9

Orthogonal Polynomials on the Unit Circle

Leonid Golinskii

One way to generalize orthogonal polynomials on subsets of $\mathbb{R}$ is to consider orthogonality on curves in the complex plane. Among these generalizations, the most developed theory is the general theory of orthogonal polynomial on the unit circle $\mathbb{T}$. The basic sources for this chapter are Grenander and Szegő (1958), Szegő (1975), Geronimus (1961), Geronimus (1962), Simon (2004a), Simon (2004b), Ismail (2005b, Chapters 8 and 17), and recent papers which will be cited in the appropriate places.

In what follows we shall use Simon's abbreviation OPUC for orthogonal polynomials on the unit circle.

9.1 Definitions and Basic Properties

The **unit circle** $\mathbb{T}$ is by far the simplest closed curve in the complex plane with a number of additional properties, so polynomials orthogonal with respect to measures on $\mathbb{T}$ are of specific interest.

Consider the class of all nontrivial probability measures $\mu(\theta)$ on $[-\pi,\pi]$ (that is, not supported on a finite set, positive Borel measures with $\mu[-\pi,\pi] = 1$). The Lebesgue decomposition of μ is the decomposition

$$\mu(\theta) = \mu_{\mathrm{ac}} + \mu_s = \mu'(\theta)\frac{d\theta}{2\pi} + \mu_s \tag{9.1.1}$$

where $\mu' \in L^1([-\pi,\pi])$ is the Radon–Nikodym derivative of μ with respect to the Lebesgue measure and μ_s is the singular part of μ.

The **moments** (Fourier coefficients) of μ are defined by

$$\mu_k = \int_{-\pi}^{\pi} e^{-ik\theta}\, d\mu(\theta), \qquad k \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}, \tag{9.1.2}$$

and form a bounded sequence. The moments of μ generate the **Toeplitz determinants**

$$D_n = D_n(\mu) = \det \|\mu_{i-k}\|_{i,k=0}^{n} > 0. \tag{9.1.3}$$

The theory of quadratic forms shows that D_n is strictly positive for all $n \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$.

The orthogonal polynomials with respect to μ arise as an outcome of the standard Gram–Schmidt procedure applied to the system of monomials $\{\zeta^n\}_{n\ge 0}$, $\zeta = e^{i\theta}$, in the Hilbert space $L^2_\mu([-\pi,\pi])$ of square-summable measurable functions on $\mathbb{T}$ with the inner product

$$(f,g)_\mu = \int_{-\pi}^{\pi} f(\zeta)\overline{g(\zeta)}\,d\mu(\theta), \quad \zeta = e^{i\theta}, \quad \|f\|^2_\mu = (f,f)_\mu.$$

There are two natural ways of normalization: the **orthonormal polynomials**

$$\phi_n(z) = \phi_n(z;\mu) = \kappa_n z^n + \text{ lower-order terms}, \quad (\phi_n,\phi_m)_\mu = \delta_{nm}, \tag{9.1.4}$$

$n, m \in \mathbb{Z}_+$, and the **monic orthogonal polynomials**

$$\Phi_n(z) = \Phi_n(z;\mu) = \kappa_n^{-1}\phi_n(z) = z^n + \ell_{n,n-1}z^{n-1} + \text{ lower-order terms}. \tag{9.1.5}$$

Both systems are uniquely determined when we demand that $\kappa_n > 0$. The monic orthogonal polynomials are characterized by the property

$$\deg(P) = n, \quad (P,\zeta)_\mu = 0, \quad 0 \le j < n \quad \text{imply} \quad P = c_n\Phi_n. \tag{9.1.6}$$

The following expressions for monic orthogonal polynomials are similar to (2.1.4) and (2.1.6):

$$\Phi_n(z) = \frac{1}{D_{n-1}}\begin{vmatrix} \mu_0 & \mu_{-1} & \dots & \mu_{-n} \\ \mu_1 & \mu_0 & \dots & \mu_{-n+1} \\ \vdots & \vdots & & \vdots \\ \mu_{n-1} & \mu_{n-2} & \dots & \mu_{-1} \\ 1 & z & \dots & z^n \end{vmatrix} \tag{9.1.7}$$

and

$$\Phi_n(z) = \frac{1}{n!D_{n-1}}\int_{\mathbb{T}^n}\prod_{j=1}^{n}(z-\zeta_j)\prod_{1\le j<k\le n}|\zeta_j-\zeta_k|^2\prod_{j=1}^{n}d\mu(\zeta_j). \tag{9.1.8}$$

Equation (9.1.7) implies an important relation

$$(\Phi_n, z^n)_\mu = \|\Phi_n\|^2_\mu = \frac{D_n}{D_{n-1}}. \tag{9.1.9}$$

Let z_0 be a zero of Φ_n. Following an elegant argument due to H. Landau (Landau, 1987), we write $\Phi_n(z) = (z - z_0)P(z)$, $\deg P = n - 1$, so $\Phi_n \perp P$ and

$$zP(z) = \Phi_n(z) + z_0P(z), \quad \|zP\|^2_\mu = \|P\|^2_\mu = \|\Phi_n\|^2_\mu + |z_0|^2\|P\|^2_\mu,$$

hence $(1 - |z_0|^2)\|P\|^2_\mu = \|\Phi_n\|^2_\mu$ and so $|z_0| < 1$. In other words, all zeros of all orthogonal polynomials lie in the open unit disk $\mathbb{D} = \{|z| < 1\}$.

The following **extremal property** of monic orthogonal polynomials is one of the highlights of OPUC theory.

Theorem 9.1.1 *The minimum of the integral*

$$\int_{-\pi}^{\pi} |P(\zeta)|^2 \, d\mu(\theta) \tag{9.1.10}$$

taken over all monic polynomials P of degree n is attained when $P = \Phi_n$. The minimum value of the integral is κ_n^{-2}.

As a straightforward consequence we have Simon's variational principle (Simon, 2007b), which proved useful in the study of Schur and related flows. Note that one can define monic OPUC for any finite positive measure, even if not normalized, and of course $\Phi_n(z; c\mu) = \Phi_n(z; \mu)$ for any positive constant c.

Theorem 9.1.2 *Let μ be a nontrivial probability measure on $[-\pi, \pi]$, and $\{z_j\}_{j=1}^k$ be among the zeros of $\Phi_n(\mu)$. Then*

$$\Phi_n(z; \mu) = \prod_{j=1}^{k} (z - z_j) \Phi_{n-k} \left(z; \prod_{j=1}^{k} |z - z_j|^2 \, d\mu \right).$$

The **reverse polynomial** f^* of a polynomial f of degree n is $f^*(z) = z^n \overline{f}(1/z)$, that is,

$$f^*(z) = \sum_{k=0}^{n} \bar{f}_{n-k} z^k \quad \text{if} \quad f(z) = \sum_{k=0}^{n} f_k z^k. \tag{9.1.11}$$

Equation (9.1.6) now shows that

$$\deg(P) \le n, \quad P \perp \zeta^j, \quad j = 1, \ldots, n \quad \text{imply} \quad P = c_n \Phi_n^*. \tag{9.1.12}$$

The polynomials Φ_n^* are called the **$*$-reverse polynomials**. Clearly, $\Phi_n^*(0) = 1$.

The next result shows how systems of orthogonal polynomials on $\mathbb{T}$ are in one-to-one correspondence with pairs of special systems of polynomials orthogonal on $[-1, 1]$. The model is $\{z^n\}$ on $\mathbb{T}$ and the Chebyshev polynomials $\{\operatorname{Re} z^n\}$ and $\{\operatorname{Im} z^{n+1} / \operatorname{Im} z\}$ on $[-1, 1]$.

Theorem 9.1.3 (Szegő's mapping theorem) *Let $d\mu(x)$ be a probability measure on $[-1, 1]$ and let ϕ_n be the polynomials orthonormal with respect to $d\mu(\cos\theta)$ on the unit circle. Assume further that $\{t_n(x)\}$ and $\{u_n(x)\}$ are orthonormal sequences of polynomials whose measures of orthogonality are $d\mu(x)$ and $c_2 \left(1 - x^2\right) d\mu(x)$, respectively. With $z \in \mathbb{T}$ and $x = (z + 1/z)/2$ we have*

$$\begin{aligned} t_n(x) &= \left[1 + \phi_{2n}(0)/\kappa_{2n}\right]^{-1/2} \left[z^{-n} \phi_{2n}(z) + z^n \phi_{2n}(1/z)\right] \\ &= \left[1 - \phi_{2n}(0)/\kappa_{2n}\right]^{-1/2} \left[z^{-n+1} \phi_{2n-1}(z) + z^{n-1} \phi_{2n-1}(1/z)\right] \end{aligned} \tag{9.1.13}$$

and

$$\begin{aligned} u_n(x) &= \frac{z^{-n-1} \phi_{2n+2}(z) + z^{n+1} \phi_{2n+2}(1/z)}{\sqrt{1 - \phi_{2n+2}(0)/\kappa_{2n+2}}\, (z - 1/z)} \\ &= \frac{z^{-n} \phi_{2n+1}(z) + z^n \phi_{2n+1}(1/z)}{\sqrt{1 + \phi_{2n+2}(0)/\kappa_{2n+2}}\, (z - 1/z)}. \end{aligned} \tag{9.1.14}$$

9.2 Szegő Recurrence Relations and Verblunsky Coefficients

A key feature of the unit circle is that the multiplication operator $Uf = zf$ in $L^2_\mu(\mathbb{T})$ is unitary. So the difference $\Phi_{n+1}(z) - z\Phi_n(z)$ is of degree at most n and orthogonal to z^j for $j = 1, 2, \ldots, n$, and by (9.1.12),

$$\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n \Phi_n^*(z) \tag{9.2.1}$$

with some complex numbers α_n, known as the **Verblunsky coefficients**,

$$\alpha_n = -\overline{\Phi_{n+1}(0)} = (-1)^n \prod_{j=1}^{n+1} \bar{z}_{j,n+1}, \quad |\alpha_n| = \prod_{j=1}^{n+1} |z_{j,n+1}| < 1, \tag{9.2.2}$$

where $\{z_{j,n+1}\}$ are zeros of $\Phi_{n+1}(\mu)$. Applying (9.1.11) to (9.2.1) yields

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z \Phi_n(z). \tag{9.2.3}$$

The recurrence relations (9.2.1) and (9.2.3) are the **Szegő recurrences**.

It follows from the unitarity of U and $\Phi_n^* \perp \Phi_{n+1}$, that

$$\|\Phi_{n+1}\|^2 = (1 - |\alpha_n|^2)\|\Phi_n\|^2, \quad \|\Phi_n\|^2 = \kappa_n^{-2} = \prod_{j=0}^{n-1}(1 - |\alpha_j|^2). \tag{9.2.4}$$

We set

$$\rho_j := \sqrt{1 - |\alpha_j|^2}, \quad \text{so that } 0 < \rho_j \leq 1. \tag{9.2.5}$$

Thus the leading coefficients κ_n satisfy $\kappa_{n+1}^{-1} = \rho_n \kappa_n^{-1}$, hence are given by

$$\kappa_n = \prod_{j=0}^{n} -1(1/\rho).$$

Combining (9.2.1) and (9.2.3) we obtain the Szegő recurrence relations in matrix form:

$$\begin{bmatrix} \Phi_{n+1}(z) \\ \Phi_{n+1}^*(z) \end{bmatrix} = A(z, \alpha_n) \begin{bmatrix} \Phi_n(z) \\ \Phi_n^*(z) \end{bmatrix}, \quad A(z, \alpha) = \begin{bmatrix} z & -\bar{\alpha} \\ -z\alpha & 1 \end{bmatrix}. \tag{9.2.6}$$

In other words,

$$\begin{bmatrix} \Phi_{n+1}(z) \\ \Phi_{n+1}^*(z) \end{bmatrix} = T_{n+1}(z) \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T_p(z) := A(z, \alpha_{p-1}) \ldots A(z, \alpha_0). \tag{9.2.7}$$

The matrix $T_p(z) =$ is called a **transfer matrix**. This leads to the **inverse Szegő recurrences**

$$z\Phi_n(z) = \rho_n^{-2} \left[\Phi_{n+1}(z) + \bar{\alpha}_n \Phi_{n+1}^*(z)\right],$$
$$\Phi_n^*(z) = \rho_n^{-2} \left[\Phi_{n+1}^*(z) + \alpha_n \Phi_{n+1}(z)\right].$$

By eliminating Φ_n^* between the direct and inverse Szegő recurrences we get the three-term recurrence relation (see Geronimus, 1962)

$$\bar{\alpha}_{n-1}\Phi_{n+1}(z) = (\bar{\alpha}_n + \bar{\alpha}_{n-1} z)\Phi_n(z) - \bar{\alpha}_n \rho_{n-1}^2 z \Phi_n(z) \tag{9.2.8}$$

for the Φ_j without any Φ_j^*, which is helpful for the study of the ratio asymptotics of orthogonal polynomials. Equation (9.2.8) has the defect that α_{n-1} can vanish.

The Szegő recurrence relations for orthonormal polynomials are

$$\begin{bmatrix}\phi_{n+1}(z)\\ \phi_{n+1}^*(z)\end{bmatrix} = \frac{1}{\rho_n}A(z,\alpha_n)\begin{bmatrix}\phi_n(z)\\ \phi_n^*(z)\end{bmatrix} = \prod_{j=0}^{n}\frac{1}{\rho_j}A(z,\alpha_j)\begin{bmatrix}1\\ 1\end{bmatrix}. \tag{9.2.9}$$

The following fundamental result is proved in Verblunsky (1935).

Theorem 9.2.1 (Verblunsky's theorem) *Let $\mathbb{D}^\infty$ be the set of complex sequences $\{\alpha_j\}_{j=0}^\infty$ with $|\alpha_j| < 1$. Let $\mathcal{S}$ be the mapping from the set of all nontrivial probability measures on $[-\pi,\pi]$ to $\mathbb{D}^\infty$ defined by $S(\mu) = \{\alpha_j(\mu)\}_{j=0}^\infty$. Then $\mathcal{S}$ is one-to-one. Moreover, $\mathcal{S}$ is a homeomorphism if the space of measures and $\mathbb{D}^\infty$ are equipped with the weak* topology and the componentwise convergence topology, respectively.*

For a detailed discussion and several proofs see Simon (2004a). In fact, the analysis in Verblunsky (1935) shows that the moments μ_n of every such measure can be parametrized by elements of $\mathbb{D}^\infty$ via

$$\mu_{n+1} = \text{polynomial in } \{\alpha_0,\bar\alpha_0,\ldots,\alpha_{n-1},\bar\alpha_{n-1}\} + \alpha_n\prod_{j=0}^{n-1}\rho_j^2.$$

Theorem 9.2.2 (Bernstein–Szegő approximation) *Let μ be a nontrivial probability measure on $[-\pi,\pi]$ with orthonormal polynomials ϕ_n. Let*

$$\mu^{(n)} = |\phi_n(\zeta;\mu)|^{-2}\,dm. \tag{9.2.10}$$

Then $\mu^{(n)}$ belongs to the same class of measures, with

$$\phi_k(z;\mu^{(n)}) = \phi_k(z;\mu),\quad k = 0,1,\ldots,n;\qquad \phi_k(z;\mu^{(n)}) = z^{k-n}\phi_n(z;\mu) \tag{9.2.11}$$

for $k \ge n$, so

$$\alpha_j(\mu^{(n)}) = \alpha_j(\mu),\quad j = 0,1,\ldots,n-1;\qquad \alpha_j(\mu^{(n)}) = 0,\quad j \ge n. \tag{9.2.12}$$

*Moreover, $\mu^{(n)} \to \mu$ as $n \to \infty$ in the *-weak topology.*

In fact, the measures with finite sequences of Verblunsky coefficients are exactly those of the form $\mu = c|P(\zeta)|^{-2}\,d\theta$, where c is picked to make μ a probability measure, and P is a monic polynomial of degree n with all zeros in $\mathbb{D}$. In this case $\Phi_k(z;\mu) = z^{k-n}P(z)$ for $k \ge n$.

The relation between measures μ and their Verblunsky coefficients in Theorem 9.2.1 is quite intricate, and very little can be said in the general setting. But there is an important situation – rotation of α – when some information about a corresponding family of measures is available. Specifically, let $\lambda \in \mathbb{T}$ and put $\alpha_{n,\lambda} = \lambda\alpha_n$, $n \in \mathbb{Z}_+$. The measures μ_λ with $\alpha_n(\mu_\lambda) = \alpha_{n,\lambda}$ (which exist by Theorem 9.2.1) are known as the **Aleksandrov measures** (or Aleksandrov–Clark measures). A representative with $\lambda = -1$ is of particular interest. The measure μ_{-1}

is called a **measure of the second kind**, and the corresponding orthogonal polynomials are called **polynomials of the second kind**. Special notation,

$$\Phi_n(z;\mu_{-1}) = \Psi_n(z), \quad \phi_n(z;\mu_{-1}) = \psi_n(z), \tag{9.2.13}$$

is standard for the monic orthogonal (orthonormal) polynomials of the second kind, respectively. The explicit formulas for Ψ_n are due to Geronimus (1961):

$$\Psi_n(z) = \int_{-\pi}^{\pi} \frac{\zeta+z}{\zeta-z}\left[\Phi_n(\zeta) - \Phi_n(z)\right] d\mu(\theta), \quad \zeta = e^{i\theta}, \tag{9.2.14}$$

$$\Psi_n^*(z) = z^n \int_{-\pi}^{\pi} \frac{\zeta+z}{\zeta-z}\left[\overline{\Phi_n(z^{-1})} - \overline{\Phi_n(\zeta)}\right] d\mu(\theta), \tag{9.2.15}$$

for $n \geq 1$. Clearly, both relations hold for orthonormal polynomials as well. There is another simple relation between Φ and Ψ:

$$\Psi_n^*(z)\Phi_n(z) + \Psi_n(z)\Phi_n^*(z) = 2z^n \prod_{j=0}^{n-1} \rho_j^2, \quad n \geq 1. \tag{9.2.16}$$

An important consequence is the following theorem.

