COEFFICIENTS OF ORTHOGONAL POLYNOMIALS ON THE UNIT CIRCLE AND HIGHER ORDER SZEGŐ THEOREMS

LEONID GOLINSKII* AND ANDREJ ZLATOŠ†

Abstract. Let μ be a non-trivial probability measure on the unit circle $\partial\mathbb{D}$, w the density of its absolutely continuous part, α_n its Verblunsky coefficients, and Φ_n its monic orthogonal polynomials. In this paper we compute the coefficients of Φ_n in terms of the α_n. If the function $\log w$ is in $L^1(d\theta)$, we do the same for its Fourier coefficients. As an application we prove that if $\alpha_n \in \ell^4$ and $Q(z) \equiv \sum_{m=0}^N q_m z^m$ is a polynomial, then with $\bar{Q}(z) \equiv \sum_{m=0}^N \bar{q}_m z^m$ and S the left shift operator on sequences we have

$$\left|Q(e^{i\theta})\right|^2 \log w(\theta) \in L^1(d\theta) \;\Leftrightarrow\; \{\bar{Q}(S)\alpha\}_n \in \ell^2$$

We also study relative ratio asymptotics of the reversed polynomials $\Phi_{n+1}^*(\mu)/\Phi_n^*(\mu) - \Phi_{n+1}^*(\nu)/\Phi_n^*(\nu)$ and provide a necessary and sufficient condition in terms of the Verblunsky coefficients of the measures μ and ν for this difference to converge to zero uniformly on compact subsets of $\mathbb{D}$.

1. Introduction

In the present paper we study certain aspects of the theory of orthogonal polynomials on the unit circle (OPUC). For background information on the subject we refer the reader to the texts [6, 17, 18, 21]. Throughout, $d\mu$ will be a non-trivial (i.e., with infinite support) probability measure on the unit circle $\partial\mathbb{D}$ in $\mathbb{C}$, identified with the interval $[0, 2\pi)$ via the map $\theta \mapsto e^{i\theta}$. We will write

$$d\mu(\theta) = w(\theta)\,\frac{d\theta}{2\pi} + d\mu_{\mathrm{sing}}(\theta)$$

with $d\theta$ the Lebesgue measure on $[0, 2\pi)$ and $d\mu_{\mathrm{sing}}$ the singular part of $d\mu$.

One usually denotes by

$$\Phi_n(z) = \kappa_{n,n} z^n + \kappa_{n,n-1} z^{n-1} + \cdots + \kappa_{n,1} z + \kappa_{n,0} \tag{1.1}$$

the *monic* (i.e., $\kappa_{n,n} = 1$) *orthogonal polynomials* for μ (with $n \geq 0$). It is standard to define the *reversed polynomials* by

$$\begin{aligned}\Phi_n^*(z) &= \lambda_{n,n} z^n + \lambda_{n,n-1} z^{n-1} + \cdots + \lambda_{n,1} z + \lambda_{n,0} \\ &\equiv \bar{\kappa}_{n,0} z^n + \bar{\kappa}_{n,1} z^{n-1} + \cdots + \bar{\kappa}_{n,n-1} z + \bar{\kappa}_{n,n}\end{aligned}$$

* The work of the first author was supported in part by INTAS Research Network NeCCA 03-51-6637.
† Corresponding author.

and let $\kappa_{n,m} = \lambda_{n,m} = 0$ whenever $m > n$. We have $\Phi_0 \equiv \Phi_0^* \equiv 1$ and for $n \geq 0$ the recurrence relations

$$\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n \Phi_n^*(z) \tag{1.2}$$

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z \Phi_n(z) \tag{1.3}$$

with $\alpha_n \in \mathbb{D}$ the *Verblunsky coefficients* of μ. A fundamental result of Verblunsky [22] says that there is a one-to-one correspondence between non-trivial probability measures μ on $\partial\mathbb{D}$ and sequences $\{\alpha_n\}_{n\geq 0} \in \mathbb{D}^{\mathbb{Z}_0^+}$. If we set $\Phi_n \equiv 0$ and $\Phi_n^* \equiv 1$ for $n \leq -1$, and

$$\alpha_{-1} \equiv -1, \qquad \alpha_n \equiv 0 \quad (n \leq -2)$$

then (1.2), (1.3) hold for all $n \in \mathbb{Z}$. We accordingly let $\kappa_{n,m} = 0$ and $\lambda_{n,m} = \delta_{m,0}$ when $n < 0$ and $m \geq 0$.

Probably the most famous OPUC result is *Szegő's Theorem.* In the form proved by Verblunsky [22] it says that $\alpha_n \in \ell^2(\mathbb{Z}_0^+)$ if and only if $\log w(\theta) \in L^1(d\theta)$. More precisely, the *sum rule*

$$\sum_{n=0}^{\infty} \log(1 - |\alpha_n|^2) = \int \log(w(\theta)) \, \frac{d\theta}{2\pi} \tag{1.4}$$

holds. Note that both sides of (1.4) are indeed non-positive since $|\alpha_n| < 1$ and by Jensen's inequality, $\int \log(w(\theta)) \frac{d\theta}{2\pi} \leq \log(\int w(\theta) \frac{d\theta}{2\pi}) \leq \log(\mu(\partial\mathbb{D})) = 0$, but they can simultaneously be $-\infty$. Recently the area of sum rules, for orthogonal polynomials as well as Schrödinger operators, saw a rapid development starting with papers by Deift-Killip [2] and Killip-Simon [9], which were followed by many others (e.g., [3, 10, 11, 12, 14, 16, 20, 23, 24]).

If $\alpha_n \in \ell^2$, one defines the *Szegő function*

$$D(z) \equiv \exp\left(\int \frac{e^{i\theta} + z}{e^{i\theta} - z} \log w(\theta) \, \frac{d\theta}{4\pi} \right)$$

which is analytic in $\mathbb{D}$. Szegő's Theorem in its full extent also shows that then

$$\frac{\Phi_n^*}{\|\Phi_n^*\|_{L^2(d\mu)}} \to D^{-1} \tag{1.5}$$

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

$$\|\Phi_n^*\|_{L^2(d\mu)} = \prod_{k=0}^{n-1} (1 - |\alpha_k|)^{1/2} = \prod_{k=0}^{n-1} \rho_k \tag{1.6}$$

where $\rho_k \equiv \sqrt{1 - |\alpha_k|^2}$ (see (1.5.13) in [17]), and so if we define d_m by

$$D(z)^{-1} \equiv \left(\prod_{k\geq 0} \rho_k \right)^{-1} (1 + d_1 z + d_2 z^2 + \dots) \tag{1.7}$$

then

$$d_m = \lim_{n\to\infty} \lambda_{n,m} \tag{1.8}$$

The first contribution of this paper is the following expression of the coefficients $\kappa_{n,m}$, $\lambda_{n,m}$, and d_m in terms of the α_k. To the best of our knowledge (and to our surprise), this result is new despite the long history and classical nature of the subject!

Theorem 1.1. *For $m \geq 1$,*

$$\bar{\kappa}_{n,n-m} = \lambda_{n,m} = \sum_{\substack{\sum_1^j a_l = m \\ j, a_l \geq 1}} \sum_{\substack{k_1 < n \\ k_2 < k_1 - a_1 \\ \dots \\ k_j < k_{j-1} - a_{j-1}}} \alpha_{k_1} \bar{\alpha}_{k_1 - a_1} \dots \alpha_{k_j} \bar{\alpha}_{k_j - a_j} \tag{1.9}$$

If $\alpha_k \in \ell^2$, then also

$$d_m = \sum_{\substack{\sum_1^j a_l = m \\ j, a_l \geq 1}} \sum_{\substack{k_2 < k_1 - a_1 \\ \dots \\ k_j < k_{j-1} - a_{j-1}}} \alpha_{k_1} \bar{\alpha}_{k_1 - a_1} \dots \alpha_{k_j} \bar{\alpha}_{k_j - a_j} \tag{1.10}$$

Remarks. 1. In the above sums $[a_1, a_2, \dots, a_j]$ runs through all 2^{m-1} ordered partitions of m, and $k_i \in \mathbb{Z}$.

2. Our choice of α_n for negative n shows that the condition "$-1 \leq k_j - a_j$" can be added under the second sum in (1.9) (which is actually finite) and (1.10). For instance, if $m = n$, then the sum in (1.9) has a single non-zero term with $j = 1$, $a_1 = n$, $k_1 = n - 1$, and so $\kappa_{n,0} = -\bar{\alpha}_{n-1}$. This can be seen from (1.2) and $\Phi_n^*(0) = 1$ as well.

3. Notice that for each partition $\{a_l\}$ with $\sum_1^j a_l = m$, the second sum in (1.10) converges when $\alpha_k \in \ell^2$. This is because then $\alpha_k \bar{\alpha}_{k-a} \in \ell^1$ for any fixed a, and so

$$|d_m| \leq \sum_{\substack{\sum_1^j a_l = m \\ j, a_l \geq 1}} \prod_{l=1}^{j} \left(\sum_k |\alpha_k \bar{\alpha}_{k-a_l}| \right)$$

Next, we describe an application of Theorem 1.1 that actually motivated our work. It involves the computation of Taylor coefficients of $\log D$. These are interesting primarily because they coincide with Fourier coefficients of $\log w$. Indeed,

$$\frac{1}{2} \frac{e^{i\theta} + z}{e^{i\theta} - z} = \frac{1}{2} + e^{-i\theta} z + e^{-2i\theta} z^2 + \dots \tag{1.11}$$

and the definition of D show that

$$\log D(z) = \frac{1}{2} w_0 + w_1 z + w_2 z^2 + \dots$$

where w_m are defined by

$$w_m \equiv \int e^{-im\theta} \log w(\theta) \frac{d\theta}{2\pi} = \bar{w}_{-m} \tag{1.12}$$

We know from (1.4) that

$$w_0 = \sum_{k \geq 0} \log(1 - |\alpha_k|^2) = 2 \sum_{k \geq 0} \log \rho_k \tag{1.13}$$

and the methods from [20] can be used to compute the first few of the other w_m. However,the corresponding computations become very complicated with increasing m (already at $m = 4$ they are close to intractable; [20] only deals with $m \leq 2$). Our method will provide w_m for all m, although the resulting formulae will obviously not be simple. That is why we postpone the exact expressions to Theorem 2.4 below and state the result here in the following form that is sufficient for our first application, Theorem 1.4 (see also Lemma 3.1 that contains a similar formula for Taylor coefficients of $\log \Phi_n^*$).

