Spectral Theory of Orthogonal Polynomials

Barry Simon
IBM Professor of Mathematics and Theoretical Physics
California Institute of Technology
Pasadena, CA, U.S.A.

Lecture 1: Introduction and Overview

- Lecture 1: Introduction and Overview
- Lecture 2: Szegö Theorem for OPUC
- Lecture 3: Three Kinds of Polynomials Asymptotics, I
- Lecture 4: Three Kinds of Polynomial Asymptotics, II

References

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

[OPUC2] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Series, **54.2**, American Mathematical Society, Providence, RI, 2005.

[SzThm] B. Simon, *Szegő's Theorem and Its Descendants: Spectral Theory for L^2 Perturbations of Orthogonal Polynomials*, M. B. Porter Lectures, Princeton University Press, Princeton, NJ, 2011.

What is spectral theory?

Spectral theory is the general theory of the relation of the fundamental parameters of an object and its "spectral" characteristics.

Spectral characteristics means eigenvalues or scattering data or, more generally, spectral measures

What is spectral theory

Examples include

- Can you hear the shape of a drum ?
- Computer tomography
- Isospectral manifold for the harmonic oscillator

What is spectral theory

The *direct problem* goes from the object to spectra.

The *inverse problem* goes backwards.

The direct problem is typically easy while the inverse problem is typically hard.

For example, the domain of definition of the harmonic oscillator isospectral "manifold" is unknown. It is not even known if it is connected!

OPs

Orthogonal polynomials on the real line (OPRL) and on the unit circle (OPUC) are particularly useful because the inverse problems are easy—indeed the inverse problem is the OP definition as we'll see.

OPs also enter in many application—both specific polynomials and the general theory.

OPs

Indeed, my own interest came from studying discrete Schrödinger operators on $\ell^2(\mathbb{Z})$

$$\big(Hu\big)_n = u_{n+1} + u_{n-1} + V u_n$$

and the realization that when restricted to $\mathbb{Z}_+$, one had a special case of OPRL.

OPRL basics

μ will be a probability measure on $\mathbb{R}$. We'll always suppose that μ has bounded support $[a,b]$ which is not a finite set of points. (We then say that μ is non-trivial.) This implies that $1, x, x^2, \dots$ are independent since $\int |P(x)|^2 \, d\mu = 0 \Rightarrow \mu$ is supported on the zeroes of P.

Apply Gram Schmidt to $1, x, \dots$ and get monic polynomials

$$P_j(x) = x^j + \alpha_{j,1} x^{j-1} + \dots$$

and orthonormal (ON) polynomials

$$p_j = P_j / \|P_j\|$$

OPRL basics

More generally we can do the same for any probability measure of bounded support on $\mathbb{C}$.

One difference from the case of $\mathbb{R}$, the linear combination of $\{x^j\}_{j=0}^\infty$ are dense in $L^2(\mathbb{R}, d\mu)$ by Weierstrass. This may or may not be true if $\mathrm{supp}(d\mu) \not\subset \mathbb{R}$.

If $d\mu = d\theta/2\pi$ on $\partial\mathbb{D}$, the span of $\{z^j\}_{j=0}^\infty$ is not dense in L^2 (but is only H^2). Perhaps, surprisingly, we'll see later that there are measures $d\mu$ on $\partial\mathbb{D}$ for which they are dense (e.g., μ purely singular).

More significantly, the argument we'll give for our recursion relation fails if $\mathrm{supp}(d\mu) \not\subset \mathbb{R}$.

OPRL basics

Since P_k is monic and $\{P_j\}_{j=0}^{k+1}$ span polynomials of degree at most $k+1$, we have

$$xP_k = P_{k+1} + \sum_{j=0}^{k} B_{k,j}\, P_j$$

Clearly

$$B_{k,j} = \langle P_j, xP_k \rangle / \|P_j\|^2$$

Now we use

$$\langle P_j, xP_k \rangle = \langle xP_j, P_k \rangle$$

(need $d\mu$ on $\mathbb{R}$!!)

If $j < k-1$, this is zero.

If $j = k-1$, $\langle P_{k-1}, xP_k \rangle = \langle xP_{k-1}, P_k \rangle = \|P_k\|^2$.

OPRL basics

Thus ($P_{-1} \equiv 0$); $\{a_j\}_{j=1}^{\infty}$, $\{b_j\}_{j=1}^{\infty}$: Jacobi recursion

$$xP_N = P_{N+1} + b_{N+1}P_N + a_N^2 P_{N-1}$$

$$b_N \in \mathbb{R}, \quad a_N = \|P_N\|/\|P_{N-1}\|$$

These are called Jacobi parameters. This implies $\|P_N\| = a_N\, a_{N-1} \dots a_1$ (since $\|P_0\| = 1$).