Theorem 9.2.3 *Let μ be a nontrivial probability measure on $[-\pi,\pi]$ and $\mu^{(n)}$ be its Bernstein–Szegő approximations* (9.2.10). *Then for $z \in \mathbb{D}$,*

$$\frac{\Psi_n^*(z)}{\Phi_n^*(z)} = F(z,\mu^{(n)}) = \int_{-\pi}^{\pi} \frac{\zeta+z}{\zeta-z}\, d\mu^{(n)}(\theta), \quad \zeta = e^{i\theta}, \tag{9.2.17}$$

so

$$\lim_{n\to\infty} \frac{\Psi_n^*(z)}{\Phi_n^*(z)} = F(z,\mu) = \int_{-\pi}^{\pi} \frac{\zeta+z}{\zeta-z}\, d\mu(\theta) \tag{9.2.18}$$

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

$$\left| F(z,\mu) - \frac{\Psi_n^*(z)}{\Phi_n^*(z)} \right| = O(z^{n+1}), \quad z \to 0. \tag{9.2.19}$$

The function $F(\mu)$ in (9.2.18) is a **Carathéodory function** of μ. An explicit formula for the error in (9.2.19) is available:

$$F(z,\mu)\Phi_n^*(z) - \Psi_n^*(z) = z^n \int_{-\pi}^{\pi} \frac{\zeta+z}{\zeta-z}\overline{\Phi_n(\zeta)}\, d\mu(\theta). \tag{9.2.20}$$

Similarly to (9.2.19) we also have

$$F(z,\mu)\Phi_n(z) + \Psi_n(z) = 2\kappa_n^{-2} z^n + O(z^{n+1}), \quad z \to 0. \tag{9.2.21}$$

There is a converse to (9.2.19)/(9.2.21) due to Peherstorfer and Steinbauer (1995).

Theorem 9.2.4 *Given a nontrivial probability measure μ on $[-\pi, \pi]$ with the Carathéodory function $F(\mu)$, let p and q be polynomials of degree at most n with*

$$p(z)F(z,\mu) + q(z) = O(z^n), \quad p^*(z)F(z,\mu) - q^*(z) = O(z^{n+1}),$$

as $z \to 0$, where p^*, q^* are the reverse polynomials for degree n. Then $p = c\Phi_n(\mu)$, $q = c\Psi_n(\mu)$ for some constant c.

It turns out that the vector

$$\begin{pmatrix} \psi_n(z) \\ -\psi_n^*(z) \end{pmatrix}$$

provides a second linearly independent solution of the Szegő recurrence (9.2.9), and the Carathéodory function $F(\mu)$ has a property analogous to a defining property of the Weyl m-function in the case of differential equations (see Geronimo, 1992; Golinskii and Nevai, 2001).

Theorem 9.2.5 *For fixed $z \in \mathbb{D}$ a number $r = F(z,\mu)$ is the unique complex number so that*

$$\begin{pmatrix} \psi_n(z) \\ -\psi_n^*(z) \end{pmatrix} + r \begin{pmatrix} \phi_n(z) \\ \phi_n^*(z) \end{pmatrix} \in \ell^2\left(\mathbb{Z}_+, \mathbb{C}^2\right). \tag{9.2.22}$$

There is another important property of $F(\mu)$ related to mass points of the orthogonality measure, which follows directly from the definition (9.2.18):

$$\lim_{r \to 1-0} (1-r)F(r\zeta, \mu) = 2\mu\{\theta\} \quad \text{for all } \zeta = e^{i\theta} \in \mathbb{T}. \tag{9.2.23}$$

The **Christoffel kernels** (reproducing kernels)

$$K_n(z,w) = \sum_{j=0}^{n} \phi_j(z)\overline{\phi_j(w)} \tag{9.2.24}$$

arise with regard to the following extremal problem.

Theorem 9.2.6 *Let μ be a nontrivial probability measure on $[-\pi, \pi]$, and $\Pi_n(w)$ be a set of all polynomials P of degree at most n subject to $P(w) = 1$. Then*

$$\lambda_n(w) = \min_{P \in \Pi_n(w)} \int_{-\pi}^{\pi} |P(\zeta)|^2 \, d\mu(\theta) = \frac{1}{K_n(w,w)}. \tag{9.2.25}$$

The minimum is attained when $P(z) = K_n(z,w)/K_n(w,w)$.

Theorem 9.2.7 (Christoffel–Darboux formula) *For any $n \in \mathbb{Z}_+$ and $z, w \in \mathbb{C}$ with $\bar{w}z \neq 1$,*

$$K_n(z,w) = \frac{\phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(z)\overline{\phi_{n+1}(w)}}{1 - \bar{w}z} \tag{9.2.26}$$

$$= \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - z\bar{w}\phi_n(z)\overline{\phi_n(w)}}{1 - \bar{w}z}. \tag{9.2.27}$$

Here are some consequences of the Christoffel–Darboux formula. Setting $z = w$ we have

$$(1 - |z|^2) \sum_{j=0}^{n} |\phi_j(z)|^2 = |\phi_{n+1}^*(z)|^2 - |\phi_{n+1}(z)|^2. \tag{9.2.28}$$

Setting $w = 0$ we come to

$$K_n(z, 0) = \sum_{j=0}^{n} \phi_j(z)\overline{\phi_j(0)} = \phi_n^*(z)\overline{\phi_n^*(0)} = \kappa_n \phi_n^*(z). \tag{9.2.29}$$

The reproducing property of the Christoffel kernels is

$$\int_{-\pi}^{\pi} P(\zeta)\overline{K_n(\zeta, w)}\, d\mu(\theta) = P(w), \tag{9.2.30}$$

which holds for an arbitrary polynomial P of degree at most n and all complex w and follows directly from the definition. As a simple consequence of (9.2.30) one has a unit circle analogue of Theorem 2.1.8.

Theorem 9.2.8 *Let $M_n(\mu) = \|\mu_{i-j}\|_{i,j=0}^n$ be the Toeplitz matrix of the moments of μ. Let*

$$K_n(z, w) = \sum_{i,j=0}^{n} a_{ij} z^i \bar{w}^j$$

be the Taylor series expansion of the Christoffel kernel about the origin. Then

$$M_n^{-1}(\mu) = A^*, \quad A = \|a_{ij}\|_{i,j=0}^n.$$

One result connected to the Christoffel circle of ideas in the case of orthogonal polynomials on the real line is the Gauss–Jacobi quadrature formula. The following is its partial analogue for the unit circle case.

Theorem 9.2.9 *Suppose that the monic orthogonal polynomial $\Phi_n(\mu)$ has n distinct roots $z_1, \ldots, z_n$. Then for each Laurent polynomial π of the form*

$$\pi(z) = \sum_{j=-n+1}^{n} \pi_j z^j$$

there exist complex numbers $\beta_1, \ldots, \beta_n$ so that

$$\int_{-\pi}^{\pi} \pi(\zeta)\, d\mu(\theta) = \sum_{j=1}^{n} \beta_j \pi(z_j). \tag{9.2.31}$$

9.3 Szegő's Theory and Its Extensions

Szegő's theorems may well be considered the most celebrated in OPUC theory. They have repeatedly served as a source for further development. For historical reasons one should state them in terms of Toeplitz determinants, $D_n(\mu)$ (cf. (9.1.3)). It follows from (9.1.9) and (9.2.4) that

$$D_n(\mu) = \prod_{j=0}^{n} \|\Phi_j\|^2 = \prod_{j=0}^{n-1} (1 - |\alpha_j|^2)^{n-j},$$

so $D_n^{1/n}(\mu)$ is monotone decreasing and

$$S(\mu) = \lim_{n\to\infty} (D_n(\mu))^{1/n} = \lim_{n\to\infty} \frac{D_n(\mu)}{D_{n-1}(\mu)} \tag{9.3.1}$$

exists and is a nonnegative number. Suppose $S(\mu) > 0$. Then $D_n(\mu)/S^{n+1}(\mu)$ is monotone increasing and

$$G(\mu) = \lim_{n\to\infty} \frac{D_n(\mu)}{S^{n+1}(\mu)} \tag{9.3.2}$$

exists and may be equal $+\infty$. Also, $G(\mu) < +\infty$ if and only if $\sum_{j=0}^{\infty} j|\alpha_j|^2 < \infty$ (Ibragimov's condition).

Szegő's theorems express S and G in terms of the absolutely continuous and singular components of μ (cf. (9.1.1)).

Theorem 9.3.1 (Szegő's theorem)

$$S(\mu) = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2) = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \mu'(\theta)\, d\theta \right). \tag{9.3.3}$$

A striking feature of this result is that the α depend heavily on the singular component μ_s, whereas the product in (9.3.3) does not!

Szegő (1915) proved this for the case $\mu_s = 0$ in 1915 (in his very first paper!). The result does not depend on μ_s – this was shown by Verblunsky (1936).

It is immediate from Szegő's theorem that

$$\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \quad \text{if and only if} \quad \log \mu' \in L^1. \tag{9.3.4}$$

The equivalent conditions (9.3.4) are known as the **Szegő condition**, and the corresponding class of measures the **Szegő class**. Within this class, the **Szegő function**

$$D(z) = D(z,\mu) = \exp\left(\frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\zeta + z}{\zeta - z} \log \mu'(\theta)\, d\theta \right), \quad \zeta = e^{i\theta}, \quad |z| < 1, \tag{9.3.5}$$

which depends only on an absolutely continuous component of the orthogonality measure, is well defined. It is clear from the definition that $D(\mu)$ lies in the Hardy space $H^2(\mathbb{D})$, and the standard boundary value theory implies $D(\zeta) = \lim_{r\uparrow 1} D(rz)$ exists almost everywhere and $|D(\zeta)|^2 = \mu'(\theta)$ a.e., so $\|D\|_{H^2} \le 1$.

Theorem 9.3.2 (Strong Szegő theorem in Ibragimov's version) *If $\mu_s = 0$ and the Szegő condition holds, then*

$$G(\mu) = \prod_{j=0}^{\infty}(1-|\alpha_j|^2)^{-j-1} = \exp\left(\sum_{n=0}^{\infty} n|w_n|^2\right) = \exp\left(\int_{\mathbb{D}} \left|\frac{D'(z)}{D(z)}\right|^2 d^2z\right),$$

where w_n are the Fourier coefficients of $\log\mu'$, *and d^2z the normalized Lebesgue measure of $\mathbb{D}$. All the values may equal $+\infty$.*

For the modern approach to this result see Simon (2004a, Chapter 6).

Simon (2004a, Section 2.8) came up with the idea of extending Szegő's theorem by allowing "Pollaczek singularities" (so all quantities in (9.3.3) may be infinite). His result can be viewed as the first-order Szegő theorem: for any $\zeta_0 \in \mathbb{T}$,

$$|\zeta-\zeta_0|^2 \log\mu' \in L^1 \quad \text{if and only if} \quad \sum_{j=0}^{\infty} |\alpha_{j+1} - \overline{\zeta}_0\alpha_j|^2 + |\alpha_j|^4 < \infty.$$

Moreover, there is a precise formula for this case similar to the second equality in (9.3.3). The second-order Szegő theorem appeared in Simon and Zlatoš (2005).

Theorem 9.3.3 *Let $\zeta_k \in \mathbb{T}$, $k = 1, 2$. Then for $\zeta_1 \ne \zeta_2$,*

$$|\zeta-\zeta_1|^2|\zeta-\zeta_2|^2 \log\mu' \in L^1 \quad \textit{if and only if}$$
$$\sum_{j=0}^{\infty} |\alpha_{j+2} - (\overline{\zeta}_1 + \overline{\zeta}_2)\alpha_{j+1} + \overline{\zeta_1\zeta_2}\alpha_j|^2 + |\alpha_j|^4 < \infty,$$

and for $\zeta_1 = \zeta_2$,

$$|\zeta-\zeta_1|^4 \log\mu' \in L^1 \ \textit{if and only if}$$
$$\sum_{j=0}^{\infty} |\alpha_{j+2} - 2\overline{\zeta}_1\alpha_{j+1} + \overline{\zeta_1^2}\alpha_j|^2 + |\alpha_j|^6 < \infty.$$

The general conjecture called the **higher-order Szegő theorem** was formulated in Simon (2004a, Section 2.8). Given $\zeta_k \in \mathbb{T}$, $k = 1, \dots, n$ and $\zeta_p \ne \zeta_q$, $p \ne q$, define a polynomial

$$P(\zeta) := \prod_{k=1}^{n}(\zeta-\zeta_k)^{m_k}, \quad m_k \in \mathbb{N} = \{1, 2, \dots\}, \quad \overline{P}(\zeta) := \prod_{k=1}^{n}(\zeta-\overline{\zeta}_k)^{m_k},$$

and put $m = 1 + \max_k m_k$. Simon conjectures that

$$|P(\zeta)|^2 \log w \in L^1 \quad \text{if and only if} \quad \left(\overline{P}(S)\right)\{\alpha_j\} \in \ell^2 \text{ and } \{\alpha_j\} \in \ell^{2m},$$

where S is the shift operator: $S(\alpha_0, \alpha_1, \dots) = (\alpha_1, \alpha_2, \dots)$.

The following particular case of Simon's conjecture which can be called the **higher-order Szegő theorem in** ℓ^4 is proved in Golinskii and Zlatoš (2007).

Theorem 9.3.4 *Assume that* $\{\alpha_j\} \in \ell^4$. *Then*

$$|P(\zeta)|^2 \log \mu' \in L^1 \quad \textit{if and only if} \quad \left(\overline{P}(S)\right)\{\alpha_j\} \in \ell^2.$$

The further advances concerning the polynomial Szegő class, that is, the class of measures μ with $|P(\zeta)|^2 \log \mu' \in L^1$, in particular, the asymptotics inside the disk and L^2 asymptotics on the circle, are in Denisov and Kupin (2006).

The celebrated Szegő asymptotic formula is one of the cornerstones of OPUC theory.

Theorem 9.3.5 (Szegő's limit theorem) *Suppose the Szegő condition* (9.3.4) *holds. Then*

$$\lim_{n\to\infty} \phi_n(z) = 0, \qquad \lim_{n\to\infty} \phi_n^*(z) = D^{-1}(z,\mu) \tag{9.3.6}$$

uniformly on compact supports of the unit disk.

The result appeared in Szegő's pioneering 1920 paper (Szegő, 1920). Another closely related result concerns the asymptotics of the Christoffel kernels (see, for example, Grenander and Szegő, 1958, Chapter 3.4)

Theorem 9.3.6 *Under the Szegő condition* (9.3.4),

$$\lim_{n\to\infty} K_n(z,w) = \sum_{j=0}^{\infty} \phi_j(z)\overline{\phi_j(w)} = \frac{1}{1-\bar{w}z}\frac{1}{\overline{D(w,\mu)}}\frac{1}{D(z,\mu)} \tag{9.3.7}$$

uniformly on compact supports of the bidisk ($|z| < 1, |w| < 1$).

In particular,

$$\lambda_\infty(w) := \min\left(\int_{-\pi}^{\pi} |P(\zeta)|^2\, d\mu(\theta) \colon\, P(w) = 1\right) = (1-|w|^2)|D(w,\mu)|^2, \tag{9.3.8}$$

where the minimum is taken over the set of all polynomials P.

As for the asymptotics on the unit circle, we begin with L^2-convergence. The argument here uses a simple equality

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} &|\phi_n^*(\zeta) - D^{-1}(\zeta)|^2 \mu'(\theta)\, d\theta + \int_{-\pi}^{\pi} |\phi_n^*(\zeta)|^2\, d\mu_s \\ &= \int_{-\pi}^{\pi} |\phi_n^*(\zeta)|^2\, d\mu + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\mu'(\theta)}{|D(\zeta)|^2}\, d\theta - 2\Re D(0)\phi^*(0)\end{aligned}$$

(cf. Simon, 2004a, Section 2.4). Since the first two terms on the right-hand side are $1+1$, and from the Szegő limit theorem

$$\phi_n^*(0) = \kappa_n = \prod_{j=0}^{n-1}(1-|\alpha_j|^2), \quad D^{-1}(0) = \prod_{j=0}^{n-1}(1-|\alpha_j|^2),$$

it is not hard to obtain the following bound:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |\phi_n^*(\zeta) - D^{-1}(\zeta)|^2 \mu'(\theta)\, d\theta + \int_{-\pi}^{\pi} |\phi_n^*(\zeta)|^2 \, d\mu_s \leq 2\sum_{j=n}^{\infty} |\alpha_j|^2.$$

In particular, we have the following theorem.