Theorem 1.2. *If $\alpha_k \in \ell^2$, then*

$$w_m = \alpha_{m-1} - \sum_{k\geq 0} \alpha_{k+m}\bar{\alpha}_k + R_m(\mu) \tag{1.14}$$

with

$$|R_m(\mu)| \leq C_m\left(\sum_{k=0}^{m-1} |\alpha_k|^2 + \sum_{k=m}^{\infty} |\alpha_k|^4\right) \tag{1.15}$$

We note that (1.14) will be obtained from (1.10) by means of expanding $\log(1 + d_1 z + d_2 z^2 + \dots)$ into its Taylor series. This is a truly remarkable fact since the sum in (1.10) is m-fold infinite and one might expect this method to only add another degree of difficulty. Nevertheless, after appropriate combinatorial manipulations it will turn out that the sum in (1.14) (as well as the one in the exact form (2.12)) has only a single infinite index!

The first application of the knowledge of w_m we present in this paper aims at the following conjecture of Simon [17] that is a higher order generalization of (1.4). Here S is the left-shift operator on sequences

$$S(x_0, x_1, ...) = (x_1, x_2, ...) \tag{1.16}$$

Conjecture 1.3. For distinct $\{\theta_m\}_{m=1}^l$ in $[0, 2\pi)$ and n_m positive integers, define $N \equiv \sum_{m=1}^l n_m$, $n \equiv 1 + \max_m n_m$, and

$$Q(z) \equiv \prod_{m=1}^{l} (z - e^{i\theta_m})^{n_m} = \sum_{m=0}^{N} q_m z^m \qquad \text{and} \qquad \bar{Q}(z) \equiv \prod_{m=1}^{l} (z - e^{-i\theta_m})^{n_m} = \sum_{m=0}^{N} \bar{q}_m z^m$$

so that

$$\bar{Q}(S) = \sum_{m=0}^{N} \bar{q}_m S^m$$

Then

$$\left|Q(e^{i\theta})\right|^2 \log w(\theta) \in L^1(d\theta) \;\Leftrightarrow\; \{\bar{Q}(S)\alpha\}_k \in \ell^2 \text{ and } \alpha_k \in \ell^{2n} \tag{1.17}$$

For $N = 0$ this is just (1.4). For $N = 1$ the conjecture was proved by Simon (Theorem 2.8.1 in [17]) and for $N = 2$ by Simon and Zlatoš [20]. It remains open for $N \geq 3$ although Denisov and Kupin [4], mimicking the work of Nazarov, Peherstorfer, Volberg, and Yuditskii [14] on Jacobi matrices, showed that for each Q there indeed is a condition in terms of finiteness of a sum involving the α_k that is equivalent to the LHS of (1.17). Unfortunately, this sum is far from transparent and its relation to the RHS of (1.17) is unclear.

Our contribution in this direction is the following *higher order Szegő theorem in* ℓ^4 which shows that Conjecture 1.3 holds if we a priori assume $\alpha_k \in \ell^4$.

Theorem 1.4. *Assume that* $\alpha_k \in \ell^4$*, and for* $q_0, q_1, \dots, q_N \in \mathbb{C}$ *define*

$$Q(z) \equiv \sum_{m=0}^{N} q_m z^m \qquad \textit{and} \qquad \bar{Q}(S) = \sum_{m=0}^{N} \bar{q}_m S^m$$

Then

$$\left|Q(e^{i\theta})\right|^2 \log w(\theta) \in L^1(d\theta) \iff \{\bar{Q}(S)\alpha\}_k \in \ell^2 \tag{1.18}$$

Remark. Of course, the most interesting is the case from Conjecture 1.3 when all zeros of Q are on the unit circle, because the validity of the LHS of (1.18) only depends on them.

Moreover, we provide in Theorem 3.3 an exact formula for the value of

$$Z_Q(\mu) \equiv \int |Q(e^{i\theta})|^2 \log w(\theta) \frac{d\theta}{4\pi}$$

in terms of the α_n. Since Z_Q is an entropy [9, 17], it is upper semi-continuous with respect to weak convergence of measures, and so $Z_Q(\mu) \geq \limsup_n Z_Q(\mu_n)$ with μ_n the *Bernstein-Szegő approximations* of μ having Verblunsky coefficients $\{\alpha_0, \dots, \alpha_n, 0, 0, \dots\}$. We show in Proposition 3.4 that, in fact, we always have $Z_Q(\mu) = \lim_n Z_Q(\mu_n)$, including the case when both sides are $-\infty$ (they cannot be $+\infty$ as each Z_Q is bounded above; see Section 3).

Finally, we apply our method to the computation of the *relative ratio asymptotics* $\Phi^*_{n+1}(\mu)/\Phi^*_n(\mu) - \Phi^*_{n+1}(\nu)/\Phi^*_n(\nu)$ where $\Phi^*_n(\mu)$ and $\Phi^*_n(\nu)$ are the reversed polynomials of measures μ and ν, respectively.

Theorem 1.5. *Let* μ *and* ν *be two non-trivial probability measures on* $\partial\mathbb{D}$*. Let* $\{\alpha_n(\mu)\}$ *and* $\{\alpha_n(\nu)\}$*, respectively, be their Verblunsky coefficients and let* $\Phi^*_n(\mu)$ *and* $\Phi^*_n(\nu)$*, respectively, be their reversed monic orthogonal polynomials. Then*

$$\frac{\Phi^*_{n+1}(\mu)}{\Phi^*_n(\mu)} - \frac{\Phi^*_{n+1}(\nu)}{\Phi^*_n(\nu)} \to 0 \tag{1.19}$$

uniformly on compact subsets of $\mathbb{D}$ *as* $n \to \infty$ *if and only if for any* $\ell \geq 1$

$$\lim_{n\to\infty} \left[\alpha_n(\mu)\bar{\alpha}_{n-\ell}(\mu) - \alpha_n(\nu)\bar{\alpha}_{n-\ell}(\nu)\right] = 0 \tag{1.20}$$

As a corollary of Theorem 1.5, we provide a simple new proof of the results of Khrushchev [8] and Barrios and López [1] on ratio asymptotics Φ^*_{n+1}/Φ^*_n as $n \to \infty$ of the reversed polynomials (Theorem 4.1), as well as their generalization (Theorem 4.2).

The paper is organized as follows. Section 2 computes the Taylor coefficients of Φ^*_n and $\log D$ in terms of the Verblunsky coefficients and proves Theorems 1.1 and 1.2. Section 3 introduces the *step-by-step sum rules* (see [9, 19, 20]) and proves Theorem 1.4. Section 4 proves Theorem 1.5.

2. Coefficients of $\Phi_n^*(z)$ and $\log D(z)$ in Terms of Verblunsky Coefficients

We start with the proof of our first result, Theorem 1.1.

Proof of Theorem 1.1. From (1.2) we have for $n \in \mathbb{Z}$ and $m \geq 0$,

$$\kappa_{n+1,m} = \kappa_{n,m-1} - \bar{\alpha}_n \lambda_{n,m}$$

with the convention $\kappa_{n,-1} \equiv 0$. Substituting this repeatedly into a similar equality obtained from (1.3), we get for $m \geq 1$,

$$\begin{aligned}
\lambda_{n+1,m} &= \lambda_{n,m} - \alpha_n \kappa_{n,m-1} \\
&= \lambda_{n,m} + \alpha_n \bar{\alpha}_{n-1} \lambda_{n-1,m-1} - \alpha_n \kappa_{n-1,m-2} \\
&= \dots \\
&= \lambda_{n,m} + \sum_{a=1}^{m-1} \alpha_n \bar{\alpha}_{n-a} \lambda_{n-a,m-a} + \alpha_n \bar{\alpha}_{n-m}
\end{aligned}$$

where in the last equality we have used $\kappa_{n-m,-1} = 0$ and $\lambda_{n-m,0} = 1$. If we now iterate this and note that $\lambda_{m-l,m} = 0$ for $m \geq 1$ and $l < 0$, we have

$$\begin{aligned}
\lambda_{n+1,m} &= \sum_{k \leq n} \sum_{a=1}^{m-1} \alpha_k \bar{\alpha}_{k-a} \lambda_{k-a,m-a} + \sum_{k \leq n} \alpha_k \bar{\alpha}_{k-m} \\
&= \sum_{k \leq n} \sum_{a=1}^{m-1} \beta_{k,a} \lambda_{k-a,m-a} + \sum_{k \leq n} \beta_{k,m} \qquad (2.1)
\end{aligned}$$

with $\beta_{k,a} \equiv \alpha_k \bar{\alpha}_{k-a}$. Of course, terms with $k < 0$ are zero.

