Orthogonal Polynomials: from Jacobi to Simon *

Leonid Golinskii and Vilmos Totik

June 29, 2005

Contents

*Supported by INTAS 03-51-6637 (first author) and by NSF grant DMS-040650 and OTKA T049448 (second author)

1 Introduction

Originally we were asked to write two separate papers. One on Barry Simon's work, and one on the state of the art in the theory of orthogonal polynomials. However, Simon's work on orthogonal polynomials is so fresh and fundamental that it will be for quite some while the state of the art of the theory on the unit circle. This conflict could have only resolved in a joint article.

This work is meant to non-experts, and therefore it contains some introductory material. We tried to list most of the actively researched fields, but because of space limitation we had one or two pages for areas where dozens of papers and several books had been published. As a result our account is necessarily incomplete.

The connection of orthogonal polynomials with other branches of mathematics is truly impressive, without even trying to be complete we mention continued fractions, operator theory (Jacobi operators), moment problems, analytic functions (Bieberbach's conjecture), interpolation, Padé approximation, quadrature, approximation theory, numerical analysis, electrostatics, statistical quantum mechanics, special functions, number theory (irrationality and transcendence), graph theory (matching numbers), combinatorics, random matrices, stochastic processes (birth and death processes; prediction theory), data sorting and compression, Radon transform and computer tomography.

The theory of orthogonal polynomials can be divided into two main but only loosely related parts. The two parts have many things in common, and the division line is quite blurred, it is more or less along algebra vs. analysis. One of the parts is the algebraic aspect of the theory, which has close connections with special functions, combinatorics and algebra, and it is mainly devoted to concrete orthogonal systems or hierarchies of systems such as the Jacobi, Hahn, Askey-Wilson, ... polynomials. All the discrete polynomials and the q-analogues of classical ones belong to this theory. We will not treat this part; the interested reader can consult the three recent excellent monographs [39], [27] and [4]. Much of the present state theory of orthogonal polynomials of several variables lies also close to this algebraic part of the theory. To discuss them would take us too far from our main direction, rather we refer the reader to the recent book [23].

The other part is the analytical aspect of the theory. Its methods are analytical, and it deals with questions that are typical in analysis, or questions that have emerged in and related to other parts of mathematical analysis. General properties fill a smaller part of the analytic theory, and the greater part falls into two main and extremely rich branches: orthogonal polynomials on the real line and on the circle. The richness is due to some special features of the real line and the circle. Classical real orthogonal polynomials, sometimes in other forms like continued fractions, can be traced back to the 18th century, but their rapid development occurred in the 19th and early 20th century. Orthogonal polynomials on the unit circle is much younger, and their existence is largely due to Szegő and Geronimus in the first half of the 20th century. Simon's recent treatise [79, 80] (see also [90]) summarizes and greatly extends what has happened since then.

The organization of the present article is as follows. First, in Part I we give a brief outline of general and real orthogonal polynomials. Then we elaborate on some recent trends and the state of the art of this branch of the analytic theory. Simon's contributions to real orthogonal polynomials will be mentioned in this part. After that, in Part II, we move to orthogonal polynomials on the circle, and, finally, Part III lists many of Simon's contributions.

Each of us has opted for his own style of exposition. Part I, prepared by the second author, deal mostly with the state of art in orthogonal polynomials and covers areas/results that are from a period of over 100 years and from a large number of people, therefore the style there is somewhat informal. In contrast, Part III discusses mostly achievements of Barry Simon, and there more formal statements are given.

Part I
General Theory

2 Orthogonal polynomials

Orthogonal polynomials with respect to measures

Let μ be a positive Borel measure on the complex plane with infinite support for which

$$\int |z|^m d\mu(z) < \infty$$

for all $m > 0$. There are unique polynomials

$$p_n(z) = p_n(\mu, z) = \kappa_n z^n + \cdots, \qquad \kappa_n > 0,\ n = 0, 1, \ldots$$

which form an orthonormal system in $L^2(\mu)$, i.e.

$$\int p_m \overline{p_n} d\mu = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n. \end{cases}$$

These p_n's are called the orthonormal polynomials corresponding to μ. κ_n is the leading coefficient, and $p_n(z)/\kappa_n = z^n + \cdots$ is called the monic orthogonal polynomial. The leading coefficients play a special and important role in the theory, many properties depend on their behavior. When $d\mu(x) = w(x)dx$ on some interval, say, then we talk about orthogonal polynomials with respect to the weight function w.

The p_n's can be easily generated: all we have to do is to make sure that

$$\int \frac{p_n(z)}{\kappa_n} \overline{z^k} d\mu(z) = 0, \qquad k = 0, 1, \ldots, n-1,$$

which is an $n \times n$ system of equations for the coefficients of $p_n(z)/\kappa_n$ with matrix $(\sigma_{i,j})_{i,j=0}^{n-1}$, where

$$\sigma_{i,j} = \int z^i \overline{z^j} d\mu(z)$$

are the complex moments of μ. This matrix is nonsingular, so the system has a unique solution, and finally κ_n comes from normalization.

In particular, the complex moments already determine the polynomials. In terms of them one can write up explicit determinant formulas:

$$p_n(z) = \frac{1}{\sqrt{D_{n-1} D_n}} \begin{vmatrix} \sigma_{0,0} & \sigma_{0,1} & \cdots & \sigma_{0,n-1} & 1 \\ \sigma_{1,0} & \sigma_{1,1} & \cdots & \sigma_{1,n-1} & z \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{n-1,0} & \sigma_{n-1,1} & \cdots & \sigma_{n-1,n-1} & z^{n-1} \\ \sigma_{n,0} & \sigma_{n,1} & \cdots & \sigma_{n,n-1} & z^n \end{vmatrix} \tag{2.1}$$

where

$$D_n = |\sigma_{i,j}|_{i,j=0}^n \tag{2.2}$$

are the so called Gram determinants.

Note that if μ is supported on the real line then

$$\sigma_{i,j} = \int x^{i+j} d\mu(x) =: \alpha_{i+j},$$

so $D_n = |\alpha_{i+j}|_{i,j=0}^n$ is a Hankel determinant, while if μ is supported on the unit circle then

$$\sigma_{i,j} = \int z^{i-j} d\mu(z) =: \beta_{i-j},$$

so $D_n = |\beta_{i-j}|_{i,j=0}^n$ is a Toeplitz determinant. In these two important cases the orthogonal polynomials have many special properties that are missing in the general theory. For example, in the real case, i.e., if μ is supported on the real line, the p_n's obey a three term recurrence formula

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1} p_{n-1}(x), \tag{2.3}$$

where

$$a_n = \frac{\kappa_n}{\kappa_{n+1}} > 0, \qquad b_n = \int x p_n^2(x) d\mu(x),$$

and conversely, any system of polynomials satisfying (2.3) with real $a_n > 0$, b_n is an orthonormal system with respect to a (not necessarily unique) measure on the real line (Favard' theorem). In the real case the zeros of p_n are real and simple and the zeros of p_n and p_{n+1} interlace, i.e. in between any two zeros of p_{n+1} there is a zero of p_n. We emphasize that the three term recurrence is a very special property of real orthogonal polynomials, and it is due to the fact that in this case the polynomials are real, hence

$$\int x p_n(x) \overline{p_m(x)} d\mu(x) = \int p_n(x)(x p_m(x)) d\mu(x) = 0$$

for $m < n-1$. This three term recurrence is missing in the general case, and it is replaced by a different recurrence for polynomials on the circle (see Part II). For example, in the real case the three term recurrence implies for the reproducing kernel the Christoffel-Darboux formula

$$\sum_{k=0}^n p_k(x) p_k(t) = \frac{\kappa_n}{\kappa_{n+1}} \frac{p_{n+1}(x) p_n(t) - p_n(x) p_{n+1}(t)}{x - t}.$$

The starting values of the recurrence (2.3) are $p_{-1} \equiv 0$, $p_0 = (\mu(\mathbf{C})))^{-1/2}$. If one starts from $q_{-1} = -1$, $q_0 \equiv 0$ and use the same recurrence (with $a_{-1} = 1$)

$$xq_n(x) = a_n q_{n+1}(x) + b_n q_n(x) + a_{n-1} q_{n-1}(x), \tag{2.4}$$

then q_n is of degree $n-1$, and by Favard's theorem the different q_n's are orthogonal with respect to some measure. The q_n's are called orthogonal polynomials of the second kind (sometimes for p_n we say that they are of the first kind). They can also be written in the form

$$q_n(z) = (\mu(\mathbf{C}))^{-1/2} \int \frac{p_n(z) - p_n(x)}{z - x} d\mu(x).$$

The Riemann-Hilbert approach

Let μ still be supported on the real line, and suppose that it is of the form $d\mu(t) = w(t)dt$ with some smooth function w. A new approach to generating orthogonal polynomials that has turned out to be of great importance was given in the early 1990's by Fokas, Its and Kitaev [26]. Consider 2×2 matrices

$$Y(z) = \begin{pmatrix} Y_{11}(z) & Y_{12}(z) \\ Y_{21}(z) & Y_{22}(z) \end{pmatrix}$$

where the Y_{ij} are analytic functions on $\mathbf{C} \setminus \mathbf{R}$, and solve for such matrices the following matrix valued Riemann-Hilbert problem:

1. for all $x \in \mathbf{R}$

$$Y_+(x) = Y_-(x) \begin{pmatrix} 1 & m(x) \\ 0 & 1 \end{pmatrix}$$

where Y_+ resp. Y_- is the limit of $Y(z)$ as z tends to x from the upper resp. lower half plane, and

2.

$$Y(z) = \left(I + O\left(\frac{1}{z}\right)\right) \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix}$$

at infinity, where I denotes the identity matrix.

There is a unique solution $Y(z)$, and its entry $Y_{11}(z)$ is precisely the monic polynomial $p_n(\mu, z)/\kappa_n$. The other entries can also be explicitly written in terms of p_n and p_{n-1}, furthermore κ_n and the recurrence coefficients a_n, b_n can be expressed from the entries of Y_1, where Y_1 is the matrix in

$$Y(z) \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} = I + Y_1 \frac{1}{z} + O\left(\frac{1}{z^2}\right).$$

For details on this Riemann-Hilbert approach see [19] or [18] in this volume.

Orthogonal polynomials with respect to inner products

Sometimes one talks about orthogonal polynomials with respect to an inner product $\langle \cdot, \cdot \rangle$ which is defined on some linear space containing all polynomials. In this case orthogonality means $\langle p_n, p_m \rangle = 0$ for $m \neq n$. When the inner product is positive definite in the sense that $\langle p, p \rangle = 0$ only for the zero polynomial

p, then the aforementioned orthogonalization process can be used, and with $\sigma_{i,j} = \langle x^i, x^j \rangle$, the determinantal formula (2.1) is still valid. The same is true if the Gram determinants (2.2) are different from zero. However, if this is not so, e.g. in the so called non-Hermitian orthogonality (see section 11), then these cannot be used. In this case we write

$$p_n(z) = \gamma_n z^n + \gamma_{n-1} z^{n-1} + \cdots,$$

and make sure that p_n is orthogonal to all powers z^k, $0 \le k < n$, i.e., solve the homogeneous system of equations

$$\sum_{j=0}^{n} \gamma_j \sigma_{j,k} = 0, \qquad k = 0, \ldots, n-1$$

for $\gamma_0, \gamma_1, \ldots, \gamma_n$. Since the number of unknowns is bigger than the number of equations, there is always a non-trivial solution, which gives rise to non-trivial orthogonal polynomials. However, we cannot assert any more $\gamma_n \neq 0$, so the degree of p_n may be smaller than n, and there may be several choices for p_n. Still, in applications where non-Hermitian orthogonality is used, these p_n play the role of orthogonal polynomials.

Varying weights

In the last 25 years orthogonal polynomials with respect to varying measures have played significant role in several problems, see e.g. the sections on exponential and Freud weights or on random matrices in Section 4. In forming them one has a sequence of measures μ_n (generally with some particular behavior), and for each n one forms the orthogonal system $\{p_k(\mu_n, z)\}_{k=0}^{\infty}$. In most cases one needs the behavior of $p_n(\mu_n, z)$ or that of $p_{n\pm k}(\mu_n, z)$ with some fixed k.

Matrix orthogonal polynomials

Orthogonality of matrix polynomials (i.e. when the entries of the fixed size matrix are polynomials of degree $n = 0, 1, \ldots$ and orthogonality is with respect to a matrix measure) is a very active area which shows extreme richness compared to the scalar case. See section 13 for a short discussion.

3 Classical orthogonal polynomials

These are

- Jacobi polynomials $P_n^{(\alpha,\beta)}$, $\alpha, \beta > -1$, orthogonal with respect to the weight $(1-x)^\alpha (1+x)^\beta$ on $[-1, 1]$,

- Laguerre polynomials $L_n^{(\alpha)}$, $\alpha > -1$, with orthogonality weight $x^\alpha e^{-x}$ on $[0, \infty)$,
- and Hermite polynomials H_n orthogonal with respect to e^{-x^2} on $(-\infty, \infty)$.

In the literature various normalizations are used for them.

They are very special, for they possess many properties that no other orthogonal polynomial system does. In particular,

- they have derivatives which form again an orthogonal polynomial system, e.g. the derivative of $P_n^{(\alpha,\beta)}$ is a constant multiple of $P_{n-1}^{(\alpha+1,\beta+1)}$:

$$(P_n^{(\alpha,\beta)})'(x) = \frac{1}{2}(n + \alpha + \beta + 1)P_{n-1}^{(\alpha+1,\beta+1)}(x),$$

- they all possess a Rodrigue's type formula

$$P_n(x) = \frac{1}{d_n w(x)} \frac{d^n}{dx^n}\{w(x)\sigma(x)^n\},$$

  where w is the weight function and σ is a polynomial that is independent of n, for example,

$$L_n^{(\alpha)}(x) = e^x x^{-\alpha} \frac{1}{n!} \frac{d^n}{dx^n} \left(e^{-x} x^{n+\alpha}\right),$$

- they satisfy a differential-difference relation of the form

$$\pi(x)P_n'(x) = (\alpha_n x + \beta_n)P_n(x) + \gamma_n P_{n-1}(x),$$

  e.g.

$$x(L_n^{(\alpha)})'(x) = nL_n^{(\alpha)}(x) - (n + \alpha)L_{n-1}^{(\alpha)}(x),$$

- they satisfy a non-linear equation of the form

$$\sigma(x)\left(P_n(x)P_{n-1}(x)\right)' = (\alpha_n x + \beta_n)P_n(x)P_{n-1}(x) + \gamma_n P_n^2(x) + \delta_n P_{n-1}^2(x),$$

  with some constants $\alpha_n, \beta_n, \gamma_n, \delta_n$, and σ a polynomial of degree at most 2, e.g.

$$(H_n(x)H_{n-1}(x))' = 2xH_n(x)H_{n-1}(x) - H_n^2(x) + 2nH_{n-1}^2(x).$$

Now every one of these has a converse, namely if a system of orthogonal polynomials possesses any of these properties, then it is (up to a change of variables) one of the classical systems [1]. See also Bochner's result in the next section claiming that the classical orthogonal polynomials are essentially the only polynomial (not just orthogonal polynomial) systems that satisfy a certain second order differential equation.

Classical orthogonal polynomials are also special in the sense that they possess a relatively simple

- second order differential equation, e.g.

$$xy'' + (\alpha + 1 - x)y' + ny = 0$$

  for $L_n^{(\alpha)}$,

- generating function, e.g.

$$\sum_n \frac{H_n(x)}{n!} w^n = \exp(2xw - w^2),$$

- integral representation, e.g.

$$(1-x)^\alpha (1+x)^\beta P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^{n+1}\pi i} \int (1-t)^{n+\alpha}(1+t)^{n+\beta}(t-x)^{-n-1} dt$$

  over an appropriate contour,

and these are powerful tools to study their behavior.

For all these results see [104].

4 Where do orthogonal polynomials come from?

In this section we mention a few selected areas where orthogonal polynomials naturally arise.

Continued fractions

Continued fractions played extremely important role in the development of several branches of mathematics, but their significance has been unjustly diminished in modern mathematics. A continued fraction is of the form

$$\frac{B_1}{A_1 + \frac{B_2}{A_2 + \cdots}},$$

and its n-th convergent is

$$\frac{S_n}{R_n} = \frac{B_1}{A_1 + \frac{B_2}{A_2 + \cdots \frac{B_n}{A_n}}}, \quad n = 1, 2, \ldots.$$

The value of the continued fraction is the limit of its convergents. The denominators and numerators of the convergents satisfy the three term recurrence relations

$$\begin{aligned} R_n &= A_n R_{n-1} + B_n R_{n-2}, \qquad R_0 \equiv 1,\ R_{-1} \equiv 0 \\ S_n &= A_n S_{n-1} + B_n S_{n-2}, \qquad S_0 \equiv 0,\ S_{-1} \equiv 1, \end{aligned}$$

which immediately connects continued fractions with 3 term recurrences and hence with orthogonal polynomials.

But the connection is deeper than just this formal observation. Many elementary functions (like $z-\sqrt{z^2-1}$) have a continued fraction development where the B_n's are constants while the A_n's are linear functions, in which case the convergents are ratios of some orthogonal polynomials of the second and first kind. An important example is that of Cauchy transforms of measures μ with compact support on the real line (so called Markov functions):

$$f(z)=\int \frac{d\mu(x)}{x-z}=-\frac{\alpha_0}{z}-\frac{\alpha_1}{z^2}-\dots. \tag{4.1}$$

The coefficients α_j in the development of (4.1) are the moments

$$\alpha_j=\int x^j d\mu(x), \quad j=0,1,\dots$$

of the measure μ. The continued fraction development

$$f(z)\sim \frac{B_1}{z-A_1+\frac{B_2}{z-A_2+\cdots}}$$

of f at infinity converges locally uniformly outside the smallest interval that contains the support of μ (A. Markov's theorem).

As has been mentioned, the numerators $S_n(z)$ and the denominators $R_n(z)$ of the n-th convergents

$$\frac{S_n(z)}{R_n(z)}=\frac{B_1}{z-A_1+\frac{B_2}{z-A_2+\cdots\frac{B_n}{z-A_n}}}, \quad n=1,2,\dots$$

satisfy the recurrence relations

$$\begin{aligned} R_n(z) &= (z-A_n)R_{n-1}(z)+B_nR_{n-2}(z), \quad R_0\equiv 1,\ R_{-1}\equiv 0 \\ S_n(z) &= (z-A_n)S_{n-1}(z)+B_nS_{n-2}(z), \quad S_0\equiv 0,\ S_{-1}\equiv 1. \end{aligned} \tag{4.2}$$

These are precisely the recurrence formulae for the monic orthogonal polynomials of the first and second kind with respect to μ, hence the n-th convergent is $cq_n(z)/p_n(z)$ with $c=\mu(\mathbf{C})^{1/2}$.

See [104, pp. 54–57] as well as [45] and the numerous references there.

Padé approximation and rational interpolation

With the preceding notation the rational function $S_n(z)/R_n(z)=cq_n(z)/p_n(z)$ with $c=\mu(\mathbf{C})^{1/2}$ of numerator degree $n-1$ and of denominator degree n interpolates $f(z)$ at infinity in the order $2n$. This is the analogue (called $[n-1/n]$ Padé approximation) of the n-th Taylor polynomial (which interpolates the function in the order n) for rational functions. The advantage of Padé approximation over Taylor polynomials lies in the fact that the poles of Padé approximants may imitate the singularities of the function in question, while Taylor polynomials are good only up to the first singularity. The error in $[n-1/n]$ Padé approximation has the form

$$f(z) - c\frac{q_n(z)}{p_n(z)} = \frac{1}{p_n^2(z)} \int \frac{p_n^2(x)}{x-z} d\mu(x).$$

Orthogonal polynomials appear in more general rational interpolation (called multipoint Padé approximation) to Markov functions, see e.g. [98, Sec. 6.1].

Moment problem

The moments of a measure μ, $\mu(\mathbf{C}) = 1$, supported on the real line, are

$$\alpha_n = \int x^n d\mu(x), \qquad n = 0, 1, \ldots.$$

The Hamburger moment problem is to determine if a sequence $\{\alpha_n\}$ (with normalization $\alpha_0 = 1$) of real numbers is the moment sequence of a measure with infinite support, and if this measure is unique (the Stieltjes moment problem asks the same, but for measures on $[0, \infty)$). The existence is easy: $\{\alpha_n\}$ are the moments of some measure supported on $\mathbf{R}$ if and only if all the Hankel determinants $|\alpha_{i+j}|_{i,j=0}^m$, $m = 0, 1, \ldots$ are positive. The unicity (usually called determinacy) depends on the behavior of the orthogonal polynomials (2.1) defined from the moments $\sigma_{i,j} = \alpha_{i+j}$. In fact, there are different measures with the same moments α_j if and only if there is a non-real z_0 with $\sum_n |p_n(z_0)|^2 < \infty$, which in turn is equivalent to $\sum_n |p_n(z)|^2 < \infty$ for all $z \in \mathbf{C}$. Furthermore, the Cauchy transforms of all solutions ν of the moment problem have the parametric form

$$\int \frac{d\nu(x)}{z-x} = \frac{C(z)F(z) + A(z)}{D(z)F(z) + B(z)},$$

where F is an arbitrary analytic function (the parameter) mapping the upper half plane $\mathbf{C}_+$ into $\overline{\mathbf{C}}_+ \cup \{\infty\}$, and A, B, C and D have explicit representation in terms of the first and second kind orthogonal polynomials p_n and q_n:

$$A(z) = z \textstyle\sum_n q_n(0) q_n(z); \qquad B(z) = -1 + z \textstyle\sum_n q_n(0) p_n(z);$$

$$C(z) = 1 + z \textstyle\sum_n p_n(0) q_n(z); \quad D(z) = z \textstyle\sum_n p_n(0) p_n(z).$$

For all these results and for an operator theoretic approach to the moment problem see Simon's survey [81] (in particular, Theorems 3 and 4.14).

