb{C}\setminus X$. Moreover, since $\mathcal{G}_{\mathsf{E}}(z,\mathbf{C})\ge 0$ for all $z\in\mathbb{C}$, we conclude that

$$\limsup_{n\to\infty}\frac{1}{n}\log|\tau_n(z)| \ge \lim_{\ell\to\infty}\frac{1}{n_\ell}\log|\tau_{n_\ell}(z)| \ge d, \qquad \forall z\in\mathbb{C}\setminus X.$$

Comparing with (5.9) shows that $B\subset X$, so μ is supported on the polar set X. □

Proof of Theorem 1.9. Defining $n(j) = j(g+1)+k_\infty$ and using Lemma 3.1 to compute a telescoping product,

$$\left(\prod_{\ell=1}^{j}\beta_\ell\right)^{1/j} = \left(\prod_{\ell=1}^{j}\frac{\kappa_{n(\ell)}}{\kappa_{n(\ell+1)}}\right)^{1/j} = \kappa_{n(1)}^{1/j}\kappa_{n(j+1)}^{-1/j}. \tag{5.11}$$

The first term on the right-hand side is independent of j, so $\kappa_{n(1)}^{1/j}\to 1$ as $j\to\infty$. For the second factor, using Theorem 1.3 we compute

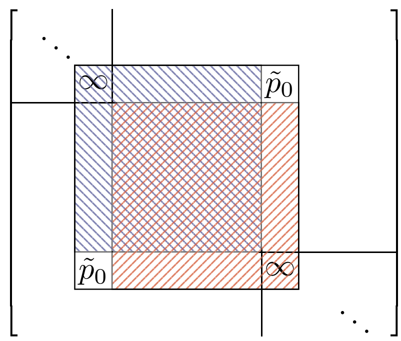
$$\liminf_{j\to\infty}\kappa_{n(j+1)}^{1/j} \ge \lambda_{k_\infty}$$

and we have the upper bound (1.9) for the lim sup of (5.11). Similarly, using the criterion Theorem 1.4(ii), it follows from (5.11) that μ is $\mathbf{C}$-regular if and only if (1.10) holds. □

6. GMP matrices 2

The proof of Theorem 1.10 will rely heavily on the results of [32]. In this section we will recall some properties of GMP matrices from [32] which we will use in the proof of Theorem 1.10. However, in order to justify the use of those constructions, we need to add some explanation of the structure of GMP matrices. This technical explanation is necessary in order to understand the action on Jacobi matrices caused by a single coefficient stripping step on GMP matrices; since such a step changes the location of ∞ in the sequence of poles, it links our GMP matrices which naturally arise from ORF expansions, and those in [32], which naturally arise from functional models of reflectionless operators. This link will allow us to use parts of the analysis of [32].

As noted in the beginning of Section 3, GMP matrices split up into blocks due to the appearance of some $c_{k_\infty} = \infty$. However, there is a choice whether to place the "window" of block size $(g+1)\times(g+1)$ so that c_{k_∞} is the last element of the previous block, or the first element of the next block. In this paper, the latter choice has been more natural (i.e., to split before ∞), because it corresponds to the choice $\tau_0 = 1$ in the rational function construction. From now on, we will call this the RF structure. On the other hand, in [32] the first choice was more natural (i.e., to split after ∞) for the functional model construction, and we will call this the FM structure. Alternatively, recall that we discussed in Remark 1 that one could view the GMP structure also as overlapping $(g+2)\times(g+2)$ blocks. The RF-structure then corresponds to placing the $(g+1)\times(g+1)$ B block at the upper left corner of the bigger block, whereas the FM structure corresponds to placing the B blocks at the lower right corner. This is shown in the figure below, where the blue lines indicate a B block corresponding to the RF structure and the red lines a B block corresponding to the FM structure. Moreover, $\tilde{p}_0$ denotes the positive entry on the outermost diagonal:



The two structures can be translated into each other, by means of the formulas (6.1). Moreover, we will show below that they are also linked by a coefficient stripping formula.

For the reader's convenience, we recall the FM structure of GMP matrices as introduced in [32]. Although the RF and FM structures are just a different interpretation of the same object, namely a GMP matrix, it will be convenient to have a separate notation. For a GMP matrix written in the FM structure we will use A, respectively for its blocks A_k, B_k, and we will use $\tilde{A}, \tilde{A}_k, \tilde{B}_k$, for GMP matrices written in the RF structure. Note that this is a change from the notation used in previous sections.

Fix a finite sequence $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_g)$ and recall that X^- denotes the upper triangular part of a matrix X (excluding the diagonal), and X^+ the lower triangular part (including the diagonal). Then we say that A acting on $\ell^2(\mathbb{Z})$ is GMP structured, and denote it by $A \in \mathbb{A}$, if it is a $(g+1)$-block Jacobi matrix

$$A = \begin{bmatrix} \ddots & \ddots & \ddots & & & \\ & A_{-1}^* & B_{-1} & A_0 & & \\ & & A_0^* & B_0 & A_1 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix}$$

such that

$$A_j = \delta_g \vec{p}_j^{\,*}, \quad B_j = (\vec{q}_j \vec{p}_j^{\,*})^- + (\vec{p}_j \vec{q}_j^{\,*})^+ + \hat{\mathbf{C}}, \quad \vec{p}_j, \vec{q}_j \in \mathbb{R}^{g+1},$$

and

$$\hat{\mathbf{C}} = \begin{bmatrix} \mathbf{c}_1 & & & \\ & \ddots & & \\ & & \mathbf{c}_g & \\ & & & 0 \end{bmatrix}, \ \vec{p}_j = \begin{bmatrix} p_0^{(j)} \\ \vdots \\ p_g^{(j)} \end{bmatrix}, \ \vec{q}_j = \begin{bmatrix} q_0^{(j)} \\ \vdots \\ q_g^{(j)} \end{bmatrix}, \quad p_g^{(j)} > 0.$$

We then say an operator $A \in \mathbb{A}$ is a two-sided GMP matrix if the resolvents $(\mathbf{c}_\ell - A)^{-1}$ exist for all $1 \le \ell \le g$ and $S^{-\ell}(\mathbf{c}_\ell - A)^{-1} S^\ell \in \mathbb{A}$. In this case we write $A \in \mathrm{GMP}(\mathbf{C})$. Again we call the generating coefficients $\{\vec{p}_j, \vec{q}_j\}_{j\in\mathbb{Z}}$ the GMP coefficients of A.

We encounter several differences compared to the RF structure presented in Section 3. First of all the 0 in $\hat{\mathbf{C}}$ is now in the last, rather than in the first, position. Moreover, in the definition of A_j, the vector $\vec{p}_j$ is now a row vector in the last row, rather than a column vector in the first column. This is exactly due to shifting the position of ∞ as described above. Extending the structure of GMP matrices to two-sided operators on $\ell^2(\mathbb{Z})$, it is not hard to see that the RF and the FM structures can be translated into each other, simply by shifting the window of size $(g+1)\times(g+1)$ by one. In this process the role of p_j and q_j changes, that is, for $1 \le k \le g$ we have

$$\tilde{p}_k^{(j)} = q_{k-1}^{(j)} p_g^{(j)}, \quad \tilde{q}_k^{(j)} = \frac{p_{k-1}^{(j)}}{p_g^{(j)}}. \tag{6.1}$$

More importantly for us is that the positive entries are the same, i.e.,

$$p_g^{(j)} = \tilde{p}_0^{(j)}.$$

Following [32], we explain how to associate to a given GMP matrix a Jacobi matrix. Let e_j denote the standard basis vectors in $\ell^2(\mathbb{Z})$; recall that $\{e_{-1}, e_0\}$ forms a spectral basis for two-sided Jacobi matrices in the sense that $\{J^n e_j \mid n \in \mathbb{N}_0, j = -1, 0\}$ is dense in $\ell^2(\mathbb{Z})$. Define the matrix resolvent function by

$$R^J(z) = \begin{bmatrix} \langle (J-z)^{-1} e_{-1}, e_{-1}\rangle & \langle (J-z)^{-1} e_0, e_{-1}\rangle \\ \langle (J-z)^{-1} e_{-1}, e_0\rangle & \langle (J-z)^{-1} e_0, e_0\rangle \end{bmatrix}.$$

Let $\ell^2_+ = \ell^2(\mathbb{N}_0)$ and $\ell^2_- = \ell^2(\mathbb{Z}) \ominus \ell^2_+$ and $\Pi_\pm$ denote the projection onto $\ell^2_\pm$. Define $J_\pm = \Pi_\pm J \Pi_\pm$ and define the half-line resolvent functions by

$$m_+^J(z) = \langle (J_+ - z)^{-1} e_0, e_0\rangle, \quad m_-^J(z) = \langle (J_- - z)^{-1} e_{-1}, e_{-1}\rangle.$$

Then, essentially due to the structure

$$J = \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix} + a_0(\langle \cdot, e_{-1}\rangle e_0 + \langle \cdot, e_0\rangle e_{-1}),$$

one can see that

$$R^J(z) = \begin{bmatrix} m_-^J(z)^{-1} & a_0 \\ a_0 & m_+^J(z)^{-1} \end{bmatrix}^{-1}; \tag{6.2}$$

cf. [9, pg 758].

For GMP matrices, we need to modify the spectral basis. Define

$$\tilde{e}_0 = \frac{1}{a_0}\Pi_+ A e_{-1}, \quad a_0 = \|\Pi_+ A e_{-1}\|, \tag{6.3}$$

with the natural embedding into $\ell^2(\mathbb{Z})$. Note that

$$a_0 \tilde{e}_0^\intercal = \begin{bmatrix} \dots & 0 & | & p_0^{(0)} & p_1^{(0)} & \dots & p_g^{(0)} & 0 & \dots \end{bmatrix}.$$

Then $\{e_{-1}, \tilde{e}_0\}$ form a spectral basis for A and similarly as for Jacobi matrices we have

$$A = \begin{bmatrix} A_- & 0 \\ 0 & A_+ \end{bmatrix} + a_0(\langle \cdot, e_{-1}\rangle \tilde{e}_0 + \langle \cdot, \tilde{e}_0\rangle e_{-1}).$$

This allows us to define

$$R^A(z) = \begin{bmatrix} \langle (A-z)^{-1}e_{-1}, e_{-1}\rangle & \langle (A-z)^{-1}\tilde{e}_0, e_{-1}\rangle \\ \langle (A-z)^{-1}e_{-1}, \tilde{e}_0\rangle & \langle (A-z)^{-1}\tilde{e}_0, \tilde{e}_0\rangle \end{bmatrix},$$

and

$$m_-^A(z) = \langle (A_- - z)^{-1}e_{-1}, e_{-1}\rangle, \quad m_+^A(z) = \langle (A_+ - z)^{-1}\tilde{e}_0, \tilde{e}_0\rangle \tag{6.4}$$

and find similar to the Jacobi case that

$$R^A(z) = \begin{bmatrix} m_-^A(z)^{-1} & a_0 \\ a_0 & m_+^A(z)^{-1} \end{bmatrix}^{-1}.$$

For a given GMP matrix A, the associated Jacobi matrix is simply defined by setting the resolvent functions to be equal, i.e.,

$$R^J(z) = R^A(z). \tag{6.5}$$

Note that this defines J uniquely. Due to the common vector e_{-1}, it follows that

$$\begin{gathered} b_{-1} = \langle Je_{-1}, e_{-1}\rangle = \langle Ae_{-1}, e_{-1}\rangle = p_g^{(-1)} q_g^{(-1)}, \\ a_0 = \|\Pi_+ J e_{-1}\| = \|\Pi_+ A e_{-1}\| = \|\vec{p}_0\|, \end{gathered} \tag{6.6}$$

which explains in hindsight the definition of a_0 in (6.3).

6.1. **Shifts on GMP matrices.** For a vector $x \in \ell^2(\mathbb{Z})$, let $|$ denote the splitting of ℓ^2_- and ℓ^2_+, i.e., we write $x^\intercal = \begin{bmatrix} \dots & x_{-1} | x_0 & \dots \end{bmatrix}$. We chose the vector of poles in the following way $\begin{bmatrix} \dots & \infty | \mathbf{c}_1 & \dots & \mathbf{c}_g & \infty & \mathbf{c}_1 & \dots \end{bmatrix}$. That is for $A_+ = \Pi_+ A \Pi_+$ the first pole is $\mathbf{c}_1 \in \mathbb{R}$. However, if we consider $\tilde{A}_+ = \Pi_+ SAS^{-1} \Pi_+$, where $Se_k = e_{k+1}$ denotes the right shift, then the first pole of $\tilde{A}_+$ is ∞.

$$\tilde{A}_+ = \left[\begin{array}{c|c} \infty & \\ \hline & \mathbf{c}_1 \\ & \qquad A_+ \\ & \end{array}\right].$$

The resolvent functions of A_+ and $\tilde{A}_+$ are related by a coefficient stripping formula:

Lemma 6.1. *Let $A \in \mathrm{GMP}(\mathbf{C})$, $A_+ = \Pi_+ A \Pi_+$, $\tilde{e}_0, a_0, b_{-1}$ as in* (6.3), (6.6) *and define $\tilde{A}_+ = \Pi_+ SAS^{-1} \Pi_+$. Then the resolvent functions*

$$m_+(z) = \langle (A_+ - z)^{-1} \tilde{e}_0, \tilde{e}_0 \rangle, \quad \tilde{m}_+(z) = \langle (\tilde{A}_+ - z)^{-1} e_0, e_0 \rangle$$

are related by the coefficient stripping formula

$$\tilde{m}_+(z) = \frac{1}{b_{-1} - z - a_0^2 m_+(z)}. \tag{6.7}$$

Proof. Recall that S_+ denotes the right shift on ℓ^2_+ and define

$$f_0 = S_+ \tilde{e}_0 = \frac{1}{a_0} \begin{bmatrix} 0 & p_0^{(0)} & p_1^{(0)} & \dots & p_g^{(0)} & 0 & \dots \end{bmatrix}.$$

Then we have

$$\tilde{A}_+ = \begin{bmatrix} b_{-1} & 0 \\ 0 & A_+ \end{bmatrix} + a_0 (\langle \cdot, e_0 \rangle f_0 + \langle \cdot, f_0 \rangle e_0).$$

Then as for Jacobi matrices this implies (6.7); cf. [24, Theorem 3.2.4]. □

Lemma 6.1 has a very natural interpretation. As we discussed above, GMP matrices split into blocks where $c_{k_\infty} = \infty$ and then there is some choice if we place ∞ as the last or the first element of a block. However, this discussion is irrelevant for Jacobi matrices, where all $c_k \equiv \infty$. Thus, if we associate to A a Jacobi matrix J by (6.5) and define $J_+ = \Pi_+ J \Pi_+$ and $\tilde{J}_+ = \Pi_+ SJS^{-1} \Pi_+$ and the associated m_+, $\tilde{m}_+$, then (6.7) becomes the standard coefficient stripping for Jacobi matrices.

There is another natural shift on GMP matrices. Namely, since the shift $A^{(1)} = S^{-(g+1)} A S^{(g+1)}$ preserves the GMP structure, one can describe how the resolvent functions of A and $A^{(1)}$ are related. This will be done by so-called elementary Blaschke-Potapov factors of the third kind with poles at $\mathbf{c}_1, \dots, \mathbf{c}_g, \infty$; cf. [1, 19]. In the following it will be convenient to use the notation

$$\mathbf{p} = (p, q), \quad \vec{\mathbf{p}} = (\vec{p}, \vec{q}).$$

Definition 6.2. For $p, q, \mathbf{c} \in \mathbb{R}$

(6.8)

$$\mathfrak{a}(z, \mathbf{c}; \mathbf{p}) = I - \frac{1}{\mathbf{c} - z} \begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p & q \end{bmatrix} \mathfrak{j} = \exp\left(-\frac{1}{\mathbf{c} - z} \begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p & q \end{bmatrix} \mathfrak{j} \right), \quad \mathfrak{j} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

represents the so-called Blaschke-Potapov factor of the third kind with a real pole $\mathbf{c}$. If $\mathbf{c} = \infty$ it is of the form

$$\mathfrak{a}(z;\mathbf{p}) = \mathfrak{a}(z,\infty;\mathbf{p}) = \begin{bmatrix} 0 & -p \\ \frac{1}{p} & \frac{z-pq}{p} \end{bmatrix}. \tag{6.9}$$

Define the matrix function

$$\mathfrak{A}(z,\vec{\mathbf{p}}) = \begin{bmatrix} \mathfrak{A}_{11} & \mathfrak{A}_{12} \\ \mathfrak{A}_{21} & \mathfrak{A}_{22} \end{bmatrix}(z,\vec{\mathbf{p}}) = \mathfrak{a}(z,\mathbf{c}_1;\mathbf{p}_0)\dots\mathfrak{a}(z,\mathbf{c}_g;\mathbf{p}_{g-1})\mathfrak{a}(z;\mathbf{p}_g). \tag{6.10}$$

The important role of the function $\mathfrak{A}(z,\vec{\mathbf{p}})$ will become clear by Theorem 6.3.

Theorem 6.3 ([32, Theorem 2.13 and Theorem 2.15])**.** *Let $A \in \mathrm{GMP}(\mathbf{C})$, $A^{(1)} = S^{-(g+1)}AS^{(g+1)}$ and A_+ and $A_+^{(1)}$ the projections onto ℓ^2_+. Let m_+^A and $m_+^{A^{(1)}}$ be the resolvent functions defined by* (6.4)*. Let $a_0^2 = \|\vec{p}_0\|^2$, $(a_0^{(1)})^2 = \|\vec{p}_0^{(1)}\|^2$. Then*

$$a_0^2 m_+^A(z) = \frac{\mathfrak{A}_{11}(z,\vec{\mathbf{p}}_0)((a_0^{(1)})^2 m_+^{A^{(1)}}(z)) + \mathfrak{A}_{12}(z,\vec{\mathbf{p}}_0)}{\mathfrak{A}_{21}(z,\vec{\mathbf{p}}_0)((a_0^{(1)})^2 m_+^{A^{(1)}}(z)) + \mathfrak{A}_{22}(z,\vec{\mathbf{p}}_0)}. \tag{6.11}$$

6.2. **Periodic GMP matrices.** We call a two-sided GMP matrix 1-periodic or simply periodic if $S^{g+1}AS^{-(g+1)} = A$. In this case $m_+^{A^{(1)}} = m_+^A$ and (6.11) is a quadratic equation for m_+^A. This allows to describe the spectrum of A in terms of the function $\mathfrak{A}(z)$.

Theorem 6.4 ([12, Theorem 1.8])**.** *Let $A = A(\vec{\mathbf{p}}) \in \mathrm{GMP}(\mathbf{C})$ be a periodic GMP matrix and $\mathfrak{A}(z,\vec{\mathbf{p}})$ be as in* (6.10) *and define the discriminant by*

$$\Delta(z) = \operatorname{tr}\mathfrak{A}(z,\vec{\mathbf{p}}).$$

Then the spectrum of A is a finite union of intervals, it is purely absolutely continuous and of multiplicity 2 *and it is given by*

$$\sigma(A) = \Delta^{-1}([-2,2]) = \{z \in \mathbb{C} |\ \Delta(z) \in [-2,2]\}.$$

The inverse problem can also be answered explicitly. Namely, given a finite union of intervals E, are there periodic GMP matrices with the given spectrum and if so can one describe the set of all such matrices? Crucially, the answer to both questions is positive for the special choice $\mathbf{C} = \mathbf{C}_\mathsf{E}$, where $\mathbf{C}_\mathsf{E}$ denotes the zeros of the Ahlfors function associated to E. We define the isospectral torus of periodic two-sided GMP matrices by

$$\mathcal{T}_\mathsf{E}(\mathbf{C}_\mathsf{E}) = \{\mathring{A} \in \mathrm{GMP}(\mathbf{C}_\mathsf{E}), \mathring{A} \text{ is periodic and } \sigma(\mathring{A}) = \mathsf{E}\}.$$

Henceforth, we will use $\mathring{A}$ for elements from $\mathcal{T}_\mathsf{E}(\mathbf{C}_\mathsf{E})$. We point out that for arbitrary finite gap sets, the isospectral torus of Jacobi matrices usually consists of almost periodic operators, whereas for GMP matrices we can always work with periodic operators. This also makes it possible to characterize the isospectral torus by a magic formula for GMP matrices. Moreover, this then can be used to describe $\mathcal{T}_\mathsf{E}(\mathbf{C}_\mathsf{E})$ also as an algebraic manifold. Recall that Λ_n denotes the outermost positive entry of the resolvents $(\mathbf{c}_\ell - A)^{-1}$. That is if $A \in \mathrm{GMP}(\mathbf{C})$ is a periodic GMP matrix let

$$\Lambda_\ell(A) = \langle e_\ell, (\mathbf{c}_{\ell+1} - A)^{-1} e_{\ell+g+1}\rangle \quad \text{for } 0 \le \ell \le g-1,$$
$$\Lambda_g(A) = \langle e_g, A e_{2g+1}\rangle.$$

The resolvent entries can again be given explicitly in terms of $\Delta(z)$.

Lemma 6.5 ([32, Theorem 2.17]). *Let $A \in \mathrm{GMP}(\mathbf{C})$ be a periodic GMP matrix. Then for $0 \le \ell \le g-1$*

(6.12)

$$\Lambda_\ell(A) = -(\operatorname{Res}_{\mathbf{c}_{\ell+1}} \Delta)^{-1}$$
$$= -\left(\operatorname{tr}\left(\prod_{k=0}^{\ell-1} \mathfrak{a}(\mathbf{c}_{\ell+1}, \mathbf{c}_{k+1}; \mathbf{p}_k)\begin{bmatrix} p_\ell \\ q_\ell \end{bmatrix}\begin{bmatrix} p_\ell & q_\ell \end{bmatrix}\mathfrak{j}\prod_{k=\ell+1}^{g-1} \mathfrak{a}(\mathbf{c}_{\ell+1}, \mathbf{c}_{k+1}; \mathbf{p}_k)\mathfrak{a}(\mathbf{c}_{\ell+1}, \mathbf{p}_g)\right)\right)^{-1}.$$

This allows to describe $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}})$ as an algebraic manifold. Let us fix a finite union of $g+1$ intervals and let Δ_{E} denote the associated discriminant defined in terms of the Ahlfors function (1.15). Then for coefficients $\vec{\mathbf{p}}$ let $A(\vec{\mathbf{p}}) \in \mathrm{GMP}(\mathbf{C}_{\mathsf{E}})$ be a periodic GMP matrix and define

$$f_0(\vec{\mathbf{p}}) = \lambda_{g+1}\langle \vec{p}, \vec{q}\rangle + d,$$
$$f_\ell(\vec{\mathbf{p}}) = \Lambda_{\ell-1}(A(\vec{\mathbf{p}}))\lambda_\ell - 1, \qquad \text{for } 1 \le \ell \le g+1,$$

and $\mathbf{F}_{\mathsf{E}} : U \subset \mathbb{R}^{2(g+1)} \to \mathbb{R}^{g+2}$ by

$$\mathbf{F}_{\mathsf{E}}(\vec{\mathbf{p}}) = \begin{pmatrix} f_0(\vec{\mathbf{p}}), & \dots & f_{g+1}(\vec{\mathbf{p}}) \end{pmatrix}. \tag{6.13}$$

We then define the isospectral manifold by

$$\mathcal{IS}_{\mathsf{E}} = \{\vec{\mathbf{p}} \in \mathbb{R}^{2g} : \mathbf{F}_{\mathsf{E}}(\vec{\mathbf{p}}) = 0\}.$$

The name is justified by Theorem 6.6:

Theorem 6.6 ([12, Theorem 1.6 and Theorem 1.10]). *Let $A \in \mathrm{GMP}(\mathbf{C}_{\mathsf{E}})$, then*

$$A \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}) \iff \Delta_{\mathsf{E}}(A) = S^{g+1} + S^{-(g+1)}. \tag{6.14}$$

Moreover, for $\vec{\mathbf{p}}$ such that $A(\vec{\mathbf{p}}) \in \mathrm{GMP}(\mathbf{C}_{\mathsf{E}})$ we have that

$$A(\vec{\mathbf{p}}) \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}}) \iff \mathbf{F}_{\mathsf{E}}(\vec{\mathbf{p}}) = 0.$$

6.3. **Resolvents in the general case and the Jacobi flow.** Similar to (6.12) one can also find explicit expressions for Λ_n for general (not necessarily periodic) GMP matrices. Let $A \in \mathrm{GMP}(\mathbf{C})$ and for $n = j(g+1) + \ell$ for $j \in \mathbb{Z}$ and $0 \le \ell \le g$ set

$$\Lambda_n(A) = \begin{cases} \langle e_n, (\mathbf{c}_{\ell+1} - A)^{-1} e_{n+g+1}\rangle, & \ell \ne g, \\ \langle e_n, A e_{n+g+1}\rangle. \end{cases}$$

Lemma 6.7 ([32, Lemma 3.2]). *Let $A \in \mathrm{GMP}(\mathbf{C})$. Then for $n = j(g+1) + \ell$ and $\ell \ne g$ we have*

(6.15)

$$\Lambda_n(A)$$
$$= -\left(\operatorname{tr}\left(\prod_{k=0}^{\ell-1} \mathfrak{a}(\mathbf{c}_{\ell+1}, \mathbf{c}_{k+1}; \mathbf{p}_k^{(j+1)})\begin{bmatrix} p_\ell^{(j+1)} \\ q_\ell^{(j+1)} \end{bmatrix}\begin{bmatrix} p_\ell^{(j)} & q_\ell^{(j)} \end{bmatrix}\mathfrak{j}\prod_{k=\ell+1}^{g-1} \mathfrak{a}(\mathbf{c}_{\ell+1}, \mathbf{c}_{k+1}; \mathbf{p}_k^{(j)})\mathfrak{a}(\mathbf{c}_{\ell+1}, \mathbf{p}_g^{(j)})\right)\right)^{-1}.$$

The explicit representation will be crucial in the following. Moreover, let us mention that due to the finite band block structure of GMP matrices, building (formal) resolvents is a purely local computation (compare e.g. [32, eq (3.8)]). This can be seen by the formula above, where only the entries of A from the blocks j and $j+1$ are needed to compute Λ_n.

Recall that we discussed already in the beginning of this section that to any GMP matrix A we can associate a Jacobi matrix J, namely by setting the resolvent functions equal to (6.5). Let us denote this map by $\mathcal{F}$. It is a deep result from [32, Proposition 5.5.] that this map is (up to a certain identification) invertible. An important question is if we can express the Jacobi parameters of $J = \mathcal{F}A$ in terms of the coefficients of A. Let $\{\vec{\mathbf{p}}_j\}$ denote the GMP coefficients and $\{a_j, b_j\}$ the Jacobi coefficients. Then we have already seen that

$$a_0 = \|\vec{p}_0\|, \quad b_{-1} = q_g^{(-1)} p_g^{(-1)}.$$

Let

$$\mathcal{S}J = S^{-1}JS,$$

and note that

$$a_0(\mathcal{S}J) = a_1(J), \quad b_{-1}(\mathcal{S}J) = b_0(J),$$

where by $a_k(J), b_k(J)$ we mean the Jacobi parameters of the Jacobi matrix J. Thus, if one understands the transform on GMP matrices which is induced by the shift action on Jacobi matrices, one can inductively obtain the Jacobi parameters by the formulas above. This leads to the definition of the Jacobi flow on GMP matrices, which is defined by the following commutative diagram:

$$\begin{array}{ccc} \text{GMP} & \xrightarrow{\mathcal{J}} & \text{GMP} \\ {}_{\mathcal{F}}\big\downarrow & & {}_{\mathcal{F}}\big\downarrow \\ \text{Jacobi} & \xrightarrow{\mathcal{S}} & \text{Jacobi} \end{array} \tag{6.16}$$

Let us mention that this is one of the reasons why it is convenient to work with two-sided operators. If in this construction we considered the shift action on ℓ^2_+, which is not unitary, then it is possible that for some m, $\mathbf{c}_k \in \sigma((S_+^*)^m J_+ S_+^m)$, and thus the corresponding half-line GMP matrix would not be well defined.

The Jacobi flow is defined and discussed in [32, Section 4]. We provide the motivating ideas of the Jacobi flow and its precise definition below. First, note that in [32], we have the ordering of the poles

$$\mathbf{C}_A := \left[\ \dots \quad \infty \mid \mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_g \quad \infty \ \vdots\ \mathbf{c}_1 \quad \dots\ \right]$$

and recall that we anchored the blocks between ∞ (at position -1) and $\mathbf{c}_1$ (at position 0). Note that for Jacobi matrices, all poles are equal to ∞, and $\mathcal{S}J$ corresponds to shifting an ∞-pole from position 0 to position -1. Now applying the spacial shift to GMP matrices would be of a different flavor, as it shifts $\mathbf{c}_1$ from 0 to -1. Thus, one first has to shift ∞, which is now at position $g+1$, to the front, and then one may apply the spacial shift. This is done in g-steps. The $\mathcal{O}$ transform defined below corresponds changing the order from $\mathbf{C}_A$ to

$$\tilde{\mathbf{C}}_A = \left[\ \dots \quad \mathbf{c}_g \mid \mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_{g-1} \quad \infty \quad \mathbf{c}_g \ \vdots\ \mathbf{c}_1 \dots\ \right].$$

Letting

$$\mathbf{o}(\phi) = \begin{bmatrix} \sin\phi & \cos\phi \\ \cos\phi & -\sin\phi \end{bmatrix},$$

we make Definition 6.8.

Definition 6.8. We define the map:

$$\mathcal{O} : \mathrm{GMP}(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_g) \to \mathrm{GMP}(\mathbf{c}_g, \mathbf{c}_1, \dots, \mathbf{c}_{g-1})$$

in the following way. Let $O = O_A$ be the block-diagonal matrix

$$O = \begin{bmatrix} \ddots & & & \\ & O_{-1} & & \\ & & O_0 & \\ & & & \ddots \end{bmatrix},$$

where O_k are the $(g+1) \times (g+1)$ orthogonal matrices

$$O_k = \begin{bmatrix} I_{g-2} & 0 \\ 0 & \mathbf{o}(\phi_k) \end{bmatrix}, \quad \begin{bmatrix} \sin\phi_k & \cos\phi_k \end{bmatrix} = \frac{\begin{bmatrix} p_{g-1}^{(k)} & p_g^{(k)} \end{bmatrix}}{\sqrt{(p_{g-1}^{(k)})^2 + (p_g^{(k)})^2}}. \tag{6.17}$$

Then

$$\mathcal{O}A := SO_A^* A O_A S^{-1}. \tag{6.18}$$

As explained above, the Jacobi flow then is defined by applying $\mathcal{O}$ g-times, in order to shift ∞ through the full block. This leads to Definition 6.9:

Definition 6.9. We define the Jacobi flow transform

$$\mathcal{J} : \mathrm{GMP}(\mathbf{C}_A) \to \mathrm{GMP}(\mathbf{C}_A)$$

by

$$\mathcal{J}A := S^{-(g+1)} \mathcal{O}^{\circ g} A S^{g+1}.$$

It is shown in [32, Equation (4.8) and Lemma 4.4] that there exists a block-diagonal unitary mapping U_A, such that

$$\mathcal{J}A = S^{-1} U_A^* A U_A S. \tag{6.19}$$

Let us also note that

$$S^{-(g+1)} \mathcal{O}(A) S^{(g+1)} = \mathcal{O}(S^{-(g+1)} A S^{(g+1)}), \tag{6.20}$$

which has the consequences

$$\mathcal{O}(\mathcal{J}^{\circ m} A) = \mathcal{J}^{\circ m}(\mathcal{O}A) \tag{6.21}$$

and

$$S^{-(g+1)} (\mathcal{J}^{\circ m} A) S^{(g+1)} = \mathcal{J}^{\circ m}(S^{-(g+1)} A S^{(g+1)}).$$

7. Proof of Theorem 1.10

Lemma 7.1 allows us to extend J_+ to a two-sided Jacobi matrix J acting on $\ell^2(\mathbb{Z})$ in a way such that $\mathbf{c}_k$ belong to the resolvent domain of J.

Lemma 7.1. *Let μ be a compactly supported probability measure such that $\mathsf{E} = \operatorname{ess\,supp} \mu$ is a union of $g+1$ intervals. Let*

$$m_+(z) = \int_{-\infty}^{\infty} \frac{1}{x-z} d\mu(x)$$

and J_+ the associated Jacobi matrix and let $\mathbf{c}_k \in \mathbb{R} \setminus \mathsf{E}$ for $1 \le k \le g$. Then there exists a two-sided Jacobi matrix J with the following properties:

(i) $J_+ = \Pi_+ J \Pi_+$;

(ii) *there exists* $\mathring{J} \in \mathcal{T}(\mathsf{E})$ *so that* $J_- := \Pi_- S J S^{-1} \Pi_-$ *obeys* $J_- = \Pi_- S \mathring{J} S^{-1} \Pi_-$*;*
(iii) $\mathbf{c}_k$ *belong to the resolvent domain of* J*;*
(iv) $\mathbf{c}_k$ *belong to the resolvent domain of* $\tilde{J}_+ := \Pi_+ S J S^{-1} \Pi_+$.

Proof. Let J denote the extended two-sided matrix. Note that J is defined by J_+, a_0, a_{-1}, b_{-1} and J_-. We fix J_+ and a_0, a_{-1} and choose J_- and b_{-1} appropriately.

By (6.7) we have

$$\tilde{m}_+(z) = \frac{1}{b_{-1} - z - a_0^2 m_+(z)}.$$

Thus, $\mathbf{c}_k$ is a pole of $\tilde{m}_+(z)$ if and only if it is a zero of $b_{-1} - z - a_0^2 m_+(z)$. Choose b_{-1} so that

$$b_{-1} - \mathbf{c}_k - a_0^2 m_+(\mathbf{c}_k) \neq 0.$$

This already defines $\tilde{J}_+$.

Let us write (6.2) at position -1 rather than at position 0 and let m_- be the resolvent function of J_- and $\tilde{m}_+$ the resolvent function of $\tilde{J}_+$. Then we see that

$$-\frac{1}{R_{-2-2}(z)} = -\frac{1}{m_-(z)} + a_0^2 \tilde{m}_+(z), \quad -\frac{1}{R_{-1,-1}(z)} = -\frac{1}{\tilde{m}_+(z)} + a_0^2 m_-(z).$$

If $\tilde{m}_+(\mathbf{c}_k) \in \{0, \infty\}$, we choose m_- so that $m_-(\mathbf{c}_k) \notin \{0, \infty\}$ and if $\tilde{m}_+(\mathbf{c}_k) \notin \{0, \infty\}$ we set $m_-(\mathbf{c}_k) = 0$. In both cases $R_{-2,-2}(\mathbf{c}_k) \neq \infty$ and $R_{-1,-1}(\mathbf{c}_k) \neq \infty$ and we obtain (iii). □

We will apply Lemma 7.1 in the following way. First we choose $\mathbf{c}_1, \dots \mathbf{c}_g$ as the zeros of the Ahlfors function of $\overline{\mathbb{C}} \setminus \mathsf{E}$. Let μ be a given Stahl-Totik regular measure and $\mathsf{E} = \operatorname{ess\,supp} \mu$. To this measure we construct J as above. Let further $\tilde{\mu}$ be the spectral measure of $\tilde{J}_+$. Clearly $\mathsf{E} = \operatorname{ess\,supp} \tilde{\mu}$ and from the characterization of regularity by existence of the limit and equality in (1.2) it follows that also $\tilde{\mu}$ is regular. Due to (iii) we can form orthogonal rational functions with respect to the periodic sequence $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_g, \infty, \mathbf{c}_1, \mathbf{c}_2, \dots)$. On the other hand (iv) allows us to associate to J a two-sided GMP matrix in the sense of [32]. In particular, $\tilde{J}_+$ satisfies the assumptions of Lemma 1.11.

It was noted in [32, Section 2.2] that

$$-\log|\Psi(z)| = \sum_{k=1}^{g+1} G_{\mathsf{E}}(z, \mathbf{c}_k) \tag{7.1}$$

and that the Yuditskii discriminant has the form (1.15) for some $\lambda_k > 0$ and $d \in \mathbb{R}$. Note that the constants λ_k can be found by computing the residue of Δ_{E} at the poles $\mathbf{c}_k$. By using (1.14) and (7.1), we find the residues to be the same constants λ_k defined in a more general setting in (1.6).

Proof of Lemma 1.11. Denote by μ the canonical spectral measure for J. Note that

$$\sigma_{\mathrm{ess}}(A) = \operatorname{ess\,supp} \mu = \mathsf{E} = \Delta_{\mathsf{E}}^{-1}([-2, 2]).$$

Since Δ_{E} maps $\mathbb{R} \setminus \{\mathbf{c}_1, \dots, \mathbf{c}_g\}$ to $\mathbb{R}$ and is piecewise strictly monotone, by a spectral mapping theorem, this implies that for $\mathbf{J} = \Delta_{\mathsf{E}}(A)$, $\sigma_{\mathrm{ess}}(\mathbf{J}) = [-2, 2]$.

As noted in Section 1, regularity of the Jacobi matrix J implies $\mathbf{C}_{\mathsf{E}}$-regularity by Corollary 1.7, and this can be characterized in terms of GMP matrix coefficients by Theorem 1.9. The GMP matrix structure together with (1.15) implies that

$\mathbf{J} = \Delta_{\mathsf{E}}(A)$ is a type 3 block Jacobi matrix (1.16); the diagonal entries of the off-diagonal blocks $\mathfrak{v}_j$ are given by $\lambda_k \Lambda_{j(g+1)+k}$ for $k = 0, \dots, g$, with the convention $\lambda_0 = \lambda_{g+1}$. Thus,

$$\det \mathfrak{v}_j = \prod_{k=0}^{g} \lambda_k \Lambda_{j(g+1)+k}.$$

By applying the criterion for regularity in Theorem 1.9 to the GMP matrix A and to its resolvents $(\mathbf{c}_k - A)^{-1}$, we conclude that $\mathbf{J}$ obeys (1.17). It follows that $\mathbf{J}$ is regular with $\sigma_{\text{ess}}(\mathbf{J}) = [-2, 2]$. □

If $\tilde{A}_+$ is such that $\sigma_{\text{ess}}(\tilde{A}_+) = \mathsf{E}$ and the corresponding measure is regular on E, then $\Delta_{\mathsf{E}}(\tilde{A}_+)$ is a block Jacobi matrix which due to Lemma 1.11 is regular for $[-2, 2]$. Therefore, if $\{\mathfrak{v}_\ell, \mathfrak{w}_\ell\}$ denote the block Jacobi coefficients of $\Delta_{\mathsf{E}}(\tilde{A}_+)$, by [23, Theorem 3.1] we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{\ell=1}^{N} \|\mathfrak{v}_\ell - I\| + \|\mathfrak{w}_\ell\| = 0. \tag{7.2}$$

We note that since $C = \sup_\ell(\|\mathfrak{v}_\ell(A) - I\| + \|\mathfrak{w}_\ell\|) < \infty$, it follows from Cauchy-Schwarz and the AM-GM inequality that

$$\left(\frac{1}{N} \sum_{\ell=1}^{N} \|\mathfrak{v}_\ell - I\| + \|\mathfrak{w}_\ell\|\right)^2 \leq \frac{2}{N} \sum_{\ell=1}^{N} \|\mathfrak{v}_\ell - I\|^2 + \|\mathfrak{w}_\ell\|^2 \leq 2C \frac{1}{N} \sum_{\ell=1}^{N} \|\mathfrak{v}_\ell - I\| + \|\mathfrak{w}_\ell\|$$

and thus

$$\lim_{N \to \infty} \frac{1}{N} \sum_{\ell=1}^{N} \|\mathfrak{w}_\ell\|^2 + \|\mathfrak{v}_\ell - I\|^2 = 0 \iff \lim_{N \to \infty} \frac{1}{N} \sum_{\ell=1}^{N} \|\mathfrak{w}_\ell\| + \|\mathfrak{v}_\ell - I\| = 0. \tag{7.3}$$

We will use this equivalence freely in the following.

In the setting of periodic Jacobi matrices and polynomial discriminants (i.e., Δ is a polynomial and $\{\mathfrak{v}_\ell, \mathfrak{w}_\ell\}$ are the coefficients of the block Jacobi matrix $\Delta(J_+)$) it is shown in [7] that

$$\sum_{\ell=1}^{\infty} \|\mathfrak{w}_\ell\|^2 + \|\mathfrak{v}_\ell - I\|^2 < \infty \iff \sum_{m=1}^{\infty} d((S_+^*)^m J S_+^m, \mathcal{T}_{\mathsf{E}}^+)^2 < \infty. \tag{7.4}$$

It was then stated in [23] that since all the arguments in [7] are local, in this setting (7.3) yields (1.13). Let us emphasize that finite gap sets whose isospectral torus consists of periodic Jacobi matrices are very special and the arguments in [23] only apply to this setting. Yuditskii [32] has extended the work of [7] and one has the same localness, but since the construction is quite involved, we will provide the main ideas of proof. In this case, the condition on the right-hand side of (7.4) is still the same, i.e., a condition for a Jacobi matrix J_+, but on the left-hand side $\{\mathfrak{v}_\ell, \mathfrak{w}_\ell\}$ are the coefficients of the block Jacobi matrix $\Delta_{\mathsf{E}}(A)$, where A is an associated GMP matrix and Δ_{E} is the rational function as defined in (1.15).

We will start with the main ingredients of the proof that the left-hand side in (7.4) implies the right-hand side and mention certain modifications to our setting. After this preparatory work will show how this can be applied to our setting.

We concluded from regularity that $\mathbf{J} = \Delta_{\mathsf{E}}(\tilde{A}_+)$ satisfies (7.2). As may be seen in [7], and [32], it is convenient to rewrite this condition into a "multiplicative form".

This leads to the notion of the Killip-Simon functional that we will define below. For a GMP matrix $A \in \mathrm{GMP}(\mathbf{C}_{\mathsf{E}})$, we define the functional as in [32, Section 6] by

$$H_+(A) = \sum_{\ell=0}^{\infty} h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}), \tag{7.5}$$

where

$$h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}) = \frac{1}{2}\operatorname{tr}(\mathfrak{v}_\ell^*\mathfrak{v}_\ell + \mathfrak{w}_\ell^2 + \mathfrak{v}_{\ell+1}\mathfrak{v}_{\ell+1}^*) - (g+1) - \log\det \mathfrak{v}_\ell\mathfrak{v}_{\ell+1}.$$

For a square matrix X its modulus is defined by $|X| := \sqrt{X^*X}$. Moreover, define $G(|X|) = |X|^2 - I - \log|X|^2$. Then we have

$$2h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}) = \operatorname{tr}\left(\mathfrak{w}_\ell^2 + G(|\mathfrak{v}_\ell|) + G(|\mathfrak{v}_{\ell+1}^*|)\right).$$

In particular, it follows from $|\mathfrak{v}_\ell|, |\mathfrak{v}_{\ell+1}| > 0$ that $h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}) > 0$. In fact even more is true. There exists $\tilde{C} > 1$ so that if $\|\mathfrak{v}_\ell - I\| < \frac{1}{2}$ then by [7, Proposition 11.12]

$$\frac{1}{\tilde{C}}\|\mathfrak{v}_\ell - I\| \leq \||\mathfrak{v}_\ell| - I\| \leq \tilde{C}\|\mathfrak{v}_\ell - I\|.$$

Thus, if $\tilde{C}\|\mathfrak{v}_\ell - I\| < \frac{1}{2}$ we conclude that $\||\mathfrak{v}_\ell| - I\| < \frac{1}{2}$ and thus the eigenvalues of $|\mathfrak{v}_\ell|$ are greater than $\frac{1}{2}$. Under this assumption (for ℓ and $\ell+1$) it is shown in [7, Theorem 11.13] that there exists a constant C so that

$$\frac{1}{C}h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}) \leq \left(\|\mathfrak{v}_\ell - I\|^2 + \|\mathfrak{w}_\ell\|^2 + \|\mathfrak{v}_{\ell+1} - I\|^2\right) \leq Ch(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}). \tag{7.6}$$

A key observation is that the functional $H_+(A)$ is related to the shift action of S^{g+1} on the GMP matrix A. But finally we want to conclude something about

$$\mathcal{S}J = S^*JS,$$

i.e., the shift action on J. This is another motivation of the Jacobi flow as defined above.

The following key lemma, which follows essentially from (6.19), allows for the computation of the "derivative" in the Jacobi flow direction, and is essential in order to extract from the finiteness of $H_+(A)$ properties of the associated Jacobi matrix J.

Lemma 7.2 ([32, Lemma 6.1]). *Let $v_{jk}^{(\ell)}, w_{jk}^{(\ell)}$ denote the matrix entries of $\mathfrak{v}_\ell, \mathfrak{w}_\ell$ and*

$$\delta_J H_+(A) = \frac{1}{2}\langle \Delta_{\mathsf{E}}(\mathcal{J}A)e_{-1}, \Delta_{\mathsf{E}}(\mathcal{J}A)e_{-1}\rangle - 1 - \log(\mathcal{J}v)_{g,g}^{(-1)}(\mathcal{J}v)_{g,g}^{(0)}.$$

Then

$$H_+(A) = H_+(\mathcal{J}A) + \delta_J H_+(A). \tag{7.7}$$

Proof. Using (6.19) the proof is based on the realization that due to the diagonal structure of U_A, conjugating A by U_A does not affect H_+. Thus, $\delta_J H_+(A)$ corrects for the term which is omitted in $H_+(\mathcal{J}A)$ due to the shift. □

For later reference let us mention that due to (6.18) we can also relate $H_+(A)$ and $H_+(\tilde{O}A)$. Moreover, it is easy to see that we can also relate $H_+(A)$ and

$H_+(S^{-(g+1)}AS^{(g+1)})$ explicitly. Moreover, Lemma 7.2 allows to obtain ℓ^2 conditions for the coefficients of $\mathcal{J}^{\circ m}(A)$ from finiteness of $H_+(A)$. We sketch the idea in the following. Let us define

$$\tilde{H}_+(A) = \sum_{m=0}^{\infty} \delta_J H_+(A(m)), \qquad \text{where } A(m) = \mathcal{J}^{\circ m}(A).$$

Since all terms are positive, iterating (7.7) yields

$$\tilde{H}_+(A) \leq H_+(A).$$

In particular, $H_+(A) < \infty$ implies $\tilde{H}_+(A) < \infty$. The vector $\Delta_{\mathsf{E}}(A(m))e_{-1}$ has only $2g+3$ nonvanishing entries which are entries of the last columns of $\mathfrak{v}_{-1}(m), \mathfrak{w}_{-1}(m)$ and $\mathfrak{v}_0(m)$. Let us denote this $2g+3$-dimensional vector by $x(m)$ and note that the first and the last components are the positive entries $x_0(m) = (v(m))^{(-1)}_{g,g}$ and $x_{2g+2}(m) = (v(m))^{(0)}_{g,g}$. With this notation we have

$$\delta_J H_+(A(m)) = \frac{1}{2}\left(G(x_0(m)) + G(x_{2g+2}(m)) + \sum_{j=1}^{2g+1} x_j(m)^2 \right). \tag{7.8}$$

Thus, $\tilde{H}_+(A) < \infty$ implies already ℓ^2-conditions for the vector $x(m)$. This is used to conclude from $\tilde{H}_+(A) < \infty$ that $A(m)$ is ℓ^2-close to be periodic and that the periodic operator is ℓ^2-close to $\mathcal{IS}_{\mathsf{E}}$. That is, if $\{\vec{p}_j(m), \vec{q}_j(m)\}_{m\in\mathbb{N}_0}$ denote the GMP parameters of $A(m)$, then [32, Theorem 1.20]

$$\begin{aligned} \{\vec{\mathbf{p}}_0(m) - \vec{\mathbf{p}}_{-1}(m)\}_{m\in\mathbb{N}_0} &\in \ell^2(\mathbb{N}_0, \mathbb{R}^{2(g+1)}), \\ \{\mathbf{F}_{\mathsf{E}}(\vec{\mathbf{p}}_0(m))\}_{m\in\mathbb{N}_0} &\in \ell^2(\mathbb{N}_0, \mathbb{R}^{g+2}). \end{aligned} \tag{7.9}$$

To show how one obtains from (7.9) convergence of $(S_+^*)^m J S_+^m$ to $\mathcal{T}_{\mathsf{E}}^+$ in the sense of (7.4), we need one more ingredient: it is well known that there are continuous functions, $\mathcal{A}, \mathcal{B}$, on $\mathbb{R}^g/\mathbb{Z}^g$, which can be expressed explicitly in terms of the Riemann theta function associated to E[1] [28, Theorem 9.4.], and a fixed element $\chi \in \mathbb{R}^g/\mathbb{Z}^g$, such that

$$\mathcal{T}_{\mathsf{E}} = \{J(\alpha) : \quad \alpha \in \mathbb{R}^g/\mathbb{Z}^g\} \tag{7.10}$$

and $J(\alpha)$ is the Jacobi matrix built from the coefficients

$$a_m(\alpha) = \mathcal{A}(\alpha - m\chi), \quad b_m(\alpha) = \mathcal{B}(\alpha - m\chi). \tag{7.11}$$

Recall that by the definition of the Jacobi flow, if J is the Jacobi matrix associated to A, then $S^{-m}JS^m$ is the Jacobi matrix associated to $A(m)$. Since every point of $\mathcal{IS}_{\mathsf{E}} = \mathbf{F}_{\mathsf{E}}^{-1}(0)$ is regular for $\mathbf{F}_{\mathsf{E}}$, by [7, Lemma 11.3] there exists a constant $C > 0$, such that

$$\operatorname{dist}(\vec{\mathbf{p}}, \mathcal{IS}_{\mathsf{E}}) \leq C\|\mathbf{F}_{\mathsf{E}}(\vec{\mathbf{p}})\|, \tag{7.12}$$

where $\vec{\mathbf{p}}$ are chosen from a fixed compact neighborhood of $\mathcal{IS}_{\mathsf{E}}$, see also [32, page 755]. Taking an element $\mathring{A}_m \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}})$ so that

$$\operatorname{dist}(\vec{\mathbf{p}}(n), \mathcal{IS}_{\mathsf{E}}) = \operatorname{dist}(A(\vec{\mathbf{p}}(n)), \mathring{A}_m),$$

[1] To be precise it is the Riemann theta function of the Riemann surface of the function $\sqrt{\prod_{k=0}^{g}(z-\mathbf{a}_k)(z-\mathbf{b}_k)}$, where $\mathbf{a}_k, \mathbf{b}_k$ denote the gap edges of E, cf. (1.11).

one can conclude from (7.9) and (7.12) that

$$\sum_{n\geq 0} \operatorname{dist}(A(\vec{\mathbf{p}}(n)), \mathring{A}_m)^2 < \infty. \tag{7.13}$$

Letting $J(\alpha_m) \in \mathcal{T}_{\mathsf{E}}$ be the Jacobi matrix with $F(\mathring{A}_m) = J(\alpha_m)$, then (6.6) implies that

$$a(m)^2 - \mathcal{A}(\alpha_m) \in \ell^2_+, \quad b(m) - \mathcal{B}(\alpha_m) \in \ell^2_+.$$

Using in addition the smoothness of the Jacobi flow, one can show that

$$\alpha_n = \sum_{j=1}^{n} \varepsilon^\alpha_m - m\chi, \quad \varepsilon^\alpha_m \in \ell^2(\mathbb{N}_0, \mathbb{R}^g).$$

This is even stronger than (7.4); cf. [32, Lemma 7.2].

Before we start with our construction, we have to mention a certain technical issue. If $\{f_m\}$ is a sequence, then clearly $\{f_m - 1\} \in \ell^2$ implies

$$\liminf_{m\to\infty} f_m > 0.$$

If $|f_m - 1|$ is only Cesáro summable, then this is not necessarily the case. However, for any $\delta > 0$ the set with $f_m < \delta$ will be sparse in the following sense. Let us introduce the notation $\{f_m\} \in \mathrm{CS}$ for sequences $\{f_m\}$ satisfying

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} |f_m| = 0,$$

and we call a set $T \subset \mathbb{N}$ sparse if

$$\lim_{N\to\infty} \frac{|T \cap \{1, 2, \ldots, N\}|}{N} = 0.$$

An elementary observation, which will be used repeatedly, is that for $f \in \mathrm{CS}$, the set $\{m \in \mathbb{N} \mid |f_m| \geq \delta\}$ is sparse for any $\delta > 0$. This follows immediately from Markov's inequality.

We have already concluded from regularity that one and hence both of the conditions in (7.3) hold. Due to the phenomena described above and the log in the definition of $h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1})$ it is not immediately clear that (7.3) also implies

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\ell=0}^{N} h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}) = 0. \tag{7.14}$$

However, using in addition once again regularity, we can show (7.14).

Lemma 7.3. *Let a Jacobi matrix satisfy the conditions of Lemma* 1.11 *and* $\{\mathfrak{v}_\ell, \mathfrak{w}_\ell\}$ *denote the coefficients of the associated block Jacobi matrix* $\mathbf{J} = \Delta_{\mathsf{E}}(A)$. *Then* (7.14) *holds.*

Proof. Recall that

$$h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1}) = \frac{1}{2} \operatorname{tr}\left((|\mathfrak{v}_\ell|^2 - I) + (|\mathfrak{v}^*_{\ell+1}|^2 - I) + \mathfrak{w}^2_\ell\right) - \log\det \mathfrak{v}_\ell \mathfrak{v}_{\ell+1}.$$

Regularity allows us to consider the terms in $h(\mathfrak{v}_\ell, \mathfrak{w}_\ell, \mathfrak{v}_{\ell+1})$ separately. It follows directly from (7.3) that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\ell=0}^{N} \operatorname{tr} \mathfrak{w}^2_\ell = 0.$$

Moreover, (1.17) implies that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\ell=0}^{N}\log\det\mathfrak{v}_\ell=0.$$

Thus it remains to show that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\ell=0}^{N}\operatorname{tr}(|\mathfrak{v}_\ell|^2-I)=0,\quad \lim_{N\to\infty}\frac{1}{N}\sum_{\ell=0}^{N}\operatorname{tr}(|\mathfrak{v}_\ell^*|^2-I)=0. \tag{7.15}$$

For a matrix $A\in\operatorname{Mat}(n,\mathbb{R})$, let $\sigma_i(A)$ denote its singular values and note that $\operatorname{tr}|A|=\sum\sigma_i(A)$. For $A,B\in\operatorname{Mat}(n,\mathbb{R})$ we will need the following inequalities

$$|\operatorname{tr}A|\leq\operatorname{tr}|A|,$$
$$\sum_{i=1}^{n}\sigma_i(AB)\leq\sum_{i=1}^{n}\sigma_i(A)\sigma_i(B), \tag{7.16}$$

which can be found for instance in [16, eq. (3.3.35) and Theorem 3.3.14]. Thus we have

$$\left|\frac{1}{N}\sum_{\ell=0}^{N}\operatorname{tr}(|\mathfrak{v}_\ell|^2-I)\right|\leq\frac{1}{N}\sum_{\ell=0}^{N}\operatorname{tr}(\big||\mathfrak{v}_\ell|^2-I\big|)=\frac{1}{N}\sum_{\ell=0}^{N}\operatorname{tr}(|(|\mathfrak{v}_\ell|-I)(|\mathfrak{v}_\ell|+I)|).$$

Using (7.16) and a uniform bound on $\sigma_i(|\mathfrak{v}_\ell|+I)$ we get

$$\operatorname{tr}(|(|\mathfrak{v}_\ell|-I)(|\mathfrak{v}_\ell|+I)|)=\sum_{j=0}^{g}\sigma_j((|\mathfrak{v}_\ell|-I)(|\mathfrak{v}_\ell|+I))\leq C\sum_{j=0}^{g}\sigma_j(|\mathfrak{v}_\ell|-I),$$

where C does not depend on ℓ. The last sum is the trace norm for $|\mathfrak{v}_\ell|-I$ and thus by the equivalence of norms on $\operatorname{Mat}(n,\mathbb{R})$ we find C_2 so that

$$\left|\frac{1}{N}\sum_{\ell=0}^{N}\operatorname{tr}(|\mathfrak{v}_\ell|^2-I)\right|\leq C_2\frac{1}{N}\sum_{\ell=0}^{N}\||\mathfrak{v}_\ell|-I\|.$$

Define the set

$$I_N=\left\{\ell:\ \tilde{C}\|\mathfrak{v}_\ell-I\|>\frac{1}{2}\right\}\cap[1,N]$$

and note that (7.2) implies

$$\lim_{N\to\infty}\frac{|I_N|}{N}=0. \tag{7.17}$$

It follows as in [7, Proposition 11.12] that for $\ell\notin I_N$, there exists a constant C_3 so that

$$\||\mathfrak{v}_\ell|-I\|\leq C_3\|\mathfrak{v}_\ell-I\|.$$

For $\ell\in I_N$ we can estimate $\||\mathfrak{v}_\ell|-I\|$ uniformly and using (7.17) and (7.3) we obtain (7.15). The proof for $\mathfrak{v}_\ell^*$ works the same by using [24, Lemma 4.6.5.] instead of [7, Proposition 11.12]. This finishes the proof. □

We are now ready to adapt Yuditskii's construction [32] to our setting. Let μ be a regular measure with $\operatorname{ess\,supp}\mu=\mathsf{E}$ and let J_+ be the associated Jacobi matrix. As already described after Lemma 7.1, we find J and $\tilde{J}_+$ such that all $\mathbf{c}_k\in\mathbf{C}_E$ belong to the resolvent set of $\tilde{J}_+$ and J and $\tilde{J}_+$ is also regular. Let $\tilde{A}_+$ and A denote the GMP matrix associated to $\tilde{J}_+$ and J respectively and $\{\mathfrak{v}_\ell,\mathfrak{w}_\ell\}$

denote the block Jacobi coefficients of $\Delta_{\mathsf{E}}(A)$. Let us further truncate A after N positive blocks before ∞ (i.e. before the position $-1+N(g+1)$) and extend it by some element $\mathring{A} \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}})$ so that $\mathbf{c}_k \notin \sigma(A_N)$. To be precise, we first truncate A and consider its resolvent function $a_0^2 r_-$, then we can extend it as in Lemma 7.1 by some reflectionless r_+ so that all $\mathbf{c}_k \in \mathbf{C}_E$ belong to the resolvent set of the associated Jacobi matrix and then we consider the associated GMP matrix by [32, Proposition 5.5]. Since elements from the isospectral torus satisfy the magic formula and computing resolvents is a purely local process, we would like to conclude from the compactness of $\mathcal{T}_{\mathsf{E}}(\mathbf{C}_{\mathsf{E}})$ that

$$H_+(A_N) = \sum_{\ell=1}^{N} h(\mathfrak{v}_{\ell-1}, \mathfrak{w}_\ell, \mathfrak{v}_\ell) + O(1),$$

where A_N denotes the truncation described above. However, due to the log-term in the definition of $h(\mathfrak{v}_{\ell-1}, \mathfrak{w}_\ell, \mathfrak{v}_\ell)$ one must be careful. At the place where we modify A by extending it by $\mathring{A}$, by formula (6.15), when computing Λ_n, in a certain range of n given precisely below, one mixes coefficients from A and $\mathring{A}$. Thus we need to argue that

$$-\log \Lambda_n$$

does not grow too fast so that we can still conclude that

$$\lim_{N\to\infty} \frac{1}{N} H_+(A_N) = 0. \tag{7.18}$$

However, looking at the formula (6.15) and the definition of the Blaschke-Potapov factors, if all the coefficients can be bounded uniformly, we see that if $p_g^{(j)} > \delta$ we find a constant C only depending on the bounds of the coefficients and of δ so that

$$\Lambda_n(A_N) \geq C. \tag{7.19}$$

Note now that

$$\Lambda_{-1+(N-1)(g+1)}(A_N) = p_g^{(N-1)},$$

which is still a coefficient of A. But

$$\Lambda_{-1+N(g+1)}(A_N) =: \mathring{p}_g$$

is already a coefficient from $\mathring{A}$. The mixing of coefficients of A and $\mathring{A}$ in computing $\Lambda_n(A_N)$ happens for $-1+(N-1)(g+1) < n < -1+N(g+1)$. But in this case the only value that can make $\Lambda_n(A_N)$ small is $\mathring{p}_g$, and for elements of the isospectral torus we know that

$$\mathring{p}_g = \frac{1}{\lambda_{g+1}}$$

and thus we can conclude (7.19) and therefore (7.18).

Together with $\tilde{H}_+(A_N) \leq H_+(A_N)$, we conclude that

$$\lim_{N\to\infty} \frac{1}{N} \tilde{H}_+(A_N) = 0. \tag{7.20}$$

Realizing that all the arguments in [32, Theorem 1.20] are local, using $2N$ blocks of A, we can obtain a local version of this theorem.

Proposition 7.4. *Let J be constructed as above and A be the associated GMP matrix. Then, there exists an N independent constant C and a sparse set I_N such that*

$$\begin{aligned} \sum_{m=1}^{N} \|\vec{\mathbf{p}}_0(m) - \vec{\mathbf{p}}_{-1}(m)\|^2 &\leq C(\tilde{H}_+(A_{2N}) + |I_N|), \\ \sum_{m=1}^{N} \|\mathbf{F}_{\mathsf{E}}(\vec{\mathbf{p}}_0(m))\|^2 &\leq C(\tilde{H}_+(A_{2N}) + |I_N|). \end{aligned} \tag{7.21}$$

We will need a more quantitative version of [32, Lemma 6.6]:

Lemma 7.5. *Let ψ_n, $\tilde{\psi}_n$, τ_n and $\tilde{\tau}_n$ be given sequences and assume that there exists $\eta > 0$ such that*

$$\cos\psi_n \geq \eta,\ \cos\tilde{\psi}_n \geq \eta,\ 0 \leq \tau_n \leq \frac{1}{\eta},\ 0 \leq \tilde{\tau}_n \leq \frac{1}{\eta}. \tag{7.22}$$

Define

$$\alpha_n := \begin{bmatrix} \tau_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sin\psi_n & \cos\psi_n \\ \cos\psi_n & -\sin\psi_n \end{bmatrix} - \begin{bmatrix} \sin\tilde{\psi}_n & \cos\tilde{\psi}_n \\ \cos\tilde{\psi}_n & -\sin\tilde{\psi}_n \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{\tau}_n \end{bmatrix}. \tag{7.23}$$

Then, there exists C depending only on η so that

$$\begin{aligned} \|\{\cos\psi_n - \cos\tilde{\psi}_n\}\|_{\ell^2(\mathbb{N},\mathbb{C})} &\leq C\|\{\alpha_n\}\|_{\ell^2(\mathbb{N},\mathbb{C})^{2\times 2}}, \\ \|\{\sin\psi_n - \sin\tilde{\psi}_n\}\|_{\ell^2(\mathbb{N},\mathbb{C})} &\leq C\|\{\alpha_n\}\|_{\ell^2(\mathbb{N},\mathbb{C})^{2\times 2}}. \end{aligned}$$

Proof. If $\|\{\alpha_n\}\|_{\ell^2(\mathbb{N},\mathbb{C})^{2\times 2}} = \infty$ the claim is trivial. If it is finite, set $S := \|\{\alpha_n\}\|_{\ell^2(\mathbb{N},\mathbb{C})^{2\times 2}}$. The constant $C > 0$ may increase throughout the proof. Directly from (7.23) we have

$$\|\{\cos\psi_n - \cos\tilde{\psi}_n\}\|_{\ell^2} \leq S \quad \text{and} \quad \|\{\tau_n\cos\psi_n - \tilde{\tau}_n\cos\tilde{\psi}_n\}\|_{\ell^2} \leq S.$$

Since

$$\tau_n\cos\psi_n - \tilde{\tau}_n\cos\tilde{\psi}_n - \tilde{\tau}_n(\cos\psi_n - \cos\tilde{\psi}_n) = (\tau_n - \tilde{\tau}_n)\cos\psi_n,$$

using $\tilde{\tau}_n \leq \frac{1}{\eta}$ and $\cos\psi_n \geq \eta$ we find $C > 0$ so that

$$\|\{\tau_n - \tilde{\tau}_n\}\|_{\ell^2} \leq CS.$$

Now, we have another two conditions

$$\|\{\tau_n\sin\psi_n - \sin\tilde{\psi}_n\}\|_{\ell^2} \leq S \quad \text{and} \quad \|\{\sin\psi_n - \tilde{\tau}_n\sin\tilde{\psi}_n\}\|_{\ell^2} \leq S.$$

Using

$$\sin\psi_n - \tilde{\tau}_n\sin\tilde{\psi}_n = \sin\psi_n - \tau_n\tilde{\tau}_n\sin\psi_n - \tilde{\tau}_n(\sin\tilde{\psi}_n - \tau_n\sin\psi_n)$$

and $\|\{\tau_n\sin\psi_n - \sin\tilde{\psi}_n\}\| \leq S$ and $\tilde{\tau}_n \leq \frac{1}{\eta}$ we conclude that

$$\|\{\sin\psi_n(1 - \tau_n\tilde{\tau}_n)\}\|_{\ell^2} \leq CS.$$

Now we have

$$1 - \tau_n^2 = 1 - \tau_n\tilde{\tau}_n + \tau_n(\tilde{\tau}_n - \tau_n)$$

and since $|\sin\psi_n| \leq 1$ and $|\tau_n| \leq \frac{1}{\eta}$, we conclude

$$\|\{(\tau_n^2 - 1)\sin\psi_n\}\| \leq CS.$$

Again by, $|\tau_n| \leq \frac{1}{\eta}$ we also get a bound for $\{(\tau_n - 1)\sin\psi_n\}$. Finally, since

$$\sin\psi_n - \sin\tilde{\psi}_n = \tau_n \sin\psi_n - \sin\tilde{\psi}_n - (\tau_n - 1)\sin\psi_n,$$

we obtain the also the estimate for $\{\sin\psi_n - \sin\tilde{\psi}_n\}$. □

Proof of Proposition 7.4. In the proof we will find constants $C > 0$ and sparse sets I_N. These quantities will change throughout the proof. Note that the union of sparse sets is clearly sparse. First we mention an important locality property of the Jacobi flow. In the following we will derive estimates for entries of $A_{2N}(m)$ in the blocks 0 and -1. Due to the locality property of the Jacobi flow, for $0 < m \leq 2N-1$, the coefficients of $A_{2N}(m)$ and $A(m)$ coincide; this is nicely visualized in the diagram [32, eq. (4.12)]. Similarly, we have already mentioned that computing entries of the resolvents, due to the band structure, can also be done locally. Thus, our estimates will be derived for the coefficients of $A_{2N}(m)$, but by restricting it to $0 < m \leq N$ they agree with the coefficients associated to A. For this reason we will also notationally not distinguish between the coefficients of A and the ones of A_{2N}.

By the explanation following Lemma 7.2 and (7.20), we conclude that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} \left(G(x_0(m)) + G(x_{2g+2}(m)) + \sum_{j=1}^{2g+1} x_j(m)^2 \right) = 0.$$

Notice that G obeys

$$c_\epsilon^{-1}(x-1)^2 \leq G(x) \leq c_\epsilon (x-1)^2, \quad \forall x \in (\epsilon, \epsilon^{-1}).$$

Thus, we find a sparse set I_N and $C > 0$ so that

$$\begin{aligned} &\sum_{m=1}^{N} \left((x_0(m)-1)^2 + (x_{2g+2}(m)-1)^2 + \sum_{j=1}^{2g+1} x_j(m)^2 \right) \\ &\quad \leq C \bigg(\sum_{m=1}^{N} \Big(G(x_0(m)) + G(x_{2g+2}(m)) + \sum_{j=1}^{2g+1} x_j(m)^2) \Big) + |I_N|) \bigg). \end{aligned} \tag{7.24}$$

Thus, for $1 \leq j \leq 2g+1$,

$$\|\{x_j(m)\}_{m=1}^N\|_2 \leq C \left(\tilde{H}_+(A_N) + |I_N| \right)$$

and

$$\|\{x_0(m) - 1\}_{m=1}^N\|_2 \leq C \left(\tilde{H}_+(A_N) + |I_N| \right),$$

$$\|\{x_{2g+2}(m) - 1\}_{m=1}^N\|_2 \leq C \left(\tilde{H}_+(A_N) + |I_N| \right).$$

We note that

$$x_{2g+2}(m) = \lambda_0 \Lambda_{-1}(m) = \lambda_0 p_g^{(0)}(m),$$

and thus

$$\frac{1}{N} \sum_{m=1}^{N} (\lambda_0 p_g^{(0)}(m) - 1)^2 = 0. \tag{7.25}$$

This is one component of F_{E}.

Let us now show the first inequality in (7.21). Denote $\hat{A} = \mathcal{O}A$, where $\mathcal{O}A$ is the transform defined in (6.18). We use the hat for all entries related to $\hat{A}$ and

$\Delta_{\mathsf{E}}(\hat{A})$, respectively. The entries of $A(m)$ are denoted by $\{p_k^{(j)}(m), q_k^{(j)}(m)\}$. Recall that m corresponds to application of the Jacobi flow, j denotes the block and k the component of the vector $\vec{p}_j(m)$. We use similar notation for $\hat{A}$, $\Delta_{\mathsf{E}}(A)$ and $\Delta_{\mathsf{E}}(\hat{A})$. Due to the definition (6.18), we find

(7.26)
$$\begin{bmatrix} v_{g-1,g-1}^{(0)}(m) & 0 \\ v_{g,g-1}^{(0)}(m) & \lambda_0 p_g^{(0)}(m) \end{bmatrix} \mathbf{o}(\phi_g^{(0)}(m)) = \mathbf{o}(\phi_g^{(-1)}(m)) \begin{bmatrix} \lambda_0 \hat{p}_g^{(0)}(m) & 0 \\ \hat{w}_{0,g}^{(0)}(m) & \hat{v}_{0,0}^{(1)}(m) \end{bmatrix}.$$

Note that $v_{g,g-1}^{(0)}(m) = x_{2g+1}(m)$. It was mentioned after Lemma 7.2 that $H_+(\hat{A})$ can be expressed in terms of $H_+(A)$. Therefore, we conclude by (6.20) that (7.25) also holds for $\hat{p}_g^{(0)}(m)$. Note that $x_{g+2}(m) = \hat{w}_{0,g}^{(-1)}(m)$. Since shifting by a full block in the very beginning only adds a fixed constant, and $\mathcal{J}$ commutes with this shift by (6.21), we can apply Lemma 7.5 to (7.26) and obtain by (7.24) that

$$\|\{\sin \phi_g^{(-1)}(m) - \sin \phi_g^{(0)}(m)\}_{m=1}^N\|_{\ell^2} \leq C(\tilde{H}_+(A_N) + |I_N|).$$

Thus, by (6.17)

$$\|\{p_{g-1}^{(-1)} - p_{g-1}^{(0)}\}_{m=1}^N\|_{\ell^2} \leq C(\tilde{H}_+(A_N) + |I_N|). \tag{7.27}$$

Since by [32, eq (4.2)] one can pass from j to $j-1$ by using $\hat{A}$, we obtain (7.27) for $0 \leq j \leq g$. Similarly, by [32, eq (4.2)], one obtains the estimates for the q_j-coefficients. This finishes the proof of the first inequality in (7.21).

It remains to prove (7.21) for the other components of F_{E}. The proof of Lemma 7.5 yields an estimate for $\|\{(v_{g-1,g-1}^{(-1)}(n) - 1)\sin \phi_g^{(-1)}(n)\}\|_2$ or, equivalently, it shows

$$\|\{(\Lambda_{-2}(m)\lambda_g - 1)p_{g-1}^{(-1)}(m)\}\| \leq C(\tilde{H}_+(A_N) + |I_N|).$$

Since $p_{g-1}^{(-1)}(m)$ may approach zero, it does not imply yet give an estimate for $\{(\Lambda_{-2}(m)\lambda_g - 1)\}$. If we can also estimate

$$\|\{(\Lambda_{-2}(m)\lambda_g - 1)q_{g-1}^{(-1)}(m)\}_{m=1}^N\|,$$

then $\inf_m \left((q_{g-1}^{(-1)}(m))^2 + (p_{g-1}^{(-1)}(m))^2\right) > 0$ yields

$$\|\{(\Lambda_{-2}(m)\lambda_g - 1)\}_{m=1}^N\| \leq C(\tilde{H}_+(A_N) + |I_N|).$$

To this end, we note that

$$\Lambda_{-2}(m+1) = \frac{\cos \phi_g^{(-1)}(m)}{\cos \phi_g^{(-2)}(m)} \Lambda_{-2}(m). \tag{7.28}$$

Indeed, by definition of the Jacobi flow

$$U(\vec{p}_{-2}(m)) \begin{bmatrix} v_{g,g}^{(-2)} & & & \\ * & v_{0,0}^{(-1)} & & \\ * & * & \ddots & \\ * & * & * & v_{g-1,g-1}^{(-1)} \end{bmatrix} (m+1) = \mathfrak{v}_{-1}(m) U(\vec{p}_{-1}(m)),$$

the second from below entry in the last column in this matrix identity means exactly (7.28). Since by the above, we can estimate $\|\{\cos \phi_g^{(-1)}(m) - \cos \phi_g^{(-2)}(m)\}_{m=1}^N\|_{\ell^2}$ we obtain

$$\|\{\Lambda_{-2}(m+1) - \Lambda_{-2}(m)\}_{m=1}^N\|_{\ell^2} \leq C(\tilde{H}_+(A_N) + |I_N|).$$

Now by [32, (4.10)] we have

$$p_{g-1}^{(-1)}(m) = -q_{g-1}^{(-1)}(m+1)f(m),$$

where $f(m)$ is an explicit function that can be small only on a sparse set. Combining this with

$$(\Lambda_{-2}(m)\lambda_g - 1)p_{g-1}^{(-1)}(m) = -(\Lambda_{-2}(m)\lambda_g - 1)q_{g-1}^{(-1)}(m+1)f(m),$$

we also get an estimate for $\|\{(\Lambda_{-2}(m)\lambda_g - 1)q_{g-1}^{(-1)}(m)\}_{m=1}^N\|_{\ell^2}$, which shows

$$\|\{(\Lambda_{-2}(m)\lambda_g - 1)\}\| \leq C(\tilde{H}_+(A_N) + |I_N|).$$

The same arguments with respect to $\mathcal{O}^k A$, $k = 1, \dots, g-1$, in a combination with (6.20), yield the estimates for all other components of F_{E}. □

Lemma 7.6. *There exist* $\{\epsilon_m^\alpha\} \in \mathrm{CS}(\mathbb{N}, \mathbb{R}^g/\mathbb{Z}^g)$ *and* $\{\epsilon_m^a\} \in \mathrm{CS}(\mathbb{N}, \mathbb{R}), \{\epsilon_n^b\} \in \mathrm{CS}(\mathbb{N}, \mathbb{R})$ *so that*

$$a_m^2 = \mathcal{A}\left(\sum_{j=1}^m \varepsilon_j^\alpha - m\chi\right) + \varepsilon_m^a,$$

$$b_m = \mathcal{B}\left(\sum_{j=1}^m \varepsilon_j^\alpha - m\chi\right) + \varepsilon_m^b,$$

where $\mathcal{A}, \mathcal{B}$ *are given in* (7.11).

Proof. Let $A(\vec{\mathbf{p}}_0(m))$ be the periodic GMP matrix with coefficients $\vec{\mathbf{p}}_0(m)$ and $A(\alpha_m) \in \mathcal{T}_{\mathsf{E}}(\mathbf{C}_E)$, so that

$$\mathrm{dist}(\vec{\mathbf{p}}_0(m), \mathcal{IS}_{\mathsf{E}}) = \mathrm{dist}(\vec{\mathbf{p}}_0(m), \mathring{\vec{\mathbf{p}}}(\alpha_m)).$$

Thus, using (7.12) we obtain

$$\sum_{m=1}^N \mathrm{dist}(\vec{\mathbf{p}}_0(m), \mathring{\vec{\mathbf{p}}}(\alpha_m))^2 \leq C(\tilde{H}_+(A_N) + |I_N|)$$

and by (6.6) we get

$$\sum_{m=1}^N (a_m^2 - \mathcal{A}(\alpha_m))^2 \leq C(\tilde{H}_+(A_N) + |I_N|),$$

$$\sum_{m=1}^N (b_m - \mathcal{B}(\alpha_m))^2 \leq C(\tilde{H}_+(A_N) + |I_N|),$$

where again $\{a_m, b_m\}_{m\in\mathbb{N}_0}$ denote the coefficients of J_+. Thus, dividing by N and sending $N \to \infty$, we obtain by (7.17) and (7.20) that

$$\{a_m^2 - \mathcal{A}(\alpha_m)\}_{m\in\mathbb{N}_0}, \{b_m - \mathcal{B}(\alpha_m)\}_{m\in\mathbb{N}_0} \in \mathrm{CS}\,. \tag{7.29}$$

The smoothness of the Jacobi flow transform, provided that $p_g^{(0)}, p_g^{(1)} > \delta$, allows for the definition of a sparse set I_N so that

$$\begin{aligned}\mathrm{dist}(\vec{\mathbf{p}}_0(m+1), \mathring{\vec{\mathbf{p}}}(\alpha_m - \chi)) &= \mathrm{dist}(\mathcal{J}(\vec{\mathbf{p}}_0(m), \vec{\mathbf{p}}_1(m)), \mathcal{J}(\vec{\mathbf{p}}(\alpha_m)) \\ &\leq C(\mathsf{E}, J, \delta)\{\mathrm{dist}(\vec{\mathbf{p}}_0(m), \mathring{\vec{\mathbf{p}}}(\alpha_m)) \\ &\qquad + \mathrm{dist}(\vec{\mathbf{p}}_0(m), \vec{\mathbf{p}}_1(m)) + |I_N|\}.\end{aligned}$$

Thus,

$$\begin{aligned}\operatorname{dist}(\mathring{\vec{\mathbf{p}}}(\alpha_{m+1}), \mathring{\vec{\mathbf{p}}}(\alpha_m - \chi)) \le C(\mathsf{E}, J, \delta)(&\operatorname{dist}(\vec{\mathbf{p}}_0(m), \mathring{\vec{\mathbf{p}}}(\alpha_m)) \\ &+ \operatorname{dist}(\vec{\mathbf{p}}_0(m), \vec{\mathbf{p}}_1(m)) + |I_N|) \\ &+ \operatorname{dist}(\vec{\mathbf{p}}_0(m+1), \mathring{\vec{\mathbf{p}}}(\alpha_{m+1})).\end{aligned}$$

Moreover, we have

$$\|\alpha - \beta\| \le C_1(E) \operatorname{dist}(\mathring{\vec{\mathbf{p}}}(\alpha), \mathring{\vec{\mathbf{p}}}(\beta)).$$

Thus, defining $\epsilon_\alpha(m) = \alpha_{m+1} - (\alpha_m - \chi)$, we conclude from (7.21) that

$$\{\epsilon_\alpha\} \in \mathrm{CS}(\mathbb{N}, \mathbb{R}^g/\mathbb{Z}^g).$$

□

Lemma 7.7. *For fixed $L \in \mathbb{N}$ and $\delta > 0$, the set*

$$B_{L,\delta} = \left\{ m : \left\| \sum_{j=m+1}^{m+\ell} \epsilon_j^\alpha \right\| \le \delta \text{ for all } \ell = 0, \dots, L-1 \right\}$$

has a sparse complement, i.e., $\frac{|B_{L,\delta} \cap \{1,\dots,N\}|}{N} \to 1$ as $N \to \infty$.

Proof. Since shifts and linear combinations of CS sequences are in CS, $\{ \sum_{j=m+1}^{m+\ell} \epsilon_j^\alpha \}_{m=0}^\infty \in \mathrm{CS}$ for any ℓ. Thus, for any ℓ, the set $\{m : \left\| \sum_{j=m+1}^{m+\ell} \epsilon_j^\alpha \right\| > \delta\}$ is sparse; the complement of $B_{L,\delta}$ is a union of finitely many sparse sets, so it is sparse. □

Proof of Theorem 1.10. It remains to prove that, for every $\epsilon > 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} \operatorname{dist}(\mathcal{T}_{\mathsf{E}}^+, (S_+^*)^m J_+ S_+^m) \le \epsilon. \tag{7.30}$$

Fix L so that $\sum_{\ell=L}^\infty e^{-\ell} \|J_+\| \le \epsilon/16$. Choose $\delta > 0$ so that

$$|\mathcal{A}(\beta_1) - \mathcal{A}(\beta_2)| \le \frac{\epsilon}{8L}, \qquad |\mathcal{B}(\beta_1) - \mathcal{B}(\beta_2)| \le \frac{\epsilon}{8L} \tag{7.31}$$

whenever $|\beta_1 - \beta_2| \le \delta$.

Since $\operatorname{dist}(\mathcal{T}_{\mathsf{E}}^+, (S_+^*)^m J_+ S_+^m)$ is uniformly bounded in m and the complement of $B_{L,\delta}$ is sparse,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{\substack{1 \le m \le N \\ m \notin B_{L,\delta}}} \operatorname{dist}(\mathcal{T}_{\mathsf{E}}^+, (S_+^*)^m J_+ S_+^m) = 0.$$

Set $\alpha_m = \sum_{j=1}^m \varepsilon_j^\alpha$. For $m \in B_{L,\delta}$, estimating the distance to $\mathcal{T}_{\mathsf{E}}^+$ by the distance to $J(\alpha_m - m\chi)$ gives

$$\begin{aligned}&\operatorname{dist}(\mathcal{T}_{\mathsf{E}}^+, (S_+^*)^m J_+ S_+^m) \\ &\qquad \le \sum_{\ell=0}^\infty e^{-\ell} (|a_{m+\ell} - \mathcal{A}(\alpha_m - (m+\ell)\chi)| + |b_{m+\ell} - \mathcal{B}(\alpha_m - (m+\ell)\chi)|).\end{aligned}$$

Using (7.31) for $\ell < L$ and using our choice of L to bound the tail of the series, we obtain

$$\begin{aligned}&\operatorname{dist}(\mathcal{T}_{\mathsf{E}}^{+}, (S_+^*)^m J_+ S_+^m)\\ &\le \frac{\epsilon}{2} + \sum_{\ell=0}^{L-1} e^{-\ell}(|a_{m+\ell} - \mathcal{A}(\alpha_{m+\ell} - (m+\ell)\chi)| + |b_{m+\ell} - \mathcal{B}(\alpha_{m+\ell} - (m+\ell)\chi)|).\end{aligned}$$

Thus, to prove (7.30), it remains to prove

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{\substack{1\le m\le N\\ m\in B_{L,\delta}}} \sum_{\ell=0}^{L-1} e^{-\ell} g_{m+\ell} \le \frac{\epsilon}{2}, \tag{7.32}$$

where $g_p = |a_p - \mathcal{A}(\alpha_p - p\chi)| + |b_p - \mathcal{B}(\alpha_p - p\chi)|$. Note $g \in \mathrm{CS}$ by (7.29). Enlarging the range of summation, we obtain

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{\substack{1\le m\le N\\ m\in B_{L,\delta}}} \sum_{\ell=0}^{L-1} e^{-\ell} g_{m+\ell} \le \limsup_{N\to\infty} \frac{1}{N} \sum_{p=1}^{N+L} \sum_{\ell=0}^{L-1} e^{-\ell} g_p.$$

Now the sum in ℓ can be separated as an explicit constant, so this lim sup is zero since $g \in \mathrm{CS}$. Then (7.32) follows, and the proof of (7.30) is complete. □

References

[1] Damir Z. Arov and Harry Dym, *J-contractive matrix valued functions and related topics*, Encyclopedia of Mathematics and its Applications, vol. 116, Cambridge University Press, Cambridge, 2008, DOI 10.1017/CBO9780511721427. MR2474532

[2] David H. Armitage and Stephen J. Gardiner, *Classical potential theory*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2001, DOI 10.1007/978-1-4471-0233-5. MR1801253

[3] Adhemar Bultheel, Pablo González-Vera, Erik Hendriksen, and Olav Njåstad, *Orthogonal rational functions*, Cambridge Monographs on Applied and Computational Mathematics, vol. 5, Cambridge University Press, Cambridge, 1999, DOI 10.1017/CBO9780511530050. MR1676258

[4] M. J. Cantero, L. Moral, and L. Velázquez, *Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle*, Linear Algebra Appl. **362** (2003), 29–56, DOI 10.1016/S0024-3795(02)00457-3. MR1955452

[5] Jacob S. Christiansen, Barry Simon, and Maxim Zinchenko, *Finite gap Jacobi matrices: a review*, Spectral analysis, differential equations and mathematical physics: a festschrift in honor of Fritz Gesztesy's 60th birthday, Proc. Sympos. Pure Math., vol. 87, Amer. Math. Soc., Providence, RI, 2013, pp. 87–103, DOI 10.1090/pspum/087/01429. MR3087900

[6] Jacob S. Christiansen, Benjamin Eichinger, and Tom VandenBoom, *Finite-gap CMV matrices: periodic coordinates and a magic formula*, Int. Math. Res. Not. IMRN **18** (2021), 14016–14085, DOI 10.1093/imrn/rnz213. MR4320803

[7] David Damanik, Rowan Killip, and Barry Simon, *Perturbations of orthogonal polynomials with periodic recursion coefficients*, Ann. of Math. (2) **171** (2010), no. 3, 1931–2010, DOI 10.4007/annals.2010.171.1931. MR2680401

[8] David Damanik, Alexander Pushnitski, and Barry Simon, *The analytic theory of matrix orthogonal polynomials*, Surv. Approx. Theory **4** (2008), 1–85. MR2379691

[9] David Damanik and Peter Yuditskii, *Counterexamples to the Kotani-Last conjecture for continuum Schrödinger operators via character-automorphic Hardy spaces*, Adv. Math. **293** (2016), 738–781, DOI 10.1016/j.aim.2016.02.023. MR3474334

[10] Karl Deckers and Doron S. Lubinsky, *Christoffel functions and universality limits for orthogonal rational functions*, Anal. Appl. (Singap.) **10** (2012), no. 3, 271–294, DOI 10.1142/S0219530512500133. MR2948895

[11] Joseph L. Doob, *Classical potential theory and its probabilistic counterpart*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1984 edition, DOI 10.1007/978-3-642-56573-1. MR1814344

[12] Benjamin Eichinger, *Periodic GMP matrices*, SIGMA Symmetry Integrability Geom. Methods Appl. **12** (2016), Paper No. 066, 19, DOI 10.3842/SIGMA.2016.066. MR3519566

[13] B. Eichinger and M. Lukić. 2020. *Stahl–Totik regularity for continuum Schrödinger operators*, `arXiv:2001.00875`.

