Journal of Approximation Theory **115**, 187–237 (2002)
doi:10.1006/jath.2001.3655, available online at http://www.idealibrary.com on IDEAL®

Cesàro Asymptotics for Orthogonal Polynomials on the Unit Circle and Classes of Measures

Leonid Golinskii[1]

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

and

Sergei Khrushchev

Department of Mathematics, Atilim University, 06836 Incek, Ankara, Turkey
E-mail: svk_49@yahoo.com

Communicated by Walter Van Assche

Received September 8, 2000; accepted in revised form September 21, 2001

The convergence in $L^2(\mathbb{T})$ of the even approximants of the Wall continued fractions is extended to the Cesàro–Nevai class CN, which is defined as the class of probability measures σ with $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |a_k| = 0$, $\{a_n\}_{n\geqslant 0}$ being the Geronimus parameters of σ. We show that CN contains universal measures, that is, probability measures for which the sequence $\{|\varphi_n|^2\, d\sigma\}_{n\geqslant 0}$ is dense in the set of all probability measures equipped with the weak-* topology. We also consider the "opposite" Szegő class which consists of measures with $\sum_{n=0}^{\infty} (1-|a_n|^2)^{1/2} < \infty$ and describe it in terms of Hessenberg matrices. © 2002 Elsevier Science (USA)

Key Words: unit circle orthogonal polynomials; Schur functions; Schur parameters; strong summability.

1. INTRODUCTION

Denote by $\mathscr{P}$ the set of all probability measures on the unit circle $\mathbb{T}$ equipped with the weak-* topology of the Banach space $M(\mathbb{T})$ of all finite complex Borel measures on $\mathbb{T}$ with variation norm. Given $\sigma \in \mathscr{P}$ the

[1] The work of the first author was partially supported by INTAS Grant 2000-272.

0021-9045/02 $35.00

sequence of orthonormal polynomials $\{\varphi_n\}_{n \geqslant 0}$ in the Hilbert space $L^2(d\sigma)$ is defined by

$$\int_{\mathbb{T}} \varphi_n \overline{\varphi_k}\, d\sigma = \delta_{n,k}, \qquad n, k \in \mathbb{Z}_+ \stackrel{\text{def}}{=} \{0, 1, ...\},$$

$$\varphi_n(\sigma, z) = \kappa_n z^n + \cdots + \varphi_n(0), \qquad \kappa_n > 0.$$

We always assume that the Borel support supp σ, that is, the smallest *closed* set with the complement having σ-measure zero, is an infinite set on $\mathbb{T}$.

A new approach to the study of Szegő, Erdős, Nevai, and Rakhmanov measures is developed in a recent paper [19]. Recall that $\sigma \in \mathscr{P}$ is called a Rakhmanov measure if

$$* - \lim_{n \to \infty} |\varphi_n|^2\, d\sigma = dm$$

in $\mathscr{P}$. Here and in what follows m stands for the normalized Lebesgue (arc) measure on $\mathbb{T}$. The importance of the Rakhmanov class R, introduced in [19], can be easily demonstrated on the example of the inverse spectral problem of wave propagation in stratified media. The problem is to recover a measure σ from a set (finite or infinite) of reflection coefficients

$$a_n \stackrel{\text{def}}{=} -\frac{\overline{\varphi_{n+1}(0)}}{\kappa_{n+1}}, \qquad n \in \mathbb{Z}_+,$$

which are also known as the *Geronimus parameters* of σ. Although by Favard's theorem for the unit circle [3, 4, 19] the sequence $\{a_n\}_{n \geqslant 0}$ uniquely determines the measure σ, the correspondence of parameters to measures is unstable. For instance, it is shown in [19, Corollary 9.2] that given any Szegő measure σ, i.e., a measure with the Geronimus parameters satisfying

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty, \tag{1}$$

and an arbitrary $\varepsilon > 0$, there exists a singular measure σ^* with the Geronimus parameters $\{a_n^*\}_{n \geqslant 0}$ which differ from $\{a_n\}_{n \geqslant 0}$ on an arbitrarily rare subset $\Lambda \subset \mathbb{Z}_+$ and $|a_n^*| < \varepsilon$ for $n \in \Lambda$. The latter means that for any set of data satisfying (1) there always exists a singular measure with "practically the same" Geronimus parameters. However, it is proved in [19, Theorem 4] that all such measures are in R. Therefore, dealing with the inverse spectral problem it is natural to expect the existence of estimates for the averages

$$\frac{1}{m(I)} \int_I |\varphi_n|^2\, d\sigma - 1,$$

I being an open arc on $\mathbb{T}$, rather than for the density $\sigma' = d\sigma/dm$.

Another important feature of R is that it is exactly the class of $\sigma \in \mathscr{P}$ admitting the ratio asymptotics of Szegő type for monic orthogonal polynomials [19, Theorem 7.4]. In the present paper we develop some tools to extend beyond R the known methods of research.

Let $\mathrm{Lim}(\sigma)$ stand for the derived set (i.e., the set of all limit points) in $\mathscr{P}$ of the sequence $|\varphi_n|^2\, d\sigma$. Regarding to the properties of $\mathrm{Lim}(\sigma)$ the set $\mathscr{P} \setminus R$ splits into two big classes. The first one is the class $\mathrm{Mar}(\mathbb{T})$ of Markoff measures σ with the property $m \notin \mathrm{Lim}(\sigma)$. The rest is the class $\mathrm{Res}(\mathbb{T})$ which is defined by $m \in \mathrm{Lim}(\sigma)$. This classification was introduced in a recent paper [21], in which more details are available. We only notice that $\mathrm{Mar}(\mathbb{T})$ is opposite in a sense to R and $\mathrm{Res}(\mathbb{T})$. For instance, it is easy to see that any measure σ with $\mathrm{supp}\,\sigma \neq \mathbb{T}$ is a Markoff measure, whereas $\mathrm{supp}\,\sigma = \mathbb{T}$ for any measure in R and $\mathrm{Res}(\mathbb{T})$ since $\mathrm{Lim}(\sigma)$ contains m.

Theorem 7.5 of [19] claims that the relation

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} |a_k| = 0 \tag{2}$$

holds for the Geronimus parameters $\{a_n\}_{n \geqslant 0}$ of any Rakhmanov's measure. We call (2) the Cesàro–Nevai condition and the corresponding class of measures the Cesàro–Nevai class or the CN class (recall that Nevai's class N is defined by $\lim_{n \to \infty} a_n = 0$).

It was Ya. L. Geronimus, who first paid attention to condition (2). In [8, p. 128] he found a simple bound for the monic orthogonal polynomials $\Phi_n \overset{\mathrm{def}}{=} \kappa_n^{-1} \varphi_n$

$$\frac{1}{n+1} \log \sup_{|z| \leqslant 1} |\Phi_{n+1}(z)| \leqslant \frac{1}{n+1} \sum_{k=0}^{n} |a_k|. \tag{3}$$

Since for an arbitrary monic polynomial P the inequality $\sup_{|z| \leqslant 1} |P| \geqslant 1$ holds we deduce from (2) and (3) that

$$\lim_{n \to \infty} \sup_{|z| \leqslant 1} |\Phi_n(z)|^{1/n} = 1. \tag{4}$$

In other words, the sequence $\{\Phi_n\}$ is an extremal sequence for the closed unit disk (see [24, Chap. 5, Proposition 5.3]).

This observation gives some hope that the orthogonal polynomials for measures $\sigma \in \mathrm{CN}$ behave regularly. As it turns out, this hope is justified to some extent. To state the results obtained we need some preliminaries as well as some basic facts from [19], which will be used repeatedly throughout the whole paper.

Recall that the Herglotz formula

$$\int_{\mathbb{T}} \frac{\zeta+z}{\zeta-z}\, d\sigma(\zeta) = \frac{1+zf(z)}{1-zf(z)} \tag{5}$$

establishes a one-to-one correspondence $f = \mathscr{H}(\sigma)$ between $\sigma \in \mathscr{P}$ and the functions from the unit ball $\mathscr{B}$ of the Hardy algebra H^∞ in the unit disk $\mathbb{D} = \{z\colon |z| < 1\}$, i.e., the set of holomorphic functions in $\mathbb{D}$ with

$$\|f\|_\infty \overset{\text{def}}{=} \sup_{|z|<1} |f(z)| \leqslant 1.$$

Equipped with the topology of uniform convergence on compact subsets of $\mathbb{D}$ the set $\mathscr{B}$ is a compact metric space. Then the mapping $\mathscr{H}\colon \mathscr{P} \to \mathscr{B}$ in (5) is a homeomorphism. The function $f = \mathscr{H}(\sigma)$ is called the *Schur function* of σ.

Given $f \in \mathscr{B}$, Schur's algorithm

$$f(z) = f_0(z) = \frac{zf_1(z)+a_0}{1+\bar{a}_0 z f_1(z)}, \ldots, \qquad f_n(z) = \frac{zf_{n+1}(z)+a_n}{1+\bar{a}_n z f_{n+1}(z)}, \ldots \tag{6}$$

is uniquely defined by Schwartz' lemma [5, 19]. The numbers $\{a_n\}_{n\geqslant 0}$ are called the *Schur parameters* of f. If $f = \mathscr{H}(\sigma)$, the Schur parameters of f agree with the Geronimus parameters of σ by Geronimus' theorem [7, 19]. The functions $\{f_n\}_{n\geqslant 0}$ in (6) are called the *direct Schur functions* of σ.

From the definition of Schur's algorithm (6) it follows that the Schur parameters

$$\mathscr{S} f = (a_0, a_1, a_2, \ldots)$$

of $f \in \mathscr{B}$ form either an infinite sequence $a = (a_n)_{n\geqslant 0}$ with the domain $\mathscr{D}(a) = \mathbb{Z}_+$ (and in this case $|a_n| < 1$ for all n) or a finite sequence $a = (a_n)_{n=0}^k$ with the domain $\mathscr{D}(a) = [0, k]$ (and in this case $|a_n| < 1$ for $n = 0, 1, \ldots, k-1$ and $|a_k| = 1$). Note that

$$\mathscr{S} f_n = (a_n, a_{n+1}, \ldots).$$

Let $\mathscr{G}^\infty$ be the set of all complex sequences a which satisfy

$$\begin{aligned} &|a_n| < 1, \qquad \forall n = 0, 1, \ldots, \qquad &&\text{if } \mathscr{D}(a) = \mathbb{Z}_+;\\ &|a_n| < 1, \qquad 0 \leqslant n < k,\ |a_k| = 1, \qquad &&\text{if } \mathscr{D}(a) = [0, k]. \end{aligned}$$

It is clear that the set $\mathscr{G}^\infty$ equipped with the topology of pointwise convergence is a compact metric space.

We proceed with two fundamental results from [19]. The first one is actually contained in [19, Lemma 4.11] and the remark after it.

THEOREM A. *The mapping* $\mathcal{S}: \mathcal{B} \to \mathcal{G}^\infty$ *is a homeomorphism of compact metric spaces.*

Thus we have a sequence of homeomorphisms

$$\mathcal{P} \xrightarrow{\mathcal{H}} \mathcal{B} \xrightarrow{\mathcal{S}} \mathcal{G}^\infty.$$

To formulate the second result we introduce the *inverse Schur functions* b_n by the equality

$$b_n(z) \stackrel{\text{def}}{=} \frac{\varphi_n(z)}{\varphi_n^*(z)}, \qquad n \in \mathbb{Z}_+,$$

where $\varphi_n^*(z) \stackrel{\text{def}}{=} z^n \overline{\varphi_n(1/\bar{z})}$. Since the zeros of φ_n lie in $\mathbb{D}$ (cf. [30, Theorem 11.4.1]), it is easy to check that b_n are finite Blaschke products. The Szegő recurrence relations for the orthonormal polynomials [30, formulae (11.4.6)–(11.4.7)] yield

$$b_{n+1}(z) = \frac{zb_n(z) - \bar{a}_n}{1 - a_n z b_n(z)}, \qquad n \in \mathbb{Z}_+,$$

and hence

$$\mathcal{S} b_n = (-\bar{a}_{n-1}, -\bar{a}_{n-2}, \ldots, 1).$$

We see that the parameters here follow in the reversed order.

We can now exhibit the following result (see [19, Theorem 3]).

THEOREM B. *Let* $\sigma \in \mathcal{P}$ *with the orthonormal polynomials* $\{\varphi_n\}_{n \geqslant 0}$ *and the Schur function* $\mathcal{H}(\sigma) = f$. *Then*

$$\mathcal{H}(|\varphi_n|^2 \, d\sigma) = f_n b_n, \qquad n \in \mathbb{Z}_+.$$

Schur's algorithm can also be represented in the form of a continued fraction

$$f(z) = a_0 + \frac{(1-|a_0|^2)\, z}{\bar{a}_0 z} \; + \frac{1}{a_1} + \frac{(1-|a_1|^2)\, z}{\bar{a}_1 z} \; + \cdots + \frac{1}{a_n} + \frac{(1-|a_n|^2)\, z}{\bar{a}_n z} \; + \cdots \tag{7}$$

(see [19, 29]). The approximants A_n/B_n of order $2n$ for (7) converge to f in $\mathscr{B}$. We call the polynomials A_n, B_n of degree n the *Wall polynomials* (of f). Since the even part of (7) is given by

$$f(z)=\frac{a_0}{1}-\frac{(1-|a_0|^2)(a_1/a_0)\,z}{1+(a_1/a_0)\,z}-\cdots-\frac{(1-|a_{n-1}|^2)(a_n/a_{n-1})\,z}{1+(a_n/a_{n-1})\,z}-\cdots, \tag{8}$$

A_n/B_n are the approximants of the Geronimus continued fraction (8).

The paper is organized as follows. The main technical tool of our investigation is the theory of strongly summable and almost convergent sequences (cf. [32, Chap. 13.7]). We collect the required results in Section 2. One of the main results herein is Theorem 2.6, which seems to be new.

The CN class is studied in Section 3. We obtain a description of this class in terms of the orthogonal polynomials as well as of functions associated with the Schur function $f=\mathscr{H}(\sigma)$. We also establish the relationship between the CN class and the Cesàro–Rakhmanov class CR, which is defined as the class of measures $\sigma\in\mathscr{P}$ with

$$*-\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}|\varphi_k|^2\,d\sigma=dm. \tag{9}$$

This class appeared in a recent research of the first author on the zeros of para-orthogonal polynomials [16]. It is shown in Theorem 3.1 that $\mathrm{CN}\subset\mathrm{CR}$. One of the main results here is the following

THEOREM 3.6. *Let $\sigma\in\mathscr{P}$ with the Schur function $f=\mathscr{H}(\sigma)$ and the Wall polynomials $\{A_n\}_{n\geqslant 0}$, $\{B_n\}_{n\geqslant 0}$. Then*

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\int_{\mathbb{T}}\left|f-\frac{A_k}{B_k}\right|^2 dm=0$$

if and only if σ is either singular or belongs to the Cesàro–Nevai class.

It is natural to expect that the number of "good" properties of the orthogonal polynomials diminishes (and, conversely, the number of weird properties increases) as soon as $\mathrm{Lim}\,(\sigma)$ is getting larger. In Section 4 we study the extreme case when $\mathrm{Lim}(\sigma)=\mathscr{P}$ or, in other words the sequence $\{|\varphi_n|^2\,d\sigma\}_{n\geqslant 0}$ is dense in $\mathscr{P}$. We call such measures *universal*. The very existence of universal measures, which may exist only in $\mathrm{Res}(\mathbb{T})$, is far from being obvious. However, not only universal measures do exist, but there are so many of them that they penetrate into the relatively small class CN.

THEOREM 4.3. *The Cesàro–Nevai class contains universal measures.*

This result coupled with (4) shows that there are universal measures with monic orthogonal polynomials forming an extremal sequence for the closed unit disk. We also prove that any sequence of parameters $\{d_n\}_{n\geqslant 0}$ (i.e., $|d_n|<1$ for all $n\in\mathbb{Z}_+$) which satisfy $\lim_{n\to\infty}|a_n-d_n|=0$, corresponds to a universal measure as long as $\{a_n\}_{n\geqslant 0}$ does.

The universal measures are singular with respect to Lebesgue measure m (Theorem 4.1), and hence their Schur functions are inner. By using Frostman's theorem we construct universal measures σ with $\mathscr{H}(\sigma)$ being a Blaschke product.

As we already mentioned the class Mar($\mathbb{T}$) is opposite to both R and Res($\mathbb{T}$). This observation allows one to obtain the opposite results to the known ones in R. The idea is to replace the condition imposed on the functionals which describe N or CN with the opposite conditions.

An important result in this respect was obtained by Máté, *et al.* in [22, Theorem 4]. Consider the following functional

$$I_n(z)\stackrel{\text{def}}{=}\frac{1}{|\varphi_n(z)|^2}\sum_{k=0}^{n}|\varphi_k(z)|^2\leqslant\infty,\qquad |z|\leqslant 1. \tag{10}$$

THEOREM [22, Theorem 4]. *If $\sigma\in N$, then*

$$\lim_{n\to\infty}\max_{|z|\leqslant 1}\frac{1}{I_n(z)}=0.$$

Conversely, if $\lim_{n\to\infty}1/I_n(z_0)=0$ *at some point $z_0\in\mathbb{D}$, then $\sigma\in N$.*

It turns out that the version of this result for the CN class is true.