We will prove (1.9) by induction on n. If $n \leq 0$ and $m \geq 1$, then it obviously holds as in that case both sides are zero. Assume therefore that (1.9) holds up to some n and all $m \geq 1$. Then (2.1) gives

$$\begin{aligned}
\lambda_{n+1,m} &= \sum_{k \leq n} \sum_{a=1}^{m-1} \beta_{k,a} \sum_{\substack{\sum_1^j a_l = m-a \\ j, a_l \geq 1}} \sum_{\substack{k_1 < k-a \\ k_2 < k_1 - a_1 \\ \dots \\ k_j < k_{j-1} - a_{j-1}}} \beta_{k_1,a_1} \dots \beta_{k_j,a_j} + \sum_{k \leq n} \beta_{k,m} \\
&= \sum_{\substack{\sum_0^j a_l = m \\ j, a_l \geq 1}} \sum_{\substack{k_0 < n+1 \\ k_1 < k_0 - a_0 \\ \dots \\ k_j < k_{j-1} - a_{j-1}}} \beta_{k_0,a_0} \dots \beta_{k_j,a_j} + \sum_{k_0 < n+1} \beta_{k_0,m} \\
&= \sum_{\substack{\sum_0^j a_l = m \\ j \geq 0 \\ a_l \geq 1}} \sum_{\substack{k_0 < n+1 \\ k_1 < k_0 - a_0 \\ \dots \\ k_j < k_{j-1} - a_{j-1}}} \beta_{k_0,a_0} \dots \beta_{k_j,a_j}
\end{aligned}$$

with $k_0 \equiv k$ and $a_0 \equiv a$. But this is (1.9) for $n+1$ in place of n. Thus (1.9) is proved, and (1.10) follows from (1.8). $\square$

Our next aim is to compute the Taylor coefficients of $\log D$. We will again assume $\alpha_k \in \ell^2$ so that D is well defined. By (1.7) we have for z close to 0,

$$\log\Big(\prod_{k\geq 0}\rho_k\Big) - \log D(z) = \sum_{j\geq 1}\frac{(-1)^{j-1}}{j}(d_1 z + d_2 z^2 + \dots)^j$$

and so w_m is the negative of the m^{th} Taylor coefficient of the RHS when $m \geq 1$. That is,

$$\begin{aligned}
w_m &= \sum_{\substack{\sum_1^j b_\ell = m\\ j,b_\ell\geq 1}} \frac{(-1)^j}{j}\prod_{\ell=1}^{j} d_{b_\ell}\\
&= \sum_{\substack{\sum_1^j b_\ell = m\\ j,b_\ell\geq 1}} \frac{(-1)^j}{j}\prod_{\ell=1}^{j}\Bigg(\sum_{\substack{\sum_1^p a_l = b_\ell\\ p,a_l\geq 1}}\ \sum_{\substack{k_2<k_1-a_1\\ \dots\\ k_p<k_{p-1}-a_{p-1}}} \beta_{k_1,a_1}\dots\beta_{k_p,a_p}\Bigg) &\quad (2.2)\\
&= \sum_{\{(k_1,a_1),\dots,(k_i,a_i)\}\in M_m} \beta_{k_1,a_1}\dots\beta_{k_i,a_i}\sum_{j=1}^{i}\frac{(-1)^j}{j}N_j\big(\{(k_1,a_1),\dots,(k_i,a_i)\}\big) &\quad (2.3)
\end{aligned}$$

with M_m and N_j defined below.

Before stating the definitions, let us first describe how (2.3) was obtained from (2.2). We multiply out the brackets in (2.2) to get a sum of products $\beta_{k_1,a_1}\dots\beta_{k_i,a_i}$ (with coefficients), and then collect terms with identical products (only differing by a permutation). The coefficient at each product $\beta_{k_1,a_1}\dots\beta_{k_i,a_i}$ obtained in this way will then equal the last sum in (2.3). For example, the product $\beta_{3,1}\beta_{1,1}^2$ appears in (2.2) for $m=3$ as $\frac{(-1)^3}{3}(\beta_{3,1})(\beta_{1,1})(\beta_{1,1})$, $\frac{(-1)^3}{3}(\beta_{1,1})(\beta_{3,1})(\beta_{1,1})$, $\frac{(-1)^3}{3}(\beta_{1,1})(\beta_{1,1})(\beta_{3,1})$, $\frac{(-1)^2}{2}(\beta_{3,1}\beta_{1,1})(\beta_{1,1})$, and $\frac{(-1)^2}{2}(\beta_{1,1})(\beta_{3,1}\beta_{1,1})$. The first three come from $j=3$ and $b_1=b_2=b_3=1$ in (2.2), the fourth from $j=2$, $b_1=2$, $b_2=1$, and the fifth from $j=2$, $b_1=1$, $b_2=2$. Therefore the coefficient at $\beta_{3,1}\beta_{1,1}^2$ in (2.3) has to be $\frac{(-1)^3}{3}3+\frac{(-1)^2}{2}2=0$.

It is obvious that the products that appear in (2.3) must satisfy $i, a_l \geq 1$ and $\sum_1^i a_l = m$, because the sum of the a_l's in any term of the ℓ^{th} bracket of (2.2) equals b_ℓ. The set M_m will therefore reflect this condition. The question now is, given any collection (i.e., set with repetitions; see below) of couples $P=\{(k_1,a_1),\dots,(k_i,a_i)\}\in M_m$, in how many ways can the corresponding product $\beta_{k_1,a_1}\dots\beta_{k_i,a_i}$ be obtained by multiplying out $j\geq 1$ brackets in (2.2). If this number is denoted $N_j'(P)$, then the correct coefficient at $\beta_{k_1,a_1}\dots\beta_{k_i,a_i}$ in (2.3) is $\sum_{j=1}^{i}\frac{(-1)^j}{j}N_j'(P)$. Hence to obtain (2.2)=(2.3), we are left with showing that $N_j(P)$, defined below, equals $N_j'(P)$.

We will call a *collection* an unordered list of elements, some of which can be identical (i.e., a collection is a set that can contain multiple identical elements, a hat with multicolored balls). Such identical elements are considered indistinguishable. A *j-tuple* will be an ordered list of j elements. Collections will be denoted by $\{\dots\}$, j-tuples by $[\dots]$. Below we will consider collections and j-tuples whose elements are couples (k,a) with $k\in\mathbb{Z}$, $a\in\mathbb{N}$.

For instance, $\{(3,1),(1,1),(1,1)\}$ is a collection ($\{(1,1),(1,1),(3,1)\}$ is the same one) and $[(3,1),(1,1),(1,1)]$, $[(1,1),(3,1),(1,1)]$, $[(1,1),(1,1),(3,1)]$ are three distinct triples. Finally, the union of collections is the collection obtained by joining their lists of elements, for instance, $\{(3,1),(1,1)\} \cup \{(1,1)\} = \{(3,1),(1,1),(1,1)\}$.

Definition 2.1. Let M_m be the set of all distinct collections $P = \{(k_1,a_1),\dots,(k_i,a_i)\}$ with $i \ge 1$, $k_l \in \mathbb{Z}$, and $a_l \ge 1$ such that $\sum_1^i a_l = m$. We let

$$\beta(P) \equiv \beta_{k_1,a_1} \dots \beta_{k_i,a_i} = \alpha_{k_1}\bar{\alpha}_{k_1-a_1} \dots \alpha_{k_i}\bar{\alpha}_{k_i-a_i} \tag{2.4}$$

We say that a collection $P = \{(k_1,a_1),\dots,(k_i,a_i)\}$ is *linear* if $k_u < k_v - a_v$ or $k_v < k_u - a_u$ whenever $u \ne v$. In particular, $k_u \ne k_v$ when $u \ne v$, which means that a linear collection P cannot contain two identical couples, and thus it is just a set.

If $P \in M_m$ and $j \ge 1$, then $N_j\big(P\big)$ is the number of distinct j-tuples $\mathcal{P} = [P_1,\dots,P_j]$ such that each P_ℓ is a non-empty linear collection and $\bigcup_{\ell=1}^j P_\ell = P$. We will call each such $\mathcal{P}$ an *admissible division* of P.

For instance, if $P = \{(3,1),(1,1),(1,1)\}$ (corresponding to $\beta_{3,1}\beta_{1,1}^2$ above), then the admissible divisions are $[\{(3,1)\},\{(1,1)\},\{(1,1)\}]$, $[\{(1,1)\},\{(3,1)\},\{(1,1)\}]$, $[\{(1,1)\},\{(1,1)\},\{(3,1)\}]$ (with $j=3$) and $[\{(3,1),(1,1)\},\{(1,1)\}]$, $[\{(1,1)\},\{(3,1),(1,1)\}]$ (with $j=2$). Hence in this case $N_3(P)=3$, $N_2(P)=2$, $N_1(P)=0$ and the last sum in (2.3) is indeed $\frac{(-1)^3}{3}3 + \frac{(-1)^2}{2}2 = 0$.

To finish the proof of (2.2)=(2.3) we need to show that $N_j(P) = N'_j(P)$ for any $P \in M_m$ (as we did for $P = \{(3,1),(1,1),(1,1)\}$), where $N'_j(P)$ is the number of times the product $\beta(P) = \beta_{k_1,a_1}\dots\beta_{k_i,a_i}$ is obtained by multiplying out j brackets in (2.2). The desired equality follows from realizing that the collection P_ℓ in the definition ($\ell = 1,\dots,j$) corresponds to the "subproduct" of $\beta(P)$ coming from the ℓ^{th} bracket in (2.2) (which is why the $\mathcal{P}$'s must be ordered, as well as why P_ℓ must be linear). With this identification in mind, it is easy to see that each admissible division $[P_1,\dots,P_j]$ of P corresponds to precisely one way of obtaining $\beta(P)$ in (2.3) from (2.2) by multiplying j subproducts (from j brackets) corresponding to $P_1,\dots,P_j$ (with b_ℓ being the sum of the a_l for which $(k_l,a_l) \in P_\ell$), and vice versa.