[14] Fritz Gesztesy, Helge Holden, Johanna Michor, and Gerald Teschl, *Soliton equations and their algebro-geometric solutions. Vol. II*, Cambridge Studies in Advanced Mathematics, vol. 114, Cambridge University Press, Cambridge, 2008. (1 + 1)-dimensional discrete models, DOI 10.1017/CBO9780511543203. MR2446594

[15] Leonid Golinskii and Sergei Khrushchev, *Cesàro asymptotics for orthogonal polynomials on the unit circle and classes of measures*, J. Approx. Theory **115** (2002), no. 2, 187–237, DOI 10.1006/jath.2001.3655. MR1901215

[16] Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original. MR1288752

[17] Helge Krüger, *Probabilistic averages of Jacobi operators*, Comm. Math. Phys. **295** (2010), no. 3, 853–875, DOI 10.1007/s00220-010-1014-y. MR2600036

[18] Paul G. Nevai, *Orthogonal polynomials*, Mem. Amer. Math. Soc. **18** (1979), no. 213, v+185, DOI 10.1090/memo/0213. MR519926

[19] V. P. Potapov, *The multiplicative structure of J-contractive matrix functions*, Amer. Math. Soc. Transl. (2) **15** (1960), 131–243, DOI 10.1090/trans2/015/07. MR0114915

[20] Thomas Ransford, *Potential theory in the complex plane*, London Mathematical Society Student Texts, vol. 28, Cambridge University Press, Cambridge, 1995, DOI 10.1017/CBO9780511623776. MR1334766

[21] Christian Remling, *The absolutely continuous spectrum of Jacobi matrices*, Ann. of Math. (2) **174** (2011), no. 1, 125–171, DOI 10.4007/annals.2011.174.1.4. MR2811596

[22] Barry Simon, *Equilibrium measures and capacities in spectral theory*, Inverse Probl. Imaging **1** (2007), no. 4, 713–772, DOI 10.3934/ipi.2007.1.713. MR2350223

[23] Barry Simon, *Regularity and the Cesàro-Nevai class*, J. Approx. Theory **156** (2009), no. 2, 142–153, DOI 10.1016/j.jat.2008.04.016. MR2494549

[24] Barry Simon, *Szegő's theorem and its descendants*, M. B. Porter Lectures, Princeton University Press, Princeton, NJ, 2011. Spectral theory for L^2 perturbations of orthogonal polynomials. MR2743058

[25] Barry Simon, *Basic complex analysis*, A Comprehensive Course in Analysis, Part 2A, American Mathematical Society, Providence, RI, 2015, DOI 10.1090/simon/002.1. MR3443339

[26] Mikhail Sodin and Peter Yuditskii, *Almost periodic Jacobi matrices with homogeneous spectrum, infinite-dimensional Jacobi inversion, and Hardy spaces of character-automorphic functions*, J. Geom. Anal. **7** (1997), no. 3, 387–435, DOI 10.1007/BF02921627. MR1674798

[27] Herbert Stahl and Vilmos Totik, *General orthogonal polynomials*, Encyclopedia of Mathematics and its Applications, vol. 43, Cambridge University Press, Cambridge, 1992, DOI 10.1017/CBO9780511759420. MR1163828

[28] Gerald Teschl, *Jacobi operators and completely integrable nonlinear lattices*, Mathematical Surveys and Monographs, vol. 72, American Mathematical Society, Providence, RI, 2000, DOI 10.1090/surv/072. MR1711536

[29] Gerald Teschl, *Mathematical methods in quantum mechanics*, 2nd ed., Graduate Studies in Mathematics, vol. 157, American Mathematical Society, Providence, RI, 2014. With applications to Schrödinger operators, DOI 10.1090/gsm/157. MR3243083

[30] J. L. Ullman, *On the regular behaviour of orthogonal polynomials*, Proc. London Math. Soc. (3) **24** (1972), 119–148, DOI 10.1112/plms/s3-24.1.119. MR291718

[31] A. Volberg and P. Yuditskii, *Kotani-Last problem and Hardy spaces on surfaces of Widom type*, Invent. Math. **197** (2014), no. 3, 683–740, DOI 10.1007/s00222-013-0495-7. MR3251833

[32] P. Yuditskii, *Killip-Simon problem and Jacobi flow on GMP matrices*, Adv. Math. **323** (2018), 811–865, DOI 10.1016/j.aim.2017.11.005. MR3725892

Institute of Analysis, Johannes Kepler University of Linz, 4040 Linz, Austria
Email address: benjamin.eichinger@jku.at

Department of Mathematics, Rice University MS-136, Box 1892, Houston, Texas 77251-1892
Email address: milivoje.lukic@rice.edu

Department of Mathematics, Rice University MS-136, Box 1892, Houston, Texas 77251-1892
Email address: gfy1@rice.edu

Constructive Approximation
https://doi.org/10.1007/s00365-023-09670-0

CONSTRUCTIVE
APPROXIMATION

Asymptotics of Chebyshev Rational Functions with Respect to Subsets of the Real Line

Benjamin Eichinger[1] · **Milivoje Lukić**[2] · **Giorgio Young**[2]

Received: 27 January 2021 / Revised: 18 May 2022 / Accepted: 29 January 2023

Abstract

There is a vast theory of Chebyshev and residual polynomials and their asymptotic behavior. The former ones maximize the leading coefficient and the latter ones maximize the point evaluation with respect to an L^∞ norm. We study Chebyshev and residual extremal problems for rational functions with real poles with respect to subsets of $\overline{\mathbb{R}}$. We prove root asymptotics under fairly general assumptions on the sequence of poles. Moreover, we prove Szegő–Widom asymptotics for sets which are regular for the Dirichlet problem and obey the Parreau–Widom and DCT conditions.

Keywords Chebyshev and residual extremal problems · Rational functions · Szegő–Widom asymptotics

Mathematics Subject Classification 41A50 · 30E15 · 30C10

Communicated by Arno Kuijlaars.

Benjamin Eichinger was supported by Austrian Science Fund FWF, projects: J 4138-N32 and P 33885.
Milivoje Lukić was supported in part by NSF grant. DMS–1700179.
Giorgio Young was supported in part by NSF grant. DMS–1745670.

Giorgio Young
gfyoung@umich.edu

Benjamin Eichinger
benjamin.eichinger@tuwien.ac.at

Milivoje Lukić
milivoje.lukic@rice.edu

1 Institute for Analysis and Scientific Computing, TU Wien, 1040 Wien, Austria

2 Department of Mathematics, University of Michigan, Ann Arbor, MI 48104, USA

Published online: 09 March 2024

1 Introduction

Chebyshev polynomials are extremal polynomials with respect to the supremum norm on a compact set E. First discovered with explicit formulas for the set $\mathsf{E} = [-1, 1]$, see [3, 4], a general theory has developed for more general sets E, with important classical and modern developments [5, 7, 11, 27, 31]. Aspects of this theory have been extended to the setting of residual polynomials [7] (which are extremizers with respect to a point evaluation rather than leading coefficient) and to the setting of Chebyshev rational functions with poles in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ [17].

To state the problems precisely, we make the following definitions. For $\mathbf{c} \in \overline{\mathbb{R}}$ we denote

$$r(z, \mathbf{c}) = \begin{cases} \frac{1}{\mathbf{c}-z}, & \mathbf{c} \neq \infty, \\ z, & \mathbf{c} = \infty. \end{cases}$$

We fix a compact proper subset $\mathsf{E} \subset \overline{\mathbb{R}}$ containing infinitely many points. Connected components of $\overline{\mathbb{R}} \setminus \mathsf{E}$ are called gaps of E. We fix a sequence of poles $\mathbf{C} = (\mathbf{c}_k)_{k=1}^{\infty}$ with $\mathbf{c}_k \in \overline{\mathbb{R}} \setminus \mathsf{E}$. The sequence $\mathbf{C}$ can have repetitions, which are used to designate multiplicity: we consider the spaces of rational functions $\mathcal{L}_n$ defined as

$$\mathcal{L}_n = \left\{ \frac{P(z)}{R_n(z)} \ : \ P \in \mathcal{P}_n \right\}, \tag{1.1}$$

where $\mathcal{P}_n$ denotes the set of polynomials of degree at most n and

$$R_n(z) = \prod_{\substack{1 \leq k \leq n \\ \mathbf{c}_k \neq \infty}} (z - \mathbf{c}_k). \tag{1.2}$$

Of course, the spaces $\mathcal{L}_n$ could also be defined iteratively, by

$$\mathcal{L}_n = \operatorname{span} \left\{ r(z, \mathbf{c}_n)^{d_n} \right\} \oplus \mathcal{L}_{n-1}, \quad \mathcal{L}_0 = \{1\},$$

where d_n denotes the multiplicity of the pole $\mathbf{c}_n$ up to that point,

$$d_n = \sum_{\substack{1 \leq k \leq n \\ \mathbf{c}_k = \mathbf{c}_n}} 1.$$

Let $\| \cdot \|_{\mathsf{E}}$ denote the supremum norm on E. We consider the two related extremal problems:

Problem 1.1 (Chebyshev Extremal Problem)

$$m_n(\mathbf{c}_n) := \sup\{\operatorname{Re} \lambda_n : \exists F_n \in \mathcal{L}_n \text{ such that } \|F_n\|_{\mathsf{E}} \leq 1 \text{ and } F_n - \lambda_n r(\cdot, \mathbf{c}_n)^{d_n} \in \mathcal{L}_{n-1}\}. \tag{1.3}$$

Problem 1.2 (Residual Extremal Problem) For $x_* \in \overline{\mathbb{R}} \setminus (\mathsf{E} \cup \{\mathbf{c}_k : 1 \le k \le n\})$,

$$m_n(x_*) := \sup\{\operatorname{Re} F_n(x_*) : F_n \in \mathcal{L}_n, \ \|F_n\|_{\mathsf{E}} \le 1\}. \tag{1.4}$$

If $\mathbf{c}_k = \infty$ for all k, Problem 1.1 is the standard extremal problem for Chebyshev polynomials on E. For this reason we refer to λ_n still as the leading coefficient. Whereas the Chebyshev extremal problem maximizes the leading coefficient at the pole $x_* = \mathbf{c}_n$, the residual extremal problem maximizes the value at a point x_* which is not a pole. We will use the notation x_* for both problems when convenient.

For both problems, an extremal function exists (i.e., the supremum is a maximum) and is unique (see Sect. 2). The goal of this paper is to study the extremal functions F_n and their asymptotics as $n \to \infty$.

Problems 1.1 and 1.2 have a conformal invariance with respect to the group $\mathrm{PSL}(2, \mathbb{R})$ of $\overline{\mathbb{R}}$-preserving, orientation-preserving Möbius transformations. This conformal invariance is obfuscated by the use of polynomials in the definitions (1.1) and (1.2), but can be made explicit in the language of divisors. Divisors on the Riemann sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ are elements of the free Abelian group over $\overline{\mathbb{C}}$. They can be implemented as formal sums or as functions $D : \overline{\mathbb{C}} \to \mathbb{Z}$ which take nonzero values only at finitely many points; we will find the second interpretation notationally convenient. The degree of D is the integer $\deg D = \sum_z D(z)$, and the divisor D is integral if $D(z) \ge 0$ for all z. We also write $D_1 \le D_2$, if $D_2 - D_1$ is integral and denote by $\operatorname{supp} D = \{z \in \overline{\mathbb{C}} : D(z) \ne 0\}$ the support of D. In particular, for a meromorphic nonconstant function $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, we denote its polar divisor by $(f)_\infty$; the polar divisor assigns to each pole the multiplicity of that pole, and takes zero values elsewhere. Similarly, for $w \in \mathbb{C}$, we define $(f)_w = (1/(f - w))_\infty$. The value $\deg(f)_w$ is independent of w and corresponds to the degree of f. We also follow the convention to set $(f)_w = 0$, if f is a constant. For any n, we define the divisor D_n^∞ by

$$D_n^\infty(\mathbf{c}) = \#\{k : \ \mathbf{c}_k = \mathbf{c}, \ 1 \le k \le n\}. \tag{1.5}$$

In other words, in the functional interpretation, $D_n^\infty = \sum_{k=1}^n \chi_{\{\mathbf{c}_k\}}$. Note that by definition $\deg D_n^\infty = n$. Any integral divisor D with degree n generates a $n+1$ dimensional vector space

$$\mathcal{L}(D) = \{f : \overline{\mathbb{C}} \to \overline{\mathbb{C}} \mid f \text{ is meromorphic and } (f)_\infty \le D\}, \tag{1.6}$$

and the definition (1.1) is equivalent to

$$\mathcal{L}_n = \mathcal{L}(D_n^\infty). \tag{1.7}$$

Now Problems 1.1, 1.2 can be unified as follows:

Problem 1.3 For a real integral divisor D_n^∞ with $\deg D_n^\infty = n$ containing only points in $\overline{\mathbb{R}} \setminus \mathsf{E}$, and a point $x_* \in \overline{\mathbb{R}} \setminus \mathsf{E}$, denote $d_n = D_n^\infty(x_*)$ and $\mathcal{L}_n = \mathcal{L}(D_n^\infty)$ and find

$$m_n(x_*) := \sup\{\operatorname{Re} \lim_{x \to x_*} \frac{F_n(x)}{r(x, x_*)^{d_n}} : F_n \in \mathcal{L}_n, \ \|F_n\|_{\mathsf{E}} \le 1\}. \tag{1.8}$$

The Chebyshev problem corresponds to $d_n > 0$ (up to a permutation of $\mathbf{c}_1, \dots, \mathbf{c}_n$) and the residual problem corresponds to $d_n = 0$. Throughout this paper, we work in the general setting of Problem 1.3.

In order to state our results in a conformally invariant form, we use the following language:

Definition 1.4 For a sequence $(t_j)_{j=0}^m$ in $\overline{\mathbb{R}}$ with $m \geq 2$, we say that the sequence is cyclically ordered if it has no repetitions and there exists $f \in \mathrm{PSL}(2, \mathbb{R})$ such that $f(t_0) = \infty$ and $f(t_1) < f(t_2) < \cdots < f(t_m)$. We will also use cyclic interval notation: for distinct $a, b \in \overline{\mathbb{R}}$, we denote

$$(a, b) = \{c \mid (a, c, b) \text{ is cyclically ordered}\}, \qquad [a, b] = \{a, b\} \cup (a, b).$$

This gives a well-defined cyclic order, since $\mathrm{PSL}(2, \mathbb{R})$ transformations preserve orientation on $\overline{\mathbb{R}}$.

Chebyshev polynomials for subsets of $\mathbb{R}$ have many universal properties; the Chebyshev alternation theorem compresses all these properties in a way that uniquely characterizes the extremizer. Namely, a polynomial P_n of degree n so that $\|P_n\|_{\mathsf{E}} \leq 1$ has a maximal set of alternation points if there are $n+1$ points $x_1 < \cdots < x_{n+1}$, $x_i \in \mathsf{E}$, so that

$$P_n(x_j) = (-1)^{n+1-j}. \tag{1.9}$$

Then P_n is the Chebyshev polynomial for the set E, if and only if it has a maximal set of alternation points. One way of viewing the alternation theorem is the following. The Chebyshev polynomial, T_n, for E has n real and simple zeros and between each of them there should be an alternation point, which gives $n-1$ of them and then there should be one at each gap edge of the extremal gap (in this case the one containing ∞) which sums up to $n+1$ points of alternation. In particular x_1 and x_{n+1} will always be counted, because of the natural order of $\mathbb{R}$. Similarly, residual polynomials have an alternation theorem, which relies on a notion of an x_* alternation set [7]. Furthermore, by [7], in the polynomial case, such a set characterizes the residual polynomial: P_n is the residual polynomial for the set E if and only if $\|P_n\|_{\mathsf{E}} \leq 1$ and P_n has an x_* alternation set.

In the setting of rational functions the counting is essentially more delicate, and the relative ordering of the poles and alternation points play an important role. The reason for this is that if between two zeros there is a gap with a pole $\mathbf{c}_j$, then the sign at the next gap edge depends on the parity of the pole. This makes it necessary to define the following *sign function*:

$$S_n(x) = \sum_{\substack{1 \leq k \leq n \\ \mathbf{c}_k \neq x_*}} \chi_{[x_*, \mathbf{c}_k)}(x) = \sum_{\mathbf{c} \in \overline{\mathbb{R}} \setminus \{x_*\}} D_n^\infty(\mathbf{c}) \chi_{[x_*, \mathbf{c})}(x).$$

Recall that a function F is called real if for all $z \in \mathbb{C}$, $\overline{F(\overline{z})} = F(z)$.

Definition 1.5 For a real function $F \in \mathcal{L}_n$ with $\|F\|_{\mathsf{E}}1$, a set of distinct points $x_1, \ldots, x_m \in \mathsf{E}$ such that the sequence $(x_*, x_1, \ldots, x_m)$ is cyclically ordered and satisfies the following alternation property

$$F(x_j) = (-1)^{m-j-S_n(x_j)} \tag{1.10}$$

for all $j = 1, \ldots, m$ is called an alternation set. We say that F has a maximal alternation set if $m = n + 1$.

It should be noted that the notion of alternation set depends on the function F, the class $\mathcal{L}_n$, the set E, and the reference point x_*. We note that in what follows, whenever we refer to extremal functions, we mean this in the sense of Problem 1.3.

Theorem 1.6 *(Alternation theorem) A real function $F \in \mathcal{L}_n$ with $\|F\|_{\mathsf{E}} \leq 1$ is an extremal function if and only if it has a maximal alternation set.*

These results generalize standard results from the polynomial case: in the Chebyshev polynomial case, $S_n(x) \equiv 0$, and in the residual polynomial case, S_n has one jump which may or may not affect the alternation criterion, depending on degree. The case of Chebyshev rational functions was also previously formulated in [17]. In all the real extremal problems, previously considered in the literature, the extremizer is seen to be nonconstant. However, in the setting of residual rational functions, the extremizer can be a constant function, and the alternation theorem lets us characterize when this happens:

Theorem 1.7 *The extremal function F_n is constant if and only if the divisor D_n^∞ is of the form (1.5) for points $\mathbf{c}_1, \ldots, \mathbf{c}_n$ such that the points $x_*, \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are in $n+1$ distinct gaps of E.*

In particular, for the Chebyshev problem, $x_* = \mathbf{c}_n$ so F_n is always nonconstant.

These results will be proved in Sect. 2, along with additional properties of F_n and its zeros. Let us assume that F_n is not constant and recall that $(F_n)_\infty \leq D_n^\infty$; we call a point x a "generalized zero" of F_n if either $(F_n)_0(x) > 0$ or if

$$D_n^\infty(x) - (F_n)_\infty(x) > 0.$$

Thus, this notion includes both actual zeros of F_n and places where there is a reduction in the order of the pole compared to the maximal allowed order. These generalized zeros are precisely counted by the divisor

$$D_n^0 := (F_n)_0 + D_n^\infty - (F_n)_\infty.$$

Since an alternation set is on E, note that changing x_* through a single gap only changes the alternation conditions up to an overall j-independent ± 1 factor. Therefore, up to $\pm$ sign, the extremizer F_n for Problem 1.3 is unchanged as x_* varies through a single gap of E. Thus, F_n should be regarded as an extremal function of a gap, rather than of a single point. In particular, the Chebyshev extremizer for Problem 1.1 is the same (up to $\pm$ sign) as the residual extremizer for Problem 1.2 for any x_* in the gap containing $\mathbf{c}_n$. Moreover, F_n might even be extremal for more than one gap. This phenomenon is already known for the so-called Widom maximizer defined below, and is the content of the following corollary.

Corollary 1.8 *Let F_n be an extremal function for $x_* \in (\mathbf{a}, \mathbf{b})$. If $(\mathbf{a}_j, \mathbf{b}_j)$ is a gap such that $|F_n(\mathbf{a}_j)| = |F_n(\mathbf{b}_j)| = 1$ and $D_n^0 = 0$ on $(\mathbf{a}_j, \mathbf{b}_j)$, then up to a ± 1 factor, F_n is an extremal function for any $x_*^j \in (\mathbf{a}_j, \mathbf{b}_j)$.*

From $\deg(F_n)_0 = \deg(F_n)_\infty$ it follows that

$$\deg D_n^0 = \deg D_n^\infty = n \tag{1.11}$$

so we can define the normalized pole counting measure

$$\mu_n := \frac{1}{n} \sum_{\mathbf{c}} D_n^\infty(\mathbf{c}) \delta_{\mathbf{c}} \tag{1.12}$$

and normalized generalized zero counting measure

$$\nu_n := \frac{1}{n} \sum_{\mathbf{c}} D_n^0(\mathbf{c}) \delta_{\mathbf{c}}. \tag{1.13}$$

In Sect. 3, we consider the asymptotics of the extremal rational functions as $n \to \infty$, extending results about root asymptotics from the polynomial setting. For a sequence of divisors D_n^∞ as in Problem 1.3 we define

$$K_{\mathbf{C}} = \overline{\bigcup_{n \geq 1} \operatorname{supp} D_n^\infty}.$$

We will use the following hypothesis repeatedly in the results that follow:

Hypothesis 1.9 *$K_{\mathbf{C}} \cap \mathsf{E} = \emptyset$ and in the topology dual to $C(\overline{\mathbb{R}})$,* w-$\lim_{n\to\infty} \mu_n = \mu$.

A similar combination of assumptions, but with poles away from the convex hull of E, is used in [28, Chapter 6] to study rational interpolation. Some of our current work mirrors our work for orthogonal rational functions [10], but that work required a periodic sequence of poles. In this sense, in addition to studying a different extremal problem, our current setting is more general. To the best of our knowledge all previous works also assumed that the sequence of divisors D_n^∞ is monotonic. Let further $(x_n^*)_{n=0}^\infty$ be a sequence in $\overline{\mathbb{R}} \setminus \mathsf{E}$ which does not accumulate on E.

The behavior of $\log|F_n|$ is governed by the zero and pole distributions. This corresponds to two Riesz representations, with $\log|F_n|$ superharmonic (respectively, subharmonic) away from the set of zeros (respectively, poles). The limiting pole distribution μ directly determines the root asymptotics of the functions F_n and the limiting zero distribution.

We assume that E is not a polar set, i.e., the domain $\Omega = \overline{\mathbb{C}}\setminus\mathsf{E}$ is Greenian, and we denote by $G(z, w) = G_{\mathsf{E}}(z, w)$ the Green function and by $\omega_{\mathsf{E}}(dz, x)$ harmonic measure for this domain.

Theorem 1.10 *(Root asymptotics) Assume that* E *is not a polar set, Hypothesis* 1.9 *holds, and* $(x_n^*)_{n=0}^{\infty}$ *be a sequence in* $\overline{\mathbb{R}}\setminus\mathsf{E}$ *not accumulating on* E*. Then uniformly on compact subsets of* $\mathbb{C}\setminus\mathbb{R}$,

$$\lim_{n\to\infty}\frac{1}{n}\log|F_n(z)| = \int G_{\mathsf{E}}(z, x)d\mu(x).$$

Moreover,

$$\operatorname*{w-lim}_{n\to\infty}\nu_n = \int \omega_{\mathsf{E}}(dz, x)d\mu(x).$$

Our proof of root asymptotics relies on an explicit representation of F_n in terms of the so-called n-extension $\mathsf{E}_n = F_n^{-1}([-1, 1])$. Representations of this type appear for instance in [7, 27]. In particular, using $\mathsf{E}\subset\mathsf{E}_n$ and monotonicity of the Green function, we obtain a Bernstein-Walsh type upper bound for F_n in terms of the Green functions $G_{\mathsf{E}}(z, \mathbf{c})$. This is the major difference between the L^2 and the L^∞ setting. In the L^2 setting [10] an asymptotic upper bound is equivalent to Stahl–Totik regularity of the measure, whereas in the L^∞ setting this bound holds for any n.

As in [25, Corollary 1.2], this can be used to describe the behavior of the leading coefficient.

Theorem 1.10 generalizes known polynomial results, which correspond to the degenerate pole distribution $\mu = \delta_\infty$. Another notable case, related to [10], is of a p-periodically repeating sequence of poles $\mu = \frac{1}{p}\sum_{j=1}^{p}\delta_{\mathbf{c}_j}$.

In Sect. 4, we prove so-called Szegő-Widom asymptotics for F_n. To the best of our knowledge, all previous results are only for polynomial extremal problems. Let Ω be a domain in $\overline{\mathbb{C}}$ which contains ∞ and $\mathsf{E} = \partial\Omega$ be an analytic Jordan curve, T_n the associated Chebyshev polynomial and B_{E} denote the Riemann map that maps $\Omega\to\mathbb{D}$ and $B_{\mathsf{E}}(\infty) = 0$, normalized so that $\lim_{z\to\infty} zB_{\mathsf{E}}(z) > 0$. Faber [13] showed that uniformly on compact subsets of Ω

$$\lim_{n\to\infty} T_n B_{\mathsf{E}}^n = 1. \tag{1.14}$$

In his landmark paper [31], Widom generalized this notion to multiply connected domains. In the following let Ω be a domain in $\overline{\mathbb{C}}$ which contains ∞ so that $\mathsf{E} = \partial\Omega$ is not polar. We will describe the type of results for multiply connected domains, but refer the reader for the precise definitions and statements to Sect. 4. The correct analog for the Riemann map for multiply connected domains is the so-called complex Green function

$$B_{\mathsf{E}}(z, \infty) = e^{-G_{\mathsf{E}}(z,\infty) - i\widetilde{G_{\mathsf{E}}(z,\infty)}}, \tag{1.15}$$

where $\widetilde{G_{\mathsf{E}}(z,\infty)}$ denotes the harmonic conjugate of $G_{\mathsf{E}}(z,\infty)$. To be more precise, since $G_{\mathsf{E}}(z,\infty)$ is harmonic, $B_{\mathsf{E}}(z,\infty)$ is first defined locally and then using the monodromy theorem [26, Theorem 11.2.1] extended to a global multivalued analytic function in Ω. Due to the multivaluedness of B_{E}, one cannot expect that $B_{\mathsf{E}}^n T_n$ converges to a single analytic function as in (1.14). For this reason, Widom considered a related character automorphic extremal problem. Let $z_0 \in \Omega$ and let $\pi_1(\Omega, z_0)$ denote the fundamental group of Ω with basepoint fixed at z_0, and $\pi_1(\Omega)^*$ the group of unitary characters of $\pi_1(\Omega, z_0)$; that is, group homomorphisms from $\pi_1(\Omega, z_0)$ into $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. If F is an analytic function on Ω, then we call F ($\pi_1(\Omega)^*$-) character-automorphic with character α, if

$$F \circ \tilde{\gamma} = e^{2\pi i \alpha(\tilde{\gamma})} F, \quad \forall \tilde{\gamma} \in \pi_1(\Omega, z_*).$$

Let $H_\Omega^\infty(\alpha)$ denote the space of analytic character-automorphic functions, F, in Ω which are uniformly bounded, i.e.,

$$\|F\|_\Omega := \sup_{z \in \Omega} |F(z)| < \infty. \tag{1.16}$$

In his '69 paper [31], Widom considered the extremal problem

$$\sup\{\operatorname{Re} F(x_*) : F \in H_\Omega^\infty(\alpha), \|F\|_\Omega \le 1\} \tag{1.17}$$

under the assumption that E is a finite union of C^2 Jordan curves and arcs and showed existence and uniqueness of the extremizer; let us call this the Widom maximizer. Let χ_n denote the character of B_{E}^n and W_n the Widom maximizer with character χ_n for the extremal point $x_* = \infty$. If E is the finite union of C^2 Jordan curves, Widom showed that uniformly on compact subsets of Ω

$$B_{\mathsf{E}}^n T_n - W_n \to 0. \tag{1.18}$$

If such type of convergence holds, we say T_n has Szegő-Widom asymptotics. The cases of arcs turned out to be essentially harder and for non-real problems only very simple cases such as one arc of the unit circle [8] are known. If $\mathsf{E} \subset \mathbb{R}$ the situation is essentially better, since in this case there are many symmetry properties, which manifests in the fact that the extremal function is real and allows for the explicit representation of the type we will derive in (2.11). If E is a finite union of intervals Christiansen, Simon and Zinchenko [7] showed that T_n has Szegő-Widom asymptotics. In 1971 Widom [32] also showed that (1.17) has a non-trivial solution as long as Ω is of Parreau–Widom type. We will define this notion in Sect. 4, but mention at this place that it also includes infinitely connected domains. Recently Christiansen, Simon, Yuditskii and Zinchenko [5] proved Szegő-Widom asymptotics for T_n if $\mathsf{E} \subset \mathbb{R}$ such that Ω is a regular Parreau–Widom domain with Direct Cauchy theorem and this was later also proved under the same assumptions for residual polynomials [7].

We point out that

$$(T_n)_\infty = n(B_{\mathsf{E}}(\cdot,\infty))_0,$$

which makes $B_{\mathsf{E}}^n T_n$ analytic and in fact a normal family. Since by definition

$$(F_n)_\infty \leq D_n^\infty,$$

in our setting B_{E}^n should be substituted by the product of complex Green functions associated to the divisor D_n^∞, i.e.

$$B_{\mathsf{E}}^{(n)}(z) = e^{i\phi_n} \prod_{\mathbf{c}} D_n^\infty(\mathbf{c}) B_{\mathsf{E}}(z, \mathbf{c}), \tag{1.19}$$

where

$$B_{\mathsf{E}}(z, \mathbf{c}) = e^{-G_{\mathsf{E}}(z,\mathbf{c}) - i\widetilde{G_{\mathsf{E}}(z,\mathbf{c})}} \tag{1.20}$$

and the phase will be specified in Sect. 4. With this modification we prove:

Theorem 1.11 *Let $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$ be a regular Parreau–Widom domain so that the Direct Cauchy theorem holds in Ω. Assume further that Hypothesis 1.9 holds, and $(x_n^*)_{n=0}^\infty$ be a sequence in $\overline{\mathbb{R}} \setminus \mathsf{E}$ without accumulation points in E. Then F_n admits Szegő-Widom asymptotics.*

In Sect. 4, we will provide necessary definitions to state the above theorem more precisely as Theorem 4.5, and provide a proof. For Chebyshev problems on Jordan arcs, it is known that it may be necessary to add an additional factor to (1.18); cf. [29, 30]. In particular, for extremal problems which are symmetric with respect to the real line, this factor is typically $\frac{1}{2}$. The same phenomena can be seen in our asymptotic statement (4.14). Since we chose to consider normalized extremal functions, the additional factor appears in the asymptotics of the extremal function rather than in the assymptotics of the extremal value as in [5, Theorem 1.3].

We want to highlight that this generalizes the known results in several ways. First of all, polynomials correspond to the case that $D_n^\infty = n\chi_{\{\infty\}}$ and so the class of functions that we allow is more general. Secondly, we allow for a sequence of extremal points x_n^*, which in particular means that depending on n, F_n might be a residual or a Chebyshev maximizer.

2 Properties of the Extremal Rational Functions

In this section we study the extremal functions for fixed n. Let us begin by acknowledging that their existence follows by usual arguments. Namely, the leading coefficient λ_n and the value $F_n(x_*)$ are continuous functions of polynomial coefficients of $F_n R_n$. Since $\mathcal{L}_n$ is finite-dimensional, the norm $\|\cdot\|_{\mathsf{E}}$ is mutually equivalent with a norm made from the polynomial coefficients, so Problem 1.3 is an extremal problem for continuous maps on the compact unit ball $\|\cdot\|_{\mathsf{E}} \leq 1$.

Next, we describe the behavior of extremal functions under $\mathrm{PSL}(2, \mathbb{R})$ transformations. This will require the following claim from [27], for which we provide a short proof.

Lemma 2.1 *For every* $z_0 \in \mathbb{C}\backslash\mathbb{R}$ *and* $x \in \overline{\mathbb{R}}$, *there exists* $t \in \overline{\mathbb{R}}$ *such that* $\max_{z\in\overline{\mathbb{R}}}\left|\frac{(z-t)(x-z_0)}{(z-z_0)(x-t)}\right| = 1$ *and* $z = x$ *is the unique maximum.*

Proof Let f be a Möbius transformation mapping $\overline{\mathbb{R}}$ to $\partial\mathbb{D}$ with $f(z_0) = 0$. Since Möbius transformations preserve cross-ratios,

$$\left|\frac{(x-z_0)(z-t)}{(x-t)(z-z_0)}\right| = \left|\frac{f(x)(f(z)-f(t))}{(f(x)-f(t))f(z)}\right| = \left|\frac{f(z)-f(t)}{f(x)-f(t)}\right|.$$

By choosing t so that $f(t) = -f(x)$, we have

$$\left|\frac{(x-z_0)(z-t)}{(x-t)(z-z_0)}\right| = \left|\frac{f(z)-f(t)}{f(x)-f(t)}\right| = \frac{|f(z)+f(x)|}{2} \le 1$$

with equality if and only if $f(z) = f(x)$, i.e., $z = x$. □

In the next lemma, we consider the effect of a conformal transformation on the extremal problems, so we will emphasize dependencies on the poles, the point x_* and the set E where appropriate. We denote by $F_n(z, \mathsf{E}, D_n^\infty; x_*)$ a maximizer for (1.8), and by $\mathcal{L}(D_n^\infty)$ the space defined in (1.6). For a divisor D and a a conformal map $f \in \mathrm{PSL}(2,\mathbb{R})$ we define the pushforward $f_* D = D \circ f^{-1}$. Lemma 2.2 is an analog of [10, Lemma 2.1] adapted to the L^∞ extremal problem (1.8).

We would like to claim that the extremizers move by a conformal map $f \in \mathrm{PSL}(2,\mathbb{R})$ by

$$F_n(f(z), f(\mathsf{E}), f_* D_n^\infty; f(x_*)) = F_n(z, \mathsf{E}, D_n^\infty; x_*).$$

However, this statement would be ambiguous until we prove uniqueness of extremizers, so we have to formulate the claim more carefully:

Lemma 2.2 *Let* $f \in \mathrm{PSL}(2,\mathbb{R})$ *and let* $F_n(z, f(\mathsf{E}), f_* D_n^\infty, f(x_*))$ *be a maximizer of* (1.8) *for* $f_* D_n^\infty$, $f(\mathsf{E})$ *and* $f(x_*)$. *Then* $F_n(f(z), f(\mathsf{E}), f_* D_n^\infty, f(x_*))$ *is a maximizer for* (1.8) *for* D_n^∞, E *and* x_*.

Proof Möbius transformations preserve zeros and their multiplicity, i.e., for any rational function F and any $w \in \overline{\mathbb{C}}$,

$$f_*^{-1}(F)_w = (F \circ f)_w.$$

Therefore, since pushforwards of integral divisors are integral, it follows from (1.6) that

$$F \in \mathcal{L}(f_* D_n^\infty) \implies F \circ f \in \mathcal{L}(D_n^\infty). \tag{2.1}$$

In particular, $F_n(f(z), f(\mathsf{E}), f_* D_n^\infty, f(x_*)) \in \mathcal{L}(D_n^\infty)$. Since $F_n(z, f(\mathsf{E}), f_* D_n^\infty, f(x_*))$ solves the extremal problem on $f(\mathsf{E})$, we have

$$\|F_n(f(\cdot), f(\mathsf{E}), f_* D_n^\infty, f(x_*))\|_{\mathsf{E}} = \|F_n(\cdot, f(\mathsf{E}), f_* D_n^\infty, f(x_*))\|_{f(\mathsf{E})} \le 1.$$

It remains then to show $F(z) := F_n(f(z), f(\mathsf{E}), f_* D_n^\infty, f(x_*))$ is an extremizer for n, E, D_n^∞ and x_*. This will follow from showing that for $d_n > 0$

$$r(f(z), f(x_*))^{d_n} - c_n r(z, x_*)^{d_n} \in \mathcal{L}(D_n^\infty - x_*), \tag{2.2}$$

$$r(f^{-1}(z), x_*)^{d_n} - \frac{1}{c_n} r(z, f(x_*))^{d_n} \in \mathcal{L}(f(D_n^\infty) - f(x_*)), \tag{2.3}$$

for constants $c_n > 0$. Indeed, given (2.2), (2.3), we suppose for the sake of contradiction there is a $\tilde{F} \in \mathcal{L}(D_n^\infty)$ with $\operatorname{Re}\lim_{x\to x_*} \frac{\tilde{F}(x)}{r(x,x_*)^{d_n}} > \operatorname{Re}\lim_{x\to x_*} \frac{F(x)}{r(x,x_*)^{d_n}}$. Then, since $\tilde{F} \circ f^{-1} \in \mathcal{L}(f_* D_n^\infty)$ by (2.1) and

$$\|\tilde{F} \circ f^{-1}\|_{f(\mathsf{E})} \leq 1,$$

we contradict extremality of $F(z)$.

To show (2.2) and (2.3), we note that for the inversions $z \mapsto -\frac{1}{z}$ and the affine transformations $z \mapsto az + b$, $b \in \mathbb{R}$ and $a > 0$, (2.2) and (2.3) follow by elementary computations. Since these generate the group $\mathrm{PSL}(2, \mathbb{R})$, by writing f in this group as $f = f_1 \circ f_2 \circ f_3$, with f_1, f_3 affine, f_2 an inversion, and applying the argument immediately above three times, we have Lemma 2.2. $\square$

Before we state one of the main theorems of the section, we recall that the set $\operatorname{supp}(f)_a$ is called the set of a-points of the function f. Polynomials or entire functions with real ± 1-points play an important role for uniform approximation problems and in the spectral theory of self adjoint operators; cf. [12, 18]. They are also intimately related with the notion of a set of alternation.

We will write

$$\mathsf{E} = \overline{\mathbb{R}} \setminus \bigcup_i (\mathbf{a}_i, \mathbf{b}_i),$$

where $(\mathbf{a}_i, \mathbf{b}_i)$ are the gaps of E, indexed by i from a countable indexing set.

Theorem 2.3 *Let F_n be a maximizer for Problem 1.3. Let $(\mathbf{a}, \mathbf{b})$ be the gap containing x_*.*

- *(i) F_n has only real generalized zeros.*
- *(ii) F_n is real.*
- *(iii) For any distinct points $x_1, x_2 \in \overline{\mathbb{R}}$ such that $D_n^0(x_i) \geq 1$, there is a point $y \in \mathsf{E} \cap (x_1, x_2)$ with $|F_n(y)| = 1$.*
- *(iv) F_n has only simple generalized zeros, i.e., $D_n^0 \leq 1$.*
- *(v) F_n has at most one generalized zero in each gap.*
- *(vi) F_n has no generalized zeros in the gap $(\mathbf{a}, \mathbf{b})$ containing x_*.*
- *(vii) There is a unique extremizer F_n.*
- *(viii) If F_n is not constant, $\{z \in \overline{\mathbb{C}} : F_n(z) \in [-1, 1]\} \subset \overline{\mathbb{R}}$. In particular, all ± 1-points of F_n lie on $\overline{\mathbb{R}}$.*

(ix) *If F_n is not constant, let $m = \deg F_n$ and let the connected components of $F_n^{-1}((-1,1))$ be called open bands of $\mathsf{E}_n := F_n^{-1}[-1,1]$. Then, there are m open bands on E_n, F_n is strictly monotonic on each of them and their endpoints account for all ± 1 points.*

(x) $F_n(\mathbf{a}) = (-1)^{\sum_{\mathbf{c}\in(\mathbf{a},x_*)} D_n^\infty(\mathbf{c})}$ *and* $F_n(\mathbf{b}) = (-1)^{\sum_{\mathbf{c}\in[x_*,\mathbf{b})} D_n^\infty(\mathbf{c})}$,

(xi) *For any gap $(\mathbf{a}_i, \mathbf{b}_i)$ containing a pole $\mathbf{c}_i$, either $|F_n(\mathbf{b}_i)| = 1$ or $|F_n(\mathbf{a}_i)| = 1$. If $D_n^0(\mathbf{c}_i) = 1$, then $|F_n(\mathbf{b}_i)| = |F_n(\mathbf{a}_i)| = 1$.*

Remark Note that (iii) is stronger than saying between two zeros of F_n, we find an extremal point on the set; this statement provides extremal points between a zero and a pole $\mathbf{c}_j$ at which F_n has a reduction in order.

Many of the statements in Theorem 2.3 will be proved by Markov correction arguments. We will call a rational function M a Markov correction term if $MF_n \in \mathcal{L}_n$ and $M(x_*) = 0$. We will define the rational function $\tilde{F}_n = (1 - \epsilon M)F_n$, and note that

$$\tilde{m}_n(x_*) = \operatorname{Re} \lim_{x\to x_*} \frac{\tilde{F}_n(x)}{r(x,x_*)^{d_n}} = \operatorname{Re} \lim_{z\to x_*} \frac{F_n(x)}{r(x,x_*)^{d_n}} = m_n(x_*).$$

If there exists ϵ so that $\|\tilde{F}_n\|_{\mathsf{E}} < 1$, then considering the rescaled function $\tilde{F}_n/\|\tilde{F}_n\|_{\mathsf{E}} \in \mathcal{L}_n$, we see that $\tilde{m}_n(x_*)/\|\tilde{F}_n\|_{\mathsf{E}} > m_n(x_*)$, contradicting the extremality of F_n.

Proof of Theorem 2.3 All the conclusions are invariant under $\mathrm{PSL}(2,\mathbb{R})$ maps, so by Lemma 2.2, it suffices to consider the case $x_* = \infty$. In this case, E is a compact subset of $\mathbb{R}$.

(i): Suppose for the sake of contradiction that there is a generalized zero $z_0 \in \mathbb{C}\backslash\mathbb{R}$. Define

$$\tilde{F}_n(z) = \left(\frac{z-t}{z-z_0}\right) F_n(z)$$

where t is selected so that $\max_{z\in\overline{\mathbb{R}}} \left|\frac{z-t}{z-z_0}\right| = 1$, using Lemma 2.1 for $x = \infty$. Since the maximum at ∞ is unique and E is compact, we have $\|\tilde{F}_n\|_{\mathsf{E}} < 1$, and by the discussion above, this would be a contradiction.

(ii): Since all poles and zeros of F_n are real, we may write $F_n = A\tilde{F}_n$, where $A \in \mathbb{C}$ with $|A| = 1$ and $\tilde{F}_n$ is real. It remains to show that $A \in \mathbb{R}$. Note that $\pm\tilde{F}_n$ are also admissible functions for the extremal problem. Since F_n is extremal and $\tilde{F}_n$ is real, we have

$$\operatorname{Re} \lim_{x\to x_*} \frac{A\tilde{F}_n(x)}{r(x,x_*)^{d_n}} = \operatorname{Re} \lim_{x\to x_*} \frac{F_n(x)}{r(x,x_*)^{d_n}} \geq \pm \operatorname{Re} \lim_{x\to x_*} \frac{\tilde{F}_n(x)}{r(x,x_*)^{d_n}} = \pm \lim_{x\to x_*} \frac{\tilde{F}_n(x)}{r(x,x_*)^{d_n}}.$$

Since $\operatorname{Re} \lim_{x\to x_*} \frac{F_n(x)}{r(x,x_*)^{d_n}} \neq 0$, we conclude that $|\operatorname{Re}(A)| \geq 1$ and therefore $A \in \{1,-1\}$.

(iii): We have $\sup_{\mathsf{E}\cap(x_1,x_2)}|F_n| = \sup_{\mathsf{E}\cap[x_1,x_2]}|F_n| = \max_{\mathsf{E}\cap[x_1,x_2]}|F_n|$. Since F_n is continuous on E we only have to explain the first equality. We only argue for x_1 since x_2 follows analogously. We distinguish two cases. If $F_n(x_1) = 0$, then clearly the sup is not changed by adding x_1. If instead $D_n^\infty(x_1) > 0$, then $x_1 \notin \mathsf{E}$ and $(x_1, x_2) \cap \mathsf{E} = [x_1, x_2) \cap \mathsf{E}$.

Now, we assume for the sake of contradiction that $\max_{\mathsf{E}\cap[x_1,x_2]}|F_n| < 1$. Recalling that $x_* = \infty$ so that $D_n^0(\infty) = 0$, define the Markov correction term

$$M(z; x_1, x_2) = \begin{cases} \frac{1}{(z-x_1)(z-x_2)}, & x_1 < x_2 \\ \frac{1}{(z-x_1)(x_2-z)}, & x_1 > x_2 \end{cases}$$

By distinguishing again the cases $F_n(x_i) = 0$ and $D_n^\infty(x_i) > 0$, we see that in either case MF_n is continuous on $\mathsf{E}\cap[x_1, x_2]$. Thus, by our assumption we find $\varepsilon > 0$ so that $\max_{[x_1,x_2]\cap\mathsf{E}}|\tilde{F}_n| < 1$. Since on the rest of E, the norm is lowered, we may conclude by contradiction.

(iv): Clearly, $D_n^0(x_*) = 0$. With our convention $x_* = \infty$ and by (i), all generalized zeros are in $\mathbb{R}$. Suppose $x \in \mathbb{R}$ with $D_n^0(x) \geq 2$. First, we take $x \notin \mathsf{E}$. We define the Markov correction term $M(z, x) = \frac{1}{(z-x)^2}$. If $x \notin \mathsf{E}$, $z \to M(z, x)$ is continuous on E and so we may find an $\epsilon > 0$ such that $\|\tilde{F}_n\|_{\mathsf{E}} < 1$. If instead, $x \in \mathsf{E}$, then we conclude as in (iii) by continuity of MF_n that we may find a small enough $\epsilon > 0$ so that $\|\tilde{F}_n\|_{\mathsf{E}} < 1$.

(v): It follows from (iii) that between any two generalized zeros there must be a point in E.

(vi): Assume there is a zero in $\mathbb{R} \setminus [\mathbf{b}, \mathbf{a}]$. We use the Markov correction term

$$M(z; x) = \begin{cases} \frac{1}{z-x}, & x < \mathbf{b} \\ \frac{1}{x-z}, & x > \mathbf{a} \end{cases}$$

which is continuous and strictly positive on E. By continuity and compactness, for all small enough $\epsilon > 0$, $\|1 - \epsilon M\|_{\mathsf{E}} < 1$, so $\tilde{F}_n = (1 - \epsilon M)F_n$ once again contradicts extremality.

(vii): Assume that there are two extremizers F_n^1, F_n^2. By convexity, $T_n = \frac{1}{2}(F_n^1 + F_n^2)$ is then also an extremizer. Let $y_i \in \mathsf{E}$ be the points given by (iii) with $|T_n(y_i)| = 1$. We note that by (iv) there are n such points. Then since $|F_n^1(y_i)|, |F_n^2(y_i)| \leq 1$ and $|T_n(y_i)| = 1$, $F_n^1(y_i) = F_n^2(y_i) = T_n(y_i)$ so that $F_n^1(y_i) - F_n^2(y_i) = 0$. Define $H_n = F_n^1 - F_n^2$ and let D_n^0 denote its divisor of generalized zeros. Then $D_n^0(x_*) \geq 1$ and $D_n^0(y_j) \geq 1$ and we conclude that $\deg D_n^0 \geq n + 1$. Since $H_n \in \mathcal{L}_n$, this implies $H_n \equiv 0$ and $F_n^1 = F_n^2$.

(viii): We write F_n in reduced form as $F_n = \frac{P}{Q}$, with $\deg(P) = m$ and note that $\deg Q \leq m$ so that $\deg(F_n) = m$. If F_n is nonconstant, we use a counting argument.

Take two consecutive zeros of F_n, x_1 and x_2. If there is no pole between them, there must be a critical value y and by (iii), it must obey $|F_n(y)| \geq 1$. Separating cases by whether $|F_n(y)| = 1$, we either obtain an (at least) double zero of $F_n^2 - 1$ at y, or zeros on intervals (x_1, y) and (y, x_2). Similarly, if there is a pole $y \in (x_1, x_2)$, by continuity there are ± 1-points on intervals (x_1, y) and (y, x_2).

Thus, counted with multiplicity, there are at least two ± 1-points on this interval. The m simple zeros of P partition $\overline{\mathbb{R}}$ into m such intervals, so we have at least $2m$ total ± 1-points. Since $\deg(F_n) = m$, this construction gives all the ± 1-points of F_n. In particular, this now also holds for the set of $\pm a$-points for any $a \in [-1, 1]$.

(ix): Let I_n^k be the connected components of the open set $F_n^{-1}((-1, 1))$. The previous argument shows that for $a \in (-1, 1)$, the $\pm a$-points are simple. Thus, if $F_n(x) = \pm a$, then $F_n'(x) \neq 0$, so by continuity, the derivative has the same sign on each open band I_n^k. In particular, $F_n(I_n^k) = (-1, 1)$ for each k and there are m connected components, $F_n^{-1}((-1, 1)) = \cup_{k=1}^m I_n^k$. That the endpoints of I_n^k account for all ± 1 points follows from the counting above.

(x): First we show the modulus is 1 at each point. If $F_n \equiv 1$, this is clear. If $\deg(F_n) \geq 1$, we will make use of the zeros of F_n. Suppose for the sake of contradiction that $|F_n(\mathbf{b})| < 1$. Then, define $x := \min\{y : F_n(y) = 0\}$, with $x \geq \mathbf{b}$ by (vi). We have $\sup_{z \in [\mathbf{b}, x]} |F_n(z)| < 1$ by (ix). Define the Markov correction term $M(z, x) = \frac{1}{x - z}$ and note that $M \leq 0$ on $[\mathbf{b}, x]$. By the same arguments as (iii) we derive a contradiction. The same argument at $\mathbf{a}$ shows $|F_n(\mathbf{a})| = 1$.

By (vi), the sign changes on $(\mathbf{a}, \infty)$ can only occur at the poles contained in this interval, which we order as $\mathbf{c}_{n_1} < \cdots < \mathbf{c}_{n_m}$. By (vi), F_n has no reduction of order at the poles at the $\mathbf{c}_{n_i}$, so for a $t \in (\mathbf{c}_{n_m}, \infty)$, $\operatorname{sgn}(F_n(\mathbf{a})) = (-1)^{\sum_{\mathbf{c} \in (\mathbf{a}, \infty)} D_n^\infty(\mathbf{c})} \operatorname{sgn}(F_n(t))$. By our definition of $r(z, \mathbf{c})$, $F_n > 0$ on $(\mathbf{c}_{n_m}, \infty)$. Since $|F_n(\mathbf{a})| = 1$ by our work above, this proves the claim at $\mathbf{a}$. Similar analysis at $\mathbf{b}$, with the modification that the parity of d_n contributes to the sign, completes the proof.

(xi): If F_n is constant, $F_n \equiv 1$ and the claim is clear. Thus, we take F_n nonconstant. By (x), it suffices to consider gaps $(\mathbf{a}_i, \mathbf{b}_i) \neq (\mathbf{a}, \mathbf{b})$. If $|F_n(\mathbf{a}_i)| = 1$ there is nothing to prove. If $|F_n(\mathbf{a}_i)| < 1$, it follows from monotonicity on the bands and $\lim_{x \to \infty} F_n(x) > 1$ that there is a $\tilde{x}_1 < \mathbf{a}_i$ with $F_n(\tilde{x}_1) = 0$. Similar considerations hold for $\mathbf{b}_i$. If $\max\{|F_n(\mathbf{a}_i)|, |F_n(\mathbf{b}_i)|\} < 1$, let $x_1 := \max\{y : y < \mathbf{c}_i, D_n^0(y) = 1\}$ and $x_2 := \min\{y : y > \mathbf{c}_i, D_n^0(y) = 1\}$. By (iii) there must be $y \in (x_1, x_2) \cap \mathsf{E}$, with $|F_n(y)| = 1$. As in the proof of (x), we conclude from monotonicity on the bands that either $y = \mathbf{a}_i$ or $y = \mathbf{b}_i$. If $D_n^0(\mathbf{c}_i) = 1$, we conclude in the same way that there is $y_1 \in (x_1, \mathbf{c}_1)$ and $y_2 \in (\mathbf{c}_1, x_2)$ with $|F_n(y_j)| = 1$ and finally that $y_1 = \mathbf{a}_i$ and $y_2 = \mathbf{b}_i$. ◻

Theorem 2.4 *Let $F \in \mathcal{L}_n$ be real and D_n^0 its generalized zero divisor. Then, any set of alternation points has at most $n + 1 - D_n^0(x_*)$ points.*

Proof Set $m = D_n^0(x_*) \geq 0$ and let $y_j \in \overline{\mathbb{R}} \setminus \{x_*\}$ be the k points with $D_n^0(y_j) > 0$, where regardless of its multiplicity each point appears only once. Since $\deg D_n^0 = n$, we see that $k \leq n - m$. Adding x_* to this list, we cyclically order the points as $(x_*, y_1, \ldots, y_k)$. We note that these points cannot be part of an alternating set, as they either are zeros of F, or coincide with some $\mathbf{c}_i \notin \mathsf{E}$ or $x_* \notin \mathsf{E}$. We also write $y_0 = y_{k+1} = x_*$.

Fix $1 \leq j \leq k+1$. On the interval (y_{j-1}, y_j), F has no generalized zeros, so the sign changes of F only occur at poles, according to the divisor D_n^∞: if $(x_1, x_2) \subset (y_{j-1}, y_j)$, and x_1, x_2 are not poles, then

$$F(x_2) = (-1)^{\sum_{\mathbf{c} \in (x_1, x_2)} D_n^\infty(\mathbf{c})} F(x_1) = (-1)^{S_n(x_2) - S_n(x_1)} F(x_1). \tag{2.4}$$

Thus, x_1, x_2 cannot be two consecutive points of the same alternation set, because by the definition of alternation set, this would imply $F(x_2) = (-1)^{1+S_n(x_2)-S_n(x_1)}F(x_1)$ and lead to contradiction. Thus, any alternation set has at most one point in each interval (y_{j-1}, y_j) for $1 \le j \le k+1$, so any alternation set has at most $k+1 \le n-m+1$ alternation points. □

The above theorem justifies the following definition.

Definition 2.5 We say that F_n has a maximal set of alternation points if it has a set of alternation points of size $n+1$.

Theorem 2.6 *If F_n is the maximizer for* (1.8)*, then it has a maximal set of alternation points.*

Proof Due to Theorem 2.3(vi), $D_n^0(y) = 0$ for all $y \in (\mathbf{a}, \mathbf{b})$ and therefore, using (1.11) and Theorem 2.3(ii),(iv), there is a cyclically ordered sequence $(\mathbf{b}, y_1, \dots, y_n, \mathbf{a})$, so that $D_n^0(y_i) = 1$. By Theorem 2.3(iii), for $2 \le j \le n$, there is a point $x_j \in (y_{j-1}, y_j)$ and $x_j \in \mathsf{E}$, so that $|F_n(x_j)| = 1$. We claim that together with $x_{n+1} = \mathbf{a}$ and $x_1 = \mathbf{b}$ these points form a maximal set of alternation points.

We start with x_{n+1} and x_1. Let $(\mathbf{a}, \mathbf{b})$ be the gap containing x_*. We have

$$S_n(\mathbf{a}) = \sum_{\mathbf{c} \in (\mathbf{a}, x_*)} D_n^\infty(\mathbf{c}).$$

Thus, it follows directly from Theorem 2.3(x) that $x_{n+1} = \mathbf{a}$ is an alternation point. Similarly, we see that

$$S_n(\mathbf{b}) = \sum_{\mathbf{c} \in (\mathbf{b}, x_*)} D_n^\infty(\mathbf{c})$$

and therefore since $\deg D_n^0 = \sum_{\mathbf{c}} D_n^\infty(\mathbf{c}) = n$ and $D_n^\infty(\mathbf{b}) = 0$,

$$n + 1 - 1 - S_n(\mathbf{b}) = \sum_{\mathbf{c}} D_n^\infty(\mathbf{c}) - \sum_{\mathbf{c} \in (\mathbf{b}, x_*)} D_n^\infty(\mathbf{c}) = \sum_{\mathbf{c} \in [x_*, \mathbf{b})} D_n^\infty(\mathbf{c}).$$

Thus, again by Theorem 2.3(x), also $x_1 = \mathbf{b}$ is an alternation point in the above sense.

Now for $j \ge 1$ take x_j, x_{j+1} and $y_j \in (x_j, x_{j+1})$ and assume that x_j is an alternation point. Note that all sign changes of F_n correspond either to a pole of F_n or to y_j. Thus,

$$F_n(x_j) = (-1)^{1+\sum_{\mathbf{c} \in (x_j, x_{j+1})} D_n^\infty(\mathbf{c})} F_n(x_{j+1}). \tag{2.5}$$

This is easily seen if $D_n^\infty(y_j) = 0$. If $D_n^\infty(y_j) > 0$, then $(F_n)_\infty(y_j) = D_n^\infty(y_j) - 1$ and (2.5) still holds. On the other hand

$$S_n(x_j) - S_n(x_{j+1}) = \sum_{\mathbf{c} \in (x_j, x_{j+1})} D_n^\infty(\mathbf{c}).$$

Therefore, x_{j+1} is also an alternating point. Thus, by induction we conclude that $\{x_i\}_{i=1}^{n+1}$ form a maximal set of alternation points for F_n. □

We also have a form of converse to Theorem 2.6, which we prove as the following theorem.

Theorem 2.7 *If $F \in \mathcal{L}_n$ is real and has a maximal alternation set, then F is the unique maximizer for Problem 1.3.*

Proof Let $F \in \mathcal{L}_n$ be real and suppose that it has a maximal set of alternation points $\{x_1, \ldots, x_{n+1}\}$. By relabeling, we assume the cyclic ordering $(x_*, x_1, \ldots, x_{n+1})$. By Theorem 2.4, if F has an alternation set with $n+1$ points, then $(F)_\infty(x_*) = d_n$. Therefore, we can define $\lim_{x\to x_*} F(x)/r(x, x_*)^{d_n} =: \alpha_n \in \mathbb{R} \setminus \{0\}$. It is convenient to rephrase our extremal problem: F_n solves (1.3) if and only if $\tilde{F}_n := \frac{1}{\lambda_n} F_n$ solves the dual problem

$$\inf\{\|\tilde{F}_n\|_{\mathsf{E}} : \lim_{x\to x_*} \frac{\tilde{F}_n(x)}{r(x, x_*)^{d_n}} = 1, \quad \tilde{F}_n \in \mathcal{L}_n\}. \tag{2.6}$$

By this duality and Theorem 2.3(vii), it will suffice to show $\tilde{F} := \frac{1}{\alpha_n} F$ is also an extremizer for (2.6); $\|\tilde{F}\|_{\mathsf{E}} = \|\tilde{F}_n\|_{\mathsf{E}}$. Suppose that $\|\tilde{F}\|_{\mathsf{E}} > \|\tilde{F}_n\|_{\mathsf{E}}$. We define $\tilde{H}_n = \tilde{F} - \tilde{F}_n$ and denote its generalized zero divisor by D_n^0. Our normalization implies that $D_n^0(x_*) \geq 1$. Since $\operatorname{sgn}(H_n(x_j)) = \operatorname{sgn}(F(x_j))$, we have $\operatorname{sgn}(H_n(x_j)) = (-1)^{n+1-j-S_n(x_j)}$ for $1 \leq j \leq n+1$. By the computation (2.4), we conclude that there must be $y_j \in (x_j, x_{j+1})$ with $D_n^0(y_j) \geq 1$ for $1 \leq j \leq n$. Thus, $\deg(D_n^0) \geq n+1$, which contradicts $H_n \in \mathcal{L}_n$. □

In particular, the proof of Theorem 1.6 is now complete and we may prove Corollary 1.8.

Proof of Corollary 1.8 We let $\{x_1, \ldots, x_{n+1}\}$ be an alternation set for F_n and the point x_*, with cyclic ordering $(x_*, x_1, \ldots, x_{n+1})$, where we recall that $x_1 = \mathbf{a}$ and $x_{n+1} = \mathbf{b}$. By definition of S_n we see that for $1 \leq \ell \leq n$ we have

$$\frac{F_n(x_\ell)}{F_n(x_{\ell+1})} = (-1)^{1+\sum_{\mathbf{c}\in(x_\ell, x_{\ell+1})} D_n^\infty(\mathbf{c})}. \tag{2.7}$$

However,

$$\frac{F_n(\mathbf{a})}{F_n(\mathbf{b})} = (-1)^{\sum_{\mathbf{c}\in(\mathbf{a},\mathbf{b})} D_n^\infty(\mathbf{c})}, \tag{2.8}$$

which is easier to see by using the expressions in Theorem 2.3(x). The difference between (2.7) and (2.8) is manifested in the fact that S_n is anchored at $x_* \in (\mathbf{a}, \mathbf{b})$. Moreover, by Theorem 1.6, if there exists a set $\{x_1, \ldots, x_{n+1}\}$ which can be cyclically ordered so that F_n satisfies (2.7) and (2.8), then for any $\tilde{x}_* \in (\mathbf{a}, \mathbf{b})$ up to a factor ± 1, F_n is the maximizer of (1.8).

Denote by x_j^* a point in the gap $(\mathbf{a}_j, \mathbf{b}_j)$. Let $S_n^j(x) := \sum_{\mathbf{c}\in\overline{\mathbb{R}}\setminus\{x_*^j\}} D_n^\infty(\mathbf{c})\chi_{[x_*^j,\mathbf{c})}(x)$. There is $1 \le k \le n$ so that $x_*^j \in (x_k, x_{k+1})$. Let us order the y_i with $D_n^0(y_i) = 1$ cyclically as $(\mathbf{a}, \mathbf{b}, y_1, \ldots, y_n)$. The assumption $D_n^0 = 0$ on $(\mathbf{a}_j, \mathbf{b}_j)$ implies $\mathbf{a}_j, \mathbf{b}_j \in (y_i, y_{i+1})$ for some $1 \le i \le n-1$. By (ix), $\mathbf{a}_j$ and $\mathbf{b}_j$ are the only points in (y_i, y_{i+1}) with $|F_n| = 1$, and since there is exactly one of the x_i in each of the (y_i, y_{i+1}), one and only one of $\mathbf{a}_j$ and $\mathbf{b}_j$ is in the alternation set $\{x_1, \ldots, x_{n+1}\}$. Without a loss of generality we take $x_k = \mathbf{a}_j$. We now claim the set $\{x_1, \ldots, x_k, \mathbf{b}_j, x_{k+1}, \ldots x_n\}$ will form our alternation set. Since S_n^j is now anchored at $x_j^* \in (\mathbf{a}_j, \mathbf{b}_j)$, we need to check (2.8) for the gap $(\mathbf{a}_j, \mathbf{b}_j)$. From the assumption that $D_n^0 = 0$ on $(\mathbf{a}_j, \mathbf{b}_j)$ it follows that

$$\frac{F_n(\mathbf{a}_j)}{F_n(\mathbf{b}_j)} = (-1)^{\sum_{\mathbf{c}\in(\mathbf{a}_j,\mathbf{b}_j)} D_n^\infty(\mathbf{c})}.$$

By the assumption that $\{x_1, \ldots, x_{n+1}\}$ form an alternation set (for S_n), (2.7) (for S_n^j) is clearly satisfied for $\{x_1, \ldots, x_k\}$ and for $\{x_{k+1}, \ldots, x_n\}$. Using again that $\{x_1, \ldots, x_{n+1}\}$ form an alternation set and that $\mathbf{a}_j = x_k$, we have

$$\begin{aligned} F_n(x_{k+1}) &= (-1)^{1+\sum_{\mathbf{c}\in(x_k,x_{k+1})} D_n^\infty(\mathbf{c})} F_n(x_k) = (-1)^{1+\sum_{\mathbf{c}\in(x_k,x_{k+1})} D_n^\infty(\mathbf{c})} \\ &\quad (-1)^{\sum_{\mathbf{c}\in(\mathbf{a}_j,\mathbf{b}_j)} D_n^\infty(\mathbf{c})} F_n(\mathbf{b}_j) \\ &= (-1)^{1+\sum_{\mathbf{c}\in(\mathbf{b}_j,x_{k+1})} D_n^\infty(\mathbf{c})} F_n(\mathbf{b}_j). \end{aligned}$$

Thus, (2.7) is also satisfied for x_{k+1} and $\mathbf{b}_j$. Similarly we can check (2.7) for x_n and x_1 and conclude that up to a factor of ± 1 F_n is also extremal for x_j^*. □

Remark In the above argument, one could have removed x_1 and kept x_{n+1} to form an alternation set for x_*^j.

Next, we describe when the extremizer is constant:

Proof of Theorem 1.7 Suppose E takes the above form. Without a loss of generality we assume that $(x_*, \mathbf{c}_1, \ldots, \mathbf{c}_n)$ are cyclically ordered. Then, $(\mathbf{b}_0, \mathbf{a}_0, x_1, \ldots, x_{n-1})$, where $x_\ell \in \mathsf{E} \cap (\mathbf{c}_\ell, \mathbf{c}_{\ell+1})$ for $1 \le \ell \le n-1$ forms a set of alternation for $F_n \equiv 1$. By Theorem 2.7, F_n is the maximizer for (1.8).

Suppose now the set is not of the above form. If there is a $\mathbf{c}_j$ with $D_n^\infty(\mathbf{c}_j) \ge 2$, by (iv), the extremizer F_n is nonconstant. If there are two distinct poles $\mathbf{c}_i$ and $\mathbf{c}_j$ in a single gap, then F_n cannot be constant by (v). In either case, F_n is nonconstant. □

We record a final corollary of Theorem 2.6.

Corollary 2.8 *If the extremal function F_n is not constant, then* $\deg F_n \ge \lceil \frac{n+1}{2} \rceil$.

Proof By Theorem 2.6, F_n has at least $\lceil \frac{n+1}{2} \rceil$ points with $|F_n| = 1$ with the same sign. Thus, if F_n is nonconstant, it has degree at least $\lceil \frac{n+1}{2} \rceil$, and we can have at most $\lfloor \frac{n-1}{2} \rfloor$ cancellations. □

The set $\mathsf{E}_n = F_n^{-1}([-1, 1])$ is called the n extension of E. Note that by definition it is an extension, i.e., $\mathsf{E} \subset \mathsf{E}_n$. Theorem 2.3, particularly the locating of ± 1 points in (viii) allows us to characterize this set with more specificity in the following theorem.

We recall our ternary order, and let $u_i, v_i \in \overline{\mathbb{R}} \setminus \mathring{\mathsf{E}}$ with $v_i \in [\mathbf{a}_i, \mathbf{b}_i]$ and $u_i \in [\mathbf{a}_i, v_i]$. Then

Theorem 2.9 *For F_n nonconstant, the n extension of E is of the form*

$$\mathsf{E}_n = \mathsf{E} \cup \bigcup_{i \geq 1} [u_i, v_i]$$

with $[u_i, v_i] \subseteq [\mathbf{a}_i, \mathbf{b}_i]$.

The following cases are possible:

(1) The gap remains unchanged, corresponding to $u_i = v_i = \mathbf{a}_i$.
(2) E is extended on one edge, corresponding to $\mathbf{a}_i = u_i$ and $v_i \neq u_i$, $v_i \neq \mathbf{b}_i$, or on the other side, $v_i = \mathbf{b}_i$ and $u_i \neq \mathbf{a}_i$, $u_i \neq v_i$.
(3) An internal interval is added, corresponding to $[u_i, v_i] \subset (\mathbf{a}_i, \mathbf{b}_i)$, $u_i \neq v_i$.
(4) The gap $(\mathbf{a}_i, \mathbf{b}_i)$ may close, corresponding to $\mathbf{a}_i = u_i$ and $\mathbf{b}_i = v_i$.

Moreover, in the following cases there is not extension into a gap:

(i) If $x_* \in (\mathbf{a}_i, \mathbf{b}_i)$, then this gap remains unchanged, i.e., $u_i = v_i = \mathbf{a}_i$.
(ii) If there is a pole $\mathbf{c}_i \in (\mathbf{a}_i, \mathbf{b}_i)$ and $D_n^0(\mathbf{c}_i) = 1$, then this gap remains unchanged, i.e., $u_i = v_i = \mathbf{a}_i$.

Remark (i) As we will see in the proof, for gaps $(\mathbf{a}_i, \mathbf{b}_i)$ containing poles of F_n, which is guaranteed for $D_n^\infty(\mathbf{c}_i) \geq 2$ by (iv), only the first three behaviors are possible.
(ii) If there is an interval added to E_n as in (3) above, then this is always related to a zero x_i of F_n and moreover $|F_n(\mathbf{a}_i)| = |F_n(\mathbf{b}_i)| = 1$. Clearly, if this zero approaches a pole, the interval around it becomes smaller. In this sense (ii) of the above theorem can be viewed as a limit of such situations, where the additional interval degenerates to a point.

Proof Applying conformal invariance of the setting, we assume again that $x_* = \infty$ and E is a compact subset of $\mathbb{R}$. Since we will prove (i) independently, we can assume that all extensions occur in bounded gaps. We first note that any internal interval cannot degenerate to a point, i.e. when $u_i = v_i$ in (3), since due to 2.3(ix) there are m open bands and their endpoints account for all ± 1 points. Thus, if the extension of the gap is not of the above form, then there would either be more then one internal interval, an extension on both sides or an extension combined with an internal interval. All cases imply that there are open bands $I_k = (y_k^-, y_k^+)$, $k = 1, 2$, so that $y_1^+, y_2^- \notin \mathsf{E}$. Let x_k denote the simple zero of F_n on these open bands. Using that $|F_n| < 1$ on I_k and $y_1^+, y_2^- \notin \mathsf{E}$, we see that $\max_{[x_1, x_2] \cap \mathsf{E}} |F_n| < 1$, contradicting Theorem 2.3(iii).

Let us now prove (i): Due to (x) and (ix) of Theorem 2.3, an extension would contain an open band that lies entirely in $(\mathbf{a}_i, \mathbf{b}_i)$. Therefore in particular, this would lead to a zero of F_n on the extremal gap contradicting (vi) of Theorem 2.3.

It remains to prove (ii): In this case again due to (xi) and (ix) of Theorem 2.3, this would lead to an open band that lies entirely in $(\mathbf{a}_i, \mathbf{b}_i)$ forcing F_n to have an additional zero in this gap. But since already $D_n(\mathbf{c}_i) = 1$, this would contradict (v) of Theorem 2.3. □

In the following let us assume that F_n is nonconstant so that $\overline{\mathbb{C}}\setminus \mathsf{E}_n$ is Greenian. Note that due to Theorem 2.9(3), E_n is a finite union of proper intervals and in particular is regular for the Dirichlet problem. We define

$$B_n(z) = e^{i\phi_n} \prod_{\mathbf{c}} (F_n)_\infty(\mathbf{c}) B_{\mathsf{E}_n}(z, \mathbf{c}), \tag{2.9}$$

and normalize the phase of B_n by the condition

$$\lim_{x\to x_*} B_n(x) r(x, x_*)^{d_n} > 0. \tag{2.10}$$

Recall that in general $B_{\mathsf{E}_n}(z, \mathbf{c})$ define multivalued functions. However, we will show that their product $B_n(z)$ is in fact single valued in $\overline{\mathbb{C}} \setminus \mathsf{E}_n$.

Theorem 2.10 *B_n is a single-valued analytic function on $\overline{\mathbb{C}}\setminus \mathsf{E}_n$ and*

$$F_n(z) = \frac{1}{2}\left(B_n(z) + \frac{1}{B_n(z)}\right). \tag{2.11}$$

Proof Recall that $\mathsf{E}_n = \{z \in \overline{\mathbb{C}} : F_n(z) \in [-1, 1]\}$. Therefore, since the Joukowsky map $J(\zeta) = \frac{1}{2}\left(\zeta + \frac{1}{\zeta}\right)$ maps $\mathbb{D}$ conformally onto $\overline{\mathbb{C}} \setminus [-1, 1]$, the function

$$\Psi_n(z) = J^{-1}(F_n(z)),$$

is well defined and single-valued in $\overline{\mathbb{C}}\setminus \mathsf{E}_n$. Moreover, for $x \in \mathsf{E}_n$, $\lim_{z\to x} |\Psi_n(z)| = 1$ and $\Psi_n(z)$ has a zero of multiplicity $(F_n)_\infty(\mathbf{c})$ at each $\mathbf{c}$. Thus, we conclude by the maximum principle that

$$-\log|\Psi_n(z)| = \sum_{\mathbf{c}} (F_n)_\infty(\mathbf{c}) G_{\mathsf{E}_n}(z, \mathbf{c}) = -\log|B_n(z)|.$$

Thus, by adding the complex conjugate, B_n is defined up to a unimodular constant c. Finally,

$$\begin{aligned} 0 < \lim_{x\to x_*} \frac{F_n(x)}{r(x, x_*)^{d_n}} &= \frac{1}{2} \lim_{x\to x_*} \left(\frac{cB_n(x)}{r(x, x_*)^{d_n}} + \frac{1}{cB_n(x) r(x, x_*)^{d_n}}\right) \\ &= \frac{1}{2} \lim_{x\to x_*} \frac{1}{cB_n(x) r(x, x_*)^{d_n}}. \end{aligned}$$

Using the normalization (2.10), we conclude $c = 1$ and obtain (2.11). □

This has the following consequence:

Lemma 2.11 *Let F_n be represented as in (2.11) and let I_n be an open band of E_n. Then*

$$1 = \sum_{\mathbf{c}} (F_n)_\infty(\mathbf{c}) \omega_{\mathsf{E}_n}(I_n, \mathbf{c}). \tag{2.12}$$

Proof Recall that

$$F_n(z) = J(B_n(z)) \tag{2.13}$$

and that F_n is strictly monotonic on I_n. That is, either F_n increases from -1 to 1 or decreases from 1 to -1 strictly monotonically. Let $I_n = (a, b)$. Since $J : \partial\mathbb{D} \cap \mathbb{C}_\pm \to (-1, 1)$ bijectively, it follows from the definition of I_n and (2.13) that

$$|\arg B_n(b) - \arg B_n(a)| = \pi.$$

By using the Cauchy-Riemann equations, we get

$$\arg B_n(b) - \arg B_n(a) = \int_a^b \frac{\partial G_n(x)}{\partial n} dx.$$

On the other hand

$$\omega_{\mathsf{E}_n}(dx, \mathbf{c}) = \frac{1}{\pi} \frac{\partial G_{\mathsf{E}_n}(x, \mathbf{c})}{\partial n} dx.$$

Thus, we get

$$\arg B_n(b) - \arg B_n(a) = \pi \sum_{\mathbf{c}} (F_n)_\infty(\mathbf{c}) \omega_{\mathsf{E}_n}(I_n, \mathbf{c})$$

and the claim follows. □

We finish this section with a Bernstein-Walsh lemma for rational functions.

Lemma 2.12 *Let $K \subset \mathbb{C}$ be a compact, nonpolar set such that $\overline{\mathbb{C}} \setminus K$ is connected. Let h be a meromorphic function on $\overline{\mathbb{C}}$ Then,*

$$\frac{|h(z)|}{\|h\|_K} \leq e^{\sum_{\mathbf{c}} (h)_\infty(\mathbf{c}) G_K(z, \mathbf{c})} \tag{2.14}$$

If we assume in addition that $K \subset \overline{\mathbb{R}}$ and that h is real, then

$$\frac{|h(z)|}{\|h\|_K} \leq \frac{1}{2} \left(e^{\sum_{\mathbf{c}} (h)_\infty(\mathbf{c}) G_K(z, \mathbf{c})} + e^{-\sum_{\mathbf{c}} (h)_\infty(\mathbf{c}) G_K(z, \mathbf{c})} \right). \tag{2.15}$$

Proof For (2.14) we follow the standard proof of the Bernstein-Walsh lemma. Set $H = h/\|h\|_K$ and consider $F(z) = \log|H(z)| - \sum_{\mathbf{c}}(h)_\infty(\mathbf{c})G_K(z, \mathbf{c})$. Then, F is subharmonic in $\Omega = \overline{\mathbb{C}} \setminus K$ and for q.e. $\zeta \in \partial\Omega$ we have $\limsup_{z\to\zeta} F(z) \le 0$. Moreover, if $\mathcal{V}_{\mathbf{c}}$ are vicinities of the points with $(h)_\infty(\mathbf{c}) > 0$ and $\mathcal{V} = \cup_c \mathcal{V}_c$, then $\log|H(z)|$ is subharmonic on $\overline{\mathbb{C}}\backslash\mathcal{V}$ and thus bounded above by [19, Theorem 2.1.2]. Since the logarithmic pole on $\mathcal{V}_{\mathbf{c}}$ is canceled, F is also bounded above on $\mathcal{V}$ and we conclude from the maximum principle [14, Theorem 8.1] that $F(z) \le 0$ in Ω.

Assume that H is real and that K is real. Define $K_H = \{z \in \overline{\mathbb{C}} : H(z) \in [-1, 1]\}$, but note that K is not necessarily a subset of $\mathbb{R}$. However, using that H is real, we have that $K \subset K_H$. Now, as in the proof of Theorem 2.10 we see that

$$H(z) = \frac{1}{2}\left(e^{G_H(z)+i\widetilde{G_H(z)}} + e^{-(G_H(z)+i\widetilde{G_H(z)})}\right), \quad G_H(z) = \sum_{\mathbf{c}}(h)_\infty(\mathbf{c})G_{K_H}(z, \mathbf{c}).$$

Let us also put $G(z) = \sum_{\mathbf{c}}(h)_\infty(\mathbf{c})G_K(z, \mathbf{c}_\ell)$. Then it follows from the monotonicity of Green functions with respect to the domain that for $z \in \overline{\mathbb{C}} \setminus K_H$, we have

$$|H(z)| = \left|\cosh\left(G_H(z) + i\widetilde{G_H(z)}\right)\right| \le \cosh G_H(z) \le \cosh G(z).$$

Note that for $z \in K_H \backslash K$, $G(z) > 0$ and thus (2.15) also holds for such z. This finishes the proof. □

We point out that (2.14) is an analog of the standard Bernstein-Walsh lemma, whereas (2.15) is a fairly recent improvement of Schiefermayr for real polynomial problems [23]. Note that this also implies that (2.15) holds for $x_* \in \mathbb{R} \setminus K$, without the extra assumption on h_n to be real. This follows from Theorem 2.3, where we showed that the residual extremizer is always real.

3 Root Asymptotics

We now turn to the study of the limiting behavior of F_n as $n \to \infty$. In this section, we will often assume Hypothesis 1.9 holds, and that (x_n^*) is a sequence in $\overline{\mathbb{R}} \setminus \mathsf{E}$ without accumulation points in E. We note that the first part of Hypothesis 1.9 implies $\operatorname{supp}\mu \cap \mathsf{E} = \emptyset$. Let further ν_n be the normalized counting measure of generalized zeros of F_n, i.e,

$$\nu_n = \frac{1}{n}\sum_x D_n^0(x)\delta_x.$$

We define the family of functions

$$h_n(z) = \frac{1}{n}\log|F_n(z)| \tag{3.1}$$

and note that h_n is subharmonic in $\overline{\mathbb{C}}\setminus \operatorname{supp} D_n^\infty$; in particular, all functions h_n are subharmonic in

$$\Omega_{\mathbf{C}} = \overline{\mathbb{C}} \setminus K_{\mathbf{C}}.$$

We start with an upper estimate:

Lemma 3.1 *Assume Hypothesis 1.9 holds. Then, for any $z \in \mathbb{C} \setminus \mathbb{R}$ we have*

$$\limsup h_n(z) \le \int G_{\mathsf{E}}(z, x) d\mu(x). \tag{3.2}$$

Proof Due to Lemma 2.12 and the definition of μ_n we have

$$h_n(z) \le \int G_{\mathsf{E}}(z, x) d\mu_n(x).$$

On the other hand, since $\mu_n \to \mu$ and by continuity of $G_{\mathsf{E}}(z, y)$ on $K_{\mathbf{C}}$ we have

$$\lim_{n\to\infty} \int G_{\mathsf{E}}(z, x) d\mu_n(x) = \int G_{\mathsf{E}}(z, x) d\mu(x).$$

□

We continue with some facts about potentials.

Lemma 3.2 *Let $\mathsf{E} \subsetneq \overline{\mathbb{R}}$ be closed and not polar so that $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$ is Greenian and μ be a probability measure supported on $\overline{\mathbb{R}}$ with $\operatorname{supp}\mu \cap \mathsf{E} = \emptyset$. Then $\int G_{\mathsf{E}}(z, x) d\mu(x)$ defines a positive superharmonic function in Ω and a harmonic function in $\Omega \setminus \operatorname{supp}\mu$. Moreover, as a harmonic function, it has a unique subharmonic extension to $\overline{\mathbb{C}}\setminus\operatorname{supp}\mu$, which vanishes q.e. on E.*

Proof If $\operatorname{supp}\mu \subset \mathbb{R}$, it follows from [21, Theorem II.5.1] and the minimum principle for superharmonic functions that $\int G_{\mathsf{E}}(z, x) d\mu(x)$ defines a positive superharmonic function in Ω and a harmonic function in $\Omega \setminus \operatorname{supp}\mu$ that vanishes q.e. on E. In particular, locally in vicinities of E it is subharmonic and vanishes away from a polar set. Thus, by [1, Theorem 5.2.1.], for $\zeta \in \mathsf{E}$

$$\int G_{\mathsf{E}}(\zeta, x) d\mu(x) = \limsup_{z\to\zeta} \int G_{\mathsf{E}}(z, x) d\mu(x)$$

defines the unique subharmonic extension to $\Omega \setminus \operatorname{supp}\mu$; since all claims are conformally invariant, the general case follows. □

Lemma 3.3 *Assume Hypothesis 1.9. Then, the set $K_{\mathbf{C}}$ intersects only finitely many open gaps.*

Proof $K_{\mathbf{C}}$ is a closed subset of $\overline{\mathbb{R}}$, so it is compact. It is contained in $\overline{\mathbb{R}} \setminus \mathsf{E}$, so its cover by the open sets $(\mathbf{a}_j, \mathbf{b}_j)$ has a finite subcover; in other words, $K_{\mathbf{C}}$ only intersects finitely many gaps. □

We obtain immediately the following corollary:

Corollary 3.4 *Assume Hypothesis 1.9. Then, for n sufficiently large, F_n is non-constant.*

Proof Because $D_n^0 \leq 1$ for any gap and there is at most one generalized zero per gap due to Theorem 2.3(iv),(v), the claim follows by Lemma 3.3 and $\deg D_n^\infty = n$. □

Since we are interested in asymptotics of F_n as $n \to \infty$, we assume from now on that F_n is non-constant.

Lemma 3.5 *Fix an open set $O \subset \overline{\mathbb{R}} \setminus \mathsf{E}$ so that $\mu(O) > 0$. Then*

$$\lim_{n\to\infty} \sum_{\mathbf{c}\in O} D_n^\infty(\mathbf{c}) = +\infty. \tag{3.3}$$

Proof By definition,

$$\mu_n(O) = \frac{1}{n} \sum_{\mathbf{c}\in O} D_n^\infty(\mathbf{c}).$$

By the Portmanteau theorem, $\liminf_{n\to\infty} \mu_n(O) \geq \mu(O) > 0$, so

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{\mathbf{c}\in O} D_n^\infty(\mathbf{c}) > 0$$

which implies (3.3). □

The following analog of Koosis's formula for the Martin or Phragmén Lindelöf function [16, Theorem on page 407] will be very useful. It was already used in [5, Proposition 4.3].

Lemma 3.6 *Let $\mathsf{E}_1 \subset \mathsf{E}_2 \subset \overline{\mathbb{R}}$ so that E_1 is not polar and let $\mathbf{c} \in \overline{\mathbb{R}} \setminus \mathsf{E}_2$. Then,*

$$G_{\mathsf{E}_1}(z, \mathbf{c}) - G_{\mathsf{E}_2}(z, \mathbf{c}) = \int_{\mathsf{E}_2\setminus\mathsf{E}_1} G_{\mathsf{E}_1}(z, x)\omega_{\mathsf{E}_2}(dx, \mathbf{c}). \tag{3.4}$$

Proof Since (3.4) is conformally invariant, by applying a conformal map we can assume that $\infty \in \mathsf{E}_1$, i.e., $\Omega_1 = \overline{\mathbb{C}} \setminus \mathsf{E}_1 \subset \mathbb{C}$. Define also $\Omega_2 = \overline{\mathbb{C}} \setminus \mathsf{E}_2$. Since the logarithmic pole at $\mathbf{c}$ is canceled, $G_{\mathsf{E}_1}(z, \mathbf{c}) - G_{\mathsf{E}_2}(z, \mathbf{c})$ defines a superharmonic function on Ω_1 which is bounded. Moreover, its Riesz measure is given by $\omega_{\mathsf{E}_2}(dx, \mathbf{c})|\Omega_1$. Since $\mathsf{E}_1 \subset \mathsf{E}_2$ it follows by the maximum principle that $G_{\mathsf{E}_1}(z, \mathbf{c}) - G_{\mathsf{E}_2}(z, \mathbf{c}) \geq 0$. Thus, in particular it has a nonnegative subharmonic minorant in Ω_1 and it follows by the Riesz decomposition theorem that

$$G_{\mathsf{E}_1}(z, \mathbf{c}) - G_{\mathsf{E}_2}(z, \mathbf{c}) = \int_{\Omega_1} G_{E_1}(z, x)\omega_{\mathsf{E}_2}(dx, \mathbf{c}) + u(z),$$

where u is the greatest harmonic minorant of $G_{\mathsf{E}_1}(z, \mathbf{c}) - G_{\mathsf{E}_2}(z, \mathbf{c})$. We have already seen that $u \geq 0$. On the other hand, since E_1 is the boundary for Ω_1 and Ω_2 it follows that for q.e. $x \in \mathsf{E}_1$ we have

$$\limsup_{z\to x} u(z) \leq \limsup_{z\to x}(G_{\mathsf{E}_1}(z, \mathbf{c}) - G_{\mathsf{E}_2}(z, \mathbf{c})) = 0.$$

Thus, u is a bounded harmonic function in Ω_1 which vanishes $q.e.$ on E_1. It follows by the maximum principle [14, Corollary 8.3] that $u = 0$ and we obtain (3.4). □

Compared to the standard Chebyshev problem, we encounter a technical difference for residual extremal functions. Let $(\mathbf{a}_i, \mathbf{b}_i)$ be a gap so that (3.3) is satisfied for $O = (\mathbf{a}_i, \mathbf{b}_i)$. We want to estimate $G_{\mathsf{E}_n}(z, \mathbf{c})$ for $\mathbf{c} \in (\mathbf{a}_i, \mathbf{b}_i)$. But since $(\mathbf{a}_i, \mathbf{b}_i)$ is not necessarily the extremal gap, there can be an extension (u_i, v_i) in this gap, which intuitively makes $G_{\mathsf{E}_n}(z, \mathbf{c})$ smaller if $[u_i, v_i]$ is close to $\mathbf{c}$. However, we have already encountered in Theorem 2.9(ii), that a cancellation of a pole can be regarded as a degenerated internal interval. Thus, we are led to expect that an additional interval can have no more "effect" than reducing the number of Green functions in the sum by one. This is the content of the following lemma:

Lemma 3.7 *Let E_n and u_i, v_i be defined as in Theorem 2.9. Fix a gap $(\mathbf{a}_i, \mathbf{b}_i)$ and define $\mathsf{E}_n^i = \mathsf{E}_n \setminus (\mathbf{a}_i, \mathbf{b}_i)$. Let $z \in \overline{\mathbb{C}}\setminus\mathsf{E}_n$ and $(F_n)_\infty(z) = 0$. Then there is a $t \in [\mathbf{a}_i, \mathbf{b}_i]$ such that $G_{\mathsf{E}}(t, z) = \max_{x\in[\mathbf{a}_i,\mathbf{b}_i]} G_{\mathsf{E}}(x, z)$ and we have*

$$\sum_{\mathbf{c}}(F_n)_\infty(\mathbf{c})G_{\mathsf{E}_n}(z, \mathbf{c}) \geq \sum_{\mathbf{c}}(F_n)_\infty(\mathbf{c})G_{\mathsf{E}_n^i}(z, \mathbf{c}) - G_{\mathsf{E}}(z, t). \tag{3.5}$$

In particular, if $z \notin (\mathbf{a}_i, \mathbf{b}_i)$, then

$$\lim_{n\to\infty}\sum_{\mathbf{c}}(F_n)_\infty(\mathbf{c})G_{\mathsf{E}_n^i}(z, \mathbf{c}) = \infty \implies \lim_{n\to\infty}\sum_{\mathbf{c}}(F_n)_\infty(\mathbf{c})G_{\mathsf{E}_n}(z, \mathbf{c}) = \infty. \tag{3.6}$$

Proof If $z \in (\mathbf{a}_i, \mathbf{b}_i)$, then $t = z$ and (3.5) is trivial. Thus, let $z \notin (\mathbf{a}_i, \mathbf{b}_i)$.