THEOREM 5.1. *For each $\sigma\in CN$ there exists a subset $\Lambda\subset\mathbb{Z}_+$ of density* 1 *such that*

$$\lim_{n\in\Lambda}\max_{|z|\leqslant 1}\frac{1}{I_n(z)}=0.$$

Conversely, if $\lim_{n\in\Lambda}1/I_n(z_0)=0$ *for some $z_0\in\mathbb{D}$ and some subset $\Lambda\subset\mathbb{Z}_+$ of density* 1, *then $\sigma\in CN$.*

For general measures from $\mathscr{P}$ the asymptotic behavior of the functionals I_n as $n\to\infty$ is rather unpredictable. By MNT's theorem within Nevai's class I_n tends to infinity uniformly on $\mathbb{T}$. Let G_∞ be the class of measures in $\mathscr{P}$ which satisfy

$$C_\sigma(\zeta)\stackrel{\text{def}}{=}\sup_n I_n(\zeta)=+\infty,\qquad \zeta\in\mathbb{T}.$$

Theorem 5.1 states that $\mathrm{CN} \subset G_\infty$. On the other hand, it is not hard to estimate $I_n(\zeta)$ in terms of the distance $\operatorname{dist}(\zeta, \operatorname{supp} \sigma)$. Indeed, consider the Szegő kernel

$$K_{n+1}(z, \zeta) \stackrel{\text{def}}{=} \sum_{k=0}^{n} \varphi_k(z)\, \overline{\varphi_k(\zeta)} = \frac{\varphi_n^*(z)\, \overline{\varphi_n^*(\zeta)} - z\bar{\zeta} \varphi_n(z)\, \overline{\varphi_n(\zeta)}}{1 - z\bar{\zeta}} \tag{11}$$

(the latter equality is the Christoffel–Darboux formula [9, formula (1.7)]). We have

$$\begin{aligned} I_n(z) &= \frac{K_{n+1}(z, z)}{|\varphi_n(z)|^2} = \frac{1}{|\varphi_n(z)|^2} \int_{\mathbb{T}} |K_{n+1}(z, \zeta)|^2\, d\sigma(\zeta) \leqslant 4 \int_{\mathbb{T}} \frac{|\varphi_n(\zeta)|^2}{|z - \zeta|^2}\, d\sigma(\zeta) \\ &\leqslant \frac{4}{\operatorname{dist}^2(z, \operatorname{supp} \sigma)} \end{aligned} \tag{12}$$

for $z \in \mathbb{T} \setminus \operatorname{supp} \sigma$. It is clear now that $\operatorname{supp} \sigma = \mathbb{T}$ for each $\sigma \in G_\infty$.

We show in Section 5 (Proposition 5.3) that for $\sigma \in \mathscr{P}$ the function C_σ cannot be finite *everywhere* on $\mathbb{T}$. That is why the "opposite" to G_∞ class G_0 is defined as the class of measures for which $C_\sigma < \infty$ *almost everywhere* with respect to Lebesgue measure m. It is clear from (12) that G_0 contains all singular measures with $m(\operatorname{supp} \sigma) = 0$. For instance, the measures with finite derived set of support [14, Section 3] belong to G_0.

Similarly, let us introduce the "opposite" Szegő class OS. A measure σ is said to belong to OS if its Geronimus parameters satisfy

$$\sum_{n=0}^{\infty} (1 - |a_n|^2)^{1/2} < \infty. \tag{13}$$

To justify the name of this class it is worth comparing (13) with (1) which determines the Szegő class. To obtain the first one from the second we replace the distance $|a_n| = \operatorname{dist}(a_n, 0)$ with $1 - |a_n| = \operatorname{dist}(a_n, \mathbb{T})$ (i.e., replace 0 with the unit circle which is "opposed" to 0 in $\mathbb{D}$) and take square root instead of squaring (the opposite operation). It is proved in Theorem 5.7 that $\mathrm{OS} \subset G_0$. A number of explicit examples of measures from OS with the Geronimus parameters tending to $\mathbb{T}$ exponentially fast are known (cf. [31, Section 6]). One of them is presented in example 5.10.

The class OS can be described in terms of Hessenberg matrices. Recall that the Hessenberg matrix of a measure $\sigma \in \mathscr{P}$ is defined by

$$\hat{U}(\sigma) = \begin{pmatrix} u_{00} & u_{01} & \cdots \\ u_{10} & u_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \qquad u_{ij} = (U(\sigma)\, \varphi_j, \varphi_i)_\mu, \tag{14}$$

where $U(\sigma)\,h=\zeta h(\zeta)$, $h\in L^2(d\sigma)$ is the unitary multiplication operator. The matrix elements are given by the formulae

$$u_{ij}=\begin{cases} -a_j\bar{a}_{i-1}\prod_{k=i}^{j-1}(1-|a_k|^2)^{1/2}, & \text{for } i<j+1\\ (1-|a_j|^2)^{1/2}, & \text{for } i=j+1\\ 0, & \text{for } i>j+1\end{cases}\tag{15}$$

(cf. [17, p. 401]) with $a_{-1}=-1$.[2] The main tool here is the (formal) series expansion introduced in [17, p. 403]

$$U(\sigma)=V^*D_{-1}(\sigma)+\sum_{j=0}^{\infty}D_j(\sigma)\,\mathrm{V}^{\mathrm{j}},\tag{16}$$

where V, V^* are the shift operators and $D_j(\sigma)$ are the diagonal operators

$$D_j(\sigma)=\operatorname{diag}(u_{0j},u_{1,j+1},\ldots)$$

in the basis $\{\varphi_n\}_{n\geqslant 0}$.

Let $T=U\,|T|=U(T^*T)^{1/2}$ be the polar decomposition of a bounded linear operator T on a Hilbert space $\mathfrak{H}$. Here U is a partial isometry mapping the range of T^* isometrically onto the range of T. Let $\mathfrak{S}_\infty$ be the set of all compact operators. If $T\in\mathfrak{S}_\infty$, then $|T|=(T^*T)^{1/2}\in\mathfrak{S}_\infty$. The eigenvalues of $|T|$ are called the s-numbers $s_n(T)$ of T. The Schatten–von Neumann classes $\mathfrak{S}_p$, $0<p<\infty$, consist of operators with

$$\|T\|_p\overset{\text{def}}{=}\left(\sum_{n=0}^{\infty}s_n^p(T)\right)^{1/p}<\infty$$

(cf. [2, Chap. XI.9]). The class $\mathfrak{S}_1$ is also known as the *trace class* or the class of nuclear operators.

THEOREM 5.8. *$\sigma\in OS$ if and only if*

$$U(\sigma)-D_0(\sigma)\in\mathfrak{S}_1.$$

The diagonal of $D_0(\sigma)$ is given by the vector

$$(\bar{a}_0,-\bar{a}_1a_0,\ldots,-\bar{a}_na_{n-1},\ldots).\tag{17}$$

Note that only a finite number of the entries in (17) may vanish for $\sigma\in\mathrm{OS}$. If $\bar{a}_na_{n-1}\neq 0$ we define $\tau_n\in\mathbb{T}$ to be the closest point to $-\bar{a}_na_{n-1}$ on $\mathbb{T}$.

[2] The notation here is slightly different from [17].

Otherwise put $\tau_n = 1$. Denote by $\mathrm{Clos}\{\tau_0, \tau_1, ...,\}$ the closure of the set $\{\tau_n\}_{n \geqslant 0}$ on $\mathbb{T}$.

THEOREM 5.9. *Let $\sigma \in OS$ with the monic orthogonal polynomials $\{\Phi_n\}_{n \geqslant 0}$. Then there is a holomorphic function H in the domain $\hat{\mathbb{C}} \setminus \mathrm{Clos}\{\tau_0, \tau_1, ...,\}$ which is not identically zero in both components of $\hat{\mathbb{C}} \setminus \mathbb{T}$ and such that*

$$\lim_{n \to \infty} \frac{\Phi_{n+1}(z)}{(z-\tau_0)(z-\tau_1)\cdots(z-\tau_n)} = H(z)$$

uniformly on compact subsets of $\hat{\mathbb{C}} \setminus \mathrm{Clos}\{\tau_0, \tau_1, ...,\}$.

This theorem shows that for measures in OS the monic orthogonal polynomials behave like polynomials with the roots on $\mathbb{T}$.

The proof is based heavily on the notion of *infinite determinants* (cf. [11, Chap. 4, Sect. 1]). Given $A \in \mathfrak{S}_1$, an infinite determinant $\det(I+A)$ is defined by the formula

$$\det(I+A) \overset{\text{def}}{=} \prod_{j=1}^{\infty} (1+\lambda_j(A)),$$

where $\lambda_j(A)$, $j \geqslant 1$ are the eigenvalues of A labelled in the decreasing order of their moduli (each is counted according to its algebraic multiplicity). If $A(z)$ is a holomorphic operator-function on a region G with the values in $\mathfrak{S}_1$, then $\det(I+A(z))$ is a holomorphic function on G.

Let $\{e_n\}_{n \geqslant 0}$ be an orthonormal basis in $\mathfrak{G}$. We let A correspond to its matrix

$$A \to \|a_{kj}\|_{k,j=0}^{\infty}, \qquad a_{kj} = (Ae_j, e_k)$$

in this basis, and a sequence of truncated $n+1$-dimensional operators (matrices) $A_n = \|a_{kj}\|_{k,j=0}^{n}$, $n \in \mathbb{Z}_+$. One of the basic properties of infinite determinants is the following limit relation

$$\det(I+A) = \lim_{n \to \infty} \det(I_n + A_n), \qquad I_n = \|\delta_{kj}\|_{k,j=0}^{n}.$$

For holomorphic operator-functions $A(z)$ we have

$$\det(I+A(z)) = \lim_{n \to \infty} \det(I_n + A_n(z))$$

uniformly on compact subsets of G.

In Section 6 we discuss some further properties of functional I_n (10). It is well known that it can be used for the solution of a mass point problem.

Given $\sigma \in \mathscr{P}$ and a Dirac measure δ_{ζ_0} at some point $\zeta_0 \in \mathbb{T}$, consider a

$$\sigma_t \stackrel{\text{def}}{=} (1-t)\,\sigma + t\,\delta_{\zeta_0}, \qquad 0 \leqslant t < 1. \tag{18}$$

A class $X \subset \mathscr{P}$ will be called *mass points invariant* if $\sigma_t \in X$ provided $\sigma_0 = \sigma \in X$. We show that CN is mass points invariant and give some partial answer regarding the class R.

In Section 6 we also relate the functional I_n to the distribution of zeros of orthogonal polynomials φ_n. Let $\lambda_{1n}, \lambda_{2n}, \ldots, \lambda_{nn}$ be the zeros of φ_n for $n \geqslant 1$ and let α_n be the argument of b_n on $\mathbb{T}$, that is, $b_n = \exp(i\alpha_n)$. Then

$$I_n(\zeta) = 1 + \sum_{j=1}^{n} \frac{1-|\lambda_{jn}|^2}{|\zeta - \lambda_{jn}|^2} = 1 + \dot{\alpha}_n(\zeta), \tag{19}$$

where $\dot{\alpha}_n = \partial\alpha_n / \partial\vartheta$.

For the reader's (and author's) convenience we complete the introduction with two well known results from Real and Complex Analysis. The first one is the famous Lebesgue theorem on differentiation (cf. [2, Theorem III.12.6]). We need a particular case related to the unit circle.

Given $\mu \in \mathscr{P}$ we call an open arc $I \subset \mathbb{T}$ μ-regular if its endpoints are not μ-masspoints.

THEOREM (Lebesgue). *Let $\mu = \mu'\,dm + \mu_s$ be Lebesgue's decomposition of $\mu \in \mathscr{P}$. Then for any sequence of μ-regular arcs $\{I_n\}$ such that $\zeta \in I_n$ and $\lim_{n\to\infty} m(I_n) = 0$ the relation*

$$\lim_{n\to\infty} \frac{\mu(I_n)}{m(I_n)} = \mu'(\zeta)$$

holds m-a.e.

The second result known as the Khinchin–A. Ostrowski theorem deals with sequences of holomorphic functions in the unit disk (cf. [26, Chap. 2, Sect. 7]). A special case pertaining to the class $\mathscr{B}$ is of our main concern.

THEOREM (Khinchin–A. Ostrowski). *Let $\{h_n\}_{n\geqslant 0}$ be a sequence of functions in $\mathscr{B}$. Assume that their boundary values $h_n(\zeta)$ converge in measure, $h_n \Rightarrow h$, on a set $E \subset \mathbb{T}$ of positive Lebesgue measure $m(E) > 0$. Then there is a function $\hat{h} \in \mathscr{B}$ such that $h_n \to \hat{h}$ in $\mathscr{B}$, i.e., uniformly on compact subsets of $\mathbb{D}$, and $\hat{h} = h$ a.e. on E.*

2. STRONG SUMMABILITY AND ALMOST CONVERGENCE

The Cesàro method of summability is closely related to the notions of strongly summable and almost convergent sequences (cf. [32, Chap. 13.7]).

DEFINITION. A sequence $\{s_n\}_{n\geqslant 0}$ of complex numbers is called *strongly summable* to limit s if

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}|s_k-s|=0.$$

In the study of strongly summable sequences an important role is played by those subsets Λ of the set $\mathbb{Z}_+=\{0, 1, 2, ...\}$ that have density 0 or 1.

DEFINITION. A sequence (subset) $\Lambda\subset\mathbb{Z}_+$ is said to have a *density d* if

$$d(\Lambda)\overset{\text{def}}{=}\lim_{n\to\infty}\frac{\operatorname{Card}(\Lambda\cap\{0, 1, ..., n\})}{n+1}=d$$

exists. Here Card X is the number of points in X.

As it has already been mentioned, we will primarily be interested in the cases $d=0, 1$. Given $\Lambda\subset\mathbb{Z}_+$ we denote by Λ^c the complement $\mathbb{Z}_+\setminus\Lambda$. The following properties of the functional $d(\Lambda)$ are quite obvious.

(i) $d(\Lambda^c)=1-d(\Lambda)$.

(ii) For $k\in\mathbb{Z}\overset{\text{def}}{=}\{..., -1, 0, 1, ...\}$ let $\Lambda+k=\{n+k : n\in\Lambda, n+k\geqslant 0\}$ be a shift of Λ. Then $d(\Lambda+k)=d(\Lambda)$.

(iii) The union (intersection) of a finite number of sequences of density 0(1) is a sequence of density 0 (1).

Let $\Lambda=\{n_1<n_2<...\}$. It is clear from the definition that

$$d(\Lambda)=0(1)\Leftrightarrow\lim_{k\to\infty}\frac{k}{n_k}=0(1). \tag{20}$$

DEFINITION. A sequence $\{s_n\}_{n\geqslant 0}$ of complex numbers is called *almost convergent* to limit s if there is a sequence $\Lambda\subset\mathbb{Z}_+$ of density 1 such that $s_n\to s$ as $n\to\infty$ along Λ

$$\lim_{n\in\Lambda}s_n=s.$$

By (iii) this type of convergence is well-defined.

The following basic results on strongly summable and almost convergent sequences can be found in [32, Theorem 13.7.2].

PROPOSITION 2.1. *Given a sequence* $\{s_n\}_{n\geqslant 0}$ *of complex numbers,* $s\in\mathbb{C}$ *and* $\varepsilon>0$ *denote by* $\Lambda(s,\varepsilon)=\{n\in\mathbb{Z}_+ : |s_n-s|\leqslant\varepsilon\}$. *Then* $\{s_n\}$ *is an almost convergent to s if and only if* $d\Lambda(s,\varepsilon)=1$ *for all* $\varepsilon>0$.

PROPOSITION 2.2. *If* $\{s_n\}_{n\geqslant 0}$ *is strongly summable to s then* $\{s_n\}$ *is almost convergent to s. Conversely, if* $\{s_n\}$ *is almost convergent to s and* $\{s_n\}$ *is bounded* then s_n is strongly summable to s.

The next helpful property of sequences of density 1 is not that apparent. We provide a proof for reader's convenience.

PROPOSITION 2.3. *Let* $\{\Lambda_k\}_{k\geqslant 1}$ *be a denumerable set of sequences of density* 1. *Then there exists a sequence* Λ *of density* 1 *such that*

$$Card(\Lambda\cap\Lambda_k^c)<\infty,\qquad k=1,2,\ldots. \tag{21}$$

Proof. Put $M_n=\bigcap_{k=1}^n\Lambda_k$. Then $d(M_n)=1$, $n=1,2,\ldots$ by (iii). Define a sequence of indices $N_1<N_2<\cdots$ by the recipe

$$\frac{\text{Card}(M_k\cap[0,n])}{n+1}>1-\frac{1}{k},\qquad n>N_k,\quad k=1,2,\ldots, \tag{22}$$

and consider the set

$$\Lambda\stackrel{\text{def}}{=}[0,N_1]\cup\left(\bigcup_{k=1}^{\infty}M_k\cap(N_k,N_{k+1}]\right).$$

Since $\{M_n\}$ is a decreasing family of subsets in $\mathbb{Z}_+$, we have for $N_j<N\leqslant N_{j+1}$

$$\Lambda\supset[0,N_1]\cup\left(\bigcup_{m=1}^{j}M_m\cap(N_m,N_{m+1}]\right)\supset M_j\cap[0,N_{j+1}],$$

which implies that $\Lambda\cap[0,N]\supset M_j\cap[0,N]$ and

$$\frac{\text{Card}(\Lambda\cap[0,N])}{N+1}\geqslant\frac{\text{Card}(M_j\cap[0,N])}{N+1}>1-\frac{1}{j}$$

by (22). Hence $d(\Lambda)=1$. Next, $M_j\cap\Lambda_k^c=\varnothing$ for $j\geqslant k$ by the definition of the sets M_n, which proves (21). ∎

The following simple consequence of Proposition 2.3 proves useful in the sequel.