Jacobi matrices and spectral theory of self-adjoint operators

Tridiagonal, so called Jacobi matrices

$$J = \begin{pmatrix} b_0 & a_0 & 0 & 0 & \cdots \\ a_0 & b_1 & a_1 & 0 & \cdots \\ 0 & a_1 & b_2 & a_2 & \cdots \\ 0 & 0 & a_2 & b_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

with bounded $a_n > 0$ and bounded real b_n define a self-adjoint bounded operator J in l_2, a so called Jacobi operator. These are the discrete analogue of second order linear differential operators of Schrödinger type on the half line. Every self adjoint operator with a cyclic vector is a Jacobi operator in an appropriate base.

The formal eigen-equation $J\pi = \lambda\pi$ is equivalent to the three term recurrence

$$a_{n-1}\pi_{n-1} + b_n\pi_n + a_n\pi_{n+1} = \lambda\pi_n, \qquad n = 1, 2, \ldots$$

$$b_0\pi_0 + a_0\pi_1 = \lambda\pi_0, \qquad \pi_0 = 1.$$

Thus, $\pi_n(\lambda)$ is of degree n in λ.

By the spectral theorem, J, as a self-adjoint operator having a cyclic vector $((1, 0, 0, \ldots))$ is unitarily equivalent to multiplication by x in some $L^2(\mu)$ with μ having compact support on the real line. More precisely, if $p_n(x) = p_n(\mu, x)$ are the orthonormal polynomials with respect to μ, and U maps the unit vector $e_n = (0, \ldots, 0, 1, 0, \ldots)$ into p_n, then U can be extended into a unitary operator from l_2 onto $L^2(\mu)$, and if $Sf(x) = xf(x)$ is the multiplication operator by x in $L^2(\mu)$, then $J = U^{-1}SU$. The recurrence coefficients for $p_n(\mu, x)$ are precisely the a_n's and b_n's from the Jacobi matrix, i.e. $p_n(x) = c\pi_n(x)$ with some fixed constant c. These show that Jacobi operators are equivalent to multiplication by x in $L^2(\mu)$ spaces if the particular basis $\{p_n(\mu)\}$ are used (see e.g. [18, Ch. 2]).

The truncated $n \times n$ matrix

$$J_n = \begin{pmatrix} b_0 & a_0 & 0 & 0 & \cdots \\ a_0 & b_1 & a_1 & 0 & \cdots \\ 0 & a_1 & b_2 & a_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & 0 & a_{n-2} & b_{n-1} \end{pmatrix}$$

has n real and distinct eigenvalues, which turn out to be the zeros of p_n, i.e. the monic polynomial $p_n(z)/\kappa_n$ is the characteristic polynomial of J_n.

Quadrature

For a measure μ an n-point quadrature is a set of points $x_1, \dots, x_n$ and a set of associated numbers $\lambda_1, \dots, \lambda_n$. It is expected that

$$\int f d\mu \sim \sum_{k=1}^{n} \lambda_k f(x_k)$$

in some sense for as large class of functions as possible. Often the accuracy of the quadrature is measured by its exactness, which is defined as the largest m such that the quadrature is exact for all polynomials of degree at most m, i.e. m is the largest number with the property that

$$\int x^j d\mu(x) = \sum_{k=1}^{n} \lambda_k x_k^j \qquad \text{for all } 0 \le j \le m.$$

For μ with support on the real line and for quadrature based on n points this exactness m cannot be larger than $2n-1$, and this optimal value $2n-1$ is attained if and only if $x_1, \dots, x_n$ are the zeros of the orthonormal polynomial $p_n(\mu, x)$ corresponding to μ, and the so called Cotes numbers λ_k are chosen to be

$$\lambda_k = \left(\sum_{j=0}^{n} p_j(\mu, x_k)^2 \right)^{-1}.$$

See [104, Ch. XV].

Random matrices

Some statistical-mechanical models in quantum systems use random matrices. Let $\mathcal{H}_n$ be the set of all $n \times n$ Hermitian matrices $M = (m_{i,j})_{i,j=1}^n$, and let there be given a probability distribution on $\mathcal{H}_n$ of the form

$$P_n(M)dM = D_n^{-1} \exp(-n\mathrm{Tr}\{V(M)\})dM,$$

where $V(\lambda)$, $\lambda \in \mathbf{R}$, is a real-valued function that increases sufficiently fast at infinity (typically an even polynomial in quantum field theory applications), $\mathrm{Tr}\{H\}$ denotes the trace of the matrix H,

$$dM = \prod_{k=1}^{n} dm_{k,k} \prod_{k<j} d\,\Re m_{k,j}\; d\,\Im m_{k,j}$$

is the "Lebesgue" measure for Hermitian matrices, and D_n is a normalizing constant so that the total integral of $P_n(M)dM$ is one.

Every matrix $M \in \mathcal{H}_n$ has n real eigenvalues which carry physical information on the system when it is in the state described by M. The quantity

$$N_n(\mathbf{D}) = \frac{\#\{\text{eigenvalues in } \mathbf{D}\}}{n}$$

is the random variable that equals the normalized number of eigenvalues in the interval $\mathbf{D}$. This model is known as the unitary ensemble associated with V.

Let $p_j(w^n, x)$ be the orthonormal polynomials with respect to the varying weight $w^n(x)$, $w(x) = \exp(-V(x))$. Then the joint probability density of the eigenvalues can be written in the form

$$d_n \left| p_{i-1}(w^n, \lambda_j) w^{n/2}(\lambda_j) \right|^2_{1 \leq i,j \leq n},$$

where d_n is a normalizing constant built up from the leading coefficients of the $p_j(w^n, \cdot)$. With the so called weighted reproducing kernel

$$K_n(t,s) = \sum_{j=0}^{n-1} p_j(w^n, t) w^{n/2}(t)\, p_j(w^n, s) w^{n/2}(s)$$

it can also be written in the form

$$\frac{1}{n!} \left| K_n(\lambda_i, \lambda_j) \right|_{1 \leq i,j \leq n}.$$

In particular, for the expected number of eigenvalues in an interval $\mathbf{D}$ we have

$$EN_n(\mathbf{D}) = \int_{\mathbf{D}} \frac{K_n(\lambda, \lambda)}{n}\, d\lambda,$$

where $1/K_n(\lambda, \lambda)$ is known in the theory of orthogonal polynomials as the n-th (weighted) Christoffel function associated with the weight w^n, while the limit of the left hand side (as $n \to \infty$) is known as the density of states.

See, e.g., [63] and [70].

5 Some questions leading to classical orthogonal polynomials

The are almost an infinite number of problems where classical orthogonal polynomials emerge. Let us just mention a few.

Electrostatics

Put to 1 and -1 two positive charges p and q, and with these fixed charges put n positive unit charge on $[-1,1]$ to the points $x_1,\dots,x_n$. On the plane the Coulomb force is proportional with the reciprocal of the distance, and so a charge generates a logarithmic potential field. Therefore, the mutual energy of all these charges is

$$I(x_1,\dots,x_n) = p\sum_{j=1}^{n}\log\frac{1}{|1-x_j|} + q\sum_{j=1}^{n}\log\frac{1}{|1+x_j|} + \sum_{i<j}\log\frac{1}{|x_i-x_j|},$$

and the equilibrium problem asks for finding $x_1,\dots,x_n$ for which this energy is minimal. The unique minimum occurs (see [104, Section 6.7]) for the zeros of the Jacobi polynomials $P_n^{(2p-1,2q-1)}$ orthogonal with respect to the weight $(1-x)^{2p-1}(1+x)^{2q-1}$.

There is a similar characterization of the zeros of Laguerre and Hermite polynomials, and even of more general orthogonal polynomials (for the latter see [39, Section 3.5]).

Polynomial solutions of eigenvalue problems

Consider the eigenvalue problem

$$f(x)\frac{d^2}{dx^2}y(x) + g(x)\frac{d}{dx}y(x) + h(x)y(x) = \lambda y(x),$$

where f, g, h are fixed polynomials and λ is a free constant, and it is required that this has a polynomial solution of exact degree n for all $n = 0, 1, \ldots$, for which the corresponding λ and $y(x)$ will be denoted by λ_n and $y_n(x)$, respectively. Bochner's theorem [13] states that except for some trivial solutions of the form $y(x) = ax^n + bx^m$ and for some polynomials related to Bessel functions, the only solutions are (in all of them we can take $h(x) = 0$)

- Jacobi polynomials $P_n^{(\alpha,\beta)}$ ($f(x) = 1 - x^2$, $g(x) = \beta - \alpha - x(\alpha + \beta + 2)$, $\lambda_n = -n(n + \alpha + \beta + 1)$)
- Laguerre polynomials $L_n^{(\alpha)}$ ($f(x) = x$, $g(x) = 1 + \alpha - x$, $\lambda_n = -n$) and
- Hermite polynomials $H_n(x)$ ($f(x) = 1$, $g(x) = -2x$, $\lambda_n = -2n$).

Harmonic analysis on spheres and balls

Harmonic analysis on spheres and balls in $\mathbf{R}^d$ is based on spherical harmonics, i.e. harmonic homogeneous polynomials. In this theory special Jacobi polynomials, so called ultraspherical or Gegenbauer polynomials $P_n^{(\alpha)}$ play a fundamental role – they are orthogonal with respect to the weight $(1-x^2)^{\alpha-1/2}$.

Let S^{d-1} be the unit sphere in $\mathbf{R}^d$ and let $\mathcal{H}_n^d$ be the restriction to S^{d-1} of all harmonic polynomials $Q(x_1,\ldots,x_n)$ of d variables that are homogeneous of degree n, i.e.

$$\sum_{k=1}^{n} \frac{\partial^2}{\partial x_k^2} Q = 0, \qquad Q(\lambda x_1,\ldots,\lambda x_n) = \lambda^n Q(x_1,\ldots,x_n), \quad \lambda > 0.$$

The dimension of $\mathcal{H}_n^d$ is

$$\binom{n+d-1}{d-1} - \binom{n+d-3}{d-1},$$

and an orthogonal basis in it can be produced as follows. With $\rho = x_{d-1}^2 + x_d^2$ let $g_{s,0} = \rho^s P_s^{(0)}(x_{d-1}/\rho)$ and $g_{s,1} = x_d \rho^s P_s^{(1)}(x_{d-1}/\rho)$. With $n_d = 0$ or $n_d = 1$ consider all multiindices $\mathbf{n} = (n_1, n_2, \ldots, n_d)$ such that $n_1 + \cdots + n_d = n$, and if for such a multiindex we define

$$Y_{\mathbf{n}}(x_1,\ldots,x_d) = g_{n_{d-1},n_d} \prod_{j=1}^{d-2} \left((x_j^2 + \cdots + x_d^2)^{n_j} P_{n_j}^{(\lambda_j)}(x_j (x_j^2 + \cdots + x_d^2)^{-1/2}) \right),$$

then these $Y_{\mathbf{n}}$ constitute an orthogonal basis in $\mathcal{H}_n^d$ (see e.g. [23, p. 35]).

Approximation theory

In the literature expansions of functions into classical orthogonal polynomial series are second only to trigonometric expansions, and numerous works have been devoted to their convergence and approximation properties, see e.g. [104, Ch. XIII].

The Chebyshev polynomials given by $\cos(n \arccos x)$ are orthogonal on $[-1,1]$ with respect to the weight $w(x) = (1-x^2)^{-1/2}$. These directly correspond to trigonometric functions, and expansions into them have virtually the same properties as trigonometric Fourier expansions. But there are many other aspects of approximation where Chebyshev polynomials appear. If one considers e.g. the best approximation on $[-1,1]$ of x^n by polynomials $P_{n-1}(x)$ of smaller degree then the smallest error appears when $x^n - P_{n-1}(x) = 2^{1-n}\cos(n \arccos x)$ is the monic n-th Chebyshev polynomial.

The monic orthogonal polynomials $p_n(\mu)/\kappa_n$ are the solutions to the extremal problem

$$\int |P_n|^2 d\mu \to \min, \tag{5.3}$$

where the minimum is taken for all monic polynomials of degree n. This extremal property makes orthogonal polynomials, in particular Chebyshev polynomials, indispensable tools in approximation theory.

Lagrange interpolation and its various generalizations like Hermite, Hermite-Fejér interpolation etc. is mostly done on the zeros of some orthogonal polynomials. In fact, these nodes are often close to optimal in the sense that the Lebesgue constant increases in the optimal rate. In many cases interpolation on zeros of orthogonal polynomials have special properties due to explicitly calculable expressions. Recall e.g. Fejér's result that if P_{2n-1} is the unique polynomial of degree at most $2n-1$ that interpolates a continuous function f at the nodes of the n-th Chebyshev polynomial and that has zero derivative at each of these nodes, then P_{2n-1} uniformly converges to f on $[-1,1]$ as $n \to \infty$. For the role of orthogonal polynomials in interpolation see the books [103] and [62].

6 Heuristics

In this section we do not state precise results, just want to indicate some heuristics on the behavior of orthogonal polynomials. For the concepts below, as well as for a more precise form of some of the heuristics see the the following sections, in particular section 7.

As we have just seen, the monic orthogonal polynomials $p_n(\mu)/\kappa_n$ minimize the $L^2(\mu)$ norm in (5.3), therefore, the polynomials try to be small where the measure is large, e.g. one expects the zeros to cluster at the support $S(\mu)$ of μ. The example of arc measure on the unit circle, for which the orthogonal polynomials are z^n, shows however, that this is not true (due to the fact that the complement of the support is not connected). The statement is true when the support lies on $\mathbf{R}$ or on some systems of arcs, and also in the general case when instead of the support one considers the polynomial convex hull of the support of μ: on any compact set outside the polynomial convex hull there can only be a fixed number of zeros of $p_n(\mu)$ for every n. When the complement of $S(\mu)$ is connected and $S(\mu)$ has no interior, then the distribution of the zeros shows a remarkable universality and indifference with respect to the size of μ, practically in every situation the distribution of the zeros is the equilibrium distribution of the support $S(\mu)$. When $S(\mu) = [-1,1]$, this means that under very weak assumptions the zero distribution is always the arcsine distribution $dx/\pi\sqrt{1-x^2}$.

The $L^2(\mu)$ minimality of $p_n(\mu)/\kappa_n$ in the sense of (5.3) is something like minimality in L^∞ norm on $S(\mu)$. Therefore, $p_n(\mu)/\kappa_n$ should behave like the monic polynomial T_n minimizing the L^∞ norm on $S(\mu)$ (so called Chebyshev polynomials for $S(\mu)$). Since

$$\frac{1}{n}\log|T_n(z)| = \int \log|z-t|d\nu_n(t)$$

where ν_n has mass $1/n$ at each zero of T_n, in the limit the behavior should be like

$$U^\nu(z) = \int \log|z-t|d\nu(t), \tag{6.1}$$

where ν is the probability measure on $S(\mu)$ for which the maximum of U^ν on $S(\mu)$ is as small as possible (this is the so called equilibrium measure of $S(\mu)$). More generally, if $d\nu = d\nu_n = w^n(x)dx$ is a varying weight in the specified way, then the same reasoning leads to a behavior like (6.1), but now ν is a measure for which the supremum of $U^\nu(z) + \log w(z)$ is as small as possible (weighted equilibrium measure).

Universal behavior can also be seen for the polynomials themselves. Usually they obey

$$\frac{1}{n}\log|p_n(\mu,z)| \to g_{\mathbf{C}\setminus S(\mu)}(z,\infty), \qquad z \notin S(\mu) \tag{6.2}$$

where $g_{\mathbf{C}\setminus S(\mu)}(z,\infty)$ is the Green function with pole at infinity associated with the complement of the support. When the unbounded component of the complement of $S(\mu)$ is simply connected, then in that component often there is a finer asymptotic behavior of $p_n(\mu)$ of the form

$$p_n(z) \sim d_n g_\mu(z)\Phi(z)^n, \qquad z \notin S(\mu) \tag{6.3}$$

where Φ is the mapping function that maps $\mathbf{C}\setminus S(\mu)$ conformally onto the outside of the unit disk, and g_μ is a function (might be called generalized Szegő function) that depends on μ. Such a fine asymptotic is restricted to the simply connected case, see e.g. section 8.

Asymptotics on orthogonal polynomials have a hierarchy, and the different types of asymptotics usually require the measure to be sufficiently strong with different degree on its support. Consider first the case of compact support $S(\mu)$. The weakest is n-th root asymptotics stating the behavior (6.2) for $|p_n(\mu,z)|^{1/n}$ outside the support of the measure. It is mostly equivalent with a corresponding distribution of the zeros, as well as asymptotical minimal behavior of $\kappa_n^{1/n}$. It holds under very weak assumption on the measure, roughly stating that the logarithmic capacity where $\mu' > 0$ (derivative with respect to equilibrium measure) be the same as the capacity of $S(\mu)$. Ratio asymptotics, i.e. asymptotic behavior of $p_{n+1}(\mu,z)/p_n(\mu,z)$ is stronger, and is equivalent with asymptotics for the ratio κ_{n+1}/κ_n of consecutive leading coefficients. It can only hold when $\mathbf{C}\setminus S(\mu)$ (more precisely its unbounded component) is simply connected, and in this case it is enough that $\mu' > 0$ almost everywhere with respect to the equilibrium measure of the support of μ (see section 8). Finally, strong asymptotics of the form (6.3) needs roughly that $\log \mu'$ be integrable (Szegő condition, see section 8).

All these are outside the support. On the support the orthogonal polynomials are of oscillatory behavior, and in the real case under smoothness assumptions on the measure often a so called Plancherel-Rotach type asymptotic formula

$$p_n(\mu, x) \sim d_n g(x) \sin(nh(x) + H(x))$$

holds, where g, h, H are fixed functions. Here $h(x)$ is directly linked with the zeros, h'/π is precisely the distribution of the zeros. When $S(\mu) = [-1, 1]$ and the measure is smooth, then $h(x) = \arccos x$.

When $S(\mu)$ is not of compact support (like Laguerre, Hermite or Freud weights), then usually the zeros are spreading out, and one has to rescale them to $[-1, 1]$ (or to $[0, 1]$) to get a distribution, which is mostly NOT the arcsine distribution. In a similar fashion, various asymptotics hold for the polynomials only after the corresponding rescaling.

7 General orthogonal polynomials

In this section μ is always of compact support $S(\mu)$. For all the results below see [98] and the references there.

The energy $V(K)$ of a compact set K is defined as the infimum of

$$I(\nu) = \int\int \log \frac{1}{|x-t|} d\nu(x) d\nu(t) \tag{7.1}$$

where the infimum is taken for all positive Borel measures on K with total mass 1. The logarithmic capacity is then $\mathrm{cap}(K) = e^{-V(K)}$. For the leading coefficients κ_n of the orthonormal polynomials $p_n(\mu)$ we have

$$\frac{1}{\mathrm{cap}(S(K))} \le \liminf_{n\to\infty} \kappa_n^{1/n}. \tag{7.2}$$

When $\mathrm{cap}(K)$ is positive, then there is a unique measure $\nu = \omega_K$ minimizing the energy in (7.1), and this measure is called the equilibrium measure of K. Green's function $g_{\mathbf{C}\setminus K}(z, \infty)$ with pole at infinity of $\mathbf{C} \setminus K$ can then be defined as

$$g_{\mathbf{C}\setminus K}(z, \infty) = \log \frac{1}{\mathrm{cap}(K)} - \int \log \frac{1}{|z-t|} d\omega_K(t). \tag{7.3}$$

We have for all μ (with $\mathrm{cap}(S(\mu)) > 0$) the estimate

$$\liminf_{n\to\infty} \frac{1}{n} \log |p_n(\mu, z)|^{1/n} \ge g_{\mathbf{C}\setminus S(\mu)}(z, \infty) \tag{7.4}$$

locally uniformly outside the convex hull of $S(\mu)$, while in the convex hull but outside $\mathrm{Pc}(S(\mu))$ (see below) (7.4) is true quasi-everywhere (i.e. with the exception of a set of zero capacity). The same is true on the outer boundary of

$S(\mu)$, which is defined as the boundary $\partial\Omega$ of the unbounded component Ω of the complement $\mathbf{C} \setminus S(\mu)$, namely for quasi-every $z \in \partial\Omega$

$$\liminf_{n\to\infty} |p_n(\mu, z)|^{1/n} \geq 1.$$

All these estimates are sharp.