Since E_n is a finite union of intervals it is clearly not polar and putting $\mathsf{E}_n \setminus \mathsf{E}_n^i = [u_i, v_i]$, we obtain from Lemma 3.6 that

$$G_{\mathsf{E}_n^i}(z, \mathbf{c}) - G_{\mathsf{E}_n}(z, \mathbf{c}) = \int_{u_i}^{v_i} G_{\mathsf{E}_n^i}(z, x)\omega_{\mathsf{E}_n}(\mathrm{d}x, \mathbf{c}). \tag{3.7}$$

By [19, Theorem 2.1.2] a subharmonic function attains its maximum on compacts and thus t is well defined. Define

$$\rho_n(dx) = \sum_{\mathbf{c}}(F_n)_\infty(\mathbf{c})\omega_{\mathsf{E}_n}(dx, \mathbf{c}),$$

and note that it follows from Theorem 2.9 and Lemma 2.11 that $\rho_n([u_i, v_i]) \leq 1$. Moreover, by the maximum principle

$$G_{\mathsf{E}_n^i}(z, x) \leq G_{\mathsf{E}}(z, x).$$

Thus,

$$\sum_{\mathbf{c}} (F_n)_\infty(\mathbf{c}) \int_{u_i}^{v_i} G_{\mathsf{E}_n^i}(z,x)\omega_{\mathsf{E}_n}(\mathrm{d}x,\mathbf{c})$$
$$= \int_{u_i}^{v_i} G_{\mathsf{E}_n^i}(z,x)\rho_n(dx) \le \int_{u_i}^{v_i} G_{\mathsf{E}}(z,x)\rho_n(dx) \le G_{\mathsf{E}}(z,t).$$

Combining this with (3.7) yields (3.5). □

By the representation (2.11), we have

$$h_n(z) = -\frac{1}{n}\log|B_n(z)| - \frac{1}{n}\log 2 + \frac{1}{n}\log|1 + B_n(z)^2|. \tag{3.8}$$

The next lemma shows that the the asymptotics of h_n for $n \to \infty$ are determined by the term $-\frac{1}{n}\log|B_n(z)|$. In fact, we even prove a stronger statement, which will be needed in Sect. 4.

Lemma 3.8 *Suppose Hypothesis 1.9 holds. Then, uniformly on compact subsets of* $\mathbb{C}\setminus\mathbb{R}$ *we have*

$$\lim_{n\to\infty} \log\left|1 + B_n(z)^2\right| = 0. \tag{3.9}$$

If we pass to a subsequence such that $\lim_{\ell\to\infty} x_{n_\ell}^* = x_\infty^*$ *and* $(\mathbf{a},\mathbf{b})$ *denotes the gap containing* x_∞^*, *then also for* $z \in (\mathbf{a},\mathbf{b})$

$$\lim_{\ell\to\infty} \log\left|1 + B_{n_\ell}(z)^2\right| = 0. \tag{3.10}$$

Proof Consider B_n as an analytic single-valued function on $\mathbb{C}_+$ or $\mathbb{C}_-$ and note that $0 < |B_n(z)| < 1$. Thus, $\log\left|1 + B_n(z)^2\right| = \operatorname{Re}\log(1 + B_n(z)^2)$ defines a family of harmonic functions which is uniformly bounded from above. Thus, by the Harnack principle, the family is precompact in the space of harmonic functions together with the function which is identically $-\infty$. Therefore, it suffices to show that pointwise for fixed z every subsequence has a subsequence so that (3.9) holds. Let us pass to a subsequence so that $\lim_{\ell\to\infty} x_{n_\ell}^* = x_\infty^*$ and let $(\mathbf{a},\mathbf{b})$ denote the gap containing x_∞^*. If necessary, we pass to a further subsequence so that $x_{n_\ell} \in (\mathbf{a},\mathbf{b})$ for all $\ell > 0$. Since for $|z| < 1$

$$|\log(|1+z|)| = |\operatorname{Re}(\log(1+z))| \le |\log(1+z)| = \left|\int_0^1 \frac{z}{1+zt}\mathrm{d}t\right| \le \frac{|z|}{1-|z|},$$

it suffices to show that

$$\lim_{\ell\to\infty} |B_{n_\ell}(z)| = 0.$$

By (2.9) and (1.20) this is equivalent to

$$\lim_{\ell\to\infty}\sum_{\mathbf{c}}(F_{n_\ell})_\infty(\mathbf{c})G_{\mathsf{E}_{n_\ell}}(z,\mathbf{c}) = +\infty. \tag{3.11}$$

Since $\mathsf{E}\cap K_{\mathbf{C}} = \emptyset$, we find a gap $(\mathbf{a}_i,\mathbf{b}_i)$ and $\varepsilon > 0$ so that $\mu((\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon)) > 0$. Thus, by Lemma 3.5 we have

$$\lim_{\ell\to\infty}\sum_{\mathbf{c}\in(\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon)} D^\infty_{n_\ell}(\mathbf{c}) = +\infty. \tag{3.12}$$

Note that it could be that $(\mathbf{a}_i,\mathbf{b}_i) = (\mathbf{a},\mathbf{b})$, which causes no problems in the following.

Set $\mathsf{E}^i = \overline{\mathbb{R}}\setminus((\mathbf{a},\mathbf{b})\cup(\mathbf{a}_i,\mathbf{b}_i))$ and $\mathsf{E}^i_{n_\ell} = \mathsf{E}_{n_\ell}\setminus(\mathbf{a}_i,\mathbf{b}_i)$. By Theorem 2.9(i), $\mathsf{E}^i_{n_\ell}\subset\mathsf{E}^i$, so the maximum principle yields

$$G_{\mathsf{E}^i}(z,\mathbf{c}) \le G_{\mathsf{E}^i_{n_\ell}}(z,\mathbf{c}). \tag{3.13}$$

Fix $z\in\mathbb{C}_+\cup\mathbb{C}_-\cup(\mathbf{a},\mathbf{b})$ and note that lower semicontinuity implies

$$0 < \delta = \min_{\mathbf{c}\in[\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon]} G_{\mathsf{E}^i}(z,\mathbf{c}).$$

Then, by Theorem 2.3(v)

$$\sum_{\mathbf{c}\in(\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon)}(F_{n_\ell})_\infty(\mathbf{c})G_{\mathsf{E}^i}(z,\mathbf{c}) \ge \delta\Bigg(-1+\sum_{\mathbf{c}\in(\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon)} D^\infty_{n_\ell}(\mathbf{c})\Bigg).$$

Thus, by (3.12) we obtain

$$\lim_{\ell\to\infty}\sum_{\mathbf{c}\in(\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon)}(F_{n_\ell})_\infty(\mathbf{c})G_{\mathsf{E}^i}(z,\mathbf{c}) = \infty. \tag{3.14}$$

Since by positivity of the Green function

$$\sum_{\mathbf{c}}(F_{n_\ell})_\infty(\mathbf{c})G_{\mathsf{E}^i}(z,\mathbf{c}) \ge \sum_{\mathbf{c}\in(\mathbf{a}_i+\varepsilon,\mathbf{b}_i-\varepsilon)}(F_{n_\ell})_\infty(\mathbf{c})G_{\mathsf{E}^i}(z,\mathbf{c})$$

we obtain together with (3.13) that

$$\lim_{\ell\to\infty}\sum_{\mathbf{c}}(F_{n_\ell})_\infty(\mathbf{c})G_{\mathsf{E}^i_{n_\ell}}(z,\mathbf{c}) = \infty.$$

By an application of Lemma 3.7 we obtain (3.11) which concludes the proof. □

Lemma 3.9 *Assume Hypothesis 1.9Then, for $z \in \mathbb{C} \setminus \mathbb{R}$, we have*

$$\liminf_{n\to\infty} \frac{1}{n} \log |F_n(z)| \geq 0. \tag{3.15}$$

Proof Fix $z \in \mathbb{C} \setminus \mathbb{R}$ and recall (3.8). Noting that $|B_n(z)| \leq 1$, the claim follows from Lemma 3.8. □

In contrast to the classical polynomial setting, our limits will be described by the difference of two potentials, one corresponding to the zeros of F_n, leading to a subharmonic part and one corresponding to the poles leading to a superharmonic part. Since in the following considerations we will work with the Riesz measures for both of them, there is no natural choice of a "coordinate system" and it will be convenient to apply conformal maps to logarithmic potentials. For a probability measure ν with $\operatorname{supp}\nu \subsetneq \overline{\mathbb{R}}$ and $z_* \in \overline{\mathbb{R}} \setminus \operatorname{supp}\nu$, let us introduce the notation

$$\Phi_\nu(z, z_*) = \int K(x, z; z_*) d\nu(x),$$

where

$$K(x, z; z_*) = \begin{cases} \log \left| 1 - \frac{z - z_*}{x - z_*} \right|, & z_* \neq \infty, \\ \log |z - x|, & z_* = \infty. \end{cases}$$

It is straightforward to see that if $z_1, z_2 \in \overline{\mathbb{R}} \setminus \operatorname{supp}\nu$, then there is $\beta \in \mathbb{R}$ so that

$$\Phi_\nu(z, z_1) = \beta + \Phi_\nu(z, z_2).$$

Lemma 3.10 *Let ν be a probability measure on $\overline{\mathbb{R}}$, $\operatorname{supp}\nu \subset \mathsf{E}$ and $f \in \mathrm{PSL}(2, \mathbb{R})$. If $f(\infty) = \infty$, then*

$$\Phi_\nu(z, z_*) = \Phi_{f_*\nu}(f(z), f(z_*)).$$

Otherwise,

$$\Phi_\nu(z, z_*) = \Phi_{f_*\nu}(f(z), f(z_*)) - \Phi_{f_*\delta_\infty}(f(z), f(z_*)).$$

Proof Let us first assume that $f(\infty) = \infty$, i.e., $f(z) = az + b$ with $a \neq 0$. Then we have

$$1 - \frac{z - z_*}{x - z_*} = 1 - \frac{f(z) - f(z_*)}{f(x) - f(z_*)}.$$

Thus, the claim follows by the transformation rule for pushforward measures.

Let now $f(\infty) \neq \infty$. Since f preserves cross-ratios, we get

$$1 - \frac{z - z_*}{x - z_*} = \frac{x - z}{x - z_*} = \frac{f(x) - f(z)}{f(x) - f(z_*)} \frac{f(z_*) - f(\infty)}{f(z) - f(\infty)}$$
$$= \left(1 - \frac{f(z) - f(z_*)}{f(\infty) - f(z_*)}\right)^{-1} \left(1 - \frac{f(z) - f(z_*)}{f(x) - f(z_*)}\right).$$

Noting that $f_* \delta_\infty = \delta_{f(\infty)}$, again the claim follows by applying the transformation rule for pushforward measures. □

Lemma 3.11 *The measures ν_n are a precompact family with respect to weak convergence on $C(\overline{\mathbb{R}})$. Any accumulation point $\nu = \lim_{\ell\to\infty} \nu_{n_\ell}$ is a probability measure and* $\operatorname{supp}\nu \subset \mathsf{E}$.

Proof Since $\deg D_n^0 = n$, precompactness follows by the Banach-Alaoglu theorem and any accumulation point is a probability measure on $\overline{\mathbb{R}}$. Let $(\mathbf{a}, \mathbf{b})$ be a connected component of $\overline{\mathbb{R}} \setminus \mathsf{E}$. Let us prove that $\nu((\mathbf{a}, \mathbf{b})) = 0$. By Möbius invariance, it suffices to assume that $(\mathbf{a}, \mathbf{b})$ is a bounded subset of $\mathbb{R}$. Due to Theorem 2.3 (vi), there is at most one generalized zero in $(\mathbf{a}, \mathbf{b})$, thus $\nu_{n_\ell}((\mathbf{a}, \mathbf{b})) \leq \frac{1}{n_\ell}$ and by the Portmanteau theorem $\nu((\mathbf{a}, \mathbf{b})) = 0$ and $\operatorname{supp}\nu \subset \mathsf{E}$. □

In the following we will need statements also for a subsequence $(h_{n_\ell})_{\ell=1}^\infty$. Therefore, for a fixed subsequence let us define

$$K' = \overline{\bigcup_{\ell\geq 1} \operatorname{supp} D_{n_\ell}^\infty}, \quad \text{and} \quad \Omega_{K'} = \overline{\mathbb{C}} \setminus K', \tag{3.16}$$

so that h_{n_ℓ} is subharmonic on $\Omega_{K'}$ for all ℓ. Since $\lim_{n\to\infty} \mu_n = \mu$, we have for any subsequence (and therefore any K'), that $\operatorname{supp}\mu \subset K' \subset K_{\mathbf{C}}$ and therefore $\Omega_{\mathbf{C}} \subset \Omega_{K'} \subset \overline{\mathbb{C}} \setminus \operatorname{supp}\mu$.

If $D_n^0(z_*^1) = D_n^\infty(z_*^2) = 0$, then by factoring F_n we see that there is $\beta_n \in \mathbb{R}$ so that

$$h_n(z) = \beta_n + \Phi_{\nu_n}(z, z_*^1) - \Phi_{\mu_n}(z, z_*^2). \tag{3.17}$$

Theorem 3.12 *Assume Hypothesis 1.9 and pass to a subsequence so that* $\lim_\ell \nu_{n_\ell} = \nu$, $\lim_\ell x_{n_\ell}^* = x_\infty$ *and* $\lim_\ell \beta_{n_\ell} = \beta \in \mathbb{R} \cup \{-\infty, +\infty\}$. *Then, in fact $\beta \in \mathbb{R}$ and for $z_* \notin K_{\mathbf{C}}$ we have uniformly on compact subsets of $\mathbb{C} \setminus \mathbb{R}$*

$$\lim_{\ell\to\infty} h_{n_\ell}(z) = \beta + \Phi_\nu(z, x_\infty) - \Phi_\mu(z, z_*) =: h(z). \tag{3.18}$$

In particular, h extends to a positive superharmonic function on $\overline{\mathbb{C}} \setminus \mathsf{E}$ and to a subharmonic function on $\overline{\mathbb{C}} \setminus \operatorname{supp}\mu$. Moreover, for q.e. every $z \in \Omega_{K'}$

$$\limsup_{\ell\to\infty} h_{n_\ell}(z) = \beta + \Phi_\nu(z, x_\infty^*) - \Phi_\mu(z, z_*).$$

Proof Let $(\mathbf{a}, \mathbf{b})$ denote the gap containing x_∞ and let us assume that ℓ is big enough so that all $x^*_{n_\ell}$ are in $(\mathbf{a}, \mathbf{b})$. Due to Theorem 2.3(vi), $\nu_{n_\ell}((\mathbf{a}, \mathbf{b})) = 0$. Thus, we can write

$$h_{n_\ell}(z) = \beta_{n_\ell} + \Phi_{\nu_{n_\ell}}(z, x^*_\infty) - \Phi_{\mu_{n_\ell}}(z, z_*). \tag{3.19}$$

Since $K(\cdot, z, x_\infty)$ is continuous on $\operatorname{supp}\nu_{n_\ell} \subset \overline{\mathbb{R}} \backslash (\mathbf{a}, \mathbf{b})$ and $K(\cdot, z, z_*)$ is continuous on $\overline{\mathbb{R}} \backslash K'$, we get

$$\lim_{\ell\to\infty} \Phi_{\nu_{n_\ell}}(z, x^*_\infty) = \Phi_\nu(z, x^*_\infty), \quad \lim_{\ell\to\infty} \Phi_{\mu_{n_\ell}}(z, x^*_\infty) = \Phi_\mu(z, x^*_\infty).$$

Since, for $z_0 \in \mathbb{C}_+$, $\Phi_\nu(z_0, x^*_\infty)$, $\Phi_\mu(z_0, z_*) \in \mathbb{R}$ the upper and lower estimates (3.2) and (3.15) imply that $\beta \in \mathbb{R}$. In fact, convergence is uniform on compact subsets of $\mathbb{C} \setminus \mathbb{R}$: since $\operatorname{supp}(\nu_{n_\ell})$, $\operatorname{supp}(\mu_{n_\ell}) \subset \overline{\mathbb{R}}$ for all ℓ and all measures are normalized, the estimate

$$\log \left| \frac{x - z_1}{x - z_2} \right| \le \log \left(1 + \frac{|z_1 - z_2|}{\operatorname{dist}(z_2, \overline{\mathbb{R}})} \right) \le \frac{|z_1 - z_2|}{\operatorname{dist}(z_2, \overline{\mathbb{R}})}, \qquad z_1, z_2 \in \mathbb{C} \setminus \mathbb{R}$$

implies uniform equicontinuity of the potentials $\int \log \left| 1 - \frac{z - x^*_\infty}{x - x^*_\infty} \right| d\nu_{n_\ell}(x)$ and $\int \log \left| 1 - \frac{z - z_*}{x - z_*} \right| d\mu_{n_\ell}(x)$ on compact subsets of $\mathbb{C} \setminus \mathbb{R}$, and the Arzelà–Ascoli theorem implies uniform convergence on compacts.

By applying a conformal map $f \in \mathrm{PSL}(2, \mathbb{R})$ and Lemma 2.2 we assume that $\infty \in (\mathbf{a}, \mathbf{b})$ so that E and K' are compact subsets of $\mathbb{R}$.

We note that Φ_ρ, for $\rho = \mu, \nu$, are subharmonic in $\mathbb{C}$ and harmonic in $\mathbb{C} \setminus \operatorname{supp}\rho$. Thus, we only need to argue why h is harmonic at ∞. Since $\operatorname{supp}\mu$ and E are bounded and μ, ν are probability measures, we have

$$\Phi_\rho(z) = \log|z| + O(1) \tag{3.20}$$

as $z \to \infty$ and therefore, $h(z) = O(1)$ there and h has a harmonic extension to ∞.

Finally, for $z \in \Omega_{K'} \backslash \{\infty\}$, $K(\cdot, z, z_*)$ is continuous on K' and thus

$$\lim_{\ell\to\infty} \Phi_{\mu_{n_\ell}}(z, z_*) = \Phi_\mu(z, z_*). \tag{3.21}$$

By the upper envelope theorem for q.e. $z \in \mathbb{C}$

$$\limsup_{\ell\to\infty} \Phi_{\nu_{n_\ell}}(z, x^*_\infty) = \Phi_\nu(z, x^*_\infty). \tag{3.22}$$

Combining (3.21) and (3.22), for q.e. $z \in \Omega_{K'}$ we have

$$\limsup_{\ell\to\infty} h_{n_\ell}(z) = \lim_{\ell\to\infty} \beta_{n_\ell} + \limsup_{\ell\to\infty} \Phi_{\nu_{n_\ell}}(z, x^*_\infty) - \lim_{\ell\to\infty} \Phi_{\mu_{n_\ell}}(z, z_*) = h(z).$$

□

Lemma 3.13 *Assume Hypothesis 1.9. Then, for $z \in \mathbb{C} \setminus \mathbb{R}$, we have*

$$\liminf_{n\to\infty} \frac{1}{n} \log |F_n(z)| \geq \int G_{\mathsf{E}}(z, x) d\mu(x).$$

Proof By applying a conformal map f, we assume $\infty \in \mathsf{E}$ so that $\Omega \subset \mathbb{C}$. Fix $z \in \mathbb{C} \setminus \mathbb{R}$ and let n_ℓ be such that

$$\lim_{\ell\to\infty} h_{n_\ell}(z) = \liminf_{n\to\infty} h_n(z)$$

and

$$\lim_{\ell\to\infty} h_{n_\ell}(z) = h(z) = \beta + \Phi_\nu(z, x_\infty) - \Phi_\mu(z, z_*)$$

in the sense of Theorem 3.12. Thus, h defines a positive superharmonic function on Ω and

$$-\Delta h = \Delta \Phi_\mu(z, z_*) = 2\pi \mu.$$

By the Riesz decomposition theorem [1, Theorem 4.4.1], we have

$$h(z) = \int G_{\mathsf{E}}(z, x) d\mu(x) + u(z),$$

where $u(z)$ is the greatest harmonic minorant of h. Since h is positive, it follows that $u \geq 0$. Thus,

$$h(z) \geq \int G_{\mathsf{E}}(z, x) d\mu(x)$$

and the claim follows. □

We can now prove the root asymptotics of F_n and convergence of generalized zero counting measures:

Proof of Theorem 1.10 Root asymptotics follow by combining Lemma 3.1 and Lemma 3.13.

By conformal invariance, we assume that $\infty \in \operatorname{supp}\mu$ so that $\Omega_\mu := \overline{\mathbb{C}} \setminus \operatorname{supp}\mu \subset \mathbb{C}$ and E is compact in $\mathbb{R}$. Due to Lemma 3.11 the family $\{\nu_n\}$ is precompact and we can consider a weakly convergent subsequence $\nu = \lim_{j\to\infty} \nu_{n_j}$. Moreover, by Lemma 3.2, $\int G_{\mathsf{E}}(z, x) d\mu(x)$ defines a subharmonic function in Ω_μ. Let us compute its Riesz measure. Take $\phi \in C_c^\infty(\Omega_\mu)$ and compute

$$\begin{aligned}\iint G_{\mathsf{E}}(z, x) d\mu(x) \Delta\phi(z) dA(z) &= \iint G_{\mathsf{E}}(z, x) \Delta\phi(z) dA(z) d\mu(x) \\ &= 2\pi \iint \omega_{\mathsf{E}}(dz, x) d\mu(x) \phi(z) dA(z),\end{aligned}$$

where Fubini's theorem is justified since $\mathrm{supp}(\phi) \subset \overline{\mathbb{C}}\backslash\mathrm{supp}(\mu)$, $\sup_{\mathrm{supp}(\phi)\times\mathrm{supp}(\mu)} |G_{\mathsf{E}}(z,x)| < \infty$. That is,

$$\frac{1}{2\pi}\Delta\left(\int G_{\mathsf{E}}(z,x)d\mu(x)\right) = \int \omega_{\mathsf{E}}(dz,x)d\mu(x) =: \rho.$$

Root asymptotics and Theorem 3.12 imply that on $\mathbb{C}\setminus\mathbb{R}$

$$\int G_{\mathsf{E}}(z,x)d\mu(x) = \beta + \Phi_\nu(z,x_\infty) - \Phi_\mu(z,z_*).$$

Applying the weak identity principle for subharmonic functions [19, Theorem 2.7.5], this equality also holds on Ω_μ. Thus, computing the distributional Laplacian on both sides yields $\nu = \rho$ and w-lim $\nu_n = \rho$. □

Lemma 3.14 *Assume Hypothesis 1.9. Then, fix a gap* $(\mathbf{a},\mathbf{b})$ *and let* $[u_n, v_n] = \mathsf{E}_n \cap [\mathbf{a},\mathbf{b}]$. *Passing to a subsequence such that there are limits* $u_\infty, v_\infty \in [\mathbf{a},\mathbf{b}]$, *i.e.,*

$$\lim_{\ell\to\infty} v_{n_\ell} = v_\infty, \quad \lim_{\ell\to\infty} u_{n_\ell} = u_\infty$$

we have

$$u_\infty = v_\infty.$$

Proof By conformal invariance we can assume that $\infty \notin (\mathbf{a},\mathbf{b})$ and consider again $\{h_n\}$ as a family of subharmonic functions in $\Omega_{\mathbf{C}}$. We have $\mathsf{E}_n \cap \mathrm{supp} D_n^\infty = \emptyset$, since either F_n has a pole at $\mathbf{c}$ or if F_n has a generalized zero at $\mathbf{c}$ then by (ii) of Theorem 2.9 there is no extension in this gap. Due to Theorem 1.10, $\lim_{n\to\infty} h_n = \int G_{\mathsf{E}}(\cdot,x)d\mu(x)$. Assume that $v_\infty - u_\infty = \delta > 0$. For any $0 < \varepsilon < \delta/2$, there exists ℓ_0 such that for all $\ell > \ell_0$, we have

$$A := [u_\infty + \varepsilon, v_\infty - \varepsilon] \subset [u_{n_\ell}, v_{n_\ell}]. \tag{3.23}$$

Therefore, defining K' as in (3.16), we have $A \cap K' = \emptyset$. Note that first we only have empty intersection without taking the closure, but since ε above can be made smaller, we also conclude that it holds for K'.

By Theorem 3.12 we have for q.e. $z \in \Omega_{K'}$

$$\limsup_{\ell\to\infty} h_{n_\ell}(z) = \int G_{\mathsf{E}}(z,x)d\mu(x).$$

Since $\mathrm{supp}\mu \subset K'$, it follows from Lemma 3.2 that $\int G_{\mathsf{E}}(z,x)d\mu(x) > 0$ for every $z \in \Omega_{K'}$ and therefore in particular for $z \in A$. On the other hand, by definition of E_n and (3.23), we have $h_{n_\ell}(z) \le 0$ there. Since A has positive capacity, this gives a contradiction. □

4 Szegő–Widom Asysmptotics

4.1 Asymptotics of log $|F_n|$

In the following in addition to the assumptions made in Sect. 3, we assume that E is a regular Parreau–Widom set. Let us recall its definition. First we assume that E is regular for the Dirichlet problem. Let $z_0 \in \overline{\mathbb{R}} \setminus \mathsf{E}$ and denote the gap containing z_0 by $(\mathbf{a}, \mathbf{b})$. Due to regularity and concavity of the Green function, $G_{\mathsf{E}}(z, z_0)$ has exactly one critical point in each gap $(\mathbf{a}_j, \mathbf{b}_j)$ except in the gap $(\mathbf{a}, \mathbf{b})$. Let us denote these critical points of $G_{\mathsf{E}}(z, z_0)$ by ξ_j. Then we call E a regular Parreau–Widom set, if

$$\mathcal{PW}_{\mathsf{E}}(z_0) = \sum_j G_{\mathsf{E}}(\xi_j, z_0) < \infty. \tag{4.1}$$

It is well known that this does not depend on the choice of z_0; see e.g. [15, Chapter V].

Denote the topological circle $\mathbb{T}_j = [\mathbf{a}_j, \mathbf{b}_j]/_{\mathbf{a}_j \sim \mathbf{b}_j}$. Since $\mathbf{a}_j, \mathbf{b}_j$ are Dirichlet regular points,

$$\lim_{x \downarrow \mathbf{a}_j} G_{\mathsf{E}}(z, x) = \lim_{x \uparrow \mathbf{b}_j} G_{\mathsf{E}}(z, x) = 0$$

so with the usual convention

$$G_{\mathsf{E}}(z, \mathbf{a}_j) = G_{\mathsf{E}}(z, \mathbf{b}_j) = 0, \tag{4.2}$$

the Green function $G_{\mathsf{E}}(z, t_j)$ depends continuously on $t_j \in \mathbb{T}_j$. We also consider the compact space

$$\mathcal{D}(\mathsf{E}) = \prod_{j=0}^{\infty} \mathbb{T}_j, \tag{4.3}$$

equipped with the product topology. As for divisors, a functional interpretation will be convenient. Thus, for an element $D \in \mathcal{D}(\mathsf{E})$, $D = (t_j)_{j=0}^{\infty}$, we also use the functional interpretation

$$D(x) = \sum_{j=0}^{\infty} \chi_{\{t_j\}}(x).$$

We want to associate to the divisor D_n^0 an element $D_n \in \mathcal{D}(\mathsf{E})$. In principle we want to define D_n as the restriction of D_n^0 to $\overline{\mathbb{R}} \setminus \mathsf{E}$. Recall that due to Theorem 2.3(v) and (vi), there is at most one generalized zero in each gap and no generalized zero in the gap containing x^*. Since $\deg D_n^0 = n$, almost all gaps do not contain a generalized zero. To overcome this, we define

$$D_n = (t_n^j)_{j=0}^{\infty}, \tag{4.4}$$

where $t_n^j = t$, if there is $t \in [\mathbf{a}_j, \mathbf{b}_j]$ such that $D_n^0(t) = 1$ and otherwise we define t_n^j to be the coset of $\mathbf{a}_j \sim \mathbf{b}_j$ in $\mathbb{T}_j$. Due to (4.2), these choices formally complete the definition of $D_n \in \mathcal{D}(\mathsf{E})$ without affecting certain sums below.

In the previous section we have described root asymptotics, i.e., asymptotics of $\frac{1}{n}\log|F_n(z)|$. The following theorem describes asymptotics of $\log|F_n(z)|$ and is the key to prove Szegő-Widom asymptotics in Theorem 4.5.

Theorem 4.1 *Assume Hypothesis 1.9 and let n_ℓ be such that $\lim_{\ell\to\infty} x_{n_\ell}^* = x_\infty^* \in (\mathbf{a},\mathbf{b})$ and $\lim_{\ell\to\infty} D_{n_\ell} = D$. Then for $z \in \overline{\mathbb{C}} \setminus (\mathbb{R} \setminus (\mathbf{a},\mathbf{b}))$, we have*

$$\lim_{\ell\to\infty}\left(\log|F_{n_\ell}(z)| - \sum_{\mathbf{c}} D_{n_\ell}^{\infty}(\mathbf{c}) G_{\mathsf{E}}(z,\mathbf{c})\right) = -\log 2 - \sum_t D(t) G_{\mathsf{E}}(z,t). \tag{4.5}$$

Moreover, $D(\mathbf{a}) = 1$.

Proof Define

$$H_{n_\ell}(z) = \sum_{\mathbf{c}} \Big((F_{n_\ell})_\infty(\mathbf{c}) G_{\mathsf{E}_{n_\ell}}(z,\mathbf{c}) - D_{n_\ell}^{\infty}(\mathbf{c}) G_{\mathsf{E}}(z,\mathbf{c}) \Big).$$

Due to (3.8) and Lemma 3.8 it remains to show that

$$\lim_{\ell\to\infty} H_{n_\ell}(z) = -\sum_t D(t) G_{\mathsf{E}}(z,t). \tag{4.6}$$

Let us assume without loss of generality that all $x_{n_\ell}^*$ lie in $(\mathbf{a},\mathbf{b})$. Recall that by Theorem 2.9(i) $D_{n_\ell}(\mathbf{a}) = 1$, showing that $D(\mathbf{a}) = 1$. Moreover, Theorem 2.9(i) implies $\mathsf{E}_{n_\ell} \cap (\mathbf{a},\mathbf{b}) = \emptyset$ and since $G_{\mathsf{E}}(z,\mathbf{c}) - G_{\mathsf{E}_n}(z,\mathbf{c}) \geq 0$ and $(F_n)_\infty = D_n^\infty$ on $(\mathbf{a},\mathbf{b})$, we conclude that $(-H_{n_\ell})_\ell$ defines a family of positive harmonic functions in $\overline{\mathbb{C}} \setminus (\mathbb{R} \setminus (\mathbf{a},\mathbf{b}))$ and is thus by the Harnack principle precompact in the space of positive harmonic functions together with the function which is identically $+\infty$ equipped with uniform convergence on compact subsets.

Let us now turn to the other gaps. Let $[u_n^j, v_n^j]$ denote the extension in the jth gap of the set E_n as in Theorem 2.9 and consider

$$G_{\mathsf{E}_{n_\ell}}(z,\mathbf{c}) - G_{\mathsf{E}}(z,\mathbf{c})$$

as a subharmonic function in $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$, which vanishes on E. Thus, by Lemma 3.6

$$G_{\mathsf{E}_{n_\ell}}(z,\mathbf{c}) - G_{\mathsf{E}}(z,\mathbf{c}) = -\sum_j \int_{u_{n_\ell}^j}^{v_{n_\ell}^j} G_{\mathsf{E}}(z,x)\,\omega_{\mathsf{E}_{n_\ell}}(\mathrm{d}x,\mathbf{c}),$$

Let us define

$$\omega_{n_\ell}(dx) = \sum_{\mathbf{c}} (F_{n_\ell})_\infty(\mathbf{c})\omega_{\mathsf{E}_{n_\ell}}(\mathrm{d}x, \mathbf{c}),$$

and recall that this is just a finite sum. We conclude that

$$\begin{aligned} H_{n_\ell}(z) &= \sum_{\mathbf{c}} (F_{n_\ell})_\infty(\mathbf{c}) \left(G_{\mathsf{E}_{n_\ell}}(z, \mathbf{c}) - G_{\mathsf{E}}(z, \mathbf{c}) \right) \\ &\quad - \sum_{\mathbf{c}} \left((D_{n_\ell}^\infty(\mathbf{c}) - (F_{n_\ell})_\infty(\mathbf{c})) G_{\mathsf{E}}(z, \mathbf{c}) \right) \\ &= - \sum_{j=0}^{\infty} \int_{u_{n_\ell}^j}^{v_{n_\ell}^j} G_{\mathsf{E}}(z, x)\omega_{n_\ell}(dx) - \sum_{\mathbf{c}} \left((D_{n_\ell}^\infty(\mathbf{c}) - (F_{n_\ell})_\infty(\mathbf{c})) G_{\mathsf{E}}(z, \mathbf{c}) \right). \end{aligned} \tag{4.7}$$

Due to Lemma 3.3 there are finitely many gaps containing poles. So by partitioning into finitely many subsequences, we can assume that for each j, for all $\ell > 0$ either $D_{n_\ell}^\infty(t_{n_\ell}^j) > 0$ or $D_{n_\ell}^\infty(t_{n_\ell}^j) = 0$, i.e., in the first case $t_{n_\ell}^j$ corresponds to a pole reduction of F_{n_ℓ}. We will show that both cases lead to the same limit.

Let us first consider a gap $(\mathbf{a}_j, \mathbf{b}_j)$ so that $D_{n_\ell}^\infty(t_{n_\ell}^j) = 0$ and let us assume that $t_{n_\ell}^j \to t_\infty^j \in (\mathbf{a}_j, \mathbf{b}_j)$. Due to Lemma 3.14,

$$\lim_\ell u_{n_\ell}^j = \lim_\ell v_{n_\ell}^j = t_\infty^j. \tag{4.8}$$

In particular for ℓ big enough we have $[u_{n_\ell}^j, v_{n_\ell}^j] \subset (\mathbf{a}_j, \mathbf{b}_j)$ and it follows then from Lemma 2.11 that

$$\omega_{n_\ell}([u_{n_\ell}^j, v_{n_\ell}^j]) = 1.$$

Hence,

$$\omega_{n_\ell}|[u_{n_\ell}^j, v_{n_\ell}^j] \to \delta_{t_\infty^j}$$

and therefore

$$\int_{u_{n_\ell}^j}^{v_{n_\ell}^j} G_{\mathsf{E}}(z, x)\omega_{n_\ell}(dx) \to G_{\mathsf{E}}(z, t_\infty^j).$$

If $t_\infty^j = \mathbf{a}_j$, using that $G_{\mathsf{E}}(z, \cdot)$ vanishes at $\mathbf{a}_j$ we conclude as above that

$$\int_{u_{n_\ell}^j}^{v_{n_\ell}^j} G_{\mathsf{E}}(z, x)\omega_{n_\ell}(dx) \to 0 = G_{\mathsf{E}}(z, t_\infty^j).$$

It remains to discuss the gaps where $D_{n_\ell}^\infty(t_{n_\ell}^j) = 1$. Due to Theorem 2.9(ii), $u_{n_\ell}^j = v_{n_\ell}^j = \mathbf{a}_j$, but in this case

$$D_{n_\ell}^\infty(t_{n_\ell}^j) - (F_{n_\ell})_\infty(t_{n_\ell}^j) = 1.$$

Thus, these are exactly the terms that contribute in the second sum in (4.7). Since G_E is continuous we conclude that $G_\mathsf{E}(z, t_{n_\ell}^j) \to G_\mathsf{E}(z, t_\infty^j)$. Hence, if we are allowed to interchange the limit and summation in (4.7), we have proved (4.6). As in [15, Chapter V], by a Harnack-type argument, the Parreau–Widom condition implies $\sum_j \sup_{x\in(\mathbf{a}_j,\mathbf{b}_j)} G_\mathsf{E}(z, x) < \infty$ and since moreover $\omega_{n_\ell}((\mathbf{a}_j, \mathbf{b}_j)) \leq 1$ interchanging the limits is justified and we are done. □

4.2 Blaschke Products, Character-automorphic Hardy Spaces and a Related H^∞ Extremal Problem

We will now pass from asymptotics of the superharmonic function $\log |F_n|$ to asymptotics of the rational function F_n. Thus, essentially in (4.5) we need to add harmonic conjugates and apply exp. Thus, the left-hand side in (4.5) will lead to complex Green functions

$$B_\mathsf{E}(z, \mathbf{c}) = e^{-(G_\mathsf{E}(z,\mathbf{c})+i\widetilde{G_\mathsf{E}(z,\mathbf{c})})},$$

as defined in (1.20). We have already mentioned that in general $B_\mathsf{E}(z, \mathbf{c})$ is a multi-valued function in Ω. Let us fix a normalization gap $(\mathbf{a}, \mathbf{b})$ and $z_0 \in (\mathbf{a}, \mathbf{b})$ and define $\mathsf{E}^j = [z_0, \mathbf{a}_j] \cap \mathsf{E}$. Let $\tilde{\gamma}_j$ be the generator of the fundamental group $\pi_1(\Omega, z_0)$, which starts at z_0 and passes through the gap $(\mathbf{a}_j, \mathbf{b}_j)$, encircling the set E^j once. If we extend $B_\mathsf{E}(z, \mathbf{c})$ analytically along γ_j, we get

$$B_\mathsf{E}(\tilde{\gamma}_j(z), \mathbf{c}) = e^{2\pi i \omega_\mathsf{E}(\mathsf{E}^j,\mathbf{c})} B_\mathsf{E}(z, \mathbf{c}). \tag{4.9}$$

When working with multi-valued functions, it is convenient to consider them as single-valued functions on the universal cover of $\Omega = \overline{\mathbb{C}} \setminus \mathsf{E}$. By means of the Koebe–Poincaré uniformization theorem, Ω is uniformized by the disk $\mathbb{D}$; that is, there exists a Fuchsian group Γ and a meromorphic function $\mathbf{z} : \mathbb{D} \to \Omega$ with the following properties:

1. $\forall z \in \Omega \;\; \exists \zeta \in \mathbb{D} : \mathbf{z}(\zeta) = z,$
2. $\mathbf{z}(\zeta_1) = \mathbf{z}(\zeta_2) \iff \exists \tilde{\gamma} \in \Gamma : \zeta_1 = \tilde{\gamma}(\zeta_2).$

We fix it by the normalization $\mathbf{z}(0) = z_0$, $\mathbf{z}'(0) > 0$. For Denjoy domains the covering map can be explicitly constructed [20, Section 4]. Moreover, there exists a Ford fundamental domain $\mathcal{F}$, so that $\mathbf{z} : \mathcal{F} \to \Omega$ is bijective. We denote by Γ^* the group of unitary characters of Γ; that is, group homomorphisms from Γ into $\mathbb{T} := \mathbb{R}/\mathbb{Z}$.

By the covering space formalism, Γ is group isomorphic to the fundamental group $\pi_1(\Omega, z_0)$. For a fixed $\zeta_1 \in \mathbb{D}$ we denote by

$$b(\zeta, \zeta_1) := \prod_{\gamma \in \Gamma} \frac{\gamma(\zeta_1)}{|\gamma(\zeta_1)|} \frac{\gamma(\zeta_1) - \zeta}{1 - \overline{\gamma(\zeta_1)}\zeta}, \tag{4.10}$$

the standard Blaschke product. Since $\mathrm{Cap}_{\mathsf{E}} > 0$, Γ is of convergent type and thus the product is indeed convergent. The functions $b(\zeta, \zeta_1)$ are character-automorphic, i.e., there exists $\chi_{z_1} \in \Gamma^*$ such that

$$b(\gamma(\zeta), \zeta_1) = e^{2\pi i \chi_{z_1}(\gamma)} b(\zeta, \zeta_1), \quad \forall \gamma \in \Gamma.$$

If $z_1 = \mathbf{z}(\zeta_1)$, then these Blaschke product are related to the Green function of Ω, by

$$-\log|b(\zeta, \zeta_1)| = G_{\mathsf{E}}(\mathbf{z}(\zeta), z_1).$$

Thus, we can regard the multi-valued functions $B_{\mathsf{E}}(z, z_1)$ as single-valued character-automorphic function on the universal cover.

Definition 4.2 Let f be analytic in $\mathbb{D}$. We call f (Γ^*-) character-automorphic with character $\alpha \in \Gamma^*$ if

$$f \circ \gamma = e^{2\pi i \alpha(\gamma)} f, \quad \forall \gamma \in \Gamma.$$

Similarly, if F is an analytic function on Ω, then we call F ($\pi_1(\Omega)^*$-) character-automorphic with character $\alpha \in \pi_1(\Omega)^*$, if

$$F \circ \tilde{\gamma} = e^{2\pi i \alpha(\tilde{\gamma})} F, \quad \forall \tilde{\gamma} \in \pi_1(\Omega).$$

Via the covering map $\mathbf{z}$, Γ^*- and $\pi_1(\Omega)^*$-character-automorphic functions are in one-to-one correspondence. The advantage is that Γ^*- character-automorphic functions on the universal cover $\mathbb{D}$ are single-valued. Therefore, we will formulate all convergence results for the corresponding single-valued lifts on $\mathbb{D}$.

Recall that $H^\infty_\Omega(\alpha)$ denotes the space of bounded analytic character-automorphic functions, F, in Ω; see (1.16). It is a fundamental result of Widom [32] that if E is a Parreau–Widom set, then $H^\infty_\Omega(\alpha) \neq \{0\}$ for every $\alpha \in \pi_1(\Omega)^*$. The Widom maximizer for x_* and character α is the unique function $W(z; \alpha, x_*)$ in the unit ball of $H^\infty_\Omega(\alpha)$ such that

$$W(x_*; \alpha, x_*) = \max\{\mathrm{Re}\, F(x_*) : F \in H^\infty_\Omega(\alpha), \|F\|_\Omega \le 1\}. \tag{4.11}$$

We are now ready to state a definition of Direct Cauchy theorem. It is usually stated as a point evaluation property for certain H^1 functions in Ω [15], and hence the name, but it can be equivalently defined by the following:

Definition 4.3 We say that the Direct Cauchy Theorem (DCT) holds in Ω, if for one and hence for all $x_* \in \Omega$, the map $\alpha \mapsto W(x_*; \alpha, x_*)$ is continuous on $\pi_1(\Omega)^*$ equipped with the topology dual to the discrete topology on Γ.

Let us for notational convenience also define $B_{\mathsf{E}}(z, z_0) \equiv 1$, if $z_0 \in \mathsf{E}$. Note that generally the harmonic conjugate is fixed up to an additive constant. So an additional normalization is required in (1.20). Since we will have varying normalizations, we will not fix it for a single function, but assume instead that for products of complex Green functions all of them are normalized to be positive at the same point. In this way, we can associate to any divisor $D \in \mathcal{D}(\mathsf{E})$ a product of complex Green functions, in other words a Blaschke product, by

$$B_{\mathsf{E}}(z, D) = B_{\mathsf{E}}(z, D, \phi) = e^{i\phi} \prod_t D(t) B_{\mathsf{E}}(z, t). \tag{4.12}$$

Note that

$$-\log |B_{\mathsf{E}}(z, D)| = \sum_t D(t) G_{\mathsf{E}}(z, t),$$

that is, these are exactly expression of the type appearing in (4.5). Moreover, the Widom condition guarantees that $B_{\mathsf{E}}(z, D)$ converges to a non-trivial function for any $D \in \mathcal{D}(\mathsf{E})$. Let us define the restriction

$$\mathcal{D}_k(\mathsf{E}) = \{D \in \mathcal{D}(\mathsf{E}) : D(\mathbf{a}_k) = 1\}.$$

For $D \in \mathcal{D}_k(\mathsf{E})$ it is natural to normalize $B_{\mathsf{E}}(z, D, \phi)$ such that $B_{\mathsf{E}}(z, D, \phi) > 0$ on $(\mathbf{a}_k, \mathbf{b}_k)$ which we fixes ϕ. To be more precise, since complex Green functions are defined locally and then extended analytically, this normalization holds only for one branch. Let us always assume that this branch corresponds to the values of the lift to $\mathbb{D}$ in the fundamental domain $\mathcal{F}$.

The Abel map is an important object in the spectral theory of self adjoint difference and differential operators. It is a map π from Divisors $\mathcal{D}_k(\mathsf{E})$ to the characters $\pi_1(\Omega)^*$. However, there is a subtle difference between this Abel map and the Abel map which we will implicitly use for Problem 1.3. It can be seen from the definition of $\mathcal{D}(\mathsf{E})$. In spectral theory one would usually take a two-fold cover of the interval $[\mathbf{a}_j, \mathbf{b}_j]$ and identify the endpoints of the two copies of the interval, whereas in our case we only took one copy and identified $\mathbf{a}_j \sim \mathbf{b}_j$. This map π is also the reason why the DCT property is needed, because this assumption makes π a bijection which is used in the proof of the following theorem. The proof relies on the fundamental construction of the generalized Abel map from Sodin and Yuditskii [22].

Theorem 4.4 ([5, Theorem 5.1], [11, Proposition 2.3]) *Let Ω be a regular Parreau–Widom domain such that DCT holds. Let $D \in \mathcal{D}_k(\mathsf{E})$ and let α be the character of $B_{\mathsf{E}}(z, D)$ defined by (4.12). Then, for $x_* \in (\mathbf{a}_k, \mathbf{b}_k)$ we have*

$$W(z; \alpha, x_*) = B_{\mathsf{E}}(z, D).$$

We see again that as for F_n, the extremal function only depends on the chosen gap $(\mathbf{a}_k, \mathbf{b}_k)$ and not the particular extremal point in the gap. Since the above theorem holds for any gap and arbitrary Blaschke products associated to divisors in $\mathcal{D}_k(\mathsf{E})$, we conclude that if $D \in \mathcal{D}_j(\mathsf{E})$ for $j \neq k$, then up to a unimodular constant $B_{\mathsf{E}}(z, D)$ is also the Widom maximizer for the gap $(\mathbf{a}_j, \mathbf{b}_j)$. This is in line with the Corollary 1.8 for F_n.

Let D_n^∞, x_n^* and d_n be as in Problem 1.3 and define

$$B_{\mathsf{E}}^{(n)}(z) = e^{i\phi_n} \prod_{\mathbf{c}} D_n^\infty(\mathbf{c}) B_{\mathsf{E}}(z, \mathbf{c}),$$

where $e^{i\phi_n}$ is chosen such that

$$\lim_{x \to x_n^*} B_{\mathsf{E}}^{(n)}(x) r(x, x_n^*)^{d_n} > 0. \tag{4.13}$$

Let χ_n denote the character of $B_{\mathsf{E}}^{(n)}$. Let further

$$W_n(z) = W(z; \chi_n, x_n^*),$$

denote the Widom maximizer for the point x_n^* and character χ_n.

For the following we follow the spirit of [5] and state convergence results on the universal cover $\mathbb{D}$ without introducing the corresponding lift of multi-valued functions on Ω. To give an example: if Q_n are $\pi_1(\Omega)^*$-character-automorphic function on Ω, we will write $Q_n \to Q$ uniformly on compact subsets of $\mathbb{D}$, meaning that there are lifts q_n of the Q_n which are Γ^*-character-automorphic functions such that $q_n \to q$ uniformly on compact subsets of $\mathbb{D}$ and Q is the projection of q.

Theorem 4.5 *Let* E *be a regular Parreau–Widom set, such that DCT holds in* Ω *and* F_n *be the extremizer of* (1.3)*. Then, assuming Hypothesis 1.9, uniformly on compact subsets of* $\mathbb{D}$*, we have*

$$\lim_{n \to \infty} \left(B_{\mathsf{E}}^{(n)}(z) F_n(z) - \frac{1}{2} W_n(z) \right) = 0. \tag{4.14}$$

We will use the following simple criterion based on normality; note that it is simpler than the corresponding criterion used in the polynomial case [5, Proposition 4.2], since our approach avoids working on multivalued functions on varying domains:

Proposition 4.6 *Let* $\{q_n\}_{n=1}^\infty$ *be a normal family on* $\mathbb{D}$*. Let* q_∞ *be analytic on* $\mathbb{D}$ *so that for some* $\zeta_0 \in \mathbb{D}$ *and some neighborhood,* V*, of* ζ_0 *we have that*

$$\lim_{n \to \infty} |q_n(\zeta)| = |q_\infty(\zeta)| \quad \textit{for all } \zeta \in V; \tag{4.15}$$

$$q_n(\zeta_0) > 0, \quad q_\infty(\zeta_0) > 0. \tag{4.16}$$

Then $q_n \to q_\infty$ *uniformly on compact subsets of* $\mathbb{D}$*.*

Proof By normality, it suffices to prove that any subsequence $(q_{n_\ell})_{\ell=1}^{\infty}$ which converges uniformly on compacts has the limit q_∞. Denote by f the limit of such a sequence. By (4.15), $|f(\zeta)| = |q_\infty(\zeta)|$ for all $\zeta \in V$. By (4.16), by possibly decreasing V, we can assume $q_\infty(\zeta) \neq 0$ for $\zeta \in V$, so by the maximum principle applied to f/q_∞, we conclude $f = e^{i\phi} q_\infty$ for some unimodular constant $e^{i\phi}$. By (4.16), $f(\zeta_0) \geq 0$ and $q_\infty(\zeta_0) > 0$, so $e^{i\phi} = 1$ and $f = q_\infty$. □

Defining

$$Q_n(z) = F_n(z) B_{\mathsf{E}}^{(n)}(z),$$

the strategy is now clear: First we need to check that $Q_n(z)$ defines a normal family. Realizing that $\log |Q_n(z)|$ is exactly the left hand-side in (4.5), Theorem 4.1 and Proposition 4.6 imply that all accumulation points are Blaschke products. Combining this with Theorem 4.4 finishes the proof of Theorem 4.5.

Lemma 4.7 *The sequence* $\{Q_n\}_{n=1}^{\infty}$ *forms a normal family in* $\mathbb{D}$.

Proof Since on Ω

$$(F_n)_\infty \leq D_n^\infty = (B_{\mathsf{E}}^{(n)})_0,$$

Q_n are analytic $\pi_1(\Omega)^*$-character automorphic functions in Ω. They have therefore Γ^*-character automorphic lifts to $\mathbb{D}$. By Montel's theorem [26, Chapter 6], it suffices to show that $|F_n B_{\mathsf{E}}^{(n)}| \leq 1$ in Ω. The functions $\log |Q_n|$ are subharmonic in Ω. Moreover, since E is regular and $|F_n| \leq 1$ on E, for every $\zeta \in \mathsf{E}$

$$\limsup_{z\to\zeta} \log |Q_n(z)| = \limsup_{z\to\zeta} \log |F_n(z)| - \lim_{z\to\zeta} \sum_t D_n^\infty(t) \lim_{z\to\zeta} G_{\mathsf{E}}(\zeta, t) \leq 0,$$

where we used Dirichlet regularity and the fact that the sum is only finite. The claim follows by the maximum principle for subharmonic functions [19, Theorem 2.3.1]. □

Lemma 4.8 *Suppose Hypothesis 1.9 holds and let n_ℓ be a subsequence such that $\lim_{\ell\to\infty} D_{n_\ell} = D$ and $\lim_{\ell\to\infty} x_{n_\ell}^* = x_\infty^* \in (\mathbf{a}_j, \mathbf{b}_j)$ for some $j \geq 0$. Then, uniformly on compact subsets of $\mathbb{D}$ we have*

$$\lim_{\ell\to\infty} Q_{n_\ell}(z) = \frac{1}{2} B_{\mathsf{E}}(z, D),$$

where $D \in \mathcal{D}_j(\mathsf{E})$ and $B_{\mathsf{E}}(x_\infty^*, D) > 0$.

Proof Let us assume without loss of generality that all $x_{n_\ell}^*$ lie in $(\mathbf{a}_j, \mathbf{b}_j)$. By (4.13), we have

$$Q_{n_\ell}(x_{n_\ell}^*) > 0.$$

Moreover, Q_{n_ℓ} are real, i.e., $Q_{n_\ell}(\overline{z}) = \overline{Q_{n_\ell}(z)}$. Since $D_n^0(t) = 0$ for every $t \in (\mathbf{a}_j, \mathbf{b}_j)$, it follows that $Q_{n_\ell}(t) > 0$. Thus, in particular at x_∞^*. Thus we can apply Proposition 4.6 in a vicinity of x_∞^* and then the claim follows from Theorem 4.1. □

Proof of Theorem 4.5 Since Q_n, W_n form normal families, by precompactness it suffices to prove that every subsequence has a subsubsequence so that $\lim_\ell Q_{n_\ell} - W_{n_\ell} = 0$. Let us pass to a subsequence such that $\lim_{\ell\to\infty} D_{n_\ell} = D$ and $\lim_{\ell\to\infty} x_{n_\ell}^* = x_\infty^* \in (\mathbf{a}_j, \mathbf{b}_j)$ as in Lemma 4.8. Then by Lemma 4.8

$$\lim_{\ell\to\infty} Q_{n_\ell}(z) = \frac{1}{2} B_{\mathsf{E}}(z, D).$$

If α is the character of $B_{\mathsf{E}}(z, D)$ this implies that $\chi_{n_\ell} \to \alpha$. By Theorem 4.4, $B_{\mathsf{E}}(z, D) = W(z, \alpha, x_\infty^*)$. On the other hand, it is proven in [5, Theorem 3.1] that DCT implies that $W(z; \chi_{n_\ell}, x_{n_\ell}^*) \to W(z, \alpha, x_\infty^*)$ uniformly on compact subsets of $\mathbb{D}$. In this reference there is no sequence of extremal points, but since the Widom maximizer only depends on the given gap and not on the particular point, the sequence $W(z; \chi_{n_\ell}, x_{n_\ell}^*)$ eventually only depends on the character. This concludes the proof.

□

References

1. Armitage, D. H., Gardiner, S. J.: Classical potential theory, Springer Monographs in Mathematics, Springer Verlag, London, (2001). MR 1801253
2. Azarin, V.: Growth theory of subharmonic functions, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, (2009). MR 2463743
3. Chebyshev, P.: Théorie des mécanismes connus sous le nom de parallélogrammes. Mémoires présentés à l'Académie Impériale des Sciences de Saint-Pétersbourg **7**, 539–586 (1854)
4. Chebyshev, P.: Sur les questions de minima qui se rattachent à la représentation approximative des fonctions. Mémoires présentés à l'Académie Impériale des Sciences de Saint-Pétersbourg, Sixiéme serie **7**, 199–291 (1859)
5. Christiansen, J.S., Simon, B., Yuditskii, P., Zinchenko, M.: Asymptotics of Chebyshev polynomials, II: DCT subsets of $\mathbb{R}$. Duke Math. J. **168**(2), 325–349 (2019) MR 3909898
6. Christiansen, J.S., Simon, B., Zinchenko, M.: Asymptotics of Chebyshev polynomials, I: subsets of $\mathbb{R}$. Invent. Math. **208**(1), 217–245 (2017) MR 3621835
7. Christiansen, J. S., Simon, B., Zinchenko, M.: Asymptotics of Chebyshev Polynomials, V. Residual Polynomials, arXiv:2008.09669 (2020)
8. Eichinger, B.: Szegő-Widom asymptotics of Chebyshev polynomials on circular arcs. J. Approx. Theory **217**, 15–25 (2017) MR 3628947
9. Eichinger, B., Lukić, M.: Stahl–Totik regularity for continuum Schrödinger operators, arXiv:2001.00875 (2020)
10. Eichinger, B., Lukić, M., Young, G.: Orthogonal rational functions with real poles, root asymptotics, and GMP matrices, arXiv:2008.11884 (2020)
11. Eichinger, B., Yuditskii, P.: The Ahlfors problem for polynomials. Mat. Sb. **209**(3), 34–66 (2018) MR 3769214
12. Eremenko, A., Yuditskii, P.: Comb functions, Recent advances in orthogonal polynomials, special functions, and their applications, Contemp. Math., vol. 578, Amer. Math. Soc., Providence, RI, pp. 99–118. (2012)MR 2964141
13. Faber, G.: Über Tschebyscheffsche Polynome. J. Reine Angew. Math. **150**, 79–106 (1920) MR 1580974
14. Garnett, J. B., Marshall, D. E.: Harmonic measure, New Mathematical Monographs, vol. 2, Cambridge University Press, Cambridge, (2005). MR 2150803

15. Hasumi, M.: Hardy classes on infinitely connected Riemann surfaces. Lecture Notes in Mathematics, vol. 1027. Springer-Verlag, Berlin (1983)
16. Koosis, P.: The logarithmic integral. I, Cambridge Studies in Advanced Mathematics, vol. 12, Cambridge University Press, Cambridge, (1998). MR 1670244
17. Lukashov, A.L.: On Chebyshev-Markov rational functions over several intervals. J. Approx. Theory **95**(3), 233–352 (1998) MR 1657679
18. Marčenko, V. A., Ostrovskii, I. V.: A characterization of the spectrum of the Hill operator, Mat. Sb. (N.S.) **97** (139) (1975), no. 4(8), 540–606, 633–634. MR 0409965
19. Ransford, T.: Potential theory in the complex plane, London Mathematical Society Student Texts, vol. 28, Cambridge University Press, Cambridge, (1995). MR 1334766
20. Rubel, L.A., Ryff, J.V.: The bounded weak-star topology and the bounded analytic functions. J. Functional Analysis **5**, 167–183 (1970) MR 0254580
21. Saff, E. B., Totik, V.: Logarithmic potentials with external fields, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 316, Springer-Verlag, Berlin, Appendix B by Thomas Bloom. (1997) MR 1485778
22. 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**(3), 387–435 (1997) MR 1674798
23. Schiefermayr, K.: The growth of polynomials outside of a compact set–the Bernstein-Walsh inequality revisited. J. Approx. Theory **223**, 9–18 (2017) MR 3707135
24. Schlag, W.: A course in complex analysis and Riemann surfaces, Graduate Studies in Mathematics, vol. 154, American Mathematical Society, Providence, RI, (2014). MR 3186310
25. Simon, B.: Equilibrium measures and capacities in spectral theory. Inverse Probl. Imaging **1**(4), 713–772 (2007) MR 2350223
26. Simon, B.: Basic complex analysis, A Comprehensive Course in Analysis, Part 2A, American Mathematical Society, Providence, RI, (2015). MR 3443339
27. Sodin, M.L., Yuditskii, P.M.: Functions deviating least from zero on closed subsets of the real axis. St. Petersburg Math. J **4**, 201–249 (1993)
28. Stahl, H., Totik, V.: General orthogonal polynomials, Encyclopedia of Mathematics and its Applications, vol. 43, Cambridge University Press, Cambridge, (1992). MR 1163828
29. Thiran, J.-P., Detaille, C.: Chebyshev polynomials on circular arcs in the complex plane, Progress in approximation theory, Academic Press, Boston, MA, pp. 771–786. (1991) MR 1114813
30. Totik, V., Yuditskii, P.: On a conjecture of Widom. J. Approx. Theory **190**, 50–61 (2015) MR 3304588
31. Widom, H.: Extremal polynomials associated with a system of curves in the complex plane. Advances in Math. **3**, 127–232 (1969) MR 239059
32. Widom, H.: $\mathcal{H}_p$ sections of vector bundles over Riemann surfaces. Ann. of Math. **2**(94), 304–324 (1971) MR 288780

J. S. Christiansen *et al.* (2021) “Finite-Gap CMV Matrices: Periodic Coordinates and a Magic Formula,”
International Mathematics Research Notices, Vol. 2021, No. 18, pp. 14016–14085
Advance Access Publication February 13, 2020
https://doi.org/10.1093/imrn/rnz213

Finite-Gap CMV Matrices: Periodic Coordinates and a Magic Formula

Jacob S. Christiansen[1,*], Benjamin Eichinger[2], and Tom VandenBoom[3]

[1]Centre for Mathematical Sciences, Lund University, Box 118, 22100 Lund, Sweden, [2]Department of Mathematics, Rice University MS-136, Box 1892, Houston, TX 77251-1892, USA, and [3]Department of Mathematics, Yale University, 10 Hillhouse Ave, New Haven, CT 06511, USA

**Correspondence to be sent to: e-mail: stordal@maths.lth.se*

We prove a bijective unitary correspondence between (1) the isospectral torus of almost-periodic, absolutely continuous CMV matrices having fixed finite-gap spectrum E and (2) special periodic block-CMV matrices satisfying a Magic Formula. This latter class arises as E-dependent operator Möbius transforms of certain generating CMV matrices that are periodic up to a rotational phase; for this reason we call them “MCMV.” Such matrices are related to a choice of orthogonal rational functions on the unit circle, and their correspondence to the isospectral torus follows from a functional model in analog to that of GMP matrices. As a corollary of our construction we resolve a conjecture of Simon; namely, that Caratheodory functions associated to such CMV matrices arise as quadratic irrationalities.

1 Introduction and Main Results

This paper studies two equivalent notions of interest, one in the spectral theory of certain unitary operators and the other in the theory of analytic functions mapping the unit disk into its closure. This connection between CMV matrices and Schur functions

Received March 1, 2019; Revised June 19, 2019; Accepted July 26, 2019
Communicated by Prof. Misha Sodin

is no mystery to experts in either field. We introduce here our results from both perspectives because we find them striking and attractive, as well as not immediately equivalent.

We begin in the context of the title:

1.1 Finite-gap CMV and MCMV matrices

Fix a sequence $\{a_k\}_{k\in\mathbb{Z}} \in \mathbb{D}^{\mathbb{Z}}$ of numbers in the unit disk $\mathbb{D}$, and define a family of 2×2 unitary matrices by

$$\Theta_k := \begin{bmatrix} \overline{a_k} & \rho_k \\ \rho_k & -a_k \end{bmatrix}, \quad \rho_k = \sqrt{1-|a_k|^2}.$$

Letting Θ_k act on the span of $\{\delta_k, \delta_{k+1}\}$ in $\ell^2(\mathbb{Z})$, define

$$L = \bigoplus_{l\in\mathbb{Z}} \Theta_{2l}, \quad M = \bigoplus_{l\in\mathbb{Z}} \Theta_{2l+1}.$$

The (whole-line) CMV matrix associated to the sequence $\{a_k\}_{k\in\mathbb{Z}}$ is the unitary operator

$$C := C(\{a_k\}) = LM.$$

Represented in the basis $\{\delta_k\}_{k\in\mathbb{Z}}$, C is a five-diagonal matrix with repeating 2×4 block structure:

$$C = \begin{bmatrix} \ddots & \ddots & \ddots & & & & \\ \rho_{2l-1}\overline{a_{2l}} & -\overline{a_{2l}}a_{2l-1} & \overline{a_{2l+1}}\rho_{2l} & \rho_{2l}\rho_{2l+1} & & & \\ \rho_{2l}\rho_{2l-1} & -\rho_{2l}a_{2l-1} & -\overline{a_{2l+1}}a_{2l} & -\rho_{2l+1}a_{2l} & & & \\ & & \rho_{2l+1}\overline{a_{2l+2}} & -\overline{a_{2l+2}}a_{2l+1} & \overline{a_{2l+3}}\rho_{2l+2} & \rho_{2l+2}\rho_{2l+3} \\ & & \rho_{2l+2}\rho_{2l+1} & -\rho_{2l+2}a_{2l+1} & -\overline{a_{2l+3}}a_{2l+2} & -\rho_{2l+3}a_{2l+2} \\ & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

We will study those CMV matrices with almost-periodic coefficients $\{a_k\}_{k\in\mathbb{Z}}$ having full a.c. spectrum consisting of a fixed finite union of non-degenerate closed circular arcs $\mathsf{E} \subset \partial\mathbb{D}$. We denote the set of all such CMV matrices by $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$. Topologically, $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ is a torus of dimension equal to the number of disjoint arcs in E.

CMV matrices are natural objects of interest in the context of orthogonal polynomials on the unit circle (OPUC) (see [30, 31]). This in part relies on the interesting fact that half-line CMV matrices $\mathcal{C}_+$, formed by setting $a_{-1} = -1$ above and restricting to $\ell^2(\mathbb{N})$, are universal for cyclic unitary operators in the sense that, for any probability measure ν with infinite support on the unit circle $\partial\mathbb{D}$, multiplication by the independent variable in $L^2(\mathrm{d}\nu)$ is unitarily equivalent to some half-line CMV matrix $\mathcal{C}_+$. This discovery was made surprisingly recently by Cantero, Moral, and Velázquez [7] by considering the basis of $L^2(\mathrm{d}\nu)$ generated by orthonormalizing $\{1, z^{-1}, z, z^{-2}, z^2, \ldots\}$. In comparison, it has long been known (see, e.g., [36]) that tridiagonal Jacobi matrices are universal models for self-adjoint operators with a cyclic vector. CMV matrices are also important, for example, in the theory of random matrices and integrable systems [15, 22], and for quantum walks [6]; see also [4, 16].

The CMV basis is far from the only generating set for $L^2(\mathrm{d}\nu)$. Letting b_w be the elementary Blaschke factor for $\mathbb{D}$ vanishing at w, that is,

$$b_w(z) := \frac{z - w}{1 - \overline{w}z} \tag{1.1}$$

and denoting by $\hat{w} = 1/\overline{w}$ reflection with respect to $\partial\mathbb{D}$, one has that $b_w(\hat{z}) = \widehat{b_w(z)} = c_w\overline{b_{\hat{w}}(z)}$, $c_w = \hat{w}w$. Setting $c_0 = 1$, this allows us to extend this notion also to $w = \infty$ by $b_\infty(z) = \overline{b_0(\hat{z})} = 1/z$ and hence we can suggestively rewrite the CMV basis above as instead being generated by orthonormalizing the sequence $\{1, b_\infty, b_0, b_\infty^2, b_0^2, \ldots\}$. For a fixed sequence of points $\{z_k\}_{k\in\mathbb{N}} \in \mathbb{D}^\mathbb{N}$ with modulus bounded uniformly away from 1, if we denote by $\{B_k\}$ and $\{B_k^*\}$ the families of Blaschke products

$$B_0(z) = 1, \quad B_k(z) = \prod_{j=1}^{k} b_{z_j}(z), \quad B_k^*(z) = \overline{B_k(\hat{z})} = \prod_{j=1}^{k} b_{z_j}(z)^{-1}, \tag{1.2}$$

one could just as well have spanned $L^2(\mathrm{d}\nu)$ by the sequence $\{B_0, B_1^*, B_1, B_2^*, B_2, \ldots\}$. In [37], Velázquez showed that the structure of multiplication by the independent variable z in $L^2(\mathrm{d}\nu)$ in the orthonormalization of this new generating set is related to CMV matrices via an operator Möbius transform; specifically, denoting by $D_+ := \operatorname{diag}_\mathbb{N}\{0, z_1, z_1, z_2, z_2, \ldots\}$, he showed that multiplication by z in $L^2(\mathrm{d}\nu)$ is unitarily equivalent to the operator

$$b_{-D_+}(\mathcal{C}_+) := \eta_{D_+}\left(1 + \mathcal{C}_+ D_+^*\right)^{-1}\left(D_+ + \mathcal{C}_+\right)\eta_{D_+}^{-1}, \quad \eta_{D_+} = \sqrt{1 - D_+ D_+^*} \tag{1.3}$$

for some half-line CMV matrix C_+. This theorem suggests we should study this new class of operator Möbius transforms of CMV matrices more closely:

Definition 1.1 (MCMV matrices). Fix $n \geq 1$ and let $\vec{z} = \{z_k\}_{k=0}^{n-1} \in \mathbb{D}^n$ with $z_0 = 0$, $\{a_k\}_{k\in\mathbb{Z}} \in \mathbb{D}^{\mathbb{Z}}$, and $\vartheta \in \mathbb{R}/2\pi\mathbb{Z}$. Denote by D_0 the $2n$-periodic diagonal matrix

$$D_0 := D_0(\vec{z}) = \mathrm{diag}_{\mathbb{Z}}\{\ldots, z_{n-2}, z_{n-1}, z_{n-1}, z_0 \,|\, z_0, z_1, z_1, z_2, \ldots\}, \tag{1.4}$$

let $\Lambda_k(\vartheta)$ be the $2n \times 2n$ diagonal matrix $\Lambda_k(\vartheta) := \mathrm{diag}_{2n}\{e^{ik\vartheta}, e^{-ik\vartheta}, \ldots, e^{ik\vartheta}, e^{-ik\vartheta}\}$, and define

$$\Lambda(\vartheta) := \bigoplus_{k\in\mathbb{Z}} \Lambda_k(\vartheta). \tag{1.5}$$

With this notation, the (whole-line) MCMV matrix for $\vec{z}$, $\{a_k\}$, and ϑ is defined by

$$A := A(\{a_k\}, \vartheta; \vec{z}) = \Lambda(\vartheta)^* b_{-D_0}(C)\Lambda(\vartheta), \tag{1.6}$$

where $C = C(\{a_k\})$ is the CMV matrix associated to $\{a_k\}_{k\in\mathbb{Z}}$.

Before proceeding, we make a few remarks regarding this definition. First, we split $\ell^2(\mathbb{Z}) = \ell^2_+ \oplus \ell^2_-$ around zero, with the zeroth entry belonging to ℓ^2_+; we denote the site of this splitting with a vertical bar in (1.4). Second, the operator Möbius transform $b_{-D_0}(C)$ in (1.6) is an operator on $\ell^2(\mathbb{Z})$ defined exactly as in (1.3) with D_0 replacing D_+. Furthermore, vectors $\vec{z} = \{z_k\}_{k=0}^{n-1} \in \mathbb{D}^n$ of interpolation points will henceforth have $z_0 = 0$ unless otherwise noted; the presence of this zero gives MCMV matrices a block CMV band-structure, see (1.7) and Section 4.3. Finally, we shall explain the role of $\Lambda(\vartheta)$ momentarily and be even more specific in Section 3; in short, this diagonal matrix enables us to change from periodicity up to a rotational phase to pure periodicity.

Like CMV matrices, an MCMV matrix A is again band-structured. If we split A into $2n \times 2n$ blocks $\mathbf{A_{ij}}$, then $\mathbf{A_{ij}} = \mathbf{0}$ if $|\mathbf{i} - \mathbf{j}| > 1$. Moreover, the off-diagonal blocks are of the form $\mathbf{A_{i,i-1}} = \mathbf{v^i}\delta_{2n-1}^{\mathsf{T}}$ and $\mathbf{A_{i,i+1}} = \mathbf{u^i}\delta_0^{\mathsf{T}}$ for some explicit vectors $\mathbf{u^i}, \mathbf{v^i} \in \mathbb{C}^{2n}$; cf. Lemma 4.8 and the figure below:

$$A = \left[\begin{array}{c|c|c|c|c} \ddots & \ddots & \ddots & & \\ \mathbf{0} & \mathbf{0}\;\; \mathbf{v^i} & \mathbf{A_{ii}} & \mathbf{u^i}\;\; \mathbf{0} & \mathbf{0} \\ & & \ddots & \ddots & \ddots \end{array}\right]. \tag{1.7}$$

Furthermore, since operator Möbius transforms preserve unitarity, MCMV matrices are likewise unitary operators. Thus MCMV matrices can be viewed as being "block-CMV." This special structure does not hold for arbitrary operator Möbius transforms of CMV matrices; it follows in our case from D_0 having periodically repeated zero entries.

We denote the class of all MCMV matrices associated to $\vec{z} \in \mathbb{D}^n$ by

$$\mathbb{A}(\vec{z}) := \left\{A(\{a_k\}, \vartheta; \vec{z}) : \{a_k\} \in \mathbb{D}^{\mathbb{Z}},\ \vartheta \in \mathbb{R}/2\pi\mathbb{Z}\right\} \tag{1.8}$$

and give special consideration to the subset $\mathbb{A}_{\mathrm{per}}(\vec{z}) \subset \mathbb{A}(\vec{z})$ of periodic operators, that is,

$$\mathbb{A}_{\mathrm{per}}(\vec{z}) := \left\{A \in \mathbb{A}(\vec{z}) : S^{2n}A = AS^{2n}\right\}, \tag{1.9}$$

where, as usual, S is the right shift operator. Notice that the usual CMV matrices are simply the special case of MCMV matrices associated to the vector $\vec{z}$ where $z_k = 0$ for all k. This realization gives rise to a natural (if somewhat ill-posed) question: is there a "best" generating set of Blaschke products for a given measure ν on $\partial\mathbb{D}$? In the context of whole-line CMV matrices $\mathcal{C} \in \mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$, we offer an affirmative answer in the form of a correspondence to certain isospectral periodic MCMV matrices. To properly state this correspondence, we first need to introduce a fundamental object in the spectral analysis of periodic MCMV matrices.

Given a value $a \in \mathbb{D}$, let

$$U(a) := \frac{1}{\rho}\begin{bmatrix} 1 & a \\ \overline{a} & 1 \end{bmatrix}, \quad \rho = \sqrt{1 - |a|^2}. \tag{1.10}$$

For a fixed $A = A(\{a_k\}, \vartheta; \vec{z}) \in \mathbb{A}_{\mathrm{per}}(\vec{z})$, define the monodromy matrix T_A by

$$\begin{aligned} T_A(z) := U(a_0)\begin{bmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_1)\begin{bmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_2)\begin{bmatrix} b_{z_2}(z) & 0 \\ 0 & 1 \end{bmatrix}\cdots \\ \cdots U(a_{2n-3})\begin{bmatrix} b_{z_{n-1}}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_{2n-2})\begin{bmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_{2n-1})\begin{bmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix}. \end{aligned} \tag{1.11}$$

Letting $B(z) = z\prod_{j=1}^{n-1} b_{z_j}(z) = \sqrt{\det T_A(z)}$, we introduce the following:

Definition 1.2 (Discriminant of an MCMV matrix). Let $A = A(\{a_k\}, \vartheta; \vec{z}) \in \mathbb{A}_{\rm per}(\vec{z})$ be a periodic MCMV matrix. The discriminant of A is the rational function defined by

$$\Delta_A(z) := \frac{1}{B(z)} \operatorname{tr}\big(T_A(z)\big). \tag{1.12}$$

Remark. Denoting by j the signature matrix $j = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, we note that $T_A(z)^* j T_A(z) \le j$ for $z \in \mathbb{D}$, while $T_A(z)^* j T_A(z) = j$ when $z \in \partial\mathbb{D}$. Functions of this type are called j-inner matrix functions and equation (1.11) represents a factorization of T_A into so-called elementary Blaschke–Potapov factors of the first kind. The study of general j-contractive matrix functions and their multiplicative structure goes back to Potapov [28].

For periodic CMV matrices, the discriminant determines the spectrum. Our discriminant does the same; specifically, we will show in Section 4.2 that the key properties of Δ_A are:

(i) $\Delta_A^{-1}([-2, 2]) = \sigma(A)$,
(ii) the poles of Δ_A are supported on $\vec{z}$ and the reflected points $\{\hat{z}_j : z_j \in \vec{z}\}$,
(iii) Δ_A is real, that is, $\overline{\Delta_A(\hat{z})} = \Delta_A(z)$.

Damanik–Killip–Simon [8] showed that the discriminant of a periodic CMV matrix also completely describes solutions to the inverse problem via a "Magic Formula." To consider the inverse problem in our setting, fix a finite-gap set $\mathsf{E} \subset \partial\mathbb{D}$, that is, a finite disjoint union of non-degenerate closed circular arcs. Let us refer to the arc-components as bands and to the connected components of $\partial\mathbb{D} \setminus \mathsf{E}$ as gaps. Then we look for a rational function Δ that satisfies the above properties (i)–(iii) with $\sigma(A)$ substituted by E. Provided such a function exists, we will see in Lemma 4.4 that these properties define Δ uniquely up to a sign. While such a function need not always exist for any $\vec{z}$, we can demonstrate by example the existence for a particular vector $\vec{z}_{\mathsf{E}}$.

Consider such a finite-gap set E having $g + 1$ gaps (and respectively bands). For any point z_0 in the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$, there exists an Ahlfors function w_{z_0} that maximizes the modulus of the derivative at z_0 (or, in the case $z_0 = \infty$, maximizes $\lim_{z\to\infty} |z w_\infty(z)|$) among all analytic functions on $\overline{\mathbb{C}} \setminus \mathsf{E}$ with modulus bounded by 1; cf. [1, 12]. This extremal property defines w_{z_0} uniquely up to a unimodular multiplier and, moreover, $w_{z_0}(z_0) = 0$. In the right normalization, these Ahlfors functions for $\mathsf{E} \subset \partial\mathbb{D}$ have the symmetry property

$$w_{\hat{z}_0}(\hat{z}) = \overline{w_{z_0}(z)}; \tag{1.13}$$

in particular, the zeros of w_∞ can be obtained by reflecting the zeros of w_0 with respect to $\partial\mathbb{D}$. In terms of these functions, we can define a special function, which we call the generalized discriminant, related to the set E:

Definition 1.3 (Generalized discriminant). For a finite union of non-degenerate closed circular arcs $\mathsf{E} \subset \partial\mathbb{D}$, the generalized discriminant is defined by

$$\Delta_{\mathsf{E}} := \frac{1}{w_0 w_\infty} + w_0 w_\infty. \tag{1.14}$$

By (1.13), we see that Δ_{E} is real-valued on $\partial\mathbb{D}$; since $|w_{z_0}(z)| = 1$ for $z \in \mathsf{E}$ in the sense of nontangential limits and $|w_{z_0}(z)| < 1$ for $z \in \overline{\mathbb{C}} \setminus \mathsf{E}$, it follows that

$$\mathsf{E} = \Delta_{\mathsf{E}}^{-1}([-2, 2]). \tag{1.15}$$

The function Δ_{E} has $2(g+1)$ poles, half of which lie inside the unit disk. Moreover, there is exactly one critical point (i.e., a zero of Δ_{E}') in each band of E and in each gap of E. While Δ_{E} maps all critical points in bands to -2, the critical points in gaps have Δ_{E}-value strictly greater than 2. For more details on the Ahlfors function and the discriminant, we refer to Appendix A (where in particular these properties are proven). As will be crucial for our analysis, we define $\vec{z}_{\mathsf{E}}$ to be some fixed ordering of the poles of Δ_{E} inside $\mathbb{D}$, that is,

$$\vec{z}_{\mathsf{E}} := \{z_0 = 0, z_1, \ldots, z_g\} \in \mathbb{D}^{g+1}, \quad z_k \in \mathbb{D} \text{ a pole of } \Delta_{\mathsf{E}}. \tag{1.16}$$

With these definitions in hand, we can now clarify precisely the relationship between almost-periodic finite-gap CMV matrices $\mathcal{C} \in \mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ and periodic MCMV matrices. For our finite-gap set $\mathsf{E} \subset \partial\mathbb{D}$, let $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$ and fix $\vec{z} \in \mathbb{D}^n$ with $z_0 = 0$. Then we define

$$\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*) := \{A \in \mathbb{A}_{\mathrm{per}}(\vec{z}) : \sigma(A) = \mathsf{E}, \Delta_A(\lambda_*) > 0\}. \tag{1.17}$$

For an independent choice of these parameters it might happen that $\mathcal{T}_{\mathrm{MCMV}}$ as defined in (1.17) is empty; this could for instance be illustrated by the fact that not every finite-gap set E is the spectrum of a periodic CMV matrix, which correspond to the choice $z_k = 0$ for all k. However, for the special choice of E and $\vec{z}_{\mathsf{E}}$ we in fact have a correspondence between $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*)$ and $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$:

Theorem 1.4 (Periodic Coordinates for finite-gap CMV matrices). Let $\mathsf{E} \subset \partial\mathbb{D}$ be a disjoint union of $g+1$ non-degenerate closed circular arcs, and let $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$. For the sequence $\vec{z}_{\mathsf{E}} := \{z_k\}_{k=0}^{g} \in \mathbb{D}^{g+1}$ of points (1.16) depending only on E, there is a unitary bijection between $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ and $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*)$; that is,

$$\mathcal{T}_{\mathrm{CMV}}(\mathsf{E}) \simeq \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*). \tag{1.18}$$

In particular, for an almost-periodic CMV matrix $\mathcal{C}$ with absolutely continuous spectrum E, there exists an associated CMV matrix $C = C(\{a_k\})$ with phase-periodic coefficients

$$a_{k+2(g+1)} = e^{-2i\vartheta} a_k, \quad k \in \mathbb{Z} \tag{1.19}$$

such that $\mathcal{C}$ is unitarily equivalent to the periodic MCMV matrix $A(\{a_k\}, \vartheta; \vec{z}_{\mathsf{E}}) \in \mathbb{A}_{\mathrm{per}}(\vec{z}_{\mathsf{E}})$ and the spectral measures of the one-sided restrictions $\mathcal{C}_+$ and A_+ coincide; cf. (4.5).

Remarks.

- (i) The above theorem shows that a periodic MCMV matrix is naturally related to two different CMV matrices: the almost-periodic CMV matrix in $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ and the "generating" phase-periodic CMV matrix. Throughout, we will denote the former by $\mathcal{C}$ with parameters $\{\mathbf{a}_k\}$ (resp. $\{\boldsymbol{\rho}_k\}$) and the latter by C with parameters $\{a_k\}$ (resp. $\{\rho_k\}$).
- (ii) As a consequence of (1.19), the operator $b_{-D_0}(C)$ is periodic up to a phase. By conjugating it with $\Lambda(\vartheta)$—and this is the main purpose of introducing such a diagonal matrix—we get that $A(\{a_k\}, \vartheta; \vec{z}_{\mathsf{E}})$ becomes periodic in the standard sense. This is particularly important in view of Theorem 1.5 below, since by Naiman's lemma [21] an operator satisfying the right-hand side of (1.20) is necessarily periodic.
- (iii) Notice that (1.19) is invariant under the substitution $\vartheta \mapsto \vartheta + \pi$. Consequently, there are in fact two MCMV matrices, $A(\{a_k\}, \vartheta; \vec{z})$ and $A(\{a_k\}, \vartheta + \pi; \vec{z})$, in $\mathbb{A}_{\mathrm{per}}(\vec{z}_{\mathsf{E}})$ with spectrum E having unitary equivalent one-sided restrictions to $\mathcal{C}_+$. However, notice that this same map sends $e^{i\vartheta}$ to $-e^{i\vartheta}$, which, by (1.11), changes the sign of the discriminant $\Delta_A \mapsto -\Delta_A$. For this reason, the normalization $\Delta_A(\lambda_*) > 0$ fixes uniquely one such MCMV matrix.

We stress that, by definition, the normalization $\Delta_{\mathsf{E}}(\lambda_*) > 0$ holds for any $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$, see (1.13), (1.14), and Appendix A; consequently, the function Δ_{E} has the key properties (i)–(iii) of a discriminant of an MCMV matrix in $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*)$. The following main result, dubbed a "Magic Formula" in deference to the analogous result of Damanik–Killip–Simon, shows that this function indeed characterizes the isospectral torus:

Theorem 1.5 (Magic Formula for MCMV matrices). Let $\mathsf{E} \subset \partial\mathbb{D}$ be a disjoint union of $g+1$ non-degenerate closed circular arcs, let $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$, and let $\vec{z}_{\mathsf{E}}$ be as in (1.16). Then, for any $A \in \mathbb{A}(\vec{z}_{\mathsf{E}})$,

$$A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*) \iff \Delta_{\mathsf{E}}(A) = S^{2(g+1)} + S^{-2(g+1)}, \tag{1.20}$$

and in this case $\Delta_A = \Delta_{\mathsf{E}}$.

On the other hand, for fixed $\vec{z} \in \mathbb{D}^n$ with $z_0 = 0$ and $A_0 \in \mathbb{A}_{\mathrm{per}}(\vec{z})$ with $\Delta_{A_0}(\lambda_*) > 0$, one has that $\sigma(A_0) = \Delta_{A_0}^{-1}([-2, 2])$, and consequently

$$\mathcal{T}_{\mathrm{MCMV}}(\sigma(A_0), \vec{z}, \lambda_*) = \left\{A \in \mathbb{A}_{\mathrm{per}}(\vec{z}) : \Delta_A = \Delta_{A_0}\right\}. \tag{1.21}$$

Remark. In light of the Magic Formula (1.20) and the observation that Δ_{E} is positive in all gaps, one may wonder why the normalization $\Delta_A(\lambda_*) > 0$ in (1.17) is required. We emphasize that the positivity of the discriminant for MCMV matrices in $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*)$ follows only from our special choice of vector $\vec{z}_{\mathsf{E}}$. If one were to fix a periodic MCMV matrix $A \in \mathbb{A}_{\mathrm{per}}(\vec{z})$ and compute its discriminant, it would not generally be the case that $\vec{z} = \vec{z}_{\sigma(A)}$, and the discriminant could then have different signs in different gaps (as demonstrated by, e.g., periodic CMV matrices). This makes it necessary to select a normalizing gap in the context of (1.21). Lemma 3.12 below will distinguish these situations.

The Magic Formula reveals further structure of MCMV matrices relating to the discriminant Δ_{E}. For simplicity, let us assume that the poles of Δ_{E} are simple (as this is typically the case). We will abuse the notation for residues and define

$$\operatorname{Res}_{\hat{z}_k} \Delta_{\mathsf{E}} := \lim_{z \to \hat{z}_k} \left(b_{z_k}^{-1} \Delta_{\mathsf{E}}\right)(z) = c_k. \tag{1.22}$$

Since Δ_{E} is real, it follows that $\operatorname{Res}_{z_k} \Delta_{\mathsf{E}} = \overline{c_k}$. By considering Δ_{E} as a meromorphic function on $\overline{\mathbb{C}}$, we see that it can be expressed in the form

$$\Delta_{\mathsf{E}}(z) = c + \sum_{k=0}^{g} \left(c_k b_{z_k}(z) + \overline{c_k} b_{z_k}(z)^{-1}\right), \quad c \in \mathbb{R}. \tag{1.23}$$

With (1.23) in mind, to understand the Magic Formula we need to understand each of the operators $b_{z_k}(A)$. In Lemma 4.9 we show that these can be represented by an operator Möbius transform of the same underlying CMV matrix $\mathcal{C}$, but related to the "shifted" diagonal operators $D_k = \left(1 - z_k D_0^*\right)^{-1}(D_0 - z_k)$. Since $(D_k)_{2k-1,2k-1} = (D_k)_{2k,2k} = 0$, it follows that $S^{-2k} b_{z_k}(A) S^{2k}$ has the same structure as an MCMV matrix. Moreover, since $b_{z_k}(A)$ is unitary, we have that $b_{z_k}(A)^{-1} = b_{z_k}(A)^*$. Hence, at the outermost diagonal of $\Delta_E(A)$ only one of the summands is non-vanishing. We have illustrated this for the off-diagonal block of $c_1 b_{z_1}(A) + \overline{c_1} b_{z_1}(A)^{-1}$ in the case $n = 4$ below:

$$\left(c_1 b_{z_1}(A) + \overline{c_1} b_{z_1}(A)^{-1}\right)_{\mathbf{i},\mathbf{i+1}} = \begin{bmatrix} 0 & & & & & & & \\ 0 & 0 & & & & & & \\ * & * & * & & & \mathbf{0} & & \\ * & * & * & 0 & & & & \\ * & * & * & 0 & 0 & & & \\ *,\star & *,\star & *,\star & \star & \star & \star & & \\ *,\star & *,\star & *,\star & \star & \star & \star & 0 & \\ *,\star & *,\star & *,\star & \star & \star & \star & 0 & 0 \end{bmatrix}. \tag{1.24}$$

Here $*$ and $\star$ indicate non-vanishing entries of $b_{z_1}(A)$ and $b_{z_1}(A)^{-1}$, respectively. In general, the outermost non-vanishing entry of $b_{z_k}(A)$ is the $(2k, 2(g+1+k))$-entry. Since all the other summands in (1.23) are vanishing at this position, the magic formula fixes the corresponding value of $b_{z_k}(A)$, that is,

$$\Delta_{\mathsf{E}}(A) = S^{2(g+1)} + S^{-2(g+1)} \implies b_{z_k}(A)_{2k,2(g+1+k)} = \frac{1}{\operatorname{Res}_{\hat{z}_k} \Delta_{\mathsf{E}}}. \tag{1.25}$$

The fact that this is a consequence of the structure of MCMV matrices will be proved in Theorem 4.10.

