Large Deviations and Sum Rules for Orthogonal Polynomials

CLAPEM XIV

Universidad de Costa Rica, December, 2016

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

Lecture 1: OPRL, OPUC and Sum Rules

- Lecture 1: OPRL, OPUC and Sum Rules
- Lecture 2: Meromorphic Herglotz Functions and Proof of KS Sum Rule
- Lecture 3: The Theory of Large Deviations
- Lecture 4: GNR Proof of Sum Rules

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.

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?

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.

Examples include

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.

Examples include

- Can you hear the shape of a drum ?

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.

Examples include

- Can you hear the shape of a drum ?
- Computer tomography

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.

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.

What is spectral theory?

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

The *inverse problem* goes backwards.

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.

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

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

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.

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

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

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 $1, x, x^2, \ldots$ are independent since $\int |P(x)|^2 \, d\mu = 0 \Rightarrow \mu$ is supported on the zeroes of P.

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 $\mathbb{C}$.

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 $\mathrm{supp}(d\mu) \not\subset \mathbb{R}$.

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 $\operatorname{supp}(d\mu) \not\subset \mathbb{R}$.

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, there are measures $d\mu$ on $\partial\mathbb{D}$ for which they are dense (e.g., μ purely singular).

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, 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

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

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 $\mathbb{R}$!!)

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.

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*

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

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*.

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 $\|P_N\| = a_N\, a_{N-1} \ldots a_1$ (since $\|P_0\| = 1$).

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.

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

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

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

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

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

Favard's Theorem

Since

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

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

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

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

If one uses Gram-Schmidt to orthonormalize $\{J^\ell (1, 0, \dots)^t\}_{\ell=0}^\infty$, 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

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:

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

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

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.

m–function

By the above construction, one has that

m–function

By the above construction, one has that

$$m(z) \equiv \int \frac{d\mu(x)}{x - z} = \langle \delta_1, (J - z)^{-1} \delta_1 \rangle$$

m–function

By the above construction, one has that

$$m(z) \equiv \int \frac{d\mu(x)}{x-z} = \langle \delta_1, (J-z)^{-1}\delta_1 \rangle$$

Let J_N be the $N \times N$ matrix obtained by keeping the top N rows and leftmost N columns of J. Then it is easy to prove that

m–function

By the above construction, one has that

$$m(z) \equiv \int \frac{d\mu(x)}{x - z} = \langle \delta_1, (J - z)^{-1} \delta_1 \rangle$$

Let J_N be the $N \times N$ matrix obtained by keeping the top N rows and leftmost N columns of J. Then it is easy to prove that

$$m(z) = \lim_{N \to \infty} \langle \delta_1, (J_N - z)^{-1} \delta_1 \rangle$$

m–function

By the above construction, one has that

$$m(z) \equiv \int \frac{d\mu(x)}{x-z} = \langle \delta_1, (J-z)^{-1}\delta_1 \rangle$$

Let J_N be the $N \times N$ matrix obtained by keeping the top N rows and leftmost N columns of J. Then it is easy to prove that

$$m(z) = \lim_{N\to\infty} \langle \delta_1, (J_N - z)^{-1}\delta_1 \rangle$$

We denote the quantity inside the limit as $m_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N)$.

Coefficient stripping

Let $D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N) = \det(J_N - z)$ so that Cramér's rule says that

Coefficient stripping

Let $D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N) = \det(J_N - z)$ so that Cramér's rule says that

$$m_N(z; a_1, \dots, b_N) = \frac{D_{N-1}(z; a_2, \dots, a_{N-1}; b_2, \dots, b_N)}{D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N)}$$

Coefficient stripping

Let $D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N) = \det(J_N - z)$ so that Cramér's rule says that

$$m_N(z; a_1, \dots, b_N) = \frac{D_{N-1}(z; a_2, \dots, a_{N-1}; b_2, \dots, b_N)}{D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N)}$$

where we look at the determinant of a once stripped matrix obtained by removing the first row and column.