Hence we have obtained an explicit expression for w_m. We will now simplify it considerably by showing that coefficients at many $\beta(P)$ in (2.3) are actually zero, as was the case for $\beta_{3,1}\beta_{1,1}^2$ (see Lemma 2.3 below).

We say that $K \in \mathbb{Z}$ is a *cut* of $P = \{(k_1,a_1),\dots,(k_i,a_i)\} \in M_m$ if $k_l \ne K$ for all l, if $\min_l\{k_l\} < K < \max_l\{k_l\}$, and if for any u,v with $k_u < K < k_v$ we have $k_u < k_v - a_v$. For instance, $P = \{(3,1),(1,1),(1,1)\}$ has one cut $K=2$. Our interest here will be mainly in "cuttless" collections as is demonstrated by the following two lemmas.

Lemma 2.2. *If $P \in M_m$ has no cut, then*

$$\max_l\{k_l\} - \min_l\{k_l - a_l\} \le m \tag{2.5}$$

Proof. Consider the union of intervals $I \equiv \bigcup_l [k_l - a_l, k_l] \subset \mathbb{R}$, with $|I| \le \sum_l a_l = m$. If P has no cut, then I is an interval (and vice versa) because otherwise the minimum of any component, except for the bottom one, were a cut. But then we obviously have

$$I = [\min_l \{k_l - a_l\}, \max_l \{k_l\}]$$

proving (2.5). $\square$

Let $|P|$ be the number of elements of a collection P, counting identical elements as many times as they are included in P. For instance, $|\{(3,1),(1,1),(1,1)\}| = 3$.

Lemma 2.3. *If $P \in M_m$ has a cut, then*

$$\sum_{j=1}^{|P|} \frac{(-1)^j}{j} N_j\big(P\big) = 0 \tag{2.6}$$

Proof. Fix $P = \{(k_1,a_1),\ldots,(k_i,a_i)\} \in M_m$ that has a cut K. First notice that if Π is the set of all admissible divisions $\mathcal{P} = [P_1,\ldots,P_{j_\mathcal{P}}]$ of P, then (2.6) is equivalent to

$$\sum_{\mathcal{P}\in\Pi} \frac{(-1)^{j_\mathcal{P}}}{j_\mathcal{P}} = 0 \tag{2.7}$$

For each $\mathcal{P}$ let $\mathcal{C}(\mathcal{P})$ be the collection (not a union!) of up to $2j_\mathcal{P}$ non-empty sets that we obtain by splitting each P_ℓ at K. That is, we define

$$P_\ell^\pm \equiv \{(k_l,a_l)\in P_\ell \,\big|\, \pm(k_l - K) > 0\}$$

so that $P_\ell = P_\ell^+ \cup P_\ell^-$, and then let $\mathcal{C}(\mathcal{P})$ be the collection of those $P_\ell^\pm$ that are not empty. Notice that the $P_\ell^\pm$ are indeed sets and they are linear — both because the same is true for P_ℓ.

Hence $\mathcal{C}$ defines an equivalence relation on Π by $\mathcal{P}\sim\mathcal{P}'$ iff $\mathcal{C}(\mathcal{P}) = \mathcal{C}(\mathcal{P}')$. We will show that the part of the sum in (2.7) corresponding to any equivalence class is zero. That is, we will prove

$$\sum_{\mathcal{C}(\mathcal{P})=\mathcal{C}_0} \frac{(-1)^{j_\mathcal{P}}}{j_\mathcal{P}} = 0 \tag{2.8}$$

for any $\mathcal{C}_0$ such that $\mathcal{C}(\mathcal{P}) = \mathcal{C}_0$ for some $\mathcal{P}\in\Pi$.

Let us fix any such $\mathcal{C}_0$. Then $\mathcal{C}_0$ is a collection of non-empty linear sets $Q_1,\ldots,Q_q$ and $R_1,\ldots,R_r$ whose union (as a union of collections) is P, such that if $(k_l,a_l)\in Q_u$, then $k_l < K$, and if $(k_l,a_l)\in R_u$, then $k_l > K$. That is, the Q_u are the non-empty P_ℓ^- and the R_v are the non-empty P_ℓ^+. Let $q\le r$, since the case $q\ge r$ is identical.

Assume first that these sets are all distinct. Then for every $0\le s\le q$ there are $\binom{q}{s}\binom{r}{s}s!(q+r-s)!$ admissible divisions $\mathcal{P}$ of P with $\mathcal{C}(\mathcal{P}) = \mathcal{C}_0$ and $j_\mathcal{P} = q+r-s$. These are created by choosing s sets from $Q_1,\ldots,Q_q$ and s from $R_1,\ldots,R_r$, taking all $s!$ pairings of the selected Q's with the selected R's, and then all $(q+r-s)!$ orderings of thus created $q+r-s$ sets (unions of the paired couples $Q_u\cup R_v$ together with the unpaired Q's and R's) — the P_ℓ's. Since all the original sets were distinct, this construction gives no repetitions. Notice also that any $P_\ell = Q_u \cup R_v$ is linear because so are Q_u and R_v and K is a cut for P. This shows that the LHS of (2.8) equals

$$\sum_{s=0}^{q} \frac{(-1)^{q+r-s}}{q+r-s}\binom{q}{s}\binom{r}{s} s!(q+r-s)! = (-1)^{q+r}(q+r-1)!\sum_{s=0}^{q}(-1)^s \frac{\binom{q}{s}\binom{r}{s}}{\binom{q+r-1}{s}} = 0$$

The last equality follows from Lerch's identity [13] (also in [7, p. 61])

$$\sum_{s=0}^{q}(-1)^s \frac{\binom{q}{s}\binom{r}{s}}{\binom{p}{s}} = \frac{\binom{p-r}{q}}{\binom{p}{q}}$$

which holds whenever $p \ge q$.

If now some Q's and/or some R's are identical, then in the above sum every $\mathcal{P}$ with $\mathcal{C}(\mathcal{P}) = \mathcal{C}_0$ is counted the same number of times T, which equals the product of the factorials of the numbers of identical sets. This is because there are T permutations of the Q's and R's that fix the classes of identical sets, and hence when we perform the above algorithm to obtain all admissible $\mathcal{P}$'s with $\mathcal{C}(\mathcal{P}) = \mathcal{C}_0$, each such $\mathcal{P}$ will be obtained T times. Therefore the LHS of (2.8) equals

$$\frac{1}{T}\sum_{s=0}^{q} \frac{(-1)^{q+r-s}}{q+r-s}\binom{q}{s}\binom{r}{s} s!(q+r-s)! = 0$$

This proves (2.8), and (2.7) follows by summing over all $\mathcal{C}_0$. $\square$

Hence the only terms that matter in (2.3) are those with no cuts (which is the main point of this section). Moreover, it is obvious that $\beta(P) = 0$ when some $k_l - a_l \le -2$. Therefore we define $\omega(P) \equiv \max_l\{k_l \,|\, (k_l, a_l) \in P\}$, $\delta(P) \equiv \min_l\{k_l - a_l \,|\, (k_l, a_l) \in P\}$,

$$N(P) \equiv \sum_{j=1}^{|P|} \frac{(-1)^j}{j} N_j(P) \tag{2.9}$$

and for $0 \le n \le \infty$

$$M_m^n \equiv \{P \in M_m \,\big|\, P \text{ has no cuts and } 0 \le \omega(P) \le n\} \tag{2.10}$$

If now $P \in M_m \setminus M_m^\infty$, then either P has a cut and so $N(P) = 0$, or $\omega(P) \le -1$ and then $\beta(P) = 0$ because $\delta(P) \le -2$. This means that the sum in (2.3) only needs to be taken over M_m^∞. Before formally stating this fact, we remark that

$$M_m^n = \{P + k \,\big|\, P \in M_m^0 \text{ and } 0 \le k \le n\} \tag{2.11}$$

where $P + k \equiv \{(k_l + k, a_l) \,|\, (k_l, a_l) \in P\}$. Also notice that $N(P+k) = N(P)$ by definition, $P - \omega(P) \in M_m^0$ for any $P \in M_m^\infty$, and M_m^0 is a finite set by Lemma 2.2. In this light the following result is an immediate consequence of (2.3) and Lemma 2.3.

Theorem 2.4. *If $\alpha_k \in \ell^2$, then for $m \ge 1$*

$$w_m = \sum_{P \in M_m^\infty} N(P)\beta(P) = \sum_{P \in M_m^0} N(P) \sum_{k=0}^{\infty} \beta(P+k) \tag{2.12}$$

Remark. The second form of w_m in (2.12) shows that for $m \neq 0$, the m^{th} Fourier coefficient of $\log w(\theta)$ (and so the m^{th} Taylor coefficient of $\log D(z)$) can be expressed as a sum over a single infinite index of products involving only "nearby" α_k's.

Now we are ready to prove Theorem 1.2.

Proof of Theorem 1.2. If

$$M'_m \equiv \{P \in M_m^\infty \,\big|\, |P| = 1\}$$

then (2.12) can be written as

$$w_m = \biggl(\sum_{P \in M'_m} + \sum_{P \in M_m^\infty \smallsetminus M'_m}\biggr) N(P)\beta(P) \tag{2.13}$$

Note that the sum in (1.14) together with $\alpha_{m-1} = -\alpha_{m-1}\alpha_{-1}$ is just the first sum in (2.13), and so $R_m(\mu)$ is the second sum in (2.13). It remains to prove (1.15).