The zeros of $p_n(\mu)$ lie in the convex hull of $S(\mu)$. When Ω is the unbounded component of the complement $\mathbf{C} \setminus S(\mu)$, then $\mathrm{Pc}(S(\mu)) = \mathbf{C} \setminus \Omega$ is called the polynomial convex hull of $S(\mu)$ (it is the union of $S(\mu)$ with all the "holes" in it, i.e. with the bounded components of $\mathbf{C} \setminus S(\mu)$). Now the zeros cluster on $\mathrm{Pc}(S(\mu))$ in the sense that for any compact subset K of Ω there is a number N_K independent of n, such that $p_n(\mu)$ can have at most N_K zeros in K. For example, if μ is supported on the real line, then $\mathrm{Pc}(S(\mu)) = S(\mu)$, and if K is a closed interval disjoint from the support, then there is at most one zero in K. Denisov and Simon [21] showed that if $x_0 \in \mathbf{R}$ is not in the support, then for some $\delta > 0$ and all n either p_n or p_{n+1} has no zero in $(x_0 - \delta, x_0 + \delta)$. Note that if μ is a symmetric measure on $[-1, -1/2] \cup [1/2, 1]$, then $p_{2n+1}(0) = 0$ for all n, so the result is sharp.

In [21] Denisov and Simon focused on attracting properties of isolated points of the $\operatorname{supp}\mu$. Let z_0 be an isolated point of $S(\mu)$, such that its distance from the convex hull of $S(\mu) \setminus \{z_0\}$ is $\delta > 0$. Then p_n has at most one zero in the disk $\{|z - z_0| < \delta/3\}$. It is also clear that for any symmetric measure μ with $S(\mu) = [-1, -1/2] \cup \{0\} \cup [1/2, 1]$ the polynomials $p_{2n}(\mu)$ have 2 zeros near 0 (in this case $\delta = 0$). Moreover, if μ lies on the unit circle, then there exist two positive constants C and a and a zero z_n of p_n such that $|z_n - z_0| \leq Ce^{-an}$.

Next put a unit mass to every zero of $p_n(\mu)$ (counting multiplicity), this gives the so called counting measure $\nu_{p_n(\mu)}$ on the zero set. Zero distribution amounts to finding the limit behavior of $\frac{1}{n}\nu_{p_n(\mu)}$. The normalized arc measure on the unit circle (for which $p_n(\mu, z) = z^n$) shows that if the interior of the polynomial convex hull $\mathrm{Pc}(S(\mu))$ is not empty, then the zeros may be far away from the outer boundary $\partial\Omega$, where the equilibrium measure $\omega_{S(\mu)}$ is supported. Thus, assume that $\mathrm{Pc}(S(\mu))$ has empty interior and also that there is no Borel set of capacity zero and full μ-measure (the case when this is not true is rather pathological, almost anything can happen with the zeros then). In this case

$$\lim_{n\to\infty} \kappa_n^{1/n} = \log \frac{1}{\mathrm{cap}(S(K))} \tag{7.5}$$

if and only if

$$\lim \frac{1}{n}\nu_{p_n(\mu)} = \omega_{S(\mu)}$$

in weak* sense, i.e. asymptotically minimal behavior of $\kappa_n^{1/n}$ (see (7.2)) is equivalent to the fact that the zero distribution is the equilibrium distribution.

(7.5) is called regular limit behavior, and in this case we write $\mu \in \mathbf{Reg}$. Thus, the important class $\mathbf{Reg}$ is defined by the property (7.5). If Ω is a regular set with respect to the Dirichlet problem, then $\mu \in \mathbf{Reg}$ is equivalent to either of

- $\lim_{n\to\infty} \|p_n(\mu)\|_{\sup,S(\mu)}^{1/n} = 1$
- For any sequence $\{P_n\}$ of polynomials of degree $n = 1, 2, \ldots$

$$\lim_{n\to\infty} \left(\frac{\|P_n\|_{\sup,S(\mu)}}{\|P_n\|_{L^2(\mu)}} \right)^{1/n} = 1.$$

The last statement expresses the fact that in n-th root sense the $L^2(\mu)$ and L^∞ norms (on $S(\mu)$) are asymptotically the same.

All equivalent formulations of $\mu \in \mathbf{Reg}$ point to a certain "thickness" of μ on its support. Regularity is an important property, and it is desirable to know "thickness" conditions under which it is true. Several regularity criteria are known, e.g. either of the conditions

- all Borel sets $B \subseteq S(\mu)$ with full measure (i.e with $\mu(B) = \mu(S(\mu))$) have capacity $\mathrm{cap}(B) = \mathrm{cap}(S(\mu))$
- $d\mu/d\omega_{S(\mu)} > 0$ (Radon-Nikodym derivative) $\omega_{S(\mu)}$-almost everywhere

is sufficient for $\mu \in \mathbf{Reg}$. Regularity holds under fairly weak assumptions on the measure, e.g. if $S(\mu) = [0, 1]$, and

$$\liminf_{r\to 0} r \log \mu([x - r, x + r]) \geq 0$$

for almost every $x \in [0, 1]$ (i.e. if μ is not exponentially small around almost every point), then $\mu \in \mathbf{Reg}$.

No necessary and sufficient condition for regularity in terms of the size of the measure μ is known. The only existing necessary condition is for the case $S(\mu) = [0, 1]$, and it reads that for every $\eta > 0$

$$\lim_{n\to\infty} \mathrm{cap}\left(\{ x \,|\, \mu([x - 1/n, x + 1/n]) \geq e^{-\eta n} \} \right) = \frac{1}{4}$$

(here $1/4$ is the capacity of $[0, 1]$).

8 Strong and ratio asymptotics

Let μ be supported on $[-1, 1]$ and suppose that the so called Szegő condition

$$\int_{-1}^{1} \frac{\log \mu'(t)}{\sqrt{1 - t^2}} dt > -\infty \tag{8.1}$$

holds, where μ' is the Radon-Nikodym derivative of μ with respect to linear Lebesgue measure. Note that this condition means that the integral is finite, for it cannot be ∞. It expresses a certain denseness of μ, and under this condition G. Szegő proved several asymptotics for the corresponding orthonormal polynomials $p_n(\mu)$. This theory was developed on the unit circle and then was translated into the real line. The Szegő function associated with μ is

$$D_\mu(z) = \exp\left(\sqrt{z^2-1}\frac{1}{2\pi}\int_{-1}^{1}\frac{\log\mu'(t)}{z-t}\frac{dt}{\sqrt{1-t^2}}\right) \tag{8.2}$$

and it is the outer function in the Hardy space on $\mathbf{C}\setminus[-1,1]$ with boundary values $|D_\mu(x)|^2 = \mu'(x)$. Outside $[-1,1]$ the asymptotic formula

$$p_n(\mu, z) = (1+o(1))\frac{1}{\sqrt{2\pi}}(z+\sqrt{z^2-1})^n D_\mu(z)^{-1} \tag{8.3}$$

holds locally uniformly, in particular, the leading coefficient κ_n of $p_n(\mu)$ is of the form

$$\kappa_n = (1+o(1))\frac{2^n}{\sqrt{2\pi}}\exp\left(\frac{-1}{2\pi}\int_{-1}^{1}\frac{\log\mu'(t)}{\sqrt{1-t^2}}dt\right). \tag{8.4}$$

For all these results see [104], Chapter 6. The Szegő condition is also necessary for these results, e.g. (8.3) and (8.4) are equivalent to (8.1).

If one assumes weaker conditions then necessarily weaker results will follow. A large and important class of measures is the Nevai class $M(b,a)$ (see [66]), for which the coefficients in the three term recurrence

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1}p_{n-1}(x)$$

satisfy $a_n \to a$, $b_n \to b$. This is equivalent to ratio asymptotics

$$\lim_{n\to\infty}\frac{p_{n+1}(z)}{p_n(z)} = \frac{z-b+\sqrt{(z-b)^2-4a^2}}{2}$$

for large z (actually, away from the support of μ), and the monograph [66] contains a very detailed treatment of orthogonal polynomials in this class. Simon [85] showed that if the limit of $p_{n+1}(z)/p_n(z)$ exists at a single non-real z, then $\mu \in M(b,a)$ for some a, b.

The classes $M(b,a)$ are scaled versions of each other, and the most important condition ensuring $M(0,1)$ is given in Rahmanov's theorem [76]: if μ is supported in $[-1,1]$ and $\mu' > 0$ almost everywhere on $[-1,1]$, then $\mu \in M(0,1)$. Conversely, Blumenthal's theorem [12] states that $\mu \in M(0,1)$ implies that the support of μ is $[-1,1]$ plus at most countably many points that converge to ± 1. Thus, in this respect the extension of Rahmanov's theorem given in [20] is of importance: if $\mu' > 0$ almost everywhere on $[-1,1]$ and outside $[-1,1]$ the measure μ has at most countably many mass points converging to ± 1, then $\mu \in M(0,1)$. However, $M(0,1)$ contains many other measures not just those that are in these theorems, e.g. in [22] a continuous singular measure in the

Nevai class was exhibited, and the result in [107] shows that the Nevai class contains practically all kinds of measures allowed by Blumenthal's theorem.

Under Rahmanov's condition $\mathrm{supp}(\mu) = [-1,1]$, $\mu' > 0$ a.e., some parts of Szegő's theory can be proven in a weaker form (see e.g. [56, 57]). In these the Turán determinants

$$T_n(x) = p_n^2(x) - p_{n-1}(x)p_{n+1}(x)$$

play significant role. In fact, then given any interval $\mathbf{D} \subset (-1,1)$ the Turán determinant T_n is positive on $\mathbf{D}$ for all large n, and $T_n^{-1}(x)dx$ converges in weak* sense to $d\mu$ on $\mathbf{D}$. Furthermore, the absolutely continuous part μ' can be also separately recovered from T_n:

$$\lim_{n\to\infty} \int \left| T_n(x)\mu'(x) - \frac{2}{\pi}(1-x^2)^{1/2} \right| dx = 0.$$

Under Rahmanov's condition we also have weak convergence, e.g.

$$\lim_{n\to\infty} \int f(x)p_n^2(x)\mu'(x)dx = \frac{1}{\pi}\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}}dx \tag{8.5}$$

for any continuous function f. Pointwise we only know a highly oscillatory behavior: for almost all $x \in [-1,1]$

$$\limsup_{n\to\infty} p_n(x) \geq \frac{2}{\pi}(\mu'(x))^{-1/2}(1-x^2)^{-1/4},$$

$$\liminf_{n\to\infty} p_n(x) \leq -\frac{2}{\pi}(\mu'(x))^{-1/2}(1-x^2)^{-1/4},$$

and if $E_n(\varepsilon)$ is the set of points $x \in [-1,1]$ where

$$|p_n(x)| \geq (1+\varepsilon)\frac{2}{\pi}(\mu'(x))^{-1/2}(1-x^2)^{-1/4},$$

then $|E_n(\varepsilon)| \to 0$ for all $\varepsilon > 0$. However, it is not true that the sequence $\{p_n(\mu,x)\}$ is pointwise bounded, since for every $\varepsilon > 0$ there is a weight function $w > 1$ on $[-1,1]$ such that $p_n(0)/n^{1/2-\varepsilon}$ is unbounded (see [77]).

Simon [85] extended (8.5) by showing that if the recurrence coefficients satisfy $b_n \to b$, $a_{2n+1} \to a'$ and $a_{2n} \to a''$, then there is an explicitly calculated measure ρ depending only on b, a', a'' such that

$$\lim_{n\to\infty} \int f(x)p_n^2(x)\mu'(x)dx = \int_{-1}^{1} f(x)d\rho(x) \tag{8.6}$$

for any continuous function f, and conversely, if (8.6) exists for $f(x) = x, x^2, x^4$, then $b_n \to b$, $a_{2n+1} \to a'$ and $a_{2n} \to a''$ with some b, a', a''.

Szegő's theory can be extended to measures lying on a single Jordan curve or arc J (see [40] where also additional outside lying mass points are allowed), in which case the role of $z + \sqrt{z^2 - 1}$ in (8.3) is played by the conformal map Φ of $\mathbf{C} \setminus J$ onto the exterior of the unit disk, and the role of 2^n in (8.4) is played by the reciprocal of the logarithmic capacity of J (see section 7). Things change considerably, if the measure is supported on a set J consisting of two or more smooth curve or arc components $J_1, \ldots, J_m$. A general feature of this case is that $\kappa_n \mathrm{cap}(J)^n$ does not have a limit, its limit points fill a whole interval. The polynomials themselves have asymptotic form

$$\frac{p_n(z)}{\kappa_n} = \mathrm{cap}(J)^n \Phi(z)^n (F_n(z) + o(1))$$

uniformly away from J, where Φ is the (multi-valued) complex Green function of the complement $\mathbf{C} \setminus J$, and where F_n is the solution of an L^2-extremal problem involving analytic functions belonging to some class Γ_n. The functions F in Γ_n are determined by an H^2 condition plus an argument condition, namely if the change of the argument of Φ as we go around J_k is $\gamma_k 2\pi$ modulo 2π, then in Γ_n we consider functions the change of the argument of which around J_k is $-n\gamma_k 2\pi$ modulo 2π. Now the point is that these function classes Γ_n change with n, and hence so does F_n, and that is the reason that a single asymptotic formula like (8.4) or (8.3) does not hold. The fundamentals of the theory was laid in H. Widom's paper [114]; and since then many results have been obtained by F. Peherstorfer and his collaborators, as well as A. I. Aptekarev, J. Geronimo and W. Van Assche. The theory has deep connections with function theory, the theory of Abelian integrals and the theory of elliptic functions. We refer the reader to the papers [6], [29], [71]–[75].

The Christoffel functions

$$\lambda_n^{-1}(\mu, x) = \sum_{k=0}^{n} p_k(\mu, x)^2$$

behave somewhat more regularly than the orthogonal polynomials. In [58] it was shown that if μ is supported on $[-1, 1]$, it belongs to the **Reg** class there (see Section 7) and $\log \mu'$ is integrable over an interval $I \subset [-1, 1]$, then for almost all $x \in I$

$$\lim_{n\to\infty} n\lambda_n(x) = \pi\sqrt{1 - x^2}\mu'(x).$$

This result is true [108] in the form

$$\lim_{n\to\infty} n\lambda_n(x) = \frac{d\mu(x)}{d\omega_{\mathrm{supp}(\mu)}(x)}, \qquad a.e.\ x \in I$$

when the support is a general compact subset of $\mathbf{R}$, $\mu \in \mathbf{Reg}$ and $\log \mu' \in L^1(I)$.

Often only a rough estimate is needed for Christoffel functions, and such one is provided in [55]: if w is supported on $[-1,1]$ and it is a doubling weight, i.e.

$$\int_{2I} w \le L \int_I w$$

for all $I \subset [-1,1]$, where $2I$ is the twice enlarged I, then uniformly on $[-1,1]$

$$\lambda_n(x) \sim \int_{\Delta_n(x)} w; \qquad \Delta_n(x) = \frac{\sqrt{1-x^2}}{n} + \frac{1}{n^2}.$$

9 Exponential and Freud weights

These are weight functions of the form $e^{-2Q(x)}$, where x is on the real line or on some subinterval of it. For simplicity we shall first assume that Q is even. We get Freud weights when $Q(x) = |x|^\alpha$, $\alpha > 0$, $x \in \mathbf{R}$ and Erdős weights if Q tends to infinity faster than any polynomial as $|x| \to \infty$. G. Freud started to investigate these weights in the sixties and seventies, but they independently appeared also in the Russian literature and in statistical physics. One can safely say that some of Freud problems and the work of P. Nevai and E. Rahmanov were the primary cause of the sudden revitalization of the theory of orthogonal polynomials since the early 1980's. In the last 20 years D. Lubinsky with coauthors have conducted systematic studies on exponential weights, see e.g. [48, 49, 52, 53, 109]. In the mid 1990's a new stimulus came from the Riemann-Hilbert approach that was used together with the steepest descent method by P. Deift and his collaborators ([19]) to give complete asymptotics when Q is analytic.

One can roughly say that because of the fast vanishing of the weight around infinity, things happen on a finite subinterval $[-a_n, a_n]$ (depending on the degree of the polynomials), and on $[-a_n, a_n]$ techniques developed for $[-1,1]$ are applied. For Freud weights one can also make the substitution $x \to n^{1/\lambda}x$ and go to orthogonality with respect to the varying weight $e^{-n|x|^\lambda}$, in which case things are automatically reduced to a finite interval which is the support of a weighted energy problem.

a_n are the so called Mhaskar-Rahmanov-Saff numbers with definition

$$n = \frac{2}{\pi}\int_0^1 \frac{a_n t Q'(a_n t)}{\sqrt{1-t^2}}\, dt. \tag{9.1}$$

The zeros of $p_n(w^2)$, $w(x) = \exp(-Q(x))$ are spreading out and the largest zero is very close to a_n, which tends to ∞.

To describe the distribution of the zeros and the behavior of the polynomials one has to make appropriate contractions. Let us consider first the case of Freud weight $w(x) = \exp(-|x|^\alpha)$, and let p_n be the n-th orthogonal polynomial with respect to w^2 (on $(-\infty, \infty)$). In this case

$$a_n = n^{1/\alpha}\gamma_\alpha, \qquad \gamma_\alpha^\alpha := \Gamma\left(\frac{\alpha}{2}\right)\Gamma\left(\frac{1}{2}\right)\Big/ 2\Gamma\left(\frac{\alpha}{2}+\frac{1}{2}\right).$$

Thus, for the largest zero $x_{n,n}$ we have $x_{n,n}/n^{1/\alpha} \to \gamma_\alpha$ as $n \to \infty$, and to describe zero distribution we divide (contract) all zeros $x_{n,i}$ by $n^{1/\alpha}\gamma_\alpha$. These contracted zeros asymptotically have the distribution

$$\frac{d\mu_w(t)}{dt} := \frac{\alpha}{\pi}\int_{|t|}^{1} \frac{u^{\alpha-1}}{\sqrt{u^2-t^2}}\,du, \quad t \in [-1,1]. \tag{9.2}$$

This measure μ_w minimizes the weighted energy

$$\int\int \log\frac{1}{|x-t|}d\mu(x)d\mu(t) + 2\int Q d\mu \tag{9.3}$$

among all probability measures compactly supported on $\mathbf{R}$. It is a general feature of exponential weights that the behavior of zeros or the polynomials is governed by the solution of a weighted energy problem (weighted equilibrium measures). If κ_n is the leading coefficient of p_n, i.e. $p_n(z) = \kappa_n z^n + \cdots$, then

$$\lim_{n\to\infty} \kappa_n \pi^{1/2} 2^{-n} e^{-n/\alpha} n^{(n+1/2)/\alpha} = 1,$$

and we have

$$\lim_{n\to\infty} |p_n(n^{1/\alpha}\gamma_\alpha z)|^{1/n}$$
$$= \exp\left(\log|z+\sqrt{z^2-1}| + \mathrm{Re}\int_0^1 \frac{zu^{\alpha-1}}{\sqrt{z^2-u^2}}\,du\right)$$

locally uniformly outside $[-1,1]$. This latter one is so called n-th root asymptotics, while the former one is strong asymptotics. Strong asymptotics for $p_n(z)$ on different parts of the complex plane was given in [46] using the Hilbert-Riemann approach (see also [18] in this volume).

Things become more complicated for non-Freud weights, but the corresponding results are of the same flavor. In this case the weight is not necessarily symmetric, but under some conditions (like Q being convex or $xQ'(x) \nearrow$ for $x > 0$ and an analogous condition for $x < 0$) the relevant weighted equilibrium measure's support is an interval, and the definition of the Mhaskar–Rahmanov–Saff numbers $a_{\pm n}$ is

$$\begin{aligned} n &= \frac{1}{\pi}\int_{a_{-n}}^{a_n} \frac{xQ'(x)}{\sqrt{(x-a_{-n})(a_n-x)}}dx, \\ 0 &= \frac{1}{\pi}\int_{a_{-n}}^{a_n} \frac{Q'(x)}{\sqrt{(x-a_{-n})(a_n-x)}}dx. \end{aligned}$$

Now one solves the weighted equilibrium problem (9.3) for all measures μ *with total mass* n, and if μ_n is the solution then $[a_{-n}, a_n]$ is the support of μ_n and μ_n/n will play the role of the measure μ_w from (9.2) above.

The weight does not even have to be defined on all $\mathbf{R}$, e.g. in [48] a theory was developed that simultaneously include far reaching generalizations of non-symmetric Freud, Erdős and Pollaczek weights (the latter are defined on $[-1,1]$ and vanish at high order at ± 1).

10 Sobolev orthogonality

In Sobolev orthogonality we consider orthogonality with respect to an inner product

$$(f, g) = \sum_{k=0}^{r} \int f^{(k)} \overline{g^{(k)}} d\mu_k, \tag{10.1}$$

where μ_k are given positive measures. There are several motivations for this kind of orthogonality, perhaps the most natural one is smooth data fitting. The Spanish school around F. Marcellán, G. Lopez and A. Martinez-Finkelshtein have been particularly active in developing this area (see the surveys [59] and [61, 60] and the references there).