Coefficient stripping

Let $D_N(z; a_1, \ldots, a_{N-1}; b_1, \ldots, b_N) = \det(J_N - z)$ so that Cramér's rule says that

$$m_N(z; a_1, \ldots, b_N) = \frac{D_{N-1}(z; a_2, \ldots, a_{N-1}; b_2, \ldots, b_N)}{D_N(z; a_1, \ldots, a_{N-1}; b_1, \ldots, b_N)}$$

where we look at the determinant of a once stripped matrix obtained by removing the first row and column. By expanding $\det(J_N - z)$ in minors in the first column we get

Coefficient stripping

Let $D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N) = \det(J_N - z)$ so that Cramér's rule says that

$$m_N(z; a_1, \dots, b_N) = \frac{D_{N-1}(z; a_2, \dots, a_{N-1}; b_2, \dots, b_N)}{D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N)}$$

where we look at the determinant of a once stripped matrix obtained by removing the first row and column. By expanding $\det(J_N - z)$ in minors in the first column we get

$$\begin{aligned} D_N(z; a_1, \dots, b_N) = (b_1 - z) & D_{N-1}(z; a_2, \dots, b_N) \\ & - a_1^2 D_{N-2}(z; a_3, \dots, b_N) \end{aligned}$$

Coefficient stripping

Let $D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N) = \det(J_N - z)$ so that Cramér's rule says that

$$m_N(z; a_1, \dots, b_N) = \frac{D_{N-1}(z; a_2, \dots, a_{N-1}; b_2, \dots, b_N)}{D_N(z; a_1, \dots, a_{N-1}; b_1, \dots, b_N)}$$

where we look at the determinant of a once stripped matrix obtained by removing the first row and column. By expanding $\det(J_N - z)$ in minors in the first column we get

$$\begin{aligned} D_N(z; a_1, \dots, b_N) = (b_1 - z) D_{N-1}(z; a_2, \dots, b_N) \\ - a_1^2 D_{N-2}(z; a_3, \dots, b_N) \end{aligned}$$

Dividing by $D_{N-1}(z; a_2, \dots, b_N)$ and taking $N \to \infty$ after using the formula for m as a ratio of D's, we see that

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix.

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m to a_1, b_1 and m_1

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m to a_1, b_1 and m_1 (and so inductively all Jacobi parameters) by looking at Taylor coefficients of $m(z)^{-1}$ near infinity.

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m to a_1, b_1 and m_1 (and so inductively all Jacobi parameters) by looking at Taylor coefficients of $m(z)^{-1}$ near infinity. By iterating, one gets a continued fraction expansion:

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m to a_1, b_1 and m_1 (and so inductively all Jacobi parameters) by looking at Taylor coefficients of $m(z)^{-1}$ near infinity. By iterating, one gets a continued fraction expansion:

$$m(z) = \cfrac{1}{b_1 - z + \cfrac{a_1^2}{b_2 - z + \cfrac{a_2^2}{b_3 - z + \cdots}}}$$

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m to a_1, b_1 and m_1 (and so inductively all Jacobi parameters) by looking at Taylor coefficients of $m(z)^{-1}$ near infinity. By iterating, one gets a continued fraction expansion:

$$m(z) = \cfrac{1}{b_1 - z + \cfrac{a_1^2}{b_2 - z + \cfrac{a_2^2}{b_3 - z + \cdots}}}$$

The convergence theorems for continued fractions lets one go from Jacobi parameters to measure.

Coefficient stripping m–function formula

$$m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$$

where m_1 is the m–function of the once stripped infinite Jacobi matrix. This provides another was to recover Jacobi paramters from a measure: go from the measure to m to a_1, b_1 and m_1 (and so inductively all Jacobi parameters) by looking at Taylor coefficients of $m(z)^{-1}$ near infinity. By iterating, one gets a continued fraction expansion:

$$m(z) = \cfrac{1}{b_1 - z + \cfrac{a_1^2}{b_2 - z + \cfrac{a_2^2}{b_3 - z + \cdots}}}$$

The convergence theorems for continued fractions lets one go from Jacobi parameters to measure. One consequence of the single stripping formula is that the poles of m_1 (i.e. the pure points of μ_1) are precisely the zeros of m.

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 $\varphi_n(z)$.

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 $z^* z = 1$.

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, $P_{n+1} - xP_n \perp \{1, x, \ldots, x^{n-2}\}$.