The above structure reveals an important property of MCMV matrices compared to their self-adjoint analog, GMP matrices. The relation (1.25) already indicates the importance of the values $\operatorname{Res}_{\hat{z}_k} \Delta_{\mathsf{E}}$. In order for $b_{z_k}(A)$ to be well-defined, they should be nonzero. In fact, for a general MCMV matrix A the values $|\operatorname{Res}_{z_k} \Delta_{S^{-2nj} A S^{2nj}}|$ are necessarily bounded away from zero. We will later see (cf. Lemma 4.3) that

$$\inf\{|\operatorname{Res}_{z_k} \Delta_A| : A \in \mathbb{A}(\vec{z})\} > 0. \tag{1.26}$$

This should be compared with [39, Theorem 3.3], where such a property was part of the definition of GMP matrices and guaranteed the existence of certain resolvents analogous to $b_{z_j}(A)$. It is natural that we do not need this condition since we are in the setting of unitary operators.

We also point out that the generalized discriminant for MCMV matrices involves the Ahlfors functions associated to two different points; in contrast, the analogous object for GMP matrices involves only the Ahlfors function at infinity. This discrepancy has the consequence that the associated MCMV matrices are even-periodic with half of the gaps closed (cf. Appendix A). While this could be avoided using a different discriminant, doing so would introduce further complications elsewhere. In particular, the benefits of defining the discriminant as we have are (1) we can treat the even- and odd-periodic CMV cases uniformly, and (2) our discriminant is a rational function.

1.2 Consequences for Schur and Caratheodory functions

Of course, one cannot discuss CMV matrices without discussing Schur functions. A Schur function is an analytic function $f : \mathbb{D} \to \overline{\mathbb{D}}$ mapping the open unit disk to its closure. We denote by $\mathcal{S}$ the class of all Schur functions. Provided $f \in \mathcal{S}$ is not a finite Blaschke product, the Schur algorithm

$$f_0(z) = f(z),$$
$$zf_{k+1}(z) = \frac{f_k(z) - \mathbf{a}_k}{1 - \overline{\mathbf{a}_k} f_k(z)}, \quad \mathbf{a}_k = f_k(0)$$

determines an infinite sequence of parameters $\{\mathbf{a}_k\} \in \mathbb{D}^{\mathbb{N}}$, also known as Schur parameters; conversely, any sequence $\{\mathbf{a}_k\} \in \mathbb{D}^{\mathbb{N}}$ determines a function $f \in \mathcal{S}$ by an associated continued fraction expansion (see, e.g., [30]). For our purposes, it is more convenient to denote the Schur algorithm in terms of equivalences of projective lines, that is,

$$\begin{bmatrix} f_k(z) \\ 1 \end{bmatrix} \sim U(\mathbf{a}_k) \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f_{k+1}(z) \\ 1 \end{bmatrix}, \tag{1.27}$$

where $v_1 \sim v_2$ if and only if there exists some nonzero $\lambda \in \mathbb{C}$ such that $v_1 = \lambda v_2$.

In correspondence to Schur functions are Caratheodory functions, analytic functions F from $\mathbb{D}$ to the right half-plane normalized such that $F(0) = 1$. A Caratheodory function F can be determined from a function $f \in \mathcal{S}$ by

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

or in terms of projective lines by

$$\begin{bmatrix} F(z) \\ 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f(z) \\ 1 \end{bmatrix}. \tag{1.28}$$

In this latter language it is clear this process is invertible, so indeed this correspondence is one-to-one. Caratheodory functions have a Herglotz integral representation as

$$F(z) = \int \frac{e^{it} + z}{e^{it} - z} \mathrm{d}\nu_F(e^{it})$$

for a unique probability measure ν_F on $\partial\mathbb{D}$, and are thus in correspondence with probability measures on the unit circle. Consequently, Schur functions can be put into correspondence with half-line CMV matrices, in the sense that for a given half-line CMV matrix $\mathcal{C}_+$, there exists $f_+ \in \mathcal{S}$ such that

$$\left\langle (\mathcal{C}_+ - z)^{-1}(\mathcal{C}_+ + z)\delta_0, \delta_0 \right\rangle = \frac{1 + zf_+(z)}{1 - zf_+(z)}, \tag{1.29}$$

and in fact one can check this f_+ is the Schur function with parameters $\{\mathbf{a}_k\}_{k\in\mathbb{N}}$ agreeing with the coefficients of $\mathcal{C}_+ = \mathcal{C}_+(\{\mathbf{a}_k\})$. Similarly, whole-line CMV matrices have two associated Schur functions, one corresponding to each half-line. Specifically, if $\{\mathbf{a}_k\}_{k\in\mathbb{Z}} \in \mathbb{D}^{\mathbb{Z}}$ and $\mathcal{C} = \mathcal{C}(\{\mathbf{a}_k\})$ is the associated CMV matrix, then one has that

$$\left\langle (\mathcal{C} - z)^{-1}(\mathcal{C} + z)\delta_0, \delta_0 \right\rangle = \frac{1 + zf_+(z)f_-(z)}{1 - zf_+(z)f_-(z)}, \tag{1.30}$$

where f_+ is the Schur function with parameters $\{\mathbf{a}_k\}_{k\in\mathbb{N}}$ and f_- is the Schur function with parameters $\{-\overline{\mathbf{a}_{-1}}, -\overline{\mathbf{a}_{-2}}, \ldots\}$; cf. [13, 25]. Any Schur function f has a natural factorization

$$f(z) = z^n \left(\prod_{w_k \neq 0} b_{w_k}(z) \frac{|w_k|}{w_k} \right) \exp\left(-\int \frac{e^{it} + z}{e^{it} - z} \mathrm{d}\nu_f(e^{it}) + i\tau \right),$$

where $n \geq 0$ is an integer, ν_f is a non-negative measure on $\partial\mathbb{D}$, $\tau \in \mathbb{R}/2\pi\mathbb{Z}$ (in fact, τ is the argument of $f(z)/z^n$ evaluated at $z = 0$), and the sequence $\{w_k\}$ of zeros satisfies the Blaschke condition $\sum(1 - |w_k|) < \infty$. We define

$$\sigma_{\mathrm{ess}}(f) := \mathrm{supp}(\mathrm{d}\nu_f) \cup \{w_k\}',$$

where $\{w_k\}' \subset \partial\mathbb{D}$ denotes the set of limit points of the sequence $\{w_k\}$ of zeros of f; cf [23, Lecture III]. Note that $\sigma_{\mathrm{ess}}(f)$ is a closed subset of $\partial\mathbb{D}$.

For a finite-gap set $\mathsf{E} \subset \partial\mathbb{D}$, a CMV matrix $\mathcal{C}$ lies in $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ precisely when its Schur functions f_+ and f_- solve the following Riemann–Hilbert problem:

$$f_-(e^{it}) = \overline{e^{it} f_+(e^{it})} \quad \text{for a.e. } e^{it} \in \mathsf{E}, \tag{1.31}$$

$$1 - e^{it} f_+(e^{it}) f_-(e^{it}) \neq 0 \text{ for } e^{it} \in \partial\mathbb{D} \setminus \mathsf{E}, \tag{1.32}$$

$$\sigma_{\mathrm{ess}}(z f_+ f_-) \subset \mathsf{E}. \tag{1.33}$$

We denote the class of all such admissible functions f_+ by

$$\mathcal{S}_+(\mathsf{E}) := \left\{ f_+ \in \mathcal{S} : \exists f_- \in \mathcal{S} \text{ s.t. } (1.31), (1.32), (1.33) \text{ hold} \right\}. \tag{1.34}$$

The description of $\mathcal{S}_+(\mathsf{E})$ as being equivalent to $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ in the finite-gap setting (and for even more general sets) was known already to Peherstorfer and Yuditskii [25]; specifically, they showed that, for $f_+ \in \mathcal{S}_+(\mathsf{E})$, the corresponding sequence $\{\mathbf{a}_k\} \in \mathbb{D}^{\mathbb{Z}}$ is almost-periodic.

This perspective gives us an alternative way of stating our main results: membership in $\mathcal{S}_+(\mathsf{E})$ is equivalent to the existence of an E-dependent Nevanlinna–Pick type interpolation, analogous to (1.27), whose coefficients are periodic up to a rotational phase.

Theorem 1.6. Fix a finite disjoint union of $g+1$ non-degenerate closed circular arcs $\mathsf{E} \subset \partial\mathbb{D}$, and let $\vec{z}_{\mathsf{E}}$ be as in (1.16). Then $f_+ \in \mathcal{S}_+(\mathsf{E})$ if and only if

$$\begin{bmatrix} f_+ \\ 1 \end{bmatrix} \sim \underbrace{U(a_0) \begin{bmatrix} b_{z_1} & 0 \\ 0 & 1 \end{bmatrix} U(a_1) \begin{bmatrix} b_{z_1} & 0 \\ 0 & 1 \end{bmatrix} U(a_2) \begin{bmatrix} b_{z_2} & 0 \\ 0 & 1 \end{bmatrix} \cdots U(a_{2g+1}) \begin{bmatrix} b_{z_0} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix}}_{= T_{A(\{a_k\},\vartheta;\vec{z}_{\mathsf{E}})}(z)} \begin{bmatrix} f_+ \\ 1 \end{bmatrix} \tag{1.35}$$

for some $\{a_k\} \in \mathbb{D}^{2(g+1)}$ and some $\vartheta \in \mathbb{R}/2\pi\mathbb{Z}$ such that

$$\frac{1}{B(z)} \operatorname{tr}\left(T_{A(\{a_k\},\vartheta;\vec{z}_{\mathsf{E}})}(z)\right) = \Delta_{\mathsf{E}}(z). \tag{1.36}$$

Using (1.28) to translate (1.35) into the language of Caratheodory functions, Theorem 1.6 immediately resolves a conjecture of Simon [31, Conjecture 11.9.6]:

Corollary 1.7. Fix a finite-gap set $\mathsf{E} \subset \partial\mathbb{D}$. For any $\mathcal{C} \in \mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$, the Caratheodory function F_+ associated to the half-line restriction $\mathcal{C}_+$ is a quadratic irrationality; that is, there exist polynomials $a(z), b(z)$, and $c(z)$ such that F_+ solves

$$a(z)F_+(z)^2 + b(z)F_+(z) + c(z) = 0 \tag{1.37}$$

for all $z \in \mathbb{C} \setminus \mathsf{E}$.

Real numbers that are quadratic irrationalities (with a, b, and c above as integers) are precisely those having eventually periodic continued fraction expansions. If one understands the interpolation of Theorem 1.6 as a special continued fraction expansion for the Schur function f_+, Corollary 1.7 should come as no surprise; indeed, our result shows that almost-periodicity of the Schur parameters associated to absolutely continuous finite-gap CMV matrices is actually a consequence of an underlying periodicity that the Schur algorithm was too naïve to see.

1.3 Methods and structure of the paper

The relationship of CMV matrices to orthogonal polynomials was discovered by Cantero, Moral, and Velázquez in 2003 [7]. Soon thereafter, the relationship of operator Möbius transforms of CMV matrices to the study of orthogonal rational functions (ORFs) on the unit circle was studied in work of Velázquez [37]. We recall these relationships in Section 2 to motivate the following construction, as well as to prove a coefficient stripping formula for Caratheodory functions associated to bases of ORFs.

Our approach to MCMV in the context of reflectionless operators is based on the functional model for the same, developed initially for Jacobi matrices by Sodin and Yuditskii [35] and later adapted for Schur functions and CMV matrices by Peherstorfer and Yuditskii [25]. Using the ideas developed by Eichinger and Yuditskii for GMP matrices (the Jacobi analog of MCMV matrices, cf. [9, 39]) and comparing this construction to that of Velázquez proves one direction of the equivalences in Theorems 1.4, 1.5, and 1.6. We review the functional model for CMV matrices and reveal the corresponding MCMV structure in Section 3.

Having motivated our class of MCMV matrices and shown that finite-gap CMV matrices correspond to periodic MCMV matrices, we perform a direct spectral analysis for periodic MCMV matrices after reviewing the corresponding classical analysis of CMV matrices (cf., e.g., [30–32]) in Section 4. We also analyze the structure of a general MCMV matrix in Section 4.3.

Finally, we use the tools developed in Sections 2 through 4 to completely resolve the proofs of Theorems 1.4, 1.5, and 1.6 in Section 5.

2 Orthogonal Rational Functions

The aim of this section is to establish a coefficient stripping formula for Caratheodory functions associated to bases of ORFs. Our main result, Theorem 2.2, is an analog of the Stieltjes expansion for m-functions of Jacobi matrices [32, Theorem 3.2.4] and of Peherstorfer's formula for OPUC [30, Theorem 3.4.2]. It will play an important role in Section 4 where we seek to solve the direct spectral problem for periodic MCMV matrices.

2.1 The Szegő recursion

Given a nontrivial (i.e., of infinite support) probability measure ν on $\partial\mathbb{D}$, one obtains the monic orthogonal polynomials $\Phi_k := \Phi_k(z,\nu)$ by orthogonalizing the family $\{1, z, z^2, \ldots\}$ in $L^2(d\nu)$. The Φ_k's are known to satisfy a recurrence relation of the form

$$\Phi_{k+1}(z) = z\Phi_k(z) - \overline{\mathbf{a}_k}\Phi_k^*(z) \tag{2.1}$$

for some sequence $\{\mathbf{a}_k\}_{k\in\mathbb{N}}$ of numbers in $\mathbb{D}$. Though it may look strange, we purposely write $-\overline{\mathbf{a}_k}$ in (2.1) so that the $\mathbf{a}_k$'s coincide with the Schur parameters (introduced in Section 1.2). Following [30], we shall also refer to the $\mathbf{a}_k$'s as Verblunsky coefficients and recall there is a one-to-one correspondence between such sequences (in $\mathbb{D}^{\mathbb{N}}$) and nontrivial probability measures on $\partial\mathbb{D}$.

Φ_k^* is the reversed polynomial, that is,

$$\Phi_k^*(z) = z^k\overline{\Phi_k(\hat{z})}.$$

While the notation of * is convenient, it is ambiguous. It depends on the class

$$\mathcal{L}_k := \operatorname{span}\{1, z, \ldots, z^k\}$$

and has a different meaning for $\mathcal{L}_k$ and $\mathcal{L}_j$ when $k \neq j$. Note that the operation $\varphi \mapsto z^k\overline{\varphi(\hat{z})}$ acts as an involution on the subspace $\mathcal{L}_k$. We shall also use this abuse of notation for bases of ORFs (where naturally z^k is substituted by the Blaschke product corresponding to the poles of the first k basis elements). Applying * for the class $\mathcal{L}_{k+1}$ to (2.1) yields the so-called Szegő recursion

$$\begin{bmatrix}\Phi_{k+1}(z)\\ \Phi_{k+1}^*(z)\end{bmatrix} = \begin{bmatrix}1 & -\overline{\mathbf{a}_k}\\ -\mathbf{a}_k & 1\end{bmatrix}\begin{bmatrix}z & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\Phi_k(z)\\ \Phi_k^*(z)\end{bmatrix}. \tag{2.2}$$

For the orthonormal polynomials $\varphi_k := \Phi_k/\|\Phi_k\|$, this recursion takes the form

$$\begin{bmatrix} \varphi_{k+1}(z) \\ \varphi_{k+1}^*(z) \end{bmatrix} = U(-\overline{\mathbf{a}_k}) \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi_k(z) \\ \varphi_k^*(z) \end{bmatrix} \tag{2.3}$$

with $U(\cdot)$ as in (1.10).

2.2 Coefficient stripping for ORFs

Let $z_0 = 0$ and fix a sequence of points $\{z_k\}_{k=1}^{\infty}$ in $\mathbb{D}$ violating the Blaschke condition, that is,

$$\sum_k (1 - |z_k|) = \infty. \tag{2.4}$$

Note that (2.4) is trivially satisfied if $\sup_k |z_k| < 1$ (which will always be the case in our setting). Recall that by $\{B_k\}$ we denote the family of finite Blaschke products

$$B_0(z) = 1, \quad B_k(z) = \prod_{j=1}^{k} b_{z_j}(z).$$

Given a nontrivial probability measure ν on $\partial\mathbb{D}$, let $\{\varphi_k\}_{k\in\mathbb{N}}$ be the corresponding sequence of orthonormal rational functions obtained by orthogonalizing the family $\{B_k\}_{k\in\mathbb{N}}$ in $L^2(\mathrm{d}\nu)$. With $\mathcal{L}_k := \mathrm{span}\{\varphi_j : 0 \le j \le k\}$, the associated *-operator is now defined by $\varphi^*(z) = B_k(z)\overline{\varphi(\hat{z})}$ for $\varphi \in \mathcal{L}_k$. Defining

$$\eta_k^2 := 1 - |z_k|^2$$

and choosing the right unimodular constants in the normalization (in particular, $\varphi_0 \equiv 1$), the φ_k's satisfy the recurrence relation (see [5, Theorem 4.1.3] and [37])

$$\begin{bmatrix} \varphi_{k+1}(z) \\ \varphi_{k+1}^*(z) \end{bmatrix} = \frac{1 - \overline{z_k} z}{1 - \overline{z_{k+1}} z} \frac{\eta_{k+1}}{\eta_k} U(-\overline{a_k}) \begin{bmatrix} b_{z_k}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi_k(z) \\ \varphi_k^*(z) \end{bmatrix}. \tag{2.5}$$

The rational functions ψ_k of the second kind are defined by

$$\psi_0 \equiv 1, \quad \psi_k(z) = \int (\varphi_k(e^{it}) - \varphi_k(z)) \frac{e^{it} + z}{e^{it} - z} \mathrm{d}\nu(e^{it}), \quad k \ge 1$$

and satisfy almost the same recurrence relation as φ_k, namely

$$\begin{bmatrix} \psi_{k+1}(z) \\ -\psi^*_{k+1}(z) \end{bmatrix} = \frac{1-\overline{z_k}z}{1-\overline{z_{k+1}}z}\frac{\eta_{k+1}}{\eta_k}U(-\overline{a_k})\begin{bmatrix} b_{z_k}(z) & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \psi_k(z) \\ -\psi^*_k(z) \end{bmatrix}. \tag{2.6}$$

Note that the coefficients a_k in (2.5)–(2.6) belong to $\mathbb{D}$ and are explicitly given by

$$a_k = \left\langle \frac{1-\overline{z_k}z}{z-z_{k+1}}, \varphi_k \right\rangle_\nu \Bigg/ \left\langle \frac{z_k - z}{z_{k+1}-z}, \varphi^*_k \right\rangle_\nu; \tag{2.7}$$

cf. [5, Theorem 4.1.2]. Conversely, starting from arbitrary coefficients $\{a_k\}_{k\in\mathbb{N}} \in \mathbb{D}^{\mathbb{N}}$ one can generate a sequence of rational functions by (2.5) and show that they are orthogonal with respect to some probability measure ν on $\partial\mathbb{D}$. This is the content of the following known results:

Theorem 2.1. [5, Theorems 8.1.4 and 9.2.1] Suppose $\{z_k\} \in \mathbb{D}^{\mathbb{N}}$ violates the Blaschke condition, that is, (2.4) holds. Given a sequence $\{a_k\} \in \mathbb{D}^{\mathbb{N}}$, define the rational functions $\{\varphi_k\}_{k\in\mathbb{N}}$ (with $\varphi_0 \equiv 1$) by (2.5) and let ν_k denote the Bernstein–Szegő approximant

$$d\nu_k = \frac{1-|z_k|^2}{|e^{it}-z_k|^2}\frac{1}{|\varphi^*_k(e^{it})|^2}\frac{dt}{2\pi}. \tag{2.8}$$

Then the associated Caratheodory function F_{ν_k} can be written as

$$F_{\nu_k}(z) = \psi^*_k(z)/\varphi^*_k(z) \tag{2.9}$$

and ν_k converges weakly to some probability measure ν, which, in turn, is the unique measure of orthogonality for $\{\varphi_k\}$. In particular,

$$\lim_{k\to\infty} F_{\nu_k}(z) = F_\nu(z) \tag{2.10}$$

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

Note that if $z_k \equiv 0$, then (2.5) reduces to the standard Szegő recursion (2.2). Just as for OPUC, there is a one-to-one correspondence between coefficient sequences $\{a_k\} \in \mathbb{D}^{\mathbb{N}}$ and nontrivial probability measures on $\partial\mathbb{D}$. In fact, the theory for ORFs generalizes the one for OPUC. We mention in passing that the assumption (2.4) ensures the measure ν of orthogonality be unique.

We are now ready to derive the promised coefficient stripping formula. Let $z_0 = 0, z_1, \dots, z_{p-1}$ be a finite number of points in $\mathbb{D}$ and consider the sequence $\{z_k\}_{k\in\mathbb{N}}$ obtained by periodic extension of the initial p values (i.e., $z_{k+p} = z_k$ for all k). Note that in this situation the Blaschke condition is trivially violated. Let

$$Y_k(z) = \begin{bmatrix} \psi_k(z) & \varphi_k(z) \\ -\psi_k^*(z) & \varphi_k^*(z) \end{bmatrix}; \quad \text{in particular,} \;\; Y_0 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

and define

$$M(z) := \begin{bmatrix} M_{11}(z) & M_{12}(z) \\ M_{21}(z) & M_{22}(z) \end{bmatrix} = Y_0 \begin{bmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_0) \cdots \begin{bmatrix} b_{z_{p-1}}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_{p-1}) Y_0^{-1}. \tag{2.11}$$

Our result then reads as follows:

Theorem 2.2. Let F_ν be the Caratheodory function associated to the sequence $\{a_k\}_{k\in\mathbb{N}}$ and suppose $F_\nu^{(1)}$ corresponds to the shifted sequence $\{a_{k+p}\}_{k\in\mathbb{N}}$. Then

$$F_\nu(z) = \frac{M_{11}(z)F_\nu^{(1)}(z) + M_{12}(z)}{M_{21}(z)F_\nu^{(1)}(z) + M_{22}(z)} \tag{2.12}$$

and M can be expressed in terms of the ORFs and the rational functions of the second kind by

$$M(z) = \frac{1}{2} \begin{bmatrix} \psi_p(z) + \psi_p^*(z) & \psi_p^*(z) - \psi_p(z) \\ \varphi_p^*(z) - \varphi_p(z) & \varphi_p(z) + \varphi_p^*(z) \end{bmatrix}. \tag{2.13}$$

Proof. Due to (2.5) and (2.6), we have

$$Y_{k+1}(z) = \frac{1 - \overline{z_k} z}{1 - \overline{z_{k+1}} z} \frac{\eta_{k+1}}{\eta_k} U(-\overline{a_k}) \begin{bmatrix} b_{z_k}(z) & 0 \\ 0 & 1 \end{bmatrix} Y_k(z).$$

Iterating p times, starting from $k = p-1$, the factors in front of the U's cancel (due to telescoping) and it follows that

$$Y_p(z) = W(z) Y_0, \tag{2.14}$$

where W is the transfer matrix

$$W(z) := W(z, \{z_k\}, \{a_k\}) = U(-\overline{a_{p-1}}) \begin{bmatrix} b_{z_{p-1}}(z) & 0 \\ 0 & 1 \end{bmatrix} \cdots U(-\overline{a_0}) \begin{bmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.15}$$

Using a superscript (l) for the objects related to to the shifted sequence $\{a_{k+lp}\}_{k\in\mathbb{N}}$, we obtain in a similar way that

$$Y^{(1)}_{np}(z) = W^{(n)}(z) \cdots W^{(1)}(z) Y_0$$

and

$$Y_{(n+1)p}(z) = W^{(n)}(z) \cdots W^{(1)}(z) W(z) Y_0.$$

Hence,

$$Y_{(n+1)p}(z) = Y^{(1)}_{np}(z) Y_0^{-1} W(z) Y_0$$

and considering the second row of this identity (or rather its transpose) yields

$$\begin{bmatrix} \psi^*_{(n+1)p}(z) \\ \varphi^*_{(n+1)p}(z) \end{bmatrix} = \left(jY_0^{-1} W(z) Y_0 j\right)^{\intercal} \begin{bmatrix} \left(\psi^{(1)}_{np}\right)^*(z) \\ \left(\varphi^{(1)}_{np}\right)^*(z) \end{bmatrix}, \quad j = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.16}$$

To see that

$$\left(jY_0^{-1} W(z) Y_0 j\right)^{\intercal} = M(z) \tag{2.17}$$

with M as defined in (2.11), we use that $Y_0 j = 2jY_0^{-1}$, $Y_0^{\intercal} = 2Y_0^{-1}$, and $jU(-\overline{a})j = U(a)^{\intercal}$. Due to (2.9) and (2.10), we now obtain (2.12) by passing to the limit as $n \to \infty$ in (2.16). Finally, the identity (2.13) follows directly from (2.14) and (2.17). ■

If the sequence $\{a_k\}$ is periodic (or periodic up to a rotational phase) and the period matches the period of the sequence $\{z_k\}$, then our result simplifies and (2.12) turns into a quadratic equation for F_ν. The result becomes particularly simple if we pass from the relation for Caratheodory functions to the one for Schur functions:

Corollary 2.3. Suppose $\{a_k\}_{k\in\mathbb{N}}$ is periodic up to a rotational phase with period p and phase -2ϑ, and let f_ν denote the associated Schur function; see (1.28). If

$$T(z) = U(a_0) \begin{bmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_1) \cdots \begin{bmatrix} b_{z_{p-1}}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_{p-1}) \begin{bmatrix} b_{z_0}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix}, \tag{2.18}$$

then

$$\begin{bmatrix} f_\nu \\ 1 \end{bmatrix} \sim T(z) \begin{bmatrix} f_\nu \\ 1 \end{bmatrix}. \tag{2.19}$$

Proof. Start by recalling that $z_0 = 0$ so that $b_{z_0}(z) = z$. Due to the Schur algorithm, we see that multiplying a Schur function by a unimodular constant leads to multiplication of the associated Schur parameters by the same constant. Using the fact that $f_\nu^{(1)}$ has Schur parameters $\{a_{k+p}\}_{k\in\mathbb{N}}$ and since

$$a_{k+p} = e^{-2i\vartheta} a_k,$$

it follows that $f_\nu = e^{2i\vartheta} f_\nu^{(1)}$. Thus (2.19) follows from (2.12) and (1.28). ■

3 The Functional Model

In the last two decades a significant amount of progress has been made in understanding reflectionless one-dimensional operators as being related to multiplication operators on certain subspaces of Hardy spaces associated to multiply-connected Riemann surfaces. We broadly refer to this construction as a "functional model" for the associated operators. In this section, we first recall the requisite definitions to develop such models, followed by the specific model of Peherstorfer–Yuditskii for almost-periodic CMV matrices, and finally associate to this our model for MCMV matrices.

3.1 Hardy spaces of character automorphic functions

Fix a finite-gap set $\mathsf{E} \subset \partial\mathbb{D}$. By means of the Koebe–Poincaré uniformization theorem, the spectral complement in the Riemann sphere $\overline{\mathbb{C}} \setminus \mathsf{E}$ is uniformized by the disk $\mathbb{D}$; that is, there exists a Fuchsian group Γ and a meromorphic function $\mathfrak{z} : \mathbb{D} \to \overline{\mathbb{C}} \setminus \mathsf{E}$ with the following properties:

1. $\forall z \in \overline{\mathbb{C}} \setminus \mathsf{E} \ \exists \zeta \in \mathbb{D} : \mathfrak{z}(\zeta) = z,$
2. $\mathfrak{z}(\zeta_1) = \mathfrak{z}(\zeta_2) \iff \exists \gamma \in \Gamma : \zeta_1 = \gamma(\zeta_2).$

We fix a normalization point $\lambda_* \in \partial\mathbb{D} \setminus \mathsf{E}$ and assume that $\mathfrak{z}(0) = \lambda_*$ and that $(-1, 1)$ is mapped onto the connected component of $\partial\mathbb{D} \setminus \mathsf{E}$ containing λ_*. Due to this choice, there exists a fundamental domain $\mathcal{F}$ for the action of Γ, which is symmetric with respect to complex conjugation, that is,

$$\mathcal{F} = \{\overline{\zeta} : \zeta \in \mathcal{F}\}, \quad \gamma^{-1}(\zeta) = \overline{\gamma}(\zeta) := \overline{\gamma(\overline{\zeta})}, \quad \gamma \in \Gamma. \tag{3.1}$$

It follows that

$$\overline{\mathfrak{z}(\overline{\zeta})} = \mathfrak{z}(\zeta)^{-1}; \tag{3.2}$$

in particular, if $\zeta_0 \in \mathcal{F}$ is such that $\mathfrak{z}(\zeta_0) = 0$, then $\mathfrak{z}(\overline{\zeta_0}) = \infty$.

We denote by Γ^* the group of unitary characters of Γ; that is, group homomorphisms from Γ into $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$. By the covering space formalism, Γ is group isomorphic to the fundamental group $\pi_1(\overline{\mathbb{C}} \setminus \mathsf{E})$, and so $\Gamma^* \cong \mathbb{T}^g$ (where $g+1$ is the number of gaps of E).

Let $H^2 = H^2(\mathbb{D})$ denote the usual Hardy space of the unit disk. For $\alpha \in \Gamma^*$, we consider the Hardy space of character automorphic functions

$$H^2(\alpha) := \left\{f \in H^2 : f \circ \gamma = e^{i\alpha(\gamma)} f \ \ \forall \gamma \in \Gamma\right\}$$

equipped with the standard H^2 inner product

$$\langle f, g\rangle = \int_0^{2\pi} f(e^{it})\overline{g(e^{it})}\frac{\mathrm{d}t}{2\pi}.$$

The group Γ also acts on $\partial\mathbb{D}$; moreover, there exists a measurable fundamental set for this action [27]. In this sense we can define the larger space $L^2(\alpha)$ as the space of those functions $f : \partial\mathbb{D} \to \mathbb{C}$, which are square integrable and α-automorphic:

$$L^2(\alpha) := \left\{f : \partial\mathbb{D} \to \mathbb{C} : \|f\|^2 < \infty, f \circ \gamma = e^{i\alpha(\gamma)} f \ \ \forall \gamma \in \Gamma\right\}.$$

Naturally, $H^2(\alpha) \subset L^2(\alpha)$ via the identification of a function $f \in H^2$ with its radial limit function on the boundary.

For finite-gap sets E, a fundamental result of Widom [38] implies that $H^2(\alpha)$ is nontrivial for all $\alpha \in \Gamma^*$. This in fact applies to all subsets $\mathsf{E} \subset \mathbb{C}$ of so-called Parreau–Widom type (see, e.g., [14] for details). By continuity of the point evaluation functional, $H^2(\alpha)$ admits a family of reproducing kernels $\{k^\alpha(\zeta, \zeta_1)\}_{\zeta_1 \in \mathbb{D}}$ such that

$$\langle f, k^\alpha(\cdot, \zeta_1)\rangle = f(\zeta_1) \quad \forall f \in H^2(\alpha). \tag{3.3}$$

By the reproducing property and nontriviality of $H^2(\alpha)$, it follows that $k^\alpha(\zeta_1,\zeta_1) > 0$. We may thus define the corresponding normalized vectors by

$$K^\alpha(\zeta,\zeta_1) = \frac{k^\alpha(\zeta,\zeta_1)}{\sqrt{k^\alpha(\zeta_1,\zeta_1)}}$$

and note that

$$\langle f, K^\alpha(\,\cdot\,,\zeta_1)\rangle = \frac{f(\zeta_1)}{K^\alpha(\zeta_1,\zeta_1)} \quad \forall f \in H^2(\alpha). \tag{3.4}$$

We will sometimes abbreviate $k^\alpha_{\zeta_1}(\zeta) := k^\alpha(\zeta,\zeta_1)$ and $K^\alpha_{\zeta_1}(\zeta) := K^\alpha(\zeta,\zeta_1)$. Since $f \in H^2(\alpha)$ implies $\overline{f(\overline{\zeta})} \in H^2(\alpha)$, we have

$$\overline{K^\alpha(\overline{\zeta},\zeta_1)} = K^\alpha(\zeta,\overline{\zeta_1}); \tag{3.5}$$

in particular, $K^\alpha(\zeta_1,\zeta_1) = K^\alpha(\overline{\zeta_1},\overline{\zeta_1})$. We shall make frequent use of (3.5) (as well as (3.8) below) in our computations.

For finite-gap sets E, the map $\alpha \mapsto K^\alpha(\zeta_1,\zeta_1)$ is continuous for every $\zeta_1 \in \mathbb{D}$. This property is in fact responsible for the almost-periodic structure of the CMV matrices in $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$, see Theorem 3.4 below. We mention in passing that this type of continuity in the character is known to hold for all Parreau–Widom sets $\mathsf{E} \subset \mathbb{C}$ satisfying the so-called Direct Cauchy Theorem (see, e.g., [14]).

For a fixed $\zeta_1 \in \mathbb{D}$ we denote by

$$\mathfrak{b}(\zeta,\zeta_1) := e^{i\phi} \prod_{\gamma\in\Gamma} \frac{\gamma(\zeta_1) - \zeta}{1 - \overline{\gamma(\zeta_1)}\zeta} \frac{|\gamma(\zeta_1)|}{\gamma(\zeta_1)} \tag{3.6}$$

the Blaschke product with zeros at the orbit of ζ_1 under Γ, and with $\phi = \phi(\zeta_1)$ normalized such that

$$\mathfrak{b}(\zeta_1,\overline{\zeta_1}) > 0. \tag{3.7}$$

As follows directly from (3.6), (3.1), and our chosen normalization (3.7), we have that

$$\overline{\mathfrak{b}(\overline{\zeta},\zeta_1)} = \mathfrak{b}(\zeta,\overline{\zeta_1}). \tag{3.8}$$

If we want to suppress the dependence on ζ, we may also write $\mathfrak{b}_{\zeta_1}(\zeta) = \mathfrak{b}(\zeta,\zeta_1)$. Note that $\mathfrak{b}_{\zeta_1}$ is related to the potential-theoretic Green's function $G_{\overline{\mathbb{C}}\setminus\mathsf{E}}(z,z_1)$ of the domain $\overline{\mathbb{C}} \setminus \mathsf{E}$ with pole at $z_1 = \mathfrak{z}(\zeta_1)$ by

$$-\log\left|\mathfrak{b}_{\zeta_1}(\zeta)\right| = G_{\overline{\mathbb{C}}\setminus\mathsf{E}}\big(\mathfrak{z}(\zeta),z_1\big). \tag{3.9}$$

Moreover, $\mathfrak{b}_{\zeta_1}$ is character automorphic with some character μ_{ζ_1}, that is,

$$\mathfrak{b}_{\zeta_1}(\gamma(\zeta)) = e^{i\mu_{\zeta_1}(\gamma)}\mathfrak{b}_{\zeta_1}(\zeta) \quad \forall\gamma \in \Gamma. \tag{3.10}$$

Interestingly, since $\mathsf{E} \subset \partial\mathbb{D}$,

$$\log|z| = G_{\overline{\mathbb{C}}\setminus\mathsf{E}}(z,\infty) - G_{\overline{\mathbb{C}}\setminus\mathsf{E}}(z,0).$$

Thus we may represent the uniformization $\mathfrak{z}$ as a ratio of distinguished Blaschke products:

$$\mathfrak{z}(\zeta) = e^{i\phi_0}\frac{\mathfrak{b}(\zeta,\zeta_0)}{\mathfrak{b}(\zeta,\overline{\zeta_0})}, \tag{3.11}$$

where $\mathfrak{z}(\zeta_0) = 0$ as before and $\phi_0 \in \mathbb{T}$ is some phase. Since $\mathfrak{z}$ is automorphic, it follows that $\mu_{\zeta_0} = \mu_{\overline{\zeta_0}}$; we will abbreviate this common character by μ_0.

In the coming subsections, we will study multiplication by this uniformization map $\mathfrak{z}$ as a linear operator on $L^2(\alpha)$ with respect to different bases. To this end, we require a technical lemma on reproducing kernels that allows us to effectively compute residues. We first recall the following orthogonal decomposition of $H^2(\alpha)$:

Lemma 3.1 ([35]). For $\zeta_1 \in \mathbb{D}$, we have

$$\mathcal{K}_{\mathfrak{b}_{\zeta_1}}(\alpha) := H^2(\alpha) \ominus \mathfrak{b}_{\zeta_1}H^2(\alpha - \mu_{\zeta_1}) = \operatorname{span}\{k^\alpha_{\zeta_1}\}. \tag{3.12}$$

Proof. Using the reproducing kernel property, it is clear that $k^\alpha_{\zeta_1} \in \mathcal{K}_{\mathfrak{b}_{\zeta_1}}(\alpha)$. Conversely, let $f \in H^2(\alpha)$ and suppose $f \perp k^\alpha_{\zeta_1}$. Then

$$0 = \langle f, k^\alpha_{\zeta_1}\rangle = f(\zeta_1)$$

and since f is character automorphic, $f(\gamma(\zeta_1)) = 0$ for all $\gamma \in \Gamma$. The standard factorization theorem for H^2 functions and a comparison of the characters now imply that $f \in \mathfrak{b}_{\zeta_1}H^2(\alpha - \mu_{\zeta_1})$. ■

Lemma 3.2. For $\zeta_1 \in \mathbb{D}$, let $\mathfrak{b}_{\zeta_1}$ and μ_{ζ_1} be as above. If $f \in L^2(\alpha)$ is such that $\mathfrak{b}_{\zeta_1}f \in H^2(\alpha + \mu_{\zeta_1})$, then for $\zeta_2 \neq \zeta_1$ we have

$$\langle f, K^\alpha_{\zeta_2}\rangle = \frac{f(\zeta_2)}{K^\alpha(\zeta_2,\zeta_2)} - \frac{(\mathfrak{b}_{\zeta_1}f)(\zeta_1)}{K^{\alpha+\mu_{\zeta_1}}(\zeta_1,\zeta_1)}\frac{K^{\alpha+\mu_{\zeta_1}}(\zeta_2,\zeta_1)}{\mathfrak{b}_{\zeta_1}(\zeta_2)K^\alpha(\zeta_2,\zeta_2)}. \tag{3.13}$$

Proof. By our assumptions and Lemma 3.1,

$$g := \mathfrak{b}_{\zeta_1} f - \left\langle \mathfrak{b}_{\zeta_1} f, K_{\zeta_1}^{\alpha+\mu_{\zeta_1}} \right\rangle K_{\zeta_1}^{\alpha+\mu_{\zeta_1}} \in \mathfrak{b}_{\zeta_1} H^2(\alpha).$$

Since $\langle K_{\zeta_1}^{\alpha+\mu_{\zeta_1}}, \mathfrak{b}_{\zeta_1} K_{\zeta_2}^{\alpha} \rangle = 0$, we have on the one hand that

$$\left\langle g, \mathfrak{b}_{\zeta_1} K_{\zeta_2}^{\alpha} \right\rangle = \left\langle \mathfrak{b}_{\zeta_1} f, \mathfrak{b}_{\zeta_1} K_{\zeta_2}^{\alpha} \right\rangle - 0 = \langle f, K_{\zeta_2}^{\alpha} \rangle.$$

On the other hand, as $g/\mathfrak{b}_{\zeta_1} \in H^2(\alpha)$, we also have

$$\langle g, \mathfrak{b}_{\zeta_1} K_{\zeta_2}^{\alpha} \rangle = \langle g/\mathfrak{b}_{\zeta_1}, K_{\zeta_2}^{\alpha} \rangle = \frac{f(\zeta_2)}{K^{\alpha}(\zeta_2, \zeta_2)} - \frac{(\mathfrak{b}_{\zeta_1} f)(\zeta_1)}{K^{\alpha+\mu_{\zeta_1}}(\zeta_1, \zeta_1)} \frac{K^{\alpha+\mu_{\zeta_1}}(\zeta_2, \zeta_1)}{\mathfrak{b}_{\zeta_1}(\zeta_2) K^{\alpha}(\zeta_2, \zeta_2)}.$$

This completes the proof. ■

3.2 The Peherstorfer–Yuditskii model for CMV matrices

To motivate the MCMV functional model, we first recall the functional model for the usual CMV matrices. Everything that follows in this section is in some way already presented in the literature. We try to be quite precise anyway, because we feel that the meaning of the additional parameter $\tau \in \mathbb{T}$ has not really been discussed yet in terms of the functional model. Moreover, it will give us an understanding of the notion of *periodicity up to a phase* in CMV matrices, which will be important in the later part of our paper.

Let $(\alpha, \tau) \in \Gamma^* \times \mathbb{T}$ and define

$$x_0^{\alpha,\tau} = K_{\overline{\zeta_0}}^{\alpha}, \quad x_1^{\alpha,\tau} = e^{i\tau} \mathfrak{b}_{\overline{\zeta_0}} K_{\zeta_0}^{\alpha-\mu_0},$$

$$y_0^{\alpha,\tau} = e^{i\tau} K_{\zeta_0}^{\alpha}, \quad y_1^{\alpha,\tau} = \mathfrak{b}_{\zeta_0} K_{\overline{\zeta_0}}^{\alpha-\mu_0}.$$

For ϕ_0 given by (3.11) we define, for every $l \in \mathbb{Z}$,

$$x_{2l}^{\alpha,\tau} = e^{-il\phi_0} \mathfrak{b}_{\zeta_0}^{l} \mathfrak{b}_{\overline{\zeta_0}}^{l} x_0^{\alpha-2l\mu_0,\tau}, \quad x_{2l+1}^{\alpha,\tau} = e^{il\phi_0} \mathfrak{b}_{\zeta_0}^{l} \mathfrak{b}_{\overline{\zeta_0}}^{l+1} x_1^{\alpha-(2l+1)\mu_0,\tau}, \tag{3.14}$$

$$y_{2l}^{\alpha,\tau} = e^{il\phi_0} \mathfrak{b}_{\zeta_0}^{l} \mathfrak{b}_{\overline{\zeta_0}}^{l} y_0^{\alpha-2l\mu_0,\tau}, \quad y_{2l+1}^{\alpha,\tau} = e^{-il\phi_0} \mathfrak{b}_{\zeta_0}^{l+1} \mathfrak{b}_{\overline{\zeta_0}}^{l} y_1^{\alpha-(2l+1)\mu_0,\tau}. \tag{3.15}$$

It is straightforward to see that for any $\tau \in \mathbb{T}$, $\{x_0^{\alpha,\tau}, x_1^{\alpha,\tau}\}$ and $\{y_0^{\alpha,\tau}, y_1^{\alpha,\tau}\}$ form two distinct orthonormal bases of the two-dimensional subspace

$$\mathcal{K}_{\mathfrak{b}_{\zeta_0}\mathfrak{b}_{\overline{\zeta_0}}}(\alpha) := \operatorname{span}\{K_{\zeta_0}^{\alpha}, K_{\overline{\zeta_0}}^{\alpha}\} = H^2(\alpha) \ominus b_{\zeta_0} b_{\overline{\zeta_0}} H^2(\alpha - 2\mu_0). \tag{3.16}$$

Iterating this decomposition exhausts $H^2(\alpha)$ (and in fact, the larger space $L^2(\alpha)$); in particular, we have the following:

Proposition 3.3. The systems $\{x_k^{\alpha,\tau}\}$ and $\{y_k^{\alpha,\tau}\}$ for $k \in \mathbb{N}$ (resp., $k \in \mathbb{Z}$) form orthonormal bases for $H^2(\alpha)$ (resp., $L^2(\alpha)$).

Almost-periodic absolutely continuous whole-line CMV matrices with spectrum E arise exactly as multiplication by $\mathfrak{z}$ in the basis $\{y_k^{\alpha,\tau}\}_{k\in\mathbb{Z}}$:

Theorem 3.4 (Peherstorfer–Yuditskii [25]). Multiplication by $\mathfrak{z}$ in the basis $\{y_k^{\alpha,\tau}\}_{k\in\mathbb{Z}}$ is a CMV matrix $\mathcal{C}(\alpha,\tau)$ with almost-periodic Verblunsky coefficients given by

$$\mathbf{a}_k(\alpha,\tau) = e^{-ik\phi_0}\mathbf{A}(\alpha - k\mu_0,\tau), \quad \boldsymbol{\rho}_k(\alpha) = \mathbf{R}(\alpha - k\mu_0), \tag{3.17}$$

where

$$\mathbf{A}(\alpha,\tau) = e^{-i\tau}\frac{K^\alpha(\zeta_0,\overline{\zeta_0})}{K^\alpha(\zeta_0,\zeta_0)}, \quad \mathbf{R}(\alpha) = \mathfrak{b}(\overline{\zeta_0},\zeta_0)\frac{K^{\alpha-\mu_0}(\zeta_0,\zeta_0)}{K^\alpha(\zeta_0,\zeta_0)} = \sqrt{1-|\mathbf{A}(\alpha,\tau)|^2}. \tag{3.18}$$

Remark. Peherstorfer and Yuditskii actually studied the family of Schur functions $f^{\alpha,\tau}$ given by

$$f^{\alpha,\tau}\circ\mathfrak{z} := e^{-i\tau}\frac{K^\alpha_{\overline{\zeta_0}}}{K^\alpha_{\zeta_0}}, \tag{3.19}$$

but this is equivalent by equality of Schur parameters and Verblunsky coefficients. This perspective explains the necessity of including the parameter τ; we wish to completely classify such Schur functions, not merely classify them up to a rotation.

We can see this theorem via the *LM* structure by alternating between the basis $\{y_k^{\alpha,\tau}\}$ and the dual basis $\{x_k^{\alpha,\tau}\}$. Denoting

$$\Theta_k(\alpha,\tau) := \begin{bmatrix} \overline{\mathbf{a}_k(\alpha,\tau)} & \boldsymbol{\rho}_k(\alpha) \\ \boldsymbol{\rho}_k(\alpha) & -\mathbf{a}_k(\alpha,\tau) \end{bmatrix},$$

we have the following:

Lemma 3.5. With notation as above,

$$\begin{bmatrix} y_0^{\alpha,\tau} \\ y_1^{\alpha,\tau} \end{bmatrix} = \Theta_0(\alpha,\tau)\begin{bmatrix} x_0^{\alpha,\tau} \\ x_1^{\alpha,\tau} \end{bmatrix}, \quad \mathfrak{z}\begin{bmatrix} x_1^{\alpha,\tau} \\ x_2^{\alpha,\tau} \end{bmatrix} = \Theta_1(\alpha,\tau)\begin{bmatrix} y_1^{\alpha,\tau} \\ y_2^{\alpha,\tau} \end{bmatrix}. \tag{3.20}$$

Proof. Since $K^{\alpha}_{\zeta_0}, K^{\alpha}_{\overline{\zeta_0}} \in \mathcal{K}_{\mathfrak{b}_{\zeta_0}\mathfrak{b}_{\overline{\zeta_0}}}(\alpha)$, it follows from the reproducing kernel property that

$$y_0^{\alpha,\tau} = \overline{\mathbf{A}(\alpha,\tau)}x_0^{\alpha,\tau} + \mathbf{R}(\alpha)x_1^{\alpha,\tau}, \quad x_0^{\alpha,\tau} = \mathbf{A}(\alpha,\tau)y_0^{\alpha,\tau} + \mathbf{R}(\alpha)y_1^{\alpha,\tau}.$$

Using (3.11), the lemma follows by algebraic manipulations. ■

Proof of Theorem 3.4. We can shift the relations in the previous lemma to see that, taking

$$L := L(\alpha,\tau) = \bigoplus_{l\in\mathbb{Z}} \Theta_{2l}(\alpha,\tau), \quad M := M(\alpha,\tau) = \bigoplus_{l\in\mathbb{Z}} \Theta_{2l+1}(\alpha,\tau),$$

then M sends the basis $\{y_k^{\alpha,\tau}\}_{k\in\mathbb{Z}}$ to $\{\mathfrak{z}(x_k^{\alpha,\tau})\}_{k\in\mathbb{Z}}$ and L sends $\{x_k^{\alpha,\tau}\}_{k\in\mathbb{Z}}$ to $\{y_k^{\alpha,\tau}\}_{k\in\mathbb{Z}}$. Thus we have that multiplication by $\mathfrak{z}$ in the basis $\{y_k^{\alpha,\tau}\}$ is given by $\mathcal{C} = LM$, which is a CMV matrix with precisely the Verblunsky coefficients $\mathbf{a}_k(\alpha,\tau)$ as above. ■

We conclude by pointing out that the CMV matrix $\mathcal{C}(\alpha,\tau)$ is periodic if and only if $\phi_0 \in 2\pi\mathbb{Q}$ and there exists $N \geq 1$ such that $\mu_0 N = \mathbf{0}_{\Gamma^*}$. If only the latter holds (i.e., $\phi_0 \notin 2\pi\mathbb{Q}$), then $\mathcal{C}(\alpha,\tau)$ is periodic up to a phase with phase $e^{-iN\phi_0}$.

3.3 A modified basis suited for periodicity

We have seen in the previous subsection that whether the isospectral torus of CMV matrices consists of periodic or almost periodic operators is related to whether there exists $N \geq 1$ such that $\left(\mathfrak{b}_{\zeta_0}\mathfrak{b}_{\overline{\zeta_0}}\right)^N$ can be lifted to a single-valued function on $\overline{\mathbb{C}} \setminus \mathsf{E}$. In this section we will study a basis associated to Blaschke products that have this property, and by definition the corresponding multiplication operator in this basis will be periodic. To fix the notation, let

$$\vec{z} := \{z_0 = 0, z_1, \ldots, z_{n-1}\} \in \mathbb{D}^n$$

and take a point $\zeta_l \in \mathfrak{z}^{-1}(z_l)$ for $l = 0, 1, \ldots, n-1$. Define

$$\mathfrak{B} := \mathfrak{B}_{\vec{z}} = \prod_{l=0}^{n-1} \mathfrak{b}_{\zeta_l} \tag{3.21}$$

and let $\beta := \beta_{\vec{z}}$ denote its character. Our condition on the vector $\vec{z}$ is that β is a half-period (i.e., $2\beta = \mathbf{0}_{\Gamma^*}$).

Remark. The Ahlfors function shows by example that a function as in (3.21) having half-period character indeed exists: recall that w_∞ denotes the Ahlfors function of $\overline{\mathbb{C}} \setminus \mathsf{E}$ and the point ∞. If E has $g+1$ gaps, w_∞ has exactly g zeros in $\mathbb{D}$, say $z_1, \ldots, z_g$, and one zero at ∞. Moreover, $|w_\infty| = 1$ on E and $|w_\infty| < 1$ in $\overline{\mathbb{C}} \setminus \mathsf{E}$. From this it follows that the pullback of zw_∞ to the uniformization, that is, $\mathfrak{w}_\infty := \mathfrak{z}(w_\infty \circ \mathfrak{z})$, is a Blaschke product as in (3.21) with $n = g+1$; furthermore, since zw_∞ is a single-valued function on $\overline{\mathbb{C}} \setminus \mathsf{E}$, the function $\mathfrak{w}_\infty$ has trivial character $\beta_{\vec{z}_\mathsf{E}} = \mathbf{0}_{\Gamma^*}$ (and, in particular, $2\beta_{\vec{z}_\mathsf{E}} = \mathbf{0}_{\Gamma^*}$). See Appendix A for a more detailed discussion.

Denoting by $\mathfrak{z}_l$ the pullback of b_{z_l} to the uniformization, that is,

$$\mathfrak{z}_l := b_{z_l} \circ \mathfrak{z}, \tag{3.22}$$

we see, by the same arguments used to prove (3.11), that there exists a certain phase ϕ_l such that

$$\mathfrak{z}_l(\zeta) = e^{i\phi_l} \frac{\mathfrak{b}(\zeta, \zeta_l)}{\mathfrak{b}(\zeta, \overline{\zeta_l})}. \tag{3.23}$$

Hence the characters of $\mathfrak{b}_{\zeta_l}$ and $\mathfrak{b}_{\overline{\zeta_l}}$ coincide. Let us abbreviate them by μ_l. If we denote $\mathfrak{B}^*(\zeta) := \overline{\mathfrak{B}(\overline{\zeta})} = \prod_{j=0}^{n-1} \mathfrak{b}_{\overline{\zeta_j}}(\zeta)$, then this implies that the character of $\mathfrak{B}\mathfrak{B}^*$ is 2β. By our assumption,

$$2\beta = 2(\mu_0 + \mu_1 + \cdots + \mu_{n-1}) = \mathbf{0}_{\Gamma^*}.$$

This allows us to decompose $H^2(\alpha)$ by iterations of the finite-dimensional subspace

$$\mathcal{K}_{\mathfrak{B}\mathfrak{B}^*}(\alpha) := \operatorname{span}\left\{K^\alpha_{\zeta_0}, K^\alpha_{\overline{\zeta_0}}, K^\alpha_{\zeta_1}, K^\alpha_{\overline{\zeta_1}}, \cdots, K^\alpha_{\zeta_{n-1}}, K^\alpha_{\overline{\zeta_{n-1}}}\right\} = H^2(\alpha) \ominus \mathfrak{B}\mathfrak{B}^* H^2(\alpha), \tag{3.24}$$

without shifting the character. This lack of shift is ultimately what will lead to periodicity up to a phase.

Our strategy is as follows: suppose we have a vector $\vec{z} \in \mathbb{D}^n$ with associated Blaschke product $\mathfrak{B}$ as above having character β a half-period. Similar to CMV matrices, we will have one step comparing symmetric pairs $\zeta_l, \overline{\zeta_l}$ corresponding to shifting from a pole z_l inside the disk to its symmetric point $\hat{z}_l$ outside the disk; this corresponds to the representation

$$\operatorname{span}\left\{K^\alpha_{\zeta_l}, K^\alpha_{\overline{\zeta_l}}\right\} = H^2(\alpha) \ominus \mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}} H^2(\alpha - 2\mu_l), \tag{3.25}$$

which we can iterate to exhaust $\mathcal{K}_{\mathfrak{B}\mathfrak{B}^*}(\alpha)$ as follows:

$$\mathcal{K}_{\mathfrak{B}\mathfrak{B}^*}(\alpha) = H^2(\alpha) \ominus \mathfrak{b}_{\zeta_0}\mathfrak{b}_{\overline{\zeta_0}}\Big(H^2(\alpha - 2\mu_0) \ominus \mathfrak{b}_{\zeta_1}\mathfrak{b}_{\overline{\zeta_1}}\big(H^2(\alpha - 2(\mu_0 + \mu_1)) \ominus \cdots\big)\Big).$$

As in the CMV case, we will be able to act on even steps by a 2×2 block-diagonal operator M to alternate between dual bases respecting the symmetric poles on each two-dimensional subspace $H^2(\alpha) \ominus \mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}}H^2(\alpha - 2\mu_l)$. However—and this is the difference relative to CMV matrices—in the odd steps we wish to pass from the pole $\overline{\zeta_l}$ to the new pole ζ_{l+1}. Of course, since $K^{\alpha}_{\zeta_k} \notin H^2(\alpha) \ominus \mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}}H^2(\alpha - 2\mu_l)$ when $\zeta_k \neq \zeta_l$, something new is required to perform this shift. In this sense, the fundamental lemma allowing for our analysis is the following simple realization:

Lemma 3.6. For any $\alpha \in \Gamma^*$ and with $z_l, z_k \in \mathbb{D}$ and ζ_l, ζ_k as above, we have

$$\frac{\mathfrak{z} - z_k}{\mathfrak{z} - z_l}\mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\overline{\zeta_k}} \in H^2(\alpha) \ominus \mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}}H^2(\alpha - 2\mu_l). \tag{3.26}$$

Proof. For $f \in H^2(\alpha - 2\mu_l)$, we have

$$\frac{1 - \overline{z_k}\mathfrak{z}}{1 - \overline{z_l}\mathfrak{z}}\mathfrak{b}_{\overline{\zeta_l}}f \in H^2(\alpha - \mu_l).$$

Since $\mathfrak{b}_{\zeta_l}$ is unimodular on the boundary, the adjoint in H^2 of multiplication by $\mathfrak{b}_{\zeta_l}$ (and consequently $\mathfrak{z}$) is multiplication by $\mathfrak{b}_{\zeta_l}^{-1}$ (respectively $\mathfrak{z}^{-1}$). Thus, by computing adjoints and applying the reproducing property, one has

$$\left\langle \mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}}f, \frac{\mathfrak{z} - z_k}{\mathfrak{z} - z_l}\mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\overline{\zeta_k}} \right\rangle = \left\langle \frac{1 - \overline{z_k}\mathfrak{z}}{1 - \overline{z_l}\mathfrak{z}}\mathfrak{b}_{\overline{\zeta_l}}f, K^{\alpha-\mu_l}_{\overline{\zeta_k}} \right\rangle = 0,$$

as claimed. ■

Now the way ahead is clear: we apply Lemma 3.6 to expand the shifted reproducing kernel in terms of the reproducing kernels for the previous pole. Let $\alpha \in \Gamma^*$ and $\zeta_l, \zeta_k \in \mathbb{D}$ be as above and define

$$c_1^{\alpha}(\zeta_k, \zeta_l) = e^{-i\phi_l}\frac{z_k - z_l}{1 - |z_l|^2}\frac{K^{\alpha}(\zeta_k, \zeta_l)}{\mathfrak{b}_{\zeta_l}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k, \zeta_k)}, \tag{3.27}$$

$$c_2^{\alpha}(\zeta_k, \zeta_l) = -c_1^{\alpha}(\zeta_k, \zeta_l)\frac{K^{\alpha}(\zeta_k, \overline{\zeta_l})}{K^{\alpha}(\zeta_k, \zeta_l)}, \tag{3.28}$$

where at the removable singularity $\zeta_k = \zeta_l$ we take

$$c_1^{\alpha}(\zeta_l, \zeta_l) = \frac{K^{\alpha}(\zeta_l, \zeta_l)}{\mathfrak{b}_{\overline{\zeta_l}}(\zeta_l)K^{\alpha-\mu_l}(\zeta_l, \zeta_l)}, \tag{3.29}$$

cf. (3.23). Then

Lemma 3.7.

$$\frac{\mathfrak{z}-z_k}{\mathfrak{z}-z_l}\mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\overline{\zeta_k}} = c_1^{\alpha}(\zeta_k,\zeta_l)K^{\alpha}_{\overline{\zeta_l}} + c_2^{\alpha}(\zeta_k,\zeta_l)K^{\alpha}_{\zeta_l}, \tag{3.30}$$

$$\frac{1-\overline{z_k}\mathfrak{z}}{1-\overline{z_l}\mathfrak{z}}\mathfrak{b}_{\overline{\zeta_l}}K^{\alpha-\mu_l}_{\zeta_k} = \overline{c_2^{\alpha}(\zeta_k,\zeta_l)}K^{\alpha}_{\overline{\zeta_l}} + \overline{c_1^{\alpha}(\zeta_k,\zeta_l)}K^{\alpha}_{\zeta_l}. \tag{3.31}$$

Proof. That such a decomposition exists is precisely the content of Lemma 3.6 and (3.25). Since $K^{\alpha}_{\zeta_l}$ is orthogonal to $\mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\zeta_k}$, we find that the coefficient in front of $K^{\alpha}_{\overline{\zeta_l}}$ in (3.30) is given by

$$\frac{\left\langle \frac{\mathfrak{z}-z_k}{\mathfrak{z}-z_l}\mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\overline{\zeta_k}}, \mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\zeta_k}\right\rangle}{\left\langle K^{\alpha}_{\overline{\zeta_l}}, \mathfrak{b}_{\zeta_l}K^{\alpha-\mu_l}_{\zeta_k}\right\rangle}.$$

Using Lemma 3.2, (3.23), and that $\frac{\mathfrak{z}-z_k}{\mathfrak{z}-z_l}K^{\alpha-\mu_l}_{\overline{\zeta_k}}$ vanishes at ζ_k, the numerator can be written as

$$\left\langle \frac{\mathfrak{z}-z_k}{\mathfrak{z}-z_l}K^{\alpha-\mu_l}_{\overline{\zeta_k}}, K^{\alpha-\mu_l}_{\zeta_k}\right\rangle = -\frac{(z_l-z_k)K^{\alpha-\mu_l}(\zeta_l,\overline{\zeta_k})}{K^{\alpha}(\zeta_l,\zeta_l)}\left(\frac{\mathfrak{b}_{\zeta_l}}{\mathfrak{z}-z_l}\right)(\zeta_l)\frac{K^{\alpha}(\zeta_k,\zeta_l)}{\mathfrak{b}_{\zeta_l}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k,\zeta_k)}$$

$$= e^{-i\phi_l}\frac{z_k-z_l}{1-|z_l|^2}\frac{\mathfrak{b}_{\overline{\zeta_l}}(\zeta_l)K^{\alpha-\mu_l}(\zeta_l,\overline{\zeta_k})}{K^{\alpha}(\zeta_l,\zeta_l)}\frac{K^{\alpha}(\zeta_k,\zeta_l)}{\mathfrak{b}_{\zeta_l}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k,\zeta_k)}.$$

We thus arrive at the expression for $c_1^{\alpha}(\zeta_k,\zeta_l)$ in (3.27).

Plugging in ζ_k to (3.30) makes the left-hand side vanish and we deduce that the coefficient in front of $K^{\alpha}_{\zeta_l}$ is given by $c_2^{\alpha}(\zeta_k,\zeta_l)$ as in (3.28). Equation (3.31) follows by applying the operation $f(\zeta)\mapsto\overline{f(\overline{\zeta})}$ to (3.30). ■

Since the decomposition in the previous lemma is not orthogonal, we do not immediately get a nice Pythagorean identity; however, if we define

$$\eta_l^2 := 1-|z_l|^2, \tag{3.32}$$

then we have that

Lemma 3.8.

$$|K^{\alpha}(\zeta_k,\zeta_l)|^2 + |\mathfrak{b}_{\zeta_l}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k,\zeta_k)|^2 = |K^{\alpha}(\zeta_k,\overline{\zeta_l})|^2 + |\mathfrak{b}_{\overline{\zeta_l}}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k,\zeta_k)|^2. \tag{3.33}$$

In particular, for c_1^α, c_2^α defined in (3.27)–(3.28),

$$|c_1^\alpha(\zeta_k,\zeta_l)|^2 - |c_2^\alpha(\zeta_k,\zeta_l)|^2 = \eta_k^2/\eta_l^2. \tag{3.34}$$

Proof. Note that $k_{\zeta_k}^\alpha$ simultaneously lives in both $H^2(\alpha) \ominus \mathfrak{b}_{\zeta_k}\mathfrak{b}_{\zeta_l}H^2(\alpha-\mu_l-\mu_k)$ and $H^2(\alpha) \ominus \mathfrak{b}_{\zeta_k}\mathfrak{b}_{\overline{\zeta_l}}H^2(\alpha-\mu_l-\mu_k)$. Equation (3.33) follows immediately from the Pythagorean identity after expanding $k_{\zeta_k}^\alpha$ in the two orthonormal bases $\{K_{\zeta_l}^\alpha, \mathfrak{b}_{\zeta_l}K_{\zeta_k}^{\alpha-\mu_l}\}$ and $\{K_{\overline{\zeta_l}}^\alpha, \mathfrak{b}_{\overline{\zeta_l}}K_{\zeta_k}^{\alpha-\mu_l}\}$.

It remains to show (3.34). With (3.23) in mind, we see that (3.33) is equivalent to

$$\frac{|K^\alpha(\zeta_k,\zeta_l)|^2 - |K^\alpha(\zeta_k,\overline{\zeta_l})|^2}{|\mathfrak{b}_{\overline{\zeta_l}}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k,\zeta_k)|^2} = 1 - |b_{z_l}(z_k)|^2.$$

A simple calculation shows that

$$1 - |b_{z_l}(z_k)|^2 = \frac{\eta_l^2\eta_k^2}{|1-\overline{z_l}z_k|^2} \tag{3.35}$$

and (3.23) implies

$$\frac{|z_k - z_l|^2}{|\mathfrak{b}_{\zeta_l}(\zeta_k)|^2} = \frac{|1-\overline{z_l}z_k|^2}{|\mathfrak{b}_{\overline{\zeta_l}}(\zeta_k)|^2}.$$

Thus we have

$$|c_1^\alpha(\zeta_k,\zeta_l)|^2 - |c_2^\alpha(\zeta_k,\zeta_l)|^2 = \frac{|1-\overline{z_l}z_k|^2}{\eta_l^4}\frac{|K^\alpha(\zeta_k,\zeta_l)|^2 - |K^\alpha(\zeta_k,\overline{\zeta_l})|^2}{\left|\mathfrak{b}_{\overline{\zeta_l}}(\zeta_k)K^{\alpha-\mu_l}(\zeta_k,\zeta_k)\right|^2} = \frac{\eta_k^2}{\eta_l^2},$$

as claimed. ■

Combining all of the above results, we arrive at

Proposition 3.9.

$$\mathfrak{z}_l^{-1}\begin{bmatrix} \mathfrak{z}-z_l & 0 \\ 0 & \mathfrak{z}-z_k \end{bmatrix}\begin{bmatrix} K_{\zeta_l}^\alpha \\ \mathfrak{b}_{\zeta_l}K_{\overline{\zeta_k}}^{\alpha-\mu_l} \end{bmatrix}$$
$$= \frac{1}{\overline{c_1^\alpha(\zeta_k,\zeta_l)}}\begin{bmatrix} -\overline{c_2^\alpha(\zeta_k,\zeta_l)} & 1 \\ \eta_k^2/\eta_l^2 & c_2^\alpha(\zeta_k,\zeta_l) \end{bmatrix}\begin{bmatrix} 1-\overline{z_l}\mathfrak{z} & 0 \\ 0 & 1-\overline{z_k}\mathfrak{z} \end{bmatrix}\begin{bmatrix} K_{\overline{\zeta_l}}^\alpha \\ \mathfrak{b}_{\overline{\zeta_l}}K_{\zeta_k}^{\alpha-\mu_l} \end{bmatrix}. \tag{3.36}$$

When $\zeta_l = \zeta_k$, this simplifies to

$$\begin{bmatrix} K^{\alpha}_{\zeta_l} \\ \mathfrak{b}_{\zeta_l} K^{\alpha-\mu_l}_{\overline{\zeta_l}} \end{bmatrix} = \frac{1}{\overline{c^{\alpha}_1(\zeta_l,\zeta_l)}} \begin{bmatrix} -\overline{c^{\alpha}_2(\zeta_l,\zeta_l)} & 1 \\ 1 & c^{\alpha}_2(\zeta_l,\zeta_l) \end{bmatrix} \begin{bmatrix} K^{\alpha}_{\overline{\zeta_l}} \\ \mathfrak{b}_{\overline{\zeta_l}} K^{\alpha-\mu_l}_{\zeta_l} \end{bmatrix}. \tag{3.37}$$

Proof. Multiplying the identity (3.36) through by $(1-\overline{z_l}\mathfrak{z})^{-1}$, the first line of the identity is simply (3.31). The second line follows from (3.30), the first line, and an application of (3.34). ■

We now have all the tools to show that multiplication by $\mathfrak{z}$ in $L^2(\alpha)$ has the appropriate structure. Fix $(\alpha,\tau) \in \Gamma^* \times \mathbb{T}$ and define the following quantities

$$A(\alpha,\tau;\zeta_k,\zeta_l) := -e^{-i\tau}\frac{c^{\alpha}_2(\zeta_k,\zeta_l)}{c^{\alpha}_1(\zeta_k,\zeta_l)} = e^{-i\tau}\frac{K^{\alpha}(\zeta_k,\overline{\zeta_l})}{K^{\alpha}(\zeta_k,\zeta_l)}, \tag{3.38}$$

$$R(\alpha;\zeta_k,\zeta_l) := \frac{1}{|c^{\alpha}_1(\zeta_k,\zeta_l)|}\frac{\eta_k}{\eta_l} = \sqrt{1-|A(\alpha,\tau;\zeta_k,\zeta_l)|^2}, \tag{3.39}$$

$$\Theta(\alpha,\tau;\zeta_k,\zeta_l) := \begin{bmatrix} \overline{A(\alpha,\tau;\zeta_k,\zeta_l)} & R(\alpha;\zeta_k,\zeta_l) \\ R(\alpha;\zeta_k,\zeta_l) & -A(\alpha,\tau;\zeta_k,\zeta_l) \end{bmatrix}. \tag{3.40}$$

Define also

$$\omega^{\alpha}_{k,l} := \arg\big(c^{\alpha}_1(\zeta_k,\zeta_l)\big) \tag{3.41}$$

and note that $\omega^{\alpha}_{l,l} = 0$ due to our normalization $\mathfrak{b}_{\overline{\zeta_l}}(\zeta_l) > 0$. Then the content of the previous proposition is that, considering the cases $\zeta_k = \zeta_l$ and $\zeta_k = \zeta_{l+1}$, respectively,

$$\begin{bmatrix} e^{i\tau}K^{\alpha}_{\zeta_l} \\ \mathfrak{b}_{\zeta_l} K^{\alpha-\mu_l}_{\overline{\zeta_l}} \end{bmatrix} = \Theta(\alpha,\tau;\zeta_l,\zeta_l) \begin{bmatrix} K^{\alpha}_{\overline{\zeta_l}} \\ e^{i\tau}\mathfrak{b}_{\overline{\zeta_l}} K^{\alpha-\mu_l}_{\zeta_l} \end{bmatrix} \tag{3.42}$$

and

$$e^{-i\phi_l}\begin{bmatrix} \frac{\mathfrak{z}-z_l}{\eta_l} & 0 \\ 0 & \frac{\mathfrak{z}-z_{l+1}}{\eta_{l+1}} \end{bmatrix} \begin{bmatrix} e^{i\tau}\mathfrak{b}_{\overline{\zeta_l}} K^{\alpha}_{\zeta_l} \\ e^{-i\omega^{\alpha}_{l+1,l}}\mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}} K^{\alpha-\mu_l}_{\overline{\zeta_{l+1}}} \end{bmatrix}$$

$$= \Theta(\alpha,\tau;\zeta_{l+1},\zeta_l) \begin{bmatrix} \frac{1-\overline{z_l}\mathfrak{z}}{\eta_l} & 0 \\ 0 & \frac{1-\overline{z_{l+1}}\mathfrak{z}}{\eta_{l+1}} \end{bmatrix} \begin{bmatrix} \mathfrak{b}_{\zeta_l} K^{\alpha}_{\overline{\zeta_l}} \\ e^{i\tau}e^{i\omega^{\alpha}_{l+1,l}}\mathfrak{b}_{\zeta_l}\mathfrak{b}_{\overline{\zeta_l}} K^{\alpha-\mu_l}_{\zeta_{l+1}} \end{bmatrix}. \tag{3.43}$$

We are finally ready to establish our basis. Let

$$\alpha_l := \alpha_{l-1} - 2\mu_l, \quad \alpha_0 := \alpha - \mu_0, \quad \vartheta_l^\alpha := \sum_{j=0}^{l} \omega_{j+1,j}^{\alpha_j+\mu_j} + \phi_j,$$

and

$$\mathfrak{B}_l := \prod_{j=1}^{l} \mathfrak{b}_{\zeta_j}, \quad \mathfrak{B}_l^* := \prod_{j=1}^{l} \mathfrak{b}_{\overline{\zeta_j}}, \quad \mathfrak{B}_0 = \mathfrak{B}_0^* = 1.$$

Taking as convention $\vartheta_{-1}^\alpha = 0$, $\mathfrak{B}_{-1} = \mathfrak{b}_{\zeta_0}^{-1}$, $\mathfrak{B}_{-1}^* = \mathfrak{b}_{\overline{\zeta_0}}^{-1}$, and $\zeta_n = \zeta_0$, we define for $0 \leq l \leq n-1$ the functions

$$x_{2l}^{\alpha,\tau} := e^{-i\vartheta_{l-1}^\alpha} \mathfrak{b}_{\overline{\zeta_0}} \mathfrak{B}_l \mathfrak{B}_{l-1}^* K_{\overline{\zeta_l}}^{\alpha_{l-1}-\mu_l}, \quad x_{2l+1}^{\alpha,\tau} := e^{i\tau} e^{i\vartheta_l^\alpha} \mathfrak{b}_{\overline{\zeta_0}} \mathfrak{B}_l \mathfrak{B}_l^* K_{\zeta_{l+1}}^{\alpha_l}, \tag{3.44}$$

$$y_{2l}^{\alpha,\tau} := e^{i\tau} e^{i\vartheta_{l-1}^\alpha} \mathfrak{b}_{\zeta_0} \mathfrak{B}_l^* \mathfrak{B}_{l-1} K_{\zeta_l}^{\alpha_{l-1}-\mu_l}, \quad y_{2l+1}^{\alpha,\tau} := e^{-i\vartheta_l^\alpha} \mathfrak{b}_{\zeta_0} \mathfrak{B}_l^* \mathfrak{B}_l K_{\overline{\zeta_{l+1}}}^{\alpha_l}. \tag{3.45}$$

In analog to the CMV case (3.16), they form two different bases of the $2n$-dimensional subspace $\mathcal{K}_{\mathfrak{B}\mathfrak{B}^*}$; cf. (3.24). Letting $p = 2n$, we extend this family of functions (for $j \in \mathbb{Z}$) by

$$x_{2l+jp}^{\alpha,\tau} = (\mathfrak{B}\mathfrak{B}^*)^j e^{-ij\vartheta_{n-1}^\alpha} x_{2l}^{\alpha,\tau}, \quad x_{2l+1+jp}^{\alpha,\tau} = (\mathfrak{B}\mathfrak{B}^*)^j e^{ij\vartheta_{n-1}^\alpha} x_{2l+1}^{\alpha,\tau}, \tag{3.46}$$

$$y_{2l+jp}^{\alpha,\tau} = (\mathfrak{B}\mathfrak{B}^*)^j e^{ij\vartheta_{n-1}^\alpha} y_{2l}^{\alpha,\tau}, \quad y_{2l+1+jp}^{\alpha,\tau} = (\mathfrak{B}\mathfrak{B}^*)^j e^{-ij\vartheta_{n-1}^\alpha} y_{2l+1}^{\alpha,\tau}. \tag{3.47}$$

By iterating the exhaustion (3.24), it isn't difficult to see that the systems of functions $\{x_k^{\alpha,\tau}\}_{k\in I}$ and $\{y_k^{\alpha,\tau}\}_{k\in I}$ each form an orthonormal basis of $H^2(\alpha)$ for $I = \mathbb{N}$. By [9, Lemma 3.5], it follows that they also form a basis of $L^2(\alpha)$ when $I = \mathbb{Z}$.

Let us comment on the meaning of the unimodular constants appearing in the definitions above. First of all, we can choose the unimodular constant freely in the normalization of the basis functions $\{x_0, x_1\}$. This explains the meaning of the additional parameter τ. Once this normalization is fixed, the normalization of the following basis functions is already determined: comparing (3.36) and (3.43), we see that—apart from the additional parameter τ—the main difference between the constant matrix on the right-hand side of (3.36) and the matrix $\Theta(\alpha, \tau; \zeta_l, \zeta_{l+1})$ in (3.43) is that the latter has positive off-diagonal entries. This has been achieved by adding the phase $e^{i\omega_{l+1,l}^\alpha}$ to the reproducing kernels. These phases accumulate with each step as the phases $e^{i\vartheta_l^\alpha}$.

Define now the periodic up to a phase Verblunsky coefficients $\{a_k(\alpha, \tau; \vec{z})\}$ by

$$\begin{aligned} a_{2l-1}(\alpha, \tau; \vec{z}) &:= e^{-2i\vartheta_{l-1}^\alpha} A(\alpha_{l-1}, \tau; \zeta_l, \zeta_l), \\ a_{2l}(\alpha, \tau; \vec{z}) &:= e^{-i(\phi_l + 2\vartheta_{l-1}^\alpha)} A(\alpha_{l-1} - \mu_l, \tau; \zeta_{l+1}, \zeta_l), \end{aligned} \qquad 0 \leq l \leq n-1, \tag{3.48}$$

and

$$a_{m+jp}(\alpha,\tau;\vec{z}) = e^{-2ij\vartheta^{\alpha}_{n-1}} a_m(\alpha,\tau;\vec{z}), \quad -1 \le m \le 2n-2, \;\; j \in \mathbb{Z}, \tag{3.49}$$

$$\rho_k(\alpha;\vec{z}) := \sqrt{1 - |a_k(\alpha,\tau;\vec{z})|^2}, \quad k \in \mathbb{Z}. \tag{3.50}$$

Then multiplication by $\mathfrak{z}$ in our modified basis is represented by an MCMV matrix with the above parameters:

Theorem 3.10. Let $C = C(\alpha,\tau;\vec{z})$ be the periodic up to a phase CMV matrix with Verblunsky coefficients $a_k(\alpha,\tau;\vec{z})$, and let D_0 be the $2n$-periodic diagonal matrix given by (1.4). Then with respect to the basis $\{y_k^{\alpha,\tau}\}_{k\in\mathbb{Z}}$ of $L^2(\alpha)$, multiplication by $\mathfrak{z}$ is represented by the MCMV matrix $b_{-D_0}(C)$.

Proof. Denote by

$$\Theta_k(\alpha,\tau;\vec{z}) := \begin{bmatrix} \overline{a_k(\alpha,\tau;\vec{z})} & \rho_k(\alpha;\vec{z}) \\ \rho_k(\alpha;\vec{z}) & -a_k(\alpha,\tau;\vec{z}) \end{bmatrix}. \tag{3.51}$$

We use liberally the following two simple observations: that diagonal matrices commute, and that, for $a \in \mathbb{D}$, $\rho = \sqrt{1-|a|^2}$, and $\theta_0, \theta_1 \in \mathbb{T}$,

$$\begin{bmatrix} e^{-i\theta_0} & 0 \\ 0 & e^{i\theta_1} \end{bmatrix} \begin{bmatrix} \overline{a} & \rho \\ \rho & -a \end{bmatrix} \begin{bmatrix} e^{-i\theta_1} & 0 \\ 0 & e^{i\theta_0} \end{bmatrix} = \begin{bmatrix} e^{-i(\theta_0+\theta_1)}\overline{a} & \rho \\ \rho & -e^{i(\theta_0+\theta_1)}a \end{bmatrix}.$$

Let $0 \le l \le n-1$. With the above facts in hand, it is then clear that (3.42) is equivalent to

$$\begin{bmatrix} e^{i\tau} e^{i\vartheta^{\alpha}_{l-1}} K^{\alpha_{l-1}}_{\zeta_l} \\ e^{-i\vartheta^{\alpha}_{l-1}} \mathfrak{b}_{\zeta_l} K^{\alpha_{l-1}-\mu_l}_{\overline{\zeta_l}} \end{bmatrix} = \Theta_{2l-1}(\alpha,\tau;\vec{z}) \begin{bmatrix} e^{-i\vartheta^{\alpha}_{l-1}} K^{\alpha_{l-1}}_{\overline{\zeta_l}} \\ e^{i\tau} e^{i\vartheta^{\alpha}_{l-1}} \mathfrak{b}_{\overline{\zeta_l}} K^{\alpha_{l-1}-\mu_l}_{\zeta_l} \end{bmatrix}.$$

Multiplying both sides by $\mathfrak{b}_{\zeta_0}\mathfrak{B}_{l-1}\mathfrak{B}^*_{l-1} = e^{-i\phi_0}\mathfrak{z}\mathfrak{b}_{\overline{\zeta_0}}\mathfrak{B}_{l-1}\mathfrak{B}^*_{l-1}$, we get that

$$e^{-i\phi_0}\mathfrak{z} \begin{bmatrix} x^{\alpha,\tau}_{2l-1} \\ x^{\alpha,\tau}_{2l} \end{bmatrix} = \Theta_{2l-1}(\alpha,\tau;\vec{z}) \begin{bmatrix} y^{\alpha,\tau}_{2l-1} \\ y^{\alpha,\tau}_{2l} \end{bmatrix}. \tag{3.52}$$

Since in the context of (3.43) we have $e^{i\phi_l} = e^{i\phi_l} I$ (where I is the 2×2 identity matrix), that equation can also be written as

$$\begin{bmatrix} \frac{\mathfrak{z}-z_l}{\eta_l} & 0 \\ 0 & \frac{\mathfrak{z}-z_{l+1}}{\eta_{l+1}} \end{bmatrix} \begin{bmatrix} e^{i\tau} e^{i\vartheta_{l-1}^\alpha} \mathfrak{b}_{\overline{\zeta_l}} K_{\zeta_l}^{\alpha_{l-1}-\mu_l} \\ e^{-i\vartheta_l^\alpha} \mathfrak{b}_{\zeta_l} \mathfrak{b}_{\overline{\zeta_l}} K_{\overline{\zeta_{l+1}}}^{\alpha_l} \end{bmatrix} = \Theta_{2l}(\alpha,\tau;\vec{z}) \begin{bmatrix} \frac{1-\overline{z}_l\mathfrak{z}}{\eta_l} & 0 \\ 0 & \frac{1-\overline{z_{l+1}}\mathfrak{z}}{\eta_{l+1}} \end{bmatrix} \begin{bmatrix} e^{-i\vartheta_{l-1}^\alpha} \mathfrak{b}_{\zeta_l} K_{\overline{\zeta_l}}^{\alpha_{l-1}-\mu_l} \\ e^{i\tau} e^{i\vartheta_l^\alpha} \mathfrak{b}_{\zeta_l} \mathfrak{b}_{\overline{\zeta_l}} K_{\zeta_{l+1}}^{\alpha_l} \end{bmatrix}.$$

Multiplying both sides by $\mathfrak{b}_{\overline{\zeta_0}} \mathfrak{B}_{l-1} \mathfrak{B}_{l-1}^* = e^{i\phi_0} \mathfrak{z}^{-1} \mathfrak{b}_{\zeta_0} \mathfrak{B}_{l-1} \mathfrak{B}_{l-1}^*$ and rearranging yields

$$e^{i\phi_0} \mathfrak{z}^{-1} \begin{bmatrix} \frac{\mathfrak{z}-z_l}{\eta_l} & 0 \\ 0 & \frac{\mathfrak{z}-z_{l+1}}{\eta_{l+1}} \end{bmatrix} \begin{bmatrix} Y_{2l}^{\alpha,\tau} \\ Y_{2l+1}^{\alpha,\tau} \end{bmatrix} = \Theta_{2l}(\alpha,\tau;\vec{z}) \begin{bmatrix} \frac{1-\overline{z}_l\mathfrak{z}}{\eta_l} & 0 \\ 0 & \frac{1-\overline{z_{l+1}}\mathfrak{z}}{\eta_{l+1}} \end{bmatrix} \begin{bmatrix} X_{2l}^{\alpha,\tau} \\ X_{2l+1}^{\alpha,\tau} \end{bmatrix}. \tag{3.53}$$

Extending to all l follows similarly from the definitions.

Denote now by $D_0 := D_0(\vec{z})$ the $2n$-periodic diagonal matrix in (1.4), let $\eta_{D_0} = \sqrt{1 - D_0 D_0^*}$, and fix

$$L := \bigoplus_{l\in\mathbb{Z}} \Theta_{2l}(\alpha,\tau;\vec{z}), \quad M := \bigoplus_{l\in\mathbb{Z}} \Theta_{2l+1}(\alpha,\tau;\vec{z}),$$

where Θ_k acts on the two-dimensional subspace $\{\delta_k, \delta_{k+1}\}$. Combining the statements (3.52) and (3.53) above, we have shown the following:

$$L(1 - \mathfrak{z}D_0^*)\eta_{D_0}^{-1} M \vec{Y}^{\alpha,\tau} = \eta_{D_0}^{-1}(\mathfrak{z} - D_0)\vec{Y}^{\alpha,\tau},$$

where $\vec{Y}^{\alpha,\tau}$ is shorthand notation for the vector $(Y_k^{\alpha,\tau})_{k\in\mathbb{Z}}$. Since the operators $\eta_{D_0}^{-1}$, $\mathfrak{z} - D_0$, and $1 - \mathfrak{z}{D_0}^*$ commute with M (for they are orthogonal sums of multiples of I along the odd terms), taking $C = LM$ we have

$$\left(1 - \eta_{D_0}(\mathfrak{z} - D_0)^{-1} C (1 - \mathfrak{z}{D_0}^*)\eta_{D_0}^{-1}\right)\vec{Y}^{\alpha,\tau} = 0, \tag{3.54}$$

which can be rearranged as

$$\left(\mathfrak{z} - \eta_{D_0}(1 + C{D_0}^*)^{-1}(C + D_0)\eta_{D_0}^{-1}\right)\vec{Y}^{\alpha,\tau} = 0.$$

Thus, in the basis $\{Y_k^{\alpha,\tau}\}$, multiplication by $\mathfrak{z}$ is given by $b_{-D_0}(C)$. ■

Of course, (3.36) and (3.37) in combination with the exhaustion (3.24) without shifted character imply a transfer matrix relation in terms of the reproducing kernels.