In this section let $Q_n(z) = z^n + \cdots$ denote the *monic* orthogonal polynomial with respect to the Sobolev inner product (10.1), and $q_n(\mu_k)$ the monic orthogonal polynomials with respect to the measure μ_k.

Most arguments for the standard theory fail in this case, e.g. it is no longer true that the zeros lie in the convex hull of the support of the measures μ_k, $k = 0, 1, \ldots, r$. It is not even known if the zeros are bounded if all the measures μ_k have compact support. Nonetheless, for the case $r = 1$, and $\mu_0, \mu_1 \in \mathbf{Reg}$ (see section 7) it was shown in [28] that the asymptotic distribution of the zeros of the *derivative* Q_n' is the equilibrium measure $\omega_{E_0 \cup E_1}$, where E_i is the support of μ_i, $i = 0, 1$ (which also have to be assumed to be regular). Furthermore, if, in addition, $E_0 \subseteq E_1$, then the asymptotic zero distribution of Q_n is ω_{E_0}.

In general, both the algebraic and the asymptotic/analytic situation is quite complicated, and there are essentially two important cases which have been understood to a satisfactory degree.

Case I: The discrete case. In this case μ_0 is some "strong" measure, e.g. from the Nevai class $M(b, a)$ (see section 8), and $\mu_1, \ldots, \mu_k$ are finite discrete measures. It turns out that then the situation is similar to adding these discrete measures to μ_0 (the new measure will also be in the same Nevai class), and considering standard orthogonality with respect to this new measure. E.g. if $r = 1$, then

$$\lim_{n\to\infty} \frac{Q_n(z)}{q_n(\mu_0 + \mu_1, z)} = 1$$

holds uniformly on compact subsets of $\mathbf{C} \setminus \mathrm{supp}(\mu_0 + \mu_1)$. Thus, the Sobolev orthogonal polynomials differ from those of the measure μ_0, but not more than what happens when adding mass points to μ_0.

In this discrete case the Q_n's satisfy a higher order recurrence relation, hence this case is also related to matrix orthogonality (see the end of the section 13).

Case II: The Szegő case. Suppose now that $\mu_0, \ldots, \mu_k$ are all supported on the same smooth curve or arc J, and they satisfy Szegő's condition there (see Section 8). In this case the k-th derivative of Q_n satisfies locally uniformly in the complement of J the asymptotic formula

$$\lim_{n\to\infty} \frac{Q_n^{(k)}(z)}{n^k q_{n-k}(\mu_m, z)} = \frac{1}{[\Phi'(z)]^{m-k}},$$

where Φ is the conformal map that maps $\mathbf{C} \setminus J$ onto the complement of the unit disk. That is, in this case the measures $\mu_0, \ldots, \mu_{r-1}$ do not appear in the asymptotic formula, only μ_m matters. The reason for this is the following: $Q = Q_n$ minimizes

$$(Q, Q) = \sum_{k=0}^{r} \int |Q^{(k)}|^2 d\mu_k \tag{10.2}$$

among all monic polynomials of degree n, while $q = q_{n-k}(\mu_k)$ minimizes

$$\int |q|^2 d\mu_k$$

among all monic polynomials of degree $n - k$. But the polynomial $Q_n^{(k)}(t) = n(n-1)\cdots(n-k+1)t^{n-k} + \cdots$ is a monic polynomial times the factor $n(n-1)\cdots(n-k+1) \sim n^k$, and this factor is dominant for $k = r$, so everything else will be negligible. There are results for compensation of this n^k factor which lead to Sobolev orthogonality with respect to varying measures.

Under the much less restrictive assumption that $\mu_0 \in \mathbf{Reg}$ and the other measures μ_k are supported in the support E of μ_0 it is true ([50]) that the asymptotic zero distribution of $Q_n^{(k)}$ is the equilibrium measure ω_E for all k,

$$\lim_{n\to\infty} \|Q_n^{(k)}\|_{\sup,E}^{1/n} = \text{cap}(E),$$

and hence, away from the zeros in the unbounded component of the complement of E, we have

$$\lim_{n\to\infty} |Q_n^k(z)|^{1/n} = e^{g_{\mathbf{C}\setminus E}(z)}$$

where $g_{\mathbf{C}\setminus E}$ is the Green function for this unbounded component.

The techniques developed for exponential weights and for Sobolev orthogonality were combined in [30] to prove strong asymptotics for Sobolev orthogonal polynomials when $r = 1$ and $\mu_0 = \mu_1$ are exponential weights.

11 Non-Hermitian orthogonality

We refer to non-Hermitian orthogonality in either of these cases:

- the measure μ is non-positive or even complex valued and we consider p_n with

$$\int p_n(z)\overline{z^k} d\mu = 0, \qquad k = 0, 1, \ldots, n-1, \tag{11.1}$$

- μ is again non-positive or complex valued, or positive but lies on a complex curve or arc and orthogonality is considered without complex conjugation, i.e. if

$$\int p_n(z)z^k d\mu = 0, \qquad k = 0, 1, \ldots, n-1. \tag{11.2}$$

More generally, one could consider non-positive inner products, but we shall restrict our attention to complex measures and orthogonality (11.2).

As an example, consider diagonal Padé approximant to the Cauchy transform

$$f(z) = \int \frac{d\mu(t)}{z-t}$$

of a signed or complex valued measure, i.e. consider polynomials p_n and q_n of degree at most n such that

$$f(z)p_n(z) - q_n(z) = O(z^{-n-1})$$

at infinity. Then p_n satisfies the non-Hermitian orthogonality relation

$$\int p_n(x)x^j d\mu(x) = 0, \qquad j = 0, 1, \ldots, n-1. \tag{11.3}$$

In this non-Hermitian case even the Gram-Schmidt process may fail, and then p_n is defined as the solution of the orthogonality condition (11.1) resp. (11.2), which give a system of homogeneous equations for the coefficients of p_n. Thus, p_n may be of smaller than n degree, and things can go pretty wild with this kind of orthogonality, e.g. in the simple case

$$d\mu(x) = (x - \cos\pi\alpha_1)(x - \cos\pi\alpha_2)(1-x^2)^{-1/2}dx, \quad x \in [-1, 1]$$

with $0 < \alpha_1 < \alpha_2 < 1$ rationally independent algebraic numbers, the zeros of p_n from (11.3) are dense on the whole complex plane (compare this with the fact that for positive μ all zeros lie in $[-1, 1]$). In [99] it was shown that it is possible to construct a complex measure μ on $[-1, 1]$, such that for an arbitrary prescribed asymptotic behavior some subsequence $\{p_{n_k}\}$ will have this zero behavior. Nonetheless, the asymptotic distribution of the zeros is again the equilibrium distribution of the support of μ under regularity conditions on μ. For example, this is the case if $|\mu|$ belongs to the **Reg** class (see section 7), and the argument of μ, i.e. $d\mu(t)/d|\mu|(t)$ is of bounded variation ([8]). In [99]–[101] H. Stahl obtained asymptotics for non-Hermitian orthogonal polynomials even for varying measures and gave several applications of them to Padé approximation. When the measure μ is of the form $d\mu(x) = g(x)(1-x^2)^{-1/2}dx$, $x \in [-1, 1]$ with an analytic g, for $z \in \mathbf{C} \setminus [-1, 1]$ strong asymptotic formula of the form

$$\frac{p_n(z)}{\kappa_n} = (1 + o(1))\frac{(z + \sqrt{z^2-1})^n}{2^n} D_\mu(z)^{-1} \exp\left(\frac{1}{2\pi}\int_{-1}^{1} \frac{\log \mu'(t)}{\sqrt{1-t^2}} dt\right)$$

(with D_μ the Szegő function (8.2)) was proved by A. I. Aptekarev and W. Van Assche via the Riemann-Hilbert approach. A similar result holds on the support of the measure, as well as for the case of varying weights, see [7].

12 Multiple orthogonality

Multiple orthogonality comes from simultaneous Padé approximation. It is a relatively new area where we have to mention the names of E. M. Nikishin, V. N. Sorokin, A. I. Aptekarev and W. Van Assche (see the survey [110] by W. Van Assche and the references there).

On $\mathbf{R}$ let there be given r measures $\mu_1, \ldots, \mu_r$ with finite moments and infinite support, and consider multiindices $\underline{n} = (n_1, \ldots, n_r)$ of nonnegative integers with norm $|\underline{n}| = n_1 + \cdots + n_r$. There are two types of multiple orthogonality corresponding to the appropriate Hermite-Padé approximation.

In type I we are looking for polynomials $Q_{\underline{n},j}$ of degree $n_j - 1$ for each $j = 1, \ldots, r$ such that

$$\sum_{j=1}^{r} \int x^k Q_{\underline{n},j}(x) d\mu_j(x) = 0, \qquad k = 0, 1, \ldots, |\underline{n}| - 2.$$

These orthogonality relations give $|\underline{n}| - 1$ homogeneous linear equations for the $|\underline{n}|$ coefficients of the r polynomials $Q_{\underline{n},j}$, so there is a non-trivial solution. If the rank of the system is $|\underline{n}| - 1$, then the solution is unique up to a multiplicative factor, in which case the index $\underline{n}$ is called normal. This happens precisely if each $Q_{\underline{n},j}$ is of exact degree $n_j - 1$.

In type II we are looking for a single polynomial $P_{\underline{n}}$ of degree $|\underline{n}|$ such that

$$\begin{aligned} \int x^k P_{\underline{n}}(x) d\mu_1(x) &= 0, \qquad k = 1, \ldots, n_1 - 1 \\ &\vdots \\ \int x^k P_{\underline{n}}(x) d\mu_r(x) &= 0, \qquad k = 1, \ldots, n_r - 1. \end{aligned}$$

These are $|\underline{n}|$ homogeneous linear equations for the $|\underline{n}|+1$ coefficients of $P_{\underline{n}}$, and again if the solution is unique up to a multiplicative constant, then $\underline{n}$ is called regular. This is again equivalent to $P_{\underline{n}}$ being of exact degree $\underline{n}$.

$\underline{n}$ is regular for type I orthogonality precisely when it is regular for type II, so we just speak of regularity. This is the case for example if the μ_j's are supported on intervals $[a_j, b_j]$ that are disjoint except perhaps for their endpoints; in fact, in this case $P_{\underline{n}}$ has n_j simple zeros on (a_j, b_j).

To describe recurrence formulae, let $\underline{e}_j = (0, \ldots, 1, \ldots, 0)$ where the single 1 entry is at position j. Then for any k

$$xP_{\underline{n}}(x) = a^*_{\underline{n},k} P_{\underline{n}+\underline{e}_k}(x) + \sum_{j=1}^{r} a_{\underline{n},j} P_{\underline{n}-\underline{\varepsilon}_j}(x).$$

Another recurrence formula is

$$xP_{\underline{n}}(x) = b^*_{\underline{n},k} P_{\underline{n}+\underline{e}_k}(x) + \sum_{j=1}^{r} b_{\underline{n},j} P_{\underline{n}-\underline{\varepsilon}_{\pi(1)}-\cdots-\underline{e}_{\pi(j)}}(x),$$

where $\pi(1), \dots, \pi(r)$ is an arbitrary but fixed permutation of $1, 2, \dots, r$. The orthogonal polynomials with different indices are strongly related to one another, e.g. $P_{\underline{n}+\underline{e}_k}(x) - P_{\underline{n}+\underline{e}_l}(x)$ is a constant multiple of $P_{\underline{n}}(x)$.

If $d\mu_j = w_j d\mu$, then similar recurrence relations hold in case of type I orthogonality for

$$Q_{\underline{n}}(x) = \sum_{j=1}^{r} Q_{\underline{n},j}(x) w_j(x).$$

Also, type I and type II are related by a biorthogonality property:

$$\int P_{\underline{n}} Q_{\underline{m}} d\mu = 0$$

except for the case when $\underline{m} = \underline{n} + \underline{e}_k$ for some k, and then the previous integral is not zero (under regularity condition).

To describe an analogue of the Christoffel-Darboux formula let $\{\underline{n}_j\}$ be a sequence of multiindices such that $\underline{n}_0$ is the identically 0 multiindex, and $\underline{n}_{j+1}$ coincides with $\underline{n}_j$ except for one component which is 1 larger than then corresponding component of $\underline{n}_j$. Set $P_j = P_{\underline{n}_j}$, $Q_j = Q_{\underline{n}_{j+1}}$ and with $\underline{n} = \underline{n}_n$

$$h^{(j)}_{\underline{n}} = \int P_{\underline{n}}(x) x^{n_j} d\mu_j(x).$$

Then [16] with $\underline{n} = \underline{n}_n$

$$(x-y) \sum_{k=0}^{n-1} P_k(x) Q_k(y) = P_{\underline{n}}(x) Q_{\underline{n}}(y) - \sum_{j=1}^{r} \frac{h^{(j)}_{\underline{n}}}{h^{(j)}_{\underline{n}-\underline{e}_j}} P_{\underline{n}-\underline{e}_j}(x) Q_{\underline{n}+\underline{e}_j}(y).$$

Thus, the left hand side depends only on $\underline{n} = \underline{n}_n$ and not on the particular choice of the sequence $\underline{n}_j$ leading to it.

There is an approach [111] to both types of multiple orthogonality in terms of matrix valued Riemann-Hilbert problem for $(r+1) \times (r+1)$ matrices $Y = (Y_{ij}(z))^r_{i,j=0}$.

Asymptotic behavior of multiple orthogonal polynomials is not fully understood yet, due to the interaction of the different measures. For all the existing results see [5, 110] and W. Van Assche's Chapter 23 in [39] and the references there.

13 Matrix orthogonal polynomials

In the last 20 years the fundamentals of matrix orthogonal polynomials have been developed mainly by A. Duran and his coauthors. The theory shows many similarities with the scalar case, but there is an unexpected richness which is still to be explored.

For all the results in this section see [51] and [24] and the numerous references there.

An $N \times N$ matrix

$$P(t) = \begin{pmatrix} p_{11}(t) & \cdots & p_{1N}(t) \\ \vdots & \ddots & \vdots \\ p_{N1}(t) & \cdots & p_{NN}(t) \end{pmatrix}$$

with polynomial entries $p_{ij}(t)$ of degree at most n is called a matrix polynomial of degree at most n. Alternatively, one can write

$$P(t) = C_n t^n + \cdots + C_0$$

with numerical matrices $C_n, \cdots, C_0$ of size $N \times N$.

$t = a$ is called a zero of P if $P(a)$ is singular, and the multiplicity of a is the multiplicity of a as a zero of $\det P(a)$. When the leading coefficient matrix C_n is non-singular, then P has nN zeros counting multiplicity.

From now on we fix the dimension to be N, but the degree n can be any natural number. I will denote the $N \times N$ unit matrix and 0 stands for all kinds of zeros (numerical or matrix).

A matrix

$$W(t) = \begin{pmatrix} \mu_{11}(t) & \cdots & \mu_{1N}(t) \\ \vdots & \ddots & \vdots \\ \mu_{N1}(t) & \cdots & \mu_{NN}(t) \end{pmatrix}$$

of measures defined on (or part of) the real line is positive definite if for any Borel set E the numerical matrix $W(E)$ is positive semidefinite. We assume that all moments of W is finite. With such a matrix we can define a matrix inner product on the space of $N \times N$ matrix polynomials via

$$(P, Q) = \int P(t) dW(t) Q^*(t),$$

and if (P, P) is nonsingular for any P with nonsingular leading coefficient, then just as in the scalar case one can generate a sequence $\{P_n\}_{n=0}^{\infty}$ of matrix polynomials of degree $n = 0, 1, \ldots$ which are orthonormal with respect to W:

$$\int P_n(t) dW(t) P_m^*(t) = \begin{cases} 0 & \text{if } n \neq m \\ I & \text{if } n = m, \end{cases}$$

and here P_n has nonsingular leading coefficient matrix. The sequence $\{P_n\}$ is determined only up to multiplication on the left by unitary matrices, i.e. if U_n are unitary matrices, then the polynomials $U_n P_n$ also form an orthonormal system with respect to W.

Just as in the scalar case, these orthogonal polynomials satisfy a three term recurrence relation

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_n P_n(t) + A_n^* P_{n-1}(t), \qquad n \geq 0, \tag{13.1}$$

where A_n are nonsingular matrices, and B_n are Hermitian. Conversely, the analogue of Favard's theorem is also true: if a sequence of matrix polynomials $\{P_n\}$ of corresponding degree $n = 0, 1, 2, \dots$ satisfy (13.1) with nonsingular A_n and Hermitian B_n, then there is a positive definite measure matrix W such that P_n are orthonormal with respect to W.

The three term recurrence formula easily yields the Christoffel-Darboux formula:

$$(w - z) \sum_{k=0}^{n-1} P_k^*(z) P_k(w) = P_{n-1}^*(z) A_n P_n(w) - P_n^*(z) A_n^* P_{n-1}(w),$$

from which e.g. it follows that

$$P_{n-1}^*(z) A_n P_n(z) - P_n^*(z) A_n^* P_{n-1}(z) = 0$$

$$\sum_{k=0}^{n-1} P_k^*(z) P_k(z) = P_{n-1}^*(z) A_n P_n'(z) - P_n^*(z) A_n^* P_{n-1}'(z).$$

The orthogonal polynomials Q_n of the second kind

$$Q_n(t) = \int \frac{P_n(t) - P_n(x)}{t - x} dW(x), \qquad n = 1, 2, \dots$$

also satisfy the same recurrence and are orthogonal with respect to some other matrix measure. For them we have

$$P_{n-1}^*(t) A_n Q_n(t) - P_n^*(z) A_n^* Q_{n-1}(t) \equiv I$$

and

$$Q_n(t) P_{n-1}^*(t) - P_n(t) Q_{n-1}^*(t) \equiv A_n^{-1}.$$

With the recurrence coefficient matrices A_n, B_n one can form the block Jacobi matrix

$$J = \begin{pmatrix} B_0 & A_0 & 0 & 0 & \cdots \\ A_0^* & B_1 & A_1 & 0 & \cdots \\ 0 & A_1^* & B_2 & A_2 & \cdots \\ 0 & 0 & A_2^* & B_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The zeros of P_n are real and they are the eigenvalues (with the same multiplicity) of the N-truncated block Jacobi matrix (which is of size nN). If a is a zero then its multiplicity p is at most N, the rank of $P_n(a)$ is $N-p$, and the space of those vectors v for which $P_n(a)v = 0$ is of dimension p. If we write $x_{n,k}$, $1 \le k \le m$ for the different zeros of P_n, and l_k is the multiplicity of $x_{n,k}$, then the matrices

$$\Gamma_k = \frac{1}{(\det(P_n(t))^{(l_k)}(x_{n,k})} \left(\mathrm{Adj}(P_n(t))\right)^{(l_k-1)}(x_{n,k})Q_n(x_{n,k}), \qquad 1 \le k \le m$$

are positive semidefinite of rank l_k, and with them the matrix quadrature formula

$$\int P(t)dW(t) = \sum_{k=1}^{m} P(x_{n,k})\Gamma_{n,k}$$

holds for all matrix polynomial P of degree at most $2n-1$.

If the matrix of orthogonality is diagonal (or similar to a diagonal matrix) with diagonal entries μ_i, then the orthogonal matrix polynomials are also diagonal with i-th entry equal to $p_n(\mu_i)$, the n-th orthogonal polynomial with respect to μ_i. Many matrix orthogonal polynomials in the literature can be reduced to this scalar case. However, recently some remarkably rich non-reducible families have been obtained by A. Duran and F. Grünbaum (see [24] and the references there), which may play the role of the classical orthogonal polynomials in higher dimension. They found families of matrix orthogonal polynomials that satisfy second order (matrix) differential equations just like the classical orthogonal polynomials. Their starting point was a symmetry property between the orthogonality measure matrix and a second order differential operator. They worked out several explicit examples, here is one of them: $N = 2$, the matrix of measure is

$$H(t) = e^{-t^2}\begin{pmatrix} 1+|a|^2t^4 & at^2 \\ \overline{a}t^2 & 1 \end{pmatrix}, \qquad t \in \mathbf{R},$$

where $a \in \mathbf{C} \setminus \{0\}$ is a free parameter. The corresponding $P_n(t)$ satisfies

$$P_N''(t) + P_n'(t)\begin{pmatrix} -2t & 4at \\ 0 & -2t \end{pmatrix} + P_n(t)\begin{pmatrix} -4 & 2a \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -2n-4 & 2a(2n+1) \\ 0 & -2n \end{pmatrix} P_n(t)$$

There is an explicit Rodrigues' type representation for the polynomials themselves, and the three term recurrence (13.1) holds with $B_n = 0$,

$$A_{n+1} = \sqrt{\frac{n+1}{2}}\begin{pmatrix} \gamma_{n+3}/\gamma_{n+2} & a\gamma_{n+2}\gamma_{n+1} \\ 0 & \gamma_n/\gamma_{n+1} \end{pmatrix},$$

where

$$\gamma_n^2 = 1 + \frac{|a|^2}{2}\binom{n}{2}.$$

Matrix orthogonality is closely connected to $(2N+1)$-term recurrences for scalar polynomials. To describe this we need the following operators on polynomials p: if $p(t) = \sum_k a_k t^k$, then

$$R_{N,m}(p) = \sum_s a_{sN+m} t^s,$$

i.e. from a polynomial the operator $R_{N,m}$ takes those powers where the exponent is congruent to m modulo N, removes the common factor t^m and changes t^N to t.