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 $j \geq 1$.

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

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

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 antiunitary and so for deg $Q = n$

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

is equivalent to

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

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)

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

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

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 $\Rightarrow \Phi_n^*(0) = 1$.

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

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

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

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 $\Rightarrow$

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

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

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

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

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

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

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

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

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

This implies that Φ_n determines $\{\alpha_j\}_{j=0}^{n-1}$ since we've seen it determines α_{n-1},

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

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

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

This implies that Φ_n determines $\{\alpha_j\}_{j=0}^{n-1}$ since we've seen it determines α_{n-1}, then by the above inverse Szegő recursion, Φ_{n-1},

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

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

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

This implies that Φ_n determines $\{\alpha_j\}_{j=0}^{n-1}$ since we've seen it determines α_{n-1}, then by the above inverse Szegő recursion, Φ_{n-1}, and then inductively all the smaller α's.

Verblunsky's Theorem

There is a one-one correspondence, called the *Verblunsky map*, from measures of infinite support and sequences in $\mathbb{D}$.

Verblunsky's Theorem

There is a one-one correspondence, called the *Verblunsky map*, from measures of infinite support and sequences in $\mathbb{D}$. We've seen how to go from measures to Verblunsky coefficients.

Verblunsky's Theorem

There is a one-one correspondence, called the *Verblunsky map*, from measures of infinite support and sequences in $\mathbb{D}$. We've seen how to go from measures to Verblunsky coefficients. One way of going in the opposite direction is to prove that for any $\{\beta_j\}_{j=0}^{n-1}$ in $\mathbb{D}^n$, if the β's are used to form OPs up to order n,

Verblunsky's Theorem

There is a one-one correspondence, called the *Verblunsky map*, from measures of infinite support and sequences in $\mathbb{D}$. We've seen how to go from measures to Verblunsky coefficients. One way of going in the opposite direction is to prove that for any $\{\beta_j\}_{j=0}^{n-1}$ in $\mathbb{D}^n$, if the β's are used to form OPs up to order n, then the measure $d\theta/2\pi|\varphi_n(e^{i\theta})|^2$ is a measure with Verblunsky coefficients

$\alpha_j = \beta_j; \quad j = 0, \dots, n-1$ and $\alpha_j = 0; \quad j \geq n$.

Verblunsky's Theorem

There is a one-one correspondence, called the *Verblunsky map*, from measures of infinite support and sequences in $\mathbb{D}$. We've seen how to go from measures to Verblunsky coefficients. One way of going in the opposite direction is to prove that for any $\{\beta_j\}_{j=0}^{n-1}$ in $\mathbb{D}^n$, if the β's are used to form OPs up to order n, then the measure $d\theta/2\pi|\varphi_n(e^{i\theta})|^2$ is a measure with Verblunsky coefficients $\alpha_j = \beta_j; \quad j = 0, \dots, n-1$ and $\alpha_j = 0; \quad j \geq n$. Given an infinite sequence in $\mathbb{D}^\infty$, one proves these measures converge to a measure with the given Verblunsky coefficients.

Verblunsky's Theorem

For measures with exactly n pure points, there are only n non–trivial OPs, and n α's. $\alpha_{n-1} \in \partial\mathbb{D}$. One has $\|\Phi_k\| = \rho_0 \dots \rho_{k-1}$ where $\rho_j = \sqrt{1 - |\alpha_j|^2}$

Verblunsky's Theorem

For measures with exactly n pure points, there are only n non–trivial OPs, and n α's. $\alpha_{n-1} \in \partial\mathbb{D}$. One has $\|\Phi_k\| = \rho_0 \dots \rho_{k-1}$ where $\rho_j = \sqrt{1 - |\alpha_j|^2}$ which explains why in the n–point case where $\|\Phi_n\| = 0$

Verblunsky's Theorem

For measures with exactly n pure points, there are only n non–trivial OPs, and n α's. $\alpha_{n-1} \in \partial\mathbb{D}$. One has $\|\Phi_k\| = \rho_0 \dots \rho_{k-1}$ where $\rho_j = \sqrt{1 - |\alpha_j|^2}$ which explains why in the n–point case where $\|\Phi_n\| = 0$ we have $|\alpha_{n-1}| = 1$.