Since each $N_j(P)$ is bounded by a constant only depending on m,

$$|R_m(\mu)| \leq C_m \sum_{P \in M_m^\infty \smallsetminus M'_m} |\beta(P)|$$

For any $P \in M_m^\infty \smallsetminus M'_m$ let $i \equiv |P| \geq 2$. If $\delta(P) \leq -2$, then $\beta(P) = 0$. If $\delta(P) = -1$, then by Lemma 2.2

$$|\beta(P)| \leq \prod_{l=1}^{i} |\alpha_{k_l}| \leq \sum_{j=0}^{m-1} |\alpha_j|^i \leq \sum_{j=0}^{m-1} |\alpha_j|^2$$

And if $\delta(P) \geq 0$, then

$$|\beta(P)| \leq \sum_{l=1}^{i} \bigl(|\alpha_{k_l}|^{2i} + |\bar\alpha_{k_l - a_l}|^{2i}\bigr) \leq m \sum_{j=\delta(P)}^{\delta(P)+m} |\alpha_j|^{2i} \leq m \sum_{j=\delta(P)}^{\delta(P)+m} |\alpha_j|^4$$

Since each $P \in M_m^\infty$ has no cuts, the number of $P \in M_m^\infty$ with any given $\delta(P)$ is a finite constant only depending on m. Hence (1.15) follows and the proof is complete. $\square$

We write here explicitly the first three w's from (2.12). Recall that $\alpha_{-1} = -1$, $\alpha_{-2} = \alpha_{-3} = \cdots = 0$, and $\rho_k = \sqrt{1 - |\alpha_k|^2}$.

$$\begin{aligned}
w_1 &= -\sum_k \alpha_k \bar\alpha_{k-1} \\
w_2 &= -\sum_k \alpha_k \bar\alpha_{k-2} \rho_{k-1}^2 + \tfrac{1}{2} \sum_k \alpha_k^2 \bar\alpha_{k-1}^2 \\
w_3 &= -\sum_k \alpha_k \bar\alpha_{k-3} \rho_{k-1}^2 \rho_{k-2}^2 + \sum_k \alpha_k^2 \bar\alpha_{k-1} \bar\alpha_{k-2} \rho_{k-1}^2 + \sum_k \alpha_k \alpha_{k-1} \bar\alpha_{k-2}^2 \rho_{k-1}^2 - \tfrac{1}{3} \sum_k \alpha_k^3 \bar\alpha_{k-1}^3
\end{aligned}$$

Finally, we note that all Taylor coefficients of $\log(D(z)/D(0))$ verify the claim of the remark after Theorem 2.4. It turns out that this is essentially the only such function of the form $F(D(z)/D(0))$.

Proposition 2.5. *Assume that F is analytic on a neighborhood of 1 and each Taylor coefficient h_m of $H(z) = F(D(z)/D(0))$ is, as a function of $\{\alpha_k\} \in \ell^2$, a sum of products of the α_k's such that if α_k and α_l both appear in the same product, then $|k-l| \le c_{F,m}$ for some $c_{F,m} < \infty$. It follows that $H(z) = a + b\log(D(z)/D(0))$ for some $a, b \in \mathbb{C}$.*

Proof. Define $G(z) = F(e^z)$ so that G is analytic on a neighborhood of 0 (with Taylor coefficients g_m) and $H(z) = G(\log(D(z)/D(0)))$. The fact that $\log(D(z)/D(0)) = \sum_{m\ge1} w_m z^m$ satisfies the proposition shows that when g_m is the first non-zero coefficient with $m \ge 2$, then h_m does not satisfy the required condition because $h_m = g_1 w_m + g_m w_1^m$. Therefore $G(z) = a + bz$ for some a, b. □

3. A Higher Order Szegő Theorem

In this section we will prove Theorem 1.4. We will do this by first deriving sum rules à la Denisov-Kupin [4] that provide us a necessary and sufficient condition for the left hand side (1.18) to hold. The difference between our Theorem 3.3 below and [4] is that in [4] this condition is expressed in terms of traces of powers of the CMV matrix (see, e.g., [17]), which is less explicit than the form we obtain here (although, obviously, the two conditions have to be equivalent). This, together with Theorem 2.4, will suffice to yield Theorem 1.4.

We start by introducing some notation. The Carathéodory and Schur functions, $F : \mathbb{D} \to i\mathbb{C}^-$ and $f : \mathbb{D} \to \mathbb{D}$, for $d\mu$ are defined by

$$F(z) \equiv \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta) \equiv \frac{1+zf(z)}{1-zf(z)}$$

It is a result of Geronimus [5] that the Verblunsky coefficients of μ coincide with the Schur parameters of f defined inductively by the Schur algorithm

$$f(z) = \frac{\alpha_0 + zf_1(z)}{1+z\bar\alpha_0 f_1(z)} \tag{3.1}$$

Here (3.1) defines $\alpha_0 \in \mathbb{D}$ and $f_1 : \mathbb{D} \to \mathbb{D}$, and iteration then yields $\alpha_1, \alpha_2, \dots$ and $f_2, f_3, \dots$. Note that $f(0) = \alpha_0$ and, by induction, m^{th} Taylor coefficient of f only depends on $\alpha_0, \dots, \alpha_m$.

In the following we will write $\Phi_n^*(\mu, z)$ and $D(\mu, z)$ for the reversed polynomials and the Szegő function. Accordingly, we will write $w_m(\mu)$ for the Taylor coefficients of $\log D(\mu, z)$, and we will also let

$$-\log \Phi_n^*(\mu, z) \equiv \sum_{m\ge1} w_{n,m}(\mu) z^m$$

We will now fix a measure μ and denote its Verblunsky coefficients α_k. For the sake of transparency, we will include $\alpha_{-1} = -1$ at the beginning of the sequence of the coefficients, so that these will be $\{-1, \alpha_0, \alpha_1, \dots\}$. We let μ_n be the n-th Bernstein-Szegő approximation of μ, with Verblunsky coefficients $\{-1, \alpha_0, \alpha_1, \dots, \alpha_n, 0, 0, \dots\}$, and $\mu^{(n)} = w^{(n)}(\theta)\frac{d\theta}{2\pi} + d\mu^{(n)}_{\text{sing}}$ the measure with Verblunsky coefficients $\{-1, \alpha_n, \alpha_{n+1}, \dots\}$.

In Section 2.9 of [17], Simon defines the *relative Szegő function*

$$(\delta D)(\mu, z) \equiv \frac{1-\bar\alpha_0 f(z)}{\rho_0}\,\frac{1-zf_1(z)}{1-zf(z)} \tag{3.2}$$

with f, f_1 from (3.1). Its advantage is that, unlike D, it is defined for any μ. If $w(\theta)$ is positive almost everywhere, then so is $w^{(1)}(\theta)$, and

$$\log \frac{w(\theta)}{w^{(1)}(\theta)} \in L^p[0, 2\pi), \quad p < \infty \tag{3.3}$$

with

$$(\delta D)(\mu, z) = \exp\left(\int \frac{e^{i\theta} + z}{e^{i\theta} - z} \log\left(\frac{w(\theta)}{w^{(1)}(\theta)}\right) \frac{d\theta}{4\pi}\right) \tag{3.4}$$

(see Theorem 2.9.3 in [17]). Obviously, in the case $\alpha_k \in \ell^2$ we have

$$(\delta D)(\mu, z) = \frac{D(\mu, z)}{D(\mu^{(1)}, z)} \tag{3.5}$$

which explains the name. We define the Fourier coefficients of $\log(w(\theta)/w^{(1)}(\theta))$ to be $\delta w_m(\mu)$ so that from (3.4) and (1.11) we obtain

$$\log(\delta D)(\mu, z) = \frac{1}{2}\delta w_0(\mu) + \delta w_1(\mu) z + \delta w_2(\mu) z^2 + \dots \tag{3.6}$$

In particular,

$$\delta w_0(\mu) = 2 \log \rho_0 \tag{3.7}$$

by (3.2) and $f(0) = \alpha_0$, and $\delta w_{-m}(\mu) = \overline{\delta w_m(\mu)}$. In the case $\alpha_k \in \ell^2$ we have by (3.5)

$$\delta w_m(\mu) = w_m(\mu) - w_m(\mu^{(1)}) \tag{3.8}$$

Finally, we define $N(P)$ by (2.9), $\beta(P)$ by (2.4), and let $\beta^{(n)}(P)$ be defined as $\beta(P)$, but with $\alpha_{-1}, \alpha_0, \dots, \alpha_{n-2}$ replaced by zeros and α_{n-1} replaced by -1. That is, $\beta^{(n)}(P)$ equals $\beta(P - n)$ for the measure $\mu^{(n)}$. For instance, $\beta^{(2)}(\{(3, 1), (1, 1)\}) = \alpha_3 \bar{\alpha}_2 (-1)\bar{0} = 0$, which is $\beta(\{(1, 1), (-1, 1)\})$ for the measure $\mu^{(2)}$ with Verblunsky coefficients $\{-1, \alpha_2, \alpha_3, \dots\}$. In particular, (2.12) for the measure $\mu^{(n)}$ and $N(P - n) = N(P)$ imply

$$w_m(\mu^{(n)}) = \sum_{P \in M_m^\infty} N(P) \beta^{(n)}(P) \tag{3.9}$$

whenever $\alpha_k \in \ell^2$. Notice also that by Lemma 2.2,

$$\beta^{(n)}(P) = \beta(P) \tag{3.10}$$

when $P \in M_m^\infty$ and $\omega(P) \geq m + n$. Since we have fixed the α_k's, it will be more transparent to use the notation $\beta(P)$, $\beta^{(n)}(P)$ rather than $\beta(\mu, P)$, $\beta(\mu^{(n)}, P - n)$.