Now suppose that $\{p_n\}_{n=0}^{\infty}$, is a sequence of scalar polynomials of corresponding degree $n = 0, 1, \ldots$ and suppose that this sequence satisfies a $(2N+1)$-term recurrence relation

$$t^N p_n(t) = c_{n,0} p_n(t) + \sum_{k=1}^{N} (\overline{c_{n,k}} p_{n-k}(t) + c_{n+k,k} p_{n+k}(t)),$$

where $c_{n,0}$ is real, $c_{n,N} \neq 0$ (and $p_k(t) \equiv 0$ for $k < 0$). Then

$$P_n(t) = \begin{pmatrix} R_{N,0}(p_{nN}) & \cdots & R_{N,N-1}(p_{nN}) \\ R_{N,0}(p_{nN+1}) & \cdots & R_{N,N-1}(p_{nN}+1) \\ \vdots & \ddots & \vdots \\ R_{N,0}(p_{nN+N-1}) & \cdots & R_{N,N-1}(p_{nN}+N-1) \end{pmatrix}$$

is a sequence of matrix orthogonal polynomials with respect to a positive definite measure matrix. Conversely, if $P_n = (P_{n,m,j})_{m,j=0}^{N-1}$ is a sequence of orthonormal matrix polynomials, then the scalar polynomials

$$p_{nN+m}(t) = \sum_{j=0}^{N-1} t^j P_{n,m,j}(t^N), \qquad 0 \leq m < N, \ n = 0, 1, 2, \ldots$$

satisfy a $(2N+1)$-recurrence relation of the above form.

Part II
Orthogonal polynomials on the unit circle

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

14 Definitions and basic properties

Orthogonality

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

If μ is a nontrivial probability measure on $\mathbf{T}$ (that is, not supported on a finite set) the monic orthogonal polynomials $\Phi_n(z,\mu)$ (or Φ_n if μ is understood) are uniquely determined by

$$\Phi_n(z) = \prod_{j=1}^{n}(z - z_{n,j}), \qquad \int_{\mathbf{T}} \zeta^{-j}\Phi_n(\zeta)\, d\mu = 0, \quad j = 0, 1, \ldots, n-1 \tag{14.1}$$

so in the Hilbert space $L^2_\mu(\mathbf{T})$, $\langle \Phi_n, \Phi_m \rangle = 0$, $n \neq m$. The orthonormal polynomials φ_n are $\varphi_n = \Phi_n/\|\Phi_n\|$. The orthonormal set $\{\varphi_n\}_{n\geq 0}$ may not be a basis in $L^2_\mu(\mathbf{T})$ (e.g., if $\mu = dm$ is the normalized Lebesgue measure, then $\varphi_n = \zeta^n$ and ζ^{-1} is orthogonal to all φ_n). A celebrated result of Szegő states that $\{\varphi_n\}$ is a basis in $L^2_\mu(\mathbf{T})$ if and only if $\log \mu' \notin L^1(\mathbf{T})$, where μ' is the Radon-Nikodym derivative of μ with respect to dm.

Clearly, (14.1) and the fact that the polynomials of degree n have dimension $n+1$ implies

$$\deg(P) = n, \quad P \perp \zeta^j, \quad j = 0, 1, \ldots, n-1 \Rightarrow P = c\Phi_n. \tag{14.2}$$

On $L^2_\mu(\mathbf{T})$ the anti-unitary map $f^*(\zeta) := \zeta^n \overline{f(\zeta)}$ (which depends on n) is naturally defined. The set of polynomials of degree at most n is left invariant:

$$P(z) = \sum_{j=0}^{n} p_j z^j \Rightarrow P^*(z) = \sum_{j=0}^{n} \bar{p}_{n-j} z^j. \tag{14.3}$$

(14.2) now implies

$$\deg(P) = n, \quad P \perp \zeta^j, \quad j = 1, \ldots, n \Rightarrow P = c\Phi_n^*. \tag{14.4}$$

Szegő recurrences and Verblunsky coefficients

A key feature of the unit circle is that the multiplication operator $Uf = zf$ in $L^2_\mu(\mathbf{T})$ is unitary. So the difference $\Phi_{n+1}(z) - z\Phi_n(z)$ is of degree n and orthogonal to z^j for $j = 1, 2, \ldots, n$, and by (14.4)

$$\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n \Phi_n^*(z) \tag{14.5}$$

with some complex numbers α_n, called the *Verblunsky coefficients*. (14.5) is known as the *Szegő recurrences* after its first occurrence in the Szegő book [104]. (14.5) at $z = 0$ implies

$$\alpha_n = -\overline{\Phi_{n+1}(0)}. \tag{14.6}$$

Applying (14.3) to (14.5) yields

$$\Phi^*_{n+1}(z) = \Phi^*_n(z) - \alpha_n z \Phi_n(z). \tag{14.7}$$

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

$$\|\Phi_{n+1}\|^2 = (1-|\alpha_n|^2)\|\Phi_n\|^2, \qquad \|\Phi_n\|^2 = \prod_{j=0}^{n-1}(1-|\alpha_j|^2), \tag{14.8}$$

and so $|\alpha_n| < 1$. Since it arises often, define

$$\rho_j := \sqrt{1-|\alpha_j|^2}, \qquad 0 < \rho \le 1, \qquad |\alpha_j|^2 + \rho_j^2 = 1. \tag{14.9}$$

Using (14.8) one can get the recursion relations for φ_n written in matrix form

$$\begin{pmatrix} \varphi_{n+1}(z) \\ \varphi^*_{n+1}(z) \end{pmatrix} = A(z,\alpha_n) \begin{pmatrix} \varphi_n(z) \\ \varphi^*_n(z) \end{pmatrix}, \quad A(z,\alpha) = \frac{1}{\rho} \begin{pmatrix} z & -\bar{\alpha} \\ -z\alpha & 1 \end{pmatrix} \tag{14.10}$$

or

$$\begin{pmatrix} \varphi_{n+1}(z) \\ \varphi^*_{n+1}(z) \end{pmatrix} = T_{n+1}(z) \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad T_p(z) = A(z,\alpha_{p-1})\ldots A(z,\alpha_0) \tag{14.11}$$

known as the *transfer matrix.*

Let $\mathbf{D}^\infty$ be the set of complex sequences $\{\alpha_j\}_{j=0}^\infty$ with $|\alpha_j| < 1$. The map $\mathcal{S}$, from $\mu \to \{\alpha_j(\mu)\}_{j=0}^\infty$ is a well defined map from the set $\mathcal{P}$ of nontrivial probability measures on $\mathbf{T}$ to $\mathbf{D}^\infty$. The following fundamental result is proved in [112].

Theorem 14.1 (Verblunsky's Theorem). *$\mathcal{S}$ is a bijection.*

As a matter of fact, $\mathcal{S}$ is a homeomorphism if $\mathcal{P}$ is given the weak* topology and in $\mathbf{D}^\infty$ the topology of component convergence is considered.

The following result is usually attributed to Szegő , Kolmogorov and Krein.

Theorem 14.2 *For any nontrivial measure μ, the following are equivalent*
(i) $\lim_{n\to\infty} \|\Phi_n\| = 0$,
(ii) $\sum_{n=0}^\infty |\alpha_n|^2 = \infty$,
(iii) $\{\varphi_n\}_{n=0}^\infty$ *is a basis for* $L^2_\mu(\mathbf{T})$,
(iv) $\int_{\mathbf{T}} \log \mu' \, dm = -\infty$, *that is* $\log \mu' \notin L^1(\mathbf{T})$.

Bernstein – Szegő approximation

An interesting problem is to identify measures μ with finite sequences of Verblunsky coefficients: $\alpha_j(\mu) = 0$ for all large enough j.

Theorem 14.3 *Let μ be a nontrivial probability measure on $\mathbf{T}$ with orthonormal polynomials φ_n. Let*

$$\mu_n := \frac{dm}{|\varphi_n(\zeta)|^2}\,. \tag{14.12}$$

Then μ_n are probability measures with

$$\alpha_j(\mu_n) = \alpha_j(\mu), \quad j = 0, 1, \ldots, n-1; \qquad \alpha_j(\mu_n) = 0, \quad j \geq n. \tag{14.13}$$

This result is often credited to Geronimus [35] even though it was proven (in different terms) by Verblunsky [113] 10 years earlier. Since, for each fixed j, $\alpha_j(\mu_n) \to \alpha_j(\mu)$ (indeed, they are equal for $n > j$), $\mu_n \to \mu$ weakly since $\mathcal{S}$ is a homeomorphism. It was Verblunsky who also found the Carathéodory function for measures (14.12):

$$F(z, \mu_n) = \int_{\mathbf{T}} \frac{\zeta + z}{\zeta - z}\, d\mu_n(\zeta) = \frac{\psi_n^*(z)}{\varphi_n^*(z)}\,,$$

where the second kind polynomials ψ_n are the orthonormal polynomials with respect to the measure μ_{-1} with $\alpha_j(\mu_{-1}) = -\alpha_j(\mu)$.

In fact, the measures with finite sequences of Verblunsky coefficients are exactly those of the form $\mu = c|P(\zeta)|^{-2}\, dm$, where c is picked to make μ a probability measure, and P is a polynomial of degree n with all zeros in $\mathbf{D}$.

15 Schur, Geronimus, Khrushchev

Schur functions and algorithm

Given a probability measure μ on $\mathbf{T}$, define the *Carathéodory function* by

$$F(z, \mu) := \int_{\mathbf{T}} \frac{\zeta + z}{\zeta - z}\, d\mu(\zeta) = 1 + 2\sum_{n=1}^{\infty} \beta_n z^n, \quad \beta_n = \int_{\mathbf{T}} \zeta^{-n} d\mu \tag{15.1}$$

the moments of μ. F is an analytic function in $\mathbf{D}$ which obeys $\Re F > 0$, $F(0) = 1$. The *Schur function* is then defined by

$$f(z, \mu) = \frac{F(z) - 1}{z(F(z) + 1)}, \qquad F(z) = \frac{1 + zf(z)}{1 - zf(z)}\,, \tag{15.2}$$

which is an analytic function in $\mathbf{D}$ with $\sup_{\mathbf{D}} |f(z)| \leq 1$. So a one-one correspondence can be easily set up between the three classes (probability measures, Carathéodory and Schur functions). Under this correspondence μ is trivial, that is, supported on a finite set, if and only if the associate Schur function is a finite Blaschke product.

Let us proceed with the Schur algorithm. Given a Schur function $f = f_0$ which is not a finite Blaschke product, define inductively

$$f_{n+1}(z) = \frac{f_n(z) - \gamma_n}{z(1 - \bar{\gamma}_n z f_n(z))}, \qquad \gamma_n = f_n(0). \tag{15.3}$$

It is clear that the sequence $\{f_n\}$ is an *infinite* sequence of Schur functions (called the n-th Schur iterates) and neither is a finite Blaschke product. The numbers $\{\gamma_n\}$ are called the *Schur parameters*.

The fundamental paper of I. Schur [78] had appeared a few years before Szegő introduced the notion of orthogonal polynomials on the unit circle (OPUC). Amazingly, neither of them benefited from the results of the other. And only 20 years later coupled Geronimus [34] these ideas together and came up with the following

Theorem 15.1 (Geronimus' Theorem) *Let μ be a nontrivial probability measure on $\mathbf{T}$, f its Schur function and $\gamma_n(f)$ the Schur parameters of f. Then $\gamma_n(f) = \alpha_n(\mu)$.*

The latter formula explains why a minus and conjugate is taken in (14.5). As a straightforward consequence of this result we see that $\gamma_j(f) = \gamma_j(g)$ for all j implies $f = g$. Furthermore, a nice relation between the moments β_n from (15.1) of the measure (Taylor coefficients of the Carathéodory function) and the Schur parameters (Verblunsky coefficients) is given by

$$\beta_n = \alpha_{n-1} \prod_{j=0}^{n-2} (1 - |\alpha_j|^2) + \text{polynomial in } (\alpha_0, \bar{\alpha}_0, , \ldots, \alpha_{n-2}, \bar{\alpha}_{n-2}).$$

Khrushchev's theory

In two remarkable papers [42, 43] Khrushchev found deep connections between Schur iterates and the structure of OPUC. A key input for the theory is

Theorem 15.2 (Khrushchev's formula) *The Schur function for the measure $|\varphi_n|^2 d\mu$ is given by the product $b_n(z) f_n(z)$, where f_n is the n-th Schur iterate, and b_n is the finite Blaschke product*

$$b_n(z) = \frac{\varphi_n(z)}{\varphi_n^*(z)}.$$

The most important consequence of Khrushchev's formula is

Theorem 15.3 *The essential support of the a.c. part of μ is all of $\mathbf{T}$ if and only if*

$$\lim_{n\to\infty} \int_{\mathbf{T}} |f_n(\zeta)|^2 \, dm = 0,$$

dm is the normalized Lebesgue measure on $\mathbf{T}$*.*

Here are some other important results of Khrushchev's theory.

Theorem 15.4

$$* - \lim_{n\to\infty} |\varphi_n|^2 d\mu = dm \Leftrightarrow \lim_{n\to\infty} \alpha_{n+j}\bar{\alpha}_n = 0, \quad j = 1, 2, \dots .$$

Define the n-th Schur approximate $f^{[n]}$ by

$$\gamma_j(f^{[n]}) = \gamma_j(f), \quad j = 1, 2, \dots, n \qquad \gamma_j(f^{[n]}) = 0, \quad j \geq n.$$

Theorem 15.5 *Let* $f^{[n]}$ *be the n-th Schur approximate. Then*

$$\lim_{n\to\infty} \int_{\mathbf{T}} |f^{[n]}(\zeta) - f(\zeta)|^2 \, dm = 0$$

if and only if either μ *is purely singular or* $\alpha_n(\mu) \to 0$*.*

16 Szegő's theory and extensions

Szegő's theorems may well be the most celebrated in OPUC. They have repeatedly served as a source for further development. For historical reasons one should state them in terms of Toeplitz determinants, $D_n(\mu)$. This is defined as the determinant of the $(n+1)\times(n+1)$ matrix $\{\beta_{k-j}\}_{0\leq k,j\leq n}$ with moments β's given in (15.1). The invariance of such determinants under triangular change of basis implies (using (14.8))

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

and so

$$\begin{aligned} S(\mu) &= \lim_{n\to\infty} (D_n(\mu))^{1/n} = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2), \\ G(\mu) &= \lim_{n\to\infty} \frac{D_n(\mu)}{S^{n+1}(\mu)} = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)^{-j-1}. \end{aligned}$$

S is always defined and is a nonnegative number. G is defined as long as $S > 0$ and is finite if and only if $\sum_{j=0}^{\infty} j|\alpha_j|^2 < \infty$.

Szegő's theorems express S and G in terms of the a.c. and singular components of the Lebesgue decomposition of μ: $\mu = w\, dm + \mu_s$, $w \in L^1(\mathbf{T})$.

Theorem 16.1 (Szegő's Theorem).

$$S(\mu) = \prod_{j=0}^{\infty}(1 - |\alpha_j|^2) = \exp\left(\frac{1}{2\pi}\int_{\mathbf{T}} \log w(\zeta)dm\right). \tag{16.1}$$

Szegő proved this when $\mu_s = 0$ in 1915 (in his very first paper!). The result does not depend on μ_s - this was shown by Verblunsky [113].

It is immediate from Szegő's Theorem that

$$\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \Leftrightarrow \log w \in L^1(\mathbf{T}). \tag{16.2}$$

The equivalent conditions (16.2) are called the *Szegő condition*, and the corresponding class of measures is known as the *Szegő class*. Within this class the *Szegő function*

$$D(z, w) = \exp\left(\frac{1}{4\pi}\int_{\mathbf{T}} \frac{\zeta + z}{\zeta - z} \log w(\zeta)dm\right), \qquad |z| < 1 \tag{16.3}$$

is well defined. Standard boundary value theory implies $D(\zeta) = \lim_{r\uparrow 1} D(rz)$ exists almost everywhere and $|D(\zeta)|^2 = w(\zeta)$ a.e. The main asymptotic result (due to Szegő) claims that

$$\lim_{n\to\infty} \varphi_n^*(z) = D^{-1}(z)$$

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

Theorem 16.2 (Strong Szegő Theorem). *If $\mu_s = 0$ and the Szegő condition holds, then*

$$G(\mu) = \prod_{j=0}^{\infty}(1 - |\alpha_j|^2)^{-j-1} = \exp\left(\sum_{n=0}^{\infty} n|w_n|^2\right),$$

where w_n are the Fourier coefficients of $\log w$.

For the up-to-date approach to this result see [91].

In a series of papers [56, 57, 58] Máté, Nevai, and Totik extended some parts of Szegő's theory to the cases where the Szegő condition fails. Their main result can be viewed as a comparative asymptotics.

Theorem 16.3 *Let $\mu = wdm + \mu_s$ be a nontrivial probability measure on $\mathbf{T}$ obeying $w > 0$ a.e. Consider another probability measure $\nu = gd\mu$ for a nonnegative function $g \in L^1(\mu)$. Suppose next, that there is a polynomial Q so that $g^{\pm}Q \in L^\infty(\mu)$. Then*

$$\lim_{n\to\infty} \frac{\varphi_n^*(z,\nu)}{\varphi_n^*(z,\mu)} = D(z, g^{-1}),$$

(D is defined in (16.3)*), uniformly on compact subsets of the unit disk* $\mathbf{D}$.

Another natural extension of Szegő's theory deals the ratio asymptotics.

Theorem 16.4 *Let μ be a nontrivial probability measure on* $\mathbf{T}$*. Suppose*

$$\lim_{n\to\infty} \frac{\Phi_{n+1}^*(z)}{\Phi_n^*(z)} = G(z) \tag{16.4}$$

exists uniformly on compacts of $\mathbf{D}$*. Then either $G \equiv 1$ or*

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

for some $\lambda \in \mathbf{T}$ and $a \in (0,1]$.

(16.4) *holds with $G = G_{a,\lambda}$ if and only if $\alpha_n(\mu)$ obeys the López condition*

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

In this case the essential support of μ is an arc, and (16.4) *holds uniformly on compact subsets of* $\mathbf{C}\backslash \operatorname{supp} \mu$.

The first statement is due to Khrushchev [43] and the second one to Barrios and López [9].

17 CMV matrices

One of the most interesting developments in the theory of OPUC in recent years is the discovery by Cantero, Moral, and Velázquez [15] of a matrix realization for multiplication by ζ on $L^2(\mathbf{T}, \mu)$ which is of finite band size (i.e., $|\langle \zeta \chi_m, \chi_n \rangle| = 0$ if $|m - n| > k$ for some k); in this case, $k = 2$ to be compared with $k = 1$ for the Jacobi matrices which correspond to the real line case. The CMV basis (complete, orthonormal system) $\{\chi_n\}$ is obtained by orthonormalizing the sequence $1, \zeta, \zeta^{-1}, \zeta^2, \zeta^{-2}, \ldots$, and the matrix, called the *CMV matrix*,

$$\mathcal{C}(\mu) = \|C_{n,m}\|_{m,n=0}^{\infty} = \langle \zeta \chi_m, \chi_n \rangle, \qquad m, n \in \mathbf{Z}_+$$

is five-diagonal. Remarkably, the χ's can be expressed in terms of φ's and φ^*'s

$$\chi_{2n} = z^{-n} \varphi_{2n}^*(z), \qquad \chi_{2n+1} = z^{-n} \varphi_{2n+1}(z), \quad n \in \mathbf{Z}_+,$$

and the matrix elements in terms of α's and ρ's: $\mathcal{C} = LM$ where L, M are block diagonal matrices

$$L = \operatorname{Diag}(\Theta_0, \Theta_2, \Theta_4, \ldots), \qquad M = \operatorname{Diag}(1, \Theta_1, \Theta_3, \ldots) \tag{17.1}$$

with

$$\Theta_j = \begin{pmatrix} \overline{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{pmatrix}, \qquad j = 0, 1, \ldots \tag{17.2}$$

(the first block of M is 1×1). By $\mathcal{C}_0$ we denote the CMV matrix for the Lebesgue measure dm.

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:

$$\Phi_n(z) = \det(zI_n - \mathcal{C}^{(n)}), \tag{17.3}$$

where $\mathcal{C}^{(n)}$ is the principal $n \times n$ block of $\mathcal{C}$. The CMV representation provides one of the proofs of Verblunsky's theorem.