Verblunsky's Theorem

For measures with exactly n pure points, there are only n non–trivial OPs, and n α's. $\alpha_{n-1} \in \partial\mathbb{D}$. One has $\|\Phi_k\| = \rho_0 \dots \rho_{k-1}$ where $\rho_j = \sqrt{1 - |\alpha_j|^2}$ which explains why in the n–point case where $\|\Phi_n\| = 0$ we have $|\alpha_{n-1}| = 1$. For this set of n point measures, the set of measures and the set of Verblunsky coefficients is $2n - 1$ (real) dimensional.

Szegő's Theorem: Toeplitz version

Szegő's Theorem concerns probability measures on $\partial\mathbb{D}$ of the form

$$d\mu(\theta) = w(\theta)\frac{d\theta}{2\pi} + d\mu_s(\theta)$$

where $d\mu_s$ is singular w.r.t. $d\theta$.

Szegő's Theorem: Toeplitz version

Szegő's Theorem concerns probability measures on $\partial\mathbb{D}$ of the form

$$d\mu(\theta) = w(\theta)\frac{d\theta}{2\pi} + d\mu_s(\theta)$$

where $d\mu_s$ is singular w.r.t. $d\theta$. The Toeplitz determinant $D_n(d\mu)$ is the n x n determinant with

$$c_{k\ell} \equiv \int e^{i(k-\ell)\theta} d\mu(\theta) = \langle e^{-ik\cdot}, e^{-i\ell\cdot}\rangle_{L^2(d\mu)}$$

Szegő's Theorem: Toeplitz version

Szegő's Theorem concerns probability measures on $\partial\mathbb{D}$ of the form

$$d\mu(\theta) = w(\theta)\frac{d\theta}{2\pi} + d\mu_s(\theta)$$

where $d\mu_s$ is singular w.r.t. $d\theta$. The Toeplitz determinant $D_n(d\mu)$ is the n x n determinant with

$$c_{k\ell} \equiv \int e^{i(k-\ell)\theta} d\mu(\theta) = \langle e^{-ik\cdot}, e^{-i\ell\cdot}\rangle_{L^2(d\mu)}$$

In 1915, Szegő proved that

$$\lim_{n\to\infty} D_n(d\mu)^{1/n} = \exp\left[\int \log(w(\theta))\tfrac{d\theta}{2\pi}\right]$$

Szegő's Theorem: Toeplitz version

Szegő's Theorem concerns probability measures on $\partial\mathbb{D}$ of the form

$$d\mu(\theta) = w(\theta)\frac{d\theta}{2\pi} + d\mu_s(\theta)$$

where $d\mu_s$ is singular w.r.t. $d\theta$. The Toeplitz determinant $D_n(d\mu)$ is the n x n determinant with

$$c_{k\ell} \equiv \int e^{i(k-\ell)\theta} d\mu(\theta) = \langle e^{-ik\cdot}, e^{-i\ell\cdot}\rangle_{L^2(d\mu)}$$

In 1915, Szegő proved that

$$\lim_{n\to\infty} D_n(d\mu)^{1/n} = \exp\left[\int \log(w(\theta))\,\tfrac{d\theta}{2\pi}\right]$$

While this is true in general, Szegő only proved it when $d\mu_s = 0$.

Szegő's Theorem: OPUC version

In 1920, Szegő realized that, because a Toeplitz matrix is just the Gram matrix of $\{z^j\}_{j=0}^{n-1}$, it is also the Gram matrix of $\{\Phi_j\}_{j=0}^{n-1}$ which is diagonal so

$$D_n = \prod_{j=0}^{n-1} \|\Phi_j\|^2$$

Szegő's Theorem: OPUC version

In 1920, Szegő realized that, because a Toeplitz matrix is just the Gram matrix of $\{z^j\}_{j=0}^{n-1}$, it is also the Gram matrix of $\{\Phi_j\}_{j=0}^{n-1}$ which is diagonal so

$$D_n = \prod_{j=0}^{n-1} \|\Phi_j\|^2$$

so using that $\|\Phi_j\|$ is monotone decreasing (by a variational argument), one has an equivalent form of his theorem, namely

$$\lim_{n\to\infty} \|\Phi_n\|^2 = \exp\left[\int \log(w(\theta) \tfrac{d\theta}{2\pi}\right]$$