To see explicitly this relation, first note that, denoting as shorthand $a_k = a_k(\alpha, \tau; \vec{z})$, $\rho_k = \rho_k(\alpha; \vec{z})$, and $U(a_k)$ as in (1.10), we can rewrite (3.52)–(3.53) in the following way:

$$\begin{bmatrix} e^{-i\phi_0}\mathfrak{z} & 0 \\ 0 & 1 \end{bmatrix} U(a_{2l+1}) \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi_0}\mathfrak{z}^{-1} \end{bmatrix} \begin{bmatrix} x^{\alpha,\tau}_{2l+2} \\ y^{\alpha,\tau}_{2l+2} \end{bmatrix} = \begin{bmatrix} y^{\alpha,\tau}_{2l+1} \\ x^{\alpha,\tau}_{2l+1} \end{bmatrix}, \tag{3.55}$$

$$\frac{1-\overline{z_{l+1}}\mathfrak{z}}{1-\overline{z_l}\mathfrak{z}} \frac{\eta_l}{\eta_{l+1}} \begin{bmatrix} 1 & 0 \\ 0 & \mathfrak{z}_l^{-1} \end{bmatrix} \begin{bmatrix} e^{i\phi_0}\mathfrak{z}^{-1} & 0 \\ 0 & 1 \end{bmatrix} U(a_{2l}) \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\phi_0}\mathfrak{z} \end{bmatrix} \begin{bmatrix} \mathfrak{z}_{l+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y^{\alpha,\tau}_{2l+1} \\ x^{\alpha,\tau}_{2l+1} \end{bmatrix} = \begin{bmatrix} x^{\alpha,\tau}_{2l} \\ y^{\alpha,\tau}_{2l} \end{bmatrix}. \tag{3.56}$$

Again using the notation that $z_n = z_0 = 0$, $\eta_n = \eta_0$, etc., and denoting

$$T(z; \vec{z}, \{a_k\}, \vartheta^\alpha_{n-1}) := U(a_0) \begin{bmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_1) \begin{bmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_2) \begin{bmatrix} b_{z_2}(z) & 0 \\ 0 & 1 \end{bmatrix} \cdots$$
$$\cdots U(a_{2n-1}) \begin{bmatrix} b_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} U(a_{2n}) \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} U(a_{2n+1}) \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^{-i\vartheta^\alpha_{n-1}} & 0 \\ 0 & e^{i\vartheta^\alpha_{n-1}} \end{bmatrix} \tag{3.57}$$

and

$$B(z) = z \prod_{j=1}^{n-1} b_{z_j}(z) = \sqrt{\det T(z; \vec{z}, \{a_k\})}, \tag{3.58}$$

we arrive at the following monodromy relation:

Theorem 3.11. The reproducing kernels satisfy the phased monodromy relation

$$\frac{1}{B(\mathfrak{z})} T\big(\mathfrak{z}; \vec{z}, \{a_k\}, \vartheta^\alpha_{n-1}\big) \begin{bmatrix} x^{\alpha,\tau}_0 \\ e^{i\phi_0} y^{\alpha,\tau}_0 \end{bmatrix} = (\mathfrak{B}\mathfrak{B}^*)^{-1} \begin{bmatrix} x^{\alpha,\tau}_0 \\ e^{i\phi_0} y^{\alpha,\tau}_0 \end{bmatrix}. \tag{3.59}$$

Proof. This follows from iterating (3.55)–(3.56) over a full period of size $p = 2n$, since the multiplier terms telescope and

$$x^{\alpha,\tau}_{2n} = \mathfrak{B}\mathfrak{B}^* e^{-i\vartheta^\alpha_{n-1}} x^{\alpha,\tau}_0, \quad y^{\alpha,\tau}_{2n} = \mathfrak{B}\mathfrak{B}^* e^{i\vartheta^\alpha_{n-1}} y^{\alpha,\tau}_0.$$

■

Remark. In terms of projective lines, (3.59) is, up to a phase, precisely the relation (2.19) for the Schur functions (3.19).

We are now ready to give a detailed explanation for introducing the matrix $\Lambda(\vartheta)$ in Definition 1.1. In our extension (3.47) of the vectors $\{y^{\alpha,\tau}_l\}_{l=0}^{2n-1}$ to a basis of $L^2(\alpha)$, we include the phase $e^{i\vartheta^\alpha_{n-1}}$ in order to represent the multiplication operator by $\mathfrak{z}$ as an operator Möbius transform of a CMV matrix $C(\alpha,\tau;\vec{z})$; specifically, the phase is needed to make the off-diagonal entries of $\Theta_l(\alpha,\tau;\vec{z})$ in its LM factorization positive. The price we pay is that the corresponding matrix is merely periodic up to a phase. If we instead chose the extension of multiplying by $(\mathfrak{B}\mathfrak{B}^*)^j$ without the phase, then the corresponding operator would be periodic. This alternative basis, say $\{Y^{\alpha,\tau}_{\mathrm{per},l}\}$, is related to $\{Y^{\alpha,\tau}_l\}$ in the following way:

$$\vec{Y}^{\alpha,\tau}_{\mathrm{per}} = \Lambda\left(\vartheta^\alpha_{n-1}\right)^* \vec{Y}^{\alpha,\tau}. \tag{3.60}$$

Thus the conjugacy by $\Lambda(\vartheta)$ in Definition 1.1 allows us to relate the periodic representation of multiplication by $\mathfrak{z}$ to an honest CMV matrix.

A particular goal of our construction is to obtain a map from $\Gamma^* \times \mathbb{T}$ to $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E},\vec{z},\lambda_*)$, see Corollary 3.13 below. For the final step we need a lemma whose full statement will not be used but we believe is of interest in its own right. It will enable us to identify the sign of Δ_A in each gap. As a by-product, we can show that Δ_{E} having the same sign in each gap is a consequence of the character of the Ahlfors function being trivial rather than merely being a half-period. First, we need to introduce some notation. Let $g+1$ be the number of gaps of E and let λ_j^-,λ_j^+ (for $0 \le j \le g$) denote the gap edges such that one can pass from λ_j^- to λ_j^+ by traversing the j-th gap counterclockwise. Moreover, let $[\lambda_j^-,\lambda_j^+]$ denote the closed arcs induced by this order and suppose that λ_* belongs to the gap $(\lambda_0^-,\lambda_0^+)$. Finally, let $\{\tilde{\gamma}_j\}_{j=1}^g$ be a set of generating loops for $\pi_1(\overline{\mathbb{C}}\setminus\mathsf{E})$ passing through the zeroth and j-th gaps and let $\{\gamma_j\}_{j=1}^g$ denote the corresponding elements in Γ; cf. [32, Section 9.6].

Lemma 3.12. With the notation from above, suppose that $\lambda \in (\lambda_j^-,\lambda_j^+)$ and pick a point $\zeta \in \mathfrak{z}^{-1}(\lambda)$. Then we have that

$$\mathfrak{B}(\zeta)\mathfrak{B}^*(\zeta) = e^{-i\beta(\gamma_j)}|\mathfrak{B}(\zeta)|^2. \tag{3.61}$$

In particular, $\mathfrak{B}(\zeta)\mathfrak{B}^*(\zeta)$ is positive in the gap containing λ_*, and it is positive in all gaps if and only if $\beta = \mathbf{0}_{\Gamma^*}$.

Proof. Recall that the action of $\tilde{\gamma}_j$ corresponds to a reflection with respect to the gap $(\lambda_j^-,\lambda_j^+)$ and the gap $(\lambda_0^-,\lambda_0^+)$. Therefore, since $(-1,1)$ is mapped onto $(\lambda_0^-,\lambda_0^+)$, we see that for $\lambda \in (\lambda_j^-,\lambda_j^+)$ and $\zeta \in \mathfrak{z}^{-1}(\lambda)$, we have that $\overline{\zeta} = \gamma_j(\zeta)$; see also [32, Theorem 9.6.3]. Thus, by means of our normalization,

$$\mathfrak{b}(\zeta,\overline{\zeta_1}) = \overline{\mathfrak{b}(\overline{\zeta},\zeta_1)} = e^{-i\mu_{\zeta_1}(\gamma_j)}\overline{\mathfrak{b}(\zeta,\zeta_1)},$$

and hence

$$\mathfrak{b}(\zeta,\zeta_1)\mathfrak{b}(\zeta,\overline{\zeta_1}) = e^{-i\mu_{\zeta_1}(\gamma_j)}|\mathfrak{b}(\zeta,\overline{\zeta_1})|^2,$$

which proves (3.61). Since the gap $(\lambda_0^-,\lambda_0^+)$ is mapped onto $(-1,1)$, it is invariant under complex conjugation. This shows by the same line of arguments the positivity statement for that gap. ■

Remark. Note that since $\{\gamma_j\}_{j=1}^g$ generate Γ, it follows that β being a half-period is equivalent to the statement that $e^{i\beta(\gamma_j)} = \pm 1$ for $1 \le j \le g$.

We are now finally ready to prove the following corollary of our construction:

Corollary 3.13. Let C and D_0 be as in Theorem 3.10, and set

$$A := A(\alpha,\tau) = \Lambda(\vartheta_{n-1}^\alpha)^* b_{-D_0}\big(C(\alpha,\tau)\big)\Lambda(\vartheta_{n-1}^\alpha). \tag{3.61}$$

Then $A \in \mathbb{A}_{\mathrm{per}}(\vec{z})$ and $\sigma(A) = \mathsf{E}$. Moreover, with Δ_A as in (1.12),

$$\Delta_A \circ \mathfrak{z} = \mathfrak{B}\mathfrak{B}^* + \frac{1}{\mathfrak{B}\mathfrak{B}^*} \tag{3.63}$$

and

$$\Delta_A(A) = S^{2n} + S^{-2n}. \tag{3.64}$$

In particular, $A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E},\vec{z},\lambda_*)$. In the special case $\vec{z} = \vec{z}_{\mathsf{E}}$, we have additionally that $\Delta_A = \Delta_{\mathsf{E}}$.

Proof. The first statement follows from the discussion above. The fact that $\sigma(A) = \mathsf{E}$ is clear since A is the matrix of multiplication by $\mathfrak{z}$. If we set $\tilde{T} = B^{-1}T$, then Theorem 3.11 states that $(\mathfrak{B}\mathfrak{B}^*)^{-1}$ is an eigenvalue of $\tilde{T}$. As $\det \tilde{T} = 1$, we thus obtain (3.63). Lemma 3.12 shows that $\Delta_A(\lambda_*) > 0$. Finally, (3.64) is a direct consequence of the fact that multiplication by $\mathfrak{B}\mathfrak{B}^*$ corresponds to the action of S^{2n} in the basis $\{Y_{\mathrm{per},l}^{\alpha,\tau}\}_{l\in\mathbb{Z}}$. ■

Remark. Formula (3.63) reiterates the relevance of our additional normalization $\Delta_A(\lambda_*) > 0$ in (1.17). In general, in our periodic extension

$$\big(Y_{\mathrm{per}}^{\alpha,\tau}\big)_{k+p} = \mathfrak{B}\mathfrak{B}^*\big(Y_{\mathrm{per}}^{\alpha,\tau}\big)_k \tag{3.65}$$

one could just as well have chosen $c\mathfrak{B}\mathfrak{B}^*$ for any unimodular constant c to obtain a periodic operator. But in order for Δ_A to be real, it is crucial that

$$\overline{\mathfrak{B}(\bar{\zeta})\mathfrak{B}^*(\bar{\zeta})} = \mathfrak{B}(\zeta)\mathfrak{B}^*(\zeta). \tag{3.66}$$

To maintain this property, we see that c must be real; hence, $c = \pm 1$ are the only possible choices for c. The choice of $c = 1$ corresponds to the aforementioned normalization.

4 Direct Spectral Theory

In the previous section, we saw that the functional model developed by Peherstorfer and Yuditskii to represent finite-gap almost-periodic CMV matrices has corresponding representations as periodic MCMV matrices. In this section, we develop the necessary tools to address the converse: that any periodic MCMV matrix arises from such a functional model.

For periodic CMV matrices $\mathcal{C} \in \mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$, the bijective nature of this correspondence is by now classical (cf. (4.4) below); we recall the elements of its construction in Section 4.1. A key component of this correspondence is a set of spectral data associated to the one-sided restriction $\mathcal{C}_+$, called the *divisor* or *Dirichlet data*, which, together with the discriminant, allows one to uniquely recover the spectral measure and hence the operator $\mathcal{C}_+$. We will adapt this construction to periodic MCMV matrices in Section 4.2, culminating in the uniqueness statement Proposition 4.6. Finally, Section 4.3 explores the block structure (1.7) of general MCMV matrices and its invariance under certain Möbius transformations.

4.1 The isospectral torus of periodic CMV matrices

We briefly recall aspects of the spectral theory for periodic CMV matrices as can be found, for instance, in [31, Chapter 11]. Let $\{a_k\}_{k\in\mathbb{Z}}$ be a periodic sequence with even period $p = 2n$ and let $\mathcal{C}$ be the corresponding whole-line CMV matrix. If we define the discriminant by

$$\Delta_{\mathcal{C}}(z) = \operatorname{tr}\left(U(a_0)\begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} U(a_1)\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{z} \end{bmatrix} \cdots U(a_{p-2})\begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} U(a_{p-1})\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{z} \end{bmatrix}\right),$$

then the spectrum of $\mathcal{C}$ is given by

$$\mathsf{E} := \sigma(\mathcal{C}) = \Delta_{\mathcal{C}}^{-1}\big([-2, 2]\big).$$

This spectrum is purely absolutely continuous and of multiplicity two. Moreover, there are p critical points $\{c_i\}_{i=1}^p$ on $\partial\mathbb{D}$ (i.e., zeros of $\Delta_{\mathcal{C}}'$) which all satisfy $|\Delta_{\mathcal{C}}(c_i)| \geq 2$.

Therefore, the set $\Delta_{\mathcal{C}}^{-1}\big((-2,2)\big)$ can be partitioned into p non-intersecting open arcs. The connected components of the complement of this set on $\partial\mathbb{D}$ are called the gaps. If a gap consists only of a single point (which is the case if $|\Delta_{\mathcal{C}}(c_i)| = 2$), we refer to it as a closed gap. Otherwise, as before, the gap is called open. Let $g+1$ denote the number of open gaps and let us fix some labeling of the open gaps as in the end of Section 3.

As we have seen, the isospectral torus $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ is a $g+1$-dimensional torus. In particular, the spectrum does not uniquely determine the operator $\mathcal{C}$. In order to get the full spectral data to solve the inverse problem, we consider the half-line operator $\mathcal{C}_+$ with spectral measure ν. One can show there are explicit rational functions u, v such that, for a suitable branch of the square root, the associated Caratheodory function is given by

$$F_\nu(z) = \frac{v(z) + \sqrt{\Delta_{\mathcal{C}}^2(z) - 4}}{u(z)}. \tag{4.1}$$

It is known that u has precisely one zero in each gap of E, and if $u(z) = 0$, then $\sqrt{\Delta_{\mathcal{C}}^2(z) - 4}$ is either $-v(z)$ or $v(z)$. A zero of u for which the numerator in (4.1) does not vanish corresponds to an eigenvalue of $\mathcal{C}_+$. Note that for closed gaps, the numerator always vanishes. Let $\{x_j\}_{j=0}^{g}$ be the set of all zeros of u, which lie in open gaps, and let us write $(x_j, 1)$ if x_j is an eigenvalue of $\mathcal{C}_+$ and $(x_j, -1)$ otherwise. Then the spectrum together with the divisor $D = \{(x_j, \varepsilon_j)\}_{j=0}^{g}$ form the full spectral data and determine $\mathcal{C}$ completely. In fact, if we define

$$\mathcal{D}(\mathsf{E}) = \left\{(x_j, \varepsilon_j) : \ x_j \in \left[\lambda_j^-, \lambda_j^+\right], \ \varepsilon_j = \pm 1, \ 0 \le j \le g\right\} / \sim \tag{4.2}$$

with the identifications $(\lambda_j^\pm, -1) \sim (\lambda_j^\pm, 1)$, then $\mathcal{D}(\mathsf{E})$ equipped with the product topology of circles is homeomorphic to $\Gamma^* \times \mathbb{T}$ and hence also to $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$. Inspired by results in the framework of Jacobi matrices [35], this has been generalized in [25] to the much more general class of Parreau–Widom sets $\mathsf{E} \subset \partial\mathbb{D}$ satisfying the Direct Cauchy Theorem. The homeomorphism

$$\mathfrak{A} : \mathcal{D}(\mathsf{E}) \to \Gamma^* \times \mathbb{T} \tag{4.3}$$

is called the generalized *Abel map*; it is the map that completes the following diagram:

$$\begin{array}{ccc} \mathcal{T}_{\mathrm{CMV}}(\mathsf{E}) & \xleftarrow{\quad \text{Thm } 3.4 \quad} & \Gamma^* \times \mathbb{T} \\ & \searrow \qquad\qquad \nearrow_{\mathfrak{A}} & \\ & \mathcal{D}(\mathsf{E}) & \end{array} \tag{4.4}$$

In the finitely connected setting, the Abel map is well understood (see, e.g., [19, 29]). The connection to spectral theory of Jacobi matrices goes back to Akhiezer [2]; see also [3, 18, 20].

4.2 Spectral theory for periodic MCMV matrices

In this section we will perform a spectral analysis for MCMV matrices that are periodic up to a phase. The spectral data will be given by the discriminant and zeros of a certain function that is explicitly defined in terms of the ORFs. For CMV matrices it is easy to see that the leading coefficient of the discriminant is positive, and the discriminant is always of maximal degree. For MCMV matrices, however, the situation is more involved; we shall clarify the degree issue in Lemma 4.3.

Let us start by introducing the concept of a half-line MCMV matrix. This will be the main object of study in the present section. Recall that given a vector $\vec{z} = \{z_0, \ldots, z_{n-1}\}$, an MCMV matrix $A = A(\{a_k\}, \vartheta; \vec{z}) \in \mathbb{A}(\vec{z})$ is defined by

$$A = \Lambda(\vartheta)^* b_{-D_0}(C)\Lambda(\vartheta),$$

where D_0 is a diagonal matrix depending only on $\vec{z}$, $C = C(\{a_k\})$ is the CMV matrix associated to $\{a_k\}_{k\in\mathbb{Z}}$, and $\Lambda(\vartheta)$ is a certain diagonal matrix; cf. Definition 1.1. Given $A \in \mathbb{A}_{\text{per}}(\vec{z})$, we define the half-line MCMV matrix A_+ by

$$A_+ := b_{-D_+}(C_+), \tag{4.4}$$

where $C_+ = C_+(\{a_k\}_{k\in\mathbb{N}})$ is the half-line CMV matrix with Verblunsky coefficients $\{a_k\}_{k\in\mathbb{N}}$ and D_+ is the diagonal operator $D_+ = \text{diag}\{z_0, z_1, z_1, \ldots, z_{n-1}, z_0, z_0, z_1, \ldots\}$. We point out that when $A \in \mathbb{A}_{\text{per}}(\vec{z})$, the corresponding representing CMV matrix C is periodic up to a phase. To see this, let us first observe that a CMV matrix is periodic up to a phase with phase $e^{-2i\vartheta}$ if and only if $\Lambda(\vartheta)^* C\Lambda(\vartheta)$ is periodic. By definition of A and with the expression (B.4) for the inverse of an operator Möbius transform in mind, we see that

$$C = b_{D_0}\big(\Lambda(\vartheta)A\Lambda(\vartheta)^*\big).$$

Using that D_0 and $\Lambda(\vartheta)$ are diagonal and therefore commute, it follows that $\Lambda(\vartheta)^* C\Lambda(\vartheta) = b_{D_0}(A)$. Since A and D_0 are periodic, so is $b_{D_0}(A)$, and thus C is periodic up to a phase.

Due to [37, Theorem 5.4], the measure ν of orthogonality for the family of orthonormal rational functions related to the poles $\{z_0, z_1, z_1, \ldots, z_{n-1}, z_0, z_0, z_1, \ldots\}$ and Verblunsky coefficients $\{a_k\}_{k\in\mathbb{N}}$ is precisely the spectral measure for A_+ (and the cyclic vector δ_0). The main result of this section will be an explicit expression for the associated Caratheodory function F_ν, analogous to (4.1).

Let $p = 2n$ and fix a sequence $\{a_j\}_{j=0}^{p-1}$ and a phase ϑ. Let T be the monodromy matrix defined by (2.18). Moreover, let us define

$$M_\vartheta(z) := M(z)\begin{bmatrix} \cos\vartheta & i\sin\vartheta \\ i\sin\vartheta & \cos\vartheta \end{bmatrix}, \quad W_\vartheta(z) := \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix} W(z), \tag{4.6}$$

where M and $W = W^{(0)}$ are the matrices defined by (2.11) and (2.15), respectively. Keeping in mind that $b_{z_0}(z) = z$, the shape of M_ϑ is such that

$$Y_0 \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} T(z) = M_\vartheta(z) Y_0 \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.7}$$

This follows from the commutant relation

$$Y_0 \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix} = \begin{bmatrix} \cos\vartheta & i\sin\vartheta \\ i\sin\vartheta & \cos\vartheta \end{bmatrix} Y_0. \tag{4.8}$$

Hence, if we consider the sequence $\{a_j\}_{j\in\mathbb{N}}$, which is obtained by extending $\{a_j\}_{j=0}^{p-1}$ in such a way that $a_{j+kp} = e^{-2ik\vartheta} a_j$, then due to Corollary 2.3 and (1.28) the associated Caratheodory function satisfies

$$\begin{bmatrix} F_\nu(z) \\ 1 \end{bmatrix} \sim M_\vartheta(z) \begin{bmatrix} F_\nu(z) \\ 1 \end{bmatrix}. \tag{4.9}$$

Using again the simple observation

$$U(e^{2i\vartheta}\overline{a}) = \begin{bmatrix} e^{i\vartheta} & 0 \\ 0 & e^{-i\vartheta} \end{bmatrix} U(\overline{a}) \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix},$$

we find that

$$W^{(k)}(z) = \begin{bmatrix} e^{ik\vartheta} & 0 \\ 0 & e^{-ik\vartheta} \end{bmatrix} W^{(0)}(z) \begin{bmatrix} e^{-ik\vartheta} & 0 \\ 0 & e^{ik\vartheta} \end{bmatrix}.$$

Thus,

$$W^{(k-1)}(z)\cdots W^{(0)}(z)=\begin{bmatrix} e^{ik\vartheta} & 0 \\ 0 & e^{-ik\vartheta}\end{bmatrix} W_\vartheta(z)^k. \tag{4.10}$$

Moreover, (2.17) and (4.8) show that $jY_0^{-1}W_\vartheta(z)^\intercal Y_0 j = M_\vartheta(z)$. So we conclude that

$$\operatorname{tr} W_\vartheta = \operatorname{tr} M_\vartheta = \operatorname{tr} T. \tag{4.11}$$

Recalling that

$$M(z)=\frac{1}{2}\begin{bmatrix} \psi_p(z)+\psi_p^*(z) & \psi_p^*(z)-\psi_p(z) \\ \varphi_p^*(z)-\varphi_p(z) & \varphi_p(z)+\varphi_p^*(z)\end{bmatrix},$$

a direct computation shows that

$$M_\vartheta(z)=\begin{bmatrix} M_{11}^\vartheta(z) & M_{12}^\vartheta(z) \\ M_{21}^\vartheta(z) & M_{22}^\vartheta(z)\end{bmatrix}=\frac{1}{2}\begin{bmatrix} \psi_{p,\vartheta}(z)+\psi_{p,\vartheta}^*(z) & \psi_{p,\vartheta}^*(z)-\psi_{p,\vartheta}(z) \\ \varphi_{p,\vartheta}^*(z)-\varphi_{p,\vartheta}(z) & \varphi_{p,\vartheta}(z)+\varphi_{p,\vartheta}^*(z)\end{bmatrix} \tag{4.12}$$

for the rotated rational functions $\varphi_{p,\vartheta} = e^{-i\vartheta}\varphi_p$, $\psi_{p,\vartheta} = e^{-i\vartheta}\psi_p$, $\varphi_{p,\vartheta}^* = e^{i\vartheta}\varphi_p^*$, and $\psi_{p,\vartheta}^* = e^{i\vartheta}\psi_p^*$. The discriminant Δ_A defined by (1.12) can therefore be written as

$$\Delta_A(z)=\frac{1}{B(z)}\operatorname{tr} M_\vartheta(z)=\frac{\psi_{p,\vartheta}(z)+\psi_{p,\vartheta}^*(z)+\varphi_{p,\vartheta}(z)+\varphi_{p,\vartheta}^*(z)}{2B(z)}, \tag{4.13}$$

where

$$B(z)=z\prod_{j=1}^{n-1} b_{z_j}(z)=\sqrt{\det M_\vartheta(z)}. \tag{4.14}$$

What follows is a detailed study of properties of Δ_A. This will enable us to give a complete description of the spectral measure of A_+ by means of a uniquely associated divisor D.

With (4.10) in mind, the analog of the Lyapunov exponent (see, e.g., [24]) in our periodic setting is given by

$$L(z)=L\big(z,\{a_j\}_{j=0}^{p-1},\vartheta\big)=\lim_{k\to\infty}\frac{1}{kp}\log\|W_\vartheta(z)^k\|, \tag{4.15}$$

provided the limit exists. This is, in fact, a Lyapunov exponent in a traditional sense, since the matrix $W_\vartheta(z)$ is, up to a unitary factor (cf. (4.10)), the periodic propagator for the ORFs $\phi_{p,\vartheta}$, $\psi_{p,\vartheta}$ analogous to, for example, (2.14). We shall shortly relate this Lyapunov exponent to the discriminant. The lemma below is critical when showing that $\Delta_A^{-1}([-2,2]) \subset \partial\mathbb{D}$.

Lemma 4.1. For every $z \in \mathbb{C} \setminus \{\hat{z}_j : 0 \leq j \leq n-1\}$, the limit $L(z)$ exists and satisfies $L(z) \geq 0$.

Remark. In fact, we will see that $L(z) = 0$ if and only if $z \in \Delta_A^{-1}([-2,2]) = \sigma(A) \subset \partial\mathbb{D}$.

Proof. Let $\lambda_1(z), \lambda_2(z)$ denote the eigenvalues of $W_\vartheta(z)$; we tacitly suppress the dependence of these values on ϑ and $\{a_j\}_{j=0}^{p-1}$. By the spectral radius formula, we have

$$\lim_{k\to\infty} \|W_\vartheta(z)^k\|^{1/kp} = \max\{|\lambda_1(z)|, |\lambda_2(z)|\}^{1/p}. \tag{4.16}$$

Moreover, since the inequality $|\det N| \leq \|N\|^2$ holds for every 2×2 matrix N and $\det W_\vartheta = B^2$, we see that on $\mathbb{C} \setminus (\mathbb{D} \cup \{\hat{z}_j : 0 \leq j \leq n-1\})$ the limit in (4.15) exists and satisfies $L \geq 0$. To show that this also holds inside $\mathbb{D}$, we first note that due to (2.14) and (4.10),

$$W_\vartheta^k = \frac{1}{2}\begin{bmatrix} \varphi_{kp,\vartheta} + \psi_{kp,\vartheta} & \varphi_{kp,\vartheta} - \psi_{kp,\vartheta} \\ \varphi^*_{kp,\vartheta} - \psi^*_{kp,\vartheta} & \varphi^*_{kp,\vartheta} + \psi^*_{kp,\vartheta} \end{bmatrix}. \tag{4.17}$$

Hence,

$$|\varphi^*_{kp,\vartheta}(z)| \leq \left\| W_\vartheta(z)^k \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right] \right\| \leq \|W_\vartheta(z)^k\|\sqrt{2},$$

and we can apply the Christoffel–Darboux formula (see [5, Theorem 3.1.3])

$$\sum_{j=0}^{l-1} |\varphi_{j,\vartheta}(z)|^2 = \frac{|\varphi^*_{l,\vartheta}(z)|^2 - |\varphi_{l,\vartheta}(z)|^2}{1 - |b_{z_l}(z)|^2}$$

to deduce that $1 - |z|^2 \leq |\varphi^*_{kp,\vartheta}(z)|^2$ since $b_{z_{kp}}(z) = z$ and $|\varphi_0| = 1$. For fixed $z \in \mathbb{D}$ we thus have a uniform lower bound on $\|W_\vartheta(z)^k\|$ and this implies that $L(z) \geq 0$. ■

The following lemma collects several important properties of Δ_A. It is the analog of [31, Theorem 11.1.1] for CMV matrices and thus we seek to keep the proofs rather short by merely indicating where adaptations are needed.

Lemma 4.2.

(i) $\overline{\Delta_A(1/\overline{z})} = \Delta_A(z)$,

(ii) $L(z) = \frac{1}{p}\log|B(z)| + \frac{1}{p}\log\left|\frac{\Delta_A(z)+\sqrt{\Delta_A(z)^2-4}}{2}\right|$,

(iii) $\Delta_A(z) \in [-2,2]$ implies that $z \in \partial\mathbb{D}$,

(iv) for all critical points $c \in \partial\mathbb{D}$ (i.e., zeros of Δ_A'), we have that $|\Delta_A(c)| \geq 2$.

Proof. *(i)* The key observation is that $B(z)^{-1}W_\vartheta(z) \in \mathbb{SU}(1,1)$ if $z \in \partial\mathbb{D}$. Using that the trace of a matrix in this class is real, the statement follows by analytic continuation.

(ii) Let again $\lambda_1(z), \lambda_2(z)$ be the eigenvalues of $W_\vartheta(z)$ and denote by $\tilde{\lambda}_1(z), \tilde{\lambda}_2(z)$ the eigenvalues of $B(z)^{-1}W_\vartheta(z)$. Then we have

$$B(z)\tilde{\lambda}_i(z) = \lambda_i(z), \quad i = 1,2$$

and thus, by (4.16),

$$L(z) = \log \lim_{k\to\infty} \|W_\vartheta(z)^k\|^{1/kp} = \frac{1}{p}\log|B(z)| + \frac{1}{p}\log(\max\{|\tilde{\lambda}_1(z)|, |\tilde{\lambda}_2(z)|\}).$$

Since $\tilde{\lambda}_1, \tilde{\lambda}_2$ are solutions of $\tilde{\lambda}^2 - \Delta_A(z)\tilde{\lambda} + 1 = 0$, we obtain *(ii)*.

(iii) Suppose that $\Delta_A(z) \in [-2,2]$. Then $\left|\frac{\Delta_A(z)+\sqrt{\Delta_A(z)^2-4}}{2}\right| = 1$ and so, by *(ii)*,

$$L(z) = \frac{1}{p}\log|B(z)|.$$

Since $L(z) \geq 0$, this implies that $|z| \geq 1$. But if $|z| > 1$ then, by *(i)*, we have

$$\Delta_A(z) = \overline{\Delta_A(1/\overline{z})} \notin [-2,2].$$

(iv) This is the same as to say that for $x \in (-2,2)$, the roots of $\Delta_A - x$ are simple. The proof is identical to that of [31, Theorem 11.1.1]. ■

It follows directly from the above lemma that $\mathsf{E} := \Delta_A^{-1}([-2,2]) \subset \partial\mathbb{D}$ is a finite-gap set with at most p gaps. As before, we denote by $g+1$ the number of open gaps in E and by $\lambda_j^\pm$ their gap edges. In what follows, it will be important to know that Δ_A is a rational function of degree p. But, in fact, an even stronger statement is true:

Lemma 4.3. Fix j and let q be the number of times z_j appears in the vector $\vec{z}$. Then

$$|(b_{z_j}^q \Delta_A)(z_j)| > C > 0, \tag{4.18}$$

where the constant C depends only on $\vec{z}$. Moreover,

$$(b_{z_j}^{-q}\Delta_A)(\hat{z}_j) = \overline{(b_{z_j}^{q}\Delta_A)(z_j)}. \tag{4.19}$$

Proof. Note that (4.19) follows from the fact that Δ_A is real. Let us first assume that $q = 1$ and $z_j = z_0 = 0$. In that case, (4.18) is equivalent to $|(\operatorname{tr} T)(0)| > C$. It will be more convenient to consider the product

$$T_r(z) := \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} T(z) \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}^{-1}$$

which clearly has the same trace as T. When $z = 0$, we have

$$T_r(0) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} U(a_0) \begin{bmatrix} -z_1 & 0 \\ 0 & 1 \end{bmatrix} \cdots \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} U(a_{p-1}) \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix}. \tag{4.20}$$

Set

$$V(z) = U(a_0) \begin{bmatrix} b_{z_1}(z) & 0 \\ 0 & 1 \end{bmatrix} \cdots U(a_{p-3}) \begin{bmatrix} b_{z_{n-1}}(z) & 0 \\ 0 & 1 \end{bmatrix}$$

and notice that V is a transfer matrix associated to the poles $z_1, \ldots, z_{n-1}$ and with coefficients $-\overline{a_0}, \ldots, -\overline{a_{p-3}}$ (in reverse order). Therefore, by (2.14), a short computation shows that

$$V = \frac{1}{2} \begin{bmatrix} \tilde{\varphi} + \tilde{\psi} & \tilde{\varphi} - \tilde{\psi} \\ \tilde{\varphi}^* - \tilde{\psi}^* & \tilde{\varphi}^* + \tilde{\psi}^* \end{bmatrix}$$

for the corresponding ORFs of degree $p - 2$. Due to (4.20), it suffices to show that

$$\left| \begin{bmatrix} 0 & 1 \end{bmatrix} V(0) U(a_{p-2}) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right| = \left| \frac{1}{2\rho_{p-2}} \Big(\tilde{\varphi}^*(0) + \tilde{\psi}^*(0) + a_{p-2}\big(\tilde{\varphi}^*(0) - \tilde{\psi}^*(0)\big) \Big) \right|$$

is uniformly bounded from below.

For arbitrary z_j, due to cyclic rotation, we would have obtained the same as above, just with ORFs associated to different coefficients, respectively a different measure. Recall that we can always "push" the matrix $\begin{bmatrix} e^{i\vartheta} & 0 \\ 0 & e^{-i\vartheta} \end{bmatrix}$ to the end of the product by use of the commutant relation

$$U(e^{2i\vartheta} a) = \begin{bmatrix} e^{i\vartheta} & 0 \\ 0 & e^{-i\vartheta} \end{bmatrix} U(a) \begin{bmatrix} e^{-i\vartheta} & 0 \\ 0 & e^{i\vartheta} \end{bmatrix}.$$

Hence it suffices to show that

$$\left|\tilde{\varphi}^*(z_j) + \tilde{\psi}^*(z_j) + a\big(\tilde{\varphi}^*(z_j) - \tilde{\psi}^*(z_j)\big)\right| \tag{4.21}$$

is uniformly bounded from below for arbitrary ORFs whose poles are supported on the set $\{\hat{z}_i : 0 \le i \le n-1, i \ne j\}$.

It is well-known that any Caratheodory function F satisfies the uniform bounds

$$\frac{1-|z|}{1+|z|} \le |F(z)| \le \frac{1+|z|}{1-|z|}, \quad \frac{1-|z|}{1+|z|} \le \operatorname{Re} F(z) \le \frac{1+|z|}{1-|z|}.$$

Since $\sup_j |z_j| < 1$, this gives positive constants c_1, c_2 such that for every Caratheodory function F and every z_j, we have that $c_1 < |F(z_j)| < c_2$. By the same reasoning, using the Christoffel–Darboux relation it is not hard to see that there exists a function $m(r)$ such that for all measures ν, all n, and all z obeying $|z| < r$, we have $|\varphi_n^*(z,\nu)| > m(r) > 0$; cf. [5, Lemma 9.3.1]. Hence we obtain a constant c_3 such that $|\tilde{\varphi}^*(z_j)| > c_3$ uniformly.

Writing $\tilde{F}$ as shorthand notation for the Caratheodory function $\tilde{\psi}^*/\tilde{\varphi}^*$, we see that

$$\left|\tilde{\varphi}^*(z_j) + \tilde{\psi}^*(z_j) + a\big(\tilde{\varphi}^*(z_j) - \tilde{\psi}^*(z_j)\big)\right| = \left|\tilde{\varphi}^*(z_j)\right|\left|(1+a) + \tilde{F}(z_j)(1-a)\right|.$$

Let us first assume that $|1-a| < (1+c_2)^{-1} \le 1/2$. Then standard estimates show that

$$\left|\tilde{\varphi}^*(z_j)\right|\left|(1+a) + \tilde{F}(z_j)(1-a)\right| > c_3.$$

If $|1-a| \ge (1+c_2)^{-1}$, we obtain the estimate

$$\begin{aligned}\left|\tilde{\varphi}^*(z_j)(1-a)\right| \left|\frac{1+a}{1-a} + \tilde{F}(z_j)\right| &\ge \left|\tilde{\varphi}^*(z_j)(1-a)\right| \operatorname{Re}\left(\frac{1+a}{1-a} + \tilde{F}(z_j)\right) \\ &\ge \left|\tilde{\varphi}^*(z_j)(1-a)\right| \operatorname{Re} F(z_j) > \frac{c_1 c_3}{1+c_2} > 0\end{aligned}$$

since the real part of $\frac{1+a}{1-a}$ is positive.

Finally, if $q > 1$ then (4.20) splits into shorter products of the form

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tilde{V}(z) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and the same arguments can be applied to each single factor. The proof is complete. ■

Combining the two previous lemmas, we obtain the following explicit representation for Δ_A:

Lemma 4.4. Let (λ_-, λ_+) be a gap of E such that $\Delta_A > 0$ in this gap and choose some point $\lambda_* \in (\lambda_-, \lambda_+)$ as normalization for the uniformization $\mathfrak{z}$ of $\overline{\mathbb{C}} \setminus \mathsf{E}$. Following the notation of (3.21), we then have that

$$\Delta_A(z) = \Psi(z) + \Psi(z)^{-1}, \tag{4.22}$$

where $\Psi \circ \mathfrak{z} = \prod_{j=0}^{n-1} \mathfrak{b}_{\zeta_j} \mathfrak{b}_{\overline{\zeta_j}}$.

Remark. The sign of Δ_A is related to our choice of $e^{i\vartheta}$ as the square root of $e^{2i\vartheta}$. Indeed, as already mentioned, by the linearity of the trace, the choice $-e^{i\vartheta}$ would have led to $-\Delta_A(z)$. Hence, if $\Delta_A < 0$ in all gaps, we simply choose from the very beginning $-e^{i\vartheta}$ to be in the setting of the lemma.

Proof. Consider the function

$$H(z) = \log \left| \frac{\Delta_A(z) + \sqrt{\Delta_A^2(z) - 4}}{2} \right| - \sum_{j=0}^{n-1} \Big(G_{\overline{\mathbb{C}} \setminus \mathsf{E}}(z, z_j) + G_{\overline{\mathbb{C}} \setminus \mathsf{E}}(z, \hat{z}_j) \Big).$$

It follows from Lemma 4.3 that H has no poles and thus is harmonic in $\overline{\mathbb{C}} \setminus \mathsf{E}$. Furthermore, since $\mathsf{E} := \Delta_A^{-1}([-2, 2])$, we have that $H(z) = 0$ for $z \in \mathsf{E}$. So by the maximum principle, $H \equiv 0$. In particular, by an application of (3.9) we get that

$$\log \left| \frac{\Delta_A\big(\mathfrak{z}(\zeta)\big) + \sqrt{\Delta_A^2\big(\mathfrak{z}(\zeta)\big) - 4}}{2} \right| = -\log |\Psi \circ \mathfrak{z}| .$$

This defines $\Psi \circ \mathfrak{z}$ up to a unimodular constant, c. Since Δ_A is real, we see that this constant must be ± 1. Due to the specific normalization and Lemma 3.12, it thus follows that $c = 1$. ∎

We are now ready to characterize the spectrum of A_+. Since $\det M_\vartheta = B^2$, (4.9) and (4.12) show that the Caratheodory function F_ν can be written as

$$F_\nu(z) = \frac{v(z) + \sqrt{\Delta_A(z)^2 - 4}}{u(z)}, \tag{4.23}$$

where u, v are explicitly given by

$$u(z) = \frac{\varphi^*_{p,\vartheta}(z) - \varphi_{p,\vartheta}(z)}{B(z)}, \quad v(z) = \frac{\psi_{p,\vartheta}(z) + \psi^*_{p,\vartheta}(z) - \varphi_{p,\vartheta}(z) - \varphi^*_{p,\vartheta}(z)}{2B(z)},$$

and where the branch of the square root is chosen such that $F_\nu(0) = 1$ and then extended analytically to $\overline{\mathbb{C}} \setminus \mathsf{E}$. In order to show that v and iu are both real-valued on $\partial\mathbb{D}$, first recall that $\varphi^*_k(z) = B_k(z)\overline{\varphi_k(\hat{z})}$ by definition. Using the fact that $B_p = B^2$, we thus see that $\varphi^*_{p,\vartheta} = B^2\overline{\varphi_{p,\vartheta}}$ on $\partial\mathbb{D}$. By a similar reasoning we obtain the same identity for $\psi^*_{p,\vartheta}$ and it follows that

$$v(e^{it}) \in \mathbb{R} \quad \text{and} \quad u(e^{it}) \in i\mathbb{R}. \tag{4.24}$$

The next step is to write $d\nu(t) = \nu_{ac}(t)\frac{dt}{2\pi} + d\nu_s(t)$, with $d\nu_s$ singular to $dt/2\pi$. Using the standard inversion formula [34, Theorem 2.5.5], we see that ν_s is a finite sum of point masses and ν_{ac} is explicitly given by

$$\nu_{ac}(t) = \lim_{r\uparrow 1} \operatorname{Re} F_\nu(re^{it}) = \frac{\sqrt{\Delta_A(e^{it})^2 - 4}}{u(e^{it})} \geq 0, \quad e^{it} \in \mathsf{E}. \tag{4.25}$$

The following lemma will be important to characterize the point masses of ν.

Lemma 4.5. $u(z)$ has all its zeros in the set of gaps of E, one in each gap.

Proof. First we show that all zeros of u lie on $\partial\mathbb{D}$. Since $|\varphi_{p,\vartheta}| < |\varphi^*_{p,\vartheta}|$ on $\mathbb{D}$ and $|\varphi_{p,\vartheta}| > |\varphi^*_{p,\vartheta}|$ on $\overline{\mathbb{C}} \setminus \overline{\mathbb{D}}$ (see [5, Corollary 3.1.4.]), the assertion $\varphi_{p,\vartheta}(z) = \varphi^*_{p,\vartheta}(z)$ (i.e., $u(z) = 0$) implies that $z \in \partial\mathbb{D}$.

Next we show that at a point $e^{it_0} \in \partial\mathbb{D}$ with $\varphi_{p,\vartheta}(e^{it_0}) = \varphi^*_{p,\vartheta}(e^{it_0})$, we have $|\operatorname{tr}\Delta_A(e^{it_0})| \geq 2$. Start by observing that for a matrix $N \in \mathbb{SL}(2,\mathbb{R})$ with $N_{21} = 0$, one has $|\operatorname{tr} N| \geq 2$. Indeed, an application of the inequality of arithmetic and geometric means shows that $|\operatorname{tr} N| \geq 2\sqrt{\det N} = 2$. Set $\widetilde{M} = B(e^{it_0})^{-1}M_\vartheta(e^{it_0})$. From (2.11) and (4.8), we see that $\widetilde{M}$ can be written as $\widetilde{M} = Y_0UY_0^{-1}$ for some $U \in \mathbb{SU}(1,1)$. Viewed as fractional linear transformations, Y_0 maps the unit disk $\mathbb{D}$ into the right half-plane $\mathbb{H}_+$ and U preserves the unit disk; thus, conjugating $\widetilde{M}$ further by the matrix

$$R = \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix},$$

we can transform $\widetilde{M}$ into an element of $\mathbb{SL}(2,\mathbb{R})$. It remains to note that, due to (4.12), $\widetilde{M}_{21}$ is zero and this property (in addition to the determinant and the trace) is preserved under conjugation by R.

Since Δ_A is of degree p, there are exactly p gaps of E. As the square root in (4.25) is analytically extended, it has alternating signs on consecutive gaps. In order to retain the positivity, the denominator must also admit a sign change and this implies there is a zero of $u(z)$ in every gap. Since u is of degree at most p, we find that it has exactly one zero per gap. ■

Inspired by the above lemma, let us define a divisor

$$D = \{(x_j, \varepsilon_j)\}_{j=0}^{g} \in \mathcal{D}(\mathsf{E}) \tag{4.26}$$

with $\{x_j\}$ accounting for the zeros of $u(z)$ in the open gaps and where $\varepsilon_j = 1$ if x_j is a mass point of the measure ν and -1 otherwise.

Proposition 4.6. We can uniquely recover the measure ν from the divisor D, the vector $\vec{z}$, and the spectrum $\mathsf{E} = \sigma(A)$. Specifically, the absolutely continuous part of ν is given by (4.25) and if $\varepsilon_j = 1$, the point mass at x_j has the weight

$$\nu(\{x_j\}) = \frac{\sqrt{|\Delta_A(x_j)^2 - 4|}}{|u'(x_j)|}. \tag{4.27}$$

Proof. Given E and $\vec{z}$, Lemma 4.4 shows that Δ_A is uniquely determined up to a sign by the key properties (i)–(iii) listed after Definition 1.2. However, this choice of sign becomes irrelevant here since the critical points and the value of the expression $\Delta_A^2(z) - 4$ as in (4.25) and (4.27) are invariant under the map $\Delta_A \mapsto -\Delta_A$.

We now explain how to recover the measure ν (and hence F_ν) from Δ_A and the divisor D. If x_j is a zero of u, then—using again that $\det M_\vartheta = B^2$—we see that $\Delta_A(x_j)^2 - 4 = v(x_j)^2$. Hence the sign of the square root determines whether or not x_j is a point mass of ν. If the numerator in (4.23) does not vanish, it is given by $2\sqrt{\Delta_A^2(x_j) - 4}$ and the weight of the point mass can be computed as

$$\nu(\{x_j\}) = \lim_{r\uparrow 1} \frac{1-r}{2} \operatorname{Re} F_\nu(rx_j) = \frac{\sqrt{|\Delta_A(x_j)^2 - 4|}}{|u'(x_j)|},$$

cf. [34, Theorem 2.5.5]. This proves (4.27), and we have already seen that the absolutely continuous part of ν is given by (4.25).

Note that up to a multiplicative constant, u is defined by its zeros. This constant is fixed by (4.24) and the fact that ν is a probability measure. Hence, D, E, and $\vec{z}$ determine ν uniquely. ■

For finite-gap sets E, Lemma 4.5 (and the comment thereafter) defines a map from $\{A \in \mathbb{A}_{\text{per}}(\vec{z}) : \sigma(A) = \mathsf{E}\}$ to $\mathcal{D}(\mathsf{E})$: namely, given $A \in \mathbb{A}_{\text{per}}(\vec{z})$, we consider the associated half-line operator $A_+ = b_{-D_+}(C_+)$, compute its Caratheodory function F_ν as in (4.23), and find and label the zeros of $u(z)$ as in (4.26). Proposition 4.6 shows the map $A_+ \mapsto D \in \mathcal{D}(\mathsf{E})$ is one-to-one:

Corollary 4.7. The map $\mathbf{G} : \mathcal{T}_{\text{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*) \to \mathcal{D}(\mathsf{E})$ assigning a divisor (4.26) to a periodic MCMV matrix $A \in \mathcal{T}_{\text{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$ is one-to-one.

Proof. Let $A_1, A_2 \in \mathcal{T}_{\text{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$ and suppose $\mathbf{G}(A_1) = \mathbf{G}(A_2)$. By Proposition 4.6, $A_{1,+}$ and $A_{2,+}$ have the same spectral measure ν (for δ_0). Thus, our choice of $\vec{z}$ determines the same sequence of ORFs, and thus the same (half-line) sequence of phase-periodic Verblunsky coefficients $\{a_k\}_{k\in\mathbb{N}}$ having phase $e^{-2i\vartheta}$ by (2.7). Extending this sequence to $\mathbb{Z}$ by phase-periodicity (which we must have by (4.6) and the surrounding discussion) generates an infinite sequence of Verblunsky coefficients determined uniquely by $\{a_k\}_{k=0}^{2n-1}$ and $e^{-2i\vartheta}$. As $\Delta_{A(\cdot,\vartheta;\vec{z})} = -\Delta_{A(\cdot,\vartheta+\pi;\vec{z})}$, precisely one of ϑ or $\vartheta + \pi$ will yield a discriminant with $\Delta_A(\lambda_*) > 0$. Since $A_1, A_2 \in \mathcal{T}_{\text{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$, it follows that they are generated by the same parameters $\{a_k\}_{k=0}^{2n-1}$, ϑ, and $\vec{z}$, that is, $A_1 = A_2$. ■

The fact that this map is also onto is true in general, and will be shown for the special choice $\vec{z} = \vec{z}_{\mathsf{E}}$ in Section 5.

4.3 The structure of a general MCMV matrix

In this section we demonstrate the "block-CMV" nature of MCMV matrices as in (1.7), as well as a structural stability under taking Möbius transformations related to the points in the generating vector $\vec{z}$. This structure will be critical to understanding our main theorems; indeed, in light of viewing the discriminant as in (1.23), the Magic Formula would be a complete mystery without developing some understanding of the structure of $b_{z_j}(A)$ for a general MCMV matrix $A \in \mathbb{A}(\vec{z})$.

Our analysis further illustrates the similarities between our MCMV matrices and their self-adjoint analog, GMP matrices. First, GMP matrices are block-Jacobi; below, we show the band structure (1.7) of an MCMV matrix in Lemma 4.8. Additionally, one of the characteristic properties of GMP matrices is that they are stable under taking resolvents (cf. [39, Definition 1.12]); the analogous statement for MCMV is Proposition 4.9. The consequences for the Magic Formula in the setting of (1.23) are the content of Theorem 4.10. Finally, we use all of this structure to prove a uniqueness result for certain rational functions of our MCMV matrices in Proposition 4.11; this will be used to prove the Magic Formula in the end.

Let us again fix a vector $\vec{z} \in \mathbb{D}^n$ having $z_0 = 0$, and recall that D_0 denotes the diagonal operator defined in (1.4) and C denotes a general CMV matrix. We begin by proving the block structure of an MCMV matrix A; up to conjugation by diagonal matrices, we may consider instead the simpler operator

$$\tilde{A} := (1 + CD_0^*)^{-1}(C + D_0). \tag{4.28}$$

Let us split $\tilde{A}$ into matrix blocks of size $2n \times 2n$ and denote the blocks by $\tilde{\mathbf{A}}_{\mathbf{ij}}$. For simplicity, we assume throughout this section that $\vec{z}$ is such that $z_j \neq z_k$ for $j \neq k$. Without this assumption, the block structure below will split into smaller blocks, but we emphasize that the below proofs can be readily adapted to the case of an arbitrary vector $\vec{z} \in \mathbb{D}^n$ with $z_0 = 0$ at the cost of adapting the results to apply to multiple sub-blocks of arbitrary even block-sizes.

Lemma 4.8. $\tilde{A}$ is band structured and $\tilde{\mathbf{A}}_{\mathbf{ij}} = 0$ if $|\mathbf{i} - \mathbf{j}| > 1$. Moreover, there exists vectors $\mathbf{u}^{\mathbf{i}}, \mathbf{v}^{\mathbf{i}} \in \mathbb{C}^{2n}$ such that

$$\tilde{\mathbf{A}}_{\mathbf{i-1,i}} = \mathbf{v}^{\mathbf{i}}\delta_{2n+1}^{\intercal}, \quad \tilde{\mathbf{A}}_{\mathbf{i,i+1}} = \mathbf{u}^{\mathbf{i}}\delta_0^{\intercal}. \tag{4.29}$$

In particular,

$$\tilde{A}_{0,2n} = \left(\frac{1}{\rho_{2n-1}} \begin{bmatrix} 1 & 0 \end{bmatrix} U(-\overline{a_{2n-2}}) \begin{bmatrix} 1 & 0 \\ 0 & -\overline{z_{n-1}} \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ 0 & -\overline{z_1} \end{bmatrix} U(-\overline{a_1}) \begin{bmatrix} -\frac{1}{\overline{z_1}} & 0 \\ 0 & 1 \end{bmatrix} U(-\overline{a_0}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)^{-1}. \tag{4.30}$$

Proof. Let

$$u = (1 + CD_0^*)^{-1}(C + D_0)\delta_k$$

and note that u satisfies the recursion relation

$$(1 + CD_0^*)u = (C + D_0)\delta_k. \tag{4.31}$$

The left-hand side of this identity can be written as

$$\begin{bmatrix} u_{2j} \\ u_{2j+1} \end{bmatrix} + \overline{z_j} \begin{bmatrix} \overline{a_{2j}} \\ \rho_{2j} \end{bmatrix} \begin{bmatrix} \rho_{2j-1} & -a_{2j-1} \end{bmatrix} \begin{bmatrix} u_{2j-1} \\ u_{2j} \end{bmatrix} + \overline{z_{j+1}} \begin{bmatrix} \rho_{2j} \\ -a_{2j} \end{bmatrix} \begin{bmatrix} \overline{a_{2j+1}} & \rho_{2j+1} \end{bmatrix} \begin{bmatrix} u_{2j+1} \\ u_{2j+2} \end{bmatrix}. \tag{4.32}$$

Introduce the matrices

$$M_j = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} + \overline{z_j} \begin{bmatrix} \overline{a_{2j}} \\ \rho_{2j} \end{bmatrix} \begin{bmatrix} \rho_{2j-1} & -a_{2j-1} \end{bmatrix}$$

and

$$N_j = -\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} - \overline{z_{j+1}} \begin{bmatrix} \rho_{2j} \\ -a_{2j} \end{bmatrix} \begin{bmatrix} \overline{a_{2j+1}} & \rho_{2j+1} \end{bmatrix}$$

so that (4.32) becomes

$$M_j \begin{bmatrix} u_{2j-1} \\ u_{2j} \end{bmatrix} - N_j \begin{bmatrix} u_{2j+1} \\ u_{2j+2} \end{bmatrix}.$$

Since $z_{kn} = 0$, we see that for $j = kn - 1$ and $j = kn$ this becomes

$$M_{kn-1} \begin{bmatrix} u_{2kn-3} \\ u_{2kn-2} \end{bmatrix} + \begin{bmatrix} 0 \\ u_{2kn-1} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} u_{2kn} \\ 0 \end{bmatrix} - N_{kn} \begin{bmatrix} u_{2kn+1} \\ u_{2kn+2} \end{bmatrix},$$

respectively. Hence we see that the recursion for the blocks $\{u_{2kn}, \dots, u_{2(k+1)n-1}\}$ is decoupled. The finite band structure and (4.29) is now a direct consequence of the structure of CMV matrices.

It remains to prove (4.30). Let now $u = (1 + CD_0^*)^{-1}(C + D_0)\delta_{2n}$. We are interested in the block $\{u_0, \dots, u_{2n-1}\}$. In this case, (4.31) leads to

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} u_0 \\ 0 \end{bmatrix} - N_0 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad M_{n-1} \begin{bmatrix} u_{2n-3} \\ u_{2n-2} \end{bmatrix} + \begin{bmatrix} 0 \\ u_{2n-1} \end{bmatrix} = \begin{bmatrix} \rho_{2n-1}\rho_{2n-2} \\ -\rho_{2n-1}a_{2n-2} \end{bmatrix}$$

and for $1 \le j < n - 1$,

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = M_j \begin{bmatrix} u_{2j-1} \\ u_{2j} \end{bmatrix} - N_j \begin{bmatrix} u_{2j+1} \\ u_{2j+2} \end{bmatrix}.$$

Iterating this leads to

$$\begin{bmatrix} u_0 \\ 0 \end{bmatrix} = N_0 M_1^{-1} N_1 \cdots N_{n-2} M_{n-1}^{-1} \left(\begin{bmatrix} \rho_{2n-1}\rho_{2n-2} \\ -\rho_{2n-1} a_{2n-2} \end{bmatrix} - \begin{bmatrix} 0 \\ u_{2n-1} \end{bmatrix} \right).$$

Using the fact that

$$M_0 \begin{bmatrix} u_{2n-1} \\ u_0 \end{bmatrix} = \begin{bmatrix} u_0 \\ 0 \end{bmatrix} \text{ and } N_{n-1} \begin{bmatrix} u_{2n-1} \\ u_0 \end{bmatrix} = \begin{bmatrix} 0 \\ -u_{2n-1} \end{bmatrix},$$

we arrive at

$$\begin{bmatrix} \rho_{2n-1}\rho_{2n-2} \\ -\rho_{2n-1} a_{2n-2} \end{bmatrix} = \left(M_{n-1} N_{n-2}^{-1} M_{n-2} \cdots N_0^{-1} M_0 - N_{n-1} \right) \begin{bmatrix} u_{2n-1} \\ u_0 \end{bmatrix}.$$

It is straightforward to see that, for $1 \le j < n-1$,

$$N_j^{-1} = \frac{-1}{\rho_{2j}\rho_{2j+1}} \left(\frac{1}{\overline{z_{j+1}}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + \begin{bmatrix} \rho_{2j+1} \\ -\overline{a_{2j+1}} \end{bmatrix} \begin{bmatrix} a_{2j} & \rho_{2j} \end{bmatrix} \right).$$

Let

$$M_{n-1} N_{n-2}^{-1} M_{n-2} \cdots N_0^{-1} = C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}.$$

Due to the simple structure of M_0 and N_{n-1}, it follows that

$$u_0 = \frac{\rho_{2n-2}\rho_{2n-1}}{c_{11}}$$

and hence it suffices to study $C \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right]$. A direct computation shows that

$$N_0^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{-1}{\rho_0 \rho_1} \begin{bmatrix} 0 & \rho_1 \\ 1 & -\overline{a_1} \end{bmatrix} \begin{bmatrix} \frac{1}{\overline{z_1}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ a_0 \end{bmatrix}$$

and combined with the identities

$$M_j \begin{bmatrix} 0 & \rho_{2j-1} \\ 1 & -\overline{a_{2j-1}} \end{bmatrix} = \begin{bmatrix} 1 & \overline{a_{2j}} \\ 0 & \rho_{2j} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \overline{z_j} \end{bmatrix} \begin{bmatrix} 1 & -\overline{a_{2j-1}} \\ -a_{2j-1} & 1 \end{bmatrix}$$

and

$$N_j^{-1}\begin{bmatrix}1 & \overline{a_{2j}}\\ 0 & \rho_{2j}\end{bmatrix} = \frac{-1}{\rho_{2j}\rho_{2j+1}}\begin{bmatrix}0 & \rho_{2j+1}\\ 1 & -\overline{a_{2j+1}}\end{bmatrix}\begin{bmatrix}\frac{1}{\overline{z_{j+1}}} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & \overline{a_{2j}}\\ a_{2j} & 1\end{bmatrix},$$

this allows us to iterate the procedure. It only remains to comment on the sign and for this we note that

$$-\begin{bmatrix}\frac{1}{\overline{z_{j+1}}} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & \overline{a_{2j}}\\ a_{2j} & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & \overline{z_j}\end{bmatrix} = \begin{bmatrix}\frac{-1}{\overline{z_{j+1}}} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & -\overline{a_{2j}}\\ -a_{2j} & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -\overline{z_j}\end{bmatrix}.$$

Therefore,

$$C\begin{bmatrix}1\\ 0\end{bmatrix} = \frac{1}{\rho_0}\begin{bmatrix}1 & \overline{a_{2n-2}}\\ 0 & \rho_{2n-2}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & \overline{z_{n-1}}\end{bmatrix}\cdots\begin{bmatrix}1 & 0\\ 0 & -\overline{z_1}\end{bmatrix}U(-\overline{a_1})\begin{bmatrix}-\frac{1}{\overline{z_1}} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1\\ a_0\end{bmatrix},$$

and hence

$$\begin{bmatrix}1 & 0\end{bmatrix}C\begin{bmatrix}1\\ 0\end{bmatrix} = \rho_{2n-2}\begin{bmatrix}1 & 0\end{bmatrix}U(-\overline{a_{2n-2}})\begin{bmatrix}1 & 0\\ 0 & -\overline{z_{n-1}}\end{bmatrix}\cdots$$

$$\cdots\begin{bmatrix}1 & 0\\ 0 & -\overline{z_1}\end{bmatrix}U(-\overline{a_1})\begin{bmatrix}-\frac{1}{\overline{z_1}} & 0\\ 0 & 1\end{bmatrix}U(-\overline{a_0})\begin{bmatrix}1\\ 0\end{bmatrix}.$$

This concludes the proof. ■

Remark. The assumption that $z_j \neq z_k$ could be weakened to $z_j \neq z_0 = 0$ for $j \neq 0$ in this proof without any further changes.

A key feature of the MCMV structure is its stability under Möbius transformations. Following the notation of Appendix B, we can write the operator Möbius transform defined in (1.3) as

$$b_S(C) = \eta_S(1 - CS^*)^{-1}(C - S)\eta_S^{-1} = \Phi_U(C),$$

where

$$U = \begin{bmatrix}1 & -S^*\\ -S & 1\end{bmatrix}\begin{bmatrix}\eta_S^{-1} & 0\\ 0 & \eta_S^{-1}\end{bmatrix}.$$

In particular, this holds for $S = z_j 1$, where 1 denotes the identity on ℓ^2, that is, $b_{z_j}(A)$ is the standard Blaschke factor evaluated at $A \in \overline{\mathbb{D}_{\ell^2}}$:

$$b_{z_j}(A) = (1 - \overline{z_j} A)^{-1}(A - z_j). \tag{4.33}$$

In addition to the usual diagonal operator D_0, we likewise define shifted diagonal operators

$$D_j = (1 - z_j D_0^*)^{-1}(D_0 - z_j) \ \text{ and } V_j = \sqrt{(1 - \overline{z_j} D_0)^{-1}(1 - z_j D_0^*)}. \tag{4.34}$$

With this notation, we can explicitly describe how the Blaschke factors associated to the generating vector $\vec{z}$ "shift" MCMV matrices:

Proposition 4.9. Let D_j and V_j be defined as above. Then for any $C \in \overline{\mathbb{D}_{\ell^2}}$, we have

$$b_{z_j}\big(b_{-D_0}(C)\big) = V_j b_{-D_j}(C) V_j. \tag{4.35}$$

Proof. Let

$$U_1 = \begin{bmatrix} 1 & -\overline{z_j} \\ -z_j & 1 \end{bmatrix} \text{ and } U_2 = \begin{bmatrix} 1 & D_0^* \\ D_0 & 1 \end{bmatrix} \begin{bmatrix} \eta_{D_0}^{-1} & 0 \\ 0 & \eta_{D_0}^{-1} \end{bmatrix}.$$

Then we have

$$b_{z_j}\big(b_{-D_0}(C)\big) = \Phi_{U_1}(\Phi_{U_2}(C)) = \Phi_{U_2 U_1}(C).$$

Recalling that $\eta_j = \sqrt{1 - |z_j|^2}$, it is straightforward to see that

$$\eta_j^{-1} U_2 U_1 = \begin{bmatrix} 1 & (D_0^* - \overline{z_j})(1 - \overline{z_j} D_0)^{-1} \\ (D_0 - z_j)(1 - z_j D_0^*)^{-1} & 1 \end{bmatrix}$$
$$\times \begin{bmatrix} \eta_j^{-1} \eta_{D_0}^{-1}(1 - z_j D_0^*) & 0 \\ 0 & \eta_j^{-1} \eta_{D_0}^{-1}(1 - \overline{z_j} D_0) \end{bmatrix}.$$

Due to (3.35), we have

$$(1 - \overline{z_j} D_0)(1 - z_j D_0^*)\eta_{D_j}^2 = \eta_j^2 \eta_{D_0}^2$$

and hence

$$\eta_j U_2 U_1 = \begin{bmatrix} 1 & D_j^* \\ D_j & 1 \end{bmatrix} \begin{bmatrix} \eta_{D_j}^{-1} & 0 \\ 0 & \eta_{D_j}^{-1} \end{bmatrix} \begin{bmatrix} V_j & 0 \\ 0 & V_j^{-1} \end{bmatrix}.$$

This concludes the proof. ■

Remark. So far we have not used that $z_j \in \vec{z}$; however, this assumption is important to retain the banded structure of an MCMV matrix. In particular, since $b_{z_j}(z_j) = 0$, applying b_{z_j} to A shifts the zeros in D_0 by $2j$. It follows that $S^{-2j} b_{z_j}(A) S^{2j}$ is again MCMV-structured (with a new generating vector). Here specifically we use the assumption that $z_j \neq z_k$ for $j \neq k$: in particular, we have that $b_{z_k}(z_j) \neq 0$ for $j \neq k$, which allows us to apply our proof of Lemma 4.8 to the Blaschke shifts $S^{-2j} b_{z_j}(A) S^{2j}$.

We have analyzed the block structure and Blaschke shifts of MCMV matrices in order to understand the Magic Formula in the context of the representation (1.23). Critical to this understanding is computing off-diagonal blocks of self-adjoint operators of the form $\mathrm{Re}(c_j b_{z_j}(A))$ (cf. (1.24)); this is essentially the content of the next theorem of this section:

Theorem 4.10. Let $\vec{z}$ be such that $z_0 = 0$ and $z_j \neq z_k$ for $j \neq k$, and let $A \in \mathbb{A}(\vec{z})$. Then

$$\left(b_{z_k}(A)\right)_{2(n+k),2k} = e^{i\vartheta} \operatorname{tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} U(a_{2k}) \begin{bmatrix} b_{z_{k+1}}(\hat{z}_k) & 0 \\ 0 & 1 \end{bmatrix} U(a_{2k+1}) \cdots \right.$$

$$\left. \cdots \begin{bmatrix} 1 & 0 \\ 0 & b_{z_{k-1}}(\hat{z}_k)^{-1} \end{bmatrix} U(a_{2(n+k)-2}) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} U(a_{2(n+k)-1}) \right)^{-1}. \quad (4.36)$$

In particular,

$$A_{2n,0} = e^{i\vartheta} \operatorname{tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} U(a_0) \begin{bmatrix} b_{z_1}(\infty) & 0 \\ 0 & 1 \end{bmatrix} U(a_1) \cdots \begin{bmatrix} 1 & 0 \\ 0 & b_{z_{n-1}}(\infty)^{-1} \end{bmatrix} U(a_{n-2}) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} U(a_{n-1}) \right)^{-1}. \quad (4.37)$$

Proof. Equation (36) follows from shifting (4.37) by Proposition 4.9; thus, we only need to prove (4.37). Note that

$$\tilde{A}_{2n,0} = \operatorname{tr}\left(\frac{1}{\rho_{2n-1}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} U(-\overline{a_{2n-2}}) \begin{bmatrix} 1 & 0 \\ 0 & -\overline{z_{n-1}} \end{bmatrix} \cdots \right.$$

$$\left. \cdots \begin{bmatrix} 1 & 0 \\ 0 & -\overline{z_1} \end{bmatrix} U(-\overline{a_1}) \begin{bmatrix} -\frac{1}{\overline{z_1}} & 0 \\ 0 & 1 \end{bmatrix} U(-\overline{a_0}) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right)^{-1}$$

$$= \operatorname{tr}\left(U(-\overline{a_{2n-1}}) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} U(-\overline{a_{2n-2}}) \begin{bmatrix} 1 & 0 \\ 0 & -\overline{z_{n-1}} \end{bmatrix} \cdots \right.$$

$$\left. \cdots \begin{bmatrix} 1 & 0 \\ 0 & -\overline{z_1} \end{bmatrix} U(-\overline{a_1}) \begin{bmatrix} -\frac{1}{\overline{z_1}} & 0 \\ 0 & 1 \end{bmatrix} U(-\overline{a_0}) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right)^{-1} .$$

Since $\operatorname{tr}(N) = \operatorname{tr}(jN^{\top}j)$, we obtain (4.37) by inserting j^2 between all factors. Note that for the second equality we also used that for arbitrary 2×2 matrices M and N, one has

$$\operatorname{tr}\left(M \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} N \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = \operatorname{tr}\left(M_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} N \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right),$$

where M_{11} is the $(1,1)$-entry of M. Recalling that $b_{z_k}(A) = \Lambda(\vartheta)^* b_{z_k}(\tilde{A})\Lambda(\vartheta)$, the result follows. ■

For the final result of this section, the adaptation of the proof for repeated points in the vector $\vec{z}$ is not straightforward, so we will drop our assumption of distinct z_j's at this place. By reordering the entries of $\vec{z}$, if needed, we can assume that the entries of $\vec{z}$ with higher multiplicity are ordered consecutively; if $z_j = z_{j+1}$ we will denote both by z_j. For a given vector $\vec{z}$, suppose there are $m \le n$ distinct entries z_j, and let m_j denote the multiplicity with which the point z_j appears in $\vec{z}$, such that $m_0 + \cdots + m_{m-1} = n$. We call a rational function Δ *suitable* for $\vec{z}$ if it is of the form

$$\Delta(z) = c + \sum_{j=0}^{m-1} \sum_{i=1}^{m_j} \left(c_{ij} b_{z_j}(z)^i + \overline{c_{ij}} b_{z_j}(z)^{-i} \right).$$

The following result will be used in the proof of Theorem 1.5 and relies fundamentally on the structure of MCMV matrices obtained in Theorem 4.10.

Proposition 4.11. Let $\vec{z} \in \mathbb{D}^n$ with $z_0 = 0$, let $A \in \mathbb{A}(\vec{z})$, and suppose Δ is a rational function that is suitable for $\vec{z}$. Then $\Delta(A) = 0$ implies that $\Delta(z) = 0$ for all z.

Proof. For the sake of simplicity, let us assume that $e^{i\vartheta} = 1$ (the general case is analogous). The key to the proof is to understand the structure of the powers A^i for $1 \leq i \leq m_0$. Due to Proposition 4.9, the structure of the powers $b_{z_j}(A)^i$ will then follow by shifting. Recall that m_0 denotes the multiplicity with which $z_0 = 0$ is represented in the vector $\vec{z}$. We analyze the structure of the $2n \times 2n$ block of A formed by the entries $\{A_{ij}\}_{i,j=0}^{2n-1}$.

Since we have seen that the block structure is obtained by repeated z_0 entries, our $2n \times 2n$ block splits up into $m_0 - 1$ diagonal blocks of size 2×2 and a possibly bigger diagonal block of size $2(n - m_0 + 1) \times 2(n - m_0 + 1)$. Let us denote the 2×2 block matrices by $\mathbf{A_0^1}, \ldots, \mathbf{A_0^{m_0-1}}$ and the larger block by $\mathbf{A}$. On each side of the diagonal blocks $\mathbf{A_0^k}$, there is a column vector of size 2; we write $\mathbf{v_0^k}$ for the vector to the left of $\mathbf{A_0^k}$ and $\mathbf{u_0^k}$ for the vector to the right. Similarly, on each side of $\mathbf{A}$ there is a column vector of size $2(n - m_0 + 1)$, which we will denote analogously by $\mathbf{v}$ and $\mathbf{u}$. In particular, for $0 \leq k \leq m_0 - 2$ we find due to (36) and (4.37) that $A_{2k,2k+2} = \rho_{2k}\rho_{2k+1} > 0$. Moreover, the first entry of $\mathbf{u}$ is the nonzero value $A_{2(m_0-1),2n} = \rho \neq 0$; this entry is the only non-vanishing entry of A on the $2(n - m_0) + 2$-th diagonal. If we consider A^2, we get exactly two non-vanishing entries on the $2(n - m_0) + 4$-th diagonal, given by

$$A_{2(m_0-2),2n} = \rho_{2(m_0-2)}\rho_{2(m_0-2)+1}\rho$$

and

$$A_{2(m_0-1),2(n+1)} = \rho\rho_{2n}\rho_{2n+1}.$$

Similarly, A^i will have i non-vanishing entries on the outermost $2(n - m_0) + 2i$-th diagonal; in particular, A^{m_0} will have m_0 non-vanishing entries on the $2n$-th diagonal given by

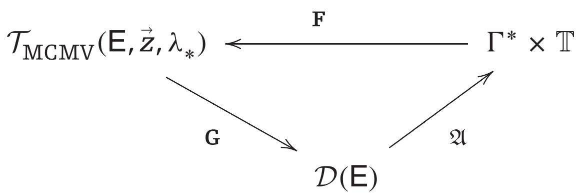
$$\begin{aligned}
A^{m_0}_{0,2n} &= \rho_0\rho_1 \cdots \rho_{2(m_0-2)}\rho_{2(m_0-2)+1}\rho, \\
A^{m_0}_{2j,2(n+j)} &= \rho_{2j}\rho_{2j+1} \cdots \rho_{2(m_0-2)}\rho_{2(m_0-2)+1}\rho\rho_{2n}\rho_{2n+1} \cdots \rho_{2(n+j)}\rho_{2(n+j)+1}, \\
A^{m_0}_{2(m_0-1),2(n+m_0-1)} &= \rho\rho_{2n}\rho_{2n+1} \cdots \rho_{2(n+m_0-2)}\rho_{2(n+m_0-2)+1}
\end{aligned}$$

for $1 \leq j < m_0 - 1$. A similar structure, but shifted, is obtained for all the matrices $b_{z_j}(A)^i$.

With this structure in mind, we can finish the proof. We first consider the entries $\Delta(A)_{j,2n+j}$. On this diagonal, only the operators $b_{z_j}(A)^{\pm m_j}$ have non-vanishing entries (and we can guarantee that they are non-vanishing). But all of them are at different positions. Hence, we see that $c_{m_j,j} = 0$ for all $0 \leq j \leq m$. In the next step, we consider the diagonal $\Delta(A)_{j,2n-2+j}$ and obtain analogously that $c_{m_j-1,j} = 0$. Inductively we see that all coefficients vanish, and consequently $\Delta \equiv 0$. ■

5 Proofs of the Main Theorems

We have laid nearly all the groundwork necessary to complete the proofs of our main theorems. Before we proceed, we recall the general strategy: in Section 3, we showed that, in the presence of a function $\mathfrak{B} = \mathfrak{B}_{\vec{z}}$ having half-period character, there is a map $\mathbf{F} : \Gamma^* \times \mathbb{T} \to \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$ taking unimodular characters and a phase (α, τ) to a periodic MCMV matrix $\mathbf{F}(\alpha, \tau) := A(\alpha, \tau)$ satisfying a Magic Formula. In Section 4, we defined a map $\mathbf{G}$ assigning a divisor to a periodic MCMV matrix. In this section we glue these constructions together via a third and final map, the Abel map $\mathfrak{A} : \mathcal{D}(\mathsf{E}) \to \Gamma^* \times \mathbb{T}$, and show that together they in fact form a commuting diagram in analog to (4.4):



(5.1)

Once this is done, our main theorems will follow as special cases for a particular choice of function $\mathfrak{B}$ and vector $\vec{z}_{\mathsf{E}}$.

To properly introduce the Abel map, we briefly recall the construction from [25] of the bijective correspondences between the isospectral torus $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$, the set of Schur functions $\mathcal{S}_+(\mathsf{E})$ defined in (1.34), and the set of divisors $\mathcal{D}(\mathsf{E})$ given by (4.2).

We first set up the correspondence $\mathcal{S}_+(E) \simeq \mathcal{D}(E)$. Consider a function $f_+ \in \mathcal{S}_+(\mathsf{E})$ and let

$$F_+(z) := \frac{1 + zf_+(z)}{1 - zf_+(z)}$$

denote the associated Caratheodory function. Due to (1.31), we have

$$\frac{1 + \overline{e^{it} f_+(e^{it})}}{1 - \overline{e^{it} f_+(e^{it})}} = \frac{1 + f_-(e^{it})}{1 - f_-(e^{it})} \quad \text{for a.e. } e^{it} \in \mathsf{E}.$$

Strictly speaking,

$$F_-(z) := \frac{1+f_-(z)}{1-f_-(z)}$$

is not a Caratheodory function since it does not admit the normalization $F_-(0) = 1$; however, it still maps $\mathbb{D}$ analytically into the right half-plane. Using (1.32), we see directly that the function

$$F(z) := \frac{1}{2}\Big(F_+(z) + F_-(z)\Big) = \frac{1 - zf_+(z)f_-(z)}{\big(1 - zf_+(z)\big)\big(1 - f_-(z)\big)}$$

has no zeros in the gaps of E. Hence, since $t \mapsto \operatorname{Im} F(e^{it})$ is decreasing on each gap, this function can have at most one sign change per gap, caused by a possible pole x_j (where j indexes the j-th gap). The measure ν in the integral representation of F is purely absolutely continuous on E, and condition (1.31) implies that ν_{ac} is split equally between F_+ and F_-. But due to (1.32), the point mass at x_j can only correspond to either F_+ or F_- and not to both. We write $(x_j, 1)$ if x_j is a pole of F_+ and $(x_j, -1)$ if x_j is a pole of F_-. Special consideration is needed for the endpoints of the gaps: by convention, we write $(\lambda_j^-, 1)$ if $\operatorname{Im} F \le 0$ in the closed gap $[\lambda_j^-, \lambda_j^+]$ and $(\lambda_j^+, 1)$ if $\operatorname{Im} F \ge 0$. With these choices, the collection

$$D = \{(x_j, \varepsilon_j)\}_{j=0}^{g}, \quad \varepsilon_j \in \{\pm 1\}$$

is the divisor in $\mathcal{D}(\mathsf{E})$ associated to f_+. Conversely, one can show that any divisor $D \in \mathcal{D}(\mathsf{E})$ leads to a function $f_+ \in \mathcal{S}_+(\mathsf{E})$ (see [25, Theorem 1.4] for details).

The correspondence $\mathcal{T}_{\mathrm{CMV}}(\mathsf{E}) \simeq \mathcal{S}_+(E)$ is implicitly given in Section 3.2. The half-line restriction $\mathcal{C}_+$ of an element $\mathcal{C} = \mathcal{C}(\alpha, \tau) \in \mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$ is linked to the Schur function $f_+^{\alpha,\tau}$ given by

$$f_+^{\alpha,\tau} \circ \mathfrak{z} = e^{-i\tau} \frac{K_{\overline{\zeta}_0}^{\alpha}}{K_{\zeta_0}^{\alpha}} \tag{5.2}$$

through the relation

$$\frac{1 + zf_+^{\alpha,\tau}(z)}{1 - zf_+^{\alpha,\tau}(z)} = \Big\langle \big(\mathcal{C}_+(\alpha,\tau) - z\big)^{-1}\big(\mathcal{C}_+(\alpha,\tau) + z\big)\delta_0, \delta_0\Big\rangle.$$

In fact, the map $(\alpha,\tau) \mapsto f_+^{\alpha,\tau}$ sets up a bijection between $\Gamma^* \times \mathbb{T}$ and $\mathcal{S}_+(\mathsf{E})$. This also enables us to define the Abel map $\mathfrak{A} : \mathcal{D}(\mathsf{E}) \to \Gamma^* \times \mathbb{T}$ by

$$\mathfrak{A}(D) := (\alpha,\tau), \tag{5.3}$$

where (α,τ) is the character and phase of the function f_+, which corresponds to the divisor D. Note that (5.3) generalizes the definition of the Abel map for periodic CMV matrices; cf. (4.4).

The first lemma of this section demonstrates that the diagram (5.1) commutes if we replace $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E},\vec{z},\lambda_*)$ with the image $\mathbf{F}(\Gamma^* \times \mathbb{T})$. Let $\tilde{D}$ be the divisor associated to the periodic operator $A(\alpha,\tau) \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E},\vec{z},\lambda_*)$ by our construction in Section 4.2, that is, $\tilde{D} = \mathbf{G}(\mathbf{F}(\alpha,\tau))$.

Lemma 5.1. Let $\tilde{f}_+$ and $\tilde{F}_+$ be the Schur and Caratheodory functions associated to the periodic operator $A(\alpha,\tau)$ as in Section 4, and let $f_+^{\alpha,\tau}$ and $F_+^{\alpha,\tau}$ be the functions associated to (α,τ) by (5.2) above. Then we have

$$\tilde{f}_+ = f_+^{\alpha,\tau}, \quad \tilde{F}_+ = F_+^{\alpha,\tau}, \quad \text{and } \mathfrak{A}(\tilde{D}) = (\alpha,\tau). \tag{5.4}$$

Proof. By (3.44) and (3.45), we have

$$f_+^{\alpha,\tau} \circ \mathfrak{z} = \frac{X_0^{\alpha,\tau}}{Y_0^{\alpha,\tau}}.$$

Therefore, the identities $\tilde{f}_+ = f_+^{\alpha,\tau}$ and $\tilde{F}_+ = F_+^{\alpha,\tau}$ follow by comparing (4.9) with Theorem 3.11.

We have now seen that $(\tilde{x}_j, 1)$ corresponds to poles of $\tilde{F}_+$. To see that $\mathfrak{A}(\tilde{D}) = (\alpha,\tau)$, we only need to show that $(\tilde{x}_j, -1)$ as defined in Section 4 corresponds to poles of the function $\tilde{F}_-$. Due to (4.23), we have

$$\tilde{F}_+(z) = \frac{v(z) + \sqrt{\Delta_A(z)^2 - 4}}{u(z)}, \tag{5.5}$$

where

$$v(e^{it}) \in \mathbb{R} \quad \text{and} \quad u(e^{it}) \in i\mathbb{R}.$$

Consider now the function

$$\tilde{F}_-(z) := -\frac{v(z) - \sqrt{\Delta_A(z)^2 - 4}}{u(z)}. \tag{5.6}$$

Since $\sqrt{\Delta_A(e^{it})^2 - 4} \in i\mathbb{R}$ on E, we obtain that $\overline{\tilde{F}_+(e^{it})} = \tilde{F}_-(e^{it})$ for all $e^{it} \in \mathsf{E}$. Hence, if $\tilde{x}_j$ is a zero of u and the numerator in (5.5) vanishes (i.e., $\varepsilon_j = -1$), then $\tilde{F}_-$ has a pole at $\tilde{x}_j$. This concludes the proof. ■

Remark. Note that (5.5) and (5.6) show that the absolutely continuous parts of the corresponding measures agree and are given by (4.25).

In order to show commutativity of (5.1), we still have to show that for a given $\vec{z}$ and λ_*, there is a functional model construction $\mathbf{F}$ that surjects onto those MCMV matrices in $\mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$ having a fixed spectrum; that is, for an arbitrary $\vec{z}$ and $A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$, there exists a function $\mathfrak{B} = \prod_{j=0}^{n-1} \mathfrak{b}_{\zeta_j}$ whose character is a half-period such that $A = A(\alpha, \tau)$ corresponds to the functional model associated to $\mathfrak{B}$ and normalization at λ_*. To show that the character of $\mathfrak{B}$ is a half-period, the representation of Δ_A from Lemma 4.4 will be crucial.

Proposition 5.2. Let $A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$ be a periodic MCMV matrix with associated discriminant Δ_A, and set $\mathsf{E} = \Delta_A^{-1}([-2, 2])$. Then there exists a unique divisor $D \in \mathcal{D}(\mathsf{E})$ such that

$$A = A(\alpha, \tau), \tag{5.7}$$

where $\mathfrak{A}(D) = (\alpha, \tau)$ and $A(\alpha, \tau)$ is defined by (3.62) for the functional model associated to $\mathfrak{B} = \prod_{j=0}^{n-1} \mathfrak{b}_{\zeta_j}$. Moreover, the diagram (5.1) commutes.

Proof. Suppose A is periodic and let D be the associated divisor. Write $(\alpha, \tau) = \mathfrak{A}(D)$ for the corresponding character and phase. Due to Lemma 4.4, we have

$$\Delta_A \circ \mathfrak{z} = \Psi + \frac{1}{\Psi}, \quad \Psi = \mathfrak{B}\mathfrak{B}^*.$$

Since the characters of $\mathfrak{B}$ and $\mathfrak{B}^*$ coincide and since Ψ is single-valued, the character of $\mathfrak{B}$ must be a half-period.

Suppose now that $A(\alpha,\tau)$ is the matrix representing multiplication by $\mathfrak{z}$ in the basis $\{Y_k^{\alpha,\tau}\}$ for the functional model associated to $\mathfrak{B}$. Then, by Lemma 5.1, the Caratheodory functions of A and $A(\alpha,\tau)$ coincide and hence we obtain (5.7). Since it was already shown in [25] that the Abel map is a bijection, this also proves the commutativity of (5.1). ■

We are now in position to complete the proofs of our main results stated in the introduction. As already mentioned in the first remark of Section 3.3, the existence of the Ahlfors function for an arbitrary finite-gap set E ensures that there always is a function $\mathfrak{B}$ with the property that its character is a half-period. To be specific, the function $\mathfrak{w}_\infty = \mathfrak{z}(w_\infty \circ \mathfrak{z})$ is an explicit choice of such a function. Our first main result, Theorem 1.4, now simply follows as the special case of Proposition 5.2 with $\vec{z} = \vec{z}_{\mathsf{E}}$ (and $\mathfrak{B} = \mathfrak{w}_\infty$).

Proof of Theorem 1.4. For the special choice of $\vec{z} = \vec{z}_{\mathsf{E}}$, we define the map $\mathbf{F} : \Gamma^* \times \mathbb{T} \to \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*)$ in the same way as was done previously. That is,

$$\mathbf{F}(\alpha,\tau) := A(\alpha,\tau),$$

where $A(\alpha,\tau)$ is the matrix representation of multiplication by $\mathfrak{z}$ in the basis $\{Y_k^{\alpha,\tau}\}$ for the functional model associated to the function $\mathfrak{B} = \mathfrak{w}_\infty$. Our considerations have shown this $\mathbf{F}$ is a bijection and that concludes the proof of (1.18).

Tracing through our construction will also lead to the more explicit version of the one-to-one correspondence between an element $A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}, \lambda_*)$ and its counterpart $C \in \mathcal{T}_{\mathrm{CMV}}(\mathsf{E})$. ■

Our analysis also leads to a quick proof of the Magic Formula, our second main result.

Proof of Theorem 1.5. Let A be an element of $\mathbb{A}(\vec{z}_{\mathsf{E}})$. If $A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_{\mathsf{E}}, \lambda_*)$, then we know from Proposition 5.2 that $A = A(\alpha,\tau)$ for some choice of character and phase. Hence the Magic Formula

$$\Delta_{\mathsf{E}}(A) = S^{2(g+1)} + S^{-2(g+1)}$$

follows by Corollary 3.13 with $n = g+1$.

Conversely, if A satisfies the Magic Formula then A must be periodic with period $2(g+1)$ due to Naiman's Lemma (see, e.g., [32, Lemma 8.2.4]). Moreover, as

$$\Delta_{\mathsf{E}}(A) = S^{2(g+1)} + S^{-2(g+1)} = \Delta_A(A),$$

Proposition 4.11 implies that $\Delta_A = \Delta_{\mathsf{E}}$. In particular, we have positivity of Δ_A in all gaps. Since $\sigma(A) = \Delta_A^{-1}([-2,2])$, it therefore follows that $\sigma(A) = \mathsf{E}$ and thus $A \in \mathcal{T}_{\mathrm{MCMV}}(\mathsf{E}, \vec{z}_E, \lambda_*)$. The second part of the theorem follows along the same lines as in the proof of Theorem 1.4. ■

Finally, Theorem 1.6 is immediate:

Proof of Theorem 1.6. This is a straightforward consequence of Theorems 1.4, 1.5, and 3.11. ■

A The Ahlfors Functions of Finitely Connected Denjoy Domains

We begin with an existence theorem:

Theorem A.1. Let Ω be a region in $\overline{\mathbb{C}}$ and fix a point $z_0 \in \Omega$. Suppose there exist nonconstant bounded analytic functions defined on Ω. Then there is a unique analytic function $w_{z_0} : \Omega \to \mathbb{D}$, which solves the Ahlfors problem, that is,

$$w'_{z_0}(z_0) = \sup\{|g'(z_0)| : g : \Omega \to \mathbb{D} \text{ analytic}\}. \tag{A.1}$$

The function w_{z_0} is called the *Ahlfors function* of Ω (and z_0) and we always have that $w_{z_0}(z_0) = 0$; cf. [33, Theorem 8.8.1]. These functions were first studied for finitely connected domains by Ahlfors [1], hence the name. Existence and uniqueness for arbitrary domains was later established by Fisher [11].