Theorem 9.3.7 *Under the Szegő condition* (9.3.4),

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |\phi_n^*(\zeta) - D^{-1}(\zeta)|^2 \mu'(\theta)\, d\theta = \lim_{n\to\infty} \int_{-\pi}^{\pi} |\phi_n^*(\zeta)|^2 \, d\mu_s = 0. \tag{9.3.9}$$

There are L^2-convergence results of a slightly different type, which deal with the limit relation

$$\lim_{n\to\infty} \frac{1}{\phi_n^*(\zeta)} = D(\zeta), \quad \zeta \in \mathbb{T}. \tag{9.3.10}$$

Theorem 9.3.8 *Under the Szegő condition, the convergence in* (9.3.10) *holds in the weak topology of* $L^2(\mathbb{T})$*, that is,*

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\overline{g(\zeta)}}{\phi_n^*(\zeta)}\, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} D(\zeta)\overline{g(\zeta)}\, d\theta, \quad g \in L^2.$$

Furthermore,

$$\lim_{n\to\infty} \left\| D - 1/\phi_n^* \right\|_{L^2}^2 = \mu_s([-\pi,\pi]).$$

In particular, the convergence in (9.3.10) *is in the* L^2*-norm if and only if* μ *is absolutely continuous* ($\mu_s = 0$).

Khruschev (2001) proved the following nice limit relation that characterizes the Szegő class

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} \left| \log |\phi_n(\zeta)|^{-2} - \log \mu'(\zeta) \right| \, d\theta = 0.$$

As far as the Christoffel function on the circle goes, one can easily prove that for an arbitrary measure μ,

$$\lim_{n\to\infty} \lambda_n(\zeta) = \lim_{n\to\infty} \frac{1}{K_n(\zeta,\zeta)} = \left(\sum_{n=0}^{\infty} |\phi_n(\zeta)|^2 \right)^{-1} = \mu\{\theta\} \tag{9.3.11}$$

for all $\zeta = e^{i\theta} \in \mathbb{T}$. A much more delicate result is due to Máté, Nevai, and Totik (1991).

Theorem 9.3.9 *For an arbitrary measure* μ *from the Szegő class one has*

$$\lim_{n\to\infty} \frac{n}{K_n(\zeta,\zeta)} = \mu'(\theta), \qquad \zeta = e^{i\theta},$$

a.e. on $[-\pi,\pi]$.

There are two natural ways to proceed from the Szegő theory. The first one is to consider proper subclasses of the Szegő class, and refine the above asymptotic results by imposing additional assumptions on the orthogonality measure, Verblunsky coefficients, ..., etc. For instance, one may ask whether the basic formula (9.3.6) holds on the unit circle (uniformly, pointwise, almost everywhere, etc.). Here is the classical result due to Szegő (see Grenander and Szegő, 1958, Theorem 3.5), which gives a rate of convergence in (9.3.6) on $\mathbb{T}$.

Given a continuous function $g \in C(\mathbb{T})$, define its modulus of continuity by

$$\omega(\delta; g) = \sup\{|g(h\zeta) - g(\zeta)| : |h-1| \le \delta,\ h \in \mathbb{T}\}.$$

Theorem 9.3.10 *Let μ be absolutely continuous, $\mu = \frac{1}{2\pi}\mu'(\theta)\,d\theta$, with a positive and continuous density μ'. Assume also that*

$$\omega(\delta; \mu') \le C\left(\log\frac{1}{\delta}\right)^{-(1+\varepsilon)}, \quad \varepsilon > 0. \tag{9.3.12}$$

Then

$$\sup_{\mathbb{T}} |\phi_n^*(\zeta) - D^{-1}(\zeta)| \le C_1 (\log n)^{-\varepsilon}.$$

Assumption (9.3.12) was relaxed by B. Golinskii (Golinskii, 1967) to

$$\int_0^a \frac{\omega(t; \mu')}{t}\,dt < \infty,$$

and some rate of convergence was found in this case.

The uniform convergence in (9.3.6) in the closed unit disk can be guaranteed by certain assumptions on Verblunsky coefficients $\alpha_n(\mu)$ (see Geronimus, 1961, Chapter 8).

Theorem 9.3.11 *Let the Verblunsky coefficients α_n of the orthogonality measure μ satisfy*

$$\sum_{n=0}^{\infty} |\alpha_n(\mu)| < \infty. \tag{9.3.13}$$

Then

$$\max_{|z|\le 1} |\phi_n^*(z) - D^{-1}(z,\mu)| \le C \sum_{k=n}^{\infty} |\alpha_k(\mu)|, \quad n \in \mathbb{N}. \tag{9.3.14}$$

In the latter two results the property $\mu' > 0$ is crucial. Indeed, the asymptotic formula (9.3.6) cannot hold uniformly on $\mathbb{T}$ as long as μ' has zeros on $[-\pi, \pi]$. Badkov (1985) suggested a modified asymptotic formula, wherein $\phi_n^*(\zeta)$ is compared with $D^{-1}(r_n\zeta)$ with $r_n \uparrow 1$ as $n \to \infty$. More precisely, he proved that in a variety of situations, when μ' has algebraic zeros on $\mathbb{T}$, the limit relation

$$\phi_n^*(\zeta) = D^{-1}((1 - cn^{-1})\zeta)[1 + o(1)]$$

holds uniformly on $\mathbb{T}$.

Denote by $E(\mu)$ the subset of $[-\pi, \pi]$ on which $D^{-1}(\zeta) = \lim_{r\to 1-0} D^{-1}(r\zeta)$, $\zeta = e^{i\theta}$. Clearly, the normalized Lebesgue measure of this set is 1. The existence of the limit in (9.3.6) at ζ can be viewed as a **tauberian problem**: under what conditions does

$$\lim_{n\to\infty} \phi_n^*(\zeta) = \lim_{n\to\infty} \lim_{r\to 1-0} \phi_n^*(r\zeta) = \lim_{r\to 1-0} \lim_{n\to\infty} \phi_n^*(r\zeta) = D^{-1}(\zeta)? \tag{9.3.15}$$

The following result is due to Geronimus (1961, Theorem 5.1). Put

$$\delta_n(\mu) = \sum_{k=n}^{\infty} |\alpha_k(\mu)|^2. \tag{9.3.16}$$

Theorem 9.3.12 *Assume that $\mu' \geq c > 0$ a.e., and $\delta_n(\mu) = o(n^{-1})$. Then on the whole unit circle,*

$$|\phi_n^*(\zeta) - D^{-1}(\zeta)| \leq |D^{-1}(\zeta) - D^{-1}(r_n\zeta)| + C(n\delta_n)^{1/3}, \quad r_n = 1 - \left(\frac{\delta_n}{n}\right)^{2/3}.$$

In particular, (9.3.15) *holds on $E(\mu)$.*

Let us now turn to the boundedness of OPUC. The simplest bound comes out of the Szegő recurrences (9.2.1)/(9.2.3).

Theorem 9.3.13 *For $\zeta \in \mathbb{T}$ one has*

$$\prod_{j=0}^{n-1}(1 - |\alpha_j|) \leq |\Phi_n(\zeta)| \leq \prod_{j=0}^{n-1}(1 + |\alpha_j|). \tag{9.3.17}$$

Moreover,

$$\sup_{|z|\leq 1} |\Phi_n(z)| \leq \exp\left(\sum_{j=0}^{n-1} |\alpha_j|\right), \tag{9.3.18}$$

so under assumption (9.3.13),

$$\sup_n \sup_{|z|\leq 1} |\Phi_n(z)| < \infty, \tag{9.3.19}$$

that is, the system Φ_n is uniformly bounded in the closed unit disk. If $B = \sup_j |\alpha_j| < 1$, then

$$\inf_{\zeta\in\mathbb{T}} |\Phi_n(\zeta)| \geq \exp\left(\sum_{j=0}^{n-1} |\alpha_j| - \frac{1}{2(1-B)^2} \sum_{j=0}^{n-1} |\alpha_j|^2\right). \tag{9.3.20}$$

One can relax (9.3.13) and get some divergent (but still useful) estimates. For instance,

$$\sum_{n=0}^{\infty} n|\alpha_n|^2 < \infty \quad \text{implies} \quad \left\|(\Phi_n^*)^{\pm 1}\right\|_\infty \leq C \exp(D\sqrt{\log n}),$$

where $\|\cdot\|_\infty$ is the L^∞-norm on $\mathbb{T}$, and C and D are suitable constants.

The uniform boundedness (9.3.19) holds under certain assumptions on the measure μ.

Theorem 9.3.14 *Let μ be absolutely continuous, and one of the following holds:*

(i) $0 < c \le \mu' \le C < \infty$ *a.e., and the moments* $\mu_n = O(1/n)$, $n \to \infty$*;*
(ii) $0 < c \le \mu'$ *a.e., and* μ' *is of bounded variation.*

Then Φ_n *is uniformly bounded in the closed unit disk.*

The study of the uniform boundedness (9.3.19) has a long history going back at least to Geronimus (1961). The uniform boundedness and uniform asymptotic representation for orthogonal polynomials is discussed in Golinskii and Golinskii (1998).

It was a conjecture of Steklov that the only condition $\mu' \ge c > 0$ yields the uniform boundedness of orthogonal polynomials. This was proven false by Rakhmanov (1980, 1982a), who showed that for any $\delta < \frac{1}{2}$, there are examples with $\limsup_{n\to\infty} |\phi_n(1)| n^{-\delta} = \infty$. On the other hand,

$$\mu' \ge c > 0 \text{ a.e.} \quad \text{implies} \quad \|\phi_n\|_\infty \le c^{-1/2}\sqrt{n+1}\,. \tag{9.3.21}$$

Statement (9.3.21) is sometimes called the Szegő estimate.

The second way to proceed from the Szegő theory is to go beyond the Szegő class and extend the above results to the case when the Szegő condition fails. The following three notable classes of measures, each of which contains the Szegő class as a proper subclass, come in naturally. These are

the **Erdös** class $\mathcal{E}$ of measures μ with $\mu' > 0$ a.e.;

the **Nevai** class $\mathcal{N}$ of measures μ with $\lim_{n\to\infty} \alpha_n(\mu) = 0$;

the **Rakhmanov** class $\mathcal{R}$ of measures μ with

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} f(\zeta)|\phi_n(\zeta)|^2 \, d\mu(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\zeta)\, d\theta, \qquad \zeta = e^{i\theta}, \qquad \text{for all } f \in C(\mathbb{T}); \tag{9.3.22}$$

in other words, $|\phi_n|^2\, d\mu \to dm$ in the *-weak topology of the space of measures.

As it turns out, each class contains the former one as a proper subclass.

The first two classes were characterized by Máté, Nevai, and Totik (1985, 1987b) (see also Nevai, 1991) by means of the quantity

$$b_{n,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \left| \frac{\phi_n(\zeta)}{\phi_{n+k}(\zeta)} \right|^2 - 1 \right| d\theta.$$

Their argument is based on the relation, which holds for arbitrary nontrivial probability measure μ,

$$\lim_{n\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| |\phi_n(\zeta)|^2 \mu'(\theta) - 1 \right|^2 d\theta \le \limsup_{k\to\infty} b_{n,k}.$$

Theorem 9.3.15 *The following are equivalent:*

(i) $\mu \in \mathcal{E}$*;*
(ii) $\lim_{n\to\infty} \sup_{k\ge 1} b_{n,k} = 0$.

Theorem 9.3.16 *The following are equivalent:*

(i) $\mu \in \mathcal{N}$;
(ii) $\lim_{n\to\infty} \inf_{k\geq 1} b_{n,k} = 0$.

As a simple consequence one has $\mathcal{E} \subset \mathcal{N}$ (Rakhmanov's theorem). Moreover, a quantitative version of this relation was proved by Denisov (2004):

$$\limsup_{n\to\infty} |\alpha_n(\mu)| \leq 2\sqrt{2}\left[1 - m^2(\Omega)\right]^{1/2}, \tag{9.3.23}$$

where $\Omega = \{\theta \in [-\pi, \pi] \colon \mu'(\theta) > 0\}$, and $m(\Omega)$ is its normalized Lebesgue measure.

The next result is also due to Máté, Nevai, and Totik (1987b).

Theorem 9.3.17 *Let $\mu \in \mathcal{E}$. Then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} \left||\phi_n(\zeta)|^2 \mu'(\zeta) - 1\right| d\theta = 0. \tag{9.3.24}$$

Moreover,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} \left(|\phi_n(\zeta)|^{-1} - \sqrt{\mu'(\theta)}\right)^2 d\theta = 0$$

if and only if μ is absolutely continuous.

Later on, Khruschev (2001) showed that (9.3.24) in fact characterizes the class $\mathcal{E}$, and suggested another characteristic property, namely

$$\lim_{n\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|\phi_n(\zeta)|^2 \mu'(\theta)\right]^a d\theta = 1 \quad \text{for all } 0 < a < 1.$$

There is another characterization of the Nevai class (Máté, Nevai, and Totik, 1987a).

Theorem 9.3.18 *Let $\mu \in \mathcal{N}$. Then*

$$\lim_{n\to\infty} \max_{|z|\leq 1} \frac{|\phi_n(z)|^2}{\sum_{k=0}^{n} |\phi_k(z)|^2} = 0.$$

Moreover, if the latter relation holds at least at one point $z_0 \in \mathbb{D}$ then $\mu \in \mathcal{N}$.

One of the most important results due to Máté, Nevai, and Totik (1987a) is the so-called **comparative** asymptotics outside the Szegő class.

Theorem 9.3.19 *Let $\mu \in \mathcal{E}$. Suppose*

$$\nu = g\mu, \quad g \geq 0, \quad \int g\, d\mu = 1,$$

and there is a polynomial Q so that $Qg^{\pm 1} \in L^\infty(\mu)$. Then uniformly for z, w in compact subsets of $\mathbb{D}$ we have

(1) $\lim_{n\to\infty} \dfrac{\phi_n^*(z,\nu)}{\phi_n^*(z,\mu)} = \exp\left(-\dfrac{1}{4\pi} \int_{-\pi}^{\pi} \dfrac{\zeta+z}{\zeta-z} \log g(\zeta)\, d\theta \right) = D(z, g^{-1});$

(2) $\lim_{n\to\infty} \dfrac{K_n^*(z,w;\nu)}{K_n^*(z,w;\mu)} = \overline{D(w,g^{-1})} D(z,g^{-1});$

(3) $\lim_{n\to\infty} \dfrac{\kappa_n(\nu)}{\kappa_n(\mu)} = \exp\left(-\dfrac{1}{4\pi} \int_{-\pi}^{\pi} \log g(\zeta)\, d\theta \right).$

Moreover,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} \left| \phi_n(\zeta;\nu)\overline{D(\zeta,g)} - \phi_n(\zeta;\mu) \right|^2 \mu'(\theta)\, d\theta = 0.$$

We now come to Rakhmanov's class, and give a characterization due to Khruschev (2001). We say that a sequence of Verblunsky coefficients obeys the **Máté–Nevai condition** if for each fixed $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \alpha_n(\mu)\alpha_{n+k}(\mu) = 0. \tag{9.3.25}$$

Let us also introduce the probability measures

$$d\nu_{n,k} = \frac{1}{2\pi} \left[\frac{|\phi_n(\zeta)|^2}{|\phi_{n+k}(\zeta)|^2} \right] d\theta. \tag{9.3.26}$$

Theorem 9.3.20 *The following are equivalent:*

(i) $\mu \in \mathcal{R}$;
(ii) *the Máté–Nevai condition holds for* $\alpha_n(\mu)$;
(iii) $\nu_{n,k}$ *converges weakly to the normalized Lebesgue measure as* $n \to \infty$ *for all* $k \in \mathbb{N}$;
(iv) *uniformly on compact subsets of* $\mathbb{D}$, *we have*

$$\lim_{n\to\infty} \frac{\Phi_{n+1}^*(z)}{\Phi_n^*(z)} = 1. \tag{9.3.27}$$

It is obvious from this theorem that $\mathcal{N} \subset \mathcal{R}$. It is also easy to manufacture examples of measures off the Nevai class which obey (9.3.25). Indeed, these are measures with sparse Verblunsky coefficients. Furthermore, the Rakhmanov measures which do not belong to the Nevai class are necessarily singular (Khruschev, 2001, Corollary 2.6).

Relation (9.3.27) is known as the **ratio asymptotics** of OPUC. As a matter of fact, there is a way to describe all possible limits for the ratio in (9.3.27). The result below is due to Khruschev (2002) and Barrios and López (1999).

Theorem 9.3.21 *Suppose*

$$\lim_{n\to\infty} \frac{\Phi_{n+1}^*(z,\mu)}{\Phi_n^*(z,\mu)} = G(z) \tag{9.3.28}$$

exists uniformly on compact subsets of $\mathbb{D}$. *Then either* $G \equiv 1$ *or*

$$G(z) = G_{a,\lambda}(z) = \frac{1 + \lambda z + \sqrt{(1 - \lambda z)^2 + 4a^2\lambda z}}{2}$$

hold for some $\lambda \in \mathbb{T}$ *and* $a \in (0, 1]$. *Equation* (9.3.28) *holds with* $G = G_{a,\lambda}$ *if and only if* $\alpha_n(\mu)$ *obeys the López condition*

$$\lim_{n\to\infty} |\alpha_n| = a, \quad \lim_{n\to\infty} \frac{\alpha_{n+1}}{\alpha_n} = \lambda. \tag{9.3.29}$$

In this case the essential support of μ *is a closed subinterval of* $[-\pi, \pi]$*, and* (9.3.28) *holds uniformly on compact subsets of* $\mathbb{C}\setminus \exp\{\operatorname{supp}\mu\}$.

The following extension of the above result, which can be viewed as the **relative ratio asymptotics**, is in Golinskii and Zlatoš (2007).