This, in turn, implies $p_n = P_n/a_1 \dots a_n$ obeys

$$xp_n = a_{n+1}p_{n+1} + b_{n+1}p_n + a_n p_{n-1}$$

OPRL basics

We have thus solved the inverse problem, i.e., μ is the spectral data and $\{a_n, b_n\}_{n=1}^\infty$ are the descriptors of the object.

In the orthonormal basis $\{p_n\}_{n=0}^\infty$, multiplication by x has the matrix

$$J = \begin{pmatrix} b_1 & a_1 & 0 & 0 & \dots \\ a_1 & b_2 & a_2 & 0 & \dots \\ 0 & a_2 & b_3 & a_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

called a Jacobi matrix.

Favard's Theorem

Since

$$b_n = \int x p_{n-1}^2(x)\, d\mu, \quad a_n = \int x p_{n-1}(x) p_n(x)\, d\mu$$

$\operatorname{supp}(\mu) \subset [-R, R] \Rightarrow |b_n| \le R,\ |a_n| \le R.$

Conversely, if $\sup_n \big(|a_n| + |b_n|\big) = \alpha < \infty$, J is a bounded matrix of norm at most 3α. In that case, the spectral theorem implies there is a measure $d\mu$ so that

$$\langle (1, 0, \ldots)^t, J^\ell (1, 0, \ldots)^t \rangle = \int x^\ell d\mu(x)$$

If one uses Gram-Schmidt to orthonormalize $\{J^\ell (1, 0, \ldots)^t\}_{\ell=0}^\infty$, one finds μ has Jacobi matrix exactly given by J.

Favard's Theorem

We have thus proven Favard's Theorem (his paper was in 1935; really due to Stieltjes in 1894 or to Stone in 1932).

Favard's Theorem. *There is a one–one correspondence between bounded Jacobi parameters*

$$\{a_n, b_n\}_{n=1}^{\infty} \in \big[(0,\infty) \times \mathbb{R}\big]^{\infty}$$

and non-trivial probability measures, μ, of bounded support via:

$$\mu \Rightarrow \{a_n, b_n\} \quad \textit{(OP recursion)}$$

$$\{a_n, b_n\} \Rightarrow \mu \quad \textit{(Spectral Theorem)}$$

There are also results for μ's with unbounded support so long as $\int x^n \, d\mu < \infty$. In this case, $\{a_n, b_n\} \Rightarrow \mu$ may not be unique because J may not be essentially self-adjoint on vectors of finite support.

OPUC basics

Let $d\mu$ be a non-trivial probability measure on $\partial\mathbb{D}$. As in the OPRL case, we use Gram-Schmidt to define monic OPs, $\Phi_n(z)$ and ON OP's $\varphi_n(z)$.

In the OPRL case, if z is multiplication by the underlying variable, $z^* = z$. This is replaced by $z^* z = 1$.

In the OPRL case, $P_{n+1} - xP_n \perp \{1, x_1, \dots, x_{n-2}\}$.

OPUC basics

In the OPUC case, $\Phi_{n+1} - z\Phi_n \perp \{z, \dots, z^n\}$, since

$$\langle z\Phi, z^j \rangle = \langle \Phi, z^{j-1} \rangle$$

if $j \geq 1$.

In the OPRL case, we used deg $P = n$ and $P \perp \{1, x, \dots, x^{n-2}\} \Rightarrow P = c_1 P_n + c_2 P_{n-1}$.

In the OPUC case, we want to characterize deg $P = n$, $P \perp \{z, z^2, \dots, z^n\}$.

OPUC basics

Define * on degree n polynomials to themselves by

$$Q^*(z) = z^n \, \overline{Q\left(\frac{1}{\bar{z}}\right)}$$

(bad but standard notation!) or

$$Q(z) = \sum_{j=0}^{n} c_j z^j \Rightarrow Q^*(z) = \sum_{j=0}^{n} \bar{c}_{n-j} \, z^j$$

Then, * is unitary and so for deg $Q = n$

$$Q \perp \{1, \dots, z^{n-1}\} \Leftrightarrow Q = c \, \Phi_n$$

is equivalent to

$$Q \perp \{z, \dots, z^n\} \Leftrightarrow Q = c \, \Phi_n^*$$

Szegő recursion and Verblunsky coefficients

Thus, we see, there are parameters $\{\alpha_n\}_{n=0}^{\infty}$ (called Verblunsky coefficients) so that

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

This is the Szegő Recursion (History: Szegő and Geronimus in 1939; Verblunsky in 1935–36)

Applying * for deg $n+1$ polynomials to this yields

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

The strange looking $-\bar{\alpha}_n$ rather than say $+\alpha_n$ is to have the α_n be the Schur parameter of the Schur function of μ (Geronimus); also the Verblunsky parameterization then agrees with α_n. These are discussed in [OPUC1].