The natural extension of $L, M, \mathcal{C}$ to doubly infinite matrices proves helpful in some problems related to periodic and stochastic Verblunsky coefficients. For the standard basis $\{e_j\}_{j\in\mathbf{Z}}$ denote by E_k the 2-dimensional subspace spanned by $\{e_k, e_{k+1}\}$, so

$$\ell^2(\mathbf{Z}) = \bigoplus_{j\in\mathbf{Z}} E_{2j} = \bigoplus_{j\in\mathbf{Z}} E_{2j+1}.$$

Let the operator Θ_k act in E_k by (17.2). We come to

$$\hat{L} = \bigoplus_{j\in\mathbf{Z}} \Theta_{2j}, \quad \hat{M} = \bigoplus_{j\in\mathbf{Z}} \Theta_{2j+1}, \qquad \hat{\mathcal{C}} = \hat{L}\hat{M}. \tag{17.4}$$

The CMV matrices play much the same role in the study of OPUC that Jacobi matrices in orthogonal polynomials on the real line.

Part III
Simon's contribution

18 Analysis of CMV matrices

CMV matrices and spectral analysis

Perturbation theory involves looking at similarities of measures when their Verblunsky coefficients are close in some suitable sense. The CMV matrices provide a powerful tool for the comparison of properties of two measures μ_1 and μ_2 if we have some information about $\alpha_n(\mu_2)$ as a perturbation of $\alpha_n(\mu_1)$. Of course, this idea is standard in the real line setting where Jacobi matrices do the job.

Put

$$\alpha_n(\mu_1,\mu_2) := |\alpha_n(\mu_1) - \alpha_n(\mu_2)|, \qquad \rho_n(\mu_1,\mu_2) := |\rho_n(\mu_1) - \rho_n(\mu_2)|.$$

An easy estimate using the LM factorization shows with $\|\cdot\|_p$ the $\mathcal{S}_p$ trace ideal norm

Lemma 18.1 *For all* $1 \le p \le \infty$

$$\|\mathcal{C}(\mu_1) - \mathcal{C}(\mu_2)\|_p \le 6 \left(\sum_{n=0}^{\infty} \alpha_n^p(\mu_1, \mu_2) + \rho_n^p(\mu_1, \mu_2) \right)^{1/p}. \tag{18.1}$$

For $p = \infty$ the right hand side of (18.1) is interpreted as $\sup_n \max(\alpha_n(\mu_1, \mu_2), \rho_n(\mu_1, \mu_2))$.

This result allows one to translate the ideas of Simon – Spencer to a new operator theoretic proof of the following result sometimes called Rakhmanov's lemma: μ is purely singular as long as $\limsup_n |\alpha_n| = 1$. It is also a key ingredient in proving the following (see [79, Section 4.3])

Theorem 18.2 *Let* $\mu_j = w_j dm + \mu_{j,s}, \;\; j = 1, 2$.

1. *If* $\lim_{n\to\infty} \alpha_n(\mu_1, \mu_2) = 0$, *then the supports of* μ_1 *and* μ_2 *have identical sets of limit points:* $(\mathrm{supp}\mu_1)' = (\mathrm{supp}\mu_2)'$.
2. *If* $\sum_n \alpha_n(\mu_1, \mu_2) < \infty$, *then up to sets of Lebesgue measure zero*

$$\{\zeta : w_1(\zeta) \ne 0\} = \{\zeta : w_2(\zeta) \ne 0\}.$$

Concerning the essential spectrum of CMV operators see [47]

CMV matrices and the Szegő function

An intimate relation between the CMV matrices and the Szegő functions (16.3) is presented in [79, Section 4.2]. Let

$$\log D(z) = \frac{1}{2} w_0 + \sum_{n=1}^{\infty} w_n z^n, \qquad w_k = \int_{\mathbf{T}} \log w(\zeta) \zeta^{-k} \, dm \tag{18.2}$$

Fourier coefficients of $\log w$.

Theorem 18.3 (i). *Assume that the Verblunsky coefficients* $\{\alpha_n\} \in \ell^2$. *Then*

1. $\mathcal{C} - \mathcal{C}_0$ *is Hilbert–Schmidt;*
2. $$\frac{D(0)}{D(z)} = \det{}_2 \left(\frac{I - z\overline{\mathcal{C}}}{I - z\overline{\mathcal{C}}_0} \right) e^{z w_1}, \quad w_1 = \alpha_0 - \sum_{n=1}^{\infty} \alpha_n \overline{\alpha}_{n-1}$$
   with $\det_2$ *being renormalized determinant for the Hilbert – Schmidt class;*
3. $$\overline{w}_n = \frac{\mathrm{Tr}(\mathcal{C}^n - \mathcal{C}_0^n)}{n}, \qquad n \ge 2.$$

(ii). *Assume that* $\{\alpha_n\} \in \ell^1$. *Then*

1. $\mathcal{C} - \mathcal{C}_0$ *is trace class;*

2.
$$\frac{D(0)}{D(z)} = \det\left(\frac{I - z\overline{\mathcal{C}}}{I - z\overline{\mathcal{C}}_0}\right);$$

3.
$$\overline{w}_n = \frac{\mathrm{Tr}(\mathcal{C}^n - \mathcal{C}_0^n)}{n}, \qquad n \geq 1.$$

In either case
$$\overline{w}_n = \sum_{j=0}^{\infty} \frac{(\mathcal{C}^n)_{jj}}{n}.$$

19 Zeros

Limit sets of zeros

The structure of zero sets for OPUC is another fascinating topic of the theory. Given a nontrivial probability measure $\mu \in \mathcal{P}$ denote by $Z_n(\mu)$ the zero set for Φ_n:

$$Z_n(\mu) = \{z_{jn}\}_{j=1}^n, \qquad |z_{nn}| \leq |z_{n-1,n}| \leq \ldots \leq |z_{1,n}|, \qquad \Phi_n(z_{jn}, \mu) = 0.$$

The basic property of zeros reads that $|z_{1,n}| < 1$, i.e., $Z_n(\mu) \subset \mathbf{D}$. Indeed, let $z_0 \in Z_n$ and define $P = \Phi_n/(z - z_0)$. Since $\deg P = n-1$, $P \perp \Phi_n$ and so

$$\|P\|^2 = \|zP\|^2 = \|z_0 P + \Phi_n\|^2 = |z_0|^2 \|P\|^2 + \|\Phi_n\|^2.$$

Hence $\|\Phi_n\|^2 = (1 - |z_0|^2)\|P\|^2$, as needed.

In 1988 Alfaro and Vigil [2] answering a question of P. Turán showed that for an arbitrary sequence of points $\{z_k\}$ in $\mathbf{D}$ there exists a unique measure $\mu \in \mathcal{P}$ with $\Phi_n(z_n, \mu) = 0$. Hence the total set of zeros of Φ_n's $Z_\infty(\mu) = \cup_n Z_n(\mu)$ can be dense in $\mathbf{D}$. The vast generalization of Alfaro – Vigil theorem is due to Simon – Totik [95].

Theorem 19.1 *For an arbitrary sequence of points* $\{z_k\}$ *in* $\mathbf{D}$ *and arbitrary sequence of positive integers* $0 < m_1 < m_2 < \ldots$ *there exists* $\mu \in \mathcal{P}$ *such that* $\Phi_{m_k}(z_j, \mu) = 0$ *for* $j = m_{k-1}+1, \ldots, m_k$.

The following consequence of this result may seem kind of amazing. Let a measure μ belong to the class of nontrivial probability measures $\mathcal{P}$. Consider the sequence $\{\nu_n(\mu)\}_{n\geq 1}$ of *normalized counting measures for zeros* of Φ_n, that is,

$$\mathrm{supp}\,\nu_n = Z_n, \quad \nu_n\{z_{jn}\} = \frac{l(z_{jn})}{n} \tag{19.1}$$

with $l(z_{jn})$ equal to the multiplicity of the zero z_{jn}. Let $\mathcal{M}_+(\bar{\mathbf{D}})$ be a space of probability measures on $\bar{\mathbf{D}}$ endowed with the weak* topology. A measure $\mu \in \mathcal{P}$ is said to be *universal* if for each $\sigma \in \mathcal{M}_+(\bar{\mathbf{D}})$ there is a sequence of indices n_j such that $\nu_{n_j}(\mu)$ converges to σ as $j \to \infty$ in weak* topology. The existence of universal measures is proved in [95, Corollary 3].

It is known that zeros of $\Phi_n(\mu)$ cluster to the support of the orthogonality measure μ as long as the support is not the whole circle. The situation changes dramatically if $\operatorname{supp} \mu = \mathbf{T}$ (see, e.g., the Lebesgue measure). By the Alfaro – Vigil theorem zeros of Φ_n can cluster to all points of $\bar{\mathbf{D}}$. Denote by

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

the point set of limit points for the zeros of all Φ_n (weakly attracting points). Let $Z_w = \{Z_w(\mu)\}_{\mu\in\mathcal{P}}$ be the class of all such point sets. So, $\bar{\mathbf{D}} \in Z_w$. It turns out that Z_w is rich enough. More precisely, each compact subset K of $\mathbf{D}$ belongs to Z_w, and the same is true if $K \supset \mathbf{T}$ ([95, Theorem 4]). On the other hand, $K = [1/2, 1]$ is not in Z_w.

Similarly, denote by

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

the point set of strongly attracting point 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 [3], that if $0 \in Z_s(\mu)$ for some $\mu \in \mathcal{P}$, then $Z_s(\mu)$ is at most a countable set which converges to the origin. So, e.g., the disk $\{|z| \le 1/2\}$ is not in Z_s.

Mhaskar – Saff and clock theorems for zeros

A remarkable theorem of Mhaskar and Saff [64] provides some information about the limit points (in the space $\mathcal{M}_+(\bar{\mathbf{D}})$) of the sequence of counting measures of zeros associated with a nontrivial probability measure $\mu \in \mathcal{P}$ in the case when Verblunsky coefficients tend to zero fast enough.

Theorem 19.2 (Mhaskar – Saff). *Let*

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

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.

Simon suggested a new approach to this result based on CMV matrices instead of potential theory. A key relation which links two subjects is obvious from (17.3):

$$\frac{1}{n} \operatorname{Tr}(\mathcal{C}^{(n)})^k = \int_{\mathbf{D}} z^k d\nu_n(\mu).$$

Theorem 19.3 *Let* $\sum_{j=0}^{n-1} |\alpha_n(\mu_1) - \alpha_n(\mu_2)| = o(n)$*,* $n \to \infty$*. Then for each* $k \in \mathbf{Z}_+$

$$\lim_{n\to\infty} \int_{\mathbf{D}} z^k \left(d\nu_n(\mu_1) - d\nu_n(\mu_2)\right) = 0.$$

In particular, if $\nu_n(\mu_1)$ *tends to* ν *and* γ *is any limit point of* $\nu_n(\mu_2)$*, then*

$$\int_{\mathbf{D}} z^k d\gamma = \int_{\mathbf{D}} z^k d\nu.$$

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

$$\alpha_n(\mu) = Cb^n + O((b\delta)^n); \qquad C \in \mathbf{C}\backslash 0, \quad 0 < b, \delta, < 1. \tag{19.2}$$

The following result is proved in [82].

Theorem 19.4 *Under the assumption* (19.2) $\Phi_n(\mu)$ *has a finite number* J *of "spurious" zeros off the circle* $|z| = b$ *for all large* n*. Furthermore, let for* $j = 1, 2, \ldots, n-J$

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

be the other zeros. Then the limit relations hold

$$\sup_{1\le j\le n-J} ||z_{jn}| - b| = O\left(\frac{\log n}{n}\right), \quad n \to \infty; \tag{19.3}$$

$$\sup_{1\le j\le n-J} n\left|\Theta_{j+1,n} - \Theta_{jn} - \frac{2\pi}{n}\right| = o(1), \quad n \to \infty; \tag{19.4}$$

$$\frac{|z_{j+1,n}|}{|z_{jn}|} = 1 + O\left(\frac{1}{n\log n}\right), \quad n \to \infty. \tag{19.5}$$

Note that (19.4)–(19.5) imply $\lim_n n|z_{j+1,n} - z_{jn}| = 2\pi b$.
Amazingly, the spurious zeros also follow the clock pattern!

In [83] Simon treats the more general case

$$\alpha_n(\mu) = \sum_{l=1}^{m} C_l e^{in\Theta_l} b^n + O((b\delta)^n).$$

20 Spectral theory in special classes

High order Szegő theorem

Simon came up with the idea to extend Szegő's Theorem by allowing "Pollaczek singularities" (so all quantities in (16.1) may be infinite). His result can be viewed as the first order Szegő's Theorem: for any $\zeta_0 \in \mathbf{T}$

$$|\zeta - \zeta_0|^2 \log w \in L^1(\mathbf{T}) \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_{j+1} - \overline{\zeta}_0 \alpha_j|^2 + |\alpha_j|^4 < \infty.$$

Moreover, there is a precise formula for this case similar to the second equality in (16.1) [79, Section 2.8]. The second order Szegő's Theorem appeared in [97]. Let $\zeta_k \in \mathbf{T}$, $k = 1, 2$. Then for $\zeta_1 \neq \zeta_2$

$$|\zeta - \zeta_1|^2 |\zeta - \zeta_2|^2 \log w \in L^1(\mathbf{T}) \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_{j+2} - (\overline{\zeta}_1 + \overline{\zeta}_2)\alpha_{j+1} + \overline{\zeta_1 \zeta_2} \alpha_j|^2 + |\alpha_j|^4 < \infty,$$

and for $\zeta_1 = \zeta_2$

$$|\zeta - \zeta_1|^4 \log w \in L^1(\mathbf{T}) \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_{j+2} - 2\overline{\zeta}_1 \alpha_{j+1} + \overline{\zeta_1^2} \alpha_j|^2 + |\alpha_j|^6 < \infty.$$

The general nicely looking conjecture called the *the higher-order Szegő's Theorem* is put forward in [79, Section 2.8]. Given $\zeta_k \in \mathbf{T}$, $k = 1, \dots, n$ and $\zeta_p \neq \zeta_q$, $p \neq q$, define a polynomial

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

and put $m = 1 + \max_k m_k$. Then

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

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

Baxter's theorem

The celebrated paper [10] by Baxter appears to be one of the cornerstones of OPUC theory. He was interested in general complex-valued weight functions and corresponding non-Hermitian Toeplitz matrices of moments in connection with two sets of Verblunsky coefficients. Applied to the case of OPUC his basic result looks as follows.

Theorem 20.1 (Baxter). *Let $\mu = w dm + \mu_s$ be a nontrivial probability measure on $\mathbf{T}$ with Verblunsky coefficients $\alpha_n(\mu)$ and moments $\beta_n(\mu)$. Then the following are equivalent:*

1. $\sum_{j=0}^{\infty} |\alpha_j(\mu)| < \infty$;

2. $\sum_{j=0}^{\infty} |\beta_j(\mu)| < \infty$, $\mu_s = 0$, w *is continuous, and* $\min w(\zeta) > 0$.

A far reaching extension of Baxter's theorem is suggested in [79, Chapter 5].

A function $\nu(n)$, $n \in \mathbf{Z}$, is called a *Beurling weight* if

$$\nu(-n) = \nu(n), \qquad \nu(n) \geq 1, \qquad \nu(m+n) \leq \nu(m)\nu(n),$$

and a *strong Beurling weight* if in addition

$$A(\nu) := \lim_{n\to\infty} \frac{\log \nu(n)}{n} = 0.$$

Examples include $\nu(n) = (1+|n|)^{\alpha}$ (strong) and $\nu(n) = e^{\alpha|n|}$, $\alpha \geq 0$.
Given a Beurling weight ν, the *Beurling algebra* ℓ_ν is the Banach algebra of two-sided sequences $\{a(n)\}_{n\in\mathbf{Z}}$ with standard operations (addition and convolution) and

$$\|a\|_\nu := \sum_{n\in\mathbf{Z}} |a(n)|\nu(n) < \infty.$$

Theorem 20.2 *Let ν be a strong Beurling weight. Then the following are equivalent:*

1. $\{\alpha_n(\mu)\} \in \ell_\nu$;

2. $\{\beta_n(\mu)\} \in \ell_\nu$, $\mu_s = 0$, w *is continuous, and* $\min w(\zeta) > 0$.

A number of equivalent conditions are displayed in [79, Theorem 5.2.2].

As an immediate consequence of this result, Simon obtains

Corollary 20.3 *Let $k \in \mathbf{N}$. Then*

1. *If* $\sum_n (n+1)^k |\alpha_n(\mu)| < \infty$, *then* $\mu = wdm$ *with* $\inf w(\zeta) > 0$ *and* w *a* C^k *function.*

2. *If* $k \geq 2$ *and* $\mu = wdm$ *with* $\inf w(\zeta) > 0$, *and* w *is a* C^k *function, then* $\sum_n (n+1)^{\beta} |\alpha_n(\mu)| < \infty$ *for any* $0 \leq \beta < k-1$.

In particular, $\mu = wdm$ *with* $\inf w(\zeta) > 0$, *and* w *a* C^∞ *function if and only if for all* k

$$|\alpha_n(\mu)| \leq C_k (n+1)^{-k}.$$

B. Golinskii – I. Ibragimov condition

In 1971 B. Golinskii and I. Ibragimov proved that if the Verblunsky coefficients $\{\alpha_n(\mu)\}$ of a measure μ obey $\sum_n n|\alpha_n(\mu)|^2 < \infty$ then μ is absolutely continuous. The extension of their ideas led Simon [94] to the class of measures for which $\sum_{n=0}^N n|\alpha_n(\mu)|^2$ diverges logarithmically. More precisely, Simon's condition is

$$\sum_{n=0}^{N} n|\alpha_n(\mu)|^2 \le A \log N + C \tag{20.1}$$

with A, C constants. He proved in [79, Theorem 6.1.7], that μ is absolutely continuous as long as (20.1) holds with some $A < 1/4$. On the other hand, for any $A > 1/4$ there are examples with $\mu_s \neq 0$, that is, $1/4$ is optimal. As a matter of fact, for $A \geq 1/4$ one should distinguish the pure point component μ_{pp} of $\mu_s = \mu_{pp} + \mu_{sc}$ from the singular continuous one, μ_{sc}. For instance, it is shown in [79, Corollary 2.7.6] that $\mu_{pp} = 0$ as long as (20.1) holds with $A = 1/4$. It turns out that for measures (20.1) with some $A < \infty$ the support of μ_{pp} consists of at most $[4A]$ points, $[x]$ being an integer part of x (see [80, Theorem 10.12.7]). Recently, Damanik [17], answering a question of Simon, proved that, indeed, $\mu_{sc} = 0$ whenever (20.1) holds for some $A < \infty$.

Sparse Verblunsky coefficients

In [41] Kiselev, Last, and Simon presented a thorough analysis of continuum and discrete Schrödinger operators with sparse potentials. In [80, Section 12.5] Simon has found analogs of these results for OPUC. A set of Verblunsky coefficients is said to be *sparse* if $\alpha_j \neq 0$ only for $j \in \{n_k\}_{k\geq 1}$ where $n_{k+1} - n_k \to \infty$ perhaps at some specific rate. The first result of the "sparse flavor" is due to Khrushchev [42], who proved that under the so-called Máté – Nevai condition

$$\lim_{n\to\infty} \alpha_n(\mu)\alpha_{n+l}(\mu) = 0$$

for each fixed $l = 1, 2, \ldots$ (which certainly holds for sparse sets and characterizes Rakhmanov's class of measures) μ is singular if $\limsup_{n\to\infty} |\alpha_n(\mu)| > 0$.

The results on sparse Verblunsky coefficients is as follows (see [80, Theorems 12.5.1, 12.5.2]).

Theorem 20.4 *Let $\{n_k\}$ be a monotone sequence of positive integers with* $\liminf_{k\to\infty} n_{k+1}/n_k > 1$. *Suppose*

$$\alpha_j(\mu) = 0, \quad j \notin \{n_k\}, \qquad \sum_{j=0}^{\infty} |\alpha_k(\mu)|^2 < \infty.$$

Then $\operatorname{supp}\mu = \mathbf{T}$ *and $\mu_s = 0$, that is, μ is absolutely continuous. Moreover, $\mu = w dm$ with $w^{\pm 1} \in L^p(\mathbf{T})$ for all $1 \leq p < \infty$.*

Theorem 20.5 *Let $\{n_k\}$ be a monotone sequence of positive integers with* $\liminf_{k\to\infty} n_{k+1}/n_k = \infty$*. Suppose*

$$\alpha_j(\mu) = 0, \quad j \notin \{n_k\}, \quad \lim_{j\to\infty} \alpha_j = 0, \qquad \sum_{j=0}^{\infty} |\alpha_k(\mu)|^2 = \infty.$$

Then $\operatorname{supp}\mu = \mathbf{T}$ *and $\mu_{pp} = \mu_{ac} = 0$, that is, μ is purely singular continuous.*

Dense embedded point spectrum

The story for Schrödinger operators goes back to Naboko [65], who constructed the operator $-\frac{d^2}{dx^2} + V(x)$ with V decaying only slightly slower than $|x|^{-1}$ but there is dense embedded point spectrum. Naboko's method extends to OPUC (see [80, Section 12.3]).

Theorem 20.6 *Let $g(n)$ be an arbitrary function with $0 < g(n) \leq g(n+1)$ and $g(n) \to \infty$ as $n \to \infty$. Let $\{\omega_j\}_{j\geq 0}$ be an arbitrary subset of $\mathbf{T}$ which are multiplicatively rationally independent, that is, for no integers $n_1, \dots, n_k$ other than zeros, is it true that $\prod_{j=1}^{k} \omega_j^{n_j} = 1$. Then there exists a sequence $\{\alpha_j(\mu)\}$ of Verblunsky coefficients with*

$$|\alpha_n(\mu)| \leq \frac{g(n)}{n}$$

for all n so that μ has pure points at each ω_j.