Theorem 9.3.22 *Let* μ *and* ν *be two nontrivial probability measures on* $\mathbb{T}$. *Let* $\{\alpha_n(\mu)\}$ *and* $\{\alpha_n(\nu)\}$*, respectively, be their Verblunsky coefficients, and let* $\Phi_n^*(\mu)$ *and* $\Phi_n^*(\nu)$*, respectively, be their reverse monic orthogonal polynomials. Then*

$$\frac{\Phi_{n+1}^*(z;\mu)}{\Phi_n^*(z;\mu)} - \frac{\Phi_{n+1}^*(z;\nu)}{\Phi_n^*(z;\nu)} \to 0 \tag{9.3.30}$$

uniformly on compact subsets of $\mathbb{D}$ *as* $n \to \infty$ *if and only if for any* $k \in \mathbb{N}$,

$$\lim_{n\to\infty} [\alpha_n(\mu)\bar{\alpha}_{n-k}(\mu) - \alpha_n(\nu)\bar{\alpha}_{n-k}(\nu)] = 0. \tag{9.3.31}$$

A closely related subject is the description of all possible limits in the Rakhmanov condition (9.3.22). A comprehensive study of this problem is in Khruschev (2002).

We conclude with a theorem of Bello and López (1998), which is analogous to Rakhmanov's theorem ($\mathcal{E} \subset \mathcal{N}$) but for any arc. Define

$$\theta_a = 2\arcsin(a), \quad 0 < a < 1, \tag{9.3.32}$$

so $0 < \theta_a < \pi$. For $\lambda \in \mathbb{T}$ we let

$$\Gamma_{a,\lambda} = \{\zeta \in \mathbb{T} \colon |\arg(\lambda\zeta)| > \theta_a\}. \tag{9.3.33}$$

Theorem 9.3.23 *Let* $\exp\{\operatorname{supp}(\mu)\} = \Gamma_{a,\lambda}$ *and* $\mu' > 0$ *a.e. Then*

$$\lim_{n\to\infty} |\alpha_n(\mu)| = a, \quad \lim_{n\to\infty} \bar{\alpha}_{n+1}(\mu)\alpha_n(\mu) = a\lambda^2. \tag{9.3.34}$$

An essential extension of this result due to Simon (2004b, Theorem 13.4.4) claims that supp can be relaxed to ess supp.

9.4 Zeros of OPUC

The structure of the zero sets for OPUC is another fascinating topic of the theory. Given a nontrivial probability measure μ, denote by $Z_n(\mu) = \{z_{jn}\}_{j=1}^n$ the zero set for the monic orthogonal polynomial $\Phi_n(\mu)$:

$$\Phi_n(z,\mu) = \prod_{j=1}^n (z - z_{jn}), \quad |z_{nn}| \le |z_{n-1,n}| \le \cdots \le |z_{1,n}|.$$

As we have already seen, $Z_n(\mu) \subset \mathbb{D}$. Conversely, the following analogue of Wendroff's theorem was proved by Geronimus (1946).

Theorem 9.4.1 *Let π_n be any monic polynomial of degree n which has all its zeros inside $\mathbb{D}$. Then $\pi_n = \Phi_n(\mu)$ is a monic orthogonal polynomial for some $\mu \in \mathcal{P}$. Moreover, if μ and ν are any two such measures, we have*

(i) $\Phi_j(\mu) = \Phi_j(\nu)$, $j = 0, 1, \dots, n$;
(ii) $\alpha_j(\mu) = \alpha_j(\nu)$, $j = 0, 1, \dots, n-1$;
(iii) $\mu_j(\mu) = \mu_j(\nu)$, $j = 0, 1, \dots, n$.

It is clear that, unlike the case of orthogonal polynomials on the real line, the zeros need not be simple. The free case ($d\mu = dm = \frac{d\theta}{2\pi}$, the normalized Lebesgue measure on $[-\pi, \pi]$), where $\Phi_j(z, dm) = z^j$, illustrates this situation. Note also that the explicit measure in Theorem 9.4.1 can be easily constructed. Namely, the Bernstein–Szegő measure

$$d\sigma = \frac{C}{|\pi_n(\zeta)|^2}\, d\theta$$

is one, which satisfies $\alpha_j(\sigma) = 0$, $j \ge n$.

The fact that $Z_n(\mu) \subset \mathbb{D}$ reflects the following quite general situation (Fejér, 1922).

Theorem 9.4.2 (Fejér's theorem) *Let μ be a nontrivial probability measure on $\mathbb{C}$ so that*

$$\int |z|^k \, d\mu(z) < \infty, \quad k = 0, 1, \dots, 2n-1.$$

Let Φ_n be the monic polynomial of degree n orthogonal to $\{1, \dots, z^{n-1}\}$ in $L^2(\mathbb{C}, \mu)$. Then all of the zeros of Φ_n lie in the convex hull of supp(μ). *Suppose further that* supp(μ) *is compact. Then no extreme point of the hull is a zero, and if support does not lie in the straight line, all zeros lie in the interior of the convex hull.*

Fejér's theorem is optimal in the following sense. For the unit circle, $\Phi_1(z) = z - \bar{\alpha}_0 = z - \bar{\mu}_1$ has its zero at $\bar{\mu}_1$. But $\int_K \zeta \, d\mu$ runs through a dense set of the convex hull of K as μ runs through all probability measures on K.

If supp$(\mu) \subset \mathbb{T}$, the interior of the convex hull is a subset of $\mathbb{D}$, so the zeros of Φ_n lie in $\mathbb{D}$. If supp(μ) is a proper subset of $\mathbb{T}$, then Fejér's theorem gives more information. For example, if $\zeta_0 \in \mathbb{T}$ and $d = \mathrm{dist}(\zeta_0, \mathrm{supp}(\mu)) > 0$, then a little geometry shows that the distance of zeros of Φ_n from ζ_0 is at least $d^2/2$.

Here is the result by Denisov–Simon (Simon, 2004a, Theorem 1.7.20) which provides some information about the zeros near isolated points of support.

Theorem 9.4.3 *Let μ and Φ_n be as in the above theorem, and ζ_0 be an isolated point of* $\mathrm{supp}(\mu)$. *Let* $\Gamma = \mathrm{supp}(\mu)\backslash\{\zeta_0\}$ *and* $\mathrm{ch}(\Gamma)$ *its convex hull. Suppose*

$$\delta = \mathrm{dist}(\zeta_0, \mathrm{ch}(\Gamma)) > 0.$$

Then Φ_n has at most one zero in $\{|z - \zeta_0| < \delta/3\}$. In particular, if μ is supported on $\mathbb{T}$ and $d = \mathrm{dist}(\zeta_0, \Gamma) > 0$*, there is at most one zero in the circle of radius $d^2/6$ about ζ_0.*

If μ is an even measure with support $\{0\} \cup [1, 2] \cup [-2, -1]$, one can show that for n large enough and even, P_n has two zeros near 0. Thus, it is not enough that ζ_0 be isolated from Γ; it must be isolated from $\mathrm{ch}(\Gamma)$.

If $\mathrm{supp}(\mu) = \mathbb{T}$, the zeros of Φ_n may stay away from the support (take, for example, $d\mu = dm$). But when this set is a proper subset of the unit circle, it attracts zeros in the following sense (see Simon, 2004a, Theorems 8.1.11 and 8.1.12).

Theorem 9.4.4 *Suppose ζ_0 is an isolated point of* $\mathrm{supp}(\mu)$. *Then for any $\delta > 0$, there is N_δ so $\{|z-\zeta_0| < \delta\}$ has exactly one zero of Φ_n for $n > N_\delta$. If this zero is called z_n, there is an $a > 0$ so for all large enough n, $|z_n - \zeta_0| \le e^{-an}$.*

Theorem 9.4.5 *Let* $\mathrm{supp}(\mu) \neq \mathbb{T}$, *and ζ_0 be a nonisolated point of* $\mathrm{supp}(\mu)$. *Then for each $\delta > 0$,*

$$\sharp\{z\colon |z - \zeta_0| < \delta,\ \Phi_n(z) = 0\} \to \infty, \quad n \to \infty.$$

The following question arises naturally: Is possible that a bulk of zeros still stay away from the support in the latter case? A negative answer was given by Widom (1967).

Theorem 9.4.6 (Widom's zero theorem) *Let* $\mathrm{supp}\,\mu \neq \mathbb{T}$. *Then, for any compact set $K \subset \mathbb{D}$, there is a positive integer n_K, so that for each $j \in \mathbb{N}$,*

$$\sharp\{z\colon z \in K,\ \Phi_j(z) = 0\} \le n_K.$$

Here is another theorem on zeros of OPUC, which appeared in Alfaro and Vigil (1988).

Theorem 9.4.7 (Alfaro–Vigil theorem) *Let $\{z_n\}_{n=1}^\infty$ be a sequence of numbers in $\mathbb{D}$. Then there exists a unique nontrivial probability measure μ on $[-\pi, \pi]$ with $\Phi_n(z_n, \mu) = 0$.*

Alfaro and Vigil were answering the following question from P. Turán (Turán, 1980): Can the set $Z_\infty(\mu) = \cup_n Z_n(\mu)$ of all zeros of the Φ_n be dense in $\mathbb{D}$? The answer is clearly yes, and follows from this theorem. Such measures are called Turán measures. It is proved in Khruschev (2003) that there are absolutely continuous Turán measures with μ' a C^∞ function. This is especially interesting since if μ' is real analytic and nonvanishing then $\overline{Z_\infty(\mu)} \neq \mathbb{D}$ (see below).

It is known (Saff and Totik, 1992) that zeros of $\Phi_n(\mu)$ cluster to $\mathrm{supp}(\mu)$ as long as this set is a proper subset of the whole circle. The situation changes dramatically if $\mathrm{supp}(\mu) = \mathbb{T}$ (see, for example, $d\mu = dm$). By the Alfaro–Vigil theorem, zeros of Φ_n can cluster to all points of $\overline{\mathbb{D}}$. Denote by

$$Z_w(\mu) := \{z \in \overline{\mathbb{D}}:\ \liminf_{n\to\infty} \operatorname{dist}(z, Z_n) = 0\}$$

the set of limit points for the zeros of all Φ_n (weakly attracting points). Let $Z_w = \{Z_w(\mu)\}_\mu$ be the class of all such point sets. So $\overline{\mathbb{D}} \in Z_w$. It turns out that Z_w is rich enough. More precisely, each compact subset K of $\mathbb{D}$ belongs to Z_w, and the same is true if $K \supset \mathbb{T}$ (Simon and Totik, 2005, Theorem 4). On the other hand, $K = [1/2, 1]$ is not in Z_w.

Similarly, denote by

$$Z_s(\mu) := \{z \in \overline{\mathbb{D}}:\ \lim_{n\to\infty} \operatorname{dist}(z, Z_n) = 0\}$$

the point set of strongly attracting points, and Z_s the class of all such point sets. The structure of the latter is quite different from that of Z_w. For instance, it is proved in Alfaro et al. (2005) that if $0 \in Z_s(\mu)$ for some measure μ, then $Z_s(\mu)$ is at most a countable set which converges to the origin. So the disk $\{|z| \le 1/2\}$ is not in Z_s.

A significant generalization of the Alfaro–Vigil theorem is due to Simon and Totik (2005).

Theorem 9.4.8 *For an arbitrary sequence of points* $\{z_k\}$ *in* $\mathbb{D}$ *and an arbitrary sequence of positive integers* $0 < m_1 < m_2 < \cdots$ *, there exists a measure* μ *on* $[-\pi, \pi]$ *such that*

$$\Phi_{m_k}(z_j, \mu) = 0, \quad j = m_{k-1} + 1, \ldots, m_k.$$

The following consequence of this result is surprising. Given a measure μ, consider the sequence $\{\nu_n(\mu)\}_{n\ge1}$ of **normalized counting measures for zeros** of Φ_n, that is,

$$\operatorname{supp} \nu_n = Z_n, \quad \nu_n\{z_{jn}\} = \frac{l(z_{jn})}{n} \tag{9.4.1}$$

with $l(z_{jn})$ equal to the multiplicity of the zero z_{jn}. Let $\mathcal{M}_+(\bar{\mathbb{D}})$ be a space of probability measures on $\bar{\mathbb{D}}$ endowed with the weak* topology. A measure μ is called **universal** if for each $\sigma \in \mathcal{M}_+(\bar{\mathbb{D}})$ there is a sequence of indices n_j such that $\nu_{n_j}(\mu)$ converges to σ as $j \to \infty$ in the weak* topology. The existence of universal measures is proved in Simon and Totik (2005, Corollary 3).

A remarkable theorem of Mhaskar and Saff (1990) provides some information about the limit points (in the space $\mathcal{M}_+(\bar{\mathbb{D}})$) of the sequence of counting measures of zeros associated with a measure μ in the case when Verblunsky coefficients tend to zero fast enough.

Theorem 9.4.9 (Mhaskar–Saff theorem) *Let*

$$A := \limsup_{n\to\infty} |\alpha_n(\mu)|^{\frac{1}{n}} = \lim_{j\to\infty} |\alpha_{n_j}(\mu)|^{\frac{1}{n_j}}. \tag{9.4.2}$$

Suppose that either $A < 1$, *or* $A = 1$ *and* $\sum_{j=0}^{n-1} |\alpha_j(\mu)| = o(n)$ *as* $n \to \infty$. *Then* $\{\nu_{n_j}(\mu)\}$ *converges to the uniform measure on the circle of radius* A.

A crucial feature of the Mhaskar–Saff theorem is its universality. Under its assumption the angular distribution is the same. To get certain quantitative bounds on the distance between zeros, Simon studied various more stringent conditions, and among them the so-called **Barrios–López–Saff** condition

$$\alpha_n(\mu) = Cb^n + O((b\Delta)^n), \quad C \in \mathbb{C}\backslash\{0\}, \ \ 0 < b, \Delta < 1. \tag{9.4.3}$$

The following result is proved in Simon (2006).

Theorem 9.4.10 *Under assumption* (9.4.3) *there is a bounded number J of "spurious" zeros of $\Phi_n(\mu)$ for all large n. Furthermore, for $j = 1, 2, \dots, n - J$ let*

$$z_{jn} = |z_{jn}|e^{i\Theta_{jn}}, \quad 0 = \Theta_{0n} < \Theta_{1n} < \cdots < \Theta_{n-J,n} < 2\pi = \Theta_{n-J+1,n}$$

be the other zeros. Then the following limit relations hold:

$$\sup_{1\le j\le n-J} \Big| |z_{jn}| - b \Big| = O\left(\frac{\log n}{n}\right), \quad n \to \infty; \tag{9.4.4}$$

$$\sup_{1\le j\le n-J} n \left| \Theta_{j+1,n} - \Theta_{jn} - \frac{2\pi}{n} \right| = o(1), \quad n \to \infty; \tag{9.4.5}$$

$$\frac{|z_{j+1,n}|}{|z_{jn}|} = 1 + O\left(\frac{1}{n\log n}\right), \quad n \to \infty. \tag{9.4.6}$$

Note that (9.4.5)–(9.4.6) imply $\lim_n n|z_{j+1,n} - z_{jn}| = 2\pi b$. Amazingly, the spurious zeros also follow the clock pattern!

Simon (2005a) treats the more general case

$$\alpha_n(\mu) = \sum_{k=1}^{m} C_k e^{in\Theta_k} b^n + O((b\Delta)^n).$$

The value A (9.4.2) is tightly related to some other characteristics in OPUC theory. Define the following "radii":

- $R(D^{-1})$ is the radius of convergence of the Taylor series for the inverse Szegő function D^{-1} about the origin, if $\mu_s = 0$ and the Szegő condition holds, and $R(D^{-1}) = 1$ otherwise;
- $R^* = \sup\{r\colon \ \sup_{n,|z|\le r} |\Phi^*(z,\mu)| < \infty\}$, if the Szegő condition holds, and $R^* = 1$ otherwise.

Let $N_n(r)$ be a number of zeros of $\Phi_n(\mu)$ in $\{r < |z| < 1\}$. Define the Nevai–Totik radius R_{NT} by

$$R_{\mathrm{NT}} = \inf\{r\colon \ N_n(r) = O(1), \ n \to \infty\}.$$

The next result is proved in Nevai and Totik (1989).

Theorem 9.4.11 (Nevai–Totik theorem) *For an arbitrary measure μ the following equalities hold:*

$$A = \limsup_{n\to\infty} |\alpha_n(\mu)|^{\frac{1}{n}} = R_{\mathrm{NT}} = \frac{1}{R(D^{-1})} = \frac{1}{R^*}. \tag{9.4.7}$$

If $A < 1$, then $\phi_n^* \to D^{-1}$ uniformly on compact subsets of $\{z\colon \ |z| < A^{-1}\}$.

9.5 CMV Matrices – Unitary Analogues of Jacobi Matrices

One of the key tools in the case of the real line, especially in perturbation theory, is the realization of a measure σ as the spectral measure of the Jacobi matrix, which comes in as a matrix of multiplication by x on $L^2_\sigma(\mathbb{R})$. Of course, in the OPUC case, μ is the spectral measure of multiplication by ζ on $L^2_\mu(\mathbb{T})$. That alone is not enough because $L^2_\mu(\mathbb{T})$ is μ-dependent, and we cannot connect different μ. What we need is a suitable matrix representation; in other words, we need to choose a convenient orthonormal basis. There is an "obvious" set to try, namely, $\{\phi_n(\mu)\}$, but the corresponding matrix, called the GGT matrix in Simon (2004a), has two defects. First, a fundamental theorem by Szegő–Kolmogorov–Krein states that $\{\phi_n(\mu)\}$ is a basis (complete, orthonormal system) if and only if μ is outside the Szegő class, that is, $\log\mu' \notin L^1$, or equivalently, $\sum_{n=0}^\infty |\alpha_n(\mu)|^2 = \infty$, and if it is not, the matrix $\mathcal{G} = \|(\zeta\phi_m, \phi_n)\|$ is not unitary. Second, even if it is, the matrix $\mathcal{G}$ is not of finite width measured from the diagonal.