COROLLARY 2.4. *Let $\{s_n\}$ be almost convergent to zero. Then there is a sequence Λ of density* 1 *such that*

$$\lim_{n \in \Lambda} s_{n+k} = 0 \tag{23}$$

for each fixed $k \in \mathbb{Z}$. In particular, $\{s_{n+k}\}_{n \geqslant 0}$ is almost convergent to zero.

Proof. By the definition there is a sequence Λ_0 with $d(\Lambda_0) = 1$ such that $s_n \to 0$, $n \in \Lambda_0$. For $k \in \mathbb{Z}$ put $\Lambda_k = \Lambda - k$. Then $s_{n+k} \to 0$, as $n \to \infty$, $n \in \Lambda_k$. Now by Proposition 2.3 we have a sequence Λ of density 1 with property (21). It is clear that $s_{n+k} \to 0$, $n \in \Lambda$ for each fixed $k \in \mathbb{Z}$, as claimed. ∎

The next result is adopted from [19, Theorem 7.5]. A simple proof of this result was suggested by the referee. We present it here for the sake of completeness. Note that the sequence $\{s_n\}_{n \geqslant 0}$ below is not assumed to be bounded.

PROPOSITION 2.5. *Let $\{s_n\}_{n \geqslant 0}$ be a sequence of complex numbers such that*

$$\lim_{n \to \infty} s_n s_{n+k} = 0 \tag{24}$$

for $k = 1, 2, \ldots$. Then $\{s_n\}_{n \geqslant 0}$ is almost convergent to zero.

Proof. Assume, on the contrary, that for some $\varepsilon > 0$, the set

$$\{n: |s_n| > \varepsilon\}$$

does not have zero density (cf. Proposition 2.1). Then, there is an integer $k > 1$ such that for every $N \geqslant 0$ there exists an integer $M > N + k$ such that, with

$$A = \{n: N < n < M \text{ and } |s_n| > \varepsilon\}$$

we have $\mathrm{Card}(A) > 2M/k$. Then the inequality $k \cdot 2M/k > M + k$ implies that the sets A, $A+1$, $A+2$, ..., $A+k-1$ cannot all be pairwise disjoint, showing that

$$\limsup_{n \to \infty} |s_n s_{n+j}| \geqslant \varepsilon^2$$

for some positive integer $j < k$. ∎

Remark. We consider it necessary to include here the following remark due to the referee which sheds some light on Proposition 2.5.

The result is certainly quite simple, but it has the following deep generalization:

Let $m \geqslant 2$ be an integer and let $\{s_n\}_{n \geqslant 0}$ be a sequence of complex numbers such that

$$\lim_{n \to \infty} \prod_{l=0}^{m-1} s_{n+kl} = 0$$

for every positive integer k. Then $\{s_n\}_{n \geqslant 0}$ is almost convergent to 0.

This is a consequence of a strengthening of Endre Szemerédi's famous theorem: *Every set S of positive upper density of positive integers includes arbitrarily long arithmetic progressions.*

The case $m = 2$ is exactly Proposition 2.5.

Our final result provides equivalent definitions of almost convergent sequences.

THEOREM 2.6. *Let $\{s_n\}_{n \geqslant 0}$ be a sequence of complex numbers. The following conditions are equivalent.*

(i) *$\{s_n\}_{n \geqslant 0}$ is almost convergent to zero.*

(ii) *There is a sequence $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that $\lim_{n \in \Lambda} s_{n+k} = 0$ for every $k \in \mathbb{Z}$.*

(iii) *There is a sequence $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that*

$$\lim_{n \in \Lambda} s_n s_{n+k} = 0, \qquad k = 1, 2, \ldots . \tag{25}$$

Proof. (i) $\Rightarrow$ (ii) by Corollary 2.4

(ii) $\Rightarrow$ (iii) is obvious as both s_n and s_{n+k} tend to zero along Λ and so does their product.

(iii) $\Rightarrow$ (i) Define an auxiliary sequence

$$s_n' = \begin{cases} s_n, & \text{for} \quad n \in \Lambda, \\ 0, & \text{for} \quad n \notin \Lambda. \end{cases}$$

It is clear that $s_n' s_{n+k}' \to 0$ as $n \to \infty$. By Proposition 2.5 the sequence $\{s_n'\}_{n \geqslant 0}$ is almost convergent to zero. Hence $s_n' \to 0$ along some sequence Λ' of density 1. By restricting n to the sequence $\Lambda \cap \Lambda'$ of density 1 (wherein $s_n' = s_n$) we see that s_n is almost convergent to zero, as needed. ∎

3. THE CESÀRO–NEVAI CLASS

Let σ be a probability measure on $\mathbb{T}$ with infinite support, $\{a_n\}_{n\geqslant 0}$ the Geronimus parameters of σ and $\{\varphi_n\}_{n\geqslant 0}$ the orthonormal polynomials in $L^2(d\sigma)$.

DEFINITION. A measure $\sigma \in \mathscr{P}$ belongs to the *Cesàro–Nevai* class $\sigma \in \mathrm{CN}$ if

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} |a_k| = 0. \tag{26}$$

Since the Geronimus parameters of each probability measure lie in the unit disk $\mathbb{D}$, Proposition 2.2 says that $\sigma \in \mathrm{CN}$ if and only if $\{a_n\}_{n\geqslant 0}$ is almost convergent to 0. Clearly, Nevai's class N is contained in CN. Moreover, in view of [19, Theorems 4 and 7.5] Rakhmanov's class R is also contained in CN.

We begin with the characterization of the class CN in terms of the direct Schur functions $\{f_n\}_{n\geqslant 0}$ and the inverse Schur functions $\{b_n\}_{n\geqslant 0}$ of σ. For $\sigma \in \mathscr{P}$ we call an open arc I *σ-regular* if its endpoints do not contain atoms of σ.

THEOREM 3.1. *The following conditions are equivalent.*

(i) *σ is in Cesàro–Nevai class.*

(ii) *The Schur functions f_n, b_n satisfy*

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} |f_k(z)\, b_k(z)| = 0.$$

(iii) *The Schur functions f_n satisfy*

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} |f_k(z)| = 0.$$

(iv) *The Schur functions b_n satisfy*

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} |b_k(z)| = 0.$$

The convergence in (ii)–(iv) *is understood in $\mathscr{B}$, i.e., uniformly on compact subsets of $\mathbb{D}$.*

(v) *There exists a subset $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that*

$$*-\lim_{n\in\Lambda} |\varphi_n|^2\, d\sigma = dm.$$

(vi) *For any continuous function g on* $\mathbb{T}$ *the sequence*

$$s_n = \int_{\mathbb{T}} g\,|\varphi_n|^2\,d\sigma, \qquad n = 0, 1, 2, \ldots$$

is strongly summable to $s = \int_{\mathbb{T}} g\,dm$.

(vii) *For any* σ*-regular open arc I the sequence*

$$s_n = \int_I |\varphi_n|^2\,d\sigma, \qquad n = 0, 1, 2, \ldots$$

is strongly summable to $s = m(I)$.

Proof. (i) $\Rightarrow$ (ii), (iii), (iv). By Proposition 2.2 and Corollary 2.4 there exists a subset $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that $\lim_{n \in \Lambda} a_{n+k} = 0$ holds for every $k \in \mathbb{Z}$. Since

$$\mathscr{S} f_n = (a_n, a_{n+1}, \ldots), \qquad \mathscr{S} b_n = (-\bar{a}_{n-1}, -\bar{a}_{n-2}, \ldots, 1) \tag{27}$$

we have by Theorem A $\lim_{n \in \Lambda} f_n = \lim_{n \in \Lambda} b_n = 0$ in $\mathscr{B}$. It is clear that all three sequences $\{f_n b_n\}_{n \geqslant 0}$, $\{f_n\}_{n \geqslant 0}$, $\{b_n\}_{n \geqslant 0}$ are strongly summable to zero by Proposition 2.2.

(iii), (iv) $\Rightarrow$ (ii) is obvious since $|f_n| \leqslant 1$, $|b_n| \leqslant 1$.

(ii) $\Rightarrow$ (i). Let $\{z_j\}_{j \geqslant 0}$ be any sequence in $\mathbb{D}$ with $z_n \to z_0$ as $n \to \infty$. For each j the sequence $\{f_n(z_j)\, b_n(z_j)\}_{n \geqslant 0}$ is almost convergent to zero. Let Λ_j be the corresponding subset of $\mathbb{Z}_+$ of density 1. By Proposition 2.3 we can choose $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that

$$\lim_{n \in \Lambda} f_n(z_j)\, b_n(z_j) = 0, \qquad j \in \mathbb{Z}_+.$$

The normal family argument provides $\lim_{n \in \Lambda} f_n b_n = 0$ in $\mathscr{B}$.

The following reasoning is only a slight modification of that in [19, Sect. 7, p. 222]. Let $\Lambda_1 = \Lambda \cap (\Lambda - 1)$. We will show that

$$\lim_{n \in \Lambda_1} a_n a_{n+k} = 0, \qquad k = 1, 2, \ldots. \tag{28}$$

By the definition of b_n and Schur's algorithm we have

$$z b_n f_n = \frac{b_{n+1} + \bar{a}_n}{1 + a_n b_{n+1}} \, \frac{z f_{n+1} + a_n}{1 + \bar{a}_n z f_{n+1}}.$$

Taking into account that $\lim_{n\in\Lambda_1} f_n b_n = \lim_{n\in\Lambda_1} f_{n+1} b_{n+1} = 0$ we obtain as in [19, formula (7.12)]

$$\lim_{n\in\Lambda_1} (a_n b_{n+1}(z) + |a_n|^2 + \bar{a}_n z f_{n+1}(z)) = 0$$

and (28) follows (see [19, Lemma 7.3]). The rest is plain by Theorem 2.6.

(i) $\Rightarrow$ (v). By Proposition 2.2 and Theorem 2.6 there exists a subset $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that

$$\lim_{n\in\Lambda} a_n a_{n+k} = 0 \tag{29}$$

holds for every $k \in \mathbb{Z}$. Next, by [19, formula (7.10)]

$$\left| \int_{\mathbb{T}} \zeta^k \, |\varphi_n|^2 \, d\sigma \right| \leqslant 2(|a_n| + |a_{n+1}| + \cdots + |a_{n+k-1}|)(|a_{n-k}| + |a_{n-k-1}| + \cdots + |a_{n-1}|) \tag{30}$$

holds for every $k = 1, 2, \ldots$ and $n > k$. Hence (30) along with (29) gives

$$\lim_{n\in\Lambda} \int_{\mathbb{T}} \zeta^k \, |\varphi_n|^2 \, d\sigma = 0, \qquad k = 1,2,\ldots,$$

which implies (v).

(v) $\Rightarrow$ (ii). By Theorem B, $f_n b_n$ is the Schur function of the measure $|\varphi_n|^2 \, d\sigma$. As above, we see that $\{f_n b_n\}_{n\geqslant 0}$ is strongly summable to zero uniformly on compact subsets of $\mathbb{D}$.

(v) $\Leftrightarrow$ (vi). Part (v) implies (vi) by Proposition 2.2. Conversely, let $\mathscr{B}_0$ be the unit ball of the space $C(\mathbb{T})$ of all continuous functions on $\mathbb{T}$. Pick any countable everywhere dense set $\{f^{(k)}\}_{k\geqslant 0}$. By Proposition 2.2. for every k there exists a subset $\Lambda_k \subset \mathbb{Z}_+$ with $d(\Lambda_k) = 1$ such that

$$\lim_{n\in\Lambda_k} \int_{\mathbb{T}} f^{(k)} \, |\varphi_n|^2 \, d\sigma = \int_{\mathbb{T}} f^{(k)} \, dm.$$

In view of Proposition 2.3 there is a subset $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that

$$\lim_{n\in\Lambda} \int_{\mathbb{T}} f^{(k)} \, |\varphi_n|^2 \, d\sigma = \int_{\mathbb{T}} f^{(k)} \, dm,$$

which implies (v).

(v) $\Leftrightarrow$ (vii). Part (v) implies (vii) by Helly's Theorem and Proposition 2.2. Conversely, pick any dense countable subset $S \subset \mathbb{T}$ of σ-measure zero and denote by $G(S)$ a countable set of open arcs with endpoints at S (all of them are σ-regular). In view of Proposition 2.3 there is a subset $\Lambda \subset \mathbb{Z}_+$ of density 1 such that

$$\lim_{n \in \Lambda} \int_I |\varphi_n|^2 \, d\sigma = m(I), \qquad \forall I \in G(S).$$

Next, given $f \in C(\mathbb{T})$, and $\varepsilon > 0$, find a finite collection of pairwise disjoint arcs $I_j \in G(S)$ that form a partition of $\mathbb{T}$ such that the oscillation of f over each I_j is less than ε. Define a piecewise constant function f_ε by the equality

$$f_\varepsilon(\zeta) = \frac{1}{m(I_j)} \int_{I_j} f \, dm, \qquad \zeta \in I_j.$$

It is clear that $|f - f_\varepsilon| < \varepsilon$. On the other hand, by the construction

$$\lim_{n \in \Lambda} \int_{\mathbb{T}} f_\varepsilon \, |\varphi_n|^2 \, d\sigma = \int_{\mathbb{T}} f_\varepsilon \, dm$$

and (v) follows. ∎

It is proved in [19, Theorem 4] that Nevai's class N is a proper subset of Rakhmanov's class R.

Definition. A measure $\sigma \in \mathscr{P}$ belongs to the *Cesàro–Rakhmanov* class $\sigma \in CR$ if

$$* - \lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} |\varphi_k|^2 \, d\sigma = dm. \tag{31}$$

Theorem 3.1 (vi) states that $\mathrm{CN} \subset \mathrm{CR}$.

Proposition 3.2. *We have $\sigma \in CR$ if and only if*

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} \frac{f_k b_k}{1 - z f_k b_k} = 0 \tag{32}$$

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

Proof. It is immediate from Theorem B that

$$\int_{\mathbb{T}} \frac{\zeta + z}{\zeta - z} \, \frac{1}{n+1} \sum_{k=0}^{n} |\varphi_k|^2 \, d\sigma = \frac{1}{n+1} \sum_{k=0}^{n} \frac{1 + z f_k(z) \, b_k(z)}{1 - z f_k(z) \, b_k(z)}.$$

The rest is plain. ∎

For the orthonormal polynomials $\{\varphi_n\}_{n\geqslant 0}$ in $L^2(d\sigma)$ we define the following functions on $\mathbb{T}$ associated with $\sigma\in\mathscr{P}$:

$$\xi_n^2 \stackrel{\text{def}}{=} \frac{1}{n+1}\sum_{k=0}^{n} |\varphi_k|^2, \qquad n\in\mathbb{Z}_+, \tag{33}$$

and

$$g_n \stackrel{\text{def}}{=} \frac{2\,|\varphi_n|^2\,\sigma'}{1+|\varphi_n|^2\,\sigma'}, \qquad n\in\mathbb{Z}_+. \tag{34}$$

Since $2x(1+x)^{-1}$ is a concave function and $1/x$ is a convex function on $(0,+\infty)$, we have by Jensen's inequality

$$\frac{1}{n+1}\sum_{k=0}^{n} g_k \leqslant \frac{2\xi_n^2\,\sigma'}{1+\xi_n^2\,\sigma'} \tag{35}$$

and

$$\frac{1}{\xi_n^2} = \frac{n+1}{\sum_{k=0}^{n}|\varphi_k|^2} \leqslant \frac{1}{n+1}\sum_{k=0}^{n}\frac{1}{|\varphi_k|^2}. \tag{36}$$

Proposition 3.3. *For any limit point μ of the sequence $\{\xi_n^{-2}\,dm\}_{n\geqslant 0}$ and every σ-regular arc I*

$$\mu(I)\leqslant\sigma(I) \tag{37}$$

holds. In particular, the set of μ mass points is included in the set of σ mass points. Moreover, if $\sigma\in CR$ then $\mu'=\sigma'$ m-a.e. on $\mathbb{T}$.

Proof. Integrate (36) over I and make $n\to\infty$ along an appropriate subsequence $\Lambda\subset\mathbb{Z}_+$, taking into account that $*-\lim_{n\to\infty}|\varphi_n|^{-2}\,dm=d\sigma$ (cf. [27, Lemma 1]). An application of Lebesgue's theorem on differentiation to (37) leads to $\mu'\leqslant\sigma'$ m-a.e. on $\mathbb{T}$.

To prove the last statement we show that the equality sign prevails in the latter inequality whenever $\sigma\in$ CR. In fact, by Cauchy's inequality

$$\frac{1}{m(I)}\int_I\sqrt{\sigma'}\,dm \leqslant \left(\frac{1}{m(I)}\int_I\frac{1}{n+1}\sum_{k=0}^{n}|\varphi_k|^2\,d\sigma\right)^{1/2}\left(\frac{1}{m(I)}\int_I\frac{dm}{\xi_n^2}\right)^{1/2}. \tag{38}$$

Since I is σ-regular, the first factor on the right-hand side of (38) tends to 1 as $n\to\infty$. Hence

$$\frac{1}{m(I)}\int_I\sqrt{\sigma'}\,dm \leqslant \left(\frac{\mu(I)}{m(I)}\right)^{1/2}.$$

Again, by Lebesgue's theorem on differentiation $\sigma' \leqslant \mu'$ m-a.e. on $\mathbb{T}$, and the statement follows. ∎

The next result can be viewed as the Cesàro version of [19, Theorem 6.3].