Next we show that Theorem 2.4 easily extends to $D(\mu_n)$, $\Phi_n^*(\mu)$, and $\delta D(\mu)$.

Lemma 3.1. *For $m \geq 1$ and any μ we have*

$$w_m(\mu_n) = \sum_{P \in M_m^n} N(P) \beta(P) = w_{n+1,m}(\mu) \tag{3.11}$$

$$\delta w_m(\mu) = \sum_{P \in M_m^\infty} N(P) \big[\beta(P) - \beta^{(1)}(P)\big] \tag{3.12}$$

Proof. The first equality in (3.11) is nothing but (2.12) for the measure μ_n instead of μ. Then (1.3), (1.5), and (1.6) show that

$$\log \Phi^*_{n+1}(\mu, z) = \log \Phi^*_{n+1}(\mu_n, z) = \sum_{k=0}^{n} \log \rho_k - \log D(\mu_n, z)$$

since $\Phi^*_{n+1}(\mu_n, z) = \|\Phi^*_{n+1}(\mu_n, z)\|_{L^2(d\mu)} D^{-1}(\mu_n, z)$, and so $w_{n+1,m}(\mu) = w_m(\mu_n)$ for $m \geq 1$.

By (3.2), the m-th Taylor coefficient of $\delta D(\mu, z)$ (and so of $\log \delta D(\mu, z)$, too) only depends on α_0, the first m Taylor coefficients of f and first $m-1$ of f_1. That is, $\delta w_m(\mu)$ is a function of $\alpha_0, \dots, \alpha_m$ only (see (1.3.48) in [17]). This means that for any $n \geq m$ we have

$$\delta w_m(\mu) = \delta w_m(\mu_n) = w_m(\mu_n) - w_m((\mu_n)^{(1)}) = \sum_{P \in M_m^n} N(P)\big[\beta(P) - \beta^{(1)}(P)\big]$$

where the second equality is (3.8) for μ_n and the third follows from (3.11) and (3.9). But the last sum equals the right hand side of (3.12) because (3.10) shows that $\beta^{(1)}(P) = \beta(P)$ when $P \in M_m^\infty \smallsetminus M_m^n$. $\square$

After this preparation we are ready to provide a characterization of sequences of Verblunsky coefficients corresponding to measures μ for which $\log w(\theta)$ is integrable with respect to some polynomial weight $|Q(e^{i\theta})|^2$. We let $Q(z) \equiv \prod_{m=0}^{N} q_m z^m$ and define p_m by

$$|Q(z)|^2 = \sum_{m=-N}^{N} p_m z^m = p_0 + \sum_{m=1}^{N} 2\,\mathrm{Re}(p_m z^m) \qquad \text{for } |z| = 1$$

(note that $p_m = q_N \bar{q}_{N-m} + \cdots + q_m \bar{q}_0 = \bar{p}_{-m}$). With the convention $\log 0 = -\infty$ we set

$$Z_Q(\mu) \equiv \int |Q(e^{i\theta})|^2 \log w(\theta) \frac{d\theta}{4\pi} \tag{3.13}$$

which is defined for any μ but can be $-\infty$. This is because with $\log_\pm x \equiv \max\{\pm \log x, 0\}$,

$$\int |Q(e^{i\theta})|^2 \log_+ w(\theta) \frac{d\theta}{4\pi} \leq \|Q\|_\infty^2 \int w(\theta) \frac{d\theta}{4\pi} \leq \frac{\|Q\|_\infty^2}{4\pi} \tag{3.14}$$

but the integral of $\log_- w(\theta)$ can be infinite, for instance, when $w(\theta) = 0$ on a set of positive measure. It is more common to let Z_Q be the negative of (3.13), so that it is bounded from below rather than above, but our definition will be more convenient here. Note also that by (3.3) with $p = 1$ subsequently applied to $\mu^{(n)}$, $n \geq 0$, in place of μ, we have either $Z_Q(\mu^{(n)}) = -\infty$ for all $n \geq 0$ or $Z_Q(\mu^{(n)}) > -\infty$ for all $n \geq 0$.

Before determining the condition for $Z_Q(\mu) > -\infty$, we prove the following *step-by-step sum rule.*

Lemma 3.2. *For any μ*

$$Z_Q(\mu) = p_0 \log \rho_0 + \sum_{m=1}^{N} \mathrm{Re}\left(\bar{p}_m \sum_{P \in M_m^\infty} N(P)\big[\beta(P) - \beta^{(1)}(P)\big]\right) + Z_Q(\mu^{(1)}) \tag{3.15}$$

Proof. The sum on the right hand side of (3.15) is always finite since it has only finitely many non-zero elements and so (3.15) holds if both Z_Q terms are $-\infty$. If both are finite, then (3.3) holds and so

$$\int |Q(e^{i\theta})|^2 \log\left(\frac{w(\theta)}{w^{(1)}(\theta)}\right)\frac{d\theta}{4\pi} = \frac{1}{2}p_0\delta w_0(\mu) + \sum_{m=1}^{N} \mathrm{Re}(\bar{p}_m \delta w_m(\mu))$$

together with (3.7) and (3.12) gives (3.15). □

Theorem 3.3. *For any μ and Q,*

$$Z_Q(\mu) = \sum_{k=0}^{\infty} \mathrm{Re}\left(p_0 \log \rho_k + \sum_{m=1}^{N} \bar{p}_m \sum_{P \in M_m^0} N(P)\beta(P+k)\right) \tag{3.16}$$

Remark. This shows that $Z_Q(\mu)$ is finite if and only if the above sum converges.

Proof. Since (1.12) and (3.13) give

$$Z_Q(\mu_n) = \frac{1}{2}p_0 w_0(\mu_n) + \sum_{m=1}^{N} \mathrm{Re}(\bar{p}_m w_m(\mu_n)) \tag{3.17}$$

it follows from (1.13) and (3.11) that

$$Z_Q(\mu_n) = \sum_{k=0}^{n} p_0 \log \rho_k + \sum_{m=1}^{N} \mathrm{Re}\left(\bar{p}_m \sum_{P \in M_m^n} N(P)\beta(P)\right) \tag{3.18}$$

Hence by (2.12) and $N(P) = N(P - \omega(P))$, the claim is equivalent to $Z_Q(\mu) = \lim_{n\to\infty} Z_Q(\mu_n)$.

It is well known that Z_Q is an entropy and therefore upper semi-continuous in μ with respect to weak convergence of measures (see Section 2.3 in [17]). In particular, since $\mu_n \rightharpoonup \mu$, we obtain

$$Z_Q(\mu) \geq \limsup_{n\to\infty} Z_Q(\mu_n)$$

Thus we are left with proving

$$Z_Q(\mu) \leq \liminf_{n\to\infty} Z_Q(\mu_n) \tag{3.19}$$

This is obviously true if $Z_Q(\mu) = -\infty$, so assume that $Z_Q(\mu^{(n)}) > -\infty$ for all $n \geq 0$. Then $w(\theta) > 0$ for a.e. θ and by Rakhmanov's theorem, $\alpha_n \to 0$ as $n \to \infty$.

The step-by-step sum rule (3.15) for $\mu^{(n)}$ in place of μ reads

$$Z_Q(\mu^{(n)}) = p_0 \log \rho_n + \sum_{m=1}^{N} \mathrm{Re}\left(\bar{p}_m \sum_{P \in M_m^\infty} N(P)\left[\beta^{(n)}(P) - \beta^{(n+1)}(P)\right]\right) + Z_Q(\mu^{(n+1)})$$

and therefore we can iterate it and cancel the terms in the telescoping sum to obtain

$$Z_Q(\mu) = p_0 \sum_{k=0}^{n} \log \rho_k + \sum_{m=1}^{N} \mathrm{Re}\left(\bar{p}_m \sum_{P \in M_m^\infty} N(P)\left[\beta(P) - \beta^{(n+1)}(P)\right]\right) + Z_Q(\mu^{(n+1)})$$

Using $\mu^{(n)} \rightharpoonup \frac{d\theta}{2\pi}$ (since $\alpha_n \to 0$), $Z_Q(\frac{d\theta}{2\pi}) = 0$, and upper semi-continuity of Z_Q, we obtain

$$Z_Q(\mu) \le \liminf_{n\to\infty} \left[p_0 \sum_{k=0}^{n} \log\rho_k + \sum_{m=1}^{N} \operatorname{Re}\left(\bar{p}_m \sum_{P\in M_m^\infty} N(P)\big[\beta(P) - \beta^{(n+1)}(P)\big] \right) \right]$$

We claim that the quantity inside the lim inf differs by $o(1)$ from (3.18) (in which case (3.19) holds and we are done). Indeed — the difference of these two is at most

$$\sum_{m=1}^{N} |p_m| \Bigg(\sum_{P\in M_m^{n+m}\setminus M_m^n} |N(P)||\beta(P) - \beta^{(n+1)}(P)| \Bigg)$$

by (3.10) and the fact that $\beta^{(n+1)}(P) = 0$ when $P \in M_m^n$. This sum has a uniformly bounded number of terms for all n, both p_m and $N(P)$ are also bounded by a constant not depending on n (only on N and Q), and

$$\lim_{n\to\infty} \sup_{P\in M_m^{n+m}\setminus M_m^n} \{|\beta(P)| + |\beta^{(n+1)}(P)|\} = 0$$

since $\alpha_n \to 0$. □

We have thus expressed $Z_Q(\mu)$ as an infinite sum in terms of the Verblunsky coefficients of μ. We can now apply Theorem 1.2 to prove Theorem 1.4.