One of the most interesting developments in the theory of OPUC in recent years is the discovery by Cantero, Moral, and Velázquez (2003) of a matrix realization for multiplication by $\zeta = e^{i\theta}$ on $L^2_\mu(\mathbb{T})$ which is of finite band size (that is, $(\zeta\chi_m, \chi_n)_\mu = 0$ if $|m - n| > k$ for some k); in this case, $k = 2$ is to be compared with $k = 1$ for the Jacobi matrices, which correspond to the real line case. The **CMV basis** $\{\chi_n\}$ is obtained by orthonormalizing the sequence $1, \zeta, \zeta^{-1}, \zeta^2, \zeta^{-2}, \ldots$, and the matrix

$$C(\mu) = \|c_{n,m}\|_{m,n=0}^\infty, \quad c_{n,m} = (\zeta\chi_m, \chi_n)_\mu,$$

called the **CMV matrix**, is unitary and pentadiagonal. Remarkably, the χ can be expressed in terms of ϕ and ϕ^* by

$$\chi_{2n}(z) = z^{-n}\phi^*_{2n}(z), \quad \chi_{2n+1}(z) = z^{-n}\phi_{2n+1}(z), \quad n \in \mathbb{Z}_+,$$

and the matrix entries in terms of α and ρ:

$$C = \mathcal{L}\mathcal{M}, \tag{9.5.1}$$

where $\mathcal{L}$, $\mathcal{M}$ are 2×2 block diagonal matrices

$$\mathcal{L} = \Theta_0 \oplus \Theta_2 \oplus \Theta_4 \oplus \cdots, \quad \mathcal{M} = 1 \oplus \Theta_1 \oplus \Theta_3 \oplus \cdots, \tag{9.5.2}$$

with

$$\Theta_j = \begin{pmatrix} \overline{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{pmatrix}, \quad j \in \mathbb{Z}_+ \ldots \tag{9.5.3}$$

(the first block of $\mathcal{M}$ is 1×1). By C_0 we will denote the CMV matrix for the Lebesgue measure $\frac{1}{2\pi}\, d\theta$. For an exhaustive exposition of the theory of CMV matrices see Simon (2004a, Chapter 4) and Simon (2007a).

Expanding out the matrix product (9.5.1)–(9.5.3) is rather laborious and leads to a quite rigid structure

$$C(\mu) = \begin{pmatrix} * & * & + & & & & & \\ + & * & * & & & & & \\ & * & * & * & + & & & \\ & + & * & * & * & & & \\ & & & * & * & * & + & \cdots \\ & & & + & * & * & * & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \tag{9.5.4}$$

where $+$ represents strictly positive entries, and $*$ generally nonzero ones. The entries marked $+$ and called the exposed entries of the CMV matrix are precisely $(2,1)$ and those of the form $(2j-1, 2j+1)$ and $(2j+2, 2j)$ with $j \in \mathbb{N}$. Matrices of the form (9.5.4) are said to have CMV shape. Naturally, CMV matrices have CMV shape, and, what is more to the point, any unitary matrix (9.5.4) is actually a CMV matrix (9.5.1)–(9.5.3). Matrices (9.5.4) (of zigzag pentadiagonal form) appeared first in Watkins (1993), who outlined the connection of such matrices with OPUC.

Yet, expanding out (9.5.1)–(9.5.3) can be carried out, and explicit formulas for the matrix entries c_{nm} in terms of the Verblunsky coefficients are available (cf. Golinskiĭ, 2006). Let $2\lambda_m := 1 - (-1)^m$, $m \in \mathbb{Z}_+$, and $\lambda_{-1} = 1$, so

$$\{\lambda_m\}_{m\ge 0} = \{0, 1, 0, 1, \dots\},$$

$$\lambda_m + \lambda_{m+1} = 1, \quad \lambda_m\lambda_{m+1} = 0, \quad \lambda_m - \lambda_{m+1} = (-1)^{m+1}.$$

One has

$$\begin{gathered} c_{mm} = -\overline{\alpha}_m \alpha_{m-1}, \\ c_{m+2,m} = \rho_m\rho_{m+1}\lambda_m, \qquad c_{m,m+2} = \rho_m\rho_{m+1}\lambda_{m+1}, \end{gathered} \tag{9.5.5}$$

and

$$\begin{aligned} c_{m+1,m} &= \overline{\alpha}_{m+1}\rho_m\lambda_m - \alpha_{m-1}\rho_m\lambda_{m+1}, \\ c_{m,m+1} &= \overline{\alpha}_{m+1}\rho_m\lambda_{m+1} - \alpha_{m-1}\rho_m\lambda_m. \end{aligned} \tag{9.5.6}$$

Given an arbitrary sequence $\{\alpha_n\} \in \mathbb{D}^\infty$ one can construct a matrix $C = C(\alpha_n)$ by (9.5.1)–(9.5.3) (which generates a unitary operator in $\ell^2(\mathbb{N})$), and make sure that a distinguished unit vector $e_0 = (1, 0, 0, \dots)'$ is **cyclic**, that is, finite linear combinations of $\{C^n e_0\}_{n=-\infty}^{\infty}$ are dense in $\ell^2(\mathbb{N})$. So, C is unitarily equivalent to the multiplication by ζ on $L^2_\mu(\mathbb{T})$, μ being a spectral measure associated to C and e_0.

Theorem 9.5.1 *For an arbitrary sequence $\{\alpha_n\} \in \mathbb{D}^\infty$ a matrix C of* (9.5.4)–(9.5.6) *is the CMV matrix associated to the measure μ, that is, C takes the form* (9.5.1)–(9.5.3) *and $\alpha_n = \alpha_n(\mu)$.*

Clearly, it is just as natural to take the ordered set $1, \zeta^{-1}, \zeta, \zeta^{-2}, \zeta^2, \dots$ in place of $1, \zeta, \zeta^{-1}, \zeta^2, \zeta^{-2}, \dots$, and come to what is called the **alternate CMV basis** $\{x_n\}$ and the alternate CMV representation

$$\tilde{c_{ij}}(\mu) = (\zeta x_j, x_i)_\mu.$$

As it turns out, $\tilde{C}$ is just the transpose of C.

To state the analogue of Stone's self-adjoint cyclic model theorem, consider a cyclic unitary model, that is, a unitary operator U on a separable Hilbert space $\mathcal{H}$ with $\dim \mathcal{H} = \infty$, along with a cyclic unit vector v_0. Two cyclic unitary models $(\mathcal{H}, U, v_0)$ and $(\tilde{\mathcal{H}}, \tilde{U}, \tilde{v}_0)$ are called equivalent if there is a unitary W form $\mathcal{H}$ onto $\tilde{\mathcal{H}}$ so that

$$WUW^{-1} = \tilde{U}, \quad Wv_0 = \tilde{v}_0.$$

Theorem 9.5.2 *Each cyclic unitary model is equivalent to a unique CMV model* $(\ell^2(\mathbb{N}), C, e_0)$.

There is an important relation between CMV matrices and monic orthogonal polynomials akin to the well-known property of orthogonal polynomials on the real line:

$$\Phi_n(z) = \det(zI_n - C^{(n)}), \tag{9.5.7}$$

where $C^{(n)}$ is the principal $n \times n$ block of C. Note that $C^{(n)}$ is no more a unitary matrix. As a matter of fact, it is quite close to unitary in the sense that $C^{(n)}$ is a contraction with one-dimensional defect. So its eigenvalues (zeros of monic orthogonal polynomial Φ_n) are inside the disk.

There is another property of CMV matrices similar to the well-known property of Jacobi matrices. Given $\zeta_0 \in \mathbb{T}$, let

$$\nu = \{\nu_n\}_{n=0}^{\infty} = \{\overline{\chi_n(\zeta_0)}\}_{n=0}^{\infty}.$$

Then $C\nu = \zeta_0 \nu$, which means $\sum_j c_{kj} \nu_j = \zeta_0 \nu_k$ for all k (because of the pentadiagonal structure of C this sum always makes sense). In general, $\nu \notin \ell^2$, but if it is, then $\zeta_0 = e^{i\theta_0}$ is an eigenvalue of C, or equivalently, θ_0 is a mass point of the measure μ (cf. (9.3.11)). Furthermore, we have the following theorem.

Theorem 9.5.3 *Let* $\nu \notin \ell^2$ *but* $\liminf |\phi_n(\zeta_0)|^{1/n} \le 1$. *Then* $\zeta_0 \in \sigma(C)$ *and it is a nonisolated point of the support of* μ.

There is an explicit formula for the resolvent of the CMV matrix C in the CMV basis. It has already proved useful in some applications of CMV matrices (see Golinskiĭ, 2006). By the spectral theorem,

$$(C - zI)^{-1}_{mn} = \int_{-\pi}^{\pi} \frac{\chi_n(\zeta)\overline{\chi_m(\zeta)}}{\zeta - z}\, d\mu(\theta).$$

Let ϕ_n and ψ_n be orthonormal polynomials of the first and second kind, respectively, and F the Carathéodory function. Define

$$p_{2k}(z) = z^{-k}(F(z)\phi_{2k}(z) + \psi_{2k}(z)), \quad p_{2k-1}(z) = z^{-k}(F(z)\phi^*_{2k-1}(z) - \psi^*_{2k-1}(z)),$$
$$\pi_{2k}(z) = z^{-k}(F(z)\phi^*_{2k}(z) - \psi^*_{2k}(z)), \quad \pi_{2k-1}(z) = z^{-k+1}(F(z)\phi_{2k-1}(z) + \psi_{2k-1}(z)).$$

The following result is in Simon (2004a, Theorem 4.4.1).

Theorem 9.5.4 *For $z \in \mathbb{D}$,*

$$[(C - zI)^{-1}]_{mn} = \begin{cases} (2z)^{-1}\chi_n(z)p_m(z), & m > n, \\ (2z)^{-1}\pi_n(z)x_m(z), & n > m, \end{cases} \tag{9.5.8}$$

and

$$\begin{aligned} [(C - zI)^{-1}]_{2n-1,2n-1} &= (2z)^{-1}\chi_{2n-1}(z)p_{2n-1}(z), \\ [(C - zI)^{-1}]_{2n,2n} &= (2z)^{-1}\pi_{2n}(z)x_{2n}(z). \end{aligned} \tag{9.5.9}$$

9.6 Differential Equations

This section is based on Ismail and Witte (2001). It will be assumed that μ is absolutely continuous, that is, the orthogonality relation becomes

$$\int\limits_{|\zeta|=1} \phi_m(\zeta)\overline{\phi_n(\zeta)}w(\zeta)\frac{d\zeta}{i\zeta} = \delta_{m,n}. \tag{9.6.1}$$

Following the notation in Section 2.8, we set

$$w(z) = e^{-v(z)}, \tag{9.6.2}$$

and assume that $w(z)$ is differentiable in a neighborhood of the unit circle, has moments of all integral orders, and the integrals

$$\int\limits_{|\zeta|=1} \frac{v'(z) - v'(\zeta)}{z - \zeta}\zeta^n w(\zeta)\frac{d\zeta}{i\zeta}$$

exist for all integers n. Let

$$A_n(z) = n\frac{\kappa_{n-1}}{\kappa_n} + i\frac{\kappa_{n-1}}{\phi_n(0)}z\int\limits_{|\zeta|=1} \frac{v'(z) - v'(\zeta)}{z - \zeta}\phi_n(\zeta)\overline{\phi^*_n(\zeta)}w(\zeta)\,d\zeta, \tag{9.6.3}$$

$$B_n(z) = -i\int\limits_{|\zeta|=1} \frac{v'(z) - v'(\zeta)}{z - \zeta}\phi_n(\zeta)\left[\overline{\phi_n(\zeta)} - \frac{\kappa_n}{\phi_n(0)}\overline{\phi^*_n(\zeta)}\right]w(\zeta)\,d\zeta. \tag{9.6.4}$$

For future reference we note that $A_0 = B_0 = 0$ and

$$A_1(z) = \kappa_1 - \phi_1(z)v'(z) - \frac{\phi_1^2(z)}{\phi_1(0)}M_1(z), \tag{9.6.5}$$

$$B_1(z) = -v'(z) - \frac{\phi_1(z)}{\phi_1(0)}M_1(z), \tag{9.6.6}$$

where M_1 is defined by

$$M_1(z) = \int_{|\zeta|=1} \zeta \frac{v'(z) - v'(\zeta)}{z - \zeta} w(\zeta) \frac{d\zeta}{i\zeta}. \tag{9.6.7}$$

Theorem 9.6.1 *Under the above stated assumptions on w, the corresponding orthonormal polynomials satisfy the differential relation*

$$\phi_n'(z) = A_n(z)\phi_{n-1}(z) - B_n(z)\phi_n(z). \tag{9.6.8}$$

Define differential operators $L_{n,1}$ and $L_{n,2}$ by

$$L_{n,1} = \frac{d}{dz} + B_n(z), \tag{9.6.9}$$

and

$$L_{n,2} = -\frac{d}{dz} - B_{n-1}(z) + \frac{A_{n-1}(z)\kappa_{n-1}}{z\kappa_{n-2}} + \frac{A_{n-1}(z)\kappa_n\phi_{n-1}(0)}{\kappa_{n-2}\phi_n(0)}. \tag{9.6.10}$$

Then the operators $L_{n,1}$ and $L_{n,2}$ are annihilation and creation operators in the sense that they satisfy

$$\begin{aligned} L_{n,1}\phi_n(z) &= A_n(z)\phi_{n-1}(z), \\ L_{n,2}\phi_{n-1}(z) &= \frac{A_{n-1}(z)}{z} \frac{\phi_{n-1}(0)\kappa_{n-1}}{\phi_n(0)\kappa_{n-2}} \phi_n(z). \end{aligned} \tag{9.6.11}$$

This establishes the second-order differential equation

$$L_{n,2}\left(\frac{1}{A_n(z)} L_{n,1}\right)\phi_n(z) = \frac{A_{n-1}(z)}{z} \frac{\phi_{n-1}(0)\kappa_{n-1}}{\phi_n(0)\kappa_{n-2}} \phi_n(z). \tag{9.6.12}$$

Note that, unlike for polynomials orthogonal on the line, $L_{n,1}^*$ is not related to $L_{n,2}$.

When $v(z)$ is a meromorphic function in the unit disk then the following functional equation holds:

$$B_n + B_{n-1} - \frac{\kappa_{n-1}}{\kappa_{n-2}} \frac{A_{n-1}}{z} - \frac{\kappa_n}{\kappa_{n-2}} \frac{\phi_{n-1}(0)}{\phi_n(0)} A_{n-1} = \frac{1-n}{z} - v'(z). \tag{9.6.13}$$

Using (9.6.13) we simplify the expanded form of (9.6.12) to

$$\begin{aligned} &\phi_n'' - \left\{\frac{A_n'}{A_n} + v'(z) + \frac{n-1}{z}\right\}\phi_n' \\ &\quad + \left\{B_n' - \frac{B_n A_n'}{A_n} + B_n B_{n-1} - \frac{\kappa_{n-1}}{\kappa_{n-2}} \frac{A_{n-1}B_n}{z} - \frac{\kappa_n}{\kappa_{n-2}} \frac{\phi_{n-1}(0)}{\phi_n(0)} A_{n-1}B_n \right. \\ &\qquad \left. + \frac{\kappa_{n-1}}{\kappa_{n-2}} \frac{\phi_{n-1}(0)}{\phi_n(0)} \frac{A_{n-1}A_n}{z}\right\}\phi_n = 0. \end{aligned} \tag{9.6.14}$$

Recall that the zeros of the polynomial $\phi_n(z)$ are denoted by $\{z_{jn}\}_{1\le j\le n}$ and are confined within the unit circle $|z| < 1$. One can construct a real function $|T(z_{1n}, \ldots, z_{nn})|$ from

$$T(z_{1n}, \ldots, z_{nn}) = \prod_{j=1}^{n} z_{jn}^{-n+1} \frac{e^{-v(z_{jn})}}{A_n(z_{jn})} \prod_{1\le j<k\le n} \left(z_{jn} - z_{kn}\right)^2, \tag{9.6.15}$$

such that the zeros are given by the stationary points of this function.

This function has the interpretation of being the total energy function for n mobile unit charges in the unit disk interacting with a one-body confining potential, $v(z) + \ln A_n(z)$, an attractive logarithmic potential with a charge $n-1$ at the origin, $(n-1)\ln z$, and repulsive logarithmic two-body potentials, $-\ln(z_i - z_j)$, between pairs of charges. However, all the stationary points are saddle points, a natural consequence of analyticity in the unit disk.

For more details and examples we refer the interested reader to Ismail (2005b, Chapter 8).

9.7 Examples of OPUC

In this section a number of examples are discussed, most of which are "exactly soluble" in the sense that there are explicit formulas for both moments and Verblunsky coefficients and, in most cases there are also explicit formulas for the actual orthogonal polynomials. A nice collection of examples is in Simon (2004a, Chapter 1.6); see also Ismail (2005b, Chapters 8 and 17).