THEOREM 3.4. *Let σ be in the Cesàro–Rakhmanov class and let $E(\sigma) \stackrel{\text{def}}{=} \{\sigma' > 0\}$. Then for g_n and ξ_n defined in* (33)–(34)

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \int_{E(\sigma)} (g_k - 1)^2 \, dm = 0 \tag{39}$$

and

$$\lim_{n\to\infty} \int_{E(\sigma)} \left(\frac{2\xi_n^2 \, \sigma'}{1+\xi_n^2 \, \sigma'} - 1 \right)^2 dm = 0. \tag{40}$$

Proof. Let h, H, y, Y be limit points of the sequences

$$h_n = \frac{1}{n+1} \sum_{k=0}^{n} g_k, \; H_n = \frac{1}{n+1} \sum_{k=0}^{n} g_k^2, \; y_n = \frac{2\xi_n^2 \, \sigma'}{1+\xi_n^2 \, \sigma'}, \; Y_n = y_n^2 \tag{41}$$

of bounded measurable functions in the weak-* topology of $L^\infty(\mathbb{T})$. We claim that for each arc I

$$\limsup_{n\to\infty} \frac{1}{m(I)} \int_I z_n \, dm \leqslant 1, \tag{42}$$

where $z_n = h_n, H_n, y_n, Y_n$. Since by Cauchy's inequality (for integrals and sums)

$$\left(\frac{1}{m(I)} \int_I h_n \, dm \right)^2 \leqslant \frac{1}{m(I)} \int_I h_n^2 \, dm \leqslant \frac{1}{m(I)} \int_I H_n \, dm,$$

$$\left(\frac{1}{m(I)} \int_I y_n \, dm \right)^2 \leqslant \frac{1}{m(I)} \int_I Y_n \, dm,$$

it suffices to prove (42) for $z_n = H_n$ and $z_n = Y_n$.

As in [19, Lemma 6.2] we introduce an auxiliary function

$$\Phi(x) = \begin{cases} x, & \text{for} \quad 0 \leqslant x < 1, \\ \dfrac{4x^2}{(1+x)^2}, & \text{for} \quad 1 \leqslant x < \infty, \end{cases} \tag{43}$$

which is an increasing, concave function on $(0, \infty)$ and satisfies

$$\Phi(x) \geqslant \frac{4x^2}{(1+x)^2}, \qquad 0 \leqslant x < \infty. \tag{44}$$

We see that $g_k^2 = \Phi(|\varphi_k|^2 \sigma')$. Now for $z_n = H_n$ in (41) we have by (44) and Jensen's inequality

$$\begin{aligned}\frac{1}{m(I)} \int_I H_n &\leqslant \frac{1}{m(I)} \int_I \frac{1}{n+1} \sum_{k=0}^{n} \Phi(|\varphi_k|^2 \sigma') \\ &\leqslant \Phi\left(\frac{1}{n+1} \sum_{k=0}^{n} \frac{1}{m(I)} \int_I |\varphi_k|^2 \, d\sigma\right) \to \Phi(1) = 1\end{aligned}$$

as $n \to \infty$, for $\sigma \in \mathrm{CR}$. The argument for $z_n = Y_n$ is quite the same, and (42) follows. Lebesgue's theorem on differentiation applied to (42) shows that

$$\max(h, H) \leqslant 1, \qquad \max(y, Y) \leqslant 1 \tag{45}$$

m-a.e. on $\mathbb{T}$.

To prove (39) take any σ-regular arc I and apply three times Cauchy's inequality (for sums and integrals)

$$\begin{aligned}\frac{1}{m(I)} \int_I \sqrt{\sigma'} \, dm &= \frac{1}{m(I)} \int_I \frac{1}{n+1} \sum_{k=0}^{n} \frac{\sqrt{2}\, |\varphi_k| \sqrt{\sigma'}}{1+|\varphi_k|^2 \sigma'^{1/2}} \, \frac{1+|\varphi_k|^2 \sigma'^{1/2}}{\sqrt{2}\, |\varphi_k|} \, dm \\ &\leqslant \frac{1}{m(I)} \int_I \left(\frac{1}{n+1} \sum_{k=0}^{n} g_k\right)^{1/2} \left(\frac{1}{n+1} \sum_{k=0}^{n} \frac{1+|\varphi_k|^2 \sigma'}{2\, |\varphi_k|^2}\right)^{1/2} \\ &\leqslant \left(\frac{1}{m(I)} \int_I \frac{1}{n+1} \sum_{k=0}^{n} g_k \, dm\right)^{1/2} \\ &\quad \times \left(\frac{1}{n+1} \sum_{k=0}^{n} \frac{1}{2m(I)} \int_I \left\{\frac{1}{|\varphi_k|^2} + \sigma'\right\} dm\right)^{1/2} \\ &\leqslant \left(\frac{1}{m(I)} \int_I \frac{1}{n+1} \sum_{k=0}^{n} g_k^2 \, dm\right)^{1/4} \\ &\quad \times \left(\frac{1}{n+1} \sum_{k=0}^{n} \frac{1}{2m(I)} \int_I \left\{\frac{1}{|\varphi_k|^2} + \sigma'\right\} dm\right)^{1/2}.\end{aligned}$$

By passing to the limit as $n \to \infty$ we obtain

$$\frac{1}{m(I)} \int_I \sqrt{\sigma'}\, dm \leqslant \left(\frac{1}{m(I)} \int_I h\, dm \right)^{1/2} \left(\frac{\sigma(I)}{m(I)} \right)^{1/2},$$

$$\frac{1}{m(I)} \int_I \sqrt{\sigma'}\, dm \leqslant \left(\frac{1}{m(I)} \int_I H\, dm \right)^{1/4} \left(\frac{\sigma(I)}{m(I)} \right)^{1/2}.$$

Lebesgue's theorem on differentiation yields $1 \leqslant \min(h, H)$ m-a.e. on $E(\sigma)$. The latter inequality coupled with the first one in (45) and the equality $z_n = 0$ on $\mathbb{T} \setminus E(\sigma)$ gives $h = H = 1_E$, 1_E being the indicator of the set $E(\sigma)$. Now

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \int_{E(\sigma)} (g_k - 1)^2\, dm = \lim_{n\to\infty} \int_{E(\sigma)} \frac{1}{n+1} \sum_{k=0}^{n} g_k^2\, dm$$

$$-2 \lim_{n\to\infty} \int_{E(\sigma)} \frac{1}{n+1} \sum_{k=0}^{n} g_k\, dm + m(E(\sigma)) = m(E) - 2m(E) + m(E) = 0.$$

The proof of (40) goes along the same line of reasoning. Again, by Cauchy's inequality

$$\frac{1}{m(I)} \int_I \sqrt{\sigma'}\, dm = \frac{1}{m(I)} \int_I \frac{\sqrt{2}\,\xi_n \sqrt{\sigma'}}{(1+\xi_n^2\sigma')^{1/2}} \frac{(1+\xi_n^2\sigma')^{1/2}}{\sqrt{2}\,\xi_n}\, dm$$

$$\leqslant \left(\frac{1}{m(I)} \int_I \frac{2\,\xi_n^2\sigma'}{1+\xi_n^2\sigma'}\, dm \right)^{1/2} \left(\frac{1}{2m(I)} \int_I \left\{ \frac{1}{\xi_n^2} + \sigma' \right\} dm \right)^{1/2}$$

$$\leqslant \left(\frac{1}{m(I)} \int_I \left\{ \frac{2\xi_n^2\sigma'}{1+\xi_n^2\sigma'} \right\}^2 dm \right)^{1/4} \left(\frac{1}{2m(I)} \int_I \left\{ \frac{1}{\xi_n^2} + \sigma' \right\} dm \right)^{1/2}.$$

Let now $n \to \infty$ along subsequences Λ_1, Λ_2 such that

$$*-\lim_{n\in\Lambda_1} \frac{2\xi_n^2\sigma'}{1+\xi_n^2\sigma'} = y, \qquad *-\lim_{n\in\Lambda_2} \left(\frac{2\xi_n^2\sigma'}{1+\xi_n^2\sigma'} \right)^2 = Y$$

in $L^\infty(\mathbb{T})$. By Proposition 3.3 we conclude that

$$\frac{1}{m(I)} \int_I \sqrt{\sigma'}\, dm \leqslant \left(\frac{1}{m(I)} \int_I y\, dm \right)^{1/2} \left(\frac{\sigma(I)}{m(I)} \right)^{1/2},$$

$$\frac{1}{m(I)} \int_I \sqrt{\sigma'}\, dm \leqslant \left(\frac{1}{m(I)} \int_I Y\, dm \right)^{1/4} \left(\frac{\sigma(I)}{m(I)} \right)^{1/2}.$$

The rest is the same as above. ∎

As a direct consequence of (40) we have $\xi_n^2 \Rightarrow 1/\sigma'$ on $E(\sigma)$, where $\Rightarrow$ stands for the convergence in measure.[3]

By using Theorem 3.4 we can give another description of the Cesàro–Nevai class.

THEOREM 3.5. *Let σ be a nonsingular measure in $\mathscr{P}$ with the Schur functions (direct and inverse) $\{f_n\}_{n\geqslant 0}$, $\{b_n\}_{n\geqslant 0}$. The following statements are equivalent.*

(i) *σ is in the Cesàro–Nevai class.*

(ii) *σ is in the Cesàro–Rakhmanov class and*

$$\frac{1}{n+1}\sum_{k=0}^{n} f_k(z)\, b_k(z)=0 \tag{46}$$

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

(iii) *There exists a subset $\Lambda\subset\mathbb{Z}_+$ with $d(\Lambda)=1$ such that*

$$\lim_{n\in\Lambda}\int_{E(\sigma)}|f_n|^2\,dm=0,\qquad E(\sigma)=\{\sigma'>0\}.$$

Proof. (i) $\Rightarrow$ (ii) is proved in (ii), Theorem 3.1.

(ii) $\Rightarrow$ (iii). We begin with the following identity (cf. [19, formula (6.4)])

$$\begin{aligned}|f_k|^2&=\frac{1-|\varphi_k|^2\sigma'}{1+|\varphi_k|^2\sigma'}+\Re(\zeta b_k f_k)+\frac{|\varphi_k|^2\sigma'-1}{1+|\varphi_k|^2\sigma'}\,\Re(\zeta b_k f_k)\\&=1-g_k+\Re(\zeta b_k f_k)+(g_k-1)\;\Re(\zeta\, \mathrm{b}_k\, \mathrm{f}_k),\end{aligned}$$

where $\zeta\in\mathbb{T}$ and g_k are defined in (34). Taking here the averages over k we obtain

$$\begin{aligned}\frac{1}{n+1}\sum_{k=0}^{n}|f_k|^2&=\frac{1}{n+1}\sum_{k=0}^{n}(1-g_k)+\Re\left\{\frac{1}{n+1}\sum_{k=0}^{n}\zeta b_k f_k\right\}\\&\quad+\frac{1}{n+1}\sum_{k=0}^{n}(g_k-1)\;\Re(\zeta b_k f_k)\\&\leqslant\frac{2}{n+1}\sum_{k=0}^{n}|g_k-1|+\Re\left\{\frac{1}{n+1}\sum_{k=0}^{n}\zeta b_k f_k\right\}.\end{aligned} \tag{47}$$

[3] Geronimus' conjecture claims that $\xi_n^2\to 1/\sigma'$ m-a.e. on $\mathbb{T}$ as long as $E(\sigma)=\mathbb{T}$.

Since $\frac{1}{n+1}\sum_{k=0}^{n} f_k b_k \in \mathscr{B}$, (46) implies the convergence of this sequence to 0 in the weak-$*$ topology of $L^\infty(\mathbb{T})$, so that we may integrate (47) over $E(\sigma)$ and make $n \to \infty$ to get

$$\begin{aligned}\frac{1}{n+1}\sum_{k=0}^{n}\int_{E(\sigma)}|f_k|^2\,dm &\leqslant \frac{2}{n+1}\sum_{k=0}^{n}\int_{E(\sigma)}|g_k-1|\,dm+o(1)\\ &\leqslant \frac{2\,|E(\sigma)|^{1/2}}{n+1}\sum_{k=0}^{n}\left(\int_{E(\sigma)}(g_k-1)^2\,dm\right)^{1/2}+o(1)\\ &\leqslant 2\,|E(\sigma)|^{1/2}\left\{\frac{1}{n+1}\sum_{k=0}^{n}\int_{E(\sigma)}(g_k-1)^2\,dm\right\}^{1/2}+o(1).\end{aligned}$$

The result is now immediate from (39) and Proposition 2.2.

(iii) $\Rightarrow$ (i). Since σ is assumed to be nonsingular, we have $m(E(\sigma))>0$ and by Khinchin–A. Ostrowski theorem $\lim_{n\in\Lambda} f_n = 0$ in $\mathscr{B}$. It remains to note that $a_n = f_n(0)$ and apply Proposition 2.2. ∎

It is worth comparing the equivalence (i) $\Leftrightarrow$ (iii) with [19, Corollary 8.3], where the similar result for Nevai's class is presented.

Theorem 3.5 can be used to obtain the following result on the convergence of Geronimus continued fractions. We adopt here the argument from [19, Sect. 8].

THEOREM 3.6. *Let $\sigma \in \mathscr{P}$ with the Schur function $f = \mathscr{H}(\sigma)$ and the Wall polynomials $\{A_n\}_{n\geqslant 0}$, $\{B_n\}_{n\geqslant 0}$. Then*

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\int_{\mathbb{T}}\left|f-\frac{A_k}{B_k}\right|^2 dm = 0 \tag{48}$$

if and only if σ is either singular or belongs to the Cesàro–Nevai class.

Proof. Suppose first that σ is nonsingular and (48) holds, so that

$$\lim_{n\in\Lambda}\int_{E(\sigma)}\left|f-\frac{A_n}{B_n}\right|^2 dm = 0 \tag{49}$$

for some Λ with $d(\Lambda)=1$. It is immediate from (49) that $A_n/B_n \Rightarrow f$, $A_{n+1}/B_{n+1} \Rightarrow f$ in measure on $E(\sigma)$ for $n \in \Lambda_1 = \Lambda \cap (\Lambda - 1)$, $d(\Lambda_1)=1$. As in [19, formula (8.11)] we have

$$\int_{E(\sigma)}\left|\frac{A_{n+1}}{B_{n+1}}-\frac{A_n}{B_n}\right| dm = \frac{|a_{n+1}|}{\sqrt{1-|a_{n+1}|^2}}\int_{E(\sigma)}\left(1-\left|\frac{A_{n+1}}{B_{n+1}}\right|^2\right)^{1/2}\left(1-\left|\frac{A_n}{B_n}\right|^2\right)^{1/2} dm,$$

which coupled with (49) and $m(E(\sigma)) > 0$ implies

$$\lim_{n \in \Lambda_1} \frac{|a_{n+1}|}{\sqrt{1-|a_{n+1}|^2}} = 0$$

and therefore a_n is almost convergent to 0, that is $\sigma \in \mathrm{CN}$.

Conversely, let σ be a nonsingular measure from the CN class. In view of the relation between the Schur functions f_n and the Wall approximants A_n/B_n (see [19, formula (8.8)]) we see that

$$\left| f - \frac{A_n}{B_n} \right|^2 = |f_{n+1}|^2 \left| 1 - \frac{\overline{A_n}}{\overline{B_n}} f \right|^2 \leqslant 4\, |f_{n+1}|^2,$$

which implies (49) for some Λ with $d(\Lambda) = 1$ by (iii) of Theorem 3.5 . By [19, Lemma 8.1] the (49) holds for integrals taken over the whole unit circle, and (48) follows.

For singular measures the Wall approximants are known to converge to the Schur function f in $L^2(\mathbb{T})$ (see [19, Lemma 8.1]), which is even stronger than (48). ∎

We complete the section with the characterization of the Cesàro–Nevai class in terms of the weak-* convergence of certain complex Borel measures. The idea here goes to P. Nevai (private communication), who proved the following[4]

THEOREM. *A measure $\sigma \in \mathcal{P}$ belongs to Nevai's class, i.e.,* $\lim_{n \to \infty} a_n(\sigma) = 0$ *if and only if*

$$*-\lim_{n \to \infty} \zeta^l \varphi_n \overline{\varphi_{n+l}}\, d\sigma = dm \qquad \forall l = 0, 1, \ldots.$$

The Cesàro version of this result is readily apparent.

THEOREM 3.7. *A measure $\sigma \in \mathcal{P}$ belongs to the Cesàro–Nevai class if and only if*

$$*-\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} \zeta^l \varphi_k \overline{\varphi_{k+l}}\, d\sigma = dm \qquad \forall l = 0, 1, \ldots. \tag{50}$$

Proof. Pick any measure σ from the CN class. Then it obviously belongs to the CR class. Note that (50) with $l = 0$ is precisely the Cesàro–Rakhmanov condition, and hence (50) holds with $l = 0$. We are going to reduce the general case $l \in \mathbb{Z}_+$ to this special one.

[4] We thank Paul Nevai for his permission to include this result here.