Proof of Theorem 1.4. The right hand side of (1.18) is equivalent to $Z_Q(\mu) > -\infty$. By Theorem 3.3, this happens precisely when $\lim_{n\to\infty}$ (3.17) $> -\infty$. But

$$\frac{1}{2} p_0 w_0(\mu_n) = p_0 \sum_{k=0}^{n} \log\rho_k = -\frac{1}{2} p_0 \sum_{k=0}^{n} \big(|\alpha_k|^2 + O(|\alpha_k|^4) \big)$$

and by Theorem 1.2 applied to μ_n, the sum in (3.17) is equal to

$$\sum_{m=1}^{N} \operatorname{Re}\Bigg(\bar{p}_m \Bigg(R_m(\mu_n) - \sum_{k\le n-m} \alpha_{k+m}\bar{\alpha}_k \Bigg) \Bigg)$$

The estimate (1.15) and the hypothesis show that $R_m(\mu_n)$ and $\sum_{k=0}^{n} O(|\alpha_k|^4)$ are uniformly bounded in n, so it only remains to show that $\{\bar{Q}(S)\alpha\}_k \in \ell^2$ is equivalent to

$$-\left[\frac{1}{2} p_0 \sum_{k\le n} |\alpha_k|^2 + \sum_{m=1}^{N} \operatorname{Re}\Bigg(\bar{p}_m \sum_{k\le n-m} \alpha_{k+m}\bar{\alpha}_k \Bigg) \right] \tag{3.20}$$

being uniformly bounded in n. We write

$$\begin{aligned}\sum_{k\le n} |\{\bar{Q}(S)\alpha\}_k|^2 &= \sum_{k\le n} |\bar{q}_N \alpha_{k+N} + \cdots + \bar{q}_0 \alpha_k|^2 \\ &= (|q_N|^2 + \cdots + |q_0|^2) \sum_{k\le n} |\alpha_k|^2 + \sum_{m=1}^{N} 2\operatorname{Re}\Bigg((\bar{q}_N q_{N-m} + \cdots + \bar{q}_m q_0) \sum_{k\le n} \alpha_{k+m}\bar{\alpha}_k \Bigg) + O(1)\end{aligned}$$

$$= p_0 \sum_{k\le n} |\alpha_k|^2 + \sum_{m=1}^{N} 2\,\mathrm{Re}\left(\bar p_m \sum_{k\le n-m} \alpha_{k+m}\bar\alpha_k\right) + O(1)$$

In the second equality the remainder $O(1)$ is bounded by a constant independent of n because it is a sum of a bounded number of terms involving only α_k with $k \le N$ or $|k-n| \le N$. And the last equality holds because $p_m = q_N\bar q_{N-m} + \cdots + q_m\bar q_0$ and $\sum_{k=n-m+1}^{n} \alpha_{k+m}\bar\alpha_k$ is uniformly bounded in n and so $O(1)$. Hence $\{\bar Q(S)\alpha\}_k \in \ell^2$ if and only if (3.20) is uniformly bounded in n. □

Recall that in the proof of Theorem 3.3 we have showed $Z_Q(\mu) = \lim_{n\to\infty} Z_Q(\mu_n)$. Here is a generalization of this fact.

Proposition 3.4. *If $f = gQ$ with $g \in C(\partial\mathbb{D})$ a positive function and Q a polynomial, then for any μ,*

$$Z_f(\mu) = \lim_{n\to\infty} Z_f(\mu_n) \tag{3.21}$$

Proof. We again have $Z_f(\mu) \ge \limsup_n Z_f(\mu_n)$ because Z_f is upper semi-continuous as well [17]. Let g_ε be a polynomial such that $g^2 \le |g_\varepsilon|^2 \le g^2 + \varepsilon$ on $\partial\mathbb{D}$. Such a polynomial exists because the functions $h(z) = \sum_{k=-K}^{K} c_k z^k$ are dense in $C(\partial\mathbb{D})$ (by the complex Stone-Weierstrass theorem) and the polynomial $z^K h(z)$ satisfies $|z^K h(z)| = |h(z)|$ for $|z| = 1$. Then $\lim_{\varepsilon\to 0} Z_{g_\varepsilon Q}(\mu) = Z_f(\mu)$ because g is bounded away from 0, and

$$Z_f(\mu_n) \ge Z_{g_\varepsilon Q}(\mu_n) - \frac{\varepsilon}{4\pi}\|Q\|_\infty^2$$

by (3.14). Since $g_\varepsilon Q$ is a polynomial, the proof of Theorem 3.3 shows $\lim_n Z_{g_\varepsilon Q}(\mu_n) = Z_{g_\varepsilon Q}(\mu)$, and we obtain $\liminf_n Z_f(\mu_n) \ge Z_f(\mu)$ by taking $\varepsilon \to 0$. □

It is an interesting open question whether this result holds for any f, not just such that vanish at only finitely many points of $\partial\mathbb{D}$ and to an even degree.

4. Relative Ratio Asymptotics

In this section we provide another application of our methods. We prove Theorem 1.5 and give a simple proof of a deep result, in part due to Khrushchev [8] and in part to Barrios and López [1], on ratio asymptotics Φ_{n+1}^*/Φ_n^* as $n \to \infty$ of the reversed polynomials (see also [18, Section 9.5]). We also give a generalization of this result.

Proof of Theorem 1.5. Let us define

$$\Omega_n(\mu,\nu) \equiv \frac{\Phi_{n+1}^*(\mu)/\Phi_n^*(\mu)}{\Phi_{n+1}^*(\nu)/\Phi_n^*(\nu)}, \qquad \log\Omega_n(\mu,\nu) \equiv \sum_{m\ge 1} \omega_{n,m}(\mu,\nu) z^m$$

(recall that $\Phi_n^*(0) = 1$). It follows from $|\Phi_n(z)| \le |\Phi_n^*(z)|$ for $z \in \mathbb{D}$ (see (1.7.1) in [17]) and from

$$\frac{\Phi_{n+1}^*}{\Phi_n^*} = 1 - \alpha_n z \frac{\Phi_n}{\Phi_n^*} \tag{4.1}$$

(see (1.3)) that the ratio Φ^*_{n+1}/Φ^*_n is bounded away from 0 and ∞ on any compact $K \subset \mathbb{D}$. Hence (1.19) is equivalent to $\Omega_n(\mu,\nu) \to 1$ as $n \to \infty$ uniformly on compact subsets of $\mathbb{D}$, which in turn is equivalent to $\omega_{n,m} \to 0$ as $n \to \infty$ for each $m \geq 1$.

Now

$$\log \Omega_n(\mu,\nu) = \log \Phi^*_{n+1}(\mu) - \log \Phi^*_n(\mu) - \log \Phi^*_{n+1}(\nu) + \log \Phi^*_n(\nu)$$

and so by (3.11) and (2.11)

$$\omega_{n,m}(\mu,\nu) = \sum_{P \in M^0_m} N(P)\big(\beta(\nu, P+n) - \beta(\mu, P+n)\big) \tag{4.2}$$

for $m \geq 1$. If (1.20) holds, then obviously $\beta(\nu, P+n) - \beta(\mu, P+n) \to 0$ for each $P \in M^0_m$, that is, $\omega_{n,m}(\mu,\nu) \to 0$ as $n \to \infty$.

Assume now that $\omega_{n,m}(\mu,\nu) \to 0$ as $n \to \infty$. When $m = 1$, then $M^0_m = \{\{(0,1)\}\}$ and (4.2) equals just $\alpha_n(\mu)\bar\alpha_{n-1}(\mu) - \alpha_n(\nu)\bar\alpha_{n-1}(\nu)$. Hence (1.20) holds for $\ell = 1$. We proceed by induction, so assume (1.20) holds for $\ell = 1, \dots, m-1$. If $P \in M^0_m$ and $|P| \geq 2$, then by the induction hypothesis $\beta(\nu, P+n) - \beta(\mu, P+n) \to 0$ (because $\max\{a_l \,|\, (k_l, a_l) \in P\} \leq m-1$). Since (4.2) converges to 0 and the only element of M^0_m with $|P| = 1$ is $P = \{(0,m)\}$ (in which case $N(P) = -1$), it follows that $\beta(\nu, \{(n,m)\}) - \beta(\mu, \{(n,m)\}) \to 0$ as well. But this is (1.20) for $\ell = m$. $\square$

For any $a \in [0,1]$ we define

$$G_a(z) \equiv \frac{1}{2}\left(1 + z + \sqrt{(1-z)^2 + 4a^2 z}\,\right)$$

with the usual branch of the square root (in particular, $G_0 \equiv 1$). We then have

Theorem 4.1 ([8] and [1]). *Let μ be a non-trivial probability measure on $\partial\mathbb{D}$. Then*

$$\Omega_n(\mu) \equiv \frac{\Phi^*_{n+1}}{\Phi^*_n} \tag{4.3}$$

converges uniformly on compact subsets of $\mathbb{D}$ as $n \to \infty$ if and only if for each $\ell \geq 1$ there is $c_\ell \in \overline{\mathbb{D}}$ such that

$$\lim_{n\to\infty} \alpha_n \bar\alpha_{n-\ell} = c_\ell \tag{4.4}$$

Moreover, if (4.4) *holds for all $\ell \geq 1$, then $c_\ell = a^2\lambda^\ell$ for some $a \in [0,1]$ and $|\lambda| = 1$ and*

$$\lim_{n\to\infty} \Omega_n(\mu; z) = G_a(\lambda z)$$

Remark. In particular, $\lim_n \Omega_n(\mu) = 1$ precisely when all $c_\ell = 0$. In this case (4.4) is called *Máté-Nevai condition.* Accordingly, one might call (1.20) *relative Máté-Nevai condition.*

Proof. Equivalence of the convergence of (4.3) and (4.4) is proved in the same way as Theorem 1.5. The only difference is that now with

$$\log \Omega_n(\mu) \equiv \sum_{m\geq 1} \omega_{n,m}(\mu) z^m$$

(4.2) reads

$$\omega_{n,m}(\mu) = -\sum_{P\in M_m^0} N(P)\beta(\mu, P+n) \tag{4.5}$$

and "$\beta(\nu, P+n) - \beta(\mu, P+n) \to 0$" and "$\omega_{n,m}(\mu,\nu) \to 0$" are replaced by "$\beta(\mu, P+n)$ converges" and "$\omega_{n,m}(\mu)$ converges", respectively, in the argument (we actually have $\beta(\mu, P+n) \to c_{a_1}\dots c_{a_i}$ when $P = \{(k_1,a_1),\dots,(k_i,a_i)\}$). Note that the proof also shows that $\lim_n \Omega_n(\mu) = 1 \equiv G_0$ precisely when all $c_\ell = 0$ (and so $a = 0$).