In [10], an explicit expression for the Ahlfors function of finitely connected domains with certain symmetries was presented. Let $0 < a_j < b_j$ for $1 \le j \le g$ and define

$$\mathsf{E}_{\mathbb{R}} := \mathbb{R}_+ \setminus \bigcup_{j=1}^{g} (a_j, b_j), \quad \Omega_{\mathbb{R}} := \mathbb{C} \setminus \mathsf{E}_{\mathbb{R}}. \tag{A.2}$$

Since the Ahlfors problem is conformally invariant, providing a solution for $\Omega_{\mathbb{R}}$ also leads to a solution for all conformally equivalent domains. In particular, for any finite systems of arcs $\mathsf{E}_{\mathbb{T}}$ on the unit circle, we can map $\mathbb{D}$ conformally onto $\mathbb{C}_+$ such that

$\mathsf{E}_{\mathbb{T}}$ corresponds to some set $\mathsf{E}_{\mathbb{R}}$ of the form (A.2). For notational simplicity, we merely present the results for sets of the form (A.2). To any such set, we associate the function

$$H(z) = \frac{1}{\sqrt{-z}} \prod_{j=1}^{g} \sqrt{\frac{z-a_j}{z-b_j}}, \tag{A.3}$$

where the square root is chosen in such a way that $H(z)$ becomes a Nevanlinna function (i.e., maps $\mathbb{C}_+$ analytically into $\mathbb{C}_+$). The following theorem generalizes a result of Pommerenke [26] (who dealt with the case of $z_0 \in \mathbb{R} \setminus \mathsf{E}_{\mathbb{R}}$).

Theorem A.2. [10, Theorem 2.3] The Ahlfors function of $\Omega_{\mathbb{R}}$ and z_0 is given explicitly by

$$w_{z_0}(z) = \frac{z-z_0}{z-\overline{z_0}} \frac{H(z)-H(\overline{z_0})}{H(z)+H(z_0)}. \tag{A.4}$$

If $\operatorname{Im} z_0 > 0$ then w_{z_0} has precisely g zeros, say $\overline{z_1}, \ldots, \overline{z_g}$, in the lower half-plane $\mathbb{C}_-$ and together with z_0, these points account for all the zeros of w_{z_0}. Moreover,

$$\log|w_{z_0}(z)| = -G_{\Omega_{\mathbb{R}}}(z, z_0) - \sum_{j=1}^{g} G_{\Omega_{\mathbb{R}}}(z, \overline{z_j}). \tag{A.5}$$

Now, fix $z_0 \in \mathbb{C}_+$ and consider the discriminant $\Delta_{\mathsf{E}_{\mathbb{R}}}$ defined by

$$\Delta_{\mathsf{E}_{\mathbb{R}}}(z) := w_{z_0}(z) w_{\overline{z_0}}(z) + \frac{1}{w_{z_0}(z) w_{\overline{z_0}}(z)}. \tag{A.6}$$

We collect the properties of $\Delta_{\mathsf{E}_{\mathbb{R}}}$ in the following theorem and point out that the conclusion for the critical points appears to be new.

Theorem A.3. $\Delta_{\mathsf{E}_{\mathbb{R}}}$ is a real rational function, that is, $\Delta_{\mathsf{E}_{\mathbb{R}}}(z) = \overline{\Delta_{\mathsf{E}_{\mathbb{R}}}(\overline{z})}$. Its poles are given by $z_0, \overline{z_0}, \ldots, z_g, \overline{z_g}$, where $\overline{z_1}, \ldots, \overline{z_g}$ are the zeros of w_{z_0} in $\mathbb{C}_-$. Moreover,

$$\mathsf{E}_{\mathbb{R}} = \Delta_{\mathsf{E}_{\mathbb{R}}}^{-1}([-2, 2]) \tag{A.7}$$

and $\Delta_{\mathsf{E}_{\mathbb{R}}}$ has exactly one critical point, say c_{2j}, in the j-th gap (a_j, b_j) with $\Delta_{\mathsf{E}_{\mathbb{R}}}(c_{2j}) > 2$ and exactly one critical point, say c_{2j+1}, in the j-th band (b_j, a_{j+1}) with $\Delta_{\mathsf{E}_{\mathbb{R}}}(c_{2j+1}) = -2$.

Proof. Since $\overline{H(\overline{z})} = H(z)$, we see that $\overline{w_{z_0}(z)} = w_{\overline{z_0}}(\overline{z})$. Also, for $z \in \mathsf{E}_{\mathbb{R}}$, one has

$$\lim_{\varepsilon \to 0} \Delta_{\mathsf{E}_{\mathbb{R}}}(z + i\varepsilon) = \lim_{\varepsilon \to 0} \Delta_{\mathsf{E}_{\mathbb{R}}}(z - i\varepsilon).$$

Hence $\Delta_{\mathsf{E}_{\mathbb{R}}}$ is a real rational function on $\overline{\mathbb{C}}$. Note that in (A.6) we have written $\Delta_{\mathsf{E}_{\mathbb{R}}}$ as the composition of $w_{z_0} w_{\overline{z_0}}$ and the Joukowsky map $u \mapsto u + 1/u$. Since $|w_{z_0}(z)| = 1$ if and only if $z \in \mathsf{E}_{\mathbb{R}}$, this proves (A.7). Moreover, since H is a Nevanlinna function, it decreases monotonically from $i\infty$ to 0 as z moves along a band $[b_j, a_{j+1}]$. Setting $H(z_0) = H_0$, it follows from the previous theorem that

$$w_{z_0}(z)w_{\overline{z_0}}(z) = \frac{H(z) - H_0}{H(z) + \overline{H_0}} \frac{H(z) - \overline{H_0}}{H(z) + H_0}.$$

Recall now that $\frac{z-H_0}{z+\overline{H_0}}$ is the Blaschke factor of the right half-plane. Therefore, as $H(z)$ decreases from $i\infty$ to 0, the values of $w_{z_0}w_{\overline{z_0}}$ run through $\partial\mathbb{D}$ precisely once, starting and ending at 1. Hence there is exactly one point $c_{2j+1} \in (b_j, a_{j+1})$ with $w_{z_0}(c_{2j+1})w_{\overline{z_0}}(c_{2j+1}) = -1$, that is, $\Delta_{\mathsf{E}_{\mathbb{R}}}(c_{2j+1}) = -2$. To analyze the behavior in the gaps, we note that in each gap $H(z)$ increases monotonically from 0 to ∞. Considering the function

$$\psi(x) = \frac{x - H_0}{x + \overline{H_0}} \frac{x - \overline{H_0}}{x + H_0}$$

on $\mathbb{R}_+$ shows that there is exactly one critical point in each gap. ■

B Operator Möbius Transforms

Let $\mathcal{H}$ be a Hilbert space and denote by $\mathcal{L}(\mathcal{H})$ the space of bounded linear operators from $\mathcal{H}$ into itself equipped with the standard operator norm. By $\mathbb{D}(\mathcal{H})$ (resp. $\overline{\mathbb{D}(\mathcal{H})}$) we denote the open (resp. closed) unit ball in $\mathcal{L}(\mathcal{H})$, that is, $\mathbb{D}(\mathcal{H})$ is the set of contractions on $\mathcal{H}$. We will write operators $U \in \mathcal{L}(\mathcal{H} \oplus \mathcal{H})$ in matrix form:

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}, \quad U_{ij} \in \mathcal{L}(\mathcal{H}). \tag{B.1}$$

To such matrices we can associate a linear fractional transformation

$$\Phi_U(S) := (SU_{12} + U_{22})^{-1}(SU_{11} + U_{21}), \tag{B.2}$$

defined for those $S \in \mathcal{L}(\mathcal{H})$ for which $SU_{12} + U_{22}$ is boundedly invertible. It is straightforward to see that $\Phi_U(\Phi_V(S)) = \Phi_{VU}(S)$ and that for any $\lambda \neq 0$, the operators U and λU generate the same transform.

In [17], a complete characterization of the class of operators U such that Φ_U is a bijection from $\mathbb{D}(\mathcal{H})$ onto $\mathbb{D}(\mathcal{H})$ was given. In this case, Φ_U is called an *operator Möbius transform*. The characterization involves the special operator

$$j = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

where as usual 1 denotes the identity operator on $\mathcal{H}$. An operator $U \in \mathcal{L}(\mathcal{H} \oplus \mathcal{H})$ is called *j-unitary* if $U^*jU = j$.

Theorem B.1. [17] Φ_U is an operator Möbius transform if and only if U is colinear with a j-unitary operator.

A complete description of all j-unitary operators (and hence all operator Möbius transforms) is known:

Theorem B.2. [17] The general form of a j-unitary operator is

$$U = \begin{bmatrix} 1 & A^* \\ A & 1 \end{bmatrix} \begin{bmatrix} \eta_{A^*}^{-1} & 0 \\ 0 & \eta_A^{-1} \end{bmatrix} \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}, \tag{B.3}$$

where $A \in \mathbb{D}(\mathcal{H})$, $\eta_A = \sqrt{1 - AA^*}$, and V_1, V_2 are unitary operators on $\mathcal{H}$.

In what follows we will derive an expression for the inverse of the operator Möbius transform corresponding to a contraction $A \in \mathbb{D}(\mathcal{H})$. Let

$$U_A = \begin{bmatrix} 1 & A^* \\ A & 1 \end{bmatrix} \begin{bmatrix} \eta_{A^*}^{-1} & 0 \\ 0 & \eta_A^{-1} \end{bmatrix} \text{ and } \tilde{U}_A = \begin{bmatrix} \eta_{A^*}^{-1} & 0 \\ 0 & \eta_A^{-1} \end{bmatrix} \begin{bmatrix} 1 & -A^* \\ -A & 1 \end{bmatrix}.$$

Then

$$\tilde{U}_A U_A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and it follows that $\Phi_{\tilde{U}_A}\big(\Phi_{U_A}(S)\big) = S$. Moreover, noticing that for $n \in \mathbb{N}$ we have $(\eta_A^2)^n A = A(\eta_{A^*}^2)^n$ and then applying the functional calculus, it follows that for every function f that is continuous on $\mathbb{R}_{>0}$, we have $f(\eta_A^2)A = Af(\eta_{A^*}^2)$. Hence, in particular, $\eta_A^{-1}A = A\eta_{A^*}^{-1}$ and $\eta_{A^*}^{-1}A^* = A^*\eta_A^{-1}$. This commutativity relation yields

$$\tilde{U}_A = \begin{bmatrix} 1 & -A^* \\ -A & 1 \end{bmatrix} \begin{bmatrix} \eta_{A^*}^{-1} & 0 \\ 0 & \eta_A^{-1} \end{bmatrix}$$

which corresponds to the fractional linear transform

$$\Phi_{\tilde{U}_A}(S) = \eta_A(1 - SA^*)^{-1}(S - A)\eta_{A^*}^{-1}. \tag{B.4}$$

Thus we see that $\tilde{U}_A = U_{-A}$, in analog to the scalar case.

Funding

This work was supported by the Swedish Research Council (VR) (grant no. 2018-03500 to J.S.C.); the Danish Council for Independent Research (project grant DFF-4181-00502 to J.S.C.); the

Austrian Science Fund FWF (project no. J 4138-N32 to B.E.); AMS Simons travel grant 2018–2020 (to T.V.).

Acknowledgments

It is a pleasure to thank Peter Yuditskii for very helpful discussions. We also wish to thank the anonymous referees whose comments greatly helped to improve the paper.

References

[1] Ahlfors, L. V. "Bounded analytic functions." *Duke Math. J.* 14 (1947): 1–11.

[2] Akhiezer, N. I. "Orthogonal polynomials on several intervals." *Dokl. Akad. Nauk SSSR 134 (1960): 9–12 (Russian)* translated as Soviet Math. Dokl 1 (1960): 989–92.

[3] Aptekarev, A. I. "Asymptotic properties of polynomials orthogonal on a system of contours, and periodic motions of Toda chains." *Math. USSR Sb. 53, no. 2 (1986): 233–60* Russian original in Mat. Sb. (N.S.) 125, no. 167 (1984): 231–58.

[4] Arlinskii, Y., L. Golinskii, and E. Tsekanovskii. "Contractions with rank one defect operators and truncated CMV matrices." *J. Funct. Anal.* 254, no. 1 (2008): 154–95.

[5] Bultheel, A., P. González-Vera, E. Hendriksen, and O. Njåstad. "Orthogonal Rational Functions." *Cambridge Monographs on Applied and Computational Mathematics*, vol. 5. Cambridge: Cambridge University Press, 1999.

[6] Cantero, M. J., F. A. Grünbaum, L. Moral, and L. Velázquez. "Matrix-valued Szegő polynomials and quantum random walks." *Comm. Pure Appl. Math.* 63, no. 4 (2010): 464–507.

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

[8] Damanik, D., R. Killip, and B. Simon. "Perturbations of orthogonal polynomials with periodic recursion coefficients." *Ann. Math.* 171, no. 3 (2010): 1931–2010.

[9] Eichinger, B. "Periodic GMP matrices." *SIGMA Symmetry Integrability Geom. Methods Appl.* 12 (2016 Paper No. 066): 19.

[10] Eichinger, B. and P. Yuditskii. "Ahlfors problem for polynomials." *Mat. Sb.* 209, no. 3 (2018): 34–66.

[11] Fisher, S. D. "On Schwarz's lemma and inner functions." *Trans. Amer. Math. Soc.* 138 (1969): 229–40.

[12] Garabedian, P. R. "Schwarz's lemma and the Szegő kernel function." *Trans. Amer. Math. Soc.* 67 (1949): 1–35.

[13] Gesztesy, F. and M. Zinchenko. "Weyl–Titchmarsh theory for CMV operators associated with orthogonal polynomials on the unit circle." *J. Approx. Theory* 139, no. 1–2 (2006): 172–213.

[14] Hasumi, M. "Hardy Classes on Infinitely Connected Riemann Surfaces." *Lecture Notes in Mathematics*, vol. 1027. Berlin: Springer, 1983.

[15] Killip, R. and I. Nenciu. "Matrix models for circular ensembles." *Int. Math. Res. Not.* 50 (2004): 2665–701.

[16] Killip, R. and I. Nenciu. “CMV: the unitary analogue of Jacobi matrices.” *Comm. Pure Appl. Math.* 60, no. 8 (2007): 1148–88.

[17] Krein, M. G. and J. L. Smuljan. “Fractional linear transformations with operator coefficients.” *Mat. Issled* 2, no. vyp. 3 (1967): 64–96.

[18] Kričever, I. M. “Algebraic curves and nonlinear difference equations.” *Uspekhi Mat. Nauk* 33, no. 4(202) (1978): 215–6.

[19] Mumford, D. “Tata Lectures on Theta. I.” *Modern Birkhäuser Classics*. Boston, MA: Birkhäuser Boston, Inc, 2007.

[20] Mumford, D. “Tata Lectures on Theta. II.” *Modern Birkhäuser Classics*. Boston, MA: Birkhäuser Boston, Inc, 2007, Jacobian theta functions and differential equations, With the collaboration of C. Musili, M. Nori, E. Previato, M. Stillman, and H. Umemura, Reprint of the 1984 original.

[21] Naiman, P. B. “On the theory of periodic and limit-periodic Jacobian matrices.” *Dokl. Akad. Nauk SSSR* 143 (1962): 277–9.

[22] Nenciu, I. “CMV matrices in random matrix theory and integrable systems: a survey.” *J. Phys. A* 39, no. 28 (2006): 8811–22.

[23] Nikolskii, N. K. “Treatise on the Shift Operator.” *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 273. Berlin: Springer, 1986, Translated from the Russian by Jaak Peetre.

[24] Pastur, L. and A. Figotin. “Spectra of Random and Almost-Periodic Operators.” *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 297. Berlin: Springer, 1992.

[25] Peherstorfer, F. and P. Yuditskii. “Almost periodic Verblunsky coefficients and reproducing kernels on Riemann surfaces.” *J. Approx. Theory* 139, no. 1–2 (2006): 91–106.

[26] Pommerenke, C. “Über die analytische Kapazität.” *Archiv der Mathematik* 11, no. 1 (1960): 270–7, German.

[27] Pommerenke, C. “On the Green’s function of Fuchsian groups.” *Ann. Acad. Sci. Fenn. Ser. A I Math.* 2 (1976): 409–27.

[28] Potapov, V. P. “The multiplicative structure of J-contractive matrix functions.” *Amer. Math. Soc. Transl. (2)* 15 (1960): 131–243.

[29] Schlag, W. “A Course in Complex Analysis and Riemann Surfaces.” *Graduate Studies in Mathematics, vol. 154*. Providence, RI: American Mathematical Society, 2014.

[30] Simon, B. “Orthogonal Polynomials on the Unit Circle. Part 1.” *American Mathematical Society Colloquium Publications*, vol. 54. Providence, RI: American Mathematical Society, 2005, Classical theory.

[31] Simon, B. “*Orthogonal Polynomials on the Unit Circle. Part 2.” American Mathematical Society Colloquium Publications*, vol. 54. Providence, RI: American Mathematical Society, 2005, Spectral theory.

[32] Simon, B. “Szegő’s Theorem and Its Descendants.” *M. B. Porter Lectures*. Princeton, NJ: Princeton University Press, 2011, Spectral theory for L^2 perturbations of orthogonal polynomials.

[33] Simon, B. “Basic Complex Analysis.” *A Comprehensive Course in Analysis, Part 2A*. Providence, RI: American Mathematical Society, 2015.

[34] Simon, B., “Harmonic analysis.” *A Comprehensive Course in Analysis*, Part 3. Providence, RI: American Mathematical Society, 2015.

[35] Sodin, M. and P. Yuditskii. “Almost periodic Jacobi matrices with homogeneous spectrum, infinite-dimensional Jacobi inversion, and hardy spaces of character-automorphic functions.” *J. Geom. Anal.* 7, no. 3 (1997): 387–435.

[36] Stone, M. H. “Linear Transformations in Hilbert Space.” *American Mathematical Society Colloquium Publications, vol. 15*. Providence, RI: American Mathematical Society, 1990, Reprint of the 1932 original.

[37] Velázquez, L. “Spectral methods for orthogonal rational functions.” *J. Funct. Anal.* 254, no. 4 (2008): 954–86.

[38] Widom, H. “$\mathcal{H}_p$ sections of vector bundles over Riemann surfaces.” *Ann. Math*. 94 (1971): 304–24.

[39] Yuditskii, P. “Killip–Simon problem and Jacobi flow on GMP matrices.” *Adv. Math.* 323 (2018): 811–65.

STAHL–TOTIK REGULARITY FOR CONTINUUM SCHRÖDINGER OPERATORS

BENJAMIN EICHINGER AND MILIVOJE LUKIĆ

ABSTRACT. We develop a theory of regularity for continuum Schrödinger operators based on the Martin compactification of the complement of the essential spectrum. This theory is inspired by Stahl–Totik regularity for orthogonal polynomials, but requires a different approach, since Stahl–Totik regularity is formulated in terms of the potential theoretic Green function with a pole at ∞, logarithmic capacity, and the equilibrium measure, notions which do not extend to unbounded spectra. For any half-line Schrödinger operator with a bounded potential (in a locally L^1 sense), we prove that its essential spectrum obeys the Akhiezer–Levin condition, and moreover, that the Martin function at ∞ obeys the two-term asymptotic expansion $\sqrt{-z} + \frac{a}{2\sqrt{-z}} + o(\frac{1}{\sqrt{-z}})$ as $z \to -\infty$. The constant a in that expansion has not appeared in the literature before; we show that it can be used to measure the size of the spectrum in a potential theoretic sense and that it should be thought of as a renormalized Robin constant suited for semibounded sets. We prove that it enters a universal inequality $a \leq \liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t$, which leads to a notion of regularity, with connections to the root asymptotics of Dirichlet solutions and zero counting measures. We also present applications to decaying and ergodic potentials.

1. INTRODUCTION

The goal of this paper is to develop a theory of Stahl–Totik regularity suitable for continuum Schrödinger operators; it is natural for this topic to work in the half-line setting, so our Schrödinger operators are unbounded self-adjoint operators on $L^2((0,\infty))$, corresponding formally to

$$L_V = -\frac{d^2}{dx^2} + V.$$

The potential V will always be real-valued and assumed to be uniformly locally integrable, i.e.

$$\sup_{x\geq 0} \int_x^{x+1} |V(t)|\mathrm{d}t < \infty \tag{1.1}$$

(in particular, 0 is a regular endpoint and $+\infty$ is a limit point endpoint in the sense of Weyl). We set the Dirichlet boundary condition at 0, so the domain of the operator is

$$D(L_V) = \{f \in L^2((0,\infty)) \mid f \in W^{2,1}_{\mathrm{loc}}([0,\infty)), -f'' + Vf \in L^2((0,\infty)), f(0) = 0\}$$

where $W^{2,1}_{\mathrm{loc}}([0,\infty))$ denotes the set of functions such that $f \in W^{2,1}([0,x])$ for all $x < \infty$, i.e., $f'' \in L^1([0,x])$ for all $x < \infty$.

The connection of orthogonal polynomials and potential theory goes back at least to the work of Faber and Szegő [33, 70]. For further references on the subject we refer to the paper of Simon [65] and the monograph of Stahl and Totik [69]. Building on important work of Ullman [71], Stahl and Totik developed a comprehensive theory for orthogonal polynomials for arbitrary measures with

B.E. was supported by Austrian Science Fund FWF, project no: J 4138-N32 and project no: P 33885.
M.L. was supported in part by NSF grant DMS–1700179.

compact support in $\mathbb{C}$. It is shown that the asymptotic behavior of the orthogonal polynomials is intimately related with so-called Stahl–Totik regularity of the measure. Regularity of the measure is then used as a reference behavior in the description of many phenomena; in spectral theory, it has important consequences through the special cases of measures supported on the real line or unit circle. For instance, on the real line, the theory provides a universal inequality between the Jacobi coefficients of a compactly supported measure and the logarithmic capacity of its topological support E, and the measure is defined to be Stahl–Totik regular if equality holds. The corresponding Jacobi matrix is then also said to be regular. This motivates the search for a similar theory for Schrödinger operators, as discussed in [65, Section 9]. However, Stahl–Totik regularity is built on potential theoretic notions, such as Green functions on the domain $\Omega = \hat{\mathbb{C}} \setminus \mathsf{E}$ with the pole at ∞, logarithmic capacity, and equilibrium measures – objects which are undefined for unbounded sets E, and therefore not applicable to continuum Schrödinger operators. For this reason, even the correct objects and extremal principles were not identified until now.

In this paper, we develop the corresponding theory for Schrödinger operators. Martin functions [3, 54] serve as the counterpart of Green functions, corresponding to boundary points $z_0 \in \partial\Omega$ instead of internal points $z_0 \in \Omega$; but whereas the Green function is defined with an explicit logarithmic singularity at z_0, the existence and behavior of Martin functions is more varied. If $\mathsf{E} \subset \mathbb{R}$ is a closed unbounded set, ∞ is a boundary point of the Denjoy domain $\Omega = \mathbb{C} \setminus \mathsf{E}$. If this domain is Greenian, associated to the boundary point ∞ is a cone of dimension 1 or 2 of positive harmonic functions in Ω which are bounded on bounded sets and vanish at every Dirichlet-regular point of E. The cone is spanned by the minimal Martin functions at ∞ [1, 2, 7, 35]. Moreover, if $\inf \mathsf{E} > -\infty$, the cone is of dimension 1, and the Martin function at ∞ is determined uniquely up to normalization; we denote it by M_{E} and simply call it the Martin function from now on.

The Akhiezer–Levin condition for semibounded sets (sets with $\inf \mathsf{E} > -\infty$) is

$$\lim_{z\to-\infty} \frac{M_{\mathsf{E}}(z)}{\sqrt{-z}} > 0 \tag{1.2}$$

(by general principles, the limit exists with a value in $[0,\infty)$). This is the semibounded version of a condition originally considered by Akhiezer–Levin [1] for arbitrary $\mathsf{E} \subset \mathbb{R}$; see also [73, Remark 1.13]. For sets obeying (1.2), we will normalize the Martin function so that the limit in (1.2) is equal to 1.

For a potential bounded in the sense (1.1), the spectrum $\sigma(L_V)$ is a closed subset of $\mathbb{R}$ bounded below but not above, so the above definitions are applicable. It will be noted that isolated points of the set don't affect the Martin function, so we can equally well use $\mathsf{E} = \sigma_{\text{ess}}(L_V)$ in what follows (more generally, $M_{\mathsf{E}_1} = M_{\mathsf{E}_2}$ if the symmetric difference of E_1 and E_2 is a polar set).

In spectral theory, Martin functions first appear implicitly, in the classical work of Marchenko–Ostrovski [55] classifying the spectra of periodic Schrödinger operators. In this work, the discriminant of a 1-periodic operator is expressed in the form $\Delta(z) = 2\cos(\Theta(z))$, and it can be recognized that $\operatorname{Im}\Theta(z)$ is the Martin function at ∞ for the periodic spectrum. The explicit use of Martin functions in spectral theory starts with works of Yuditskii and coauthors [27, 30, 68], through inverse spectral theoretic studies associated to Dirichlet-regular spectra obeying a Widom condition and finite gap length conditions.

In contrast to the previous works, our first theorem is a set of universal properties of the spectra of Schrödinger operators obeying (1.1); note that a boundedness condition such as (1.1) is essential for the following theory, since potentials going to $-\infty$ or $+\infty$ can give spectrum equal to $\mathbb{R}$ or spectrum which is a polar set.

Theorem 1.1. *For any potential V obeying* (1.1) *and* $\mathsf{E} = \sigma_{\mathrm{ess}}(L_V)$*, the domain* $\Omega = \mathbb{C} \setminus \mathsf{E}$ *is Greenian,* ∞ *is a Dirichlet-regular point for* Ω*,* Ω *obeys the Akhiezer–Levin condition, and there exists* $a_{\mathsf{E}} \in \mathbb{R}$ *such that the Martin function has the asymptotic behavior*

$$M_{\mathsf{E}}(z) = \mathrm{Re}\left(\sqrt{-z} + \frac{a_{\mathsf{E}}}{2\sqrt{-z}}\right) + o\left(\frac{1}{\sqrt{|z|}}\right), \tag{1.3}$$

as $z \to \infty$, $\arg z \in [\delta, 2\pi - \delta]$*, for any* $\delta > 0$.

Each of the conclusions of this theorem is strictly stronger than the previous; we will point out examples in Section 2. In particular, the second term of the expansion (1.3) is not an automatic property of Akhiezer–Levin sets, but rather an added feature corresponding to spectra of Schrödinger operators. It should be emphasized that spectra of Schrödinger operators with bounded potentials can be very thin in the sense that they can even have zero Hausdorff dimension [17] and zero lower box counting dimension [18], while our result is a universal "thickness" result in the perspective of the Martin function.

In the references given above, the Martin function was used in spectral theory as a positive harmonic function in Ω that vanishes on the boundary. In fact, Martin theory provides a whole kernel $M(z, x)$ on $\Omega \times (\hat{\Omega} \setminus \{z_*\})$, where $\hat{\Omega}$ denotes the Martin compactification of Ω and $z_* \in \Omega$ is a normalization point. If $\partial_1^M \Omega$ denotes the so-called minimal Martin boundary of Ω, then for every positive harmonic function h on Ω there exisists a unique finite measure ν such that

$$h(z) = \int_{\partial_1^M \Omega} M(z, x) \mathrm{d}\nu(x).$$

We will provide more details and precise definitions in Section 2. It is new to combine this theory with the spectral theory of unbounded self-adjoint operators and this was crucial for the proof of Theorem 1.1.

It is crucial that Theorem 1.1 associates to the essential spectrum E the real-valued constant a_{E}, which will serve as a substitute for the Robin constant from potential theory. Expansions of the form (1.3) have previously appeared in the spectral theory literature [55] only under strong a priori assumptions on the spectrum. Namely, the set E is closed so it can be written in the form

$$\mathsf{E} = [b_0, \infty) \setminus \bigcup_{j=1}^{N} (a_j, b_j) \tag{1.4}$$

where j indexes the "gaps", i.e., connected components of $[b_0, \infty) \setminus \mathsf{E}$, and N is finite or ∞. If $\sum_j (b_j - a_j) < \infty$, the Martin function has an expansion (1.3) with $a_{\mathsf{E}} = b_0 + \sum_j (a_j + b_j - 2c_j)$, where c_j denotes the (unique) location of the maximum of the restriction of M_{E} to the interval $[a_j, b_j]$ (see Lemma 6.2) by harmonic/complex theoretic arguments. Instead, our Theorem 1.1 applies even when the spectrum E is very thin and this is not a purely complex theoretic result; its proof is a combination of spectral theoretic arguments and the theory of the Martin boundary of Denjoy domains.

The renormalized Robin constant a_{E} obeys a decreasing property on the spectra of Schrödinger operators, so it should be interpreted as an inverse measure of the size of E. For instance, our next result is a universal inequality involving a_{E}, which should be seen as a *lower* bound on the size of the essential spectrum:

Theorem 1.2. *If V is a potential obeying* (1.1) *and* $\mathsf{E} = \sigma_{\text{ess}}(L_V)$*, then*

$$a_{\mathsf{E}} \le \liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t. \tag{1.5}$$

The perspective on a_{E} as an inverse measure of the size of E will be most explicitly illustrated later, in the proof of Theorem 1.12, which will use the argument that if $\mathsf{E} \subset [0,\infty)$ and $a_{\mathsf{E}} \le a_{[0,\infty)}$, then $\mathsf{E} = [0,\infty)$. This kind of argument wasn't available before in this generality, because there was no known quantity with the correct properties: any quantity based on Lebesgue measure or dimension would sometimes give infinite or trivial values.

For any $z \in \mathbb{C}$, the Dirichlet eigensolution is the solution of the initial value problem

$$-\partial_x^2 u(x,z) + V(x)u(x,z) = zu(x,z), \qquad u(0,z) = 0, \qquad (\partial_x u)(0,z) = 1.$$

Our next result is that the Martin function provides a universal lower bound on the growth rate of the Dirichlet solution.

Theorem 1.3. *If V is a potential obeying* (1.1) *and* $\mathsf{E} = \sigma_{\text{ess}}(L_V)$*, then*

$$M_{\mathsf{E}}(z) \le \liminf_{x\to\infty} \frac{1}{x}\log|u(x,z)|, \qquad \forall z \in \mathbb{C}\setminus[\min \mathsf{E},\infty).$$

Exclusion of $[\min \mathsf{E},\infty)$ in Theorem 1.3 is necessary because for $z \in (\min \mathsf{E},\infty)$, by Sturm oscillation theory [64], the Dirichlet solution has infinitely many zeros.

Definition 1.4. The potential V is regular if

$$a_{\mathsf{E}} = \lim_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t. \tag{1.6}$$

Of course, due to (1.5), this is equivalent to requiring

$$a_{\mathsf{E}} \ge \limsup_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t.$$

In our next theorem, we will characterize regularity in terms of root asymptotics for the Dirichlet eigensolutions. We say that a property holds a.e. on E with respect to harmonic measure if it holds away from a set $A \subset \mathsf{E}$ such that $\omega_{\mathsf{E}}(A, z_0) = 0$, where $\omega_{\mathsf{E}}(\cdot, z_0)$ denotes the harmonic measure of Ω evaluated at some $z_0 \in \Omega$. This condition is independent of the choice of $z_0 \in \Omega$ since the harmonic measures are mutually absolutely continuous.

Theorem 1.5. *If V is a potential obeying* (1.1) *and* $\mathsf{E} = \sigma_{\text{ess}}(L_V)$*, the following are equivalent:*

(i) V is regular;
(ii) For every Dirichlet-regular $z \in \mathsf{E}$, $\limsup_{x\to\infty} \frac{1}{x}\log|u(x,z)| \le 0$;
(iii) For a.e. $z \in \mathsf{E}$ with respect to harmonic measure,

$$\limsup_{x\to\infty} \frac{1}{x}\log|u(x,z)| \le 0;$$

(iv) There exists $z \in \mathbb{C}_+$ such that $\limsup_{x\to\infty} \frac{1}{x}\log|u(x,z)| \le M_{\mathsf{E}}(z)$;
(v) For all $z \in \mathbb{C}$, $\limsup_{x\to\infty} \frac{1}{x}\log|u(x,z)| \le M_{\mathsf{E}}(z)$;
(vi) $\lim_{x\to\infty} \frac{1}{x}\log|u(x,z)| = M_{\mathsf{E}}(z)$ uniformly on compact subsets of $\mathbb{C}\setminus[\min \mathsf{E},\infty)$.

Since (v) or (vi) trivially imply (iv), (iv) is of interest as a criterion for establishing regularity of V, whereas (v), (vi) are of interest as consequences of regularity. Similarly, (ii) implies (iii), so (ii) is of interest as a consequence of regularity and (iii) as a condition for regularity. Instead of conditions (ii) and (iii), it would be customary to state the single condition that the inequality holds quasi-everywhere; this is between our conditions since the set of Dirichlet-irregular points is polar and polar sets have harmonic measure 0. The benefit of (ii) is that it can be used pointwise (in particular, for a Dirichlet-regular set E, the inequality holds everywhere on E). More importantly, the benefit of (iii) is that the characterization in terms of harmonic measure will be essential for our proof of Theorem 1.8 below.

There are no previous results on Stahl–Totik regularity for continuum Schrödinger operators, even in special cases. This topic was previously considered by Simon [65, Section 9], who formulated several conjectures. The first is that for semibounded spectra that are "close" to $[0,\infty)$ (e.g. $[0,\infty)\setminus\mathsf{E}$ of finite Lebesgue measure) there should be a version of equilibrium measure ν_{E} and equilibrium potential Φ_{E}, characterized by several properties including a normalization $\Phi_{\mathsf{E}}(z)\sim\operatorname{Re}(\sqrt{-z})(1+o(1))$ as $z\to-\infty$. It was suggested that regularity for continuum Schrödinger operators can be defined by the condition $\limsup_{x\to\infty}\frac{1}{x}\log|u(x,z)|=\Phi_{\mathsf{E}}(z)$, and that this would have equivalent characterizations similar to the orthogonal polynomial case. Our work does not use a finite Lebesgue measure assumption for $[0,\infty)\setminus\mathsf{E}$, so it solves these conjectures in a far greater generality than they were even previously conjectured. Moreover, our work provides the correct potential theoretic interpretation for the function Φ_{E} (now understood as the Martin function M_{E}), and that interpretation is crucial in the proofs.

Simon also conjectured that the asymptotics $\Phi_{\mathsf{E}}(z)=\operatorname{Re}(\sqrt{-z})(1+o(1))$ should improve to the asymptotic behavior $\operatorname{Re}\sqrt{-z}+o(1)$; this is motivated by the asymptotic behavior $\sqrt{-z}+o(1)$ of m-functions, proved by Atkinson [4]. While that asymptotic statement for individual m-functions cannot be improved for locally integrable potentials, we discover that due to averaging effects, the asymptotic behavior of our quantities improves even more, to the form (1.3). This discovery of (1.3) has enabled us to introduce the constant a_{E}, which was not previously conjectured, and to use it for the robust general definition of regularity given above.

We also define the correct "equilibrium measure" which will be related to a deterministic density of states. The Martin function can be extended to a subharmonic function on $\mathbb{C}$, so it has a Riesz measure, given by

$$\rho_{\mathsf{E}}=\frac{1}{2\pi}\Delta M_{\mathsf{E}},$$

which we will call the Martin measure of the set E. Conversely, the Martin function has a Hadamard representation of the form

$$M_{\mathsf{E}}(z)=M_{\mathsf{E}}(z_*)+\int_{\mathsf{E}}\log\left|1-\frac{z-z_*}{t-z_*}\right|\mathrm{d}\rho_{\mathsf{E}}(t)$$

where $z_*<\min\mathsf{E}$ is an arbitrary normalization point. The Martin measure will serve the same role in this theory that the logarithmic equilibrium measure serves for orthogonal polynomials. However, ρ_{E} is not defined with respect to any extremal property (and it is not even a finite measure), so different proofs will be needed in the current setting.

For any $x>0$, let ρ_x denote the zero counting measure for $u(x,z)$ divided by x,

$$\rho_x=\frac{1}{x}\sum_{z:u(x,z)=0}\delta_z. \tag{1.7}$$

Note that ρ_x is the Riesz measure of $\frac{1}{x}\log|u(x,z)|$. The limit of ρ_x as $x\to\infty$, when it exists, is interpreted as a deterministic density of states associated to V. The convergence of measures will be understood in the weak-$*$ sense, i.e., when integrated against continuous functions with compact support. The Martin measure and the zero counting measures are related by the following pair of results:

Theorem 1.6. *Assume V is regular. Then ρ_x converges to ρ_{E} as $x \to \infty$, in the weak-$*$ sense.*

The following is a continuum analog of a result of Stahl–Totik [69]:

Theorem 1.7. *Assume that V obeys* (1.1) *and let μ be a maximal spectral measure for L_V. Suppose that ρ_x converges to ρ_{E} as $x \to \infty$ in the weak-$*$ sense. Then, either V is regular, or there exists a polar Borel set X such that $\mu(\mathbb{R} \setminus X) = 0$.*

Of course, the statement $\mu(\mathbb{R} \setminus X) = 0$ can be restated in the language of the Borel functional calculus as $\chi_{\mathbb{R}\setminus X}(L_V) = 0$.

So far, we have seen that regularity of V can be established from the root asymptotics of Dirichlet solutions. The next theorem shows that it can be established from spectral properties of the operator. It is the continuum counterpart of a theorem of Widom [76].

Theorem 1.8. *Let μ be a maximal spectral measure for L_V. If $\omega_{\mathsf{E}}(\cdot, z_0)$ for some $z_0 \in \mathbb{C} \setminus \mathsf{E}$ is absolutely continuous with respect to μ, then V is regular.*

This theory leads to several new results even for the special case of half-line essential spectrum $[0, \infty)$; we present those as our first applications. If V is a decaying potential in the sense

$$\lim_{x\to\infty} \int_x^{x+1} |V(t)| \mathrm{d}t = 0 \tag{1.8}$$

then $\mathsf{E} = \sigma_{\mathrm{ess}}(L_V) = [0, \infty)$ by Blumenthal–Weyl [9, 75]. It follows that $M_{\mathsf{E}}(z) = \operatorname{Re} \sqrt{-z}$. In particular, $a_{\mathsf{E}} = 0$, so immediately from the definition:

Corollary 1.9. *If V is a decaying potential in the sense* (1.8)*, then V is regular with $\sigma_{\mathrm{ess}}(L_V) = [0, \infty)$.*

Since harmonic measure for $\mathsf{E} = [0, \infty)$ is mutually absolutely continuous with $\chi_{(0,\infty)}(x)\mathrm{d}x$, the following is an immediate consequence of Theorem 1.8:

Corollary 1.10. *Assume that V obeys* (1.1) *and denote by μ a maximal spectral measure for L_V. Denote by $\mathrm{d}\mu = f\mathrm{d}x + \mathrm{d}\mu_s$ the Radon–Nikodym decomposition of μ with respect to Lebesgue measure. If $\sigma_{\mathrm{ess}}(L_V) = [0, \infty)$ and $f(x) > 0$ for Lebesgue-a.e. $x > 0$, then V is regular.*

More generally, a version of Corollary 1.10 holds, whenever the harmonic measure for the domain $\mathbb{C} \setminus \mathsf{E}$ is absolutely continuous with respect to the Lebesgue measure $\chi_{\mathsf{E}}(x)\mathrm{d}x$. In particular, it holds for finite gap sets (i.e., when N is finite in (1.4)) and regular Parreau-Widom sets. If E is Dirichlet-regular, the Green function $G_{\mathsf{E}}(z, z_0)$, for $z_0 < \min \mathsf{E}$, has exactly one critical point $c_j \in (a_j, b_j)$ in each gap. If, in addition, the critical values of $G_{\mathsf{E}}(z, z_0)$ are summable, i.e.,

$$\sum_{j=1}^{\infty} G_{\mathsf{E}}(c_j, z_0) < \infty,$$

we call E a regular Parreau-Widom set. In fact, the harmonic measure for the domain $\mathbb{C} \setminus \mathsf{E}$ is absolutely continuous with respect to the Lebesgue measure if and only if E satisfies a certain sector condition [32, Theorem 4]. We will describe this generalization in Section 6.

Sparse potentials are not covered by Corollary 1.9 or Corollary 1.10, but nonetheless provide additional examples of regular potentials:

Example 1.11. Let $W \in L^1((0,\infty))$ be compactly supported, $W \geq 0$, let $x_n \geq 0$ be an increasing sequence such that $x_{n+1} - x_n \to \infty$ as $n \to \infty$ and $V(x) = \sum_n W(x - x_n)$. Then V is regular with $\sigma_{\text{ess}}(L_V) = [0,\infty)$.

The sparse potentials from Example 1.11 are not decaying in the sense (1.8), so Corollary 1.9 does not have a converse; sparse potentials have purely singular spectrum by [46, 60], so Corollary 1.10 does not have a converse.

However, we prove that Corollary 1.9 has the following partial converse; we have already described Theorem 1.1 as a universal thickness result about the spectrum, and the following result similarly guarantees presence of essential spectrum.

Theorem 1.12. *Assume that V obeys* (1.1) *and that $\sigma_{\text{ess}}(L_V) \subset [0,\infty)$. Then:*
(a) $\liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \geq 0$*;*
(b) If $\liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \leq 0$*, then* $\sigma_{\text{ess}}(L_V) = [0,\infty)$*;*
(c) If $\limsup_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \leq 0$*, then* $\sigma_{\text{ess}}(L_V) = [0,\infty)$ *and V is regular.*

Part (a) can also be established by other means, but we include it for completeness. Parts (b) and (c) generalize known results giving sufficient conditions for $\sigma_{\text{ess}}(L_V) = [0,\infty)$. In particular, Damanik–Remling [25, Theorem 1.2] showed that $\sigma_{\text{ess}}(L_{\pm V}) \subset [0,\infty)$ implies $\sigma_{\text{ess}}(L_V) = [0,\infty)$. Part (b) of our theorem is a strict generalization of that result; strict because it applies, e.g., to the sparse potentials of Example 1.11 where [25] does not (for a positive sparse potential V, $\min \sigma_{\text{ess}}(L_{-V}) < 0$), and a generalization because $\sigma_{\text{ess}}(L_{-V}) \subset [0,\infty)$ implies $\limsup_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \leq 0$ (by (a) applied to $-V$), so our parts (b), (c) also apply to the potentials in [25]. In particular, $\sigma_{\text{ess}}(L_{\pm V}) \subset [0,\infty)$ implies that V is regular and $\sigma_{\text{ess}}(L_V) = [0,\infty)$.

In the theory of Jacobi matrices, a result of Simon [67] shows that a regular Jacobi matrix with essential spectrum $[-2,2]$ obeys a Cesàro–Nevai condition. The analog for Schrödinger operators is false – the continuum setting allows rapid oscillations which can break any Cesàro-type decay in an L^1 sense:

Example 1.13. The potential defined piecewise by $V(x) = (-1)^{\lfloor 2n(x-n)\rfloor}$ on $x \in [n-1,n)$ for integer n is regular with $\sigma_{\text{ess}}(L_V) = [0,\infty)$, but $\frac{1}{x}\int_0^x |V(t)|\mathrm{d}t \not\to 0$ as $x \to \infty$.

All objects considered above are deterministic (defined only in terms of a single half-line potential V), but for ergodic families of Schrödinger operators, they can be recognized almost surely as ergodic notions such as the Lyapunov exponent and the ergodic density of states, so our results can be interpreted in the ergodic setting. In the ergodic setting, it is natural to work with whole line potentials: let us consider a family $(V_\eta)_{\eta\in S}$ of real-valued potentials on $\mathbb{R}$ on a probability space S which is metrically transitive with respect to a group of measure preserving transformations τ_y such that $V_{\tau_y \eta}(x) = V_\eta(x-y)$ and such that any measurable subset A of S which is invariant under all τ_y has probability 0 or 1. The group of transformations can be continuous (indexed by $y \in \mathbb{R}$) or discrete (indexed by $y \in \ell\mathbb{Z}$ for some $\ell > 0$); the former case includes almost periodic Schrödinger operators and the latter case includes many Anderson-type models studied in the literature [26, 42], including those with a periodic background. We also assume that V_η almost surely obeys

$$\sup_{x\in\mathbb{R}} \int_x^{x+1} |V_\eta(t)|\mathrm{d}t < \infty; \tag{1.9}$$

in fact, much of the literature on ergodic Schrödinger operators is focused on bounded potentials. Let us denote by H_{V_η} the self-adjoint operators on $L^2(\mathbb{R})$ given by

$$D(H_{V_\eta}) = \{f \in L^2(\mathbb{R}) \mid f \in W^{2,1}_{\text{loc}}(\mathbb{R}), -f'' + V_\eta f \in L^2(\mathbb{R})\}$$

and recall the basic properties of this ergodic family (see textbooks [11, 15, 57] and a paper of Kirsch [42] addressing some nuances for locally L^1 ergodic potentials with a discrete group of transformations). There is an almost sure spectrum $\mathsf{E} \subset \mathbb{R}$,

$$\mathsf{E} = \sigma(H_{V_\eta}) = \sigma_{\text{ess}}(H_{V_\eta}), \qquad \text{for a.e. } \eta \in S,$$

and the potentials V_η have an almost sure Birkhoff average $\mathbb{E}(V) \in \mathbb{R}$,

$$\mathbb{E}(V) = \lim_{x\to\infty} \frac{1}{x} \int_0^x V_\eta(t)\mathrm{d}t, \qquad \text{for a.e. } \eta \in S.$$

If L_{V_η} denotes the half-line operator corresponding to the restriction of V_η to $[0,\infty)$, then $\mathsf{E} = \sigma_{\text{ess}}(L_{V_\eta})$ almost surely, so as a direct consequence of our deterministic results, E corresponds to a Martin function with an expansion (1.3), and

$$a_{\mathsf{E}} \le \mathbb{E}(V). \tag{1.10}$$

This inequality is new; several cases of the equality $a_{\mathsf{E}} = \mathbb{E}(V)$ are well known and among the most studied classes of ergodic Schrödinger operators (periodic, reflectionless almost periodic with finite gap length), and we can now interpret this through the fact that the corresponding potentials are regular.

In the ergodic setting, two central objects is the Lyapunov exponent $\gamma(z)$ and the density of states $\mathrm{d}\rho$; both are almost sure ergodic averages of important spectral quantities. The transfer matrix $T_\eta(x,z)$ is the 2×2-matrix valued solution of the initial value problem

$$(\partial_x T_\eta)(x,z) = \begin{pmatrix} 0 & V_\eta(x) - z \\ 1 & 0 \end{pmatrix} T_\eta(x,z), \qquad T_\eta(0,z) = I,$$

and the corresponding Dirichlet solution is $u_\eta(x,z) = (T_\eta)_{2,1}(x,z)$. If $\rho_{\eta,x}$ denotes the measure corresponding to u_η as in (1.7), then

$$\gamma(z) = \lim_{x\to+\infty} \frac{1}{x} \log\|T_\eta(x,z)\|, \qquad \text{for a.e. } \eta \in S, \tag{1.11}$$

and

$$\mathrm{d}\rho = \operatorname*{w-lim}_{x\to+\infty} \mathrm{d}\rho_{\eta,x}, \qquad \text{for a.e. } \eta \in S.$$

Thus Theorem 1.5, specialized to the ergodic setting, immediately gives the following:

Corollary 1.14. *For any ergodic family of Schrödinger operators obeying* (1.9)*, the following are equivalent:*

(i) $a_{\mathsf{E}} = \mathbb{E}(V)$*;*
(ii) For every Dirichlet-regular $z \in \mathsf{E}$*,* $\gamma(z) = 0$*;*
(iii) For almost every $z \in \mathsf{E}$ *with respect to harmonic measure,* $\gamma(z) = 0$*;*
(iv) For all $z \in \mathbb{C}_+$*,* $\gamma(z) \le M_{\mathsf{E}}(z)$*;*
(v) For all $z \in \mathbb{C} \setminus \mathsf{E}$*,* $\gamma(z) \le M_{\mathsf{E}}(z)$*;*
(vi) $\gamma(z) = M_{\mathsf{E}}(z)$ *for all* $z \in \mathbb{C} \setminus [\min \mathsf{E}, \infty)$*.*

We say that a family of ergodic Schrödinger operators is regular if one (and therefore all) of the statements of Corollary 1.14 holds. Although this notion is new, let us point out that it contains several of the most well studied families of almost periodic Schrödinger operators known to have zero Lyapunov exponent on the spectrum, such as quasiperiodic operators at small coupling [19, 21, 22, 23, 31] and limit-periodic potentials superexponentially well approximated by periodic operators [13, 34, 58, 59]. In fact, the question of when the Lyapunov exponent is zero or positive on E is one of the basic questions for an almost periodic family of operators and an important dichotomy in their study; this is especially well studied in the setting of discrete Schrödinger operators, see e.g. [5, 16, 40]. In inverse spectral theory one considers reflectionless Schrödinger operators on Dirichlet-regular Widom spectra with the DCT property and associated solutions of the KdV equation [8, 20, 28, 29, 30, 37, 68]; those operators have zero Lyapunov exponent on the spectrum so they are regular in the sense of this paper.

For a 1-periodic potential V, it is well known that the discriminant has an asymptotic expansion at ∞ whose coefficients are equal to averages of differential polynomials in V (under the appropriate regularity assumptions on V). The first of those equalities, rewritten for the Martin function, give the equality $a_{\mathsf{E}} = \int_0^1 V(x)\mathrm{d}x$. This can now be interpreted through the fact that periodic potentials are regular.

For an almost periodic potential V, Johnson–Moser [41] introduced the spatial average of m-functions, whose real part is the Lyapunov exponent γ. Their construction relies heavily on almost periodicity through compactness of the hull, so their methods would not extend to our setting; [41] noted as a consequence of their results, the spectrum of any almost periodic Schrödinger operator is not a polar set (i.e. Ω is Greenian), but further consequences of Theorem 1.1 were not previously known even in the almost periodic case.

The next theorem is a specialization of Theorems 1.6, 1.7 to the ergodic setting:

Theorem 1.15. *Let $(V_\eta)_{\eta\in S}$ be an ergodic family of Schrödinger operators obeying* (1.9)*. If this ergodic family is regular, then its density of states ρ is equal to the Martin measure ρ_{E}. Conversely, if $\rho = \rho_{\mathsf{E}}$, then either the ergodic family is regular, or for a.e. η, the maximal spectral measure μ_η is supported on a polar set.*

Although positive Lyapunov exponents don't always correspond to localization, we can now prove that they always correspond to very thin spectral type. This is the analog of a Jacobi matrix result which has been described as the ultimate Pastur–Ishii theorem.

Theorem 1.16. *Let γ denote the Lyapunov exponent associated to the ergodic family $(V_\eta)_{\eta\in S}$ and let μ_η denote a maximal spectral measure for H_{V_η}. Let $Q \subset \mathbb{R}$ be the Borel set of $\lambda \in \mathbb{R}$ with $\gamma(\lambda) > 0$. Then for a.e. $\eta \in S$, there exists a polar set X_η such that $\mu_\eta(Q \setminus X_\eta) = 0$. In particular, the measure $\chi_Q \mathrm{d}\mu_\eta$ is of local Hausdorff dimension zero.*

It is known in great generality [26] that one-dimensional random Schrödinger operators give rise to positive Lyapunov exponent throughout the spectrum. In particular, random Schrödinger operators provide examples of non-regular operators.

Throughout this paper, we follow the dominant literature by working with locally integrable potentials; we expect that the theory presented here can be extended to potentials which are in the negative Sobolev space $H^{-1}([0,x])$ for $x < \infty$, with an appropriate uniform bound replacing (1.1), and that it can be adapted to certain other classes of one-dimensional differential operators.

We expect that the notion of regularity introduced in this paper will pave the way to new kinds of results on Schrödinger operators which were previously beyond reach. For instance, regularity of measures is used as the standard reference behavior in the study of the local distribution of zeros of orthogonal polynomials, through so-called clock behavior and universality [51, 56, 66]; we conjecture that similar results hold for regular Schrödinger operators. Without regularity, the only currently available Schrödinger result is inevitably more limited in scope to certain perturbations of periodic Schrödinger operators [53]. Likewise, logarithmic capacity is used to formulate the generalization of the Shohat–Nevai theorem to measures whose essential supports are regular Parreau–Widom sets

[12]; the Schrödinger counterpart of this result couldn't even be formulated without the renormalized Robin constant a_{E}. We expect the theory in this paper to be an integral part of its eventual proof, and of the broader program of investigating sum rules for Schrödinger operators with regular Parreau–Widom essential spectra.

2. The Martin function and Akhiezer–Levin sets

In this section we consider in more detail the general Martin theory for Denjoy domains $\Omega = \mathbb{C}\setminus\mathsf{E}$ with $\min \mathsf{E} = b_0 > -\infty$. Clearly, we have in mind the application that E is the essential spectrum of some continuum Schrödinger operator, L_V, where V satisfies (1.1).

Recall that the capacity of a Borel set A is defined by

$$\mathrm{Cap}(A) = \sup\{\mathrm{Cap}(K) : \ K \text{ compact}, K \subset A\}$$

and we call a Borel set, A, polar, if $\mathrm{Cap}(A) = 0$. Moreover, a property holds quasi-everywhere on a set B, if there exists a polar set A such that the property holds on $B \setminus A$. We start with a discussion of the Green function $G_{\mathsf{E}}(z, z_0)$, $z_0 \in \Omega$. For standard references on potential theory see [3, 36, 62]. If $z_0 \in \mathbb{R}$, then $G_{\mathsf{E}}(z, z_0)$ is symmetric, that is, $G_{\mathsf{E}}(\overline{z}, z_0) = G_{\mathsf{E}}(z, z_0)$. Let us fix $z_0 < b_0$. Then there exists a comb domain

$$\Pi_{z_0} = \{x + iy : \ 0 < x < \pi, \ y > s(x)\}, \tag{2.1}$$

where s is a positive upper semicontinuous function, bounded from above, and vanishes Lebesgue-a.e., and a conformal mapping $\theta_{z_0} : \mathbb{C}_+ \to \Pi_{z_0}$ such that

$$G_{\mathsf{E}}(z, z_0) = \mathrm{Im}\, \theta_{z_0}(z), \quad z \in \mathbb{C}_+. \tag{2.2}$$

(such a representation was proved in [32] in the case that E is compact and $z_0 = \infty$; by a simple transformation $\lambda = \frac{1}{z_0 - z}$ this yields a corresponding representation for the current setting). Note that $\theta_{z_0}(b_0) = i \limsup_{u\to 0} s(u)$ and $\theta_{z_0}(\infty) = i \limsup_{u\to\pi} s(u)$. Moreover, harmonic measure $\omega_{\mathsf{E}}(\cdot, z_0)$ corresponds to the pullback of the normalized (by π) Lebesgue measure on the base of the comb. The mapping can be extended by symmetry to $\mathbb{C} \setminus [b_0, \infty)$ such that (2.2) still holds there. In fact, any such function s leads to a Green function of a certain domain.

The Martin kernel normalized at $z_* < b_0$ is defined on $\Omega \times (\Omega \setminus \{z_*\})$ by

$$M_{\mathsf{E}}(z, z_0) = \frac{G_{\mathsf{E}}(z, z_0)}{G_{\mathsf{E}}(z_*, z_0)}. \tag{2.3}$$

The *Martin compactification* $\widehat{\Omega}$ is the smallest metric compactification of Ω such that $M_{\mathsf{E}}(z, \cdot)$ can be continuously extended to the boundary $\partial^M \Omega = \widehat{\Omega} \setminus \Omega$ for each z. We will also write $M_{\mathsf{E}}(z, z_0)$ for the extended function. Note that by the Harnack principle the family $\{M_{\mathsf{E}}(z, z_0)\}$ is precompact in the space of positive harmonic functions equipped with uniform convergence on compacts. We call a positive harmonic function, M, *minimal* if any harmonic function, h, which satisfies $0 \le h \le M$, is a multiple of M, i.e., $h = cM$, $c \ge 0$. Finally, let $\partial_1^M \Omega \subset \partial^M \Omega$ denote the subset of the Martin boundary, which consists of minimal harmonic functions. In this case, for every positive harmonic function h, there exists a unique finite measure ν such that

$$h(z) = \int_{\partial_1^M \Omega} M_{\mathsf{E}}(z, x) \mathrm{d}\nu(x), \quad h(z_*) = \nu(\partial_1^M \Omega). \tag{2.4}$$

In general $\partial_1^M \Omega$ can be quite abstract, but the situation is rather intuitive for Denjoy domains. In [35, Theorem 6] it is shown that there exists a map $\pi : \partial_1^M \Omega \to \mathsf{E} \cup \{\infty\}$ such that for every $x \in \mathsf{E} \cup \{\infty\}$, $\#\pi^{-1}(\{x\})$ is either one or two, depending on how "thin" $\mathbb{R} \cap \Omega$ is at x. To state this precisely we need some definitions. If A is a subset of the Martin boundary $\partial^M\Omega = \hat{\Omega} \setminus \Omega$, then we say a property, P, holds near A if there is a Martin-neighborhood $A \subset W$ such that P holds on $W \cap \Omega$. Then, for $A \subset \hat{\Omega}$ and a positive superharmonic function h on Ω we define the reduced function

$$R_h^A(x) = \inf\{u(x) : \ u \geq 0 \text{ is superharmonic, } h \leq u \text{ on } A \cap \Omega \text{ and } h \leq u \text{ near } A \cap \partial^M\Omega\} \tag{2.5}$$

and $\hat{R}_h^A$ denotes its lower semicontinuous regularization. A set $A \subset \Omega$ is said to be minimally thin at $y \in \partial_1^M\Omega$ if

$$\hat{R}^A_{M_{\mathsf{E}}(\cdot,y)} \neq M_{\mathsf{E}}(\cdot, y).$$

Then $\#\pi^{-1}(\{x\}) = 2$ if and only if there is $y \in \pi^{-1}(\{x\})$ such that $\Omega \cap \mathbb{R}$ is minimally thin at y. Informally, if E is sufficiently "dense" at x, then Ω locally splits into the two half spaces $\mathbb{C}_+$ and $\mathbb{C}_-$ and we obtain a Martin function for each of them.

A reformulation of the above statement can be given in the following way. For $x \in \mathsf{E}$, let $\mathcal{P}_{\mathsf{E}}(x)$ denote the set of positive harmonic functions that are bounded outside every neighborhood of x and vanish quasi-everywhere on E. As in the proof of [38, Lemma 2.9] one can see, that $\mathcal{P}_{\mathsf{E}}(x)$ is spanned by the Martin functions related to x. Hence, the above question is whether $\mathcal{P}_{\mathsf{E}}(x)$ is one- or two-dimensional. We will provide a simplified proof for the case that there is only one Martin function associated to x below. This question has attracted much interest and several conditions have been obtained, [2, 7, 43, 50]. To note two extreme cases, if $x \in (a, b) \subset \mathsf{E}$, then $\mathcal{P}_{\mathsf{E}}(x)$ is two-dimensional, whereas if x is a endpoint of a gap of E, then $\mathcal{P}_{\mathsf{E}}(x)$ is one-dimensional, as discussed in [35] after Theorem 6.

We are particularly interested in the Martin kernel related to ∞. Since E is semibounded, $\mathcal{P}_{\mathsf{E}} = \mathcal{P}_{\mathsf{E}}(\infty)$ is one-dimensional and we can talk about the Martin function $M_\infty(z) = M_{\mathsf{E}}(z, \infty)$ related to ∞ which is known to be symmetric, i.e., $M_\infty(\overline{z}) = M_\infty(z)$. Moreover, all limits with $z_n \to -\infty$ must lead to M_∞ and we have

$$M_\infty(z) = \lim_{z_0 \to -\infty} M(z, z_0) = \lim_{z_0 \to -\infty} \frac{\operatorname{Im} \theta_{z_0}(z)}{G_{\mathsf{E}}(z_*, z_0)}$$

Note that M_∞ is not exactly M_{E} from the introduction, because in the general situation we cannot use the normalization (1.2). For this reason, we keep the normalization at z_*, but once we have specified to sets where the limit in (1.2) is positive, we can pass to this normalization. Since $M_\infty(z)$ is positive and harmonic in Ω, setting $\lambda^2 = z - b_0$ it defines a positive harmonic function for $\lambda \in \mathbb{C}_+$ by

$$f(\lambda) = M_\infty(z).$$

Since f can be represented as

$$f(x + iy) = ay + \int \frac{y}{(x-t)^2 + y^2} \mathrm{d}\nu(t), \quad \int \frac{\mathrm{d}\nu(t)}{1+t^2} < \infty \tag{2.6}$$

and

$$0 \leq a = \lim_{y \to \infty} \frac{f(iy)}{y}, \tag{2.7}$$

we see that $M_\infty(z)$ can grow at most as $\sqrt{-z}$ as $z \to -\infty$. In case of two-sided unbounded sets, where the Martin function can grow at most linearly, Akhiezer and Levin showed that $\mathcal{P}_{\mathsf{E}}$ is two-dimensional whenever the Martin function admits the maximal possible growth. This explains why we call E an *Akhiezer–Levin set* if

$$\lim_{z\to-\infty}\frac{M_\infty(z)}{\sqrt{-z}}>0. \tag{2.8}$$

Note that by (2.7) this limit indeed exists in $[0,\infty)$. Since in (2.6), the integral $\int\frac{y}{(x-t)^2+y^2}\mathrm{d}\nu(t)$ defines again a positive harmonic function it follows that

$$a\operatorname{Re}\sqrt{b_0-z}\le M_\infty(z) \tag{2.9}$$

in Ω. The following theorem presents a list of equivalent characterizations of M_∞. We say that h vanishes continuously at a point $x\in\mathsf{E}$ if $\lim_{\substack{z\to x\\ z\in\Omega}}h(z)=0$. We call a subset of Ω bounded if it is bounded as a subset of $\mathbb{C}$.

Theorem 2.1. *Let $H_{+,b}(\Omega)$ denote the set of positive harmonic functions on Ω that are bounded on every bounded subset of Ω. Then, the following are equivalent:*

(i) $h\in H_{+,b}(\Omega)$ and h vanishes continuously for every Dirichlet-regular point of E;
(ii) $h\in H_{+,b}(\Omega)$ and h vanishes continuously quasi-everywhere on E;
(iii) $h\in H_{+,b}(\Omega)$ and h vanishes continuously $\omega_{\mathsf{E}}(\cdot,z_0)$-a.e.;
(iv) $h=cM_\infty$, where $c\ge 0$;

Proof. Due to [35, Remark 5, Theorem 6] $(iv)\implies(i)$. Kellogg's theorem [36, Corollary 6.4] yields $(i)\implies(ii)$ and by [36, Theorem III.8.2] we get that $(ii)\implies(iii)$. It remains to show that $(iii)\implies(iv)$. Due to (2.4) there exists ν such that

$$h(z)=\int_{\partial_1^M\Omega}M_{\mathsf{E}}(z,x)\mathrm{d}\nu(x).$$

Let $K\subset\partial^M\Omega\setminus\{M_\infty\}$ be compact. Then K has an open neighborhood U in $\hat{\Omega}$ such that $U\cap\Omega$ is bounded. As in the proof of [3, Theorem 8.4.1]

$$R_h^K(z)=\int_K M_{\mathsf{E}}(x,z)d\nu(x).$$

Since $h\in H_{+,b}(\Omega)$, h is majorized by a constant in $U\cap\Omega$, so R_h^K is a bounded harmonic function in Ω which vanishes $\omega_{\mathsf{E}}(\cdot,z_0)$-a.e. on the boundary. By the maximum principle [36, Theorem III.8.1] it follows that $R_h^K=0$. In particular, $R_h^K(z_*)=\nu(K)=0$. The claim follows. ☐

In his series of papers [48, 49, 50], Levin first systematically established the relation between extremal problems and comb mappings imposing Dirichlet-regularity on the set E. Eremenko and Yuditskii [32] provided a modern approach to it, giving a detailed proof for comb mappings for Green functions as discussed above. It relies on the representation of Green functions for a compact set E, as

$$G_E(z,\infty)=\int_E\log|z-t|\mathrm{d}\rho_E(t)+\gamma_E, \tag{2.10}$$

where $\mathrm{Cap}(E)=e^{-\gamma_E}$ and $\rho_E(X)=0$ for sets of zero capacity. It is also discussed that the proof carry over for Martin functions and the corresponding description is given. Since we were not able to find in our generality a reference for a representation of the type (2.10), which is certainly known to experts, for the readers convenience we survey the corresponding theory in the following.

Since M_∞ vanishes quasi-everywhere, we can extend M_∞ to a subharmonic function to all of $\mathbb{C}$ by

$$M_\infty(x) = \limsup_{\substack{z\to x\\ z\in\Omega}} M_\infty(z), \quad x \in \mathsf{E}, \tag{2.11}$$

see [3, Theorem 5.2.1]. Hence, we obtain a subharmonic, symmetric function in $\mathbb{C}$, which is positive and harmonic in $\mathbb{C}_+$ and $\mathbb{C}_-$. For the following result we refer to [49, Lemma 2.3] and its corollary. It was initially proved for majorants of subharmonic functions, but it is mentioned that it extends to the version stated below:

Lemma 2.2. *Let v be a subharmonic, symmetric function in $\mathbb{C}$, which is positive and harmonic in $\mathbb{C}\setminus[b_0,\infty)$ for some $b_0\in\mathbb{R}$. Then*

$$v(z) = v(z_*) + \int_{b_0}^\infty \log\left|1-\frac{z-z_*}{t-z_*}\right| \mathrm{d}\nu(t), \quad \int_{b_0}^\infty \frac{\mathrm{d}\nu(t)}{t-z_*} < \infty, \tag{2.12}$$

and for $y>0$

$$\frac{\partial v(x+iy)}{\partial y} = \int_{b_0}^\infty \frac{y}{(t-x)^2+y^2}\mathrm{d}\nu(t) > 0. \tag{2.13}$$

Remark. (2.12) is essentially the Hadamard representation for the subharmonic function v and ν is its *Riesz measure*. Usually the Hadamard representation would include a normalization term $\frac{\operatorname{Re} z}{t}$, which is not needed due to the convergence property of ν in (2.12).

Lemma 2.3. *Let Θ be such that $\operatorname{Im}\Theta = M_\infty$ for $z\in\mathbb{C}_+$ and ρ be the Riesz measure for M_∞. Then, the functions Θ and $i\Theta'$ are Herglotz functions and in particular*

$$i\Theta'(z) = \int_{\mathsf{E}} \frac{\mathrm{d}\rho(t)}{t-z}.$$

They can be analytically extended to $\mathbb{C}\setminus[b_0,\infty)$ and $\Theta'\neq 0$ there.

Proof. Applying Lemma 2.2 to M_∞ gives a representation of the form (2.12) in terms of the Riesz measure ρ supported on E and, in particular, $\int_{\mathsf{E}} \frac{\mathrm{d}\rho(t)}{t-z_*} < \infty$. Moreover,

$$i\Theta'(z) = c_0 + \int_{\mathsf{E}} \frac{\mathrm{d}\rho(t)}{t-z} \tag{2.14}$$

for some $c_0\in\mathbb{R}$, since the imaginary parts of the two sides are equal by (2.13). Since Θ is also a Herglotz function, for some measure μ supported on E,

$$i\Theta'(z) = i\int \frac{\mathrm{d}\mu(t)}{(t-z)^2}, \quad \int \frac{\mathrm{d}\mu(t)}{1+t^2} < \infty. \tag{2.15}$$

Using monotone convergence and taking the limit as $z\to-\infty$ in (2.14) and (2.15) yields $\lim_{z\to-\infty} i\Theta'(z) = 0 = c_0$. Since $i\Theta'$ is Herglotz, $\Theta'\neq 0$ in $\mathbb{C}_+$ and $\mathbb{C}_-$. Moreover, since it is increasing on $(-\infty,b_0)$ and vanishes at $-\infty$ we obtain the final claim. □

The following lemma shows that, like the harmonic measure, ρ gives zero measure to polar sets. Of course, once we introduce the Martin measure ρ_{E}, it will be a scalar multiple of ρ, so the following claim will also hold for ρ_{E}.

Lemma 2.4. *Let $X\subset\mathbb{C}$ be a Borel polar set. Then $\rho(X)=0$.*

Proof. By [62, Theorem 3.2.3] it suffices to show that for each $s > b_0$ we have

$$\int_{b_0}^{s}\int_{b_0}^{s} \log|x-t|\mathrm{d}\rho(x)\mathrm{d}\rho(t) > -\infty \tag{2.16}$$

By means of the subharmonic extension (2.11), M_∞ is non-negative on $\mathbb{C}$ and we get

$$0 \le \int_{b_0}^{s} M_\infty(x)\mathrm{d}\rho(x) = d + I_1 + I_2,$$

where

$$d = \rho(b_0,s)\left(1 - \int_{b_0}^{s}\log|t-z_*|\mathrm{d}\rho(t)\right), \quad I_1 = \int_{b_0}^{s}\int_{b_0}^{s}\log|x-t|\mathrm{d}\rho(t)\mathrm{d}\rho(x),$$
$$I_2 = \int_{b_0}^{s}\int_{s}^{\infty}\log\left|1-\frac{x-z_*}{t-z_*}\right|\mathrm{d}\rho(t)\mathrm{d}\rho(x).$$

Since $I_2 \le 0$, it follows that $-\infty < -d \le I_1$, i.e., we have (2.16). □

It was already encountered in [49, Lemma 2.4] that there is an explicit connection between ρ and the conformal map Θ defined in Lemma 2.3, see also [32]. Note that although in [49] Dirichlet-regularity is assumed for the set E the proof of the following lemma holds also in our setting. Namely, the Lebesgue measure on the base of the comb corresponds to the measure ρ on E. To be more precise, $\operatorname{Re}\Theta$ extends continuously to $\mathbb{R}$ and we have

$$\operatorname{Re}\Theta(b) - \operatorname{Re}\Theta(a) = \pi\rho((a,b)). \tag{2.17}$$

These are all the ingredients needed to describe the comb domains related to the conformal mapping Θ. There exists a positive upper semicontinuous function s on $(0,b)$, where $b \in (0,\infty]$ such that Θ maps $\mathbb{C}_+$ conformally onto

$$\Pi = \{x+iy : \ 0 < x < b, y > s(x)\}.$$

If $b < \infty$ then $\limsup_{x\to b} s(x) = \infty$. We will show in Corollary 2.8 that b being finite corresponds to ∞ being not Dirichlet-regular.

Example 2.5. In their classical work [55] Marchenko–Ostrovskii studied the relation between spectra of 1-periodic L^2 potentials on the real line and corresponding data of the mapping Θ_{E}. They showed that E is the spectrum of a Schrödinger operator of this type if and only if the corresponding comb domain is of the form

$$\Pi_{\mathsf{E}} = \{x+iy, x>0, y>0\} \setminus \{k\pi + iy : k \in \mathbb{N}, 0 \le y \le s_k\},$$

and the slit heights s_k satisfy $\sum_{k=1}^{\infty} k^2 s_k^2 < \infty$.

The next example demonstrates Akhiezer–Levin sets which don't have an expansion of the form (1.3).

Example 2.6. We will construct an explicit expression for the conformal map, $\Theta : \mathbb{C}_+ \to \Pi = \mathbb{C}_+ \setminus \{n+iy : n \in \mathbb{Z}, 0 < y \le y_0\}$, where $y_0 > 0$ is an arbitrary but fixed parameter. We will show that along the imaginary axis we have

$$\Theta(iy) = iy + ic(y_0) + o(1), \quad \text{as } y \to \infty,$$

where, $c(y_0)$ is a real constant that depends monotonically on y_0 and can attain in fact any real value. Note that Θ can be continuously extended to $\mathbb{R}$ and that $E := \Theta^{-1}(\mathbb{R})$ is symmetric, $E = -E = \{-x, x \in E\}$. Hence, again by defining $\tilde{\Theta}(z) = \Theta(\lambda^2)$, the function $M(z) = \operatorname{Im}\tilde{\Theta}(z)$

is an example for a Martin function of an Akhiezer–Levin set, which has a constant term in its asymptotic expansion. The Christoffel–Darboux transformation

$$f_1(w) = \frac{1}{\pi}\int_{-1}^{w} \frac{\mathrm{d}x}{\sqrt{1-x^2}}$$

maps $\mathbb{C}_+$ onto $\Pi_1 = \{\vartheta = \xi + i\eta : \eta > 0, 0 < \xi < 1\}$. In particular $f_1(-1) = 0$ and $f_1(1) = 1$. We choose $\ell > 1$ so that $iy_0 = f_1(-\ell)$ and consider

$$f_2(w) = \frac{1}{\pi}\int_{-\ell}^{w} \frac{\mathrm{d}x}{\sqrt{\ell^2-x^2}} = f_1(w/\ell).$$

Then $\Theta = f_1 \circ f_2^{-1}$ defines a conformal map $\Theta : \Pi_1 \to \Pi_1$ such that $\Theta(0) = iy_0$. By symmetry, we can extend Θ to a conformal map from $\Theta : \mathbb{C}_+ \to \Pi$. Calculations of f_1, f_2 along the imaginary axis give $\Theta(iy) = i\cosh^{-1}(\ell\cosh(y))$, so

$$\Theta(iy) = iy + i\log(\ell) + o(1), \quad \text{as } y \to \infty.$$

We emphasize that in order to show that the limit in (2.8) is always finite for the Martin function, it was only used that M_∞ represents a positive harmonic function in Ω. This shows that the same conclusion holds for any such function. In view of (2.4) this growth should also be reflected in the corresponding asymptotic behavior of M_∞, leading to the following criterion for E to be an Akhiezer–Levin set.

Lemma 2.7. *Assume that there exists a positive harmonic function in Ω such that*

$$\lim_{z\to-\infty} \frac{h(z)}{\sqrt{-z}} = 1.$$

Then Ω is Greenian and E is an Akhiezer–Levin set. Moreover, in this case we have

$$M_{\mathsf{E}}(z) \le h(z), \tag{2.18}$$

for all $z \in \Omega$, where M_{E} is normalized by $\lim_{z\to-\infty} M_{\mathsf{E}}(z)/\sqrt{-z} = 1$.

Proof. By Myrberg's theorem [3, Theorem 5.3.8] the existence of a non-constant positive harmonic function on Ω implies that Ω is Greenian. Since h is a positive harmonic function in Ω there exists a unique measure ν with $\nu(\partial_1^M\Omega) = h(z_*)$ such that

$$h(z) = \int_{\partial_1^M\Omega} M(z,x)\mathrm{d}\nu(x).$$

In particular, $\nu(\{\infty\}) < \infty$. Recall that $\#\pi^{-1}(\{\infty\}) = 1$. Since $(-\infty, b_0) \subset \Omega$, the negative half axis is clearly not minimally thin at ∞ so it follows by [3, Theorem 9.2.6] that

$$\liminf_{z\to-\infty} \frac{h(z)}{M_\infty(z)} \le \nu(\{\infty\}) < \infty. \tag{2.19}$$

Let $\lambda^2 = z - b_0$ and $g(\lambda) = h(z)$ and $f(\lambda) = M_\infty(z)$. Then f defines a positive harmonic function in $\mathbb{C}_+$ and

$$f(x+iy) = ay + \int \frac{y}{(x-t)^2+y^2}\mathrm{d}\mu(t), \quad a = \lim_{y\to\infty}\frac{f(iy)}{y}.$$

Hence,

$$0 < \limsup_{z\to-\infty}\frac{M_\infty(z)}{h(z)} = \limsup_{y\to\infty}\frac{f(iy)}{g(iy)} = \limsup_{y\to\infty}\frac{f(iy)}{y} = a.$$

Hence, E is an Akhiezer–Levin set. Due to [3, Theorem 9.3.3] we have

$$\nu(\{\infty\}) = \inf_{z\in\Omega} \frac{h(z)}{M_\infty(z)} \leq \frac{h(z)}{M_\infty(z)} \tag{2.20}$$

and the second claim follows. Finally, (2.20) shows that we actually have equality in (2.19) and it follows that $\nu(\{\infty\})$ corresponds to the normalization of M_∞ at ∞. $\square$

Carleson and Totik [10] showed that $\mathcal{P}_{\mathsf{E}}(x_0)$ being two-dimensional is equivalent to the fact that $G_{\mathsf{E}}(z, z_0)$ is Lipschitz continuous at x_0, where z_0 is some arbitrary interior point. As a corollary of the comb mapping representation for Θ, we show that E being an Akhiezer–Levin set implies continuity at infinity. Note that by the aforementioned equivalence, one cannot hope for Lipschitz continuity for semibounded sets, since in this case $\mathcal{P}_{\mathsf{E}}(\infty)$ is always one-dimensional. Alternatively, this could be seen from the fact that often, at a gap edge a, the Green function has behaviour $G_{\mathsf{E}}(z, z_0) \sim \sqrt{z-a}$ and thus is not Lipschitz continuous. Moreover, as discussed in [72] the set $\mathsf{E} = \mathbb{R}_+ \setminus \cup_{n\in\mathbb{Z}} r^n(a_1, b_1)$, where $0 < a_1 < b_1$ and $r > 1$ provides an example of a set for which ∞ is Dirichlet-regular, but which is not an Akhiezer–Levin set. In this sense the following result is optimal.

Corollary 2.8. *Let $\mathsf{E} \subset \mathbb{R}$ be closed and semibounded and Θ the corresponding comb-mapping. If $\sup\{\operatorname{Re}\Theta(z) : z \in \mathbb{C}_+\} = \infty$, then ∞ is a Dirichlet-regular point of E. This holds in particular if E is an Akhiezer–Levin set.*

Proof. We will assume that $\limsup_{z_0\to-\infty} G_{\mathsf{E}}(z_0, z_*) = \varepsilon > 0$ in order to obtain a contradiction. Note that $\sup\{\operatorname{Re}\theta_{z_0}(z) :\ z \in \mathbb{C}_+\} = \pi$, so for any $z \in \mathbb{C}_+$,

$$\lim_{z_0\to-\infty} \frac{\operatorname{Re}\theta_{z_0}(z)}{G_{\mathsf{E}}(z_*, z_0)} \leq \liminf_{z_0\to-\infty} \frac{\pi}{G_{\mathsf{E}}(z_*, z_0)}.$$

Since $\Theta(z) = \lim_{z_0\to-\infty} \frac{\theta_{z_0}(z)}{G_{\mathsf{E}}(z_*,z_0)}$, taking the supremum over $z \in \mathbb{C}_+$ gives $\sup\{\operatorname{Re}\Theta(z) :\ z \in \mathbb{C}_+\} \leq \varepsilon^{-1}\pi < \infty$. Now, as already mentioned in [32], using upper semicontinuity of h it follows that vanishing of the radial limit of $G_{\mathsf{E}}(z_0, z_*)$ implies Dirichlet-regularity. Let $\operatorname{Im}\theta_{z_*} = G_{\mathsf{E}}(z, z_*)$ and it will be more convenient to shift the mapping by $-\pi$. Then, $\lim_{z_0\to-\infty} G_{\mathsf{E}}(z_0, z_*) = 0$ implies that $\limsup_{u\to 0} h(u) = 0$. Therefore, $(-\infty, z_*)$ is mapped by θ_{z_*} onto $i\mathbb{R}_+$ and we can extend θ_{z_*} by symmetry to $\mathbb{C} \setminus (\mathbb{R} \setminus (-\infty, z_*))$. In particular $i\mathbb{R}_+$ is an interior ray of the image, $\Pi_e = \Pi_{z_*} \cup i\mathbb{R}_+ \cup \{-x+iy : x+iy \in \Pi_{z_*}\}$, of this extended map. $\limsup_{u\to 0} h(u) = 0$ now implies that Π_e is locally connected at 0 and hence, θ_{z_*} can be continuously extended to 0 which implies that ∞ is a Dirichlet-regular point. This finishes the proof of the first claim.

In view of (2.17), $\sup\{\operatorname{Re}\Theta(z) : z \in \mathbb{C}_+\} < \infty$ means that ρ is finite. We show that this implies that M_∞ can grow at most like $\rho(\mathbb{R})\log|z|$ and therefore E is not an Akhiezer–Levin set. Let's assume that $|z_* - b_0| > 1$ and $z_* < 0$. Then, using (2.12) we see that for $z < z_*$ we have

$$M_\infty(z) - \rho(\mathbb{R})\log|z| = M_\infty(z_*) + \int_{b_0}^{\infty} \log\left|\frac{1}{t-\lambda_*}\left(1 - \frac{z_*}{z}\right)\right| \mathrm{d}\rho(t) \leq M_\infty(z_*). \qquad \square$$

For Akhiezer–Levin sets one could also use the result of Carleson and Totik and the substitution $\lambda^2 = z - b_0$ to see that G_{E} is Hölder continuous with exponent $1/2$ at ∞.

3. Asymptotic behavior of eigensolutions

We now turn our attention to the Schrödinger operator L_V and associated objects. Fundamental solutions at $z \in \mathbb{C}$ are defined as solutions $u(x,z), v(x,z)$ of the initial value problems

$$-\partial_x^2 u + (V(x) - z)u = 0, \qquad u(0,z) = 0, \qquad (\partial_x u)(0,z) = 1 \tag{3.1}$$

$$-\partial_x^2 v + (V(x) - z)v = 0, \qquad v(0,z) = 1, \qquad (\partial_x v)(0,z) = 0 \tag{3.2}$$

The natural regularity class for the solutions are functions which are in $W^{2,1}([0,x])$ for every $x < \infty$, and the differential equations are interpreted as equality of L^1 functions, i.e., equality Lebesgue-a.e.. It is useful to substitute

$$k = \sqrt{-z}$$

and view the initial value problems as perturbations by V of $-\partial_x^2 + k^2$. We will always assume that $\operatorname{Re} k \geq 0$; this can be done pointwise throughout $\mathbb{C}$, and later we will view k as a branch of the square root such that $\operatorname{Re} k > 0$ if $z \in \mathbb{C} \setminus [0,\infty)$. Note also that this makes $\operatorname{Im} k < 0$ if $z \in \mathbb{C}_+$. By choosing the branch $\sqrt{z} = ik$, we see that $\sqrt{z} \in \mathbb{C}_+$ if $z \in \mathbb{C} \setminus [0,\infty)$. In particular, $\operatorname{Im} \sqrt{z} = \operatorname{Re} k$.

The fundamental solutions for $V = 0$ are the functions

$$c(x,k) = \cosh(kx), \qquad s(x,k) = \begin{cases} \frac{\sinh(kx)}{k} & k \neq 0 \\ x & k = 0 \end{cases}.$$

By standard arguments, for general $V \in L^1([0,1])$, the initial value problems (3.1), (3.2) are rewritten as integral equations, and by Volterra-type arguments, convergent series representations are then found for the fundamental solutions. With the notation $\Delta_n(x) = \{t \in \mathbb{R}^n \mid x \geq t_1 \geq t_2 \geq \cdots \geq t_n \geq 0\}$, the series representations for fundamental solutions and their first derivatives are

$$u(x,z) = s(x,k) + \sum_{n=1}^{\infty} \int_{\Delta_n(x)} s(x - t_1, k) \left(\prod_{j=1}^{n-1} V(t_j) s(t_j - t_{j+1}, k) \right) V(t_n) s(t_n, k) \mathrm{d}^n t \tag{3.3}$$

$$v(x,z) = c(x,k) + \sum_{n=1}^{\infty} \int_{\Delta_n(x)} s(x - t_1, k) \left(\prod_{j=1}^{n-1} V(t_j) s(t_j - t_{j+1}, k) \right) V(t_n) c(t_n, k) \mathrm{d}^n t \tag{3.4}$$

$$(\partial_x u)(x,z) = c(x,k) + \sum_{n=1}^{\infty} \int_{\Delta_n(x)} c(x - t_1, k) \left(\prod_{j=1}^{n-1} V(t_j) s(t_j - t_{j+1}, k) \right) V(t_n) s(t_n, k) \mathrm{d}^n t \tag{3.5}$$

$$(\partial_x v)(x,z) = k^2 s(x,k) + \sum_{n=1}^{\infty} \int_{\Delta_n(x)} c(x - t_1, k) \left(\prod_{j=1}^{n-1} V(t_j) s(t_j - t_{j+1}, k) \right) V(t_n) c(t_n, k) \mathrm{d}^n t \tag{3.6}$$

These expansions are derived, e.g., in [61] for $V \in L^2([0,x])$, but they hold for $V \in L^1([0,x])$ as well, due to the estimate

$$\begin{aligned} &\left| \int_0^x \int_0^{t_1} \cdots \int_0^{t_{n-1}} e^{\operatorname{Re} k (x - t_1)} \left(\prod_{j=1}^{n} V(t_j) e^{\operatorname{Re} k (t_j - t_{j+1})} \right) V(t_n) e^{\operatorname{Re} k t_n} \, dt_n \dots dt_2 \, dt_1 \right| \\ &\leq \frac{1}{n!} \left(\int_0^x |V(s)| \mathrm{d}s \right)^n e^{\operatorname{Re} kx} \end{aligned} \tag{3.7}$$

which is proved by combining the exponentials and using permutations of t and symmetry, and the elementary estimates which follow directly from Euler's formula,

$$|c(x,k)| \le e^{\operatorname{Re} kx}, \qquad |s(x,k)| \le |k|^{-1} e^{\operatorname{Re} kx}. \tag{3.8}$$

The same estimates which guarantee convergence, provide exponential upper bounds on eigensolutions; these are often stated over a fixed interval, but we will need a kind of uniformity in x:

Lemma 3.1. *For all $z = -k^2 \in \mathbb{C}$ and $x > 0$,*

$$|u(x,-k^2)| \le e^{(1+\operatorname{Re} k)x + \int_0^x |V(t)| dt}. \tag{3.9}$$

Proof. Using $|s(x,k)| = |\int_0^x c(t,k)\mathrm{d}t| \le x e^{\operatorname{Re} kx} \le e^{(1+\operatorname{Re} k)x}$ and then applying (3.7) to each term of (3.3) implies that

$$|u(x,-k^2)| \le e^{(1+\operatorname{Re} k)x} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\int_0^x |V(t)| \right)^n. \qquad \square$$

Corollary 3.2. *If V obeys (1.1), for each $R > 0$ there exists C_R such that for all $|z| \le R$ and $x \ge 1$ we have $\frac{1}{x} \log |u(x,z)| \le C_R$.*

Proof. This is an immediate consequence of the previous lemma together with $\int_0^x |V(t)| \mathrm{d}t \le C(x+1) \le 2Cx$ for $x \ge 1$, where $C = \sup_{x \ge 0} \int_x^{x+1} |V(t)| \mathrm{d}t$. $\square$

We will need asymptotic statements about m-functions. Such statements are ubiquitous, especially for smooth potentials; we need an asymptotic expansion which doesn't assume any smoothness.

Lemma 3.3. *For fixed $x > 0$, as $z \to \infty$, $\arg z \in [\delta, 2\pi - \delta]$,*

$$-\frac{v(x,z)}{u(x,z)} = -k - \int_0^x V(t) e^{-2kt}\, dt + \frac{1}{k} \int_0^x \int_0^{t_1} e^{-2kt_1}(1 - e^{-2kt_2}) V(t_1) V(t_2)\, dt_2\, dt_1 + O(|k|^{-2})$$

uniformly in V in bounded subsets of $L^1([0,x])$.