Example 9.7.1 (Free case) Let $\mu = \frac{1}{2\pi}\,d\theta$; then for the moments, Verblunsky coefficients, and orthogonal polynomials we have, respectively,

$$\mu_n = \delta_{n0}, \quad \alpha_n \equiv 0, \quad \Phi_n = \phi_n = z^n, \quad n \in \mathbb{Z}_+.$$

In this case $\phi_n^* = 1$ for all n, so the Szegő function $D(z, dm) = 1$.

Example 9.7.2 (Bernstein–Szegő measures and polynomials) Let T be a positive trigonometric (Laurent) polynomial of degree n on $\mathbb{T}$. By the Fejér–Riesz theorem there is a unique algebraic polynomial p_n of degree n with a positive leading coefficient and all zeros inside $\mathbb{D}$ so that $T(\zeta) = |p_n(\zeta)|^2$. The measures $\mu = cT^{-1}(\zeta)\,d\theta$, $\zeta = e^{i\theta}$ constitute the class of **Bernstein–Szegő measures**, $c > 0$ is a normalizing constant,

$$c^{-1} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{d\theta}{|p_n(\zeta)|^2}.$$

The orthonormal polynomials and Verblunsky coefficients are

$$\phi_k(z,\mu) = c^{-1/2} z^{k-n} p_n(z), \quad \alpha_k(\mu) = 0, \quad k = n, n+1, \ldots.$$

Since $\phi_k^* = c^{-1/2} p_n^*$, $k \ge n$, we have $D(\mu) = \sqrt{c}\,(p_n^*)^{-1}$.

An important particular case is $p_1(z) = z - \bar{w}$, $w \in \mathbb{D}$. Now $c = 1 - |w|^2$,

$$\phi_k(z,\mu) = \frac{z^k - \bar{w}z^{k-1}}{\sqrt{1-|w|^2}}, \quad d\mu(\zeta) = \frac{1-|w|^2}{|1-w\zeta|^2}\,dm, \quad D(z,\mu) = \frac{\sqrt{1-|w|^2}}{1-wz}.$$

The Verblunsky coefficients are $\alpha_0 = w$, $\alpha_j = 0$ for $j \ge 1$. The moments $\mu_j = w^j$ for $j \ge 0$ and $\mu_j = \bar{w}^{|n|}$ for $n \le 0$.

The Bernstein–Szegő measures had already arisen in Szegő's work in the early 1920s (Szegő, 1920, 1921). Translated to the real line, they were studied by Bernstein about 10 years later.

Example 9.7.3 (Single nontrivial moment) This example goes back to Grenander and Szegő (1958, Section 5.3).

Let $\mu = |1-\zeta|^2 \frac{d\theta}{4\pi}$ and $\Phi_n(\mu)$ be monic orthogonal polynomials that satisfy

$$\int_{-\pi}^{\pi} \Phi_n(\zeta)\zeta^{-j}(2-\zeta-\zeta^{-1})\,d\theta = 0, \quad j = 0,1,\dots,n-1.$$

If

$$\Phi_n(z,\mu) = \sum_{k=0}^{n} f_{kn} z^k, \quad f_{nn} = 1,$$

we come to a simple boundary value problem for the second-order difference equation

$$2f_{kn} = f_{k-1,n} + f_{k+1,n}, \quad k = 0,1,\dots,n-1, \qquad f_{-1,n} = 0, \quad f_{nn} = 1,$$

so $f_{kn} = (k+1)(n+1)^{-1}$, and

$$\Phi_n(z,\mu) = \frac{1}{n+1}\sum_{k=0}^{n}(k+1)z^k, \quad \alpha_n(\mu) = -\frac{1}{n+2}, \quad n \in \mathbb{Z}_+.$$

By (9.2.4),

$$\|\Phi_n\|^2 = \prod_{k=0}^{n-1}(1-|\alpha_k|^2) = \frac{n+2}{2(n+1)}$$

so

$$\phi_n(z,\mu) = k_n\sum_{k=0}^{n}(k+1)z^k, \quad \phi_n^*(z,\mu) = k_n\sum_{k=0}^{n}(n-k+1)z^k, \quad k_n = \sqrt{\frac{2}{(n+1)(n+2)}},$$

and

$$D^{-1}(z,\mu) = \lim_{n\to\infty}\phi_n^*(z,\mu) = \sqrt{2}\sum_{k=0}^{\infty} z^k = \frac{\sqrt{2}}{1-z},$$

initially in the sense of Taylor coefficients, but then using the Szegő limit theorem, on all of $\mathbb{D}$. The Szegő function is $D(z,\mu) = (1-z)/\sqrt{2}$.

The general case $\mu = |1-r\zeta|^2 \frac{d\theta}{2\pi(1+r^2)}$, $0 < r \le 1$ can be handled in the same way (cf. Simon, 2004a, Example 1.6.4). For instance,

$$\alpha_n(\mu) = -\frac{r^{-1}-r}{r^{-n-2}-r^{n+2}}$$

so α_n decays exponentially,

$$\Phi_n(z,\mu) = \frac{1}{d_n^-}\sum_{k=0}^{n} d_k^- z^k, \quad d_k^- = \frac{r^{-k-1}-r^{k+1}}{r^{-1}-r},$$

the Szegő function is $D(z,\mu) = (1+r^2)^{-1/2}(1-rz)$.

Example 9.7.4 (Circular Jacobi polynomials) Let $\mu = w(\zeta)\,d\theta$ with

$$w(\zeta) = \frac{\Gamma^2(a+1)}{2\pi\Gamma(2a+1)}|1-\zeta|^{2a}, \quad a > -1$$

which for $a = 1$ is Example 9.7.3. Now the orthogonal polynomials are expressed in terms of the hypergeometric function

$$\phi_n(z,\mu) = \frac{(a)_n}{\sqrt{n!(2a+1)_n}}\, {}_2F_1(-n, a+1; -n+1-a; z), \quad (a)_n = a(a+1)\dots(a+n-1),$$

and the Verblunsky coefficients are

$$\alpha_n(\mu) = -\frac{a}{n+a+1}, \quad n \in \mathbb{Z}_+.$$

Example 9.7.5 (Rogers–Szegő polynomials) The example is from Szegő (1926) and the name comes from earlier consideration of Rogers (1894, 1895). Ismail (2005b) has a whole Chapter 17 on this example (see also Simon, 2004a, Example 1.6.5). This class of polynomials is parametrized by a number $q \in (0,1)$ (although the extension to $q \in \mathbb{D}$ is easy). The weight function is a "wrapped Gaussian." Let

$$q = e^{-a}, \quad a = \log\frac{1}{q} > 0.$$

The Gaussian measure on the real line of variance a is given by

$$d\nu_a(x) = (2\pi a)^{-1/2} e^{-x^2/2a}\,dx.$$

The wrapped Gaussian measure on $[-\pi,\pi]$ is defined by

$$\mu = \mu(q,\zeta) = v_q(\zeta)\,d\theta, \quad v_q(e^{i\theta}) = \frac{1}{\sqrt{2\pi a}}\sum_{j=-\infty}^{\infty} e^{-(\theta-2\pi j)^2/2a}. \tag{9.7.1}$$

It is a matter of direct calculation to find the moments $\mu_n = q^{n^2/2}$.

Identifying the orthogonal polynomials depends on the use of q-binomial coefficients defined by

$$\begin{bmatrix} n \\ j \end{bmatrix}_q = \frac{[n]_q}{[j]_q[n-j]_q}, \quad [n]_q = (1-q)(1-q^2)\cdots(1-q^n),\ [0]_q = 1.$$

The monic orthogonal with respect to the wrapped Gaussian measure (9.7.1) polynomials, known as the Rogers–Szegő polynomials, are

$$\Phi_n(z,\mu) = \sum_{j=0}^{n} (-1)^{n-j} \begin{bmatrix} n \\ j \end{bmatrix}_q q^{(n-j)/2} z^j, \tag{9.7.2}$$

so

$$\alpha_n(\mu) = (-1)^n q^{(n+1)/2}, \quad \|\Phi_n\|^2 = [n]_q.$$

The Szegő function is now

$$D(z,\mu) = \prod_{j=1}^{\infty}(1-q^j)^{1/2}\prod_{j=0}^{\infty}(1+zq^{j+1/2}).$$

An amazing property of zeros of the Rogers–Szegő polynomials (9.7.2) is proved in Mazel, Geronimo, and Hayes (1990): all of them have their zeros on the same circle $|z| = q^{1/2}$.

Example 9.7.6 (Geronimus measures and polynomials) This example is perhaps the most notable example of a measure outside the Szegő class. In this (and the next) examples it is more convenient to view a measure as one supported on the unit circle $\mathbb{T}$.

The Geronimus polynomials are those associated with constant Verblunsky coefficients $\alpha_n \equiv \alpha$, $\alpha \in \mathbb{D}\backslash\{0\}$. By Verblunsky's theorem (see Section 9.8) the corresponding measure μ_α, called the Geronimus measure, is uniquely determined. The measures and polynomials appeared in Geronimus (1977), and have been extensively studied over the past fifteen years (see Simon, 2004a, Example 1.6.12).

The Szegő recurrence relations (9.2.9) for orthonormal Geronimus polynomials and their reverse take the form

$$\begin{bmatrix}\phi_n(z)\\ \phi_n^*(z)\end{bmatrix} = T^n(z,\alpha)\begin{bmatrix}1\\1\end{bmatrix}, \quad T(z,\alpha) = \frac{1}{\sqrt{1-|\alpha|^2}}\begin{bmatrix} z & -\bar{\alpha}\\ -z\alpha & 1\end{bmatrix}. \tag{9.7.3}$$

It is not hard now to derive the expressions for Geronimus polynomials and their reverse. Denote by $r_{1,2}$ the eigenvalues of matrix T (cf. (9.7.3)), which are the roots of characteristic equation

$$r^2 - \frac{z+1}{\rho}r + z = 0, \quad \rho = \sqrt{1-|\alpha|^2},$$

so

$$r_{1,2}(z) = \frac{z+1\pm\sqrt{(z+1)^2-4\rho^2 z}}{2\rho} = \frac{z+1\pm\sqrt{(z-\zeta_\tau)(z-\zeta_\tau^{-1})}}{2\rho} \tag{9.7.4}$$

with $\zeta_\tau = e^{i\tau}$ and $\sin\frac{\tau}{2} = |\alpha|$, $0 < \tau < \pi$, and the branch of the square root is taken so that $r_2(0) = 0$. It is clear that the spectrum of T depends only on $|\alpha|$. Define a circular arc Δ_τ closely related to T by

$$\Delta_\tau = \{\zeta = e^{it} : \tau \le t \le 2\pi - \tau\} \tag{9.7.5}$$

so

$$|r_2(z)| < 1 < |r_1(z)|, \quad z \in \mathbb{C}\backslash\Delta_\tau; \qquad |r_2(\zeta)| = |r_1(\zeta)| = 1, \quad \zeta \in \Delta_\tau,$$

and $r_1 = r_2$ only at the endpoints of Δ_τ. It follows from (9.7.3) that

$$\phi_n(z,\mu_\alpha) = \frac{z-\bar{\alpha}}{\rho}\frac{r_1^n - r_2^n}{r_1 - r_2} - z\frac{r_1^{n-1}-r_2^{n-1}}{r_1-r_2}, \tag{9.7.6}$$

$$\phi_n^*(z,\mu_\alpha) = \frac{1-\alpha z}{\rho}\frac{r_1^n - r_2^n}{r_1 - r_2} - z\frac{r_1^{n-1}-r_2^{n-1}}{r_1-r_2}. \tag{9.7.7}$$

There is another expression for Geronimus polynomials which holds on the arc Δ_τ. Indeed, for $e^{it} \in \Delta_\tau$,

$$r_{1,2}(e^{it}) = \frac{e^{i\frac{t}{2}}}{\rho}\left(\cos\frac{t}{2} \pm i\sqrt{\cos^2\frac{\tau}{2} - \cos^2\frac{t}{2}}\right) = e^{i\frac{t}{2}}(\cos\lambda \pm i\sin\lambda), \quad \cos\lambda = \frac{\cos\frac{t}{2}}{\cos\frac{\tau}{2}},$$

$0 \le \lambda \le \pi$, so one has

$$\phi_n(e^{it}, \mu_\alpha) = e^{in\frac{t}{2}}\left(\frac{e^{i\frac{t}{2}} - \bar{\alpha}e^{-i\frac{t}{2}}}{\rho}U_{n-1}(\cos\lambda) - U_{n-2}(\cos\lambda)\right),$$

where U_k are the Chebyshev polynomials of the second kind. In particular, there is a bound for Geronimus polynomials on Δ_τ,

$$|\phi_n(\zeta, \mu_\alpha)| \le C(\alpha)\min(n, v^{-1}(\zeta)), \quad v(z) = \sqrt{(z - \zeta_\tau)(z - \zeta_\tau^{-1})}, \quad n \in \mathbb{Z}_+,$$

and hence they are uniformly bounded inside the arc Δ_τ and

$$|\phi_n(e^{\pm i\tau})| = \left|\left(\frac{e^{i\frac{\tau}{2}} - \bar{\alpha}e^{-i\frac{\tau}{2}}}{\rho} \mp 1\right)n \pm 1\right|. \tag{9.7.8}$$

It is clear from the definition that the second kind measures and polynomials are also Geronimus measures and polynomials for the parameter $-\alpha$, so for $\psi_n(z, \mu_\alpha)$, $\psi_n^*(z, \mu_\alpha)$ the same formulas as (9.7.6) hold.

The Carathéodory function (9.2.18) can be now computed explicitly

$$F(z, \mu_\alpha) = 1 + \frac{z + 2\alpha z - 1 + \sqrt{(z - \zeta_\tau)(z - \zeta_\tau^{-1})}}{(1 + \alpha)(\zeta_\beta - z)}, \quad \zeta_\beta = e^{i\beta} = \frac{1 + \bar{\alpha}}{1 + \alpha}. \tag{9.7.9}$$

Thus there is at most one mass point at $\zeta_\beta \notin \Delta_\tau$, and the actual value of this mass can be found from (9.2.23) and (9.7.9):

$$\mu_\alpha\{\zeta_\beta\} = \begin{cases} \frac{2}{|1+\alpha|^2}\left(\left|\alpha + \frac{1}{2}\right|^2 - \frac{1}{4}\right), & |\alpha + \frac{1}{2}| > \frac{1}{2}, \\ 0, & |\alpha + \frac{1}{2}| \le \frac{1}{2}. \end{cases} \tag{9.7.10}$$

As follows from (9.7.9), the measure μ_α is supported on Δ_τ along with a possible mass point at ζ_β, and $\mu_\alpha = \mu_\alpha'\, dm$ with

$$\mu_\alpha'(e^{it}) = \frac{1}{|1 + \alpha|}\frac{\sqrt{\cos^2\frac{\tau}{2} - \cos^2\frac{t}{2}}}{\sin\frac{t-\beta}{2}}, \quad e^{it} \in \Delta_\tau.$$

Example 9.7.7 (Perturbed Geronimus measures) A measure μ is called a perturbed Geronimus measure if

$$\lim_{n\to\infty} \alpha_n(\mu) = \alpha, \quad 0 < |\alpha| < 1. \tag{9.7.11}$$

The following fundamental result about such measures is due to Geronimus (1941).

Theorem 9.7.8 *Let μ be a perturbed Geronimus measure. Then Δ_τ (cf. (9.7.5)) belongs to the support of μ, and the part of the support outside Δ_τ is at most a countable point set which can accumulate only to the endpoints of Δ_τ.*

Much more can be said about μ as long as some additional assumptions are imposed upon the rate of convergence in (9.7.11). The following results are in Golinskii, Nevai, and Van Assche (1995).

Theorem 9.7.9 *Let μ be a perturbed Geronimus measure.*

(i) *If $\{\alpha_n(\mu) - \alpha\} \in \ell^1$ then μ is absolutely continuous inside Δ_τ, and $1/\mu' \in L^\infty(\Delta)$ for any interior closed arc $\Delta \subset \Delta_\tau$.*
(ii) *If $\{\log n(\alpha_n(\mu) - \alpha)\} \in \ell^1$ then μ satisfies the Szegő condition for the arc*

$$\int_{\Delta_\tau} \frac{|\log \mu'(\zeta)|}{\sqrt{|\zeta - \zeta_\tau||\zeta - \zeta_\tau^{-1}|}}\, dm < \infty.$$

(iii) *If $\{n(\alpha_n(\mu) - \alpha)\} \in \ell^1$, then μ is absolutely continuous on the whole Δ_τ, and $\mu'(\zeta) \geq C|\zeta - \zeta_\tau||\zeta - \zeta_\tau^{-1}|$ a.e. on Δ_τ.*

The bounds for perturbed Geronimus polynomials are also available.

Theorem 9.7.10 *Let $\{\phi_n\}$ be perturbed Geronimus polynomials. If $\{\alpha_n(\mu) - \alpha\} \in \ell^1$, then*

$$\sup_n \max_{\zeta\in\Delta} |\phi_n(\zeta,\mu)| = C(\Delta) < \infty$$

for any interior closed arc $\Delta \subset \Delta_\tau$. If $\{n(\alpha_n(\mu) - \alpha)\} \in \ell^1$, then

$$\sup_n \max_{\zeta\in\Delta_\alpha} \frac{|\phi_n(\zeta,\mu)|}{n} < \infty.$$

Equation (9.7.8) shows that the latter bound is optimal.

The following result (Golinskii, 2000) provides a sufficient condition for the perturbed Geronimus measure to have finitely many mass points outside Δ_τ.