Hence assume (4.4) holds with not all $c_\ell = 0$. It is obvious that if $c_1 = 0$, then the existence of the limit c_ℓ implies $c_\ell = 0$ for all ℓ. Thus we must have $c_1 = a^2\lambda$ for some $a \in (0,1]$ and $|\lambda| = 1$. In particular, $\liminf_n |\alpha_n| > 0$. But then $|\alpha_{n+3}||\alpha_{n+2}| - |\alpha_{n+1}||\alpha_n| \to 0$ and $|\alpha_{n+3}||\alpha_{n+1}| - |\alpha_{n+2}||\alpha_n| \to 0$ (both by (4.4)) give $|\alpha_{n+2}| - |\alpha_{n+1}| \to 0$, which together with $|\alpha_{n+2}||\alpha_{n+1}| \to a^2$ gives $|\alpha_n| \to a$. This and (4.4) imply $\alpha_{n+1}\alpha_n^{-1} \to \lambda$, and then $\alpha_n\alpha_{n-\ell}^{-1} \to \lambda^\ell$ so that $c_\ell = a^2\lambda^\ell$ for all ℓ.

It remains to prove that in the case $a \neq 0$ the limit $G(z)$ of (4.3) is $G_a(\lambda z)$. Let ν be the measure with Verblunsky coefficients $\alpha_n(\nu) \equiv a\lambda^n$ if $0 < a < 1$ and $\alpha_n(\nu) \equiv a_n\lambda^n$ with $a_n \uparrow 1$ if $a = 1$. Then Theorem 1.5 applies and so the limit function of $\Phi^*_{n+1}(\nu)/\Phi^*_n(\nu)$ is also $G(z)$. By (4.1) we know that the limit $H(z)$ of $\alpha_n(\nu)\Phi_n(\nu)/\Phi^*_n(\nu)$ also must exist and

$$G = 1 - zH \tag{4.6}$$

From (1.2) we have

$$\lambda z\alpha_n(\nu)\,\frac{\Phi_n(\nu)}{\Phi^*_n(\nu)} = \lambda\alpha_n(\nu)\,\frac{\Phi_{n+1}(\nu)}{\Phi^*_n(\nu)} + a^2\lambda = \alpha_{n+1}(\nu)\,\frac{\Phi_{n+1}(\nu)}{\Phi^*_{n+1}(\nu)}\frac{\Phi^*_{n+1}(\nu)}{\Phi^*_n(\nu)} + a^2\lambda$$

Therefore $\lambda zH = HG + a^2\lambda$. We substitute $H(G-\lambda z) = -a^2\lambda$ into (4.6) multiplied by $G-\lambda z$ to obtain $G^2 - (1+\lambda z)G + (1-a^2)\lambda z = 0$. Using $G(0) = 1$, it follows that $G(z) = G_a(\lambda z)$. □

We conclude with using (4.5) to obtain the following generalization of Theorem 4.1.

Theorem 4.2. *Let μ be a non-trivial probability measure on $\partial\mathbb{D}$, let $\{j_n\}$ be an increasing sequence of integers and $k' \in \mathbb{Z}\cup\{\infty\}$. Then $\Omega_{k+j_n}(\mu)$ from (4.3) converges for any $k < k'$ uniformly on compact subsets of $\mathbb{D}$ as $n \to \infty$ if and only if for each $\ell \geq 1$ and $k < k'$ there is $c_{k,\ell} \in \overline{\mathbb{D}}$ such that*

$$\lim_{n\to\infty} \alpha_{k+j_n}\bar{\alpha}_{k+j_n-\ell} = c_{k,\ell} \tag{4.7}$$

Remark. For the special case $j_n = np$ with $p \geq 1$ see [18, Theorem 9.5.10].

Proof. We again follow the lines of the two previous proofs. In one direction we have that the existence of all the $c_{k,\ell}$ with $k < k'$ gives the convergence of $\beta(\mu, P+k+j_n)$ for any $P \in \bigcup_m M_m^0$ and $k < k'$ as $n \to \infty$ (note that if $(k_l, a_l) \in P \in M_m^0$, then $k_l \leq 0$). This in turn gives the convergence of $\omega_{k+j_n,m}(\mu)$ for any $k < k'$ and m by (4.5), and thus that of $\Omega_{k+j_n}(\mu)$ for $k < k'$.

In the opposite direction, the reverse of this argument and induction on ℓ as in the proof of Theorem 1.5 shows that the convergence of $\Omega_{k+j_n}(\mu)$ for any $k < k'$ implies (4.7). This again uses the fact that if $(k_l, a_l) \in P \in M_m^0$, then $k_l \leq 0$. □

References

[1] D. Barrios Rolania and G. López Lagomasino, *Ratio asymptotics for polynomials orthogonal on arcs of the unit circle*, Constr. Approx. **15** (1999), 1–31.

[2] P.A. Deift and R. Killip, *On the absolutely continuous spectrum of one-dimensional Schrödinger operators with square summable potentials*, Comm. Math. Phys. **203** (1999), 341–347.

[3] S.A. Denisov, *On the coexistence of absolutely continuous and singular continuous components of the spectral measure for some Sturm-Liouville operators with square summable potential*, J. Differential Equations **191** (2003), 90–104.

[4] S.A. Denisov and S. Kupin, *Asymptotics of the orthogonal polynomials for the Szegő class with a polynomial weight*, to appear in J. Approx. Theory.

[5] Ya. L. Geronimus, *On polynomials orthogonal on the circle, on trigonometric moment problem, and on allied Carathéodory and Schur functions*, Mat. Sb. **15** (1944), 99–130. [Russian]

[6] Ya. L. Geronimus, *Orthogonal Polynomials: Estimates, Asymptotic Formulas, and Series of Polynomials Orthogonal on the Unit Circle and on an Interval*, Consultants Bureau, New York, 1961.

[7] J Kaucký, *Kombinatorické Identity*, Veda, Bratislava, 1975. [Czech]

[8] S.V. Khrushchev, *Classification theorems for general orthogonal polynomials on the unit circle*, J. Approx. Theory **116** (2002), 268–342.

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

[10] S. Kupin, *On sum rules of special form for Jacobi matrices*, C. R. Math. Acad. Sci. Paris **336** (2003), 611–614.

[11] S. Kupin , *On a spectral property of Jacobi matrices*, Proc. Amer. Math. Soc. **132** (2004), 1377–1383.

[12] A. Laptev, S. Naboko, and O. Safronov, *On new relations between spectral properties of Jacobi matrices and their coefficients*, Comm. Math. Phys. **241** (2003), 91–110.

[13] M. Lerch, *Z počtu integrálního*, Rozpravy České akad. věd a umění. Řada mat.-fys. **9** (1893). [Czech]

[14] F. Nazarov, F. Peherstorfer, A. Volberg, and P. Yuditskii, *On generalized sum rules for Jacobi matrices*, Int. Math. Res. Not. **2005**, 155–186.

[15] E.A. Rakhmanov, *The asymptotic behavior of the ratio of orthogonal polynomials II*, Mat. Sb. **118** (1982), 104–117.

[16] B. Simon, *A canonical factorization for meromorphic Herglotz functions on the unit disk and sum rules for Jacobi matrices*, J. Funct. Anal. **214** (2004), 396–409.

[17] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part. 1: Classical Theory*, AMS Colloquium Publications, Vol. 54.1, American Mathematical Society, Providence, RI, 2005.

[18] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part. 2: Spectral Theory*, AMS Colloquium Publications, Vol. 54.2, American Mathematical Society, Providence, RI, 2005.

[19] B. Simon and A. Zlatoš, *Sum rules and the Szegő condition for orthogonal polynomials on the real line*, Comm. Math. Phys. **242** (2003), 393–423.

[20] B. Simon and A. Zlatoš, *Higher order Szegő theorems with two singular points*, J. Approx. Theory. **134** (2005), 114–129.

[21] G. Szegő, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ., Vol. 23, American Mathematical Society, Providence, R.I., 1939; 3rd edition, 1967.

[22] S. Verblunsky, *On positive harmonic functions* (*second paper*), Proc. London Math. Soc. (2) **40** (1936), 290–320.

[23] A. Zlatoš, *Sum rules for Jacobi matrices and divergent Lieb-Thirring sums*, J. Funct. Anal. **225** (2005), 371–382.

[24] A. Zlatoš, *The Szegő condition for Coulomb Jacobi matrices*, J. Approx. Theory **121** (2003), 119–142.

Mathematics Division, Institute for Low Temperature Physics and Engineering, 47 Lenin Avenue, Kharkov 61103, Ukraine
E-mail address: golinskii@ilt.kharkov.ua

Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI 53706, USA
E-mail address: zlatos@math.wisc.edu