Proof. Assume that $\int_0^x |V(t)| \mathrm{d}t \le C$. Denote

$$A_n = 2k^{n+1} e^{-kx} \int_{\Delta_n(x)} s(x - t_1, k) \left(\prod_{j=1}^{n-1} V(t_j) s(t_j - t_{j+1}, k) \right) V(t_n) s(t_n, k)\, d^n t,$$

$$B_n = 2k^n e^{-kx} \int_{\Delta_n(x)} s(x - t_1, k) \left(\prod_{j=1}^{n-1} V(t_j) s(t_j - t_{j+1}, k) \right) V(t_n) c(t_n, k)\, d^n t,$$

From (3.8) and (3.7) it follows that $|A_n|, |B_n| \le \frac{2C^n}{n!}$. In the nontangential limit $z \to \infty$, $\arg z \in [\delta, 2\pi - \delta]$, we have the elementary estimates

$$\frac{s(x,k)}{\frac{e^{kx}}{2k}} = 1 - e^{-2kx} = 1 + O(|k|^{-3}), \qquad \frac{c(x,k)}{\frac{e^{kx}}{2}} = 1 + e^{-2kx} = 1 + O(|k|^{-3}),$$

so the series expansions for $u(x,z)$, $v(x,z)$ imply

$$u(x,z) = \frac{e^{kx}}{2k} \left(1 + \frac{A_1}{k} + \frac{A_2}{k^2} + O(|k|^{-3}) \right),$$

$$v(x,z) = \frac{e^{kx}}{2}\left(1 + \frac{B_1}{k} + \frac{B_2}{k^2} + O(|k|^{-3})\right),$$

with the error $O(|k|^{-3})$ depending only on C and δ. Dividing,

$$-\frac{v(x,z)}{u(x,z)} = -k\left(1 + \frac{B_1 - A_1}{k} + \frac{B_2 - A_2 - A_1(B_1 - A_1)}{k^2} + O(|k|^{-3})\right). \tag{3.10}$$

Moreover,

$$B_1 - A_1 = \int_0^x (1 - e^{-2k(x-t)})V(t)e^{-2kt}\,dt = \int_0^x V(t)e^{-2kt}\,dt + O(e^{-2\operatorname{Re} kx}) \tag{3.11}$$

Multiplying by $A_1 = \frac{1}{2}\int_0^x (1 - e^{-2k(x-s)})V(s)(1 - e^{-2ks})\,ds$ gives a formula for $A_1(B_1 - A_1)$ as a double integral over $[0,x]^2$, and using the substitution $t_1 = \max\{s,t\}$, $t_2 = \min\{s,t\}$ gives

$$\begin{aligned} A_1(B_1 - A_1) = {} & \frac{1}{2}\int_0^x \int_0^{t_1} (e^{-2kt_1} + e^{-2kt_2} - 2e^{-2k(t_1+t_2)} - e^{-2k(x-t_1+t_2)})V(t_1)V(t_2)\,dt_2\,dt_1 \\ & + O(e^{-2\operatorname{Re} kx}) \end{aligned}$$

(some terms are grouped into the error $O(e^{-2\operatorname{Re} kx})$ since, e.g., $x - t_2 + t_1 \geq x$). Similarly,

$$\begin{aligned} B_2 - A_2 &= \frac{1}{2}\int_0^x \int_0^{t_1} (1 - e^{-2k(x-t_1)})V(t_1)(1 - e^{-2k(t_1-t_2)})V(t_2)e^{-2kt_2}\,dt_2\,dt_1 \\ &= \frac{1}{2}\int_0^x \int_0^{t_1} (e^{-2kt_2} - e^{-2kt_1} - e^{-2k(x-t_1+t_2)})V(t_1)V(t_2)\,dt_2\,dt_1 + O(e^{-2\operatorname{Re} kx}) \end{aligned}$$

Substituting these formulas into (3.10) concludes the proof. □

Returning to the half-line setting from the introduction, we recall that half-line potentials obeying the boundedness assumption (1.1) are in the limit point case at $+\infty$, i.e., for every $z \in \mathbb{C} \setminus \mathsf{E}$, the set of solutions of

$$-\partial_x^2 \psi + V\psi = z\psi, \qquad \psi \in L^2((0,\infty))$$

is one-dimensional. Any such nontrivial solution is called the Weyl solution; it is uniquely determined up to normalization and we will not fix any particular normalization. We will use

$$m(x,z) = \frac{(\partial_x \psi)(x,z)}{\psi(x,z)}. \tag{3.12}$$

Proposition 3.4. *As $z \to \infty$, $\arg z \in [\delta, \pi - \delta]$,*

$$\begin{aligned} m(s,z) = {} & -k - \int_0^1 V(s+t)e^{-2kt}\,dt \\ & + \frac{1}{k}\int_0^1 \int_0^{t_1} e^{-2kt_1}(1 - e^{-2kt_2})V(s+t_1)V(s+t_2)\,dt_2\,dt_1 + O(|k|^{-2}) \end{aligned}$$

and the error is uniform in $s \in [0,\infty)$ if V obeys (1.1).

Proof. By an argument of Atkinson [4], for $\arg z \in [\delta, \pi - \delta]$, the Weyl circle at x has radius

$$r = \frac{2|k|^2}{|\operatorname{Im} k|}e^{-2x\operatorname{Re} k}(1 + O(|k|^{-1}))$$

which decays exponentially as $z \to \infty$, $\arg z \in [\delta, \pi - \delta]$; the error term $O(|k|^{-1})$ is uniform for V in bounded subsets of $[0,x]$, since this term is derived by arguments like those in the proof of

Lemma 3.3. Since $m_+(0,z)$ lies inside the Weyl circle and $-v(1,z)/u(1,z)$ lies on the circle, this radius allows us to estimate

$$\left| m(0,z) + \frac{v(1,z)}{u(1,z)} \right| \le \frac{4|k|^2}{|\operatorname{Im} k|} e^{-2\operatorname{Re} k}(1 + O(|k|^{-1})).$$

In the nontangential limit as $\arg z \in [\delta, \pi-\delta]$, this error is $O(|k|^{-2})$, so the previous lemma implies

$$m(0,z) = -k - \int_0^1 V(t) e^{-2kt}\, dt + \frac{1}{k}\int_0^1 \int_0^{t_1} e^{-2kt_1}(1 - e^{-2kt_2}) V(t_1) V(t_2)\, dt_2\, dt_1 + O(|k|^{-2}).$$

Applying this for an arbitrary $s \ge 0$ to the translated half-line potential $V_s(x) = V(x+s)$ on $[0,\infty)$ concludes the proof. □

For the half-line operator L_V, the Dirichlet solution can be interpreted as the Weyl solution corresponding to the endpoint 0. Therefore, the Atkinson argument can be applied also "in reverse", to produce uniform asymptotics on the logarithmic derivative of $u(x,z)$. To produce uniform asymptotics, we fix the interval length 1, as in the previous proof:

Corollary 3.5. *As $z \to \infty$, $\arg z \in [\delta, \pi - \delta]$, for all $s \ge 1$,*

$$\begin{aligned} -\frac{(\partial_x u)(s,z)}{u(s,z)} &= -k - \int_0^1 V(s-t) e^{-2kt}\, dt \\ &\quad + \frac{1}{k}\int_0^1 \int_0^{t_1} e^{-2kt_1}(1 - e^{-2kt_2}) V(s-t_1) V(s-t_2)\, dt_2\, dt_1 + O(|k|^{-2}) \end{aligned}$$

and the error is uniform in $s \in [1,\infty)$ if V obeys (1.1).

To make some uniform statements for a family of Herglotz functions, we will use the Carathéodory inequality for the half-plane [47, Proof of Theorem I.8]: for any Herglotz function f,

$$|f(z)| \le |f(i)| + \operatorname{Im} f(i) \frac{2|z-i|}{|z+i| - |z-i|}, \qquad \forall z \in \mathbb{C}_+. \tag{3.13}$$

Lemma 3.6. *Fix a potential V which obeys* (1.1)*. For each $z \in \mathbb{C}_+$,*

$$\sup_{x \ge 1} \left| \frac{(\partial_x u)(x,z)}{u(x,z)} \right| < \infty. \tag{3.14}$$

Proof. The ratio $-(\partial_x u)(x,z)/u(x,z)$ is a Herglotz function and obeys the nontangential asymptotics in Corollary 3.5. The error is uniform in $x \ge 1$ since V obeys (1.1). In particular, for $z = iy_0$ with some fixed $y_0 > 0$ large enough, Corollary 3.5 implies an upper bound independent of x and therefore (3.14). By rescaling by y_0 and using (3.13), the upper bound at iy_0 implies uniform upper bounds for z in compact subsets of $\mathbb{C}_+$. □

For $z \notin \sigma(L_V)$, ψ decays exponentially as $x \to \infty$. The Weyl solution ψ and the Dirichlet solution u are related by the Wronskian

$$W(\psi, u) = (\partial_x u)(x,z)\psi(x,z) - (\partial_x \psi)(x,z) u(x,z)$$

which is independent of x and nonzero, since u, ψ are linearly independent (otherwise they would give an eigenvalue of L_V). This strongly suggests that u should grow at the same rate at which ψ decays, but a proof based only on the Wronskian is difficult due to the derivative, especially if a pointwise statement is desired. We therefore use a different argument:

Lemma 3.7. *Fix a potential V which obeys* (1.1). *For each $z \in \mathbb{C}_+$, there exists C such that for all $x \in [1,\infty)$,*

$$C^{-1} \le |u(x,z)\psi(x,z)| \le C.$$

Proof. We use the diagonal (spectral theoretic) Green's function for L_V,

$$g(x,x;z) = \frac{u(x,z)\psi(x,z)}{W(\psi,u)}, \tag{3.15}$$

which can be written as

$$-\frac{1}{g(x,x;z)} = \frac{(\partial_x\psi)(x,z)}{\psi(x,z)} - \frac{(\partial_x u)(x,z)}{u(x,z)}. \tag{3.16}$$

Using the above asymptotics for m-functions gives a well known asymptotic statement,

$$g(x,x;z) = \frac{1}{2\sqrt{-z}} + O(|z|^{-1}), \qquad z\to\infty, \arg z \in [\delta, \pi-\delta],$$

and the proof given here shows that this asymptotic behavior is uniform in $x \in [1,\infty)$, since V obeys (1.1). In particular, for some fixed $z = iy$ with y large enough, this implies

$$\sup_{x\in[1,\infty)} |g(x,x;iy)| < \infty, \qquad \inf_{x\in[1,\infty)} |g(x,x;iy)| > 0.$$

Rescaling z by a factor y and applying (3.13) to the Herglotz functions $g(x,x;z)$ and $-1/g(x,x;z)$ implies uniform upper and lower bounds on compact subsets of $\mathbb{C}_+$.

For any $z \in \mathbb{C}_+$, the Wronskian is nonzero and independent of x, so by (3.15), uniform bounds in x for $g(x,x;z)$ imply uniform bounds in x (for each $z \in \mathbb{C}_+$) for $u(x,z)\psi(x,z)$. ☐

The growth rate of $u(x,z)$ can now be expressed in terms of averages of the m-functions:

Corollary 3.8. *For any $z \in \mathbb{C}_+$,*

$$\limsup_{x\to\infty} \left| \frac{1}{x}\log u(x,z) + \frac{1}{x}\int_0^x m(s,z)\mathrm{d}s \right| = 0. \tag{3.17}$$

Proof. This follows from Lemma 3.7 since $m(x,z)$ is the logarithmic derivative of $\psi(x,z)$. ☐

Expansions for $m(s,z)$ are often stated in terms of values of V and its derivatives at s, but such expansions assume some regularity of V, and the error terms in such expansions are usually not uniform in the appropriate local norm for V. By working directly with the expansion in Prop. 3.4, we can obtain uniform expansions for the averages without imposing any regularity on V.

Corollary 3.9. *If V obeys* (1.1),

$$\limsup_{x\to\infty} \left| \frac{1}{x}\int_0^x m(s,z)\mathrm{d}s + k + \frac{1}{2kx}\int_0^x V(s)\mathrm{d}s \right| = O(|k|^{-2}), \tag{3.18}$$

as $z = -k^2 \to \infty$, $\arg z \in [\delta,\pi-\delta]$, for any $\delta > 0$.

Proof. Due to the uniformity of the error in the asymptotic expansion from Prop. 3.4,

$$\begin{aligned}\frac{1}{x}\int_0^x m(s,z)\,ds = -k - \frac{1}{x}\int_0^x\int_0^1 V(s+t)e^{-2kt}\mathrm{d}t\mathrm{d}s \\ + \frac{1}{kx}\int_0^x\int_0^1\int_0^{t_1} e^{-2kt_1}(1-e^{-2kt_2})V(s+t_1)V(s+t_2)\mathrm{d}t_2\mathrm{d}t_1\mathrm{d}s + O(|k|^{-2})\end{aligned}$$

with the error term independent of x. For the term linear in V, we use $p = s + t$ to rewrite the iterated integral as $\int_0^1 \int_t^{x+t} V(p)e^{-2kt} dpdt$. Then we wish to note that

$$\frac{1}{x}\int_0^1 \int_t^{x+t} V(p)e^{-2kt}\mathrm{d}p\mathrm{d}t = \frac{1}{x}\int_0^1 \int_0^{x} V(p)e^{-2kt}\mathrm{d}p\mathrm{d}t + O(x^{-1}), \qquad x \to \infty, \tag{3.19}$$

for any k. This is because the two iterated integrals describe similar regions in $\mathbb{R}^2$: the symmetric difference of the regions $\{(t,p) \mid 0 \le t \le 1, t \le p \le x + t\}$ and $\{(t,p) \mid 0 \le t \le 1, 0 \le p \le x\}$ is contained in $[0,1] \times ([0,1] \cup [x, x+1])$, and the double integral over that region is bounded uniformly in x due to (1.1). Now the integral in (3.19) separates and simplifies using $\int_0^1 e^{-2kt}\mathrm{d}t = \frac{1}{2k} + O(e^{-2\operatorname{Re} k})$. By analogous arguments, using $q = s + t_2$ to rewrite the quadratic term and comparing the regions $\{(t_1,t_2,q) \mid 0 \le t_2 \le t_1 \le 1, t_2 \le q \le x + t_2\}$ and $\{(t_1,t_2,q) \mid 0 \le t_2 \le t_1 \le 1, 0 \le q \le x\}$,

$$\begin{aligned}
&\frac{1}{kx}\int_0^x \int_0^1 \int_0^{t_1} e^{-2kt_1}(1 - e^{-2kt_2})V(s+t_1)V(s+t_2)\mathrm{d}t_2\mathrm{d}t_1\mathrm{d}s \\
&= \frac{1}{kx}\int_0^1 \int_0^{t_1} \int_0^{x} e^{-2kt_1}(1 - e^{-2kt_2})V(q+t_1-t_2)V(q)\mathrm{d}q\mathrm{d}t_2\mathrm{d}t_1 + O(x^{-1}) \\
&= \frac{1}{kx}\int_0^1 \int_0^{x} h(u)V(q+u)V(q)\mathrm{d}q\mathrm{d}u + O(x^{-1})
\end{aligned}$$

as $x \to \infty$, for any k. For the last step we introduced $u = t_1 - t_2 \in [0,1]$ and $h(u) = \int_0^{1-u} e^{-2k(u+t_2)}(1 - e^{-2kt_2})\mathrm{d}t_2$. For the remaining double integral, it is elementary to estimate that $h(u) = O(|k|^{-1})$ uniformly in $u \in [0,1]$ and that

$$\frac{1}{x}\int_0^1 \int_0^x |V(q+u)V(q)|\mathrm{d}q\mathrm{d}u \le C^2$$

where C denotes the sup in (1.1), so (3.18) follows. □

4. Regular measures for half-line Schrödinger operators

The main part of this section is devoted to the study of limits of the function

$$h(x,z) := \frac{1}{x}\log|u(x,z)|, \tag{4.1}$$

as $x \to \infty$. Our first goal is to show that for $z \in \mathbb{C}_+$ we have that $\liminf_{x\to\infty} h(x,z) \ge 0$.

Lemma 4.1. *Fix $z \in \mathbb{C}_+$. Then*

$$\liminf_{x\to\infty} \frac{1}{x}\log|u(x,z)| \ge 0.$$

Proof. Note first of all that $u(x,z) \ne 0$ whenever $x > 0$, because the converse would correspond to an complex eigenvalue for the self-adjoint realization of L_V on $[0,x]$ with Dirichlet boundary conditions. The Weyl solution $\psi(x,z)$ is an eigensolution and is in $L^2((0,\infty))$; the condition (1.1) is sufficient to conclude that ψ decays pointwise [52, Theorem 1.1], i.e.

$$\lim_{x\to\infty} \psi(x,z) = 0.$$

Combining with Lemma 3.7 shows that $|u(x,z)| \to \infty$ as $x \to \infty$, which completes the proof. □

Let $\mathsf{E} = \sigma_{\mathrm{ess}}(L_V)$ written in the form (1.4). That is $b_0 = \min \mathsf{E}$ and (a_j, b_j) denote the gaps of E.

Lemma 4.2. *For any $\varepsilon > 0$ there exists $x_0 > 0$ such that $u(x,z) \neq 0$ for $x > x_0$ and $z \leq b_0 - \varepsilon$. Moreover, let $n_j(\varepsilon)$ denote the finite number of eigenvalues in $(a_j + \varepsilon, b_j - \varepsilon)$. Then, for any $x > 0$, $u(x,z)$ has at most $n_j(\varepsilon) + 1$ zeros in $(a_j + \varepsilon, b_j - \varepsilon)$.*

Proof. Since L_V is semibounded there are at most finitely many eigenvalues below $b_0 - \varepsilon$. Hence, the first statement follows by Sturm oscillation theory.

As in the proof of Lemma 3.7, we use the spectral theoretic Green's function $g(x,x;z)$. By the Weyl M-matrix representation for L_V centered at x, $g(x,x;\cdot)$ is analytic on $\mathbb{C} \setminus \sigma(L_V)$ and, since it is Herglotz, it is strictly increasing on intervals in $\mathbb{R} \setminus \sigma(L_V)$. In particular, every pole of $g(x,x;\cdot)$ is an eigenvalue of L_V, so it has at most $n_j(\varepsilon)$ poles in $(a_j+\varepsilon, b_j-\varepsilon)$. By (3.16), every zero of $u(x,z)$ is a pole of $-(\partial_x u)(x,z)/u(x,z)$ and a zero of $g(x,x;\cdot)$. Since zeros and poles of the Herglotz function $g(x,x;\cdot)$ strictly interlace on intervals in the domain of meromorphicity, it follows that $u(x,z)$ has at most $n_j(\varepsilon) + 1$ zeros in $(a_j + \varepsilon, b_j - \varepsilon)$. ☐

We are now ready to study the existence of limit points for the family of functions $\mathcal{F} = \{h(x,z)\}_{x\in[1,\infty)}$. Since $u(x,\cdot)$ are entire functions, the functions $h(x,\cdot)$ are subharmonic in $\mathbb{C}$, and they can be viewed as elements of the space of distributions $\mathcal{D}'(\mathbb{C})$ with nonnegative distributional Laplacian.

Theorem 4.3. *(a) The family $\mathcal{F} = \{h(x,z)\}_{x\in[1,\infty)}$ is precompact in $\mathcal{D}'(\mathbb{C})$.*
(b) For any sequence $(x_j)_{j=1}^\infty$ with $x_j \to \infty$ such that $h(x_j,\cdot)$ converges in $\mathcal{D}'(\mathbb{C})$, the limit $h = \lim_{j\to\infty} h(x_j,\cdot)$ is also a subharmonic function on $\mathbb{C}$, harmonic on $\mathbb{C} \setminus \mathsf{E}$, and $h(x_j,\cdot)$ also converge to h uniformly on compact subsets of $\mathbb{C} \setminus \mathsf{E}$.

Proof. (a) By Corollary 3.2, $h(x,z)$ is uniformly bounded from above on compact subsets of $\mathbb{C}$. Moreover, Lemma 4.1 implies a pointwise lower bound at some arbitrary point $z_0 \in \mathbb{C}_+$. Hence, [39, Theorem 4.1.9] shows that $\mathcal{F}$ is precompact in the topology of $\mathcal{D}'(\mathbb{C})$.

(b) On $\mathbb{C}_+$ and on $\mathbb{C}_-$, the functions $h(x,z)$ are harmonic and uniformly bounded above. Since they are also pointwise bounded below, they are uniformly bounded and uniformly equicontinuous on each compact subset of $\mathbb{C}_\pm$. Therefore, they are precompact in the topology of uniform convergence on compact subsets of $\mathbb{C}_\pm$. Since this convergence implies convergence in $L^1_{\mathrm{loc}}(\mathbb{C}_\pm)$, it follows that if the sequence $h(x_j,\cdot)$ converges in $\mathcal{D}'(\mathbb{C})$ to h, then it also converges to h uniformly on compact subsets of $\mathbb{C}_\pm$.

Next, we show that h has a harmonic extension through an arbitrary gap (a_m, b_m) of E. Fix $\varepsilon > 0$. By Lemma 4.2, there are at most $n_m(\varepsilon) + 1$ zeros of $u(x_j, z)$ in $(a_m + \varepsilon, b_m - \varepsilon)$. Let p_j be the monic polynomial of degree at most $n_m(\varepsilon) + 1$ which vanishes exactly at these zeros. Now consider

$$f_j(z) = \frac{1}{x_j} \log \left| \frac{u(x_j,z)}{p_j(z)} \right|,$$

which is harmonic on $\mathbb{C}_+ \cup \mathbb{C}_- \cup (a_m+\varepsilon, b_m-\varepsilon)$. On the boundary of the rectangle $(a_m-1, b_m+1) \times (-1,1)$, p_j is uniformly bounded below by 1, so by the maximum principle, the analytic functions $\frac{u(x_j,z)}{p_j(z)}$ are also bounded above by e^{cx_j} in this rectangle for some constant c. Hence, $f_j(z)$ is locally uniformly bounded above on $R_m = (a_m + \epsilon, b_m - \epsilon) \times (-1,1)$. Since all zeros of p_j are in (a_m, b_m), there is still a pointwise lower bound for $z_0 \in \mathbb{C}_+$. Hence, the functions f_j are harmonic on R_m and precompact in the topology of uniform convergence on compacts. For any $z \in R_m \setminus \mathbb{R}$,

$$\lim_{j\to\infty} (h_j(z) - f_j(z)) = \lim_{j\to\infty} \frac{1}{x_j} \log|p_j(z)| = 0$$

since $|\operatorname{Im} z|^{n_m(\varepsilon)+1} \le |p_j(z)| \le (b_m - a_m + 1)^{n_m(\varepsilon)+1}$. Hence, any subsequential limit of the $f_j(z)$ is a harmonic function on R_m which agrees with h on $R_m \setminus \mathbb{R}$. It follows that f_j converge in R_m uniformly on compacts, so it provides a harmonic extension for h through $(a_m + \varepsilon, b_m - \varepsilon)$. Since $\varepsilon > 0$ was arbitrary and the extensions must coincide on their common domain, we obtain an extension through (a_m, b_m) by letting $\varepsilon \to 0$. It follows from the weak identity principle for subharmonic functions [62, Theorem 2.7.5] that the harmonic extension coincided with h.

Consider a compact $K \subset \mathbb{C} \setminus [b_0, \infty)$. By possibly increasing K, assume that $K \not\subset \mathbb{R}$. Choose an open set U such that $K \subset U \subset \overline{U} \subset \mathbb{C} \setminus [b_0, \infty)$. By Lemma 4.2, for all sufficiently large j, $h_j(z)$ is harmonic in U. The functions h_j are uniformly bounded above and pointwise bounded below at $z_0 \in K \cap (\mathbb{C}_+ \cup \mathbb{C}_-)$, so they form a precompact sequence with respect to uniform convergence on K. As before, every limit is equal to h, so h_j converge to h uniformly on compacts. □

Collecting our results now yields that the limits define a positive harmonic function in $\Omega = \mathbb{C} \setminus \mathsf{E}$.

Theorem 4.4. *Let $x_j \to \infty$ be a sequence such that $h_j = h(x_j, \cdot)$ converge in $\mathcal{D}'(\mathbb{C})$. Then $h = \lim_{j\to\infty} h_j$ defines a positive harmonic function in Ω, the limit*

$$a = \lim_{j\to\infty} \frac{1}{x_j} \int_0^{x_j} V(x) \mathrm{d}x \tag{4.2}$$

exists, and h has the nontangential asymptotic behavior

$$h(z) = \operatorname{Re}\left(k + \frac{a}{2k}\right) + O(|k|^{-2}), \tag{4.3}$$

$z \to \infty$, $\delta \le \arg z \le 2\pi - \delta$, for any $\delta > 0$.

Proof. Harmonicity of h was proved in Theorem 4.3 and positivity in $\mathbb{C}_+ \cup \mathbb{C}_-$ follows from Lemma 4.1. That h is also positive in $\mathbb{R} \setminus \mathsf{E}$ follows by the maximum principle for harmonic functions, and by Corollary 3.8,

$$h(z) = -\lim_{j\to\infty} \frac{1}{x_j} \operatorname{Re} \int_0^{x_j} m(x,z) \mathrm{d}x. \tag{4.4}$$

Denote $c = \min \sigma(L_V)$. By general spectral theory, then $m(x,z)$ are analytic functions on $\mathbb{C} \setminus [c, \infty)$ and $m(x,z) < 0$ on $(-\infty, c)$. Since convergence of analytic functions follows from convergence of their real parts together with convergence at one point, from $\operatorname{Im} m(x,z) = 0$ for $z < c$ together with (4.4), it follows that the limit

$$w(z) = \lim_{j\to\infty} \frac{1}{x_j} \int_0^{x_j} m(x,z) \mathrm{d}x$$

converges uniformly on compact subsets of $\mathbb{C} \setminus [c, \infty)$. If a denotes some accumulation point of the sequence $\frac{1}{x_j} \int_0^{x_j} V(x) \mathrm{d}x$, applying Corollary 3.9 along the subsequence and using uniformity of the error term, it follows that

$$w(z) = -k - \frac{a}{2k} + O(|k|^{-2}) \tag{4.5}$$

nontangentially as $z \to \infty$, with $\arg z \in [\delta, \pi - \delta]$. This asymptotic behavior can only hold for one value of a, so it follows that the limit (4.2) exists.

We know that (4.5) holds as $z \to \infty$ with $\arg z \in [\delta, \pi - \delta]$ and, by symmetry, for $\arg z \in [\pi + \delta, 2\pi - \delta]$. It remains to extend this asymptotic behavior to a sector of the form $\arg z \in [\pi - \delta, \pi + \delta]$.

Without loss of generality assume $c = 0$. Since $\operatorname{Re} w = -h \le 0$, the function $f(\lambda) = -iw(\lambda^2)$ is Herglotz, and obeys

$$f(\lambda) = \lambda - \frac{a}{2\lambda} + O(|\lambda|^{-2}), \quad |\lambda| \to \infty, \tag{4.6}$$

along the rays $\arg \lambda = \pi/2 - \delta/2$ and $\arg \lambda = \pi/2 + \delta/2$. In the sector $T = \{\lambda : \pi/2 - \delta/2 \le \arg \lambda \le \pi/2 + \delta/2\}$, the function $g(\lambda) = \lambda^2 \left(f(\lambda) - \lambda + \frac{a}{2\lambda}\right)$ is analytic. It has a continuous extension to $\overline{T}$ with $g(0) = 0$, because $f(\lambda) = O(1/\lambda)$ as $\lambda \to 0$ nontangentially. By (4.6), g is bounded on the boundary of T. Finally, since f is Herglotz, f, g grow at most polynomially as $\lambda \to \infty$, $\lambda \in T$, so by Phragmén–Lindelöf, g is bounded in T. This implies that f has the asymptotic behavior (4.6) also in the sector T. Rewriting the conclusion for w and $h = -\operatorname{Re} w$ completes the proof. ☐

We need the following variant of the Herglotz representation:

Lemma 4.5. *Let f be a Herglotz function. Assume that $\operatorname{Im} f(iy) = O(y^{-1})$ as $y \to \infty$. Then for some $\beta \in \mathbb{R}$*

$$f(\lambda) = \beta + \int_{\mathbb{R}} \frac{\mathrm{d}\mu(t)}{t - \lambda}, \quad \text{with } \lim_{y\to\infty} y \operatorname{Im} f(iy) = \mu(\mathbb{R}) < \infty,$$

and

$$f(\lambda) = \beta - \frac{\mu(\mathbb{R})}{\lambda} + o(|\lambda|^{-1}), \tag{4.7}$$

$\lambda \to \infty$, $\delta \le \arg \lambda \le \pi - \delta$, for any $\delta > 0$.

Proof. Starting from the Herglotz representation, we can write $\operatorname{Im} f(iy) = ay + \int \frac{y}{t^2+y^2} \mathrm{d}\mu(t)$, with $\lim\limits_{y\to\infty} \frac{\operatorname{Im} f(iy)}{y} = a$. Hence, by our assumption, $a = 0$. Moreover, by monotone convergence

$$\lim_{y\to\infty} y \operatorname{Im} f(iy) = \lim_{y\to\infty} \int \frac{y^2}{t^2 + y^2} \mathrm{d}\mu(t) = \mu(\mathbb{R}).$$

By our assumption, this shows that $\mu(\mathbb{R}) < \infty$. We have $\lambda \int_{\mathbb{R}} \frac{\mathrm{d}\mu(t)}{t-\lambda} + \mu(\mathbb{R}) = \int_{\mathbb{R}} \frac{t}{t-\lambda} \mathrm{d}\mu(t) \to 0$ as $\lambda \to \infty$, by dominated convergence since $\left|\frac{t}{t-\lambda}\right| \le \frac{1}{\sin \delta}$. ☐

We are now ready to prove an asymptotic expansion (1.3) of higher order for M_{E}.

Proof of Theorem 1.1. By translation, we may assume that $0 = \min \mathsf{E}$. By precompactness of the family $\{h(x, z)\}_{x\ge1}$, there is a sequence $x_n \to \infty$ for which the limit $h = \lim_{n\to\infty} \frac{1}{x_n} \log|u(x_n, \cdot)|$ is convergent in $\mathcal{D}'(\mathbb{C})$. By Theorem 4.4, h is a positive harmonic function in Ω and $h(z)/\sqrt{-z} \to 1$ as $z \to -\infty$, so by Lemma 2.7, Ω is Greenian, obeys the Akhiezer–Levin condition, and $h \ge M_{\mathsf{E}}$ in Ω. Using (2.9), we obtain for $z \in \Omega$

$$\operatorname{Re} \sqrt{-z} \le M_{\mathsf{E}}(z) \le h(z). \tag{4.8}$$

Hence, the difference $M_{\mathsf{E}}(-k^2) - \operatorname{Re} k$ defines a positive harmonic function in Ω and (4.3), (4.8) imply that $M_{\mathsf{E}}(-k^2) - \operatorname{Re} k = O(|k|^{-1})$. Set $z = \lambda^2$ and $v(\lambda) = M_{\mathsf{E}}(-k^2) - \operatorname{Re} k$. We thus obtain a positive harmonic function in $\mathbb{C}_+$ such that $v(iy) = O(y^{-1})$. By Lemma 4.5 there is a constant c such that

$$v(\lambda) = -\operatorname{Im}\left(\frac{c}{\lambda}\right) + o(|\lambda|^{-1})$$

as $\lambda \to \infty$ nontangentially in $\mathbb{C}_+$. Recalling that $\lambda = ik$, this shows that

$$M_{\mathsf{E}}(-k^2) - \operatorname{Re} k = \operatorname{Re}\left(\frac{c}{k}\right) + o(|k|^{-1}).$$

This completes the proof. □

Proof of Theorem 1.2. Consider a sequence $x_n \to \infty$ such that

$$\lim_{n\to\infty} \frac{1}{x_n} \int_0^{x_n} V(t)\,dt = \liminf_{x\to\infty} \frac{1}{x} \int_0^x V(t)\,dt.$$

Due to Theorem 4.3, this sequence has a subsequence for which the limit $h = \lim_{j\to\infty} \frac{1}{x_{n_j}} \log|u(x_{n_j},\cdot)|$ is convergent in $\mathcal{D}'(\mathbb{C})$. As in the proof of Theorem 1.1, we have $h \geq M_{\mathsf{E}}$ in Ω. Theorem 1.1 and Theorem 4.4 yield

$$a_{\mathsf{E}} = \lim_{k\to+\infty} 2k(M_{\mathsf{E}}(-k^2) - k) \leq \lim_{k\to+\infty} 2k(h(-k^2) - k) = \lim_{j\to\infty} \frac{1}{x_{n_j}} \int_0^{x_{n_j}} V(s)\,ds. \qquad \square$$

Proof of Theorem 1.3. Fix $z_0 \in \mathbb{C} \setminus [\min \mathsf{E}, \infty)$ and consider a sequence $x_n \to \infty$ such that

$$\lim_{n\to\infty} \frac{1}{x_n} \log|u(x_n, z_0)| = \liminf_{x\to\infty} \frac{1}{x} \log|u(x, z_0)|.$$

We can again pass to a subsequence such that $h = \lim_{j\to\infty} \frac{1}{x_{n_j}} \log|u(x_{n_j},\cdot)|$ and $h \geq M_{\mathsf{E}}$ in Ω. In particular,

$$\liminf_{x\to\infty} \frac{1}{x} \log|u(x, z_0)| = h(z_0) \geq M_{\mathsf{E}}(z_0). \qquad \square$$

Proof of Theorem 1.5. By inclusions, we have $(vi) \implies (iv)$ and $(v) \implies (iv)$.

$(iv) \implies (vi)$: Consider any sequence $x_j \to \infty$ such that the limit $h = \lim_{j\to\infty} h(x_j,\cdot)$ converges. The limit h obeys $h \geq M_{\mathsf{E}}$ on $\mathbb{C}_+$ by Theorem 1.3 and obeys $h(z) \leq M_{\mathsf{E}}(z)$ for some $z \in \mathbb{C}_+$. By the maximum principle, $h = M_{\mathsf{E}}$ on $\mathbb{C}_+$, and then on Ω by harmonic continuation. Thus, M_{E} is the only possible subsequential limit of $h(x,\cdot)$ as $x \to \infty$, so by precompactness, $\lim_{x\to\infty} h(x,z) = M_{\mathsf{E}}(z)$ uniformly on compact subsets of $\mathbb{C} \setminus [b_0, \infty)$.

$(vi) \implies (v)$: Given (vi), we know that for any convergent sequence $h(x_n, z)$ the limit is M_{E}. For $z \in [b_0, \infty)$ we have by [6, Theorem 2.7.4.1] that

$$\limsup_{n\to\infty} h(x_n, z) \leq (\limsup_{n\to\infty} h(x_n, z))\check{} = M_{\mathsf{E}}(z),$$

where $\check{f}$ denotes the upper semicontinuous regularization of f. The first inequality follows by the general fact that $f \leq \check{f}$.

$(v) \implies (ii)$: This follows from Theorem 2.1.

$(ii) \implies (iii)$: Due to [36, Corollary 6.4] the set of Dirichlet-irregular points is polar and thus, by [36, Theorem 8.2] it is of harmonic measure zero and the claim follows.

$(iii) \implies (vi)$: Take a sequence $x_n \to \infty$ such that $\lim_{n\to\infty} h(x_n, z) = h(z)$ in $\mathcal{D}'(\mathbb{C})$ and uniformly on compact subsets of $\mathbb{C} \setminus [b_0, \infty)$. Due to the upper envelope theorem [6, Theorem 2.7.4.1], there is a polar set X_1 such that for any $z \in \mathbb{C} \setminus X_1$,

$$\limsup_{n\to\infty} h(x_n, z) = h(z).$$

On the other hand, assuming (iii), there exists X_2 with $\omega_{\mathsf{E}}(X_2, z_0) = 0$, such that for $t \in \mathsf{E}\setminus(X_1 \cup X_2)$ by upper semicontinuity

$$0 \leq \liminf_{\substack{z\to t\\ z\in\Omega}} h(z) \leq \limsup_{\substack{z\to t\\ z\in\Omega}} h(z) \leq h(t) \leq 0.$$

Since $\omega_{\mathsf{E}}(X_1 \cup X_2, z_0) = 0$, Theorem 2.1 gives $h = cM_{\mathsf{E}}$. Comparing the leading order asymptotic behavior at ∞ shows that $c = 1$. Thus, M_{E} is the only possible subsequential limit of $h(x, \cdot)$ as $x \to \infty$, so by precompactness, $\lim_{x\to\infty} h(x, z) = M_{\mathsf{E}}(z)$ uniformly on compact subsets of $\mathbb{C}\setminus[b_0, \infty)$.

$(vi) \implies (i)$: By Theorem 4.4, (vi) implies that $\frac{1}{x_j}\int_0^{x_j} V(t)\mathrm{d}t \to a_{\mathsf{E}}$ for every sequence $x_j \to \infty$, so (i) follows.

$(i) \implies (vi)$: Take a sequence $x_n \to \infty$ such that $h = \lim_{n\to\infty} h(x_n, \cdot)$ converges in $\mathcal{D}'(\mathbb{C})$. Define $v(\lambda) = h(-k^2) - M(-k^2)$. Similarly to the proof of Theorem 1.1, this yields a positive harmonic function in $\mathbb{C}_+$. By Theorem 4.4 and Theorem 1.1, $v(iy) = o\left(y^{-1}\right)$ as $y \to \infty$. By Lemma 4.5, $\lim_{y\to\infty} yv(iy) = 0$ implies that $v \equiv 0$. This shows that M_{E} is the only subsequential limit of $h(x, \cdot)$ as $x \to \infty$. By precompactness, (vi) follows. □

The functions $u(x, z)$ are entire functions of order $\frac{1}{2}$ and as such admit a product representation

$$u(x, z) = u(x, z_*) \prod_{j=1}^{\infty} \left(1 - \frac{z - z_*}{z_j - z_*}\right),$$

where z_j depend on x and z_* is some normalization point. Then the Riesz measure, ρ_x, of the subharmonic function $\log|u(x, z)|$ is a rescaled zero counting measure of $u(x, z)$. That is,

$$\frac{1}{x}\log|u(x, z)| = \frac{1}{x}\log|u(x, z_*)| + \int \log\left|1 - \frac{z - z_*}{t - z_*}\right| \mathrm{d}\rho_x(t),$$

where ρ_x is defined in (1.7).

Proof of Theorem 1.6. By Theorem 1.5 and Theorem 4.3, $h(x, \cdot) \to M_{\mathsf{E}}$ in $\mathcal{D}'(\mathbb{C})$ as $x \to \infty$. By the definition of the Riesz measure, for any $\phi \in C_c^\infty(\mathbb{C})$,

$$\begin{aligned}\lim_{x\to\infty} 2\pi \int \phi(z)\mathrm{d}\rho_x(z) &= \lim_{x\to\infty} \int h(x, z)\Delta\phi(z)\mathrm{d}\lambda(z)\\ &= \int M_{\mathsf{E}}(z)\Delta\phi(z)\mathrm{d}\lambda(z) = 2\pi \int \phi(z)\mathrm{d}\rho_{\mathsf{E}}(z),\end{aligned}$$

where $\mathrm{d}\lambda$ denotes the Lebesgue measure on $\mathbb{C}$. The rest follows from density of $C_c^\infty(\mathbb{C})$ in $C_c(\mathbb{C})$. □

Proposition 4.6. *Let* $\mathrm{d}\mu$ *be the spectral measure of* L_V*, where* V *satisfies* (1.1) *and* $\sigma_{\mathrm{ess}}(L_V) = \mathsf{E}$*. Suppose that along a sequence* $x_n \to \infty$ *the Riesz measure* $\mathrm{d}\rho_{x_n}$ *converge to* ρ_{E} *in the weak-∗ sense. Then, either* $h(x_n, z)$ *converge to* $M_{\mathsf{E}}(z)$ *or there exists a polar Borel set* X *such that* $\mu(\mathbb{R}\setminus X) = 0$*.*

Proof. Assume that $h(x_n, \cdot)$ do not converge to M_{E} and consider a subsequence x_{n_j} such that $h(x_{n_j}, \cdot) \to h$ in $\mathcal{D}'(\mathbb{C})$ with some limit h not equal to M_{E}. By the upper envelope theorem [6, Theorem 2.7.4.1] there is a polar set X_1 such that for any $z \in \mathbb{C}\setminus X_1$,

$$\limsup_{j\to\infty} h(x_{n_j}, z) = h(z).$$

The subharmonic function h has some Riesz measure ρ and by the same arguments as in the proof of Theorem 1.6, $\rho_{x_{n_j}}$ converges to ρ in the weak-$*$ sense. Hence, by uniqueness of the limits our assumption implies that $\rho = \rho_{\mathsf{E}}$ and, by Lemma 2.2 applied to h and M_{E},

$$h(z) = h(z_*) + \int \log\left|1 - \frac{z - z_*}{t - z_*}\right| \mathrm{d}\rho_{\mathsf{E}}(t) = d + M_{\mathsf{E}}(z)$$

where $d = h(z_*) - M_{\mathsf{E}}(z_*)$. Recall that M_{E} has a unique subharmonic extension to $\mathbb{C}$ which vanishes q.e. on E. Therefore, there is a polar set X_2 such that $h(z) = d$ for $z \in \mathsf{E} \setminus X_2$. Moreover, since $M_{\mathsf{E}} \le h$ on Ω we see that $d \ge 0$, and since h is not equal to M_{E}, $d > 0$. In particular,

$$\limsup_{j\to\infty} h(x_{n_j}, z) = d > 0, \qquad \forall z \in \mathsf{E} \setminus (X_1 \cup X_2).$$

However, by Schnol's theorem [74], for μ-a.e. $z \in \mathsf{E}$, the Dirichlet solution decays at most polynomially and, in particular,

$$\limsup_{j\to\infty} h(x_{n_j}, z) \le 0.$$

Thus $\mu(\mathsf{E} \setminus (X_1 \cup X_2)) = 0$, which implies the claim with $X = X_1 \cup X_2$. □

In particular, Theorem 1.7 is now proved.

Proof of Theorem 1.8. By Schnol's theorem [74] for μ-a.e. $z \in \mathsf{E}$

$$\limsup_{x\to\infty} h(x, z) \le 0. \tag{4.9}$$

Hence, by assumption, (4.9) holds $\omega_\Omega(\cdot, z_0)$-a.e.. Therefore, V is regular by Theorem 1.5. □

5. Applications

Proof of Theorem 1.12. (a) Denoting $\mathsf{E} = \sigma_{\mathrm{ess}}(L_V)$, it follows from $\mathsf{E} \subset [0,\infty)$ that M_{E} is a positive harmonic function on $\mathbb{C} \setminus [0,\infty)$. Since the Martin function for the domain $\mathbb{C} \setminus [0,\infty)$ is $\operatorname{Re}\sqrt{-z}$, it follows from Lemma 2.7 that $M_{\mathsf{E}}(z) \ge \operatorname{Re}\sqrt{-z}$. Comparing this with the asymptotic expansion (1.3) as $z \to -\infty$ shows that $a_{\mathsf{E}} \ge 0$ so, by (1.5), $\liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \ge 0$.

(b) As in (a), $a_{\mathsf{E}} \ge 0$. By (1.5) and $\liminf_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \le 0$, this implies that $a_{\mathsf{E}} = 0$. Moreover, $M_{\mathsf{E}}(z) - \operatorname{Re}\sqrt{-z} = o(\sqrt{|z|}^{-1})$ defines a positive harmonic function in $\mathbb{C} \setminus [0,\infty)$ so, by Lemma 2.7, $M_{\mathsf{E}}(z) = \operatorname{Re}\sqrt{-z}$. If E was a proper subset of $[0,\infty)$, since E is closed, there would exist a gap $(a,b) \subset [0,\infty) \setminus \mathsf{E}$, and on this gap M_{E} would be strictly positive, contradicting $M_{\mathsf{E}}(z) = \operatorname{Re}\sqrt{-z}$.

(c) Again by $a_{\mathsf{E}} \ge 0$ and (1.5), $\limsup_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t \le 0$ implies that V is regular. □

We now turn to the construction of a potential which is regular for $\mathsf{E} = [0,\infty)$ but not decaying, even in the Cesàro sense. The potential will be constructed piecewise, so we begin by considering a 2δ-periodic potential defined by

$$W_\delta(x) = \begin{cases} 1 & x \in [0,\delta) \\ -1 & x \in [\delta, 2\delta) \end{cases}$$

Let us compute the discriminant $\Delta_\delta(z)$ and the smallest eigenvalue for the periodic problem,

$$\lambda_\delta = \min\{\lambda \in \mathbb{R} \mid \Delta_\delta(\lambda) = 2\}.$$

Lemma 5.1. $\lim_{\delta\downarrow 0} \lambda_\delta = 0$.

Proof. Since $|W_\delta| \leq 1$ and λ_δ is the minimum of the periodic spectrum, by standard variational principles, $\lambda_\delta \in [-1,1]$ for all $\delta > 0$. The transfer matrix corresponding to W_δ at energy $\lambda \in (-1,1)$ is

$$T_\delta(\lambda) = \begin{pmatrix} \cosh(\delta\sqrt{1-\lambda}) & \frac{\sinh(\delta\sqrt{1-\lambda})}{\sqrt{1-\lambda}} \\ \sqrt{1-\lambda}\sinh(\delta\sqrt{1-\lambda}) & \cosh(\delta\sqrt{1-\lambda}) \end{pmatrix} \begin{pmatrix} \cos(\delta\sqrt{1+\lambda}) & \frac{\sin(\delta\sqrt{1+\lambda})}{\sqrt{1+\lambda}} \\ -\sqrt{1+\lambda}\sin(\delta\sqrt{1+\lambda}) & \cos(\delta\sqrt{1+\lambda}) \end{pmatrix}$$

From this it is elementary to obtain the asymptotic behavior for the discriminant, $\Delta_\delta(\lambda) = \operatorname{tr} T_\delta(\lambda)$, in the form

$$\Delta_\delta(\lambda) = 2 - 4\lambda\delta^2 + O(\delta^3), \qquad \delta \downarrow 0, \tag{5.1}$$

uniformly in $\lambda \in (-1,0)$ (and then, by continuity, for $\lambda \in [-1,0]$). From this, it follows that for any $t < 0$, there exists $\delta_0 > 0$ such that $\delta \in (0,\delta_0)$ and $\lambda \in [-1,t)$ implies $\Delta_\delta(\lambda) > 2$ and therefore $\lambda_\delta \geq t$. It follows that $\liminf_{\delta\downarrow 0} \lambda_\delta \geq 0$.

Meanwhile, $\Delta_\delta(0) = 2\cosh\delta\cos\delta = 2 - \delta^4/3 + o(\delta^4)$ as $\delta \to 0$ implies that $\limsup_{\delta\downarrow 0} \lambda_\delta \leq 0$. □

Proof of Example 1.13. Consider the Dirichlet solution $u(x,t)$ corresponding to the given potential at some $t < 0$. There exists n_0 such that for all $n \geq n_0$, $\lambda_{1/(2n)} > t$. At energies below the periodic spectrum, transfer matrices have strictly positive entries; applying this on intervals $[n, n+1]$ and since products of matrices with positive entries have positive entries, we conclude that $u(x,t)$ has at most one zero with $x > n_0 - 1$. Since zeros of an eigensolution are isolated, it follows that $u(\cdot,t)$ has finitely many zeros, so by Sturm oscillation theory, $\min\sigma_{\mathrm{ess}}(L_V) \geq t$. Since this holds for arbitrary $t < 0$, we conclude $\min\sigma_{\mathrm{ess}}(L_V) \geq 0$.

Conversely, since V obeys $\lim_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t = 0$, the statement is completed by Theorem 1.12. □

Proof of Example 1.11. For $x \in [x_n, x_{n+1}]$ we have $\frac{1}{x}\int_0^x V(t)\mathrm{d}t \leq \int W(t)\mathrm{d}t\frac{n+1}{x_n}$. Since the condition on x_n implies that $\frac{x_n}{n} \to \infty$ we see that $\lim_{x\to\infty} \frac{1}{x}\int_0^x V(t)\mathrm{d}t = 0$. Since $V \geq 0$, we have $\sigma_{\mathrm{ess}}(L_V) \subset \sigma(L_V) \subset [0,\infty)$, so by Theorem 1.12, V is regular and $\sigma_{\mathrm{ess}}(L_V) = [0,\infty)$.

Let H_W be the whole-line operator with the potential $W(x)$. Since $W \geq 0$, we have $\sigma(H_W) \subset [0,\infty)$. Hence, we conclude that $\min\sigma(H_{-W}) < 0$, for otherwise [24, Corollary 1] would imply that $W \equiv 0$. Now by [45, Theorem 7.1] it follows that $\sigma_{\mathrm{ess}}(H_{-V}) = \sigma(H_{-W})$ (where H_{-V} is the full line operator with potential V extended to $\mathbb{R}_-$ by $V \equiv 0$). Since $\sigma_{\mathrm{ess}}(H_{-V}) = \sigma_{\mathrm{ess}}(L_0) \cup \sigma_{\mathrm{ess}}(L_{-V})$ this shows that $\min\sigma_{\mathrm{ess}}(L_{-V}) < 0$. □

Proof of Theorem 1.16. The Lyapunov exponent γ is harmonic in $\mathbb{C}_+ \cup \mathbb{C}_-$ and subharmonic in $\mathbb{C}$. By (1.11) for a.e. $\eta \in S$

$$\lim_{x\to\infty} \frac{1}{x}\log|u_\eta(x,z)| = \gamma(z)$$

converges pointwise in $\mathbb{C}_+ \cup \mathbb{C}_-$; by the weak identity principle for subharmonic functions and precompactness, convergence to γ is also in $\mathcal{D}'(\mathbb{C})$. By Schnol's theorem, for μ_η-a.e. z,

$$\limsup_{x\to\infty} \frac{1}{x}\log|u_\eta(x,z)| \leq 0. \tag{5.2}$$

Fix a sequence $x_n \to \infty$. By the upper envelope theorem [6, Theorem 2.7.4.1] there is a polar set X_η such that for any $z \in \mathbb{C} \setminus X_\eta$,

$$\limsup_{n\to\infty} \frac{1}{x_n}\log|u_\eta(x_n,z)| = \gamma(z).$$

On Q, $\gamma > 0$. Hence, since (5.2) holds for μ_η-a.e. z, we have $\mu_\eta(Q \setminus X_\eta) = 0$. □

6. Conformal maps

In view of Corollary 1.10 and the subsequent discussion, it is of great interest if the harmonic measure of the domain $\mathbb{C}\setminus\mathsf{E}$ is absolutely continuous with respect to the Lebesgue measure $\chi_{\mathsf{E}}(x)\mathrm{d}x$. Let $z_0 < \min E$ and $G_{\mathsf{E}}(z,z_0)$ be the Green function of $\mathbb{C}\setminus\mathsf{E}$ with pole at z_0 and Π_{z_0} the associated comb domain, defined by the upper semicontinuous function s. We say that Π_{z_0} satisfies the sector condition if

$$S_{z_0}(x) = \sup_{y\in(0,\pi)} \frac{s_{z_0}(y)}{|x-y|}$$

is finite for Lebesgue-a.e. $x \in (0,\pi)$. Then, $\omega_{\mathsf{E}}(\cdot, z_0)$ is absolutely continuous with respect to the Lebesgue measure if and only if Π_{z_0} satisfies the sector condition.

The proceeding discussion holds for general semibounded sets E and does not assume that E is an Akhiezer-Levin set. Let M be the Martin function with pole at ∞, normalized at some internal point z_*, ρ its Riesz measure and Π and Θ the corresponding comb and comb mapping. There is a similar characterization for absolute continuity of ρ. Let s be the upper semicontinuous function defining Π. Then ρ is absolutely continuous with respect to $\chi_{\mathsf{E}}(x)\mathrm{d}x$ if and only if the domain contains a Stolz angle at a.e. point at the base of the comb, i.e.

$$S(x) = \limsup_{y\to x} \frac{s(y)}{|x-y|} \tag{6.1}$$

is finite for Lebesgue-a.e. $x \in (0,b)$.

Under various conditions on the set E, it is known that the conformal map $i\Theta'$ has a product representation. We now provide a general proof which does not assume Dirichlet-regularity or any other additional assumptions.

Lemma 6.1. *Let E be a closed non-polar set of the form (1.4). For each j there exists $c_j \in [a_j,b_j]$ such that M is strictly increasing on (a_j,c_j) and strictly decreasing on (c_j,b_j), and $\Theta'(z)$ is given on $z \in \mathbb{C}\setminus[b_0,\infty)$ by*

$$i\Theta'(z) = \frac{C}{\sqrt{b_0 - z}} e^{\int_{[b_0,\infty)\setminus\mathsf{E}} \xi(x) \frac{1+xz}{x-z} \frac{\mathrm{d}x}{1+x^2}} \tag{6.2}$$

where $\xi(x) = 1/2$ for $x \in (a_j,c_j)$, $\xi(x) = -1/2$ for $x \in (c_j,b_j)$, $\xi(x) = 0$ for $x \notin [b_0,\infty)\setminus\mathsf{E}$, and $C > 0$ is a normalization constant.

Proof. For finite-gap sets, this is a reformulation of the Schwarz–Christoffel mapping. If E has infinitely many gaps, we consider them labelled by $j \in \mathbb{N}$ in an arbitrary way and denote $\mathsf{E}_n = [b_0,\infty)\setminus\cup_{j=1}^n(a_j,b_j)$. Denote by M_n the Martin functions at ∞ corresponding to the sets E_n, normalized by $M_n(z_*) = 1$ for some fixed $z_* < b_0$. Since the functions M_n are all positive harmonic on $\mathbb{C}\setminus[b_0,\infty)$, for any $R > |b_0|$, by Harnack's principle they are uniformly bounded on the line segments parametrized by $-R+it$, $t+iR$, $t-iR$, with $t \in [-R,R]$. Since $M_n(x+iy)$ are increasing in $y > 0$ and symmetric, it follows that M_n are uniformly bounded above on the boundary of $(-R,R)\times(-R,R)$ for any R large enough. Since they are also nonnegative, they are a precompact sequence of subharmonic functions on $\mathbb{C}$. By the upper envelope theorem, for any subsequential limit $h = \lim_{k\to\infty} M_{n_k}$, quasi-everywhere on E, $h(z) = \lim_{k\to\infty} M_{n_k}(z) = 0$, so by Theorem 2.1, h is Martin function for the domain $\mathbb{C}\setminus\mathsf{E}$ with $h(z_*) = 1$. It follows that M_n converge to h in $\mathcal{D}'(\mathbb{C})$.

It follows that Θ_n converge to Θ since their real parts converge and their imaginary parts are zero on $(-\infty,b_0)$. In particular, the Herglotz functions $i\Theta_n'$ converge to $ci\Theta'$ uniformly on compact subsets of $\mathbb{C}_+$, so by interpreting this convergence in terms of their exponential Herglotz representations,

$$\lim_{n\to\infty}\int_{\mathbb{R}} g(x)\xi_n(x)\frac{\mathrm{d}x}{1+x^2} = \int_{\mathbb{R}} g(x)\xi(x)\frac{\mathrm{d}x}{1+x^2}, \qquad \forall g\in C(\mathbb{R}\cup\{\infty\})$$

where ξ is determined by $\lim_{y\downarrow 0}\arg\Theta'(x+iy)=\pi\xi(x)$ Lebesgue-a.e. $x\in\mathbb{R}$. By using test functions g supported in (a_j,b_j), it follows that for each j, the critical points $c_{j,n}$ must converge to a point $c_j\in[a_j,b_j]$. Then ξ_n converge pointwise to the function $\tilde\xi$ which is 1 on intervals (a_j,c_j), -1 on (c_j,b_j), and 0 on $[b_0,\infty)$, so by dominated convergence with dominating function $\|g\|_\infty\frac{1}{1+x^2}\chi_{[b_0,\infty)\setminus\mathsf{E}}$,

$$\lim_{n\to\infty}\int_{\mathbb{R}} g(x)\xi_n(x)\frac{\mathrm{d}x}{1+x^2} = \int_{\mathbb{R}} g(x)\tilde\xi(x)\frac{\mathrm{d}x}{1+x^2}, \qquad \forall g\in C(\mathbb{R}\cup\{\infty\}).$$

Of course, this implies $\xi=\tilde\xi$ which implies (6.2). Finally, by separating the contribution from the gap (a_j,b_j) from the remainder of the integral, (6.2) can be extended into the gap (a_j,b_j) to show that $i\Theta'>0$ on (a_j,c_j) and $i\Theta'<0$ on (c_j,b_j). It follows that $M'>0$ on (a_j,c_j) and $M'<0$ on (c_j,b_j), so our construction of c_j as limits of $c_{j,n}$ satisfies the property in the lemma. □

As the final topic of this section, we describe a class of Akhiezer–Levin sets for which it can be seen by purely complex theoretic arguments that the Martin function has the two-term expansion (1.3). While this is not as general as Theorem 1.1, within its scope of applicability, it provides a formula for a_E in terms of critical points of the Martin function.

Lemma 6.2. *Let $\mathsf{E}\subset\mathbb{R}$ be of the form* (1.4)*. If $\sum_{j=1}^N (b_j-a_j)<\infty$, then E is an Akhiezer–Levin set, the Martin function obeys the two-term expansion* (1.3)*, and*

$$a_\mathsf{E} = b_0 + \sum_{j=1}^N (a_j+b_j-2c_j). \tag{6.3}$$

Proof. Finite gap length can be restated as $\int\chi_{[b_0,\infty)\setminus\mathsf{E}}(x)\mathrm{d}x<\infty$ and it implies that the exponent in (6.2) can be split into two separately integrable integrands, of which one is z-independent, to give

$$i\Theta_\mathsf{E}'(z) = \frac{C_\mathsf{E}}{\sqrt{b_0-z}}e^{\int_{[b_0,\infty)\setminus\mathsf{E}}\xi(x)\frac{1}{x-z}\mathrm{d}x}.$$

For any $\delta>0$, using finite gap length and dominated convergence,

$$\int_{[b_0,\infty)\setminus\mathsf{E}}\xi(x)\frac{1}{x-z}\mathrm{d}x = -\frac{1}{z}\int_{[b_0,\infty)\setminus\mathsf{E}}\xi(x)\mathrm{d}x + o(|z|^{-1}),$$

as $z\to\infty$, $\arg z\in[\delta,2\pi-\delta]$. Evaluating the integral $\int_{[b_0,\infty)\setminus\mathsf{E}}\xi(x)\mathrm{d}x$ and substituting into $\Theta'(z)$,

$$i\Theta_\mathsf{E}'(z) = C_\mathsf{E}\left(\frac{1}{\sqrt{-z}}+\frac{1}{2}(b_0+\sum_{j=1}^N(a_j+b_j-2c_j))\frac{1}{\sqrt{-z}^3}+o(|z|^{-3/2})\right)$$

and integrating along rays shows that, as $z\to\infty$ with $\arg z\in[\delta,2\pi-\delta]$,

$$i\Theta_\mathsf{E}(z) = C_\mathsf{E}\left(-2\sqrt{-z}+(b_0+\sum_{j=1}^N(a_j+b_j-2c_j))\frac{1}{\sqrt{-z}}+o(|z|^{-1/2})\right).$$

Taking imaginary parts gives a two-term expansion of M_E, which matches (1.3) with $C_\mathsf{E}=\frac{1}{2}$. Reading off the second term gives (6.3). □

References

1. N. I. Akhiezer and B. Ya. Levin, *Generalizations of the Bernstein inequalities for derivatives of entire functions*, Studies in the Modern Problems of the Theory of Functions of a Complex Variable (Moscow) (1960), 111–165.
2. A. Ancona, *Une propriété de la compactification de Martin d'un domaine euclidien*, Ann. Inst. Fourier (Grenoble) **29** (1979), no. 4, ix, 71–90. MR 558589
3. D. H. Armitage and S. J. Gardiner, *Classical potential theory*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2001. MR 1801253
4. F. V. Atkinson, *On the location of the Weyl circles*, Proc. Roy. Soc. Edinburgh Sect. A **88** (1981), no. 3-4, 345–356. MR 616784
5. A. Avila, *Global theory of one-frequency Schrödinger operators*, Acta Math. 215 (2015), 1–54.
6. V. Azarin, *Growth theory of subharmonic functions*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 2009. MR 2463743
7. M. Benedicks, *Positive harmonic functions vanishing on the boundary of certain domains in* $\mathbf{R}^n$, Ark. Mat. **18** (1980), no. 1, 53–72. MR 608327
8. I. Binder, D. Damanik, M. Goldstein, M. Lukic, *Almost Periodicity in Time of Solutions of the KdV Equation*, Duke Math. J. 167 (2018), 2633–2678.
9. O. Blumenthal, *Über die Entwicklung einer willkürlichen Funktion nach den Nennern des Kettenruches für* $\int_{-\infty}^{0} \frac{\phi(\xi)d\xi}{z-\xi}$, Ph.D. thesis, Göttingen, 1898.
10. L. Carleson and V. Totik, *Hölder continuity of Green's functions*, Acta Sci. Math. (Szeged) **70** (2004), no. 3-4, 557–608. MR 2107529
11. R. Carmona, J. Lacroix, *Spectral theory of random Schrödinger operators*, Probab. Appl. Birkhäuser Boston, Inc., Boston, MA, 1990. xxvi+587 pp.
12. J. S. Christiansen, *Szegő's theorem on Parreau–Widom sets*, Adv. Math. 229 (2012), no. 2, 1180–1204.
13. V. Chulaevskii, *Perturbations of a Schrödinger operator with periodic potential*, Uspekhi Mat. Nauk **36** (1981), 203–204.
14. W. Craig, *The trace formula for Schrödinger operators on the line*, Comm. Math. Phys. **126** (1989), no. 2, 379–407. MR 1027503
15. H. L. Cycon, R. G. Froese, W. Kirsch, B, Simon, *Schrödinger operators with application to quantum mechanics and global geometry.* Texts Monogr. Phys. Springer Study Ed. Springer-Verlag, Berlin, 1987. x+319 pp.
16. D. Damanik, *Schrödinger operators with dynamically defined potentials*, Ergodic Theory and Dynamical Systems, 37(6), 1681–1764.
17. D. Damanik, J. Fillman, and M. Lukic, *Limit-periodic continuum Schrödinger operators with zero measure Cantor spectrum*, J. Spectr. Theory **7** (2017), no. 4, 1101–1118. MR 3737889
18. D. Damanik, J. Fillman, and A. Gorodetski, *Multidimensional almost-periodic Schrödinger operators with Cantor spectrum*, Ann. Henri Poincaré **20** (2019), no. 4, 1393–1402. MR 3928381
19. D. Damanik, M. Goldstein, *On the inverse spectral problem for the quasi-periodic Schrödinger equation*, Publ. Math. Inst. Hautes Études Sci. **119** (2014), 217–401.
20. D. Damanik, M. Goldstein, *On the existence and uniqueness of global solutions of the KdV equation with quasi-periodic initial data*, J. Amer. Math. Soc., to appear (arXiv:1212.2674).
21. D. Damanik, M. Goldstein, M. Lukic, *A Multi-Scale Analysis Scheme on Abelian Groups with an Application to Operators Dual to Hill's Equation*, Trans. Amer. Math. Soc. 369 (2017),

1689–1755.

22. D. Damanik, M. Goldstein, and M. Lukic, *The isospectral torus of quasi-periodic Schrödinger operators via periodic approximations*, Invent. Math. **207** (2017), no. 2, 895–980. MR 3595938
23. D. Damanik, M. Goldstein, M. Lukic, *The Spectrum of a Schrödinger Operator With Small Quasi-Periodic Potential is Homogeneous*, J. Spectr. Theory 6 (2016), 415–427.
24. D. Damanik, R. Killip, and B. Simon, *Schrödinger operators with few bound states*, Comm. Math. Phys. **258** (2005), no. 3, 741–750. MR 2172016
25. D. Damanik and C. Remling, *Schrödinger operators with many bound states*, Duke Math. J. **136** (2007), no. 1, 51–80. MR 2271295
26. D. Damanik, R. Sims, and G. Stolz, *Lyapunov exponents in continuum Bernoulli-Anderson models*, Operator methods in ordinary and partial differential equations (Stockholm, 2000), Oper. Theory Adv. Appl., vol. 132, Birkhäuser, Basel, 2002, pp. 121–130. MR 1924975
27. D. Damanik and P. Yuditskii, *Counterexamples to the Kotani-Last conjecture for continuum Schrödinger operators via character-automorphic Hardy spaces*, Adv. Math. **293** (2016), 738–781. MR 3474334
28. I. E. Egorova, *Almost periodicity of some solutions of the KdV equation with Cantor spectrum*, Dopov./Dokl. Akad. Nauk Ukraini, (1993), 26–29.
29. I. E. Egorova, *The Cauchy problem for the KdV equation with almost periodic initial data whose spectrum is nowhere dense*, Adv. Soviet Math., **19** (1994), 181–208.
30. B. Eichinger, T. VandenBoom, and P. Yuditskii, *KdV hierarchy via abelian coverings and operator identities*, Trans. Amer. Math. Soc. Ser. B **6** (2019), 1–44. MR 3894927
31. H. Eliasson, *Floquet solutions for the* 1*–dimensional quasiperiodic Schrödinger equation.* Commun. Math. Phys. **146** (1992), 447–482.
32. A. Eremenko and P. Yuditskii, *Comb functions*, Recent advances in orthogonal polynomials, special functions, and their applications, Contemp. Math., vol. 578, Amer. Math. Soc., Providence, RI, 2012, pp. 99–118. MR 2964141
33. G. Faber, *Über Tschebyscheffsche Polynome*, J. Reine Angew. Math. **150** (1920), 79–106. MR 1580974
34. J. Fillman, M. Lukic, *Spectral Homogeneity of Limit-Periodic Schrödinger Operators*, J. Spectr. Theory 7 (2017), 387–406.
35. S. J. Gardiner and T. Sjödin, *Potential theory in Denjoy domains*, Analysis and mathematical physics, Trends Math., Birkhäuser, Basel, 2009, pp. 143–166. MR 2724611
36. J. B. Garnett and D. E. Marshall, *Harmonic measure*, New Mathematical Monographs, vol. 2, Cambridge University Press, Cambridge, 2005. MR 2150803
37. F. Gesztesy and P. Yuditskii, *Spectral properties of a class of reflectionless Schrödinger operators*, J. Funct. Anal. **241** (2006), no. 2, 486–527. MR 2271928
38. K. Hirata, *Martin boundary points of cones generated by spherical John regions*, Ann. Acad. Sci. Fenn. Math. **32** (2007), no. 2, 289–300. MR 2337478
39. L. Hörmander, *The analysis of linear partial differential operators. I*, Classics in Mathematics, Springer-Verlag, Berlin, 2003, Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)]. MR 1996773
40. C. Marx, S. Jitomirskaya, *Dynamics and spectral theory of quasi-periodic Schrödinger-type operators*, Ergodic Theory and Dynamical Systems, 37(8), 2353–2393.
41. R. Johnson and J. Moser, *The rotation number for almost periodic potentials*, Comm. Math. Phys. **84** (1982), no. 3, 403–438. MR 667409

42. W. Kirsch, *On a class of random Schrödinger operators*, Adv. in Appl. Math. **6** (1985), no. 2, 177–187. MR 789852
43. P. Koosis, *The logarithmic integral. I*, Cambridge Studies in Advanced Mathematics, vol. 12, Cambridge University Press, Cambridge, 1998, Corrected reprint of the 1988 original. MR 1670244
44. S. Kotani, *Generalized Floquet theory for stationary Schrödinger operators in one dimension*, Chaos Solitons Fractals **8** (1997), no. 11, 1817–1854. MR 1477262
45. Y. Last and B. Simon, *The essential spectrum of Schrödinger, Jacobi, and CMV operators*, J. Anal. Math. **98** (2006), 183–220. MR 2254485
46. ———, *Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*, Invent. Math. **135** (1999), no. 2, 329–367. MR 1666767
47. B. Ya. Levin, *Distribution of zeros of entire functions*, revised ed., Translations of Mathematical Monographs, vol. 5, American Mathematical Society, Providence, R.I., 1980, Translated from the Russian by R. P. Boas, J. M. Danskin, F. M. Goodspeed, J. Korevaar, A. L. Shields and H. P. Thielman. MR 589888 (81k:30011)
48. ———, *Majorants in classes of subharmonic functions*, Teor. Funktsii Funktsional. Anal. i Prilozhen. (1989), no. 51, 3–17. MR 1009141
49. ———, *The connection of a majorant with a conformal mapping. II*, Teor. Funktsii Funktsional. Anal. i Prilozhen. (1989), no. 52, 3–21. MR 1018326
50. ———, *Classification of closed sets on* $\mathbf{R}$ *and representation of a majorant. III*, Teor. Funktsii Funktsional. Anal. i Prilozhen. (1989), no. 52, 21–33. MR 1018327
51. D. S. Lubinsky, *A new approach to universality limits involving orthogonal polynomials*, Ann. of Math. (2) 170 (2009), no. 2, 915–939. MR 2552113
52. M. Lukic, *Derivatives of* L^p *eigenfunctions of Schrödinger operators*, Math. Model. Nat. Phenom. **8** (2013), no. 1, 170–174. MR 3022987
53. A. Maltsev, *Universality limits of a reproducing kernel for a half-line Schrődinger operator and clock behavior of eigenvalues*, Comm. Math. Phys. 298 (2010), no. 2, 461–484.
54. R. S. Martin, *Minimal positive harmonic functions*, Trans. Amer. Math. Soc. **49** (1941), 137–172. MR 3919
55. V. A. Marčenko and I. V. Ostrovskii, *A characterization of the spectrum of the Hill operator*, Mat. Sb. (N.S.) **97(139)** (1975), no. 4(8), 540–606, 633–634. MR 0409965
56. A. Mate, P. Nevai, and V. Totik, *Szegő's extremum problem on the unit circle*, Ann. of Math. (2) **134** (1991), no. 2, 433–453. MR 1127481
57. L. Pastur, A. Figotin, *Spectra of random and almost-periodic operators.* Grundlehren Math. Wiss., 297[Fundamental Principles of Mathematical Sciences] Springer-Verlag, Berlin, 1992. viii+587 pp.
58. L. Pastur, V. A. Tkachenko, *On the spectral theory of the one-dimensional Schrödinger operator with limit-periodic potential* (Russian), Dokl. Akad. Nauk SSSR **279** (1984) 1050–1053.
59. L. Pastur, V. A. Tkachenko, *Spectral theory of a class of one-dimensional Schrödinger operators with limit-periodic potentials* (Russian), Trudy Moskov. Mat. Obshch. **51** (1988) 114–168.
60. D. B. Pearson, *Singular continuous measures in scattering theory*, Comm. Math. Phys. **60** (1978), no. 1, 13–36. MR 484145
61. J. Pöschel and E. Trubowitz, *Inverse spectral theory*, Pure and Applied Mathematics, vol. 130, Academic Press, Inc., Boston, MA, 1987. MR 894477
62. T. Ransford, *Potential theory in the complex plane*, London Mathematical Society Student Texts, vol. 28, Cambridge University Press, Cambridge, 1995. MR 1334766

63. C. Remling, *The absolutely continuous spectrum of one-dimensional Schrödinger operators*, Math. Phys. Anal. Geom. **10** (2007), no. 4, 359–373. MR 2386257
64. B. Simon, *Sturm oscillation and comparison theorems*, Sturm-Liouville theory, Birkhäuser, Basel, 2005, pp. 29–43. MR 2145076
65. ______, *Equilibrium measures and capacities in spectral theory*, Inverse Probl. Imaging **1** (2007), no. 4, 713–772. MR 2350223
66. B. Simon, *Two extensions of Lubinsky's universality theorem*, J. Anal. Math. 105 (2008), 345–362.
67. ______, *Regularity and the Cesàro-Nevai class*, J. Approx. Theory **156** (2009), no. 2, 142–153. MR 2494549
68. M. Sodin and P. Yuditskii, *Almost periodic Sturm-Liouville operators with Cantor homogeneous spectrum*, Comment. Math. Helv. **70** (1995), no. 4, 639–658. MR 1360607
69. H. Stahl and V. Totik, *General orthogonal polynomials*, Encyclopedia of Mathematics and its Applications, vol. 43, Cambridge University Press, Cambridge, 1992. MR 1163828
70. G. Szegő, *Bemerkungen zu einer Arbeit von Herrn M. Fekete: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten*, Math. Z. **21** (1924), no. 1, 203–208. MR 1544698
71. J. L. Ullman, *On the regular behaviour of orthogonal polynomials*, Proc. London Math. Soc. (3) **24** (1972), 119–148. MR 0291718
72. A. Volberg and P. Yuditskii, *Mean type of functions of bounded characteristic and Martin functions in Denjoy domains*, Adv. Math. **290** (2016), 860–887. MR 3451940
73. P. Yuditskii, *Direct Cauchy theorem and Fourier integral in Widom domains*, J. Anal. Math.141(2020), no.1, 411–439.
74. I. É. Šnol', *On the behavior of eigenfunctions*, Doklady Akad. Nauk SSSR (N.S.) **94** (1954), 389–392. MR 0062902
75. H. Weyl, *Über beschränkte quadratische Formen, deren Differenz vollstetig ist*, Rend. Circ. Mat. Palermo **27** (1909), 373–392.
76. H. Widom, *Polynomials associated with measures in the complex plane*, J. Math. Mech. **16** (1967), 997–1013. MR 0209448

Institute for Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040, Vienna, Austria and Department of Mathematics and Statistics, Flyde College, Lancaster University, Lancaster LA1 4YF, U.K.

Email address: `benjamin.eichinger@tuwien.ac.at`

Department of Mathematics, Rice University MS-136, Box 1892, Houston, TX 77251-1892, USA.

Email address: `milivoje.lukic@rice.edu`

ASYMPTOTICS FOR CHRISTOFFEL FUNCTIONS ASSOCIATED TO CONTINUUM SCHRÖDINGER OPERATORS

By

BENJAMIN EICHINGER*

Abstract. We prove asymptotics of the Christoffel function, $\lambda_L(\xi)$, of a continuum Schrödinger operator for points in the interior of the essential spectrum under some mild conditions on the spectral measure. It is shown that $L\lambda_L(\xi)$ has a limit and that this limit is given by the Radon–Nikodym derivative of the spectral measure with respect to the Martin measure. Combining this with a recently developed local criterion for universality limits at scale $\lambda_L(\xi)$, we compute universality limits for continuum Schrödinger operators at scale L and obtain clock spacing of the eigenvalues of the finite range truncations.

1 Introduction

The goal of this paper is to derive asymptotics for Christoffel functions of continuum Schrödinger operators. It is natural for this topic to work in the half-line setting, so our Schrödinger operators are unbounded self-adjoint operators, H_V, on $L^2((0,\infty))$, corresponding formally to the differential expression

$$-\frac{d^2}{dx^2} + V.$$

We always assume that the potential V is real-valued and uniformly locally integrable, i.e.,

$$\sup_{x\geq 0}\int_x^{x+1} |V(t)|\mathrm{d}t < \infty. \tag{1.1}$$

In particular, 0 is a regular endpoint and $+\infty$ is a limit point endpoint in the sense of Weyl. We set a Neumann boundary condition at 0, so the domain of H_V is

$$D(H_V) = \left\{ f \in L^2((0,\infty)) \mid f, f' \in \mathrm{AC}_{\mathrm{loc}}([0,\infty)),\ -f'' + Vf \in L^2((0,\infty)),\ f'(0) = 0 \right\}$$

*This work was supported by the Austrian Science Fund FWF, project no: P33885

JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. TBD (2023)
DOI 10.1007/s11854-023-0319-7

where $AC_{loc}([0,\infty))$ denotes the set of functions which are absolutely continuous on all bounded intervals.

For any $z \in \mathbb{C}$ the Neumann solution, $v(x,z)$, is the solution of the initial value problem

$$-\partial_x^2 v(x,z) + V(x)v(x,z) = zv(x,z), \quad v(0,z) = 1, \ \partial_x v(0,z) = 0. \tag{1.2}$$

The Christoffel function is defined by

$$\lambda_L(z) = \left(\int_0^L |v(x,z)|^2 dx \right)^{-1}, \quad z \in \mathbb{C}, L \geq 0. \tag{1.3}$$

As a function of L, it measures the growth rate of eigensolutions, which is known to be an important quantity in spectral theory. For instance, growth rates of eigensolutions are used in subordinacy theory developed by Gilbert and Pearson [20], or by Last and Simon [24], to describe the absolutely continuous spectrum of H_V. In our main result, Theorem 1.1 below, we will prove asymptotics for $\lambda_L(\xi)$ as $L \to \infty$ and as a consequence obtain universality limits for Christoffel–Darboux kernels of continuum Schrödinger operators and clock spacing of the eigenvalues of finite range truncations of H_V. Asymptotics of λ_L as well as universality limits and zero spacing of eigenvalues has received much attention in recent years; see [4, 6, 9, 13, 22, 25, 27, 26, 31, 28, 32] for a partial list of references.

In order to formulate Theorem 1.1 we need to recall the construction of a maximal spectral measure using Weyl theory and the concept of the Martin function from potential theory.

Since ∞ is a limit point endpoint, there is (up to a scalar multiple unique) $\psi(x,z)$ satisfying

$$-\partial_x^2 \psi(x,z) + V(x)\psi(x,z) = z\psi(x,z) \tag{1.4}$$

and $\psi \in L^2((0,\infty))$, which is called the Weyl solution at ∞. On the upper half-plane $\mathbb{C}_+$, the Weyl m-function is defined by

$$m(z) = -\frac{\psi(0,z)}{\partial_x \psi(0,z)}. \tag{1.5}$$

The function m is a Herglotz function, i.e., it maps $\mathbb{C}_+$ analytically into itself. It is a general fact that Herglotz functions admit an integral representation. That is, there exist $a \geq 0$, $b \in \mathbb{R}$ and a positive Borel measure μ with

$$\int \frac{d\mu(\xi)}{1+\xi^2} < \infty$$

such that

$$m(z) = az + b + \int_{\mathbb{R}} \Big(\frac{1}{\xi - z} - \frac{\xi}{1+\xi^2} \Big) d\mu(\xi).$$

From the perspective of operator theory, the measure μ represents a maximal spectral measure of H_V. Let $d\mu = f_\mu dx + d\mu_s$ be the Lebesgue decomposition of μ with respect to the Lebesgue measure.

In [12] a theory of Stahl–Totik regularity was developed for continuum Schrödinger operators. To introduce this theory we will use some standard objects from potential theory which can be found in [34, 2]. Let $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$ and $\Omega = \mathbb{C} \setminus \mathsf{E}$. For potentials V satisfying (1.1), E is bounded from below but not from above. Therefore, one can show that the cone of positive harmonic functions in Ω which vanish quasi-everywhere (q.e.) on E is one-dimensional. Elements of this cone are called Martin functions of Ω at ∞. For an excellent survey on the Martin theory for Denjoy domains we refer to [17]. In [12] it is shown that for any such Martin function

$$\lim_{z \to -\infty} \frac{M(z)}{\sqrt{-z}} > 0.$$

Existence of the limit follows by standard arguments for positive harmonic functions. The important point of the above statement is that the limit is positive. It allows to normalize at ∞ and we obtain a unique Martin function, $M_{\mathsf{E}}(z)$, such that the limit above is equal to 1. Due to [12, Theorem 1.1] there exists $a_{\mathsf{E}} \in \mathbb{R}$ such that the Martin function has the asymptotic behavior

$$M_{\mathsf{E}}(z) = \mathrm{Re}\Big(\sqrt{-z} + \frac{a_{\mathsf{E}}}{2\sqrt{-z}} \Big) + o\Big(\frac{1}{\sqrt{|z|}} \Big),$$

as $z \to \infty$, $\arg z \in [\delta, 2\pi - \delta]$, for any $\delta > 0$. This higher asymptotic expansion is then used to characterize regularity in the sense of Stahl and Totik. A potential V satisfying (1.1) is called Stahl–Totik regular if

$$\lim_{L \to \infty} \frac{1}{L} \int_0^L V(s) ds = a_{\mathsf{E}}. \tag{1.6}$$

The Martin function can be extended to a subharmonic function on $\mathbb{C}$ and thus its distributional Laplacian defines a positive measure. We call

$$\rho_{\mathsf{E}} = \frac{1}{2\pi} \Delta M_{\mathsf{E}}$$

the Martin measure of the domain $\mathbb{C} \setminus \mathsf{E}$. It plays the role of the equilibrium measure from the theory of orthogonal polynomials for compactly supported measures. Again we write its Lebesgue decomposition $d\rho_{\mathsf{E}}(\xi) = f_{\mathsf{E}}(\xi) d\xi + d\rho_{\mathsf{E},s}(\xi)$.

Assuming regularity, we are able to characterize the asymptotic behavior of λ_L at interior points of $\sigma_{\mathrm{ess}}(H_V)$.

Theorem 1.1. *Let V be a Stahl–Totik regular potential such that $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$ is Dirichlet regular and μ the corresponding spectral measure. Let $I \subset \mathrm{int}(\mathsf{E})$ be a closed interval such that μ is absolutely continuous in a neighborhood of I and its density f_μ is positive and continuous at every point of I. Then we have*

$$\lim_{L\to\infty} L\lambda_L(\xi) = \frac{f_\mu(\xi)}{f_{\mathsf{E}}(\xi)}, \tag{1.7}$$

uniformly for $\xi \in I$.

Stahl–Totik regularity is a quite general property. For sufficiently nice sets E, regularity follows from

$$f_\mu > 0 \quad \text{Lebesgue a.e. on } \mathsf{E}. \tag{1.8}$$

To be precise, by the Widom criterion [12, Theorem 1.8] V is regular if the harmonic measure of $\mathbb{C} \setminus \mathsf{E}$ is absolutely continuous with respect to μ. Thus, if E is such that the harmonic measure is mutually absolutely continuous with the Lebesgue measure restricted to E, (1.8) implies Stahl–Totik regularity. By [42], mutual absolute continuity of the harmonic and the Lebesgue measure holds for regular Parreau–Widom sets. These sets are well studied in inverse spectral theory [41, 7, 14]. Every set which is homogeneous in the sense of Carleson, i.e., sets E for which there exists $\tau > 0$ so that

$$|\mathsf{E} \cap [\xi_0 - \varepsilon, \xi_0 + \varepsilon]| \geq \tau\varepsilon, \quad \forall \xi_0 \in \mathsf{E}, \forall \varepsilon \in (0, 1],$$

is a regular Parreau–Widom set [23]. In particular, for finite gap sets or for spectra of Schrödinger operators with periodic potentials (1.8) implies that V is Stahl–Totik regular.

Christoffel functions are well studied in the setting of orthogonal polynomials. In this case, the Christoffel function is defined similar to (1.3), but the Neumann solution at L is substituted by the orthonormal polynomial of degree n. For compactly supported measures, typical results show that under certain assumptions

$$\lim_{n\to\infty} n\lambda_n(\xi) = \frac{f_\mu(\xi)}{f_{\mathsf{E}}(\xi)}, \tag{1.9}$$

where $\lambda_n(\xi)$ is the Christoffel function associated to the orthonormal polynomials and $f_{\mathsf{E}}(\xi)$ denotes the density of the equilibrium measure. A fundamental result of Máté–Nevai–Totik [31] shows (1.9) for the case $\mathsf{E} = [-2, 2]$. More precisely, it is shown that (1.9) holds provided that μ is Stahl–Totik regular on $[-2, 2]$, $f_\mu(\xi) > 0$, $\log f_\mu$ is integrable in a neighborhood of ξ, and ξ is a Lebesgue point of both the measure μ and the Szegő function associated to f_μ. This has been extended by Totik to arbitrary compact sets by using the polynomial preimage method [43].

Our approach is inspired by a method used by Simon and we obtain a full analog for continuum Schrödinger operators of all results in [40]. Let us mention that the assumptions in [40] or in Theorem 1.1 are stronger than the ones in [31]. However, the conclusion is also stronger, since uniformity in (1.7) require continuity of f_μ, see also [44]. It is an interesting question if our method could also be used to prove (1.7) under Lebesgue point and local Szegő conditions as used by Máté–Nevai–Totik.

Limits of Christoffel functions for continuum Schrödinger operators were first studied by Maltsev in [28]. At that time the notion of regularity for continuum Schrödinger operators was not available and Maltsev proved (1.7) for potentials $V = \mathring{V} + \tilde{V}$, where $\mathring{V}$ is a periodic continuous potential, $\tilde{V}$ is so that $\sigma_{\mathrm{ess}}(V) = \sigma_{\mathrm{ess}}(\mathring{V})$ and $\tilde{V}$ is Césaro decaying, i.e.,

$$\lim_{L\to\infty} \frac{1}{L} \int_0^L |\tilde{V}(x)| dx = 0. \tag{1.10}$$

Thus, our result generalizes [28] in several directions. First of all, if E is the spectrum of a continuum Schrödinger operator, then generically there is no periodic potential so that the essential spectrum of the associated operator is E. Moreover, even if E is the spectrum of a periodic Schrödinger operator, a regular potential does not necessarily satisfy (1.10). A counterexample can be found even in the simplest case $\mathsf{E} = [0, \infty)$ with $\mathring{V} \equiv 0$. It is shown in [12, Example 1.13] that the potential defined piecewise by $V(x) = (-1)^{\lfloor 2n(x-n) \rfloor}$ on $x \in [n-1, n)$ for integer n is regular with $\sigma_{\mathrm{ess}}(L_V) = [0, \infty)$, but (1.10) does not hold. On the other hand, since periodic potentials are regular, it follows from (1.6) and (1.10) that the potentials considered in [28] are Stahl–Totik regular.

We turn to applications of our main theorem. The Christoffel–Darboux kernel is defined by

$$K_L(z, w) = \int_0^L v(x, z)\overline{v(x, w)} dx.$$

For $\xi \in \mathbb{R}$, bulk universality limits are double scaling limits of the type

$$\lim_{L\to\infty} \frac{K_L(\xi + \frac{z}{\tau_L(\xi)}, \xi + \frac{w}{\tau_L(\xi)})}{K_L(\xi, \xi)} = \frac{\sin(\eta\pi(z - \overline{w}))}{\eta\pi(z - \overline{w})}.$$

Universality limits were often studied at explicit polynomial scales. Recent results suggest that universality limits at such scales are a combination of two different phenomena. One is universality at scale $\lambda_L(\xi)$, studied in great generality in [13]; the other is the explicit asymptotics of $\lambda_L(\xi)$ now provided by Theorem 1.1. Thus, we can combine Theorem 1.1 with the results from [13] to obtain universality limits at scale L.

Theorem 1.2. *With the assumptions of Theorem 1.1 we have*

$$\lim_{L\to\infty} \frac{K_L(\xi + \frac{z}{L}, \xi + \frac{w}{L})}{K_L(\xi, \xi)} = \frac{\sin(\pi f_{\mathsf{E}}(\xi)(z - \overline{w}))}{\pi f_{\mathsf{E}}(\xi)(z - \overline{w})}, \tag{1.11}$$

uniformly for $\xi \in I$.

As a consequence of Theorem 1.2 we will obtain asymptotic equal eigenvalue spacing of the eigenvalues of the finite range truncations. It is a common scheme already observed by Wigner for random matrix ensembles that the global asymptotic distribution of the eigenvalues depends on the particular model, however the local microscopic scale exhibits universal behavior.