Theorem 9.7.11 *The portion of the support of μ outside Δ_τ is a finite set as long as $\{n(\alpha_n(\mu) - \alpha)\} \in \ell^1$.*

9.8 Modification of Measures

By Verblunsky's theorem each transformation in the class of nontrivial probability measures on $[-\pi,\pi]$ gives rise to a certain transformation in the space $\mathbb{D}^\infty$ of the Verblunsky coefficients and vice versa. We consider here the simplest such transformations when the explicit expressions are available. Again we will deal with measures on $\mathbb{T}$ rather than on $[-\pi,\pi]$.

Let S be a Borel transformation of $\mathbb{T}$ into itself. Such a transformation acts in the space of measures by $S\mu = \mu_S$, $\mu_S(E) = \mu(S^{-1}E)$. A key role is played by the change of variables formula

$$\int_{\mathbb{T}} h(\zeta)\, d\mu_S = \int_{\mathbb{T}} h(S(\zeta))\, d\mu.$$

9.8.1 Rotation of the Circle and Parameters

Let $\lambda \in \mathbb{T}$ and $S(\zeta) = \lambda\zeta$. It is clear that

$$\Phi_n(z, \mu_S) = \lambda^n \Phi_n(\bar{\lambda} z, \mu), \quad \alpha_n(\mu_S) = \lambda^{-n-1}\alpha_n(\mu)$$

for the monic orthogonal polynomials and Verblunsky coefficients, respectively. For the Carathéodory functions one has $F(z, \mu_S) = F(\bar{\lambda} z, \mu)$.

Conversely, the rotation of parameters leads to Aleksandrov measures $\{\mu_\lambda\}_{\lambda\in\mathbb{T}}$ with $\alpha_n(\mu_\lambda) = \lambda\alpha_n$. The second kind measure is included in the family with $\lambda = -1$. The rotation $\alpha_n \to \lambda\alpha_n$ can be viewed as a change of boundary conditions since (cf. (9.2.7))

$$\begin{bmatrix} \Phi_{n,\lambda}(z) \\ \bar{\lambda}\Phi^*_{n,\lambda}(z) \end{bmatrix} = T_{n+1}(z) \begin{bmatrix} 1 \\ \bar{\lambda} \end{bmatrix}, \tag{9.8.1}$$

where $\Phi_{n,\lambda}$ are monic orthogonal polynomials for μ_λ. Since the space of solutions of (9.2.6) is 2-dimensional, any solution can be written in terms of Φ and Ψ:

$$2\Phi_{n,\lambda}(z) = (1 + \bar{\lambda})\Phi_n(z) + (1 - \bar{\lambda})\Psi_n(z). \tag{9.8.2}$$

For the corresponding Carathéodory functions one has

$$F_\lambda(z) = \frac{(1 - \lambda) + (1 + \lambda)F(z)}{(1 + \lambda) + (1 - \lambda)F(z)}, \quad F_{-1}(z) = \frac{1}{F(z)}. \tag{9.8.3}$$

It is sometimes advisable to study spectral properties of the entire family of Aleksandrov measures. The following result is in Simon (2004a, Theorem 3.2.16).

Theorem 9.8.1 *Let the Lebesgue decomposition for Aleksandrov measures be*

$$d\mu_\lambda = w_\lambda(\zeta)\, dm + d\mu_{s,\lambda}.$$

Then

(i) μ_λ *have the same essential support, and* $\{\zeta : w_\lambda(\zeta) \neq 0\}$ *is a.e. independent of* λ.
(ii) *If* $\operatorname{supp}(\mu_1) \cap (\zeta_0, \zeta_1)$ *is a finite set, the same is true for* $\operatorname{supp}(\mu_\lambda) \cap (\zeta_0, \zeta_1)$ *for each* λ.
(iii) *The singular components* $\mu_{\lambda,s}$ *and* $\mu_{\lambda',s}$ *are mutually singular for* $\lambda \neq \lambda'$.

There is another important property of Aleksandrov measures, known as the "spectral averaging," which states that roughly speaking the average of μ_λ over λ is always the Lebesgue measure (Golinskii and Nevai, 2001). More precisely, for any Borel set $B \subset \mathbb{T}$,

$$\int_{\mathbb{T}} \mu_\zeta(B,)\, dm(\zeta) = m(B).$$

9.8.2 Sieved Measures and Polynomials

Let $N \geq 2$ be a positive integer, and $S(\zeta) = \zeta^N$. Now $\mu_S = \mu^{(N)}$ puts scaled copies of μ on each of the arcs $[\zeta_j, \zeta_{j+1}]$ with

$$\zeta_j = \exp(2\pi i j/N), \quad j = 0, 1, \ldots, N-1.$$

One can easily show that for Verblunsky coefficients,

$$\alpha_n(\mu^{(N)}) = \begin{cases} \alpha_r(\mu), & n = rN + N - 1, \\ 0, & \text{otherwise.} \end{cases}$$

For the monic orthogonal polynomials one has

$$\Phi_n(z, \mu^{(N)}) = z^k \Phi_r(z, \mu), \quad n = rN + k, \quad k = 0, 1, \ldots, N-1.$$

The Carathéodory function is $F(z, \mu^{(N)}) = F(z^N, \mu)$.

A process in this example is known as **sieving**, and these Φ are the sieved polynomials. They were systematically discussed in Badkov (1987) and Ismail and Li (1992b).

We complete with a particular example of the Al-Salam–Carlitz q-polynomials on the unit circle. Let

$$A_n(x) = \frac{U_n^{(-1)}(x; q)}{q^{\frac{n(n-1)}{4}} \sqrt{(q; q)_n}}$$

be orthonormal Al-Salam–Carlitz q-polynomials (see Section 6.7 of this volume). The orthogonality measure γ is concentrated on two sequences $\{\pm q^j\}$, which converge to zero, and is symmetric with respect to the origin: $\gamma\{q^j\} = \gamma\{-q^j\}$. The three-term recurrence relation is

$$x A_n(x) = a_{n+1} A_{n+1}(x) + a_n A_{n-1}(x), \quad a_n^2 = q^{n-1}(1 - q^n).$$

Going over first to the unit circle by the Szegő mapping theorem we end up with $\mu \in \mathcal{P}$ concentrating on a discrete point set $\{e^{\pm i\theta_j^\pm}\}$ with $\cos\theta_j^\pm = \pm q^j$, which has two limit points ± 1. The corresponding Verblunsky coefficients are

$$\alpha_{2k}(\mu) = 0, \quad \alpha_{2k+1} = 1 - 2q^{k+1}, \quad k \in \mathbb{Z}_+.$$

9.8.3 Inserting Point Mass

There is an interesting problem of comparing Verblunsky coefficients and orthogonal polynomials of two measures μ and ν. We consider here an obvious way of building ν from μ by adding a mass point (finitely many mass points). Such a transformation is known as the Jost–Kohn perturbation. Explicitly,

$$\nu = t\mu + (1-t)\sigma, \quad 0 < t < 1, \quad \sigma = \sum_{j=1}^{p} k_j \delta(\zeta_j) \tag{9.8.4}$$

is a finite linear combination of pure point masses adjusted so that σ is a probability measure. Jost–Kohn theory for OPUC appeared in Golinskii (1966), Geronimus (1961), Cachafeiro and Marcellán (1988, 1993), Marcellán and Maroni (1992), and Peherstorfer and Steinbauer (1999). In particular, the phenomenon discovered in Peherstorfer and Steinbauer (1999) says that it can happen that adding a point mass to a case with $\alpha_n(\mu) \to a$ can result with a ν obeying $\alpha_n(\nu) \to a' \neq a$.

The relation between the Carathéodory functions is simple:

$$F(z,\nu) = tF(z,\mu) + (1-t)\sum_{j=1}^{p} k_j \frac{\zeta_j + z}{\zeta_j - z}.$$

For the case $p = 1$, $\nu = \nu(t,\zeta_1)$ the relation between orthogonal polynomials was obtained in Geronimus (1961, formula (3.30)),

$$\Phi_n(z,\nu) = \Phi_n(z,\mu) - \frac{s\Phi_n(\zeta_1,\mu)K_{n-1}(z,\zeta_1;\mu)}{1 + sK_{n-1}(\zeta_1,\zeta_1;\mu)}, \qquad s = \frac{t}{1-t}, \tag{9.8.5}$$

where K_n is the Christoffel kernel (9.2.24). By using the complex conjugate of (9.2.29) we have for $z = 0$,

$$\alpha_n(\nu) - \alpha_n(\mu) = \frac{s\overline{\Phi_{n+1}(\zeta_1,\mu)}\kappa_n\phi_n(\zeta_1,\mu)}{1 + sK_n(\zeta_1,\zeta_1;\mu)}. \tag{9.8.6}$$

But

$$|\Phi_{n+1}(\zeta_1,\mu)| = \left|\zeta_1\Phi_n(\zeta_1,\mu) - \bar{\alpha}_n(\mu)\Phi_n^*(\zeta_1,\mu)\right| \le 2\,|\Phi_n(\zeta_1,\mu)|$$

so

$$|\alpha_n(\nu) - \alpha_n(\mu)| \le \frac{2s|\phi_n(\zeta_1,\mu)|^2}{1 + sK_n(\zeta_1,\zeta_1;\mu)}. \tag{9.8.7}$$

Let us say that a class X of nontrivial probability measures on $\mathbb{T}$ is invariant with regard to addition of the mass points if $\mu \in X$ implies $\nu \in X$ for all $0 < t < 1$ and $\zeta_1 \in \mathbb{T}$. Clearly, both Szegő and Erdös classes are invariant (the addition of a mass point does not affect the absolutely continuous part of the measure). As a consequence of (9.8.7) and Theorem 9.3.18 one has a much more delicate result that the Nevai class is also invariant.

As far as the Rakhmanov class goes, the problem is still open. There is partial result in this direction (Golinskii and Khrushchev, 2002) which claims that a proper subclass $\mathcal{R}_0 \subset \mathcal{R}$, which consists of measures $\mu \in \mathcal{R}$ with $\sup_n |\alpha_n(\mu)| < 1$, is invariant with regard to addition of the mass points.

9.8.4 Modification by a Rational Function

Let G be a rational function regular on $\mathbb{T}$ such that

$$\int_{-\pi}^{\pi} |G(\zeta)|^2\, d\mu(\theta) = 1, \quad \zeta = e^{i\theta}.$$

Put $\nu = |G|^2\mu$, also known as the Christoffel–Bargmann perturbation.

We start with the unit circle analogue of the Christoffel formula.

Theorem 9.8.2 *Let $\{\phi_n\}$ be orthonormal with respect to μ, and G_{2m} be a polynomial of precise degree $2m$ such that*

$$\zeta^{-m} G_{2m}(\zeta) = |G_{2m}(\zeta)|, \quad |\zeta| = 1.$$

Let $\phi = \phi_{n+m}$. Define polynomials $\{\xi_n\}$ by

$$G_{2m}(z)\xi_n(z) \tag{9.8.8}$$
$$= \begin{vmatrix} \phi^*(z) & z\phi^*(z) & \cdots & z^{m-1}\phi^*(z) & \phi(z) & z\phi(z) & \cdots & z^m\phi(z) \\ \phi^*(z_1) & z_1\phi^*(z_1) & \cdots & z_1^{m-1}\zeta^*(z_1) & \phi(z_1) & z_1\phi(z_1) & \cdots & z_1^m\phi(z_1) \\ \phi^*(z_2) & z_2\phi^*(z_2) & \cdots & z_2^{m-1}\phi^*(z_2) & \phi(z_2) & z_2\phi(z_2) & \cdots & z_2^m\phi(z_2) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \phi^*(z_{2m}) & z_{2m}\phi^*(z_{2m}) & \cdots & z_{2m}^{m-1}\phi^*(z_{2m}) & \phi(z_{2m}) & z_{2m}\phi(z_{2m}) & \cdots & z_{2m}^m\phi^*(z_{2m}) \end{vmatrix},$$

where $z_1, z_2, \ldots, z_{2m}$ are the zeros of G_{2m}.

For zeros of multiplicity r, $r > 1$, replace the corresponding rows in (9.8.8) *by derivatives of order* $0, 1, \ldots, r-1$ *of the polynomials in the first row evaluated at that zero.*

Then $\{\xi_n(z)\}$ are orthogonal with respect to $C\,|G_{2m}(\zeta)|\, d\mu(\theta)$.

A similar result holds when G goes in the denominator.

Theorem 9.8.3 *Let μ, $\{\phi_n\}$ be as in the above theorem. Let H_{2k} be a polynomial of precise degree $2k$ such that*

$$\zeta^{-k} H_{2k}(\zeta) = |H_{2k}(\zeta)| > 0, \qquad |\zeta| = 1,$$

and put $\phi = \phi_{n+k}$. Define a new system of polynomials $\{\eta_n\}$, $n = 2k, 2k+1, \ldots$ by

$$\eta_n(z) \tag{9.8.9}$$
$$= \begin{vmatrix} \phi^*(z) & z\phi^*(z) & \cdots & z^{k-1}\phi^*(z) & \phi(z) & z\phi(z) & \cdots & z^k\phi(z) \\ L_{w_1}(\phi^*) & L_{w_1}(z\phi^*) & \cdots & L_{w_1}(z^{k-1}\phi^*) & L_{w_1}(\phi) & L_{w_1}(z\phi) & \cdots & L_{w_1}(z^k\phi) \\ L_{w_2}(\phi^*) & L_{w_2}(z\phi^*) & \cdots & L_{w_2}(z^{k-1}\phi^*) & L_{w_2}(\phi) & L_{w_2}(z\phi) & \cdots & L_{w_2}(z^k\phi) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ L_{w_{2k}}(\phi^*) & L_{w_{2k}}(z\phi^*) & \cdots & L_{w_{2k}}(z^{k-1}\phi^*) & L_{w_{2k}}(\phi) & L_{w_{2k}}(z\phi) & \cdots & L_{w_{2k}}(z^k\phi) \end{vmatrix},$$

where the zeros of H_{2k} are $w_1, w_2, \ldots, w_{2k}$, and we define

$$L_\beta(p) := \int_{-\pi}^{\pi} p(\zeta)\overline{\left(\frac{\zeta^k}{\zeta-\beta}\right)}\, d\mu(\theta), \quad \beta \notin \mathbb{T}.$$

For zeros of multiplicity h, $h > 1$, we replace the corresponding rows in the determinant (9.8.9) *by*

$$L_\beta^j(p) := \int_{-\pi}^{\pi} p(\zeta)\overline{\left(\frac{\zeta^k}{(\zeta-\beta)^j}\right)}\, d\mu(\theta), \quad j = 1, 2, \ldots, h$$

acting on the first row.

Under the above assumptions, for $n \ge 2k$, $\{\eta_n(z)\}$ are orthogonal with respect to $C\,|H_{2k}(\zeta)|^{-1}\,d\mu$.

A combination of Theorems 9.8.2 and 9.8.3 leads to the following result which covers the modification by a rational function.

Theorem 9.8.4 *Let μ, $\{\phi_n(z)\}$, G_{2m}, H_{2k}, and $z_1, \dots, z_{2m}$, $w_1, \dots, w_{2k}$ be as in Theorems* 9.8.2 *and* 9.8.3. *Let ϕ denote ϕ_{n+m-k} and $s = m + k$. For $n \ge 2k$ define ψ_n by*

$$G_{2m}(z)\psi_n(z) \tag{9.8.10}$$
$$= \begin{vmatrix} \phi^*(z) & z\phi^*(z) & \cdots & z^{s-1}\phi^*(z) & \phi(z) & z\phi(z) & \cdots & z^s\phi(z) \\ \phi^*(z_1) & z_1\phi^*(z_1) & \cdots & z_1^{s-1}\phi^*(z_1) & \phi(z_1) & z_1\phi(z_1) & \cdots & z_1^s\phi(z_1) \\ \phi^*(z_2) & z_2\phi^*(z_2) & \cdots & z_2^{s-1}\phi^*(z_2) & \phi(z_2) & z_2\phi(z_2) & \cdots & z_2^s\phi(z_2) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \phi^*(z_{2m}) & z_{2m}\phi^*(z_{2m}) & \cdots & z_{2m}^{s-1}\phi^*(z_{2m}) & \phi(z_{2m}) & z_{2m}\phi(z_{2m}) & \cdots & z_{2m}^s\phi(z_{2m}) \\ L_{w_1}(\phi^*) & L_{w_1}(z\phi^*) & \cdots & L_{w_1}(z^{s-1}\phi^*) & L_{w_1}(\phi) & L_{w_1}(z\phi) & \cdots & L_{w_1}(z^s\phi) \\ L_{w_2}(\phi^*) & L_{w_2}(z\phi^*) & \cdots & L_{z_2}(z^{s-1}\phi^*) & L_{w_2}(\phi) & L_{z_2}(z\phi) & \cdots & L_{w_2}(z^s) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ L_{w_{2k}}(\phi^*) & L_{w_{2k}}(z\phi^*) & \cdots & L_{w_{2k}}(z^{s-1}\phi^*) & L_{w_{2k}}(\phi) & L_{w_{2k}}(z\phi) & \cdots & L_{w_{2k}}(z^s\phi) \end{vmatrix},$$

where we define

$$L_\beta(p) := \int_{-\pi}^{\pi} p(\zeta)\overline{\left(\frac{\zeta^s}{\zeta - \beta}\right)}\,d\mu(\theta), \quad \beta \notin \mathbb{T}.$$

For zeros of H_{2k} of multiplicity h, $h > 1$, we replace the corresponding rows in the determinant (9.8.10) *by*

$$L_\beta^j(p) := \int_{-\pi}^{\pi} p(\zeta)\overline{\left(\frac{\zeta^s}{(\zeta - \beta)^j}\right)}\,n d\mu(\theta), \quad j = 1, 2, \dots, h$$

acting on the first row.