Let $\varphi_n = \kappa_n z^n + \ldots$, $\kappa_n > 0$ be orthonormal polynomials in $L^2(\sigma)$. We start out with the following known relation which stems directly from the orthogonality conditions and the formula for the leading coefficients (see [9, formula (8.6)]):

$$\int_{\mathbb{T}} \zeta^l \varphi_k(\zeta)\, \overline{\varphi_{k+l}(\zeta)}\, d\sigma = \frac{\kappa_k}{\kappa_{k+l}} = \prod_{j=1}^{l} (1 - |a_{k+j}|^2)^{1/2}. \tag{51}$$

Let $l \geqslant 1$ and $f \in C(\mathbb{T})$. Then by Cauchy's inequality

$$\left| \int_{\mathbb{T}} f \zeta^l \varphi_k \overline{\varphi_{k+l}}\, d\sigma - \int_{\mathbb{T}} f\, |\varphi_k|^2\, d\sigma \right|$$
$$\leqslant \left(\int_{\mathbb{T}} |f \varphi_k|^2\, d\sigma \right)^{1/2} \left(\int_{\mathbb{T}} |\zeta^l \varphi_k - \varphi_{k+l}|^2\, d\sigma \right)^{1/2}.$$

The first factor on the right-hand side does not exceed $\|f\|_\infty$. As for the second one, it can be computed explicitly in view of (51)

$$\int_{\mathbb{T}} |\zeta^l \varphi_k - \varphi_{k+l}|^2\, d\sigma = 2\left(1 - \frac{\kappa_k}{\kappa_{k+l}}\right) = 2\left(1 - \prod_{j=1}^{l} (1 - |a_{k+j}|^2)^{1/2}\right).$$

By Corollary 2.4, $\sigma \in \mathrm{CN}$ implies $\lim_{k \in \Lambda_1} a_{k+j} = 0$ for some subset $\Lambda_1 \subset \mathbb{Z}_+$ of density 1 and for all integers j. Therefore

$$\lim_{k \in \Lambda_1} \left| \int_{\mathbb{T}} f \zeta^l \varphi_k \overline{\varphi_{k+l}}\, d\sigma - \int_{\mathbb{T}} f\, |\varphi_k|^2\, d\sigma \right| = 0.$$

On the other hand, by (v), Theorem 3.1 there is a subset $\Lambda_2 \subset \mathbb{Z}_+$ of density 1 such that

$$\lim_{k \in \Lambda_2} \int_{\mathbb{T}} f\, |\varphi_k|^2\, d\sigma = \int_{\mathbb{T}} f\, dm, \qquad \forall f \in C(\mathbb{T}).$$

Hence

$$\lim_{k \in \Lambda_3} \int_{\mathbb{T}} f \zeta^l \varphi_k \overline{\varphi_{k+l}}\, d\sigma = \int_{\mathbb{T}} f\, dm$$

for the subset $\Lambda_3 = \Lambda_1 \cap \Lambda_2$ of density 1 and for all nonnegative integers l. (50) now follows from Proposition 2.2.

The converse statement is much easier. Take in (50) $l=1$ and $f=1$. By (51)

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\int_{\mathbb{T}}\zeta\varphi_k\overline{\varphi_{k+1}}\,d\sigma=\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}(1-|a_{k+1}|^2)^{1/2}=1,$$

whence it follows that a_n is strongly convergent to zero, as needed. ∎

4. UNIVERSAL MEASURES

DEFINITION. A probability measure σ on $\mathbb{T}$ is called *universal* if the sequence $\{|\varphi_n|^2\,d\sigma\}_{n\geqslant 0}$ is dense in $\mathscr{P}$.

In view of Theorem B the definition can be paraphrased in terms of the direct and inverse Schur functions $\{f_n\}_{n\geqslant 0}$ and $\{b_n\}_{n\geqslant 0}$ as follows.

A probability measure σ on $\mathbb{T}$ is universal if and only if the sequence $\{f_n b_n\}_{n\geqslant 0}$ is dense in $\mathscr{B}$.

The very existence of universal measures is far from being obvious. However, it is easy to establish some properties of such measures.

THEOREM 4.1. *Let the Dirac measure δ_λ at some point $\lambda\in\mathbb{T}$ be a limit point of the sequence $\{|\varphi_n|^2\,d\sigma\}_{n\geqslant 0}$. Then σ is singular. In particular, each universal measure is singular.*

Proof. Since the Schur function of δ_λ is constant $\bar{\lambda}$, Theorem B implies $\lim_{n\in\Delta} f_n b_n=\bar{\lambda}$ uniformly on compact subsets of $\mathbb{D}$ for some infinite subsequence $\Delta\subset\mathbb{Z}_+$. But $f_n(0)=a_n$, $b_n(0)=-\bar{a}_{n-1}$ (see (27)), so that $\lim_{n\in\Delta} a_n\bar{a}_{n-1}=-\bar{\lambda}$. Hence $\limsup_{n\to\infty}|a_n|=1$ and σ is singular by Rakhmanov's Lemma [28, Lemma 4, p. 110] (see also [19, Corollary 9.5]). ∎

It is clear that for universal measures the derived set of the sequence $\{a_n\bar{a}_{n-1}\}_{n\geqslant 1}$ contains the unit circle, for each Dirac mass δ_λ is a limit point of $\{|\varphi_n|^2\,d\sigma\}_{n\geqslant 0}$.

PROPOSITION 4.2. *Let σ be a universal measure. Then*

(i) *The set $\{a_n\bar{a}_{n-1}\}_{n\geqslant 1}$ is dense in $\mathbb{D}$. In particular,* $\liminf_{n\to\infty}|a_n|=0$.

(ii) $\operatorname{supp}\sigma=\mathbb{T}$.

Proof. (i) By the definition of universal measures and Theorem B each function $s\in\mathscr{B}$ is a limit point of the sequence $\{f_n b_n\}_{n\geqslant 0}$ and hence, as in Theorem 4.1, for each $w\in\mathbb{D}$ $\lim_{n\in\Delta} a_n\bar{a}_{n-1}=w$ for some $\Delta\subset\mathbb{Z}_+$.

(ii) If there were an open arc I with $I\cap\operatorname{supp}\sigma=\varnothing$, then for $\lambda\in I$ the Dirac measure δ_λ could not be the limit point of $\{|\varphi_n|^2\,d\sigma\}_{n\geqslant 0}$. ∎

There are obviously no universal measures in Rakhmanov's class R (and the more so, in Nevai's class N). Amazingly enough, the Cesàro–Nevai class which is presumably quite close to N, contains such measures (in fact, plenty of them).

THEOREM 4.3. *There are universal measures in the Cesàro–Nevai class.*

Proof. We proceed in two steps.

Step 1. Here is the main ingredient of the whole construction.
Let $P, Q \in \mathscr{B}$ with the infinite number of Schur parameters

$$\mathscr{S}P = (p_0, p_1, \ldots), \qquad \mathscr{S}Q = (q_0, q_1, \ldots)$$

that is, P, Q are not finite Blaschke products, and put $H = PQ$. Take $\Gamma \subset \mathbb{Z}_+$ to be a union of disjoint intervals of nonnegative integers

$$\Gamma = \bigcup_{k=1}^{\infty} [n_k - m_k, n_k + m_k]$$

with $m_k \to \infty$ as $k \to \infty$. Define a function

$$f \in \mathscr{B}, \qquad \mathscr{S}f = \{a_0, a_1, \ldots\}$$

as

$$a_j \stackrel{\text{def}}{=} p_{j-n_k}, \quad n_k \leqslant j \leqslant n_k + m_k; \qquad a_j \stackrel{\text{def}}{=} -\bar{q}_{n_k - j - 1}, \quad n_k - m_k \leqslant j \leqslant n_k - 1$$

and arbitrarily on the complement Γ^c. Let f_n and b_n be the Schur functions (direct and inverse) of f. Then for $n = n_k$

$$\mathscr{S}f_n = (p_0, p_1, \ldots, p_{m_k}, \ldots), \qquad \mathscr{S}b_n = (q_0, q_1, \ldots, q_{m_k - 1}, \ldots).$$

By Theorem A

$$\lim_{n \in \Delta} f_n(z) = P(z), \qquad \lim_{n \in \Delta} b_n(z) = Q(z), \qquad \Delta \stackrel{\text{def}}{=} \{n_1 < n_2 < \cdots\}$$

in $\mathscr{B}$, and hence

$$\lim_{n \in \Delta} f_n b_n = H. \tag{52}$$

Let $\sigma, \nu \in \mathscr{P}$ be measures which correspond to f and H, respectively, and let $\{\varphi_n\}_{n \geqslant 0}$ be the orthonormal system in $L^2(d\sigma)$. By Theorem B (52) means that $\nu = \lim_{n \in \Delta} |\varphi_n|^2 \, d\sigma$.

Step 2. Given positive integers r and s denote by Λ_{rs} the set of integers

$$\Lambda_{rs} \stackrel{\text{def}}{=} [2^{2^r(2s-1)}, 2^{2^r(2s-1)}+2s] = [n_{rs}-m_{rs}, n_{rs}+m_{rs}].$$

Let us first check that the sets Λ_{rs} are pairwise disjoint. Indeed, suppose that there is an integer k with

$$2^{2^r(2s-1)} < 2^k \leqslant 2^{2^r(2s-1)}+2s$$

or

$$2^r(2s-1) < k < 2^r(2s-1)+\log_2\left(1+\frac{2s}{2^r(2s-1)}\right).$$

But

$$\log_2\left(1+\frac{2s}{2^r(2s-1)}\right) < \frac{2s}{2^r\log 2(2s-1)} < \frac{1}{2^{r-1}\log 2} < 1$$

for $r \geqslant 2$. The same is true for $r=1$ and $s \geqslant 2$ as well. Finally, the segment $\Lambda_{11}=[4, 6]$ contains no points 2^k with $k>2$, as claimed.

It is clear that Card $\Lambda_{rs}=2s+1$. We show that

$$d(\Lambda)=0, \qquad \Lambda \stackrel{\text{def}}{=} \bigcup_{r,s=1}^{\infty} \Lambda_{rs}.$$

Fix a big positive number N. The number of indices k which satisfy $2^k<N$ does not exceed $\log_2 N$. If $k=2^r(2s-1)$ then $s \leqslant k < \log_2 N$. It follows that

$$\operatorname{Card}(\Lambda \cap [0, N]) \leqslant (2\log_2 N+1)\log_2 N,$$

which yields $d(\Lambda)=0$.

Next, take a sequence of polynomials $\{H_l\}_{l \geqslant 0}$ dense in $\mathscr{B}$. The corresponding sequence of measures $\{\nu_l\}_{l\geqslant 0}$ from Step 1 is clearly dense in $\mathscr{P}$. Each H_l can be factored as $H_l=P_lQ_l$, where none of the functions P_l, Q_l is a finite Blaschke product. Put

$$\Lambda_l \stackrel{\text{def}}{=} \bigcup_{s=1}^{\infty} \Lambda_{ls} = \bigcup_{s=1}^{\infty} [n_{ls}-m_{ls}, n_{ls}+m_{ls}], \qquad l=1, 2, \ldots$$

and apply the procedure from Step 1 to each collection Λ_l, P_l, Q_l, H_l with $l=1, 2, \ldots$. Thereby the Schur parameters a_n of the function $f \in \mathscr{B}$ (these are the same as the Geronimus parameters of the measure σ) are already determined for $n \in \Lambda$. For the rest of indices we define a_n to meet $\lim_{n \in \Lambda^c} a_n=0$. According to the construction in Step 1, for each $l \geqslant 1$ there is a subset $\Delta_l \subset \mathbb{Z}_+$ such that

$$\lim_{n\in\Delta_l} f_n(z)=P_l(z), \qquad \lim_{n\in\Delta_l} b_n(z)=Q_l(z), \qquad \lim_{n\in\Delta_l} f_n(z)\,b_n(z)=H_l(z)$$

in $\mathscr{B}$ (cf. (52)). By Theorem B the latter means that the each measure ν_l is the limit point of the measures $|\varphi_n|^2\, d\sigma$, and thus σ is the universal measure. Finally, $\lim_{n \in \Lambda^c} a_n = 0$ for the subset $\Lambda^c \subset \mathbb{Z}_+$ of density 1, that is $\sigma \in \mathrm{CN}$. The proof is complete. ∎

Remark. There is a large arbitrariness in the definition of Geronimus parameters on Λ^c. We can take advantage of such freedom to have universal measures with some added properties. For instance, we can find universal measures in CN for which $\{a_n\}_{n \geqslant 0}$ is dense in $\mathbb{D}$.[5] Indeed, write

$$\Lambda^c = I_0 \cup \bigcup_{k=1}^{\infty} I_k$$

with pairwise disjoint and nonempty subsets I_k and $d(I_0) = 1$. Pick any dense set $\{w_k\}$ in $\mathbb{D}$ and define $a_n = 0$, $n \in I_0$,

$$a_n = w_s, \qquad n \in \bigcup_{r=0}^{\infty} I_{2^r(2s-1)}, \qquad s = 1, 2, \ldots.$$

Next, if we put $a_n = a$, $a \neq 0$, for $n \in \Lambda^c$, we arrive at the universal measure σ in the *Cesàro–Geronimus* class:

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} |a_k(\sigma) - a| = 0.$$

For the further discussion of this class as well as the related class of Jacobi matrices see [15]. Note that if $\lim_{n \to \infty} |a_n - a| = 0$ for some $a \neq 0$, then by [6, Theorem 1′] the support $\operatorname{supp} \sigma$ of the corresponding measure is a proper subset of $\mathbb{T}$, that is, the measure cannot be universal (see (ii), Proposition 4.2).

We show next that the class of universal measures is stable under certain perturbations of the Geronimus parameters.

THEOREM 4.4. *Let σ be a universal measure with the Geronimus parameters $\{a_n\}_{n \geqslant 0}$. Let $\{c_n\}_{n \geqslant 0}$ satisfy $|a_n + c_n| < 1$ and $c_n \to 0$ as $n \to \infty$. Then the measure μ with the Geronimus parameters $a_n + c_n$ is also universal.*

Proof. Let $f \in \mathscr{B}$ correspond to σ with the Schur functions f_n and b_n. By the definition of universal measures and Theorem B, given $H \in \mathscr{B}$ there is a sequence $\Lambda' = \Lambda'(H)$ such that $\lim_{n \in \Lambda'} f_n b_n = H$ in $\mathscr{B}$. The normal family argument applied to both f_n and b_n shows that in fact

$$\lim_{n \in \Lambda} f_n(z) = P(z), \qquad \lim_{n \in \Lambda} b_n(z) = Q(z)$$

[5] Note that this property does not necessarily follow from (i), Proposition 4.2.

in $\mathscr{B}$ for some subsequence $\Lambda \subset \Lambda'$. By Theorem A for each fixed $k = 0, 1, \ldots$

$$\lim_{n \in \Lambda} a_{n+k} = p_k, \qquad \lim_{n \in \Lambda} -\bar{a}_{n-k-1} = q_k.$$

Since $c_n \to 0$, we have

$$\lim_{n \in \Lambda} (a_{n+k} + c_{n+k}) = p_k, \qquad \lim_{n \in \Lambda} (-\bar{a}_{n-k-1} - \bar{c}_{n-k-1}) = q_k.$$

Let $g \in \mathscr{B}$ correspond to μ with the Schur functions g_n and d_n. Again, application of Theorem A yields

$$\lim_{n \in \Lambda} g_n(z) = P(z), \qquad \lim_{n \in \Lambda} d_n(z) = Q(z)$$

so that $\lim_{n \in \Lambda} g_n d_n = PQ = H$ in $\mathscr{B}$, as needed. ∎

We prove now that the class of universal measures is closed with respect to certain operations on Schur parameters.

Proposition 4.5. *Let σ be a universal measure with the Geronimus parameters $\{a_n\}_{n \geqslant 0}$, and let $\lambda \in \mathbb{T}$. Consider the family of measures $\{\sigma_\lambda\}$ with the Geronimus parameters $\{\lambda a_n\}_{n \geqslant 0}$. Then $\{\sigma_\lambda\}$ is universal for all λ.*

Proof. Let f_n, b_n and f_n^λ, b_n^λ be the direct and inverse Schur functions of σ and σ_λ, respectively. Then $f_n^\lambda = \lambda f_n$, $n \in \mathbb{Z}_+$. For the Schur parameters of b_n^λ we have

$$\mathscr{S} b_n^\lambda = (-\overline{\lambda a}_{n-1}, -\overline{\lambda a}_{n-2}, \ldots, -\overline{\lambda a}_0, 1),$$

and hence

$$b_n^\lambda(z) = \bar{\lambda} b_n(z) + O(z^n), \qquad f_n^\lambda(z)\, b_n^\lambda(z) = f_n(z)\, b_n(z) + O(z^n), \qquad z \to 0.$$

Therefore the sequences $\{f_n^\lambda b_n^\lambda\}_{n \geqslant 0}$ and $\{f_n b_n\}_{n \geqslant 0}$ have the same limit points in $\mathscr{B}$. This completes the proof. ∎

There are two shift operators acting on the set of all infinite sequences of the Geronimus (Schur) parameters:

$$S_l(a_0, a_1, \ldots) = (a_1, a_2, a_3, \ldots),$$
$$S_r(a_0, a_1, \ldots) = (\alpha, a_0, a_1, \ldots), \qquad |\alpha| < 1.$$

Denote by $\sigma_l, \varphi_{l,n}, \psi_{l,n}$ and $\sigma_r, \varphi_{r,n}, \psi_{r,n}$ the transformed measures and orthonormal polynomials of the first and second kinds related to S_l and S_r, respectively.[6] The explicit relations, which follow from the Szegő recurrence relations, are known for these polynomial systems

$$
\begin{aligned}
2z\varphi_{l,n}(z) &= \varphi_{n+1}(z)(\psi_1(z)+\psi_1^*(z))+\psi_{n+1}(z)(\varphi_1^*(z)-\varphi_1(z)),\\
2z\varphi_{l,n}^*(z) &= \varphi_{n+1}^*(z)(\psi_1(z)+\psi_1^*(z))+\psi_{n+1}^*(z)(\varphi_1(z)-\varphi_1^*(z)),
\end{aligned}
\tag{53}
$$

and

$$
\begin{aligned}
2\varphi_{r,n+1}(z) &= (\varphi_n(z)+\psi_n(z))\,\varphi_{1,r}(z)+(\varphi_n(z)-\psi_n(z))\,\varphi_{1,r}^*(z),\\
2\varphi_{r,n+1}^*(z) &= (\varphi_n^*(z)-\psi_n^*(z))\,\varphi_{1,r}(z)+(\varphi_n^*(z)+\psi_n^*(z))\,\varphi_{1,r}^*(z),
\end{aligned}
\tag{54}
$$

where $\varphi_n^*(z)=z^n\overline{\varphi_n(1/\bar{z})}$ (see [25] regarding (53)).