Note that if $g(n) \leq n^{1/2-\varepsilon}$, then $\alpha_n \in \ell^2$ so by Szegő's Theorem, $w > 0$ a.e., that is, the point spectrum is embedded in a.c. spectrum.

Fibonacci subshifts

There is an extensive literature on subshifts for discrete Schrödinger operators. Simon [80, Section 12.8] has analyzed the OPUC analogs of the mostly studied of these subshifts, defined as follows. Pick $\alpha \neq \beta$ in the open unit disk and let $F_1 = \alpha$, $F_2 = \alpha\beta$, and $F_{n+1} = F_n F_{n-1}$ for $n = 2, 3, \dots$. These are finite strings built up from α and β. As F_{n+1} starts with F_n, there is a limit $F(\alpha, \beta)$ which is a one way infinite string $\alpha\alpha\beta\alpha\beta\alpha\cdots$. Let $\mu(\alpha, \beta)$ be the measure that has Verblunsky coefficients (in an infinite string form) $\alpha_0\alpha_1\cdots = F(\alpha, \beta)$.

Theorem 20.7 *The essential support of the measure $\mu(\alpha, \beta)$ is a closed perfect set of Lebesgue measure zero. For fixed α_0, β_0 and a.e. $\lambda \in \mathbf{T}$, the measure $\mu(\lambda\alpha_0, \lambda\beta_0)$ is a pure point measure, with each pure point isolated and the limit points of the pure points form a perfect set of Lebesgue measure zero.*

21 Periodic Verblunsky coefficients

The theory of one-dimensional periodic Schrödinger operators (also known as Hill's equations) and of periodic Jacobi matrices has been extensively developed. The theory up to the 1950's is summarized in the Magnus – Winkler book [54]. There was an explosion of ideas following the KdV revolution, including spectrally invariant flows and Abelian functions on hyperelliptic Riemann surfaces. The ideas have been carried over to the discrete setting of orthogonal polynomials on the real line, see, e.g. Toda [106] and Flaschka – McLaughlin [25].

In the 1940's Geronimus [33] found the earliest results on OPUC with periodic Verblunsky coefficients, that is, for some $p \geq 1$

$$\alpha_{j+p} = \alpha_j, \qquad j \in \mathbf{Z}_+. \tag{21.1}$$

In particular, the case $\alpha_j \equiv \alpha \in \mathbf{D}\backslash\{0\}$ yields OPUC called Geronimus polynomials (see Example 1.6.12 of [79]). Many of the general features for OPUC obeying (21.1) were found in a fundamental series of papers by Peherstorfer and collaborators.

The aforementioned literature on OPUC used little from the work on Hill's equation. A partial link is Geronimo – Johnson [31], which discussed almost periodic Verblunsky coefficients using Abelian functions. See also Geronimo – Gesztesy – Holden [32], which includes work on isospectral flows.

Some new results and approaches for periodic Verblunsky coefficients are presented in [80, Chapter 11].

Discriminant

For Schrödinger operators and Jacobi matrices the discriminant is known to be just the trace of the transfer matrix, which has determinant one in this case. For OPUC, the transfer matrix T_p (14.11) has $\det(T_p(z)) = z^p$, so it is natural to define the *discriminant* by

$$\Delta(z) := \mathrm{Tr}(z^{-p/2}\, T_p(z)) \tag{21.2}$$

which explains why p is normally assumed to be an even number. Δ is known to be real on $\mathbf{T}$, so

$$\Delta(1/\bar{z}) = \overline{\Delta(z)},$$

and $z^{-p/2}\, T_p(z)$ has eigenvalues $\frac{\Delta}{2} \pm i\sqrt{1 - \left(\frac{\Delta}{2}\right)^2}$. In particular, these eigenvalues have magnitude 1, that is, $\sup_m \|T_{mp}(z)\| < \infty$ exactly when $\Delta \in [-2, 2]$.

Theorem 21.1 *There exist* $\{x_j\}_{j=1}^{2p}$, $\{y_j\}_{j=1}^{2p}$ *with*

$$x_1 < y_1 \leq x_2 < y_2 \leq \ldots \leq x_p < y_p \leq x_1 + 2\pi \equiv x_{p+1}$$

so that the solutions of $\Delta(z) = 2$ *(resp.* -2*) are exactly* e^{ix_1}*,* e^{iy_2}*,* $e^{ix_3}, \ldots, e^{ix_p}$ *(resp.* e^{iy_1}*,* e^{ix_2}*,* $e^{iy_3}, \ldots, e^{iy_p}$*) and* $\Delta(z) \in [-2, 2]$ *exactly on the bands*

$$B = \bigcup_{j=1}^{p} B_j, \qquad B_j = \{e^{i\theta} : x_j \le \theta \le y_j\}. \tag{21.3}$$

B *is the essential support of* μ_{ac} *and the only possible singular spectrum are mass points which can occur in open gaps (i.e., nonempty sets of the form* $\{e^{i\theta} : y_j \le \theta \le x_{j+1}\}$*) with at most one mass point in each gap.*

Further properties of Δ are given in the following result.

Theorem 21.2 (i) *For all nonzero* $z \in \mathbf{C}$ *the Lyapunov exponent is*

$$\gamma(z) = \lim_{n\to\infty} \|T_n(z)\|^{1/n} = \frac{1}{2}\log(z) + \frac{1}{p}\left|\frac{\Delta(z)}{2} + \sqrt{\frac{\Delta^2(z)}{4} - 1}\right|, \tag{21.4}$$

where the branch of square root is taken that maximizes the logarithm.
(ii) *The logarithmic capacity of* B (21.3) *is given by*

$$C_B = \prod_{j=0}^{p-1} (1 - |\alpha_j|^2)^{1/p}$$

and $-[\gamma + \log C_B]$ *is the equilibrium potential for* B*.*
(iii) *The density of zeros is the equilibrium measure for* B *and is given in terms of* Δ *by* $d\nu = V d\theta/2\pi$*, where*

$$V(\theta) = \frac{1}{p}\, \frac{|\Delta'(e^{t\theta})|}{\sqrt{4 - \Delta^2(e^{i\theta})}}$$

on B_j*, and* $V = 0$ *on the complement of* B*.*
(iv) $\nu(B_j) = 1/p$*.*

For any $\{\alpha_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ one can define a discriminant $\Delta(z, \{\alpha_j\}_{j=0}^{p-1})$ for the p-periodic Verblunsky coefficients that agree with $\{\alpha_j\}_{j=0}^{p-1}$ for $j = 0, \ldots, p-1$. It turns out (see [80, Theorem 11.10.2]) that the set $\{\alpha_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ for which Δ has all gaps open, i.e., all B_j in (21.3) are disjoint, is a dense open set in $\mathbf{D}^p$ (the case of generic potentials). [80] has two proofs of this result: one is based on Sard's theorem and one is perturbation theoretic calculation.

In 1946 Borg [14] proved two celebrated theorems for Schrödinger operators about the implication of closed gaps. In [80, Section 11.14] Simon gives the analogs of these results for OPUC.

Theorem 21.3 (i) *Let α_j be a periodic sequence of Verblunsky coefficients. Suppose all gaps are closed, i.e., the support of the orthogonality measure is the whole unit circle. Then $\alpha_j \equiv 0$.*
(ii) *Let p be even and $\{\alpha_j\}$ has period $2p$. Then all gaps with $\Delta(z) = -2$ are closed if and only if $\alpha_{j+p} = \alpha_j$. All the gaps with $\Delta(z) = 2$ are closed if and only if $\alpha_{j+p} = -\alpha_j$.*

The isospectral results constitute one of the highlights of the OPUC theory.

Theorem 21.4 *Let $\{\alpha_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ such that $\Delta(z, \{\alpha_j\}_{j=0}^{p-1})$ has k open gaps. Then the set of all $\{\beta_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ with $\Delta(z, \{\beta_j\}_{j=0}^{p-1}) = \Delta(z, \{\alpha_j\}_{j=0}^{p-1})$ is a k-dimensional torus.*

There is one important difference between OPUC and the Jacobi case. In the latter, the infinite gap does not count in the calculation of the dimension of the torus, so it has a dimension equal to the genus of the Riemann surface for the m-function. In the OPUC case, all gaps count and the torus has dimension one more than the genus.

The torus can be defined explicitly in terms of natural additional data associated to $\{\alpha_j\}_{j=0}^{p-1}$. One way to define the data is to analytically continue the Carathéodory function F for the periodic sequence. One cuts $\mathbf{C}$ on the connected components of $\{|\Delta| \leq 2\}$, and forms the two-sheeted Riemann surface associated to $\sqrt{\Delta^2 - 4}$. On this surface F is meromorphic with exactly one pole on each "extended gap". The ends of the gap are branch points and join the two images of the gap into a circle. The p points, one on each gap, are thus p-dimensional torus, and the refined version of the above result is that there is exactly one Carathéodory function associated to a period p set of Verblunsky coefficients with specified poles.

Alternately, the points in the gap are solutions of $\Phi_p = \Phi_p^*$ with sheets determined by whether the points are pure points of the associated measure or not.

22 Random coefficients

Let Ω be the space $\prod_{n=0}^{\infty} \mathbf{D}$, σ_n measures on $\mathbf{D}$ and $\sigma = \prod_{j=0}^{\infty} \sigma_j$ the product measure on Ω. In other words, sequences $\{\alpha_n\}_{n=0}^{\infty}$ of Verblunsky coefficients considered in this section are independent random variables. If in addition $\sigma_j = \sigma_0$ for all j, then they are identically distributed. One is interested in statements that hold when $\alpha_j = \omega_j$ for a.e. $\omega \in \Omega$ with respect to σ. The main result is that, typically, the associated measure μ_ω with $\alpha_j(\mu_\omega) = \omega_j$ is a pure point measure with pure points often dense in $\mathbf{T}$.

For $\lambda \in \mathbf{T}$ let $\mu_{\lambda,\omega}$ be the measure with $\alpha_j(\mu_{\lambda,\omega}) = \lambda\omega_j$. The main result for the i.i.d. case is [80, Theorem 12.6.3].

Theorem 22.1 *Suppose that σ_0 is not supported at a single point. Then for a.e. $(\lambda, \omega) \in \mathbf{T} \times \Omega$ the measure $\mu_{\lambda,\omega}$ has pure point spectrum. If σ_0 is absolutely continuous with respect to the area Lebesgue measure on $\mathbf{D}$, then μ_ω is pure point for a.e. $\omega \in \Omega$.*

Decaying random Verblunsky coefficients

Decaying random potentials were studied starting with Simon [86] who found the first example of Jacobi matrices with $|a_n - 1| + |b_n| \to 0$ and μ purely singular. The pioneering results on decaying Verblunsky coefficients are due to Nikishin [69] and Teplyaev [105]. The detailed account of the subject is presented in [80, Sections 12.6-7].

Now Verblunsky coefficients $\{\alpha_j(\omega)\}$ are assumed to be independent random but not necessarily identically distributed variables, which decay to zero in some sense; at a minimum, the mean value $\mathbf{E}(|\alpha_j(\omega)|^2) \to 0$. The main result states that there is no singular spectrum.

Theorem 22.2 *Let the Verblunsky coefficients $\{\alpha_j(\omega)\}$ be independent random variables with $\mathbf{E}(\alpha_j(\omega)) = 0$ and*

$$\sum_{j=0}^{\infty} \mathbf{E}(|\alpha_j(\omega)|^2) < \infty,$$

that is, the Szegő condition holds a.e.. Then the corresponding orthogonality measure μ_ω is absolutely continuous for a.e. ω.

The situation changes dramatically for the slowly decaying Verblunsky coefficients. Assume that

$$\sup_{\omega,j} |\alpha_j(\omega)| < 1, \qquad \lim_{j\to\infty} \sup_{\omega} |\alpha_j(\omega)| = 0, \tag{22.1}$$

$$\mathbf{E}(\alpha_j(\omega)) = \mathbf{E}(\alpha_j^2(\omega)) = 0, \tag{22.2}$$

and

$$\mathbf{E}(|\alpha_j(\omega)|^2)^{1/2} = \Gamma j^{-\gamma}, \qquad j \geq j_0, \quad 0 < \gamma \leq \frac{1}{2}. \tag{22.3}$$

Theorem 22.3 (i) *Suppose that independent Verblunsky coefficients $\{\alpha_j(\omega)\}$ satisfy* (22.1) – (22.3) *with $\gamma < 1/2$. Then for a.e. pairs $(\lambda, \omega) \in \mathbf{T} \times \Omega$ the measure $\mu_{\lambda,\omega}$ is pure point and dense in $\mathbf{T}$.*
(ii) *Suppose that* (22.1) – (22.3) *hold with $\gamma = 1/2$ and in addition*

$$\sup_{\omega,j} n^{1/2} |\alpha_j(\omega)| < \infty.$$

Then for a.e. pairs $(\lambda, \omega) \in \mathbf{T} \times \Omega$ the measure $\mu_{\lambda,\omega}$ has dense pure point spectrum as long as $\Gamma > 1$. If $\Gamma \leq 1$, then for a.e. pairs $(\lambda, \omega) \in \mathbf{T} \times \Omega$ the measure $\mu_{\lambda,\omega}$ has purely singular continuous spectrum of exact Hausdorff dimension $1 - \Gamma^2$.

23 Miscellanea

Exponential decay estimates

In [67] Nevai and Totik proved that $\limsup_{n\to\infty} |\alpha_n(\mu)|^{1/n} = R^{-1} < 1$ if and only if μ is absolutely continuous and the reciprocal of the Szegő function (16.3) D^{-1} is analytic in the disk $\{|z| < R\}$ providing a formula for the exact rate of exponential decay in terms of properties of D^{-1}. Moreover, the following result is true (see [84, Theorem 2.1]).

Theorem 23.1 *Let* $\limsup_{n\to\infty} |\alpha_n(\mu)|^{1/n} = R^{-1} < 1$ *and define*

$$S(z) := -\sum_{j=0}^{\infty} \alpha_{j-1}(\mu) z^j, \qquad r(z) := \overline{D(1/\bar{z})} D^{-1}(z), \tag{23.1}$$

so both S *and* D^{-1} *are analytic in* $\{|z| < R\}$*. Then the difference* $r - S$ *is analytic in* $\{z : 1 - \delta < |z| < R^3\}$ *for some* $\delta > 0$*.*

The point of this theorem is that both S and D^{-1} have singularities on the circle of radius R so the fact that the combination has analytic continuation is subtle.

To study the more specific exponential decay of Verblunsky coefficients Simon came up with the following
Definition. A sequence $\{A_n\}_{n\geq -1}$ of complex numbers is said to have an *asymptotic series with error* R^{-n} for some $R > 1$ if there exists a finite number of points $\{\mu_j\}_{j=1}^J$, $1 < |\mu_j| < R$, and polynomials $\{P_j\}_{j=1}^J$ so that

$$\limsup_{n\to\infty} \Big| A_n - \sum_{j=1}^{J} \frac{P_j(n)}{\mu_j^{n+1}} \Big|^{1/n} \leq R^{-1}. \tag{23.2}$$

A_n has a *complete asymptotic series* if the left hand side in (23.2) is zero.

It is not hard to see that $\{A_n\}_{n\geq -1}$ has an asymptotic series with error R^{-n} if and only if

$$L(z) := \sum_{n=0}^{\infty} A_{n-1} z^n$$

is meromorphic in $\{|z| < R\}$ with a finite number of poles, all in $\{1 < |z| < R\}$. In particular, $\{A_n\}_{n\geq -1}$ has a complete asymptotic series if and only if L is an entire meromorphic function, that is, meromorphic on the whole $\mathbf{C}$.

Here is the main result of [84].

Theorem 23.2 *Let a nontrivial probability measure* $\mu \in \mathcal{P}$ *with Verblunsky coefficients* $\alpha_n(\mu)$*. Then* α_n *has a complete asymptotic series if and only if* μ *is an absolutely continuous measure from the Szegő class and* D^{-1} *is an entire meromorphic function. The latter is equivalent to the function* S (23.1) *being an entire meromorphic function.*

Simon also provides the relation between μ_j's that enter in the asymptotic series and the poles of D^{-1}.

Rakhmanov's theorem on an arc

Let $\mu = wdm + \mu_s$ be a measure from $\mathcal{P}$. It was Rakhmanov [76] who proved that $w > 0$ a.e. on $\mathbf{T}$ implies $\alpha_n(\mu) \to 0$ as $n \to \infty$. Later on Bello and López [11] took this result over to an arc of the unit circle.

Theorem 23.3 (Bello – López). *Let Δ_α be an arc on $\mathbf{T}$*

$$\Delta_\alpha := \{\zeta = e^{it} : \alpha \le t \le 2\pi - \alpha\}, \qquad 0 < \alpha < \pi.$$

Suppose that $\operatorname{supp}\mu = \Delta_\alpha$ *and $w > 0$ a.e. on this arc. Then*

$$\lim_{n\to\infty} |\alpha_n(\mu)| = \sin\frac{\alpha}{2}, \qquad \lim_{n\to\infty} \overline{\alpha_{n+1}(\mu)}\alpha_n(\mu) = \sin^2\frac{\alpha}{2}.$$

On the other hand the recent result of Denisov [20] states

Theorem 23.4 (Denisov). *Let $J = J(\{a_n\},\{b_n\})$ be a Jacobi matrix with the spectral measure $\sigma = vdm + \sigma_s$. Suppose that* $\operatorname{ess\,supp}\sigma = [-2,2]$ *and $v > 0$ a.e. on $[-2,2]$. Then $a_n \to 1$, $b_n \to 0$ as $n \to \infty$.*

The original Rakhmanov theorem (more precisely, its real line analog) corresponds to the case $\operatorname{supp}\sigma = [-2,2]$. There is also an alternate proof of Denisov's theorem due to Nevai – Totik [68], which appeared to be a starting point for Simon to prove

Theorem 23.5 *The result of Bello – López holds under the relaxed assumption* $\operatorname{ess\,supp}\mu = \Delta_\alpha$.

Measures with VC's from ℓ^p, $p > 2$

For the class $\mathcal{P}$ of nontrivial probability measures on $\mathbf{T}$ Totik [107] established the following

Theorem 23.6 *Let $\mu \in \mathcal{P}$ with* $\operatorname{supp}\mu = \mathbf{T}$. *Then there exists a measure ν mutually absolutely continuous with μ and such that $\alpha_n(\nu) \to 0$ as $n \to \infty$.*

Simon proved the stronger (see [79, Theorem 2.10.1])

Theorem 23.7 *Let $\mu \in \mathcal{P}$ with* $\operatorname{supp}\mu = \mathbf{T}$. *Then there exists a measure ν mutually absolutely continuous with μ and such that for all $p > 2$*

$$\sum_{n=0}^{\infty} |\alpha_n(\nu)|^p < \infty.$$

Khrushchev [44] gave examples of singular continuous measure μ on $\mathbf{T}$ with $\{\alpha_n(\mu)\} \in \ell^p$ for all $p > 2$.

Counting eigenvalues in gaps

Suppose that $\mu^{(0)}$ and μ are measures from $\mathcal{P}$ with Verblunsky coefficients $\{\alpha_n^{(0)}\}$ and $\{\alpha_n\}$, respectively, and let $|\alpha_n - \alpha_n^{(0)}| \to 0$ with some information on the rate. If an open circular arc Γ is disjoint from $\sigma_{ess}(\mu^{(0)})$ then the same happens for $\sigma_{ess}(\mu)$ by Theorem 20.2. If the number of pure points of μ in Γ is infinite, they can only have the endpoints of Γ as limit points, and one can ask about the growth of the number of pure points near the endpoints of Γ.

The analogous problem for Schrödinger operators and Jacobi matrices is heavily studied. Given a Jacobi matrix $J(\{a_n\}_{n\ge1}, \{b_n\}_{n\ge1})$ with $a_n \to 1$ and $b_n \to 0$ as $n \to \infty$, denote by $N(J)$ a number of eigenvalues of J off $[-2, 2]$.

Theorem 23.8 (Geronimo – Case, Chihara – Nevai). *Suppose that*

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

Then $N(J) < \infty$.

Geronimo and later Hundertmark and Simon found a quantitative bound (see [38, Theorem A1])

$$N(J) \le \sum_{n=1}^{\infty} (n|b_n| + (4n+2)(a_n - 1)_+), \qquad (x)_+ = \max(x, 0), \tag{23.3}$$

the result known as the *Barmann-type bound.* Next, denote by $\{E_n^{\pm}\}$ the eigenvalues of J with

$$E_1^+ > E_2^+ > \ldots > 2 > -2 > \ldots > E_2^- > E_1^-.$$

Then (see [38, Theorem 2])

$$\sum_{n=1}^{\infty} \left(|E_n^+ - 2|^{p-1/2}| + |E_n^- + 2|^{p-1/2}| \right) \le C_p \left(\sum_{n=1}^{\infty} |b_n|^p + 4 \sum_{n=1}^{\infty} |a_n - 1|^p \right) \tag{23.4}$$

for all $p \ge 1$, where C_p is an explicitly given constant depending only on p. The estimate is usually called *Lieb – Thirring-type bound.*

In [80, Section 12.2] Simon has found similar bounds in the OPUC setting.