Szegő recursion and Verblunsky coefficients

Φ_n monic $\Rightarrow$ constant term in Φ_n^* is 1 $\Rightarrow \Phi_n^*(0) = 1$.

This plus $\Phi_{n+1} = z\Phi_n - \bar{\alpha}_n \Phi_n^*(z)$ implies

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

i.e., Φ_n determines α_{n-1}.

Szegő recursion and Verblunsky coefficients

For OPRL, we saw $\|P_{n+1}\|/\|P_n\| = a_{n+1}$. We are looking for the analog for OPUC.

Szegő Recursion $\Rightarrow \Phi_{n+1} + \bar{\alpha}_n \Phi_n^* = z\Phi_n$

$$\Phi_{n+1} \perp \Phi_n^* \Rightarrow \|\Phi_{n+1}\|^2 + |\alpha_n|^2 \|\Phi_n^*\|^2 = \|z\Phi_n\|^2$$

Multiplication by z unitary plus * antiunitary $\Rightarrow$

$$\|\Phi_{n+1}\|^2 = \rho_n^2 \|\Phi_n\|^2; \quad \rho_n^2 = 1 - |\alpha_n|^2$$

which implies $|\alpha_n| < 1$ (i.e., $\alpha_n \in \mathbb{D}$) and

$$\|\Phi_n\| = \rho_{n-1} \cdots \rho_0$$

Szegő recursion and Verblunsky coefficients

$$\begin{pmatrix} \varphi_{n+1} \\ \varphi_{n+1}^* \end{pmatrix} = A_n(z) \begin{pmatrix} \varphi_n \\ \varphi_n^* \end{pmatrix} x; \quad A_n = \rho_n^{-1} \begin{pmatrix} z & -\bar{\alpha}_n \\ -\alpha_n \, z & 1 \end{pmatrix}$$

$\det A_n \neq 0$ if $z \neq 0$, so we can get φ_n (Φ_n) from φ_{n+1} (Φ_{n+1}) by

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

$$\Phi_n^* = \rho_n^{-2} \left[\Phi_{n+1} + \alpha_n \Phi_{n+1} \right]$$

Szegő recursion and Verblunsky coefficients

We see that Φ_{n+1} determines α_n, so by induction and inverse recursion,

Theorem. *If two measures have the same Φ_n, they have the same $\{\Phi_j\}_{j=0}^{n-1}$ and $\{\alpha_j\}_{j=0}^{n-1}$.*

Szegő recursion and Verblunsky coefficients

A similar argument to the one that led to $|\alpha_n| < 1$ yields

Theorem. *All zeros of* Φ_n *lie in* $\mathbb{D}$*.*

Proof. $\Phi_n(z_0) = 0 \Rightarrow \Phi_n = (z - z_0)p$, $\deg p = n - 1$

$zp = \Phi_n + z_0 p$ and $p \perp \Phi_n \Rightarrow \|p\|^2 = \|\Phi_n\|^2 + |z_0|^2\|p\|^2$

$\Rightarrow |z_0| < 1$

Corollary. *All zeros of* $\Phi_n^*(z)$ *lie in* $\mathbb{C} \setminus \overline{\mathbb{D}}$*.*

Szegő recursion and Verblunsky coefficients

Here is a second proof that only uses Szegő recursion. By induction, suppose that all zeros of Φ_n are in $\mathbb{D}$. Then, for $|\beta| < 1$

$$z\Phi_n + \beta\Phi_n^* \neq 0 \text{ on } \partial\mathbb{D}$$

since $|z\Phi_n(z)| = |\Phi_n^*(z)|$ on $\partial\mathbb{D}$. $\left(\frac{1}{\bar{z}} = z\right)$

If $\Phi_{n+1}^{(\beta)} = z\Phi_n + \beta\Phi_n^*$, then at $\beta = 0$, all zeros of $\Phi_{n+1}^{(\beta)}$ are in $\mathbb{D}$.

As β varies in $\mathbb{D}$, all zeros of $\Phi_{n+1}^{(\beta)}$ are trapped in $\mathbb{D}$. QED.

Bernstein–Szegő Approximation

We are heading towards a proof that any $\{\alpha_n\}_{n=0}^{\infty} \subset \mathbb{D}$ are the Verblunsky coefficients of a measure on $\partial\mathbb{D}$ (analog of Favard's Theorem). It will depend on

Theorem (Bernstein–Szegő measures). *Let* $\{\alpha_j^{(0)}\}_{j=0}^{n-1} \in \mathbb{D}^n$. *Let* $\varphi_n(z)$ *be the normalized degree* n *polynomial obtained by Szegő recursion. Let*

$$d\mu_n(\theta) = \frac{d\theta}{2\pi|\varphi_n(e^{i\theta})|^2}$$

Then $d\mu_n$ *has Verblunsky coefficients*

$$\alpha_j(d\mu_n) = \begin{cases} \alpha_j^{(0)} & j = 0, \dots, n-1 \\ 0 & j \geq n \end{cases}$$