Proposition 4.6. *If σ is a universal measure, then σ_l and σ_r are also universal.*

Proof. Let $f_n(f_{l,n}, f_{r,n})$ and $b_n(b_{l,n}, b_{r,n})$ be the direct and inverse Schur functions for $\sigma(\sigma_l, \sigma_r)$, respectively. We begin with the left shift operator and use (53) to compute $b_{l,n}$:

$$
b_{l,n}(z)=\frac{\varphi_{l,n}(z)}{\varphi_{l,n}^*(z)}=\frac{z(1+a_0)(\varphi_{n+1}(z)-\psi_{n+1}(z))+(1+\bar{a}_0)(\varphi_{n+1}(z)+\psi_{n+1}(z))}{z(1+a_0)(\varphi_{n+1}^*(z)+\psi_{n+1}^*(z))+(1+\bar{a}_0)(\varphi_{n+1}^*(z)-\psi_{n+1}^*(z))}.
$$

In terms of the Wall polynomials we have (see [19, formulae (5.5)])

$$
\begin{aligned}
b_{l,n}(z) &= \frac{(1+\bar{a}_0)\,B_n^*(z)-(1+a_0)\,A_n^*(z)}{(1+a_0)\,B_n(z)-(1+\bar{a}_0)\,A_n(z)},\\
b_{n+1}(z) &= \frac{\varphi_{n+1}(z)}{\varphi_{n+1}^*(z)}=\frac{zB_n^*(z)-A_n^*(z)}{B_n(z)-zA_n(z)}.
\end{aligned}
$$

Next, it is clear that $f_{l,n}=f_{n+1}$, $n\geqslant 0$, so that it seems natural to take the difference

$$
\begin{aligned}
&f_{l,n}(z)\,b_{l,n}(z)-f_{n+1}(z)\,b_{n+1}(z)\\
&\quad = f_{n+1}(z)(b_{l,n}(z)-b_{n+1}(z))\\
&\quad = \frac{f_{n+1}(z)((1+\bar{a}_0)-z(1+a_0))}{((1+a_0)\,B_n(z)-(1+\bar{a}_0)\,A_n(z))(B_n(z)-zA_n(z))}\\
&\qquad \times B_n^*B_n-A_n^*A_n=O(z^n)
\end{aligned}
$$

[6] The polynomials $\varphi_{l,n}$, $\psi_{l,n}$ are known as the *associated polynomials*.

in $\mathcal{B}$. Therefore the sequences $\{f_{l,n}b_{l,n}\}_{n\geqslant 0}$ and $\{f_n b_n\}_{n\geqslant 0}$ have the same limit points. This completes the proof for the operator S_l. The arguments for the right shift operator S_r are quite the same. We use (54) to obtain

$$f_{r,n+1}(z)\, b_{r,n+1}(z) - f_n(z)\, b_n(z) = O(z^n)$$

in $\mathcal{B}$. ∎

We see that the property "being universal" does not depend on a finite number of the Geronimus parameters:

Corollary 4.7. *Let the Geronimus parameters of two measures σ, $\mu \in \mathcal{P}$ agree from some point on. Then the both are universal simultaneously.*

Since every universal measure is singular, its Schur function is an inner function. We show now that there are universal measures whose Schur functions are Blaschke products (cf. [20, Corollary 1]).

Let $\tau_\alpha(z) = (z+\alpha)(1+\bar{\alpha}z)^{-1}$ be a Möbius transform, $\alpha \in \mathbb{D}$. Given $f \in \mathcal{B}$ consider the composition $f(\alpha) \stackrel{\text{def}}{=} \tau_\alpha \circ f$. Since $f(\alpha) \in \mathcal{B}$, there is a unique $\sigma(\alpha)$ in $\mathcal{P}$ such that $f(\alpha)$ is the Schur function of $\sigma(\alpha)$.

Proposition 4.8. *Let $\sigma \in \mathcal{P}$ be a universal measure with the Schur function f. Then for all $\alpha \in \mathbb{D}$ except for a set of logarithmic capacity zero, $\sigma(\alpha) \in \mathcal{P}$ are universal measures and their Schur functions $f(\alpha)$ are infinite Blaschke products.*

Proof. Let $\mathcal{S}f = (a_0, a_1, ...)$. The idea is to compute explicitly the Schur parameters of $f(\alpha)$. It is a matter of a routine (but rather lengthy) calculation to verify that

$$f(\alpha,z) = \lambda_\alpha \tau_\xi(zf_1(z)) = \lambda_\alpha \frac{zf_1(z)+\xi}{1+\bar{\xi}zf_1(z)},$$

where

$$\xi = \tau_{a_0}(\alpha) = \frac{\alpha + a_0}{1+\bar{a}_0\alpha}, \qquad \lambda_\alpha = \frac{1+\bar{a}_0\alpha}{1+a_0\bar{\alpha}}, \quad |\lambda_\alpha| = 1,$$

and f_1 is the first Schur function of f, $\mathcal{S}f_1 = (a_1, a_2, ...)$. It is not hard to see that

$$\mathcal{S}f(\alpha) = \lambda_\alpha \xi, \lambda_\alpha a_1, \lambda_\alpha a_2, ...$$

In view of Corollary 4.7 and Proposition 4.5, $\sigma(\alpha)$ are universal for all $\alpha \in \mathbb{D}$. Now, an application of Frostman's theorem (cf. [5, Theorem 6.4]) completes the proof. ∎

5. THE CLASSES WHICH ARE "OPPOSITE" TO G_∞ AND SZEGŐ CLASS

We begin with the Cesàro version of MNT's theorem [22, Theorem 4]. Let

$$I_n(z)=\frac{1}{|\varphi_n(z)|^2}\sum_{k=0}^{n}|\varphi_k(z)|^2.$$

THEOREM 5.1. *For each $\sigma\in CN$ there exists a subset $\Lambda\subset\mathbb{Z}_+$ of density* 1 *such that*

$$\lim_{n\in\Lambda}\max_{|z|\leqslant 1}\frac{1}{I_n(z)}=0. \tag{55}$$

Conversely, if $\lim_{n\in\Lambda}1/I_n(z_0)=0$ for some $z_0\in\mathbb{D}$ and some subset $\Lambda\subset\mathbb{Z}_+$ of density 1, then $\sigma\in$ CN.

Proof. Consider the Szegő kernel

$$K_{n+1}(z,\zeta)=\sum_{k=0}^{n}\varphi_k(z)\,\overline{\varphi_k(\zeta)}=\frac{\varphi_n^*(z)\,\overline{\varphi_n^*(\zeta)}-z\bar{\zeta}\varphi_n(z)\,\overline{\varphi_n(\zeta)}}{1-z\bar{\zeta}},$$

where $\varphi_n^*(z)=z^n\overline{\varphi_n(1/\bar{z})}$. For $|z|<1$ we have

$$\frac{1}{I_n(z)}=\frac{|\varphi_n(z)|^2}{K_{n+1}(z,z)}=\frac{|\varphi_n(z)|^2\,(1-|z|^2)}{|\varphi_n^*(z)|^2-|z\varphi_n(z)|^2}=\frac{|b_n(z)|^2\,(1-|z|^2)}{1-|zb_n(z)|^2}. \tag{56}$$

We see that $\lim_{n\in\Lambda}1/I_n(z_0)=0$ at $z_0\in\mathbb{D}$ if and only if $\lim_{n\in\Lambda}b_n(z_0)=0$. Put $\Lambda_1=\Lambda\cap(\Lambda-1)$, which is also of density 1. In the identity

$$b_{n+1}(z_0)=\frac{z_0b_n(z_0)-\bar{a}_n}{1-a_nz_0b_n(z_0)}$$

let $n\to\infty$ with $n\in\Lambda_1$. We conclude that $\lim_{n\in\Lambda}a_n=0$, i.e., $\{a_n\}_{n\geqslant 0}$ is almost convergent to zero. Hence $\sigma\in$ CN by Proposition 2.2.

Conversely, let $\sigma\in$ CN. We invoke the inequality (see [22, p. 58]))

$$\max_{|z|\leqslant 1}\frac{|\varphi_n(z)|^2}{K_{n+1}(z,z)}\leqslant\frac{2}{M}+2\left(\sum_{j=n-M}^{n}|a_{j-1}|\right)^2,\qquad n\geqslant M+1, \tag{57}$$

which holds for an arbitrary measure and an arbitrary positive integer M. By Corollary 2.4 there is a subset $\Lambda \subset \mathbb{Z}_+$ of density 1, such that $\lim_{n \in \Lambda} a_{n-k} = 0$ for every integer k. Hence by (57)

$$\limsup_{n \in \Lambda} \max_{|z| \leqslant 1} \frac{|\varphi_n(z)|^2}{K_{n+1}(z, z)} \leqslant \frac{2}{M}$$

for every positive M, and the result follows immediately upon letting $M \to \infty$. ▮

It is well known [10, formula (20.16)] that $\sum_{n=0}^{\infty} |\varphi_n(z_0)|^2 < \infty$ at each mass point $z_0 \in \mathbb{T}$ of a measure $\sigma \in \mathscr{P}$. Hence $\lim_{n \to \infty} 1/I_n(z_0) = 0$ at each such point, so that the converse statement in Theorem 5.1 is false for $z_0 \in \mathbb{T}$.

Theorem 5.1 says that in the Cesàro–Nevai class

$$\limsup_{n \to \infty} I_n(\zeta) = +\infty$$

uniformly on $\mathbb{T}$. In other words, CN is contained in the class G_∞ of measures $\sigma \in \mathscr{P}$ with

$$C_\sigma(z) = \sup_n I_n(\zeta) = +\infty \qquad \forall \zeta \in \mathbb{T}. \tag{58}$$

The formal opposite of condition (58) would look like

$$C_\sigma(z) < \infty \qquad \forall \zeta \in \mathbb{T}. \tag{59}$$

It turns out that there are no measures in $\mathscr{P}$ which satisfy (59) *everywhere* on $\mathbb{T}$. To show this, we need the following elementary result.

LEMMA 5.2. *Let $\{d_n\}_{n \geqslant 0}$ be a sequence of positive numbers such that $d_0 = 1$ and*

$$d_0 + d_1 + \cdots + d_n < D d_n, \qquad n = 0, 1, 2, \ldots \tag{60}$$

for some $D > 0$. Then

$$\left(\frac{D}{D-1}\right)^{n-1} < (D-1)\, d_n, \qquad n = 1, 2, \ldots . \tag{61}$$

Proof. It is clear from (60) with $n = 0$ that $D > 1$. Note also that (60) is equivalent to

$$d_0 + d_1 + \cdots + d_n < (D-1)\, d_{n+1}, \qquad n = 0, 1, 2, \ldots . \tag{62}$$

We prove (61) by induction. For $n = 1$ (61) turns into (62). Suppose that (61) holds for every $n = 1, 2, \ldots, k$. Then, by the induction hypothesis

$$(D-1)\left(\left(\frac{D}{D-1}\right)^k - 1\right) = 1 + \frac{D}{D-1} + \left(\frac{D}{D-1}\right)^2 + \cdots + \left(\frac{D}{D-1}\right)^{k-1}$$
$$< (D-1)\,(d_1 + \cdots + d_k),$$

which gives (61) with $n = k+1$ by (62). ∎

Proposition 5.3. *For each $\sigma \in \mathscr{P}$ there is a point $\zeta = \zeta_\sigma \in \mathbb{T}$ with $C_\sigma(\zeta) = +\infty$.*

Proof. We observe that C_σ is lower semi-continuous and hence is Borel measurable. Suppose, on the contrary, that there is a measure $\mu \in \mathscr{P}$ such that C_μ is finite everywhere on $\mathbb{T}$. Then, for $k = 2, 3, \ldots,$ the sets $F_k = \{\zeta \in \mathbb{T} : C_\mu(\zeta) \leqslant k\}$ form an increasing sequence of closed sets and

$$\mathbb{T} = \bigcup_{k=2}^{\infty} F_k.$$

An integer k clearly exists with $\mu(F_k) > 0$. We have

$$\sum_{j=0}^{n} |\varphi_j(\zeta)|^2 < (k+1)\,|\varphi_n(\zeta)|^2$$

for all $n \geqslant 0$ and all $\zeta \in F_k$. By Lemma 5.2 we see that

$$\frac{1}{k}\left(\frac{k+1}{k}\right)^{n-1} < |\varphi_n(\zeta)|^2, \qquad n = 1, 2, \ldots, \quad \zeta \in F_k. \tag{63}$$

Upon integrating (63) with respect to μ over F_k we arrive at the relation

$$\frac{1}{k}\left(\frac{k+1}{k}\right)^{n-1} \mu(F_k) < \int_{F_k} |\varphi_n|^2\, d\mu \leqslant 1,$$

which leads to contradiction as $n \to \infty$. ∎

Definition. A measure $\sigma \in \mathscr{P}$ is said to belong to G_0 (the opposite of the G_∞ class) if

$$C_\sigma(\zeta) = \sup_n I_n(\zeta) < \infty$$

m-a.e. on $\mathbb{T}$.

DEFINITION. A measure $\sigma \in \mathscr{P}$ is said to belong to OS (the opposite Szegő class) if

$$\sum_{n=0}^{\infty} (1-|a_n|^2)^{1/2} < \infty, \tag{64}$$

$\{a_n\}_{n \geqslant 0}$ being the Geronimus parameters of σ.

The class OS is obviously nonempty, for it is defined in terms of independent parameters a_n. In view of Proposition 5.3 the same conclusion for the class G_0 is not quite apparent. However it is not hard to derive some properties of measures from G_0.

THEOREM 5.4. *Every measure $\sigma \in G_0$ is singular.*

Proof. We proceed in the same fashion as in the proof of Proposition 5.3. The only difference is that now

$$\mathbb{T} = F \cup \bigcup_{k=2}^{\infty} F_k, \qquad F \stackrel{\text{def}}{=} \{\zeta \in \mathbb{T} : C_\sigma(\zeta) = +\infty\}.$$

Again, (63) is true for all n, k and $\zeta \in F_k$. As above, we have

$$\sigma(F_k) < k\left(\frac{k}{k+1}\right)^{n-1} \int_{F_k} |\varphi_n|^2 \, d\mu \leqslant k\left(\frac{k}{k+1}\right)^{n-1},$$

which gives $\sigma(F_k) = 0$ for every $k = 2, 3, \ldots$, since n is an arbitrary positive integer. Hence $\mathbb{T} = G \cup F$ with $\sigma(G) = m(F) = 0$, i.e., σ is singular. ▮

Remark 5.5. The argument in the proof of Theorem 5.4 is of a quite general nature and has nothing to do with the specific case of the unit circle. As a matter of fact, the following statement concerning general orthogonal polynomials is true.

Let K be a compact set of the complex plane $\mathbb{C}$ and ν be a probability measure on K. There exists a unique system of orthonormal polynomials $p_n(z, \nu) = \gamma_n z^n + \ldots$, $\gamma_n > 0$. If

$$\sup_n \frac{1}{|p_n(z, \nu)|^2} \sum_{k=0}^{n} |p_k(z, \nu)|^2 < \infty$$

m-a.e., where m is the normalized Lebesgue measure on K, then ν is singular with respect to m.

The proof of G_0 being nonempty is much more delicate. The following result from the ergodic theory of inner functions is crucial for this.

LEMMA 5.6. *Let J be an inner function with $J(0)=0$. Then $J\colon \mathbb{T} \longrightarrow \mathbb{T}$ is a measure preserving endomorphism of the measure space* $(\mathbb{T}, dm)$.

Proof. By Fatou's theorem the boundary values of any inner function J satisfy $|J|=1$ m-a.e. on $\mathbb{T}$ and hence $J\colon \mathbb{T} \to \mathbb{T}$ is a measurable mapping of the measure space $(\mathbb{T}, dm)$.

We define a linear operator U on the space of trigonometric polynomials by $Up(\zeta)=p(J(\zeta))$. It is clear that $U\zeta^n = J^n(\zeta)$, $n \in \mathbb{Z}$, which implies

$$\int_{\mathbb{T}} U\zeta^n \overline{U\zeta^k}\, dm = \int_{\mathbb{T}} J^{n-k}(\zeta)\, dm = J^{n-k}(0) = 0$$

for $n > k$. Hence the operator U transforms the orthonormal basis $\{\zeta^n\}_{n\in\mathbb{Z}}$ into the orthonormal system $\{J^n\}_{n\in\mathbb{Z}}$ in the Hilbert space $L^2(\mathbb{T})$. It follows that U admits the extension to an isometric operator in $L^2(\mathbb{T})$ and for each $h \in L^2(\mathbb{T})$

$$\int_{\mathbb{T}} |h \circ J|^2\, dm = \int_{\mathbb{T}} |h|^2\, dm.$$

The result follows by putting $h = 1_E$ with a measurable set E. ∎

THEOREM 5.7. $OS \subset G_0$.