For any $L > 0$, let ν_L denote the zero counting measure for $\partial_L v(L, \cdot)$ divided by L,

$$\nu_L = \frac{1}{L} \sum_{\xi:\ \partial_L v(L,\xi)=0} \delta_\xi.$$

The measure ν_L is intimately related to the eigenvalues of the finite range truncation of H_V. Namely, if H_V^L denotes the restriction of H_V onto $(0, L)$ with Neumann boundary condition at L, then H_V^L has purely discrete spectrum given by the zeros of $\partial_L v(L, \cdot)$.

The global distribution is given by the Martin measure ρ_{E}. That is, regularity of V implies that ν_L has a weak-$*$ limit and that this limit is given by ρ_{E}. However, it follows from the Freud–Levin theorem [16, 25] that (1.11) implies equal eigenvalue spacing at scale $1/L$.

For $L > 0$ and $\xi \in I$, we denote by $\xi_j^L(\xi)$ for $j \in \mathbb{Z}$ the zeros of $\partial_L v(L, \cdot)$ counted from ξ, i.e.,

$$\cdots < \xi_{-2}^L(\xi) < \xi_{-1}^L(\xi) < \xi \le \xi_0^L(\xi) < \xi_1^L(\xi) < \cdots$$

with no zeros of $\partial_L v(L, \cdot)$ between ξ_j^L and ξ_{j+1}^L.

Theorem 1.3. *With the assumptions of Theorem 1.1 we have that the zeros of* $\partial_L v(L, \cdot)$ *admit uniform clock behavior on* I, *i.e., for fixed* $j \in \mathbb{Z}$

$$\lim_{L\to\infty} L f_{\mathsf{E}}(\xi)(\xi_{j+1}^L(\xi) - \xi_j^L(\xi)) = 1 \tag{1.12}$$

uniformly for $\xi \in I$.

The organization of the paper is as follows. In Section 2 we recall concepts from the theory of Stahl–Totik regularity for continuum Schrödinger operators and prove that the additional assumption of Dirichlet regularity of E leads to uniformity in the asymptotic estimates. This is crucial to control the exponential growth of the

Neumann solution close to E. In Section 3 we recall aspects of the spectral theory for continuum Schrödinger operators and show how this can be viewed as a special case of the general theory of canonical systems. In particular, we show that the Christoffel function can also be defined through an extremal problem. In Section 4 we prove asymptotics of the Christoffel function for finite gap potentials. Section 5 is devoted to the proofs of the main theorems. We also provide an appendix in which we recall some parts from the theory of canonical systems.

Acknowledgements. I would like to thank Milivoje Lukić and Brian Simanek for helpful discussions.

2 Stahl–Totik regularity

From the discrete setting it is known that in order to show (1.9), in addition to local properties of the measure, some additional global assumption is needed [43]. A common sufficient assumption is to assume that the measure is regular in the sense of Stahl and Totik. In [12] a corresponding theory was developed for continuum Schrödinger operators with uniformly locally integrable potential. In the following, we show that an additional uniformity is obtained in the estimates under the additional assumption that the underlying spectrum is regular for the Dirichlet problem.

For $z \in \mathbb{C}$ the Dirichlet solution, $u(x, z)$, is the solution of (1.2) with initial condition $u(0, z) = 0$, $\partial_x u(0, z) = 1$. Stahl–Totik regularity as defined in the introduction was linked to exponential growth of the Dirichlet solutions. However, the same proofs also characterize the exponential growth of the Neumann solution. By $\mathcal{D}'(\mathbb{C})$ we denote the space of distributions and recall that subharmonic functions can be viewed as representatives of real-valued distributions with non-negative Laplacian.

Lemma 2.1. *Let V obey* (1.1) *and $v(x, z)$ denote the Neumann solution. Then:*

(a) *For any $x > 0$, the function $\log|v(x, z)|$ is a subharmonic function on $\mathbb{C}$.*

(b) *The family of functions $\{\log|v(x, z)|\}_{x\geq 1}$ is precompact in $\mathcal{D}'(\mathbb{C})$.*

Moreover, if V is Stahl–Totik regular, then:

(i) *The functions $\log|v(x, z)|$ converge as $x \to \infty$ to $M_{\mathsf{E}}(z)$ in the distributional sense as well as uniformly on compact subsets of $\mathbb{C}_+$.*

(ii) *For all $z \in \mathbb{C}$, $\limsup_{x\to\infty} \frac{1}{x} \log |v(x, z)| \leq M_{\mathsf{E}}(z)$.*

Proof. (a) By general principles, for any $x > 0$, $v(x, z)$ is an entire function of z [33], so $\log|v(x, z)|$ is subharmonic.

(b) Is analogous to the proof of [12, Theorem 4.3]: it is a consequence of locally uniform upper bounds which follow from general principles, and a pointwise lower bound which follows from boundedness of the diagonal Green function for the Schrödinger operator with Neumann boundary conditions.

(i) If V is Stahl–Totik regular and $u(x, z)$ denotes the Dirichlet solution, then by [12, Theorem 1.5] for any $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\lim_{x\to\infty} \frac{1}{x} \log|u(x, z)| = M_{\mathsf{E}}(z).$$

In words, $u(x, z)$ grows exponentially with rate $M_{\mathsf{E}}(z) > 0$. Moreover, the Weyl solution $\psi(x, z)$ decays, $\lim_{x\to\infty} \psi(x, z) = 0$. The Neumann solution $v(x, z)$ is a linear combination of $u(x, z)$ and $\psi(x, z)$ and it is not a multiple of $\psi(x, z)$ (since z is not real, it is not an eigenvalue of the self-adjoint operator). Thus, $v(x, z)$ also obeys

$$\lim_{x\to\infty} \frac{1}{x} \log|v(x, z)| = M_{\mathsf{E}}(z), \quad \forall z \in \mathbb{C} \setminus \mathbb{R}.$$

This implies that M_{E} is the only possible subsequential limit of $\frac{1}{x} \log|v(x, z)|$ in $\mathcal{D}'(\mathbb{C})$. By precompactness of the family, this implies convergence to M_{E} in the topology of $\mathcal{D}'(\mathbb{C})$.

(ii) Follows from convergence in $\mathcal{D}'$ together with the principle of descents for subharmonic functions. □

We will show that for Dirichlet regular sets (ii) holds uniformly on compact subsets of $\mathbb{C}$. This will follow from showing that (ii) holds not only pointwise but in the following stronger sense: Given a sequence $(z_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} z_n = z_\infty \in \mathbb{C}$$

and increasing sequence $(x_n)_{n\in\mathbb{N}}$, $\lim_{n\to\infty} x_n = \infty$, then

$$\limsup_{n\to\infty} \frac{1}{x_n} \log|u(x_n, z_n)| \le M_{\mathsf{E}}(z_\infty).$$

If then, in addition, M_{E} is continuous (or merely lower semicontinuous), this implies that (ii) holds uniformly on compact subsets of $\mathbb{C}$. Since for Dirichlet regular sets E, M_{E} is continuous [12, Theorem 2.1], this will prove the above claim.

We will need a version of principle of descents for Green potentials. Let $\Omega \subset \mathbb{C}$ be a Greenian domain and denote its Green function by $G_\Omega(z, w)$, cf. [2, Section 4]. Let ν be a measure supported in Ω and define the Green potential of ν by

$$\Phi_\nu^G(z) = \int G_\Omega(z, t)d\nu(t)$$

provided that there exists $z_0 \in \Omega$ such that $\Phi_\nu^G(z_0) < \infty$. It defines a superharmonic function in Ω. By [2, Lemma 4.2.2] this holds in particular, if ν is supported on a compact subset of Ω.

Lemma 2.2. *Let Ω be a Greenian domain and let ν_n, ν_∞ be finite measures with support in a common compact subset of Ω and $\lim_{n\to\infty} \nu_n = \nu_\infty$ in the weak-$*$ sense. Let $z_n \in \Omega$ with $\lim_{n\to\infty} z_n = z_\infty \in \Omega$. Then*

$$\liminf_{n\to\infty} \Phi_{\nu_n}^G(z_n) \geq \Phi_{\nu_\infty}^G(z_\infty).$$

Proof. By assumption there exists a compact set $K \subset \Omega$ such that

$$\operatorname{supp} \nu_n \cup \operatorname{supp} \nu_\infty \cup \{z_n\}_{n\in\mathbb{N}} \cup \{z_\infty\} \subset K.$$

For $M > 0$ define

$$G_\Omega^M(z,t) = \min\{M, G_\Omega(z,t)\}.$$

The claim will follow from uniform continuity of G_Ω^M on $K \times K$. By [2, Theorem 4.1.9.] $G_\Omega(z,t)$ is continuous on $K \times K$ in the extended sense (i.e., with the value $+\infty$ allowed). Thus it follows that $G_\Omega^M(z,t)$ is continuous on $K \times K$ and $G_\Omega^M(z,t) \leq M$. We conclude that $G_\Omega^M(z,t)$ is continuous in the standard sense and since K is compact, uniform continuity follows. Since ν_∞ is a finite measure and $\nu_n \to \nu_\infty$ we have $\sup \nu_n(K) < \infty$. Now it follows from the monotone convergence theorem that

$$\begin{aligned}
\Phi_{\nu_\infty}^G(z_\infty) &= \lim_{M\to\infty} \int G_\Omega^M(z_\infty, t) d\nu_\infty(t) \\
&= \lim_{M\to\infty} \lim_{n\to\infty} \int G_\Omega^M(z_n, t) d\nu_n(t) \\
&\leq \lim_{M\to\infty} \liminf_{n\to\infty} \int G_\Omega(z_n, t) d\nu_n(t) \\
&= \liminf_{n\to\infty} \Phi_{\nu_n}^G(z_n).
\end{aligned}$$

□

Following the proof of [5, Theorem 2.7.4.1] we obtain:

Theorem 2.3. *Let $(u_n)_{n\in\mathbb{N}}$ and u_∞ be superharmonic functions in $\mathbb{C}$ such that $\lim_{n\to\infty} u_n = u_\infty$ in $\mathcal{D}'(\mathbb{C})$. Let $z_n \in \mathbb{C}$ with $\lim_{n\to\infty} z_n = z_\infty \in \mathbb{C}$. Then*

$$\liminf_{n\to\infty} u_n(z_n) \geq u_\infty(z_\infty).$$

Proof. Fix $R > 0$ such that $z_n, z_\infty \in B_R(0)$. By [5, Theorem 2.7.1.1] the u_n are uniformly bounded from below on $B_{2R}(0)$ and thus we can assume that the u_n and u_∞ are non-negative there.

For a non-negative superharmonic function u in a domain Ω and $E \subset \Omega$ let $\hat{R}_u^E$ denote the regularized reduced function [2, Section 5.3].[1] View u_n, u as non-negative superharmonic functions in $B_{2R}(0)$. Set $K = \overline{B_R(0)}$ and

$$v_n = \hat{R}^K_{u_n}, \quad v = \hat{R}^K_u.$$

Since $K \subset B_{2R}(0)$ is compact we see that v_n and v are Green potentials [2, Theorem 5.3.5], i.e., $v_n = \Phi^G_{\nu_n}$, $v = \Phi^G_\nu$ for some measures ν_n, ν and the Green function of the domain $B_{2R}(0)$. By [2, Theorem 5.3.4] v_n, v are harmonic outside of K and it follows that ν_n, ν are supported on K. Due to [5, Theorem 2.7.2.2], $v_n \to v$ in $\mathcal{D}'(B_{2R}(0))$ and thus since by [2, Theorem 4.3.8] ν_n, ν are the Riesz measures of v_n and v we conclude that $\nu_n \to \nu$ in $C(K)^*$. Therefore, we obtain from Lemma 2.2 that

$$\liminf_{n\to\infty} \Phi^G_{\nu_n}(z_n) \geq \Phi^G_\nu(z).$$

Since $v_n = u_n$ and $v = u$ on $B_R(0)$ the claim follows. □

We immediately get the following:

Theorem 2.4. *If V is Stahl–Totik regular then for any $z_n \to z$ and $x_n \to \infty$ we have*

$$\limsup_{n\to\infty} \frac{1}{x_n} \log |v(x_n, z_n)| \leq M_{\mathsf{E}}(z). \tag{2.1}$$

In particular, if E is Dirichlet regular, uniformly on compact subsets of $\mathbb{C}$ we have

$$\limsup_{x\to\infty} \frac{1}{x} \log |v(x, z)| \leq M_{\mathsf{E}}(z).$$

Proof. As in the proof of Lemma 2.1 regularity implies that for any $x_n \to \infty$ the family of subharmonic functions $(\frac{1}{x_n} \log |v(x_n, z_n)|)$ converges to $M_{\mathsf{E}}(z)$ in $\mathcal{D}'(\mathbb{C})$ and is bounded from above on compact subsets of $\mathbb{C}$. Thus, the first claim follows from Theorem 2.3 applied to $(-\frac{1}{x_n} \log |v(x_n, z_n)|)$. If E is Dirichlet regular, M_{E} is continuous on $\mathbb{C}$. The second statement then follows from (2.1) and lower semi-continuity of M_{E}. □

We need to control the Neumann solution for large real energies. We only need bounds for real spectral parameter ξ, but ξ can be negative. We will always assume that $\operatorname{Im} \sqrt{z} \geq 0$ for $z \in \mathbb{C}$. The Neumann and Dirichlet solutions for $V = 0$ are the functions

$$c(x, z) = \cos(\sqrt{z}x), \quad s(x, z) = \begin{cases} \frac{\sin(\sqrt{z}x)}{\sqrt{z}}, & z \neq 0, \\ x, & z = 0. \end{cases}$$

[1] In [5, Theorem 2.7.2.1] the regularized reduction was introduced in a different way, but it follows from its characteristic properties that these two notions, up to switching from super- to subharmonic functions, coincide.

By standard arguments, for general $V \in L^1([0,x])$, the initial value problem (1.2) is rewritten as integral equations, and by Volterra-type arguments, convergent series representations are then found for the fundamental solutions. With the notation $\Delta_n(x) = \{t \in \mathbb{R}^n \mid x \geq t_1 \geq t_2 \geq \cdots \geq t_n \geq 0\}$, the series representation for the Neumann solution is

$$
\begin{aligned}
v(x,z) &= c(x,z) \\
&\quad + \sum_{n=1}^{\infty} \int_{\Delta_n(x)} s(x-t_1,z)\left(\prod_{j=1}^{n-1} V(t_j)s(t_j-t_{j+1},z)\right)V(t_n)c(t_n,z)\mathrm{d}^n t,
\end{aligned} \tag{2.2}
$$

see [33] or [12, Section 3].

Lemma 2.5. *Let $x > 0$. For $\xi \geq 1$ we have*

$$
|v(x,\xi)| \leq e^{\frac{\int_0^x |V(t)|dt}{\sqrt{\xi}}}. \tag{2.3}
$$

For $\xi < 1$ we have

$$
|v(x,\xi)| \leq e^{(1+\operatorname{Im}\sqrt{\xi})x + \int_0^x |V(t)|dt}. \tag{2.4}
$$

Proof. For $\xi \geq 1$ we use the estimates $|c(x,\xi)| \leq 1$ and $|s(x,\xi)| \leq \sqrt{\xi}^{-1}$ to get from (2.2)

$$
|v(x,\xi)| \leq 1 + \sum_{n=1}^{\infty} \frac{\int_{\Delta_n(x)} \prod_{j=1}^n |V(t_j)|\mathrm{d}^n t}{\sqrt{\xi}^n} = 1 + \sum_{n=1}^{\infty} \frac{\left(\int_0^x |V(t)|dt\right)^n}{\sqrt{\xi}^n n!} = e^{\frac{\int_0^x |V(t)|dt}{\sqrt{\xi}}}.
$$

Similarly, for $\xi < 1$ we use $|s(x,\xi)| = |\int_0^x c(t,\xi)\mathrm{d}t| \leq xe^{\operatorname{Im}\sqrt{\xi}x} \leq e^{(1+\operatorname{Im}\sqrt{\xi})x}$ and $|c(x,\xi)| \leq e^{\operatorname{Im}\sqrt{\xi}x} \leq e^{(1+\operatorname{Im}\sqrt{\xi})x}$ and get

$$
|v(x,\xi)| \leq e^{(1+\operatorname{Im}\sqrt{\xi})x} e^{\int_0^x |V(t)|dt}.
$$

□

Combining these estimates for large real energies and Theorem 2.4 on compact subsets of $\mathbb{C}$, we get a uniform growth estimate for the Neumann solutions for $\xi \in \mathbb{R}$ which are close to E. To this end let us introduce for $0 < \delta < 1$ the extension

$$
\mathsf{E}_\delta = \{\xi \in \mathbb{R} : \operatorname{dist}(\xi, \mathsf{E}) < \delta\} \cup [1/\xi, \infty). \tag{2.5}
$$

Having in mind that ∞ is a boundary point of Ω it is natural to also add a half-line. Note that this makes E_δ a finite gap set, which will be crucial in Section 5.

Theorem 2.6. *Let V be Stahl–Totik regular such that $\mathsf{E} = \sigma_{\mathrm{ess}}(V)$ is Dirichlet regular. Then for any $\varepsilon > 0$ sufficiently small there exists $0 < \delta < 1$ and $C > 0$ such that for any $\xi \in \mathsf{E}_\delta$ and $x > 0$ we have*

$$
|v(x,\xi)| \leq Ce^{\varepsilon x}.
$$

Proof. Let

$$c_1 = \sup_{x \geq 1} \int_x^{x+1} |V(t)| dt$$

and fix $0 < \varepsilon < 2c_1$. We have $\int_0^x |V(t)|dt \leq c_1(x+1) \leq 2c_1 x$ for $x \geq 1$. Setting $\delta_1 = (\frac{\varepsilon}{2c_1})^2$ we obtain the estimate for $\xi \geq 1/\delta_1$ from (2.3).

Let $K = [\inf \mathsf{E} - 1, 1/\delta_1]$. Since M_{E} is uniformly continuous on K and vanishes on E, we find $\delta_2 > 0$ such that for $\xi \in K$ with $\operatorname{dist}(\xi, \mathsf{E}) \leq \delta_2$, $M_{\mathsf{E}}(\xi) < \frac{\varepsilon}{2}$. Thus, by Theorem 2.4 there exists $L_0 > 0$ such that for $x \geq L_0$ and $\xi \in K$ with $\operatorname{dist}(\xi, \mathsf{E}) < \delta_2$

$$\frac{1}{x} \log |v(x, \xi)| \leq \varepsilon.$$

For $x \leq L_0$ we use (2.4) to get a uniform bound for $|v(x, \xi)|$. The claim follows with $\delta = \min\{\delta_1, \delta_2\}$. □

3 Christoffel–Darboux kernel, Christoffel function and Fourier transform

In this section we recall aspects of the spectral theory of continuum Schrödinger operators on the half-line. By relating it to the rich theory of canonical systems developed by Krein and de Branges, we will characterize the Christoffel–Darboux kernel as a reproducing kernel of a certain Hilbert space of entire functions. This viewpoint will also give an expression of $\lambda_L(z)$ in terms of an extremal problem. The connection between canonical systems and continuum Schrödinger operators has been extensively discussed by Remling in [35, 36].

For a given potential $V : [0, \infty) \to \mathbb{R}$, $V \in L^1_{\text{loc}}$, let H_V be the associated Schrödinger operator on the half-line, where we again assume a Neumann boundary condition at 0 and that V is so that ∞ is a limit point endpoint. If we consider the truncation to $[0, L]$, then we assume in addition a Neumann boundary condition at L

$$f'(L) = 0$$

and denote the corresponding operator by H_V^L.

Let us start by considering H_V^L. It is well known that the spectrum is purely discrete and simple. Let $v(x, z)$ and $u(x, z)$ denote the Neumann and Dirichlet solution, respectively. Define the measure

$$\mu_L = \sum_{\xi : \partial_L v(L, \xi) = 0} \frac{\delta_\xi}{\|v(\cdot, \xi)\|_{L^2_{dx}((0,L))}},$$

where δ_ξ denotes the Dirac measure. Note that this is not the same measure as ν_L defined in the introduction. The operator $U : L^2_{dx}((0, L)) \to L^2_{d\mu_L}(\mathbb{R})$

$$(Uf)(z) = \int_0^L f(x)v(x, z)dx \tag{3.1}$$

is unitary and

$$UH_V^L U^* = S_z,$$

where S_z denotes the operator of multiplication with the independent variable in $L^2_{d\mu_L}(\mathbb{R})$. The adjoint of U is given by

$$(U^*F)(x) = \int v(x, \xi)F(\xi)d\mu_L(\xi).$$

Although in the above discussion $U(L^2_{dx}((0, L)))$ was considered as $L^2_{d\mu_L}(\mathbb{R})$, (3.1) allows to interpret for $f \in L^2_{dx}((0, L))$, $F(z) = (Uf)(z)$ as a function on $\mathbb{C}$. This leads to the theory of de Branges spaces. Define the transfer matrix

$$T(x, z) = \begin{pmatrix} v(x, z) & -u(x, z) \\ -\partial_x v(x, z) & \partial_x u(x, z) \end{pmatrix}$$

and note that it solves the differential equation

$$j\partial_x T(x, z) = \left(-z \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} V(x) & 0 \\ 0 & -1 \end{pmatrix} \right) T(x, z), \quad j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.2}$$

Such a system is called a canonical system with coefficient functions $A(x) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $B(x) = \left(\begin{smallmatrix} V(x) & 0 \\ 0 & -1 \end{smallmatrix}\right)$; see Appendix A. It follows from (A.3) that

$$E_L(z) = v(L, z) + i\partial_L v(L, z)$$

is a Hermite–Biehler function. This means that E is entire, has no zeros in $\mathbb{C}_+$ and satisfies $|E(z)| \geq |E(\overline{z})|$ there. To a Hermite–Biehler function we can associate a Hilbert space of entire functions $B(E)$ with scalar product

$$\langle F, G \rangle_{B(E)} = \frac{1}{\pi} \int F(x)\overline{G(x)} \frac{dx}{|E(x)|^2}; \tag{3.3}$$

see Appendix A.

Theorem 3.1 ([35, Theorem 3.1])**.** *The Hilbert space $B(E_L)$ and $L^2_{d\mu_L}(\mathbb{R})$ are identical. More precisely, if $F \in B(E_L)$, then the restriction of F to $\mathbb{R}$ belongs to $L^2_{d\mu_L}(\mathbb{R})$ and $F \mapsto F|_{\mathbb{R}}$ is unitary.*

For an entire function F we denote $F^{\#}(z) = \overline{F(\overline{z})}$. The Hilbert spaces $B(E_L)$ are reproducing kernel Hilbert spaces and the kernel is given by

$$K_L(z,w) = \frac{\overline{E_L(w)}E_L(z) - \overline{E_L^{\#}(w)}E_L^{\#}(z)}{2i(\overline{w}-z)} = \frac{\overline{v(L,w)}v'(L,z) - \overline{v'(L,w)}v(L,z)}{\overline{w}-z}. \tag{3.4}$$

On the other hand, using (A.4), it follows that

$$K_L(z,w) = \int_0^L v(x,z)\overline{v(x,w)}dx. \tag{3.5}$$

Using the terminology common in the orthognal polynomials literature, we call $K_L(z,w)$ the Christoffel–Darboux kernel. In the setting of orthogonal polynomials, the equivalence of (3.4) and (3.5) is called the Christoffel–Darboux formula. Evaluating (3.4) on the diagonal for $\xi \in \mathbb{R}$ gives

$$K_L(\xi,\xi) = (\partial_L v)(L,\xi)(\partial_\xi v)(L,\xi) - v(L,\xi)(\partial_\xi \partial_L v)(L,\xi). \tag{3.6}$$

It is a remarkable property that as sets $B(E_L)$ do not depend on the potential; see [35, Theorem 4.1]. That is, for any $V \in L^1([0,L])$,

$$B(E_L) = S_L := \left\{ \int_0^L f(x)\cos(\sqrt{z}x)dx : f \in L^2((0,L)) \right\} \tag{3.7}$$

where this is understood as set equality. Of course the topology depends on the potential through (3.3). Recall that $c(x,z) = \cos(\sqrt{z}x)$ is the Neumann solution for $V = 0$. Using the variable $k^2 = z$ it will be convenient to also consider the set

$$\mathfrak{S}_L = \left\{ \int_0^L f(x)\cos(kx)dx : f \in L^2((0,L)) \right\}. \tag{3.8}$$

It is mentioned in [29, page 101] and it follows from the Paley–Wiener theorem that

$$\mathfrak{S}_L = \{g \in L^2_{dk}(\mathbb{R}) : g \text{ is entire, even and of exponential type at most } L\}.$$

From (3.7) it follows directly that $B(E_{L_1}) \subset B(E_{L_2})$ for $L_1 \leq L_2$ and in fact this inclusion is isometric. But even more is true:

Theorem 3.2 ([35, Theorem 3.2])**.** *Suppose* $0 < L_1 \leq L_2$. *Then* $B(E_{L_1})$ *is isometrically contained in* $B(E_{L_2})$. *Moreover, if* μ *denotes the spectral measure for the half-line problem, then for every* $L > 0$, $B(E_L)$ *is isometrically contained in* $L^2_{d\mu}(\mathbb{R})$ *in the sense that for* $F \in B(E_L)$, $F|_{\mathbb{R}} \in L^2_{d\mu}$ *and* $\|F\|_{B(E_L)} = \|F|_{\mathbb{R}}\|_{L^2_{d\mu}}$.

In the following, we will drop the restriction map and interpret F either as a function in $L^2_{d\mu}$ or as an entire function in $B(E_L)$ depending on the context. This shouldn't lead to any confusion.

Using that $B(E_L)$ is isometrically contained in $L^2_{d\mu}(\mathbb{R})$ and the equality (3.7) we can define the Christoffel function:

Definition 3.3. For $L > 0$ and $z \in \mathbb{C}$ we define the Christoffel function associated to V by

$$\lambda_L(z) = \inf\{ \|F\|^2_{L^2_{d\mu}} : F \in S_L, F(z) = 1\}.$$

Since $B(E_L)$ is a reproducing kernel Hilbert space it follows from the Cauchy–Schwarz inequality that the infimum is in fact a minimum and that the extremizer is given by

$$Q_L(z, z_0) = \frac{K_L(z, z_0)}{K_L(z_0, z_0)}.$$

Note that $K_L(z_0, z_0) = \|K_L(\cdot, z_0)\|^2_{B(E_L)} = \|K_L(\cdot, z_0)\|^2_{L^2_{d\mu}}$. In particular,

$$\lambda_L(z) = \frac{1}{K_L(z, z)}.$$

Combining this with (3.5) shows that this definition of $\lambda_L(z)$ coincides with the one given in (1.3).

4 Asymptotics for finite gap potentials

The goal of this section is to show (1.7) for so-called finite gap potentials. We will need some preliminary observations:

4.1 The isospectral torus. In this section we assume that E is a finite gap set of the form

$$\mathsf{E} = [b_0, \infty) \setminus \bigcup_{j=1}^{g} (a_j, b_j),$$

where $b_0 < a_j < b_j < a_{j+1} < b_{j+1}$, for $0 < j < g - 1$. In this setting it is more natural to consider Schrödinger operators on $\mathbb{R}$. Let V be a continuous and bounded potential on $\mathbb{R}$ and H_V acting on its natural domain in $L^2(\mathbb{R})$. Clearly $H_{V_\pm}$ with potential $V_+ = V|_{[0,\infty)}$ and $V_- : [0, \infty) \to \mathbb{R}$ defined by $V_-(x) := V(-x)$ have ∞ as a limit point endpoint and thus we can associate Weyl m-functions $m_\pm$ by (1.5).

Definition 4.1. We say that V is reflectionless on $A \subset \mathbb{R}$ if for almost every $\xi \in A$

$$m_+(\xi + i0) = -m_-(\xi - i0).$$

For a given finite gap set E we define the isospectral torus by

$$\mathcal{T}(\mathsf{E}) = \{ V \in C_b(\mathbb{R}) : \sigma(H_V) = \mathsf{E} \text{ and } V \text{ is reflectionless on } \mathsf{E}\}.$$

This class has been considered for essentially more general sets [19, 41].

4.2 Abelian integrals. For finite gap potentials many important spectral theoretical objects can be given explicitly in terms of the Abelian integrals on the associated hyperelliptic Riemann surface $\mathcal{R}_\mathsf{E}$ associated to $\sqrt{(z-b_0)\prod(z-a_j)(z-b_j)}$, i.e.,

$$\mathcal{R}_\mathsf{E} = \left\{(z, w) \in \mathbb{C}^2 : w^2 = (z-b_0)\prod_{j=1}^{g}(z-a_j)(z-b_j)\right\} \cup \{\infty\}.$$

Typically $\mathcal{R}_\mathsf{E}$ is visualized as two copies of $\mathbb{C} \setminus \mathsf{E}$, corresponding to the two branches of the square root, glued together along E. For more details see [18]. These representations and the properties which follow from them allow us to compute the limit of the Christoffel functions explicitly.

For finite gap sets the Martin function can be given in terms of a Schwarz–Christoffel mapping. Define

$$\theta_\mathsf{E}(z) = \int_{b_0}^{z} \frac{-1}{2\sqrt{u-b_0}} \prod_{j=1}^{g} \frac{(u-c_j)}{\sqrt{(u-a_j)(u-b_j)}} du,$$

where $c_j \in (a_j, b_j)$ is uniquely determined by

$$\theta_\mathsf{E}(b_j) = \theta_\mathsf{E}(a_j).$$

The function θ_E is a conformal mapping of $\mathbb{C}_+$ to a comb

$$\Pi_\mathsf{E} = \{\xi + iy : \ \xi, y > 0\} \setminus \bigcup_{j=1}^{g}\{\eta_j + iy : 0 < y < h_j\},$$

where η_j are called frequencies and h_j heights. Note that θ_E can be extended to $\mathbb{R}$,

$$\theta_\mathsf{E}(z) = \sqrt{z}, \quad \text{as } z \to -\infty, \quad \theta_\mathsf{E}(b_0) = 0$$

and that $\theta_\mathsf{E}^{-1}(\mathbb{R}_+) = \mathsf{E}$. It follows from these properties that $M_\mathsf{E}(z) = \operatorname{Im}\theta_\mathsf{E}(z)$ is the Martin function of the domain. These type of comb mappings are commonly used in inverse spectral theory and in uniform approximation problems [1, 30] or [15] for a modern approach to the subject. Since M_{E} can be extended to a subharmonic function on $\mathbb{C}$, its distributional Laplacian is a positive measure and we can define $\rho_{\mathsf{E}} = \frac{1}{2\pi}\Delta M_{\mathsf{E}}$. The Riesz representation then yields

$$M_{\mathsf{E}}(z) = M_{\mathsf{E}}(z_*) + \int_{\mathsf{E}} \log\left|1 - \frac{z - z_*}{t - z_*}\right| d\rho_{\mathsf{E}},$$

where $z_* < b_0$ is some normalization point. Computing $\frac{1}{2\pi}\Delta M_{\mathsf{E}}$ we find that

$$d\rho_{\mathsf{E}}(\xi) = \frac{1}{\pi}\theta'_{\mathsf{E}}(\xi)d\xi = \frac{-1}{2\pi\sqrt{\xi - b_0}} \prod_{j=1}^{g} \frac{(\xi - c_j)}{\sqrt{(\xi - a_j)(\xi - b_j)}} d\xi. \tag{4.1}$$

That is, ρ_{E} is purely a.c. and its density is given by

$$\frac{d\rho_{\mathsf{E}}(\xi)}{d\xi} = f_{\mathsf{E}}(\xi) = \frac{-1}{2\pi\sqrt{\xi - b_0}} \prod_{j=1}^{g} \frac{(\xi - c_j)}{\sqrt{(\xi - a_j)(\xi - b_j)}}.$$

In particular, we see that $f_{\mathsf{E}}(\xi)$ is real analytic inside any band of E, see also Lemma B.2. We will also constantly use that $\theta_{\mathsf{E}}(\xi) \in \mathbb{R}$ for $\xi \in \mathsf{E}$.

Let ψ_+ denote the Weyl solution at $+\infty$, cf. (1.4). Due to [18, p. 462], ψ_+ corresponds to the restriction of the Baker–Akhiezer function to the upper sheet and thus by [18, Theorem 1.20] it can be represented as

$$\psi_+(x, z) = e^{i\theta_{\mathsf{E}}(z)x} f(x, z),$$

where f is given in terms of Theta functions on $\mathcal{R}_{\mathsf{E}}$. In the following let $[a, b] = I \subset \operatorname{int}(\mathsf{E})$. We can extend $\psi_+(x, z)$ analytically to I. Moreover, we have

(i) $x \mapsto f(x, \xi)$ is almost periodic,[2]

(ii) $\xi \mapsto f(x, \xi)$ is analytic on I and all derivatives are uniformly bounded for $x > 0$ and $\xi \in I$.

4.3 Asymptotics of $\lambda_L(\xi)$. In the following let $[a, b] = I \subset \operatorname{int}(\mathsf{E})$. If we write $f(\xi)$ or $m_+(\xi)$, $\xi \in I$, we mean the corresponding limits $\xi + i\varepsilon$ as $\varepsilon \to 0$. All of them can be analytically extended to I. Moreover, we have $\operatorname{Im} m_+ > 0$ there. To avoid confusion, we mention that for m_+, this does not correspond to the extension $\overline{m(\overline{z})}$, which is an extension through $\mathbb{R} \setminus \mathsf{E}$.

Let

$$W(f, g)(x) = f(x)g'(x) - f'(x)g(x)$$

[2] This follows from continuity of the Theta function and the linearization of the Abel map in [18, Theorem 1.20].

denote the Wronskian of f and g. For $\xi\in I$, $\psi_+(x,\xi)$ and $\overline{\psi_+(x,\xi)}$ both solve (1.4) and thus their Wronskian is constant. For $\xi\in I$, $\partial_x\psi_+(0,\xi)\neq 0$, since this would lead to an eigenvalue and $\sigma(H_V)$ is purely absolutely continuous there. Recall that m_+ is given by

$$m_+(z)=-\frac{\psi_+(0,z)}{\partial_x\psi_+(0,z)}.$$

Thus we see that by constancy of the Wronskian

$$\begin{aligned}W_\xi(\psi_+,\overline{\psi_+}) &:= W(\psi_+(\cdot,\xi),\overline{\psi_+(\cdot,\xi)})(x)=W(\psi_+(\cdot,\xi),\overline{\psi_+(\cdot,\xi)})(0)\\ &=-2i|\partial_x\psi_+(0,\xi)|^2\operatorname{Im}m_+(\xi)\neq 0.\end{aligned}\tag{4.2}$$

It follows now by direct verification of (1.2) that

$$v(x,\xi)=\frac{\psi_+(x,\xi)\overline{\partial_x\psi_+(0,\xi)}-\overline{\psi_+(x,\xi)}\partial_x\psi_+(0,\xi)}{W_\xi(\psi_+,\overline{\psi_+})}$$

defines the Neumann solution for H_V.

Let us set

$$c(\xi)=\frac{\overline{\partial_x\psi_+(0,\xi)}}{W_\xi(\psi_+,\overline{\psi_+})}$$

and

$$h(x,\xi)=c(\xi)\psi_+(x,\xi).$$

It will also be convenient to take off the exponential part and consider

$$g(x,\xi)=e^{-i\theta_{\mathsf{E}}(\xi)x}h(x,\xi).$$

Using $\overline{W_\xi(\psi_+,\overline{\psi_+})}=-W_\xi(\psi_+,\overline{\psi_+})$ and $\theta_{\mathsf{E}}(x)\in\mathbb{R}$, we see that

$$v(x,\xi)=h(x,\xi)+\overline{h(x,\xi)}=e^{i\theta_{\mathsf{E}}(\xi)x}g(x,\xi)+e^{-i\theta_{\mathsf{E}}(\xi)x}\overline{g(x,\xi)}.$$

The Herglotz function m_+ can be continuously extended to I. Thus, if

$$d\mu(\xi)=f_\mu(\xi)d\xi+d\mu_s(\xi)$$

denotes the spectral measure of H_V, we have

$$f_\mu(\xi)=\frac{1}{\pi}\operatorname{Im}m_+(\xi).$$

Lemma 4.2. *For $\xi\in I$, we have*

$$W(h(\cdot,\xi),\overline{h(\cdot,\xi)})(x)=-2i\theta_{\mathsf{E}}(\xi)|g(x,\xi)|^2+W(g(\cdot,\xi),\overline{g(\cdot,\xi)})(x)=\frac{1}{2\pi i f_\mu(\xi)}.$$

Proof. The first equality follows by direct computation. Using the definition of h and (4.2), we see that

$$\begin{aligned} W(h(\cdot,\xi),\overline{h(\cdot,\xi)})(x) &= W(c(\xi)\psi_+(\cdot,\xi),\overline{c(\xi)\psi_+(\cdot,\xi)})(x) \\ &= |c(\xi)|^2 W_\xi(\psi_+(\cdot,\xi),\overline{\psi_+}(\cdot,\xi))(x) \\ &= \frac{|\partial_x\psi_+(0,\xi)|^2}{|W_\xi(\psi_+,\overline{\psi_+})|^2} W_\xi(\psi_+(\cdot,\xi),\overline{\psi_+}(\cdot,\xi))(x) = \frac{1}{2\pi i f_\mu(\xi)}. \qquad \square \end{aligned}$$

We are now ready to prove the main theorem of this section:

Theorem 4.3. *Let* $\xi \in I = [a,b] \subset \mathrm{int}(\mathsf{E})$. *Then*

$$K_x(\xi,\xi) = x\frac{f_{\mathsf{E}}(\xi)}{f_\mu(\xi)} + O(1)$$

as $x \to \infty$,*where the* $O(1)$ *is uniform for* $\xi \in I$.

Proof. We use the representation

$$v(x,\xi) = e^{i\theta_{\mathsf{E}}(\xi)x} g(x,\xi) + e^{-i\theta_{\mathsf{E}}(\xi)x}\overline{g(x,\xi)}.$$

All O notation is related to $x \to \infty$. We have

$$\begin{aligned} \partial_\xi v(x,\xi) &= i\theta'_{\mathsf{E}}(\xi)x(e^{i\theta_{\mathsf{E}}(\xi)x}g(x,\xi) + e^{-i\theta_{\mathsf{E}}(\xi)x}\overline{g(x,\xi)}) \\ &\quad + e^{i\theta_{\mathsf{E}}(\xi)x}\partial_\xi g(x,\xi) - e^{-i\theta_{\mathsf{E}}(\xi)x}\partial_\xi\overline{g(x,\xi)} \\ &= i\theta'_{\mathsf{E}}(\xi)x(e^{i\theta_{\mathsf{E}}(\xi)x}g(x,\xi) - e^{-i\theta_{\mathsf{E}}(\xi)x}\overline{g(x,\xi)}) + O(1). \end{aligned}$$

and

$$\begin{aligned} \partial_x v(x,\xi) &= i\theta_{\mathsf{E}}(\xi)e^{i\theta_{\mathsf{E}}(\xi)x}g(x,\xi) - i\theta_{\mathsf{E}}(\xi)e^{-i\theta_{\mathsf{E}}(\xi)x}\overline{g(x,\xi)} \\ &\quad + e^{i\theta_{\mathsf{E}}(\xi)x}\partial_x g(x,\xi) + e^{-i\theta_{\mathsf{E}}(\xi)x}\partial_x\overline{g(x,\xi)} \end{aligned}$$

and

$$\begin{aligned} \partial_x\partial_\xi v(x,\xi) &= i\theta'_{\mathsf{E}}(\xi)x(i\theta_{\mathsf{E}}(\xi)e^{i\theta_{\mathsf{E}}(\xi)x}g(x,\xi) + i\theta_{\mathsf{E}}(\xi)e^{-i\theta_{\mathsf{E}}(\xi)x}\overline{g(x,\xi)} \\ &\quad + e^{i\theta_{\mathsf{E}}(\xi)x}\partial_x g(x,\xi) - e^{-i\theta_{\mathsf{E}}(\xi)x}\partial_x\overline{g(x,\xi)}) + O(1). \end{aligned}$$

Thus, having in mind (3.6) we compute

$$\begin{aligned} &\partial_\xi v(x,\xi)\partial_x v(x,\xi) - v(x,\xi)\partial_x\partial_\xi v(x,\xi) \\ &\qquad = 2i\theta'_{\mathsf{E}}(\xi)x(-2i\theta_{\mathsf{E}}(\xi)|g|^2 + W(g(\cdot,\xi),\overline{g(\cdot,\xi)})(x)) + O(1). \end{aligned}$$

Using Lemma 4.2 and (4.1) we see that

$$2i\theta_{\mathsf{E}}'(\xi)x(-2i\theta_{\mathsf{E}}(\xi)|g|^2 + W(g(\cdot,\xi), \overline{g(\cdot,\xi)})(x)) = \frac{2\pi i f_{\mathsf{E}}(\xi)}{2\pi i f_{\mu}(\xi)}x$$

which finishes the proof by noting that the $O(1)$ is uniform for $\xi \in I$. □

We will use later that this implies in particular for any $\varepsilon > 0$

$$\lim_{L\to\infty} \frac{K_{(1+\varepsilon)L}(\xi,\xi)}{K_L(\xi,\xi)} = 1+\varepsilon \tag{4.3}$$

uniformly for $\xi \in I$.

5 Proofs of the main theorems

The goal of this section is to prove Theorems 1.1, 1.2 and 1.3. We will need some preparatory work.

Recall the set

$$S_L = \left\{ \int_0^L \cos(\sqrt{z}t) f(t)dt, \quad f \in L^2((0,L)) \right\}$$

and that

$$\lambda_L(\xi) = \min\{ \|F\|^2_{L^2_{d\mu}} : F \in S_L,\ F(\xi) = 1\}, \tag{5.1}$$

with minimizer

$$Q_L(z,z_0) = \frac{K_L(z,z_0)}{K_L(z_0,z_0)}. \tag{5.2}$$

Lemma 5.1. *Let $d_0 \in \mathbb{R}$ and $\xi_0 > d_0$. Then there is $c > 0$ such that*

$$F_c(z) = \frac{\sin(c(\sqrt{z-d_0} - \sqrt{\xi_0 - d_0}))}{\sqrt{z-d_0} - \sqrt{\xi_0-d_0}} + \frac{\sin(c(\sqrt{z-d_0} - \sqrt{\xi_0 - d_0}))}{\sqrt{z-d_0} + \sqrt{\xi_0-d_0}}$$

satisfies

(i) $|F_c(\xi)| \le F_c(\xi_0)$ *for all* $\xi \ge d_0$;

(ii) *for any* $\delta > 0$ *there exists* $\varepsilon > 0$ *such that for any* $\xi \in [d_0,\infty) \setminus (\xi_0-\delta, \xi_0+\delta)$ *we have that* $|F_c(\xi)| \le F_c(\xi_0) - \varepsilon$;

(iii) $F_c \in S_c$.

Proof. First we find c so that $|F_c|$ has a global maximum at ξ_0. Define

$$G(u) = \frac{\sin(u-u_0)}{u-u_0} + \frac{\sin(u+u_0)}{u+u_0},$$

where u_0 is the first positive root of $\tan(2u_0) - 2u_0$. It can be directly verified that G is even and that $|G(u)| \leq G(u_0)$ for all $u \in \mathbb{R}$. From this it follows that for given $\zeta_0 > 0$, $c = u_0/\zeta_0$ and

$$H(\zeta) = \frac{\sin(c(\zeta - \zeta_0))}{\zeta - \zeta_0} + \frac{\sin(c(\zeta + \zeta_0))}{\zeta + \zeta_0},$$

that $|H|$ has a global maximum at ζ_0. Since $\zeta(z) = \sqrt{z - d_0}$ maps $[d_0, \infty)$ onto $[0, \infty)$ the first claim follows by substitution and setting

$$\zeta_0 = \zeta(\xi_0) = \sqrt{\xi_0 - d_0}, c = u_0/\sqrt{\xi_0 - d_0}.$$

It remains to show that $F \in S_c$.

We claim that F_c is the reproducing kernel $K_c(z, \xi_0)$ for the constant potential $V = d_0$. The Christoffel–Darboux formula (3.4) applied to $v(x, z) = \cos(x\sqrt{z - d_0})$ yields

$$K_c(z, \xi_0) = \frac{-\sqrt{z - d_0}\cos(c\sqrt{\xi_0 - d_0})\sin(c\sqrt{z - d_0})}{\xi_0 - z} + \frac{\sqrt{\xi_0 - d_0}\sin(c\sqrt{\xi_0 - d_0})\cos(c\sqrt{z - d_0})}{\xi_0 - z}.$$

Using trigonometric identities we get that $K_c(z, z_0)$ is given by

$$\frac{1}{2}\Big(\frac{\sin(c(\sqrt{z - d_0} - \sqrt{\xi_0 - d_0}))(\sqrt{z - d_0} + \sqrt{\xi_0 - d_0})}{z - d_0 - (\xi_0 - d_0)} + \frac{\sin(c(\sqrt{z - d_0} + \sqrt{\xi_0 - d_0}))(\sqrt{z - d_0} - \sqrt{\xi_0 - d_0})}{z - d_0 - (\xi_0 - d_0)}\Big).$$

Thus,

$$F_c(z) = 2K_c(z, \xi_0)$$

and in particular, $F_c \in S_c$. □

Remark. For later reference we mention that the c is explicitly constructed in the proof. Let u_0 be the first positive solution of $2u = \tan(2u)$, i.e., u_0 is a constant not depending on d_0, ξ_0. Then c is given by

$$c = \frac{u_0}{\sqrt{\xi_0 - d_0}}.$$

In particular, if $d_0(\varepsilon) = \min \mathsf{E} - \varepsilon$ and $\xi_0 \in \operatorname{int}(\mathsf{E})$, then

$$\lim_{\varepsilon \to 0} \varepsilon c(\varepsilon) = 0 \tag{5.3}$$

and this limit is uniform for $\xi_0 \in [a, b] \subset \operatorname{int}(\mathsf{E})$.

Recall the δ-extension (2.5). The following estimate is the crucial bound which allows to prove (1.7) for regular potentials:

Lemma 5.2. *Let V be a Stahl–Totik regular potential such that $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$ is Dirichlet regular, μ the associated spectral measure and Q_L as in* (5.2)*. Then for any $\varepsilon > 0$ there is $C > 0$ and $\delta > 0$ such that for any $\xi \in \mathsf{E}_\delta$ and $L > 0$*

$$|Q_L(\xi, \xi_0)| \le Ce^{\varepsilon L}\sqrt{\lambda_L(\xi_0)}.$$

Proof. By Theorem 2.6 we find $\delta > 0$ so that for $\xi \in \mathsf{E}_\delta$ and $x > 0$

$$|v(x, \xi)| \le \tilde{C}e^{\varepsilon x}.$$

Thus,

$$\int_0^L v(x,\xi)^2 dx \le \tilde{C}^2 \int_0^L e^{2\varepsilon x} dx = \frac{\tilde{C}^2}{2\varepsilon}(e^{2\varepsilon L} - 1) \le \frac{\tilde{C}^2}{2\varepsilon} e^{2\varepsilon L}.$$

Thus with $C^2 = \tilde{C}^2/2\varepsilon$ we have

$$\left(\int_0^L v(x,\xi)^2 dx\right)^{1/2} \le Ce^{\varepsilon L}.$$

On the other hand, by the reproducing kernel property

$$\begin{aligned} |K_L(\xi, \xi_0)| &= |\langle K_L(\cdot, \xi_0), K_L(\cdot, \xi)\rangle| \le \|K_L(\cdot, \xi_0)\| \|K_L(\cdot, \xi)\| \\ &= \sqrt{K_L(\xi_0, \xi_0)}\sqrt{K_L(\xi, \xi)} \end{aligned}$$

Using

$$K_L(\xi, \xi) = \int_0^L v(x, \xi)^2 dx, \quad \lambda_L(\xi_0) = \frac{1}{K_L(\xi_0, \xi_0)}$$

and (5.2) the claim follows. □

We are now ready to prove the main comparising result that allows to lift the results from Section 4 to arbitrary regular potentials.

Theorem 5.3. *Let $V, \tilde{V}$ be potentials satisfying* (1.1) *and μ, $\tilde{\mu}$ the associated spectral measures and $\mathsf{E} = \sigma_{\mathrm{ess}}(H_V)$, $\tilde{\mathsf{E}} = \sigma_{\mathrm{ess}}(H_{\tilde{V}})$. Suppose that V is a Stahl–Totik regular potential and E Dirichlet regular. Let I be a closed interval such that $I \subset \mathrm{int}(\mathsf{E}) \cap \mathrm{int}(\tilde{\mathsf{E}})$, μ and $\tilde{\mu}$ are absolutely continuous in a neighborhood of I and its densities $f_\mu, f_{\tilde{\mu}}$ are positive and continuous at every point of I. For any $\varepsilon > 0$, let $\delta > 0$ be as in Lemma 5.2. If there exists $\delta_1 > 0$ such that*

$$\tilde{\mathsf{E}}_{\delta_1} \subset \mathsf{E}_\delta \tag{5.4}$$

then there exist $\delta_2 = \delta_2(I) > 0$, $D = D(\varepsilon) > 0$ *and* $\gamma = \gamma(\varepsilon, \delta_2) < 1$ *such that*

$$\frac{\lambda_M(\xi_0, \tilde{V})}{\lambda_L(\xi_0, V)} \leq \sup_{|\xi-\xi_0|<\delta_2} \frac{f_{\tilde{\mu}}(\xi)}{f_\mu(\xi)} + D\gamma^{4N} e^{2\varepsilon L} + De^{2\varepsilon L} 2^{-2N},$$

where $M = L + 2Nc$ *and* $N = N(\varepsilon)$ *is sufficiently large.*

Remark. Let us comment on the meaning of (5.4). In the proof we will need to estimate the extremizer for $\lambda_L(\xi_0, V)$ on supp $\tilde{\mu}$. Close to the spectrum this can be done due to regularity by Lemma 5.2. However, there may be point masses of $\tilde{\mu}$ in the gaps. By extending $\tilde{E}$ we ensure that there are only finitely many point masses in $\mathbb{R} \setminus \tilde{\mathsf{E}}_{\delta_1}$, since there can only be finitely many eigenvalues in each gap of $\tilde{\mathsf{E}}_{\delta_1}$ and $\tilde{\mathsf{E}}_{\delta_1}$ is a finite gap set.

Proof. Let $Q_L^V(\xi, \xi_0)$ be the minimizer for V and the point ξ_0. Then by Lemma 5.2 for $\xi \in \mathsf{E}_\delta$ we have

$$|Q_L^V(\xi, \xi_0)| \leq Ce^{\varepsilon L} \sqrt{\lambda_L(\xi_0, V)}. \tag{5.5}$$

Thus by assumption this also holds on $\tilde{\mathsf{E}}_{\delta_1}$. Let $d_0 = \min \tilde{\mathsf{E}}_{\delta_1}$, F_c as in Lemma 5.1 and

$$G(z) = \frac{F_c(z)}{F_c(\xi_0)}.$$

Then

(i) $G(\xi_0) = 1$;

(ii) for any $r > 0$ there is $\gamma < 1$ such that for every $\xi > d_0$ with $|\xi - \xi_0| > r$, $|G(\xi)| < \gamma$;

(iii) there is $C_1 > 0$ such that for $\xi > \xi_0 + 1$,

$$|G(\xi)| \leq \frac{C_1}{\sqrt{\xi - \xi_0}}.$$

(iv) $G \in S_c$.

Since $\tilde{\mathsf{E}}_{\delta_1}$ has only finitely many gaps, there are only finitely many point masses of $\tilde{\mu}$ in $\mathbb{R} \setminus \tilde{\mathsf{E}}_{\delta_1}$. Let these points be denoted by $\zeta_1, \ldots, \zeta_n$. Let P be a polynomial of degree $n + 1$ that vanishes at these points and P has a local maximum at ξ_0 such that $P(\xi_0) = 1$. Let $N > n + 1$ and define

$$Q = Q_L^V G^{2N} P.$$

We claim that $Q \in S_{L+2Nc}$. Use $k^2 = z$ and define

$$H(k) = Q(k^2).$$

Since $Q_L^V \in S_L$ and $F \in S_c$ it follows that H is an even entire function of exponential type at most $L + 2Nc$. Thus, by (3.8) it remains to show that $H \in L^2_{dk}$. By (iii) $G^{2N}P$ are bounded on $\mathbb{R}_+$. Moreover, since $Q_L^V \in S_L$, $Q_L(k^2) \in L^2_{dk}$ and we conclude that $H \in L^2_{dk}$. Moreover, we have $Q(\xi_0) = 1$.

Thus, by (5.1) we get

$$\lambda_{L+2Nc}(\xi_0, \tilde{V}) \leq \|Q\|^2_{L^2_{d\tilde{\mu}}}. \tag{5.6}$$

We will split the integral into several parts. First let $\delta_2 > 0$ such that on $I_0 = (\xi_0 - \delta_2, \xi_0 + \delta_2)$, μ, $\tilde{\mu}$ are purely absolutely continuous and both are positive there. This can be achieved since they are continuous at every point of I. Moreover, let δ_2 be sufficiently small so that $|G|, |P| \leq 1$ on I_0. Then

$$\begin{aligned}
\int_{I_0} |Q(\xi)|^2 d\tilde{\mu}(\xi) &\leq \int_{I_0} |Q_L^V(\xi, \xi_0)|^2 d\tilde{\mu}(\xi) \\
&\leq \sup_{t \in I_0} \frac{f_{\tilde{\mu}}(t)}{f_\mu(t)} \int_{I_0} |Q_L^V(\xi, \xi_0)|^2 d\mu(\xi) \\
&\leq \sup_{t \in I_0} \frac{f_{\tilde{\mu}}(t)}{f_\mu(t)} \lambda_L(\xi_0, V).
\end{aligned}$$

Let us note that on $\operatorname{supp}(\tilde{\mu})$ we have

$$|Q(\xi)| \leq C e^{\varepsilon L} \sqrt{\lambda_L(\xi_0)} |G(\xi)|^{2N} |P(\xi)|. \tag{5.7}$$

For, we have already argued that (5.5) holds on $\tilde{\mathsf{E}}_{\delta_1}$. Thus, the only points where (5.7) may fail are the finite point masses of $\tilde{\mu}$ in $\mathbb{R} \setminus \tilde{\mathsf{E}}_{\delta_1}$. But this is where P vanishes and thus we obtain (5.7) on $\operatorname{supp}(\tilde{\mu})$.

Let $I_1 = (\operatorname{supp}(\tilde{\mu}) \setminus I_0) \cap (-\infty, \xi_0 + 1]$ and $I_2 = (\operatorname{supp}(\tilde{\mu})) \cap (\xi_0 + 1, \infty)$. Then

$$\begin{aligned}
\int_{I_1} |Q(\xi)|^2 d\tilde{\mu}(\xi) &\leq C^2 e^{2\varepsilon L} \lambda_L(\xi_0, V) \int_{I_1} |G(\xi)|^{4N} |P(\xi)|^2 d\tilde{\mu}(\xi) \\
&\leq \gamma^{4N} C^2 e^{2\varepsilon L} \lambda_L(\xi_0, V) \int_{I_1} |P(\xi)|^2 d\tilde{\mu}(\xi) \\
&= \gamma^{4N} C^2 e^{2\varepsilon L} C_1 \lambda_L(\xi_0, V).
\end{aligned}$$

Since

$$\int_{\mathbb{R}} \frac{d\tilde{\mu}(\xi)}{1 + \xi^2} < \infty,$$

we get for $n \geq 2$

$$\int_2^\infty \frac{d\mu(\xi)}{\xi^n} \leq \frac{K}{2^n}, \quad K = 4 \int_2^\infty \frac{d\mu(\xi)}{\xi^2}.$$

We conclude

$$\begin{aligned}\int_{I_2} |Q(\xi)|^2 d\tilde{\mu}(\xi) &\leq C^2 e^{2\varepsilon L} \lambda_L(\xi_0, V) \int_{I_2} |G(\xi)|^{4N} |P(\xi)|^2 d\tilde{\mu}(\xi) \\ &\leq C^2 C_2 e^{2\varepsilon L} \lambda_L(\xi_0, V) \int_{I_2} \frac{\xi^{2(n+1)}}{(\xi - \xi_0)^{2N}} d\tilde{\mu}(\xi) \\ &\leq C_3 e^{2\varepsilon L} \lambda_L(\xi_0, V) 2^{-2N}.\end{aligned}$$

Combining the integrals over I_1, I_2, I_3 and using (5.6) yields the claim. □

Remark. In the proof we have $d_0 = \min \tilde{\mathsf{E}}_{\delta_1}$ and thus d_0 depends on ε. By the definition of G via F in Lemma 5.1 this shows that $\gamma < 1$, which is the maximum of G outside of $(\xi_0 - \delta_2, \xi_0 - \delta_2)$, also depends on ε. However, for fixed δ_2 we have

(5.8) $$\gamma_1 = \sup_{\varepsilon \in (0,1)} \gamma(\varepsilon, \delta_2) < 1.$$

This remains true, if $\xi_0 \in [a, b] \subset \mathrm{int}(\mathsf{E})$. This will be important in the following.

We are now ready to prove Theorem 1.1:

Proof of Theorem 1.1. Assume that $\mathsf{E} = \sigma_{\mathrm{ess}}(\mu)$ and μ is regular and let E_r be the extension as defined above. Clearly E_r is a finite gap set and E_r decreases monotonically to E. We conclude from Lemma B.2 that $d\rho_{\mathsf{E}_r}$ is absolutely continuous on I and that the densities $f_{\mathsf{E}_r}(\xi)$ increase with r and are bounded above by $f_{\mathsf{E}}(\xi)$. We claim that

(5.9) $$\lim_{r\to\infty} f_{\mathsf{E}_r}(\xi) = f_{\mathsf{E}}(\xi)$$

uniformly on I. By monotonicity and boundedness for every ξ,

$$\lim_{r\to\infty} f_{\mathsf{E}_r}(\xi) = g(\xi)$$

exists and since f_{E_r} are in particular continuous the convergence is uniform and g is continuous. On the other hand, using the upper envelope theorem, we conclude as in [12, Lemma 6.1] that $M_{\mathsf{E}_r} \to M_{\mathsf{E}}$, which implies $\rho_{\mathsf{E}_r} \to \rho_{\mathsf{E}}$ in the weak-$*$ sense. We conclude that $g = f_{\mathsf{E}}$ on I.

For fixed r let V_r be a finite gap potential as discussed in Section 4 and μ_r its spectral measure. Note that V_r is Stahl–Totik regular. We will apply Theorem 5.3 with $\mu = \mu_r$ and $\tilde{\mu} = \mu$. Since $\mathsf{E} \subset \mathsf{E}_r$, (5.4) is satisfied for arbitrary $\varepsilon > 0$. Let $\varepsilon > 0$ be fixed, and δ_2, D be as in Theorem 5.3 and γ_1 be defined by (5.8). Then

$$\frac{\lambda_M(\xi_0, V)}{\lambda_L(\xi_0, V_r)} \leq \sup_{|\xi - \xi_0| < \delta_1} \frac{f_\mu(t)}{f_{\mu_r}(t)} + D\gamma_1^{4N} e^{2\varepsilon L} + De^{2\varepsilon L} 2^{-2N},$$

where $M = L + 2Nc$. Choose η so that $\max\{\gamma_1, 1/\sqrt{2}\}^\eta \le e^{-1}$ and $N = \eta\varepsilon L$. Note that by the definition of γ_1, η does not depend on ε. It follows that

$$D\gamma_1^{4N} e^{2\varepsilon L} + De^{2\varepsilon L} 2^{-2N} = O(e^{-2\varepsilon L}).$$

Thus,

$$\limsup_{M\to\infty} \frac{\lambda_M(\xi_0, V)}{\lambda_L(\xi_0, V_r)} \le \sup_{|\xi-\xi_0|<\delta_1} \frac{f_\mu(\xi)}{f_{\mu_r}(\xi)}.$$

Since $M = L(1 + 2\eta\varepsilon c)$ we obtain by (4.3)

$$\lim_{L\to\infty} \frac{\lambda_L(\xi_0, V_r)}{\lambda_{(1+2\eta\varepsilon c)L}(\xi_0, V_r)} = 1 + 2\eta\varepsilon c.$$

Therefore,

$$\limsup_{M\to\infty} \frac{\lambda_M(\xi_0, V)}{\lambda_M(\xi_0, V_r)} \le \sup_{|\xi-\zeta_0|<\delta_1} \frac{f_\mu(\xi)}{f_{\mu_r}(\xi)}(1 + 2\eta\varepsilon c).$$

Taking first $\varepsilon \to 0$ and using that $\varepsilon c(\varepsilon) \to 0$ by (5.3) and then $\delta_1 \to 0$ we get

$$\limsup_{M\to\infty} \frac{\lambda_M(\xi_0, V)}{\lambda_M(\xi_0, V_r)} \le \frac{f_\mu(\xi_0)}{f_{\mu_r}(\xi_0)}. \tag{5.10}$$

Since, on the other hand, E_r is a finite gap set we have

$$\lim_{M\to\infty} M\lambda_M(\xi_0, V_r) = \frac{f_{\mu_r}(\xi_0)}{f_{\mathsf{E}_r}(\xi_0)}.$$

Plugging this into (5.10) yields

$$\limsup_{M\to\infty} M\lambda_M(\zeta_0, V) \le \frac{f_\mu(\xi_0)}{f_{\mathsf{E}_r}(\xi_0)}.$$

By (5.9) sending $r \to 0$ we conclude that

$$\limsup_{M\to\infty} M\lambda_M(\zeta_0, V) \le \frac{f_\mu(\xi_0)}{f_{\mathsf{E}}(\xi_0)}. \tag{5.11}$$

To get the opposite inequality we would like to switch the roles of V and V_r. It can be seen from the proof of Theorem 5.3 that in this case we can even take $\delta_1 = 0$, since μ_r only has finitely many point masses outside E_r. In Theorem 5.3 we have to estimate the eigensolutions of H_V on E_r. This shows that for fixed $r > 0$, we cannot take $\varepsilon \to 0$. However, for fixed $r > 0$ we find $\varepsilon(r)$ and note that $\varepsilon(r) \to 0$ as $r \to 0$.

Now switching the roles of V and V_r we get

$$\frac{\lambda_M(\xi_0, V_r)}{\lambda_L(\xi_0, V)} \le \sup_{|\xi-\xi_0|<\delta_1} \frac{f_{\mu_r}(\xi)}{f_\mu(\xi)} + D\gamma_1^{4N} e^{2\varepsilon L} + De^{2\varepsilon L} 2^{-2N},$$

where $N = \eta\varepsilon L$ and, as before,

$$\limsup_{L\to\infty} \frac{\lambda_L(\xi_0, V_r)}{\lambda_L(\xi_0, V)} \leq \frac{f_{\mu_r}(\xi_0)}{f_\mu(\xi_0)} \frac{1}{1+2\eta\varepsilon c}$$

and

$$\limsup_{L\to\infty} \frac{1}{L\lambda_L(\zeta_0, V)} \leq \frac{f_{\mathsf{E}_r}(\xi_0)}{f_\mu(\xi_0)} \frac{1}{1+2\eta\varepsilon c}.$$

But now for fixed r we cannot take $\varepsilon \to 0$. However, since $\varepsilon(r) \to 0$ as $r \to 0$ we get by (5.9)

$$\limsup_{L\to\infty} \frac{1}{L\lambda_L(\xi_0, V)} \leq \frac{f_{\mathsf{E}}(\zeta_0)}{f_\mu(\xi_0)}$$

and thus

$$\frac{f_\mu(\xi_0)}{f_{\mathsf{E}}(\zeta_0)} \leq \liminf_{L\to\infty} L\lambda_L(\xi_0, V). \tag{5.12}$$

Combining (5.11) and (5.12) and noting that all the arguments are uniform in $\xi_0 \in I$ yields the claim. □

Proof of Theorem 1.2. Let m be the Weyl m-function associated to H_V. We showed in Section 3 that H_V can be written as a canonical system. Thus, by [13, Theorem 9] it follows that if for $\xi \in I$ the limit

$$\frac{1}{\pi} \lim_{z\to\xi} \operatorname{Im} m(z) = f_\mu(\xi) \in (0, \infty) \tag{5.13}$$

exists non-tangentially, then

$$\lim_{L\to\infty} \lambda_L(\xi) K_L(\xi + \lambda_L(\xi)z, \xi + \lambda_L(\xi)w) = \frac{\sin(\pi f_\mu(\xi)(\overline{w} - z))}{\pi f_\mu(\xi)(\overline{w} - z)} \tag{5.14}$$

uniformly for $\xi \in I$ and z, w in compact subsets of $\mathbb{C}$. It follows from properties of Poisson integrals [38, Theorem 11.22, Theorem 11.23] that (5.13) holds on I under the assumptions of Theorem 1.2. From Theorem 1.1 we conclude that

$$\lim_{L\to\infty} L\lambda_L(\xi) = \frac{f_\mu(\xi)}{f_{\mathsf{E}}(\xi)}.$$

The claim follows from continuity of the sinc kernel and the fact that (5.14) holds uniformly for z, w in compact subsets of $\mathbb{C}$. □

We finish this section with the proof of Theorem 1.3. There are several proofs in the orthogonal polynomials case that show how to conclude from universality clock spacing for the zeros of orthogonal polynomials, which only use interlacing properties of the zeros of orthogonal polynomials; cf. [16, 25, 39] The same proof carries over to the setting of continuum Schrödinger operators (or even canonical systems) without any change. We supply the proof for the reader's convenience.

Proof of Theorem 1.3. We start by a well known fact: The function

$$m_L(z) = \frac{v(L,z)}{v'(L,z)}$$

is a Herglotz function and since $v(L,\cdot)$, $v'(L,\cdot)$ are entire, the measure in its integral representation is purely discrete and supported at the zeros of $v'(L,z)$. Since m_L is increasing between poles, the zeros of $v(L,z)$ and $v'(L,z)$ interlace.

By (3.4) and the fact that $v(L,\cdot)$ and $v'(L,\cdot)$ cannot vanish simultaneously, we see that for $z \neq \overline{w}$, $K_L(z,w) = 0$ if and only if $m_L(z) = m_L(\overline{w})$ (the value of $m_L(z)$ can also be ∞ corresponding to a zero of $v'(L,z)$). Fix $\xi \in I$ and define

$$\tilde{f}_L(z) = \frac{K_L(\xi, \xi + \frac{z}{Lf_{\mathsf{E}}(\xi)})}{K_L(\xi,\xi)}.$$

Let $\cdots < \tilde{z}_{-1}^L < \tilde{z}_0^L = 0 < \tilde{z}_1^L < \cdots$ denote the zeros of $\tilde{f}_L(z)$. By (1.11) we see that $\tilde{z}_{\pm 1}^L \to \pm 1$ and inductively we get

$$\tilde{z}_{\pm j}^L \to \pm j.$$

Let $0 \le z_0^L < \tilde{z}_1^L$ be so that $\xi + \frac{z_0^L}{f_{\mathsf{E}}(x_0)L}$ is the first pole of m_L to the right of ξ. Set

$$f_L(z) = \frac{K_L(\xi + \frac{z_0^L}{Lf_{\mathsf{E}}(\xi)}, \xi + \frac{z_0^L + z}{Lf_{\mathsf{E}}(\xi)})}{K_L(\xi,\xi)}.$$

If we denote the zeros of f_L by $\cdots < z_{-1}^L < z_0^L = 0 < z_1^L < \cdots$ we see as before that

$$z_{\pm j}^L \to \pm j$$

By our definition of z_0^L, $\xi_j^L = \xi + \frac{z_0^L + z_j^L}{f_{\mathsf{E}}(x_0)L}$ are the zeros of $v'(L,z)$, which finishes the proof. Since the convergence in (1.11) is uniform this shows that (1.12) holds uniformly on I. □

Appendix A de Branges spaces and canonical systems

It has already been realized in [27] and also in [13, 8] that the theory of canonical systems is useful for the understanding of universality limits for Christoffel–Darboux kernels. The inverse theory developed by de Branges is based on a theory of Hilbert spaces of entire functions [11]. These spaces are called de Branges spaces. We recall some part of the general theory, to highlight that the objects discussed in Section 3 are only a special case of this rich theory. We follow the presentation in [37].

For an entire function F we denote $F^{\#}(z) = \overline{F(\overline{z})}$. Moreover, let $H^2 = H^2(\mathbb{C}_+)$ denote the standard Hardy space of the upper half-plane.

Definition A.1. A Hermite–Biehler function is an entire function E with no zeros in $\mathbb{C}_+$ satisfying $|E(z)| \geq |E^{\#}(z)|$ for $z \in \mathbb{C}_+$. Given a Hermite–Biehler function E we define the de Branges space $B(E)$ as

$$B(E) = \{F \mid F \text{ entire}, F/E, F^{\#}/E \in H^2\}.$$

In the following it will be useful to decompose E into its real and imaginary part. We define $a(z) = \frac{1}{2}(E(z) + E^{\#}(z))$ and $c(z) = \frac{1}{2i}(E^{\#}(z) - E(z))$. The notation a and c and the unexpected minus sign in the definition of c will become clear below.

Example A.2. The motivating example for de Branges was the Hermite-Biehler function $E_a(z) = e^{-iaz}$ for $a > 0$, in which case $B(E_a)$ denotes the standard Paley–Wiener space of entire square-integrable functions of exponential type at most a.

Theorem A.3 ([37, Theorem 4.4])**.** *Let E be a Hermite–Biehler function. Then $B(E)$ becomes a Hilbert space when endowed with the scalar product*

$$\langle F, G\rangle_{B(E)} = \frac{1}{\pi}\int_{-\infty}^{\infty} F(t)\overline{G(t)}\frac{dt}{|E(t)|^2}.$$

$B(E)$ is a reproducing kernel Hilbert space with reproducing kernel:

$$K_E(z, w) = \frac{\overline{E(w)}E(z) - \overline{E^{\#}(w)}E^{\#}(z)}{2i(\overline{w} - z)} = \frac{a(z, x)\overline{c(w, x)} - c(z, x)\overline{a(w, x)}}{\overline{w} - z}.$$

De Branges spaces arise naturally when discussing canonical systems:

Definition A.4. Consider matrix-valued functions $A, B : [0, N) \to \mathrm{Mat}(2, \mathbb{C})$ for some $N > 0$ or $N = \infty$ which are locally integrable in the sense that their entries are in $L^1([0, x])$ for all $x < N$, and have the property that $A(x) \geq 0$, $B(x)^* = B(x)$ and $\operatorname{tr} A(x)j = \operatorname{tr} B(x)j = 0$ for Lebesgue-a.e. $x \in [0, N)$, with j defined below. A canonical system is a differential equation of the form

$$j\partial_x y(x, z) = (-zA(x) + B(x))y(x, z), \quad j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

A solution $T : [0, N) \times \mathbb{C} \to \mathrm{Mat}(2, \mathbb{C})$ which satisfies the initial value problem

$$j\partial_x T(x, z) = (-zA(x) + B(x))T(x, z), \qquad T(0, z) = I_2$$

is called the transfer matrix of the canonical system. It is a fundamental object in the theory of canonical systems. Differentiating $T(s, w)^* jT(s, z)$ we see that

$$T(x, w)^* jT(x, z) - j = (\overline{w} - z) \int_0^x T(s, w)^* A(s)T(s, z)ds. \tag{A.1}$$

The expression on the left-hand side above is called the j form of T [3, 21]. For fixed x as a function of z it is entire and satisfies

$$i\left(T(x, z)^* jT(x, z) - j\right) = \begin{cases} \geq 0, & \text{if } z \in \mathbb{C}_+, \\ = 0, & \text{if } z \in \mathbb{R}. \end{cases}$$

We say that T is j-expanding in $\mathbb{C}_+$ and j-unitary on $\mathbb{R}$. Moreover, from (A.1) it also follows that for $z \in \mathbb{C}_+$, T satisfies the j-monotonicity property

$$i(T^*(x_2, z)jT(x_2, z) - T^*(x_1, z)jT(x_1, z)) \geq 0 \tag{A.2}$$

for $x_2 \geq x_1$. Thus, $\{T(x, z)\}_{x\in[0,N)}$ forms a j-monotonic family of entire matrix functions. It can be shown that vice-versa to every such family one can associate a canonical system [10, Remark 2.3]. The j-form is invariant under multiplying T from the left by some $U \in \mathrm{SL}(2, \mathbb{R})$, since such U satisfies $U^* jU = j$. This gives a certain freedom, which is called gauge freedom. A common gauge normalization, which was used by Potapov and de Branges, is to assume that for all $x \in [0, N)$, $T(x, 0) = I$, which leads on the level of canonical systems to $B = 0$. In this case A is usually denoted by H and called the Hamiltonian of the system. We call this the Potapov–de Branges gauge. If T is normalized arbitrarily, then passing to $\tilde{T}(x, z) = T(x, 0)^{-1}T(x, z)$ we obtain a transfer matrix in the Potapov–de Branges gauge. The corresponding Hamiltonian is given by

$$H(x) = T(x, 0)^* A(x)T(x, 0).$$

In particular, any canonical system obtained from a Schrödinger equation as in (3.2) can be rewritten into the Potapov–de Branges gauge. We found it more convenient to provide this gauge independent presentation and work directly with (3.2).

Let us write

$$T(x, z) = \begin{pmatrix} a(z, x) & b(z, x) \\ c(z, x) & d(z, x) \end{pmatrix}.$$

It follows from (A.2) that for fixed x

$$m(z, x) = -\frac{a(z, x)}{c(z, x)}$$

is a generalized Herglotz function, i.e., the map $z \mapsto m(z,x)$ maps $\mathbb{C}_+$ analytically into $\overline{\mathbb{C}_+}$, where $\overline{\mathbb{C}_+} = \mathbb{C}_+ \cup \mathbb{R} \cup \{\infty\}$; see [37, Lemma 4.15].[3]

Using this we see that

$$E(z,x) := a(z,x) - ic(z,x) \tag{A.3}$$

satisfies

$$\left|\frac{E^\#(z,x)}{E(z,x)}\right| = \left|\frac{a(z,x)+ic(z,x)}{a(z,x)-ic(z,x)}\right| = \left|\frac{-\frac{a(z,x)}{c(z,x)} - i}{-\frac{a(z,x)}{c(z,x)} + i}\right| \le 1$$

and thus $E(z,x)$ is a Hermite–Biehler function. This also explains in hindsight the minus sign in the definition of c. Dividing by $(\overline{w} - z)$ in (A.1) the (1,1)-entry gives exactly the reproducing kernel

$$K_{E_x}(z,w) = \frac{a(z,x)\overline{c(w,x)} - c(z,x)\overline{a(w,x)}}{\overline{w} - z} \tag{A.4}$$

$$= \begin{pmatrix}1\\0\end{pmatrix}^* \int_0^x T(s,w)^* A(s) T(s,z)\,ds \begin{pmatrix}1\\0\end{pmatrix}, \tag{A.5}$$

where $E_x(z) = E(z,x)$.

The canonical system is said to be limit point at N if

$$\int_0^N \operatorname{tr} T(s,0)^* A(s) T(s,0)\,ds = \infty.$$

Due to (A.2), Weyl disks can be introduced in this setting. By a standard abuse of notation, we will use the same notation for an $\mathrm{SL}(2,\mathbb{C})$ matrix and for the Möbius transformation it generates on the Riemann sphere $\hat{\mathbb{C}}$, with the standard projective identification of $w \in \mathbb{C}$ with the coset of $\binom{w}{1}$ and ∞ with the coset of $\binom{1}{0}$. For any $z \in \mathbb{C}_+$, the Weyl disks are defined by

$$D(x,z) = \{ w \in \hat{\mathbb{C}} \mid T(x,z)w \in \overline{\mathbb{C}_+} \}.$$

Due to (A.2), the Weyl disks are nested, $D(x_2,z) \subset D(x_1,z)$ for $x_1 \le x_2$. Thus, for each $z \in \mathbb{C}_+$, the intersection $\bigcap_{0 \le x < N} D(x,z)$ is a disk or a point. The assumption of being limit point at N exactly means that this intersection is a point. In this case, the Weyl disks define an analytic map $m : \mathbb{C}_+ \to \overline{\mathbb{C}_+}$ by

$$\{m(z)\} = \bigcap_{0 \le x < N} D(x,z).$$

Note that since $\mathrm{SL}(2,\mathbb{R})$ matrices leave $\overline{\mathbb{C}_+}$ invariant, also $D(x,z)$ and thus m do not depend on the gauge normalization.

[3]This definition differs slightly from the definition of Herglotz functions given in the introduction, since we allow values in $\overline{\mathbb{C}_+}$ rather than in $\mathbb{C}_+$. By the maximum principle, an analytic function that attains a value in $\mathbb{R} \cup \{\infty\}$ must already be constant and therefore the set of generalized Herglotz functions consists of Herglotz functions and constant functions with values in $\mathbb{R} \cup \{\infty\}$.

Appendix B Martin measure

Let E be a semibounded set so that for any Martin function for the domain $\mathbb{C} \setminus \mathsf{E}$ and the point ∞

$$\lim_{z \to -\infty} \frac{M(z)}{\sqrt{-z}} > 0.$$

Sets with this property are called Akhiezer–Levin sets. Let again M_{E} be normalized so that the limit is equal to 1. Since M_{E} vanishes q.e. on E it can be extended to a subharmonic function on $\mathbb{C}$. Let ρ_{E} be the associated Riesz measure, defined by

$$\rho_{\mathsf{E}} := \frac{1}{2\pi} \Delta M_{\mathsf{E}}.$$

Since M_{E} is a harmonic function in $\mathbb{C}_+$, we find an analytic function, Θ_{E}, with $\operatorname{Im} \Theta_{\mathsf{E}} = M_{\mathsf{E}}$ and since M_{E} is positive, Θ_{E} is a Herglotz function. Moreover, it can be shown that also $i\Theta'_{\mathsf{E}}$ is a Herglotz function and the measure in its integral representation is exactly ρ_{E}:

Lemma B.1 ([12, Lemma 2.3])**.** *The measure ρ_{E} is such that*

$$\int_{\mathbb{R}} \frac{d\rho_{\mathsf{E}}(t)}{1+|t|} < \infty.$$

Moreover, $i\Theta'_{\mathsf{E}}$ is a Herglotz function and we have

$$i\Theta'_{\mathsf{E}}(z) = \int_{\mathbb{R}} \frac{d\rho_{\mathsf{E}}(t)}{t-z}.$$

Recall that $d\rho_{\mathsf{E}}(t) = f_{\mathsf{E}} dt + d\rho_{E,s}(t)$ denotes the Lebesgue decomposition of ρ_{E}.

Lemma B.2. *Let $(a, b) = I \subset \mathsf{E}^\circ$. Then $\rho_{\mathsf{E}}(\xi)|_I$ is absolutely continuous and $f_{\mathsf{E}}(\xi)$ is real analytic. Moreover, if $I \subset \mathsf{E}_1 \subset \mathsf{E}_2$ we have*

$$f_{\mathsf{E}_2}(\xi)|_I \leq f_{\mathsf{E}_1}(\xi)|_I. \tag{B.1}$$

Proof. Since I only contains Dirichlet regular points by [34, Theorem 4.2.2], for every $\xi \in I$, $\lim_{z\to\xi} M_{\mathsf{E}}(z) = 0$, by [12, Theorem 2.1]. Since Θ_{E} is a Herglotz function, this implies that it can be analytically extended through I. Hence, also $\Theta'_{\mathsf{E}}(z)$ has an analytic extension through I. Since $i\Theta'_{\mathsf{E}}$ is a Herglotz function, this shows that the measure in its integral representation is purely absolutely continuous and moreover,

$$\frac{1}{\pi} \operatorname{Re} \Theta'_{\mathsf{E}}(\xi) = f_{\mathsf{E}}(\xi).$$

Since $M_{\mathsf{E}}|_I = 0$, $\operatorname{Im} \Theta'_{\mathsf{E}}(\xi) = 0$ and we conclude that f_{E} is real analytic on I.

If $\mathsf{E}_1 \leq \mathsf{E}_2$ we have that $M_{\mathsf{E}_2} \leq M_{\mathsf{E}_1}$ as follows, e.g., from [12, Lemma 2.7]. Since by the Cauchy–Riemann equations we get

$$\pi f_{\mathsf{E}}(\xi) = \lim_{y\to 0} \frac{M_{\mathsf{E}}(\xi + iy)}{y}$$

and we conclude (B.1).

Open access funding provided by Austrian Science Fund (FWF).

References

[1] N. I. Ahiezer and B. Ja. Levin, *Generalization of S.N. Bernšteĭn's inequality for derivatives of entire functions*, *Issledovanija po Sovremennym Problemam Teorii Funkciĭ Kompleksnogo Peremennogo*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1960, pp. 111–165.

[2] D. H. Armitage and S. J. Gardiner, *Classical Potential Theory*, Springer, London, 2001.

[3] D. Z. Arov and H. Dym, *J-Contractive Matrix Valued Functions and Related Topics*, Cambridge University Press, Cambridge, 2008.

[4] A. Avila, Y. Last and B. Simon, *Bulk universality and clock spacing of zeros for ergodic Jacobi matrices with absolutely continuous spectrum*, Anal. PDE **3** (2010), 81–108.

[5] V. Azarin, *Growth Theory of Subharmonic Functions*, Birkhäuser, Basel, 2009.

[6] R. Bessonov, *Entropy function and orthogonal polynomials*, J. Approx. Theory **272** (2021), Article no. 105650.

[7] I. Binder, D. Damanik, M. Goldstein and M. Lukic, *Almost periodicity in time of solutions of the KdV equation*, Duke Math. J. **167** (2018), 2633–2678.

[8] J. Breuer, *Scaling limits of Jacobi matrices and the Christoffel–Darboux kernel*, Constr. Approx. **53** (2021), 349–379.

[9] J. Breuer and D. Weissman, *Level repulsion for Schrödinger operators with singular continuous spectrum*, J. Spectr. Theory **9** (2019), 429–451.

[10] D. Damanik, B. Eichinger, and P. Yuditskii, *Szego's theorem for canonical systems: the Arov gauge and a sum rule*, J. Spectr. Theory **11** (2021), 1255–1277.

[11] L. de Branges, *Hilbert Spaces of Entire Functions*, Prentice-Hall, Englewood Cliffs, NJ, 1968.

[12] B. Eichinger and M. Lukić, *Stahl–Totik regularity for continuum Schrödinger operators*, arXiv:2001.00875 [math.SP]

[13] B. Eichinger, M. Lukić, and B. Simanek, *An approach to universality using Weyl m-functions*, arXiv:2108.01629 [math.CA]

[14] B. Eichinger, T. VandenBoom and P. Yuditskii, *KdV hierarchy via abelian coverings and operator identities*, Trans. Amer. Math. Soc. Ser. B **6** (2019), 1–44.

[15] A. Eremenko and P. Yuditskii, *Comb functions*, in *Recent Advances in Orthogonal Polynomials, Special Functions, and Their Applications*, American Mathematical Society, Providence, RI, 2012, pp. 99–118.

[16] G. Freud, *Orthogonal Polynomials*, Pergamon Press, New York, 1971.

[17] S. J. Gardiner and T. Sjödin, *Potential theory in Denjoy domains*, in *Analysis and Mathematical Physics*, Birkhäuser, Basel, 2009, pp. 143–166.

[18] F. Gesztesy and H. Holden, *Soliton Equations and Their Algebro-Geometric Solutions. Vol. I*, Cambridge University Press, Cambridge, 2003,

[19] F. Gesztesy and P. Yuditskii, *Spectral properties of a class of reflectionless Schrödinger operators*, J. Funct. Anal. **241** (2006), 486–527.

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

[21] L. Golinskii and I. Mikhailova, *Hilbert spaces of entire functions as a J theory subject [Preprint No. 28-80, Inst. Low Temp. Phys. Engrg., Kharkov, 1980]*, in *Topics in Interpolation Theory (Leipzig, 1994)*, Birkhäuser, Basel, 1997, pp. 205–251.

[22] P. Gubkin, *Mate–Nevai–Totik theorem for Krein systems*, Integral Equations Operator Theory **93** (2021), Article no. 33.

[23] P. W. Jones and D. E. Marshall, *Critical points of Green's function, harmonic measure, and the corona problem*, Ark. Mat. **23** (1985), 281–314.

[24] Y. Last and B. Simon, *Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*, Invent. Math. **135** (1999), 329–367.

[25] E. Levin and D. S. Lubinsky, *Applications of universality limits to zeros and reproducing kernels of orthogonal polynomials*, J. Approx. Theory **150** (2008), 69–95.

[26] D. S. Lubinsky, *A new approach to universality limits involving orthogonal polynomials*, Ann. of Math. (2) **170** (2009), 915–939.

[27] D. S. Lubinsky, *Universality limits for random matrices and de Branges spaces of entire functions*, J. Funct. Anal. **256** (2009), 3688–3729.

[28] A. Maltsev, *Universality limits of a reproducing kernel for a half-line Schrödinger operator and clock behavior of eigenvalues*, Comm. Math. Phys. **298** (2010), 461–484.

[29] V. A. Marchenko, *Sturm–Liouville Operators and Applications*, AMS Chelsea Publishing, Providence, RI, 2011.

[30] V. A. Marčenko and I. V. Ostrovskii, *A characterization of the spectrum of the Hill operator*, Mat. Sb. (N.S.) **97(139)** (1975), 540–606, 633–634.

[31] A. Máté, P. Nevai and V. Totik, *Szegő's extremum problem on the unit circle*, Ann. of Math. (2) **134** (1991), 433–453.

[32] M. Mitkovski, *Universality limits for entire functions*, Proc. Amer. Math. Soc. **141** (2013), 3119–3124.

[33] J. Pöschel and E. Trubowitz, *Inverse Spectral Theory*, Academic Press, Boston, MA, 1987.

[34] T. Ransford, *Potential Theory in the Complex Plane*, Cambridge University Press, Cambridge, 1995.

[35] C. Remling, *Schrödinger operators and de Branges spaces*, J. Funct. Anal. **196** (2002), 323–394.

[36] C. Remling, *Inverse spectral theory for one-dimensional Schrödinger operators: the A function*, Math. Z. **245** (2003), 597–617.

[37] C. Remling, *Spectral Theory of Canonical Systems*, De Gruyter, Berlin, 2018.

[38] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

[39] B. Simon, *The Christoffel–Darboux kernel*, in *Perspectives in Partial Differential Equations, Harmonic Analysis and Applications*, American Mathematical Society, Providence, RI, 2008, pp. 295–335.

[40] B. Simon, *Two extensions of Lubinsky's universality theorem*, J. Anal. Math. **105** (2008), 345–362.

[41] M. Sodin and P. Yuditskii, *Almost periodic Sturm–Liouville operators with Cantor homogeneous spectrum*, Comment. Math. Helv. **70** (1995), 639–658.

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

[43] V. Totik, *Asymptotics for Christoffel functions for general measures on the real line*, J. Anal. Math. **81** (2000), 283–303.

[44] V. Totik, *Universality and fine zero spacing on general sets*, Ark. Mat. **47** (2009), 361–391.

Benjamin Eichinger
INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING
VIENNA UNIVERSITY OF TECHNOLOGY
WIEN, A-1040, AUSTRIA
email: benjamin.eichinger@tuwien.ac.at

and
FLYDE COLLEGE
LANCASTER UNIVERSITY
LANCASTER LA1 4YF, UNITED KINGDOM

(Received July 5, 2022)