For zeros of G_{2m} of multiplicity h, $h > 1$, we replace the corresponding row in the determinant (9.8.10) *by the derivatives of order $0, 1, 2, \dots, h-1$ of the polynomials in the first row, evaluated at that zero.* (*As usual, $p_r^*(z) = z^r\,\bar{p}_r(z^{-1})$, for ψ_r a polynomial of degree r.*)

Then $\{\psi_n\}$ are orthogonal with respect to $C|G_{2m}/H_{2k}|\,d\mu$ on the unit circle.

The results of this section are in Ismail and Ruedemann (1992), which contains explicit formulas for certain polynomials. For earlier partial results see Golinskii (1958), Mikaelyan (1978), and Godoy and Marcellán (1991).

9.8.5 Bessel Transformations and Schur Flows

Throughout the rest of the section we will view a nontrivial probability measure μ supported on $\mathbb{T}$. Define a family of measures which depends on parameter $t \ge 0$ by

$$\mu(\zeta, t) = C(t)e^{t(\zeta+\zeta^{-1})}\mu(\zeta, 0), \quad C^{-1}(t) = \int_{\mathbb{T}} e^{t(\zeta+\zeta^{-1})}\, d\mu(\zeta, 0) \tag{9.8.11}$$

is a normalizing factor. We refer to (9.8.11) as the **Bessel transformation** of the initial measure $\mu = \mu(\cdot, 0)$. The main problem we deal with here is the dynamics of corresponding orthogonal polynomials $\Phi_n(\cdot, t)$ and Verblunsky coefficients $\alpha_n(\mu(t)) = \alpha_n(t)$.

As far as the polynomials go, the following result is proved in Golinskiĭ (2006).

Theorem 9.8.5 *The monic polynomials $\Phi_n(\cdot, t)$ orthogonal with respect to $\mu(t)$ (cf.* (9.8.11)*) satisfy the first-order differential equation*

$$\frac{d}{dt}\Phi_n(z, t) = \Phi_{n+1}(z, t) - (z + \overline{\alpha}_n(t)\alpha_{n-1}(t))\Phi_n(z, t) - (1 - |\alpha_{n-1}(t)|^2)\Phi_{n-1}(z, t).$$

A comprehensive study of the asymptotic behavior of Verblunsky coefficients $\alpha_n(t)$ for each fixed n and $t \to \infty$ is accomplished in Simon (2007b). Moreover, in Simon (2007b) the asymptotics of the zeros $\{z_{j,n}(t)\}_{j=1}^n$ of Φ_n is examined, which yields the information about α_n via

$$\alpha_{n-1}(t) = (-1)^{n-1} \prod_{j=1}^{n} \overline{z_{j,n}(t)} \tag{9.8.12}$$

The key tool is Theorem 9.1.2. As it turns out, the limit behavior of the α_n depends heavily on whether the point 1 belongs to the essential support of the initial measure $\mu(\zeta, 0)$, that is, any punctured neighborhood of 1 has nonempty intersection with the support of μ, or not. The former case is rather simple, and here is the result.

Theorem 9.8.6 *Let* $1 \in \operatorname{supp}_{\text{ess}} \mu$. *Then*

$$\lim_{t\to\infty} z_{j,n}(t) = 1 \ \textit{ for all } n \in \mathbb{N},\ j = 1, 2, \ldots, n \quad \textit{implies} \quad \lim_{t\to\infty} \alpha_{n-1}(t) = (-1)^{n-1}.$$

The latter case is much more complicated, and a complete picture is available only for the case when μ is symmetric (and then so are all $\mu(t)$), and $\alpha_n(t)$ are real-valued functions. Now, there exists a unique open arc $\Gamma(\mu) = (\bar{\Theta}, \Theta)$, $\Theta = \Theta(\mu)$ so that $\Im\Theta > 0$ and

(i) its endpoints $\bar{\Theta}, \Theta$ belong to the essential support of μ, and $1 \in \Gamma(\mu)$;
(ii) the portion of $\operatorname{supp}\mu$ on $\Gamma(\mu)$ is at most a countable set of mass points $\{\zeta_j\}_{j=1}^N$, $N \le \infty$, with no limit points inside Γ.

One can label ζ_j so that $\Re\zeta_1 \ge \Re\zeta_2 \ge \cdots$, and it is clear that this series of inequalities cannot have two equality signs in a row. Specifically, $\Re\zeta_n = \Re\zeta_{n+1}$ if and only if $\zeta_{n+1} = \bar{\zeta}_n$.

Theorem 9.8.7 *Suppose that* $1 \notin \operatorname{supp}_{\text{ess}} \mu$*, and* $\Gamma(\mu)$ *has an infinity of mass points* ζ_j.

(i) *If* $1 \in \operatorname{supp}\mu$, *then* $1 = \Re\zeta_1 > \Re\zeta_2 = \Re\zeta_3 > \Re\zeta_4 = \Re\zeta_5 > \cdots$ *and*

$$\lim_{t\to\infty} \alpha_{2n}(t) = 1, \quad \lim_{t\to\infty} \alpha_{2n+1}(t) = -\Re\zeta_{2n+2}, \quad n \in \mathbb{Z}_+.$$

(ii) *If* $1 \notin \operatorname{supp}\mu$, *then* $1 > \Re\zeta_1 = \Re\zeta_2 > \Re\zeta_3 = \Re\zeta_4 > \cdots$ *and*

$$\lim_{t\to\infty} \alpha_{2n}(t) = \Re\zeta_{2n+1}, \quad \lim_{t\to\infty} \alpha_{2n+1}(t) = -1, \quad n \in \mathbb{Z}_+.$$

Theorem 9.8.8 *Suppose that* $1 \notin \operatorname{supp}_{\mathrm{ess}}\mu$, *and* $\Gamma(\mu)$ *has* $N < \infty$ *mass points* ζ_j.

(i) *If* $N = 2m+1$, *then* $1 \in \operatorname{supp}\mu$,

$$1 = \Re\zeta_1 > \Re\zeta_2 = \Re\zeta_3 > \cdots > \Re\zeta_{N-1} = \Re\zeta_N > \Re\Theta(\mu)$$

and

$$\begin{aligned} &\lim_{t\to\infty} \alpha_{2n}(t) = 1, && n \in \mathbb{Z}_+, \\ &\lim_{t\to\infty} \alpha_{2n+1}(t) = -\Re\zeta_{2n+2}, && n = 0, 1, \ldots, m-1, \\ &\lim_{t\to\infty} \alpha_{2n+1}(t) = -\Re\Theta(\mu), && n = m, m+1, \ldots. \end{aligned}$$

(ii) *If* $N = 2m$, *then* $1 \notin \operatorname{supp}\mu$,

$$1 > \Re\zeta_1 = \Re\zeta_2 > \Re\zeta_3 = \cdots > \Re\zeta_{N-1} = \Re\zeta_N > \Re\Theta(\mu)$$

and

$$\begin{aligned} &\lim_{t\to\infty} \alpha_{2n+1}(t) = -1, && n \in \mathbb{Z}_+, \\ &\lim_{t\to\infty} \alpha_{2n}(t) = -\Re\zeta_{2n+1}, && n = 0, 1, \ldots, m-1, \\ &\lim_{t\to\infty} \alpha_{2n}(t) = \Re\Theta(\mu), && n = m, m+1, \ldots. \end{aligned}$$

Some particular results for the general case are also obtained in Simon (2007b). For instance, an example of a measure μ with $\Theta(\mu) = i$ and no mass points in $\Gamma(\mu)$ is given for which $\alpha_0(t)$ has no limit as $t \to \infty$.

Note that the distinguished role of the point 1 is quite obvious: this is the only global maximum for the function $\Re\zeta$ on $\mathbb{T}$. If 1 is in the essential support, it attracts all zeros of all polynomials Φ_n. If 1 is an isolated mass point, it can attract only one zero by Theorem 9.4.4. The behavior of other zeros is in general rather chaotic.

One can think of the Bessel transformation (9.8.11) as the unit circle analogue of a Toda-type transformation from Theorem 2.5.3. Instead of Jacobi parameters and matrices the Verblunsky coefficients $\alpha_n(t)$ and CMV matrices $C(t)$ (9.5.1)–(9.5.3) appear on the central stage. So (9.8.11) plays the same role in the theory of discrete integrable systems as the Toda transformation. The result below is in Golinskiĭ (2006).

Theorem 9.8.9 (Schur flows) *Let* $\mu(\cdot, t)$ *be a family of measures that depend on a real parameter* $t \geq 0$, *with Verblunsky coefficients* $\alpha_n(t)$. *The following three statements are equivalent:*

(i) $\mu(\cdot,t)$ *satisfy* (9.8.11)*;*

(ii) $\alpha_n(t)$ *solve the system of differential-difference equations*

$$\frac{d}{dt}\alpha_n(t) = (1-|\alpha_n(t)|^2)(\alpha_{n+1}(t)-\alpha_{n-1}(t)), \quad t>0 \tag{9.8.13}$$

known as the Schur flow;

(iii) *the CMV matrices* $C(t)$ *satisfy the Lax equation*

$$\frac{d}{dt}C(t) = [A,C], \tag{9.8.14}$$

where $A(t)$ *is an upper-triangular and tridiagonal matrix*

$$A = \begin{pmatrix} \Re\overline{\alpha}_0 & \rho_0\overline{\Delta}_0 & \rho_0\rho_1 & & & \\ & -\Re\overline{\alpha}_1\alpha_0 & \rho_1\Delta_1 & \rho_1\rho_2 & & \\ & & -\Re\overline{\alpha}_2\alpha_1 & \rho_2\overline{\Delta}_2 & \rho_2\rho_3 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{9.8.15}$$

where $\Delta_n = \alpha_{n+1}(t) - \alpha_{n-1}(t)$.

There is an equivalent form of (9.8.14):

$$\frac{d}{dt}C(t) = [B,C], \tag{9.8.16}$$

$$\begin{aligned} B = \frac{(C+C^*)_+ - (C+C^*)_-}{2} &= \frac{1}{2}\begin{pmatrix} 0 & \rho_0\overline{\Delta}_0 & \rho_0\rho_1 & & & \\ -\rho_0\Delta_0 & 0 & \rho_1\Delta_1 & \rho_1\rho_2 & & \\ -\rho_0\rho_1 & -\rho_1\overline{\Delta}_1 & 0 & \rho_2\overline{\Delta}_2 & \rho_2\rho_3 & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \\ &= A - \frac{C+C^*}{2} = -B^*, \end{aligned} \tag{9.8.17}$$

where we use the standard notation $X_\pm$ for the upper (lower) projection of a matrix X. The latter form of the Lax equation is closer to its counterpart in the Toda lattices setting.

Thereby, the solution of the initial–boundary value problem for the Schur flow (9.8.13) with arbitrary initial data

$$|\alpha_n(0)| < 1, \quad n \in \mathbb{Z}_+, \quad \alpha_{-1} \equiv -1 \tag{9.8.18}$$

amounts to a combination of the direct and inverse spectral problems for the unit circle (from Verblunsky coefficients to orthogonality measures and backwards) with (9.8.11) in between.

The Schur flow (9.8.13) appeared in Ablowitz and Ladik (1976a,b) under the name "discrete modified KdV equation," as a spatial discretization of the modified Korteweg–de Vries equation

$$\partial_t f = 6f^2\partial_x f - \partial_x^3 f.$$

The name "Schur flow" is suggested in Faybusovich and Gekhtman (1999), where the authors consider finite real Schur flows and suggest two more Lax equations based upon the Hessenberg matrix representation of the multiplication operator (see also Ammar and Gragg, 1994). In Mukaihira and Nakamura (2000, 2002) the Bessel modification of measures appeared, and a part of the results from Theorem 9.8.9 is proved. In Nenciu (2005) (see also Killip and Nenciu, 2005) the authors deal with the Poisson structure and Lax pairs for the Ablowits–Ladik systems closely related to the Schur flows. The latter can also be viewed as the zero-curvature equation for the Szegő matrices (cf. Geronimo, Gesztesy, and Holden, 2005)

$$\frac{d}{dt}T_n(z,t) + T_n(z,t)W_n(z,t) - W_{n+1}(z,t)T_n(z,t) = 0,$$

$$W_n(z,t) := \begin{pmatrix} z+1-\alpha_{n-1}\overline{\alpha}_n & -\overline{\alpha}_n - \overline{\alpha}_{n-1}z^{-1} \\ -\alpha_{n-1}z - \alpha_n & 1-\overline{\alpha}_{n-1}\alpha_n + z^{-1} \end{pmatrix}.$$

It might be worth pointing out that some properties of Verblunsky coefficients for the Bessel transformed measures (such as the rate of decay) are inherited from those of the initial data (see Golinskiĭ, 2006).

Theorem 9.8.10 *Let $\alpha_n(t)$ solve the Schur flow equations* (9.8.13). *Then*

(i) $\{\alpha_n(0)\} \in \ell^p$ *implies* $\{\alpha_n(t)\} \in \ell^p$ *for all* $t > 0$, $p = 1, 2$;
(ii) $|\alpha_n(0)| \le Ke^{-an}$ *implies* $|\alpha_n(t)| \le K(t)e^{-an}$ *for all* $t > 0$, $\alpha > 0$.

Because of the boundary condition $\alpha_{-1} \equiv -1$ the initial–boundary value problem (9.8.13)/(9.8.18) with zero initial conditions

$$\alpha_0(0) = \alpha_1(0) = \cdots = 0$$

has a nontrivial solution. We are now dealing with the Bessel transformation of the Lebesgue measure

$$\mu(\zeta,t) = C(t)e^{t(\zeta+\zeta^{-1})}\,dm,$$

called the **modified Bessel measures** on the unit circle. Denote by $\beta_n(t)$ the Verblunsky coefficients of $\mu(\cdot,t)$, which are clearly real. The corresponding system of orthogonal polynomials has arisen from studies of the length of longest increasing subsequences of random words (Baik, Deift, and Johansson, 1999) and matrix models (Periwal and Shevitz, 1990).

Note first that the normalizing constant $C(t)$ can be easily computed

$$C^{-1}(t) = \int_{\mathbb{T}} e^{t(\zeta+\zeta^{-1})}\,dm = \frac{1}{2\pi}\int_0^{2\pi} e^{2t\cos x}\,dx = \sum_{n=0}^{\infty}\frac{t^{2n}}{(n!)^2} = I_0(2t),$$

where I_k is the modified Bessel function of order k. Similarly, for the moments of the measure we have

$$\mu_p(t) = \int_{\mathbb{T}} \zeta^{-p}\,d\mu(\zeta,t) = \frac{I_p(2t)}{I_0(2t)}, \quad p \in \mathbb{Z}_+, \quad \mu_{-p} = \mu_p. \tag{9.8.19}$$

The explicit expression for Verblunsky coefficients as a ratio of two determinants follows from (9.1.7) with $z = 0$ and (9.8.19),

$$\beta_n(t) = (-1)^n \frac{\det \|I_{k-j-1}(2t)\|_{0\le k,j\le n}}{\det \|I_{k-j}(2t)\|_{0\le k,j\le n}}, \quad n \in \mathbb{Z}_+. \tag{9.8.20}$$

There is an important feature of the modified Bessel measures proved in Periwal and Shevitz (1990).

Theorem 9.8.11 (Periwal–Shevitz) *The Verblunsky coefficients $\beta_n(t)$ for the modified Bessel measures satisfy a form of the discrete Painlevé II equation*

$$-(n+1)\frac{\beta_n(t)}{t(1-\beta_n^2(t))} = \beta_{n+1}(t) + \beta_{n-1}(t), \quad n \in \mathbb{Z}_+, \tag{9.8.21}$$

with $\beta_{-1} = -1$, $\beta_0 = I_1(2t)/I_0(2t)$.

There are also differential relations satisfied by modified Bessel polynomials, their leading coefficients, and Verblunsky coefficients, specific for this particular case. For instance (see Ismail, 2005b, Lemma 8.3.6),

$$\begin{aligned}\frac{2}{\kappa_n(t)}\frac{d}{dt}\kappa_n(t) &= \frac{I_1(t)}{I_0(t)} + \alpha_n(t)\alpha_{n-1}(t),\\ \frac{d}{dt}\alpha_n(t) &= \frac{I_1(t)}{I_0(t)} + \alpha_{n+1}(t) - (1-|\alpha_n(t)|^2)\alpha_{n-1}(t).\end{aligned}$$

Concerning the long-time behavior of Verblunsky coefficients, the following result is proved in Simon (2007b).

Theorem 9.8.12 *Let $\mu(\cdot, t)$ be the Bessel transformation* (9.8.11)*. Suppose that $\mu(\zeta, 0) = w(\zeta)dm$, w is a positive and continuous function on $\mathbb{T}$. Then for the Verblunsky coefficients one has*

$$(-1)^n\alpha_n(t) = 1 - \frac{n+1}{4t} + O\left(\frac{1}{t}\right), \quad t \to \infty. \tag{9.8.22}$$

In particular, (9.8.22) holds for $\beta_n(t)$. It might be a challenging problem to find an asymptotic series expansion for β_n from (9.8.20) and the expansion for the modified Bessel function

$$I_k(t) \simeq \frac{e^t}{\sqrt{2\pi t}} \sum_{j=0}^{\infty} (-1)^j \frac{(4k^2-1^2)\cdots(4k^2-(2j-1)^2)}{j!(8t)^j}, \quad t \to \infty.$$