Proof. Our starting point is the relation

$$\left|\frac{\varphi_k(\zeta)}{\varphi_{k+1}(\zeta)}\right|^2 = \frac{1-|a_k|^2}{|1-a_k\zeta b_k(\zeta)|^2}, \tag{65}$$

which is actually a direct consequence of the Szegő recurrencies for orthonormal polynomials on the unit circle. Let I_k be an open arc on $\mathbb{T}$ centered at $\bar{a}_k/|a_k|$ of the circular length $m(I_k) = 10(1-|a_k|^2)^{1/2}$, and let

$$E_k \stackrel{\text{def}}{=} \{\zeta \in \mathbb{T} : \zeta b_k(\zeta) \in I_k\}, \qquad E_k^c = \mathbb{T} \setminus E_k.$$

Suppose that $|a_k| > 1/2$ for $k \geqslant k_0$. It is a matter of routine computation to verify that for $w \notin I_k$ and $k \geqslant k_0$

$$|1-a_k w|^2 > 4(1-|a_k|^2).$$

Hence it follows from (65) that

$$\left|\frac{\varphi_k(\zeta)}{\varphi_{k+1}(\zeta)}\right|^2 < \frac{1}{4}, \qquad \zeta \in E_k^c. \tag{66}$$

We claim that the set

$$\Omega \overset{\text{def}}{=} \bigcap_{s=1}^{\infty} \bigcup_{j=s}^{\infty} E_j$$

has m-measure zero. Indeed, $\Omega \subset \bigcup_{j=s}^{\infty} E_j$ for each $s \geqslant 1$. But according to Lemma 5.6 we have $m(E_k) = m(I_k)$ and hence

$$m(\Omega) \leqslant m\left(\bigcup_{j=s}^{\infty} E_j\right) \leqslant \sum_{j=s}^{\infty} m(E_j) = \sum_{j=s}^{\infty} m(I_j) < 10 \sum_{j=s}^{\infty} (1-|a_j|^2)^{1/2},$$

which tends to zero as $s \to \infty$, as claimed.

Thus $m(\Omega^c) = 1$ for

$$\Omega^c = \bigcup_{s=1}^{\infty} \bigcap_{j=s}^{\infty} E_j^c = \liminf_{n \to \infty} E_n^c.$$

The latter is the set of all points ζ such that $\zeta \in E_n^c$ for all but finite number of indices n. In other words, for each $\zeta \in \Omega^c$ there is a positive integer $l_0(\zeta)$ such that $\zeta \in E_l^c$, $l \geqslant l_0$. Hence (66) holds for $k \geqslant l_1 = \max(l_0, k_0)$. Put

$$M(\zeta) \overset{\text{def}}{=} \max_{0 \leqslant k \leqslant l_1} \left| \frac{\varphi_k(\zeta)}{\varphi_{k+1}(\zeta)} \right|^2 < \infty.$$

Then for $n \geqslant l_1$

$$\begin{aligned} I_n(\zeta) = \sum_{j=0}^{n} \left| \frac{\varphi_j(\zeta)}{\varphi_n(\zeta)} \right|^2 &= \sum_{j=0}^{l_1-1} \left| \frac{\varphi_j(\zeta)}{\varphi_n(\zeta)} \right|^2 + \sum_{j=l_1}^{n} \left| \frac{\varphi_j(\zeta)}{\varphi_n(\zeta)} \right|^2 \\ &\leqslant l_1 M^{l_1}(\zeta) + \sum_{j=0}^{\infty} \left(\frac{1}{4}\right)^j < \infty. \end{aligned}$$

The proof is complete. ▮

It is not hard to observe that the inclusion in Theorem 5.7 is proper. To this end take any $\sigma \in \mathrm{OS}$ with Geronimus parameters $\{a_n\}_{n \geqslant 0}$ and form the so-called "sieved" measure $\hat{\sigma}$ with Geronimus parameters

$$\hat{a}_{2n} = a_n, \qquad \hat{a}_{2n+1} = 0, \qquad n \in \mathbb{Z}_+ .$$

For the corresponding orthonormal polynomials we then have

$$\hat{\varphi}_{2n}(z) = \varphi_n(z^2), \qquad \hat{\varphi}_{2n+1}(z) = z\varphi_n(z^2).$$

It is easy to see by the definition of the classes G_0 and OS that $\hat{\sigma} \in G_0 \setminus \mathrm{OS}$.

The class OS has a nice interpretation in terms of Hessenberg matrices (see (14)–(15)).

THEOREM 5.8. *$\sigma \in OS$ if and only if*

$$U(\sigma)-D_0(\sigma) \in \mathfrak{S}_1. \tag{67}$$

Proof. Suppose first that $\sigma \in \mathrm{OS}$. We proceed as in the proof of [13, Lemma 14]. The key idea is to show that series (16) converges in the trace norm. We have

$$U(\sigma)-D_0(\sigma)=V^*D_{-1}(\sigma)+\sum_{j=1}^{\infty} D_j(\sigma)\, \mathrm{V}^{\mathrm{j}}, \tag{68}$$

where V, V^* are the shift operators and $D_j(\sigma)$ are the diagonal operators

$$D_j(\sigma)=\mathrm{diag}(u_{0j}, u_{1, j+1}, \ldots)$$

in the basis $\{\varphi_n\}_{n\geqslant 0}$. Then (64) implies $D_j(\sigma) \in \mathfrak{S}_1$ for $j=-1, 1, 2, \ldots$. Take a big enough N to obey

$$q \stackrel{\text{def}}{=} \sum_{n=N}^{\infty} (1-|a_n|^2)^{1/2}<1$$

and let $j>N$. By (15)

$$|u_{n, j+n}| \leqslant \prod_{k=n}^{j+n-1} (1-|a_k|^2)^{1/2}$$

and

$$\sum_{n=N}^{\infty} |u_{n, j+n}| \leqslant \sum_{n=N}^{\infty} \prod_{k=n}^{j+n-1} (1-|a_k|^2)^{1/2} \leqslant q^{j-1} \sum_{n=N}^{\infty} (1-|a_n|^2)^{1/2} \leqslant q^j.$$

Next, for $n<N$

$$|u_{n, j+n}| \leqslant \prod_{k=N}^{j-1} (1-|a_k|^2)^{1/2} \leqslant q^{j-N}.$$

Hence $\|D_j(\sigma)\|_1=\sum_{n\geqslant 0}|u_{n, j+n}| \leqslant Nq^{j-N}+q^j$, which means that the series (68) converges in the trace norm and (67) holds.

Conversely, it follows from (67) that $T=V^*(U-D_0) \in \mathfrak{S}_1$, so that

$$\sum_{j=0}^{\infty} |(T\varphi_j, \varphi_j)| \leqslant \|T\|_1.$$

(cf. [2, Lemma XI.9.13]). By the definition of V we have

$$(T\varphi_j, \varphi_j) = ((U - D_0)\, \varphi_j, V\varphi_j) = ((U - D_0)\, \varphi_j, \varphi_{j+1}) = (U\varphi_j, \varphi_{j+1}) = u_{j+1,\, j},$$

and (64) follows, as claimed. ▮

The same argument leads to the following conclusion

$$\sum_{n=0}^{\infty} (1 - |a_n|^2)^{p/2} < \infty \Leftrightarrow U(\sigma) - D_0(\sigma) \in \mathfrak{S}_p, \qquad 0 < p < \infty.$$

One may consider the opposite ON of Nevai's class. Following the convention accepted in the definition of OS, we can define ON as the class of probability measures on $\mathbb{T}$ with $\lim_{n \to \infty} |a_n| = 1$ (cf. [14, p. 62]). By Rakhmanov's lemma [28, Lemma 4], every measure in ON is singular. As in Theorem 5.8, it is not hard to ascertain that $\sigma \in \mathrm{ON}$ if and only if $U(\sigma) - D_0(\sigma) \in \mathfrak{S}_\infty$.

Assume now that for $\sigma \in \mathscr{P}$ the relation

$$\sum_{n=0}^{\infty} (1 - |a_n|) < \infty$$

holds (the latter is true for all $\sigma \in \mathrm{OS}$). Then the zeros $\{\lambda_{j,\,n}\}_{j=1}^{n}$ of the polynomials φ_n behave rather regularly. Specifically, they satisfy the Blaschke condition

$$\sum_{n=1}^{\infty} \sum_{j=1}^{n} (1 - |\lambda_{j,\,n}|) < \infty. \tag{69}$$

Indeed, by the equation $|a_{n-1}| = |\lambda_{1,\,n}| \cdots |\lambda_{n,\,n}|$ we have

$$1 - |a_{n-1}| = \sum_{j=1}^{n} (1 - |\lambda_{j,\,n}|)\, |\lambda_{j+1,\,n}| \cdots |\lambda_{n,\,n}| > |a_{n-1}| \sum_{j=1}^{n} (1 - |\lambda_{j,\,n}|),$$

so that (69) follows.

We complete the section with sort of strong asymptotics off the unit circle that holds for the monic orthogonal polynomials in the OS class. Put $a_{-1} = -1$ and define for $n \in \mathbb{Z}_+$

$$\tau_n = \begin{cases} -\dfrac{\bar{a}_n a_{n-1}}{|\bar{a}_n a_{n-1}|}, & \text{for} \quad a_n a_{n-1} \neq 0, \\ 1, & \text{for} \quad a_n a_{n-1} = 0. \end{cases}$$

Denote by Clos $\{\tau_0, \tau_1, ...,\}$ the closure of the set $\{\tau_n\}_{n\geqslant 0}$ on $\mathbb{T}$.

THEOREM 5.9. *Let* $\sigma \in OS$ *with the monic orthogonal polynomials* $\{\Phi_n\}_{n\geqslant 0}$. *Then there is a holomorphic function* H *in the domain* $\hat{\mathbb{C}} \setminus Clos\{\tau_0, \tau_1, ...,\}$ *with no zeros inside either components of* $\hat{\mathbb{C}} \setminus \mathbb{T}$ *and such that*

$$\lim_{n\to\infty} \frac{\Phi_{n+1}(z)}{(z-\tau_0)(z-\tau_1)\cdots(z-\tau_n)} = H(z) \tag{70}$$

uniformly on compact subsets of $\hat{\mathbb{C}} \setminus Clos\{\tau_0, \tau_1, ...,\}$.

Proof. Put $\tilde{D} = \operatorname{diag}(\tau_0, \tau_1, ...,)$ in the basis $\{\varphi_n\}_{n\geqslant 0}$. We state that $U(\sigma) - \tilde{D} \in \mathfrak{S}_1$. By Theorem 5.8 it suffices to prove that $D_0(\sigma) - \tilde{D} \in \mathfrak{S}_1$. The latter is equivalent to the convergence of the series

$$\sum_{n=0}^{\infty} |\tau_n + \bar{a}_n a_{n-1}| = \sum_{n=0}^{\infty} (1 - |a_n|\,|a_{n-1}|).$$

The statement now follows from (64) and the elementary inequality

$$1 - xy \leqslant (1-x^2) + (1-y^2), \qquad 0 \leqslant x, y \leqslant 1.$$

An intimate relation between monic orthogonal polynomials and Hessenberg matrices is well-known (cf. [1, pp. 193–194; 18])

$$\begin{aligned}\Phi_{n+1}(z) &= \det(zI_n - U_n(\sigma))\\ &= \det(zI_n - \tilde{D}_n)\det(I_n + (zI_n - \tilde{D}_n)^{-1}(\tilde{D}_n - U_n(\sigma)).\end{aligned}$$

Here I_n, U_n, D_n are $n+1$-dimensional truncated operators (see Section 1). Observe that $\det(zI_n - \tilde{D}_n) = (z-\tau_0)\cdots(z-\tau_n)$. By the basic property of infinite determinants (cf. [11, p. 207])

$$\lim_{n\to\infty} \frac{\Phi_{n+1}(z)}{(z-\tau_0)(z-\tau_1)\cdots(z-\tau_n)} = H(z) \tag{71}$$

uniformly on compact subsets of $\hat{\mathbb{C}} \setminus \operatorname{Clos}\{\tau_0, \tau_1, ...,\}$, where [7]

$$H(z) = \det(I + (zI - \tilde{D})^{-1}(\tilde{D} - U(\sigma))).$$

Next, it is clear that

$$\lim_{|z|\to\infty} \|(zI - \tilde{D})^{-1}(\tilde{D} - U(\sigma))\|_1 = 0,$$

[7] Note that $(AB)_n = A_n B_n$ whenever A is a diagonal operator.

and hence H does not vanish in a neighborhood of ∞. On the other hand,

$$|H(0)| = \lim_{n\to\infty} |\det(D_n^{-1}U_n)| = \lim_{n\to\infty} |\det U_n| = \lim_{n\to\infty} |\Phi_{n+1}(0)| = 1,$$

and hence H is not identically zero in $\mathbb{D}$ either.

Assume that $H(z_0)=0$ for some $z_0 \in \mathbb{D}$. Then by (71) and Hurwitz's theorem z_0 is a limit point of the zeros of Φ_n. But the latter is incompatible with (69). The same argument is applied to the exterior of the unit disk (there are no zeros at all at the exterior of the unit disk). The proof is complete. ∎

EXAMPLE 5.10. Consider the so-called symmetric Wall measure on the unit circle (cf. [31, p. 89]). This measure is known to be a pure point measure with $\zeta=-1$ being the only accumulating point of its support. The Geronimus parameters are of the form

$$a_{2n} = 1-2bq^n, \qquad a_{2n-1} = 1-2q^n, \qquad 0<b, q<1, \quad n=0, 1, \ldots .$$

Hence $\tau_0 = 1$, $\tau_k = -1$ for $k \geqslant 1$ and (70) turns into

$$\lim_{n\to\infty} \frac{\Phi_{n+1}(z)}{(z-1)(z+1)^n} = H(z)$$

uniformly inside $\hat{\mathbb{C}} \setminus \{\pm 1\}$, where φ_n are the symmetric Wall polynomials on the unit circle.

6. SOME MORE RESULTS ON I_n

We begin with the mass points problem as posed in the introduction.

THEOREM 6.1. *Let $\sigma \in CN$ and $\zeta_0 \in \mathbb{T}$. Then σ_t (18) is also in CN.*

Proof. Let $\Phi_n^{(t)}$ be the monic orthogonal polynomials in $L^2(\sigma_t)$. It is known (see [12, p. 36]) that

$$\Phi_{n+1}^{(t)}(z) = \Phi_{n+1}(z) - \frac{s\Phi_{n+1}(\zeta_0)\, K_{n+1}(z, \zeta_0)}{1+sK_{n+1}(\zeta_0, \zeta_0)}, \qquad s = \frac{t}{1-t}. \tag{72}$$

Let $a_n^{(t)}$ be the Geronimus parameters of σ_t. By taking $z=0$ in (72) and using the relation $K_{n+1}(0, \zeta_0) = \kappa_n \overline{\varphi_n^*(\zeta_0)}$ (see (11)) we come to

$$|a_n^{(t)} - a_n| = \frac{s\, |\Phi_{n+1}(\zeta_0)\, \kappa_n \varphi_n(\zeta_0)|}{1+sK_{n+1}(\zeta_0, \zeta_0)}.$$

But $|\Phi_{n+1}(\zeta_0)| = |\zeta_0 \Phi_n(\zeta_0) - \bar{a}_n \Phi_n^*(\zeta_0)| \leqslant 2\,|\Phi_n(\zeta_0)|$, so that

$$|a_n^{(t)} - a_n| \leqslant \frac{2s\,|\varphi_n(\zeta_0)|^2}{1 + sK_{n+1}(\zeta_0, \zeta_0)} < \frac{2}{I_n(\zeta_0)}. \tag{73}$$

By Theorem 5.1 there exists a subset $\Lambda \subset \mathbb{Z}_+$ with $d(\Lambda) = 1$ such that $1/I_n(\zeta_0) \to 0$ as $n \to \infty$ along Λ. By (73) we obtain

$$\lim_{n \in \Lambda} |a_n^{(t)} - a_n| = 0, \qquad 0 \leqslant t < 1.$$

As $\sigma \in \mathrm{CN}$, the sequence $\{a_n\}_{n \geqslant 0}$ is almost convergent to zero, and so is $\{a_n^{(t)}\}_{n \geqslant 0}$ for all $t \in [0, 1)$, which gives $\sigma_t \in \mathrm{CN}$. ∎

It is clear from (73) and MNT's theorem that $\sigma_t \in \mathrm{N}$ as long as $\sigma \in \mathrm{N}$. The question arises whether the same is true for the intermediate Rakhmanov's class. We can prove this only for a certain proper subclass of R that contains N.

THEOREM 6.2. *Let $\sigma \in R$ and assume that* $\sup_n |a_n| < 1$. *Then*

$$\lim_{n \to \infty} \max_{\zeta \in \mathbb{T}} \frac{1}{I_n(\zeta)} = 0. \tag{74}$$

In particular, σ_t is a Rakhmanov measure for $t \in [0, 1)$ as long as $\sigma = \sigma_0$ is.

Proof. We proceed in three steps.

Step 1. The following property of the class R is worth mentioning. Let $\sigma \in \mathrm{R}$ with the Geronimus parameters $\{a_n\}_{n \geqslant 0}$. Fix a positive integer l and consider the l-dimensional vector

$$v_n = \{a_{n+1}, a_{n+2}, \ldots, a_{n+l}\}.$$

Rearrange the entries in decreasing order of their moduli

$$v_n = \{a'_{n+1}, a'_{n+2}, \ldots, a'_{n+l}\}, \qquad |a'_{n+1}| \geqslant |a'_{n+2}| \geqslant \cdots \geqslant |a'_{n+l}|.$$

We wish to show that $\lim_{n \to \infty} |a'_{n+2}| = 0$.[8] Indeed, assume that $\limsup_{n \to \infty} |a'_{n+2}| = 2\delta > 0$. Then for some subset $\Lambda \subset \mathbb{Z}_+$ and $n \in \Lambda$ we have

$$|a'_{n+2}| > \delta, \qquad |a'_{n+1}| > \delta, \qquad |a'_{n+1}|\,|a'_{n+2}| > \delta^2.$$

But the latter contradicts [19, Theorem 4], according to which $\lim_{n \to \infty} a_n a_{n+k} = 0$ holds in the class R for each $k = 1, 2, \ldots$.