Theorem 23.9 *Let $\eta_{j+N} = \eta_j$, $j \in \mathbf{Z}_+$, be a periodic sequence of complex numbers, and suppose the Verblunsky coefficients $\alpha_n(\mu)$ of a measure $\mu \in \mathcal{P}$ satisfy*

$$\sum_{j=0}^{\infty} j|\alpha_j(\mu) - \eta_j| < \infty.$$

Then μ has an essential support whose complement has at most N gaps, and each gap has only finitely many mass points.

Theorem 23.10 *Suppose α's and η's are as in the above Theorem, and*

$$\sum_{j=0}^{\infty} |\alpha_j(\mu) - \eta_j|^p < \infty$$

holds for some $p \ge 1$. Then for the mass points z_j in the gaps we have

$$\sum_{z_j} \operatorname{dist}(z_j, \text{ess supp } \mu)^q < \infty,$$

where $q > 1/2$ if $p = 1$ and $q \ge p - 1/2$ if $p > 1$.

Remark. In a recent paper [36] L. Golinskii has found precise quantitative analogs of (23.3) and (23.4) in the OPUC setting for the case of a constant background ($N = 1$).

Jitomirskaya–Last inequalities

In a fundamental paper intended to understand and extend the subordinacy theory of Gilbert – Pearson, Jitomirskaya and Last proved some basic inequalities about singularities of the m-function as the energy approaches the spectrum. In [80, Section 10.8] an analog of their results for OPUC is established.

For a sequence of complex numbers $\{a_j\}_{j\ge 0}$ and a positive $0 < x < \infty$ define

$$\|a\|_x^2 := \sum_{j=0}^{[x]} |a_j|^2 + (x - [x])|a_{[x]+1}|^2,$$

so $\|a\|_n^2 = \sum_{j=0}^n |a_j|^2$ and $\|a\|_x^2$ is linearly interpolated in between. Let $\varphi_n(\zeta,\mu)$ and $\psi_n(\zeta,\mu)$ be the orthonormal polynomials of the 1st and 2nd kind, respectively, with respect to μ. Since either $\{\varphi_n\}$ or $\{\psi_n\}$ is not in ℓ^2 for $\zeta \in \mathbf{T}$, the product $\|\varphi\|_x \|\psi\|_x$ increases monotonically from 1 to $+\infty$, $0 < x < \infty$. Hence for $\zeta \in \mathbf{T}$ and $0 \le r < 1$ there is a unique solution $x(r)$ of

$$\|\varphi(\zeta)\|_{x(r)} \|\psi(\zeta)\|_{x(r)} = \frac{\sqrt{2}}{1-r}.$$

The unit circle analogue of Jitomirskaya – Last inequalities takes the following form.

Theorem 23.11 *Let $F = F(z,\mu)$ be the Carathéodory function of μ. Then*

$$A^{-1}\left[\frac{\|\psi(\zeta)\|_{x(r)}}{\|\varphi(\zeta)\|_{x(r)}}\right] \le |F(r\zeta,\mu)| \le A\left[\frac{\|\psi(\zeta)\|_{x(r)}}{\|\varphi(\zeta)\|_{x(r)}}\right],$$

where $1 < A < \infty$ is a universal constant.

Corollary 23.12 *Let $G \subset \mathbf{T}$ be a Borel set on the unit circle so that for all $\zeta \in G$ the transfer matrix T_n defined in* (14.11) *satisfies*

$$\sup_n \|T_n(\zeta)\| < \infty.$$

Then μ is absolutely continuous on G.

The OPUC analogs of the Gilbert – Pearson subordinacy theory and the Simon – Wolf theory appeared first in [37] and were elaborated in [80, Section 10].

References

[1] W.A. Al-Salam, Characterization theorems for orthogonal polynomials, in: Orthogonal Polynomials: Theory and Practice, M. Ismail and P. Nevai (eds), NATO ASI Series C, **294**, Cluwer, Dordrecht, 1990.

[2] M.P. Alfaro and L. Vigil, Solution of a problem of P. Turán, *J. Approx. Theory* **53**(1988), 195–197.

[3] M.P. Alfaro, M. Bello, J.M. Montaner and J.L. Varona, Some asymptotic properties for orthogonal polynomials with respect to varying measures, preprint.

[4] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, Encyclopedia of Mathematics and its Applciations, Cambridge University Press, New York, 1999.

[5] A. I. Aptekarev, Multiple othogonal polynomials, *J. Comput. Appl. Math.*, **99**(1988), 423–447.

[6] A. I. Aptekarev, Asymptotic properties of polynomials orthogonal on a system of contours, and periodic motions of Toda chains. (Russian) *Mat. Sb. (N.S.)*, **125(167)**(1984), 231–258.

[7] A. I. Aptekarev and W. Van Assche, Scalar and matrix Riemann-Hilbert approach to the strong asymptotics of Padé approximants and complex orgthogonal polynomials with varying weigts.

[8] L. Baratchart, R. Küstner and V. Totik, Zero distribution via orthogonality, *Annales de l'Institut Fourier* (to appear)

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

[10] G. Baxter, A convergence equivalence related to polynomials orthogonal on the unit circle, *Trans. Amer. Math. Soc.* **99**(1961), 471–487.

[11] M. Bello and G. López, Ratio and relative asymptotics of polynomials orthogonal on an arc of the unit circle, *J. Approx. Theory*, **92**(1998), 216–244.

[12] O. Blumenthal, *Über die Entwicklung einer willkürlichen Funktion nach den Nennern des Kettenbruches für* $\int_{-\infty}^{0}[\phi(\xi)/(z-\xi)]\,d\xi$*,* Inaugural Dissertation, Göttingen, 1898.

[13] S. Bochner, Über Sturm-Liouvillesche Polynomsysteme, *Math. Zeitschrift*, **29**(1929), 730–736.

[14] G. Borg, Eine Umkehrung der Sturm-Liouvilleschen Eigenwertaufgabe, *Acta Math.*(1946), 1–96.

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

[16] E. Daems and A.B.J. Kuijlaars, A Christoffel-Darboux formula for multiple orthogonal polynomials, *J. Approx. Theory*, **130**(2004), 190–202.

[17] D. Damanik, Verblunsky coefficients with Coulomb-type decay, to appear in JAT.

[18] P. Deift, this volume

[19] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, Courant Lecture Notes, 1999.

[20] S. A. Denisov, On Rakhmanov's theorem for Jacobi matrices, *Proc. Amer. Math. Soc.* **132**(2004), 847–852.

[21] S. A. Denisov and B. Simon, Zeros of orthogonal polynomials on the real line, *J. Approx. Theory*, **121**(2003), 357–364.

[22] F. Delyon, B. Simon, and B. Souillard, From power pure point to continuous spectrum in disordered systems, *Annales Inst. Henri Poincaré* **42** (1985), 283–309.

[23] C.F. Dunkl and Y. Xu, *Orthogonal polynomials of several variables*, Encyclopedia of Mathematics and its Applciations, Cambridge University Press, New York, 2001.

[24] A. J. Durán and F. A. Grünbaum, A survey on orthogonal matrix polynomials satisfying second order differential equations, *J. Comput. Appl. Math.,*

[25] H. Flaschka and D.W. McLaughlin, Canonically conjugate variables for the Korteweg-de Vries equation and the Toda lattice with periodic boundary conditions, *Progr. Theoret. Phys.*, **55**(1976), 438–456.

[26] A.S. Fokas, A.R. Its and A.V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, *Comm. Math. Phys.*, **147**(1992), 395–430.

[27] W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2004.

[28] W. Gautchi and A. B. J. Kuijlaars, Zeros and critical points of Sobolev orthogonal polynomials, *J. Approx. Theory*, **91**(1997), 117–137.

[29] J. S. Geronimo and W. Van Assche, Orthogonal polynomials on several intervals via a polynomial mapping, *Trans. Amer. Math. Soc.*, **308**(1988), 559–581.

[30] J. Geronimo, D. S. Lubinsky and F. Marcellan, Asymptotics for Sobolev orthogonal polynomials for exponential weights, (manuscript)

[31] J.S. Geronimo and R. Johnson, An inverse problem associated with polynomials orthogonal on the unit circle, *Comm. Math. Phys.*, **193**(1998), 125–150.

[32] J.S. Geronimo, F. Gesztesy, and H. Holden, Algebro-geometric solutionof the Baxter-Szegő difference equation, preprint.

[33] Ya.L. Geronimus, On the character of the solution of the moment problem in the case of the periodic in the limit associated fraction, *Bull. Acad. Sci. USSR Math.*, **5**(1941), 203–210. [Russian]

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

[35] Ya.L. Geronimus, On the trigonometric moment problem, *Ann. of Math.*, **47**(1946), 742–761.

[36] L. Golinskii, Lieb-Thirring and Bargman-type inequalities for circular arc, to appear in JAT.

[37] L. Golinskii and P. Nevai, Szegő difference equations, transfer matrices and orthogonal polynomials on the unit circle, *Comm. Math. Phys.*, **223**(2001), 223–259.

[38] D. Hundertmark and B. Simon, Lieb-Thirring inequalities for Jacobi matrices, *J. Approx. Theory*, **118**(2002), 106–130.

[39] M.E.H. Ismail, Classical and quantum orthogonal polynomials in one variable, Encyclopedia in Mathematics, Cambridge University Press, 2005.

[40] V. Kaliaguine, A note on the asymptotics of orthogonal polynomials on a complex arc: the case of a measure with a discrete part., *J. Approx. Theory*, **80**(1995), 138–145.

[41] A. Kiselev, Y. Last and B. Simon, Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators, *Comm. Math. Phys.*, **194**(1998), 1–45.

[42] S. Khrushchev, Schur's algorithm, orthogonal polynomials, and convergence of Wall's fractions in $L^2(\mathbf{T})$, *J. Approx. Theory*, **108**(2001), 161–248.

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

[44] S. Khrushchev, A singular Riesz product in the Nevai class and inner functions with the Schur parameters in $\cap_{p>2}\ell^p$, *J. Approx. Theory*, **108**(2001), 249–255.

[45] S. Khrushchev, Continued fractions and orthogonal polynomials on the unit circle, Proc. Conf. OPSFA, Copenhagen, 2003.

[46] T. Kriecherbauer and K. T-R. McLaughlin, Strong asymptotics of polynomials orthogonal with respect to a Freud weight, *Internat. Math. Research Notices*, **6**(1999), 299–333.

[47] Y. Last and B. Simon, The essential spectrum of Schrödinger, Jacobi, and CMV operators, preprint.

[48] A. Levin and D. S. Lubinsky, *Orthogonal polynomials for exponential weights*, CMS Books in Mathematics/Ouvrages de Mathmatiques de la SMC, **4**, Springer-Verlag, New York, 2001

[49] A. Levin and D. S. Lubinsky, *Christoffel Functions and Orthogonal Polynomials for Exponential Weights on* $[-1,1]$, Memoirs Amer. Math. Soc., **535**, (1994)

[50] G. L. Lagomasino, H. Pijeira-Cabrera, I. P. Izquierdo, Sobolev orthogonal polynomials in the complex plane, *J. Comput. Appl. Math.,* **127**(2001), 219-230.

[51] P. López-Rodriguez and A. J. Durán, Orthogonal matrix polynomials, Laredo Lecture Notes (Lecture notes of the 1st SIAG Summer School on Orthogonal Polynomials and Special Functions, Laredo, Spain, July 24-29 2000). Nova Science Publisher (2003). R. Álvarez-Nodarse, F. Marcellán and W. Van Assche (Eds.), 2004

[52] D. S. Lubinsky, *Strong Asymptotics for Extremal Errors Associated with Erdös-type Weights*, Pitman Res. Notes, Vol. **202**, Longmans, Nurn Mill, Essex, 1989.

[53] D. S. Lubinsky and E. B. Saff, *Strong Asymptotics for Extremal Polynomials Associated with Weights on* $(-\infty, \infty)$, Springer Lecture Notes, Vol. **1305**, Springer Verlag, Berlin, 1988.

[54] W. Magnus and S. Winkler, *Hill's Equation*, Interscience Tracts in Pure and Applied Mathematics, No. 20, Interscience Publishers, NY, 1966.

[55] G. Mastroianni and V. Totik, Polynomial inequalities with doubling and A_∞ weights, *Constr. Approx.*, **16** (2000), 37–71.

[56] A. Máté, P. Nevai and V. Totik, Extensions of Szegő's theory of orthogonal polynomials, *Constructive Approximation*, **3**(1987), 51–72.

[57] A. Máté, P. Nevai and V. Totik, Strong and weak convergence of orthogonal polynomials, *Amer. J. Math.*, **109**(1987), 239–282.

[58] A. Máté, P. Nevai and V. Totik, Szegő's extremum problem on the unit circle, *Annals of Math.*, **134**(1991), 433–453.

[59] F. Marcellán, M. Alfaro and M. L. Rezola, Orthogonal plynomials on Sobolev spaces: old and new directions, *J. Comput. Appl. Math.,* **48**(1993), 113–131.

[60] A. Martinez-Finkelshtein, Analytic aspects of Sobolev orthogonal polynomials, *J. Comput. Appl. Math.,* **99**(1998), 491–510.

[61] A. Martinez-Finkelshtein, Analytic properties of Sobolev orthogonal polynomials revisited, *J. Comput. Appl. Math.,* **127**(2001), 255-266.

[62] G. Mastroianni and G. Milovanovic, *Approximation of Functions and Functionals with Application*, World Scientific (to appear)

d

[63] M. Mehta, *Random matrices*, second ed., Academic Press, Inc., Boston, 1991.

[64] H.N. Mhaskar and E.B. Saff, On the distribution of zeros of polynomials orthogonal on the unit circle, *J. Approx. Theory*, **63**(1990), 30–38.

[65] S.N. Naboko, On the dense point spectrum of Schrödinger and Dirac operators, *Theoret. and Math. Phys.*, **68**(1986), 646–653.

[66] P. Nevai, *Orthogonal Polynomials,* Memoirs Amer. Math. Soc. **213** (1979)

[67] P. Nevai and V. Totik, Orthogonal polynomials and their zeros, *Acta Sci. Math.*(Szeged), **53**(1989), 99–104.

[68] P. Nevai and V. Totik, Denisov's theorem on recurrent coefficients, to appear in JAT.

[69] E.M. Nikishin, An estimate for orthogonal polynomials, *Acta Sci. Math.*(Szeged), **48**(1985), 395–399. [Russian]

[70] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*, Grundlehren de mathematischen Wissenschaften, Vol. **297**, Springer Verlag, Berlin, 1992.

[71] F. Peherstorfer and R. Steinbauer, On polynomials orthogonal on several intervals, Special functions (Torino, 1993), *Ann. Numer. Math.*, **2**(1995), 353–370.

[72] F. Peherstorfer, Zeros of polynomials orthogonal on several intervals. *Int. Math. Res. Notes*, **7**(2003), 361–385.

[73] F. Peherstorfer, On Bernstein-Szegő orthogonal polynomials on several intervals, *SIAM J. Math. Anal.*, **21**(1990), 461–482.

[74] F. Peherstorfer, On Bernstein-Szegő orthogonal polynomials on several intervals. II. Orthogonal polynomials with periodic recurrence coefficients, *J. Approx. Theory*, **64**(1991), 123–161.

[75] F. Peherstorfer, Orthogonal and extremal polynomials on several intervals, *J. Comput. Appl. Math.*, **48**(1993), 187–205.

[76] E. A. Rakhmanov, On the asymptotics of the ratio of orthogonal polynomials, *Math. USSR Sbornik,* **46** (1983), 105–117; Russian original, *Mat. Sb.*, **118** (1982), 104–117.

[77] E. A. Rahmanov, On V. A. Steklov's problem in the theory of orthogonal polynomials, (Russian) *Dokl. Akad. Nauk SSSR*, **254**(1980), 802–806.

[78] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt send, I, II, *J. Reine Angew. Math.*, **147**(1917), 205–232; **148**(1918), 122–145.

[79] B. Simon, *Orthogonal Polynomials on the Unit Circle*, V.1: Classical Theory, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[80] B. Simon, *Orthogonal Polynomials on the Unit Circle*, V.2: Spectral Theory, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[81] B. Simon, The classical moment problem as a self-adjoint finite difference operator, *Advances in Math.*, **137**(1998), 82–203.

[82] B. Simon, Fine structure of the zeros of orthogonal polynomials, I. A tale of two pictures, Proc. Constructive Functions Tech-04.

[83] B. Simon, Fine structure of the zeros of orthogonal polynomials, II. OPUC with competing exponential decay, to appear in JAT.

[84] B. Simon, Meromorphic Szegő functions and asymptotic series for Verblunsky coefficients, preprint.

[85] B. Simon, Ratio asymptotics and weak asymptotic measures for orthogonal polynomials on the real line, *J. Approx. Theory*, **126**(2004), 198–217.

[86] B. Simon, Some Jacobi matrices with decaying potential and dense point spectrum, *Comm. Math. Phys.*, **87**(1982), 253–258.

[87] B. Simon, Fine structure of the zeros of orthogonal polynomials, III. Periodic recursion coefficients, to appear in Comm. Pure Appl. Math.

[88] B. Simon, OPUC on one foot, preprint.

[89] B. Simon, Aizenman's theorem for orthogonal polynomials on the unit circle, to appear in Const. Approx.

[90] B. Simon, Orthogonal polynomials on the unit circle: New results, *Int. Math. Res. Not.*, **53**(2004), 2837–2880.

[91] B. Simon, The sharp form of the strong Szegő theorem, to appear in J. Conf. on Geometry and Spectral Theory.

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

[93] B. Simon, Analogs of the m-functions in the theory of orthogonal polynomials on the unit circle, *J. Comp. Appl. Math.*, **171**(2004), 411–424.

[94] B. Simon, The Golinskii-Ibragimov method and a theorem of Damanik-Killip, *Int. Math. Res. Not.*, **36**(2003), 1973–1986.

[95] B. Simon and V. Totik, Limits of zeros of orthogonal polynomials on the circle, to appear in Math. Nachr.

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

[97] B. Simon and A. Zlatos, Higher-order Szegő theorems with two singular points, to appear in JAT.

[98] H. Stahl and V. Totik, *General Orthogonal Polynomials*, Encyclopedia of Mathematics and its Applications, **43**, Cambridge University Press, New York, 1992

[99] H. Stahl, Orthogonal polynomials with respect to complex-valued measures. Orthogonal polynomials and their applications (Erice, 1990), *IMACS Ann. Comput. Appl. Math.*, **9**(1991), 139–154.

[100] H. Stahl, Orthogonal polynomials with complex-valued weight function. I, II. *Constr. Approx.*, **2**(1986), 225–240, 241–251.

[101] H. Stahl, Divergence of diagonal Padé approximants and the asymptotic behavior of orthogonal polynomials associated with nonpositive measures, *Constr. Approx.* **1**(1985), 249–270.

[102] M. Stoiciu, The statistical distribution of the zeros of random paraorthogonal polynomials on the unit circle, to appear in JAT.

[103] J. Szabados and P. Vértesi, *Interpolation of Functions*, World Scientific, Singapore, 1990.

[104] G. Szegő, *Orthogonal Polynomials*, Coll. Publ., **XXIII**, Amer. Math. Soc., Providence, 1975.

[105] A.V. Teplyaev, Properties of polynomials that are orthogonal on the circle with random parameters, *J. Soviet Math.*, **61**(1992), 1931–1935.

[106] M. Toda, *Theory of Nonlinear Lattices*, second edition, Springer Series in Solid-State Sciences, **20**, Springer-Verlag, Berlin, 1989.

[107] V. Totik, Orthogonal polynomials with ratio asymptotics, *Proc. Amer. Math. Soc.*, **114** (1992), 491–495.

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

[109] W. Van Assche, *Asymptotics for Orthogonal Polynomials*, Spriger Lecture Notes, Vol. **1265**, Springer Verlag, Berlin, 1987.

[110] W. Van Assche, Hermite-Padé approximation and multiple orthogonal polynomials (manuscript)

[111] W. Van Assche, J.S. Geronimo, A.B.J. Kuijlaars, Riemann-Hilbert problems for multiple orthogonal polynomials, in: J. Bustoz et al. (eds), Special Functions 2000: Current Perspectives and Future Directions, Kluwer, Dordrecht, 2001, 23–59.

[112] S. Verblunsky, On positive harmonic functions: Acontribution to the algebra of Fourier series, *Proc. London Math. Soc.*, **38**(1935), 125–157.

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

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

Leonid Golinskii
Verkin Institute for Low Temperature Physics,
Kharkov
Lenin ave. 47, 61103
Ukraine

Vilmos Totik
Bolyai Institute, University of Szeged
Szeged
Aradi v. tere 1, 6720
Hungary, and

Department of Mathematics
University of South Florida
Tampa, FL, 33620
USA