Bernstein–Szegő Approximation

The first step of the proof is to show that

$$k, \ell, n \text{ with } k < n + \ell \Rightarrow \int_{z=e^{i\theta}} \bar{z}^k \, z^\ell \, \varphi_n(z) d\mu_n(\theta) = 0$$

For $z \in \partial\mathbb{D} \Rightarrow \overline{\varphi_n(z)} = \overline{\varphi_n\left(\frac{1}{\bar{z}}\right)} = z^{-n}\varphi_n^*(z)$.
Thus the integral above is

$$\oint \frac{\bar{z}^k \, z^\ell \, \varphi_n(z)}{z^{-n} \, \varphi_n(z) \, \varphi_n^*(z)} \frac{dz}{2\pi i z} = \frac{1}{2\pi} \oint z^{\ell+n-k-1} \frac{dz}{\varphi_n^*(z)}$$

is zero since $\left[\varphi_n^*(z)\right]^{-1}$ is analytic on a neighborhood of $\overline{\mathbb{D}}$ and $\ell + n - k - 1 \geq 0$.

Bernstein–Szegő Approximation

Thus, $z^\ell \varphi_n$ is a multiple of the OP's for $d\mu_n$.

Since $\int |z^\ell \varphi_n|^2 \, d\mu = 1$, we see that

$$\varphi_{n+k}(z; d\mu) = z^k \varphi_n(z); \; k > 0.$$

As we saw, Φ_n determines $\{\alpha_j\}_{j=0}^{n-1}$ and Φ_j by inverse Szegő recursion and $-\overline{\Phi_{j+1}(0)} = \alpha_j$. This shows that

$$\varphi_j(z; d\mu) = \begin{cases} \varphi_j(x) & j = 0, \dots, n \\ z^{j-n} \varphi_n(z) & j = n, n+1, \dots \end{cases}$$

implying the claimed result.

Bernstein–Szegő Approximation

Given $\{\alpha_j\}_{j=0}^{\infty} \subset \mathbb{D}^{\infty}$, we can form $d\mu_n$ as above. Via $\int \Phi_j(e^{i\theta})d\mu(e^{i\theta}) = 0$, $\{\Phi_j\}_{j=0}^{n}$ determines $\{\int z^j d\mu\}_{j=0}^{n}$ inductively (actually they determine more moments). Thus

$$\int z^j d\mu_n = \int z^j d\mu_m \quad j \leq \min(n, m)$$

and $\int \overline{z^j}\, d\mu_n = \overline{\left(\int z^j\, d\mu_n\right)}$.

Thus, $d\mu_n$ has a weak limit $d\mu_\infty$. Clearly, $\alpha_j(d\mu_\infty) = \alpha_j$.

We have thus proven

Verblunsky's Theorem. $\mu \to \{\alpha_j(\mu)\}_{j=0}^{\infty}$ sets up a 1–1 correspondence between non-trivial probability measures on $\partial\mathbb{D}$ and $\mathbb{D}^{\infty}$.

Carmona Simon Formula

Simon [CRM Proc. and Lecture Notes **42** (2007), 453–463] has proven an analog of the Bernstein–Szegő approximation for OPRL (the analog for Schrödinger operators is due to Carmona; hence the name):

Let $d\rho$ be a probability measure on $\mathbb{R}$ with $\int |x|^n \, d\rho < \infty$ for all n. Let $\{p_n\}_{n=0}^\infty$ be its orthonormal polynomials and $\{a_n, b_n\}_{n=1}^\infty$ its Jacobi parameters. Let

$$d\nu_n(x) = dx / \left[\pi(a_n^2 p_n^2(x) + p_{n-1}^2(x))\right]$$

Then, for $\ell = 0, \ldots, 2n-2$, $\int x^\ell \, d\nu_n = \int x^\ell \, d\rho$.

If the moment problem for $d\rho$ is determinate, then $d\nu_n \to d\rho$ weakly.

Carmona Simon Formula

One important consequence of this result is

Theorem. If $I \subset \mathbb{R}$ is an interval and for all $x \in I$ and some $c > 0$, we have that

$$c \leq a_n^2 p_n^2(x) + p_{n-1}^2(x) \leq c^{-1}$$

then $d\rho \restriction I$ has a.c. part and no singular spectrum.

Similarly, for $I \subset \partial\mathbb{D}$ and μ a probability measure

$$c \leq |\varphi_n(z)| \leq c^{-1} \quad \text{all } z \in I$$

implies $d\mu \restriction I$ has a.c. part and no singular spectrum.

Remark. A much stronger result is known (see e.g., Simon [Proc AMS **124** (1996), 3361]); I can be any set and c can be x-dependent.