[8] For $\sigma \in N$ we have $\lim_{n \to \infty} |a'_{n+1}| = 0$.

Step 2. The argument here is of general nature. Write

$$I_n(\zeta)=\frac{K_{n+1}(\zeta,\zeta)}{|\varphi_n(\zeta)|^2}=\sum_{k=0}^{n}\left|\frac{\varphi_{n-k}(\zeta)}{\varphi_n(\zeta)}\right|^2.$$

By (65)

$$\left|\frac{\varphi_j(\zeta)}{\varphi_{j+1}(\zeta)}\right|^2=\frac{1-|a_j|^2}{|1-a_j\zeta b_j(\zeta)|^2}\geqslant\frac{1-|a_j|}{1+|a_j|},$$

and hence

$$\left|\frac{\varphi_{n-k}(\zeta)}{\varphi_n(\zeta)}\right|^2\geqslant\prod_{j=1}^{k}\frac{1-|a_{n-j}|}{1+|a_{n-j}|}=\prod_{p=n-k}^{n-1}h(|a_p|),\qquad h(x)\stackrel{\text{def}}{=}\frac{1-x}{1+x}.$$

In terms of the rearranged values this inequality can be displayed as

$$\left|\frac{\varphi_{n-k}(\zeta)}{\varphi_n(\zeta)}\right|^2\geqslant\prod_{p=n-k}^{n-1}h(|a'_p|)=h(|a'_{n-k}|)\prod_{p=n-k+1}^{n-1}h(|a'_p|).$$

The function h decreases from 1 to 0 on $[0,1]$. If now $\sup_n|a_n|<r<1$, then

$$\left|\frac{\varphi_{n-k}(\zeta)}{\varphi_n(\zeta)}\right|^2\geqslant h(r)\,h^{k-1}(|a'_{n-k+1}|).$$

Take a positive integer N and let $n>N$. We have

$$I_n(\zeta)\geqslant\sum_{k=0}^{N}\left|\frac{\varphi_{n-k}(\zeta)}{\varphi_n(\zeta)}\right|^2\geqslant h(r)\sum_{k=0}^{N}h^{k-1}(|a'_{n-k+1}|).$$

By Step 1 the right-hand side tends to $h(r)(N+1)$ as $n\to\infty$, and since N is arbitrary and $h(r)>0$, we conclude that

$$\lim_{n\to\infty}\min_{\zeta\in\mathbb{T}}I_n(\zeta)=+\infty,$$

which is equivalent to (74).

Step 3. To prove the last statement note that (73) and (74) imply

$$\lim_{n\to\infty}|a_n^{(t)}-a_n|=0,\qquad\lim_{n\to\infty}|a_n^{(t)}a_{n+k}^{(t)}-a_na_{n+k}|=0$$

for each $k=1,2,\ldots$. The result then follows from [19, Theorem 4]. ∎

By comparing Theorem 6.2 with MNT's theorem we see that, for measures $\sigma \in \mathrm{R}\setminus\mathrm{N}$ (all of them are singular by [19, Corollary 2.6]) with $\sup_n |a_n| < 1$, (74) holds but $1/I_n$ tends to zero at no point in $\mathbb{D}$.

In the same fashion as Theorem 6.1, the following result can be established.

THEOREM 6.3. *Given $\sigma \in \mathcal{P}$ and $\zeta_0 \in \mathbb{T}$, define $d\mu = C\,|\zeta - \zeta_0|^2\, d\sigma \in \mathcal{P}$, where C is a normalizing positive factor. Then μ is in the CN class as long as σ is.*

There is a nice formula which relates I_n to the zeros $\lambda_{1n}, \lambda_{2n}..., \lambda_{nn}$ of orthogonal polynomials φ_n. As is known, $|\lambda_{kn}| < 1$ for all $n = 1, 2, ...$ and $k = 1, 2, ..., n$. Let us denote by $\dot{x}$ the derivative of a function x on $\mathbb{T}$ with respect to the arc length, i.e., $\dot{x} = \partial x/\partial \vartheta$. We have

$$\frac{\partial}{\partial\vartheta}\,\frac{e^{i\vartheta}-\lambda}{1-\bar{\lambda}e^{i\vartheta}} = ie^{i\vartheta}\,\frac{1-|\lambda|^2}{(1-\bar{\lambda}e^{i\vartheta})^2} = ie^{i\vartheta}\,\frac{\overline{1-\bar{\lambda}e^{i\vartheta}}}{1-\bar{\lambda}e^{i\vartheta}}\,\frac{1-|\lambda|^2}{|1-\bar{\lambda}e^{i\vartheta}|^2} = i\,\frac{e^{i\vartheta}-\lambda}{1-\bar{\lambda}e^{i\vartheta}}\,\frac{1-|\lambda|^2}{|1-\bar{\lambda}e^{i\vartheta}|^2}.$$

Now, letting $b_n(\zeta) = e^{i\alpha_n(\zeta)}$, we see that

$$\dot{\alpha}_n(\zeta) = \sum_{j=1}^{n} \frac{1-|\lambda_{jn}|^2}{|\zeta-\lambda_{jn}|^2}. \tag{75}$$

Hence any continuous branch α_n of the argument of the Blaschke product $b_n(\zeta)$ increases as ζ moves along $\mathbb{T}$ counterclockwise.

THEOREM 6.4. *For every $\sigma \in \mathcal{P}$ and $n \in \mathbb{Z}_+$ the relation*

$$I_n(\zeta) = 1 + \sum_{j=1}^{n} \frac{1-|\lambda_{jn}|^2}{|\zeta-\lambda_{jn}|^2} \tag{76}$$

holds.

Proof. By (56)

$$I_n(z) = \frac{1-|zb_n(z)|^2}{(1-|z|^2)\,|b_n(z)|^2}. \tag{77}$$

It is clear that

$$zb_n(z) = \prod_{j=0}^{n} \beta_{jn}(z), \qquad \beta_{jn}(z) = \frac{z-\lambda_{jn}}{1-\bar{\lambda}_{jn}z}, \qquad \lambda_{0n} = 0.$$

Since

$$1-|zb_n(z)|^2=\sum_{j=0}^{n}|\beta_0(z)\,\beta_1(z)\cdots\beta_{j-1}(z)|^2\,(1-|\beta_j(z)|^2)$$

we have by (77)

$$I_n(\zeta)=\lim_{r\to 1-0}\frac{1-|r\zeta b_n(r\zeta)|^2}{(1-r^2)\,|b_n(r\zeta)|^2}=1+\sum_{j=1}^{n}\lim_{r\to 1-0}\frac{1-|r\zeta\beta_{jn}(r\zeta)|^2}{1-r^2}.$$

It is a matter of the routine computation to verify that

$$\lim_{r\to 1-0}\frac{1-|r\zeta\beta_{jn}(r\zeta)|^2}{1-r^2}=\frac{1-|\lambda_{jn}|^2}{|\zeta-\lambda_{jn}|^2},$$

and (76) is proved. ∎

It follows from (75) and (76) that for $\zeta\in\mathbb{T}$

$$I_n(\zeta)=1+\dot{\alpha}_n(\zeta), \tag{78}$$

which is (19). By the definition of the class G_0 we see that $\sup_n\dot{\alpha}_n<\infty$ m-a.e. for each $\sigma\in G_0$.

To examine the convergence of

$$\frac{\dot{\alpha}_n(\zeta)}{n}=\frac{1}{n\,|\varphi_n(\zeta)|^2}\sum_{k=0}^{n-1}|\varphi_k(\zeta)|^2, \tag{79}$$

we need the following simple result.

LEMMA 6.5. *Let $\{h_n\}$ be a sequence of nonnegative measurable functions on a probability space (X,ν), and let $h_n\Rightarrow 1$ in measure as $n\to\infty$ and*

$$\int_X h_n\,d\nu=1,\qquad \forall n=1,2,\ldots.$$

Then $h_n\to 1$ in $L^1(\nu)$.

Proof. Given $\varepsilon>0$ put $E_n(\varepsilon)\overset{\text{def}}{=}\{x\in X:|h_n-1|>\varepsilon\}$. By the hypothesis of the lemma $\nu(E_n)\to 0$ as $n\to\infty$. We have

$$\int_X|h_n-1|\,d\nu\leqslant\varepsilon+\int_{E_n}|h_n-1|\,d\nu$$

$$\leqslant\varepsilon+\nu(E_n)+\int_{E_n}h_n\,d\nu=\varepsilon+2\nu(E_n)+\int_{E_n}(h_n-1)\,d\nu.$$

But

$$\left|\int_{E_n} (h_n - 1)\, d\nu\right| = \left|\int_{E_n^c} (h_n - 1)\, d\nu\right| \leqslant \int_{E_n^c} |h_n - 1|\, d\nu \leqslant \varepsilon,$$

so that by letting $n \to \infty$ we see that

$$\limsup_{n\to\infty} \int_X |h_n - 1|\, d\nu \leqslant 2\varepsilon.$$

The rest is plain. ∎

PROPOSITION 6.6. *For each Erdős measure σ (i.e., $\sigma' > 0$ m-a.e. on $\mathbb{T}$) the relation*

$$\lim_{n\to\infty} \int_{\mathbb{T}} \left|\frac{\dot{\alpha}_n(\zeta)}{n} - 1\right| dm = 0$$

holds.

Proof. By [23, Corollary 2.2], in the Erdős class

$$\lim_{n\to\infty} \int_{\mathbb{T}} ||\varphi_n(\zeta)|^2 \sigma' - 1|\, dm = 0,$$

which implies that $|\varphi_n|^2 \Rightarrow 1/\sigma'$ in measure m on $\mathbb{T}$. Hence $\dot{\alpha}_n/n \Rightarrow 1$ by (79). Next,

$$\int_{\mathbb{T}} \dot{\alpha}_n(\zeta)\, dm = n$$

by (75), and application of Lemma 6.5 completes the proof. ∎

Given $\zeta \in \mathbb{T}$ and $s > 0$, consider the set

$$W_s(\zeta) = \left\{ z : \frac{|\zeta - z|^2}{1 - |z|^2} < s \right\},$$

on the complex plane, which is known as *oricycle*. It is clear that $\{W_s\}_{s>0}$ is an increasing family of disks inside the unit disk which touch it at the point ζ, and $\mathbb{D} = \bigcup_{s>0} W_s(\zeta)$. Relation (76) leads to the following rather curious conclusion about the zeros distribution of orthogonal polynomials on the unit circle.

PROPOSITION 6.7. *Let $C = \sup_n I_n(\zeta_0) < \infty$. Then the number of zeros of φ_n on each oricycle $W_s(\zeta_0)$ is uniformly bounded in n.*

Proof. Let $\Lambda_n \stackrel{\text{def}}{=} \{\lambda_{1n}, \lambda_{2n} \ldots, \lambda_{nn}\}$. By (76)

$$C > \sum_{j=1}^{n} \frac{1-|\lambda_{jn}|^2}{|\zeta-\lambda_{jn}|^2} \geqslant \sum_{\lambda_{jn} \in W_s(\zeta)} \frac{1-|\lambda_{jn}|^2}{|\zeta-\lambda_{jn}|^2} > \frac{1}{s} \operatorname{Card}(\Lambda_n \cap W_s(\zeta_0)),$$

as needed. ∎

Note that nonasymptotic results can be obtained provided we know a quantitative bound for I_n. For instance, if, say, $I_n(1) \leqslant 2$ for all n then there are no zeros of φ_n in the oricycle $W_1(1)$ (and, in particular, in the interval $[0, 1]$).

ACKNOWLEDGMENTS

The authors thank Paul Nevai for helpful discussions. The authors are also grateful to the referee for a careful reading of the manuscript and for important remarks made in Section 2.

REFERENCES

1. G. S. Ammar, D. Calvetti, and L. Reichel, Continuation methods for the computation of zeros of Szegő polynomials, *in* "Orthogonal Polynomials on the Unit Circle: Theory and Applications" (M. Alfaro, A. Garcia, and F. Marcellan, Eds.), pp. 173–205, Universidad Carlos III de Madrid, Leganes, 1994.
2. N. Dunford and J. Schwartz, "Linear Operators," Interscience, New York/London, 1958, 1963.
3. T. Erdélyi, J. S. Geronimo, P. Nevai, and J. Zhang, A simple proof of "Favard's Theorem" on the unit circle, *Atti Sem. Mat. Fis. Univ. Modena* **29** (1991), 41–46.
4. J. Favard, Sur les polynômes de Tchebicheff, *C. R. Acad. Sci.* **200** (1935), 2052–2053.
5. J. B. Garnett, "Bounded Analytic Functions," Academic Press, London/New York, 1981.
6. Ya. L. Geronimus, On the character of the solutions of the moment problem in case of a limit-periodic associated fraction, *Izv. Akad. Nauk SSSR Ser. Mat.* **5** (1941), 203–210. [In Russian]
7. Ya. L. Geronimus, On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathéodory and Schur functions, *Mat. Sb.* **15**, No. 57 (1944), 99–130. [In Russian]
8. Ya. L. Geronimus, On asymptotic properties of polynomials, orthogonal on the unit circle, and on some properties of positive harmonic functions, *Izv. Akad. Nauk SSSR* **14** (1950), 123–144. [In Russian]
9. L. Ya. (aka Ya. L.) Geronimus, "Orthogonal Polynomials," Consultants Bureau, New York, 1961.
10. Ya. L. Geronimus, Polynomials orthogonal on a circle and their applications, *in* "Series and Approximations," Amer. Math. Soc. Transl. (1), Vol. 3, pp. 1–78, Providence, 1962.

11. I. Gohberg and M. G. Krein, "Introduction to the Theory of Linear Nonselfadjoint Operators in Hilbert Space," Nauka, Moscow, 1965. [In Russian]
12. B. L. Golinskii, On certain estimates for Christoffel kernels and moduli of orthogonal polynomials, *Izv. Vuz. Mat.* **1**, No. 50 (1966), 30–42. [In Russian]
13. L. Golinskii, Operator theoretic approach to orthogonal polynomials on an arc of the unit circle, *Math. Phys. Anal. Geom.* **7** (2000), 3–34.
14. L. Golinskii, Singular measures on the unit circle and their reflection coefficients, *J. Approx. Theory* **103** (2000), 61–77.
15. L. Golinskii, On the spectra of infinite Hessenberg and Jacobi matrices, *Math. Phys. Anal. Geom.* **7** (2000), 284–298.
16. L. Golinskii, Quadrature formula and zeros of para-orthogonal polynomials on the unit circle, manuscript.
17. L. Golinskii, P. Nevai, and W. Van Assche, Perturbation of orthogonal polynomials on an arc of the unit circle, *J. Approx. Theory* **83** (1995), 392–422.
18. W. B. Gragg, Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle, *J. Comput. Appl. Math.* **46** (1993), 183–198.
19. S. Khrushchev, Schur's algorithm, orthogonal polynomials and convergence of Wall's continued fractions in $L^2(\mathbb{T})$, *J. Approx. Theory* **108** (2001), 161–248.
20. S. Khrushchev, A singular Riesz product in Nevai's class and inner functions with Schur parameters in $\bigcap_{p>2}\ell^p$, *J. Approx. Theory* **108** (2001), 249–255.
21. S. Khrushchev, Classification theorems for general orthogonal polynomials on the unit circle, *J. Approx. Theory*, to appear.
22. A. Máté, P. Nevai, and V. Totik, Extensions of Szegő's theory of orthogonal polynomials, II, *Constr. Approx.* **3** (1987), 51–72.
23. A. Máté, P. Nevai, and V. Totik, Strong and weak convergence of orthogonal polynomials, *Amer. J. Math.* **109** (1987), 239–281.
24. E. M. Nikishin and V. N. Sorokin, "Rational Approximations and Orthogonality," Nauka, Moscow, 1988 [In Russian]; English translation, *in* Translations of Mathematical Monographs, Vol. 92, 1991.
25. F. Peherstorfer, A special class of polynomials orthogonal on the unit circle including the associated polynomials, *Constr. Approx.* **12** (1996), 161–185.
26. I. I. Privalov, "Boundary Properties of Analytic Functions," Gostekhizdat, Moscow, 1950. [In Russian]
27. E. A. Rahmanov, On the asymptotics of the ratio of orthogonal polynomials, *Math. USSR–Sb.* **32** (1977), 199–213; *Russian original*, *Mat. Sb.* **103**, No. 145 (1977), 237–252.
28. E. A. Rahmanov, On the asymptotics of the ratio of orthogonal polynomials, II, *Math. USSR–Sb.* **46** (1983), 105–117; *Russian original*, *Mat. Sb.* **118**, No. 160 (1982), 104–117.
29. H. S. Wall, Continued fractions and bounded analytic functions, *Bull. Amer. Math. Soc.* **50** (1944), 110–119.
30. G. Szegő, "Orthogonal Polynomials," 4th ed., Amer. Math. Soc. Colloq. Publ., Vol. 23, Amer. Math. Soc., Providence, 1975.
31. A. Zhedanov, On some classes of polynomials orthogonal on arcs of the unit circle connected with symmetric orthogonal polynomials on an interval, *J. Approx. Theory* **94** (1998), 73–106.
32. A. Zygmund, "Trigonometric Series," Cambridge Univ. Press, Cambridge, 1977.