Szegő's Theorem: OPUC version

In 1920, Szegő realized that, because a Toeplitz matrix is just the Gram matrix of $\{z^j\}_{j=0}^{n-1}$, it is also the Gram matrix of $\{\Phi_j\}_{j=0}^{n-1}$ which is diagonal so

$$D_n = \prod_{j=0}^{n-1} \|\Phi_j\|^2$$

so using that $\|\Phi_j\|$ is monotone decreasing (by a variational argument), one has an equivalent form of his theorem, namely

$$\lim_{n\to\infty} \|\Phi_n\|^2 = \exp\left[\int \log(w(\theta)\frac{d\theta}{2\pi}\right]$$

But the recursion relation was only published by Szegő in 1939, so he didn't have a form in term's of α_n and ρ_n.

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood)

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood) first of all extended Szegő's theorem to allow a singular part,

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood) first of all extended Szegő's theorem to allow a singular part, introduced the α_n in a different form than as recursion coefficients

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood) first of all extended Szegő's theorem to allow a singular part, introduced the α_n in a different form than as recursion coefficients and wrote Szegő's theorem as a sum rule

$$\sum_{j=0}^{\infty} \log(1 - |\alpha_j|^2) = \int \log(w(\theta)) \frac{d\theta}{2\pi}$$

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood) first of all extended Szegő's theorem to allow a singular part, introduced the α_n in a different form than as recursion coefficients and wrote Szegő's theorem as a sum rule

$$\sum_{j=0}^{\infty} \log(1 - |\alpha_j|^2) = \int \log(w(\theta)) \frac{d\theta}{2\pi}$$

It is critical that this always holds although both sides may be $-\infty$.

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood) first of all extended Szegő's theorem to allow a singular part, introduced the α_n in a different form than as recursion coefficients and wrote Szegő's theorem as a sum rule

$$\sum_{j=0}^{\infty} \log(1-|\alpha_j|^2) = \int \log(w(\theta))\frac{d\theta}{2\pi}$$

It is critical that this always holds although both sides may be $-\infty$. This implies what I've called a "spectral theory gem"

$$\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \iff \int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$$

Szegő's Theorem: Szegő-Verblunsky sum rule

In two remarkable 1935-36 papers, long unappreciated, Samuel Verblunsky (then just past his PhD. under Littlewood) first of all extended Szegő's theorem to allow a singular part, introduced the α_n in a different form than as recursion coefficients and wrote Szegő's theorem as a sum rule

$$\sum_{j=0}^{\infty} \log(1 - |\alpha_j|^2) = \int \log(w(\theta)) \frac{d\theta}{2\pi}$$

It is critical that this always holds although both sides may be $-\infty$. This implies what I've called a "spectral theory gem"

$$\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \iff \int \log(w(\theta)) \frac{d\theta}{2\pi} > -\infty$$

In particular, $\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \Rightarrow \Sigma_{ac} = \partial\mathbb{D}$.

Szegő's Theorem: Szegő-Verblunsky sum rule

What makes the gems so interesting is that they allow arbitrary singular parts of the measures so long as the Szegő condition holds, i.e. $\int \log(w(\theta)) \frac{d\theta}{2\pi} > -\infty$.

Szegő's Theorem: Szegő-Verblunsky sum rule

What makes the gems so interesting is that they allow arbitrary singular parts of the measures so long as the Szegő condition holds, i.e. $\int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$. If $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, one can show that there is a scattering theory and strong asymptotic completeness holds in that there is only a.c. spectrum.

Szegő's Theorem: Szegő-Verblunsky sum rule

What makes the gems so interesting is that they allow arbitrary singular parts of the measures so long as the Szegő condition holds, i.e. $\int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$. If $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, one can show that there is a scattering theory and strong asymptotic completeness holds in that there is only a.c. spectrum. The VS sum rules implies in going from ℓ^1 to ℓ^2 Verblunsky coefficients, one can have arbitrary mixed spectral types.

Szegő's Theorem: Szegő-Verblunsky sum rule

What makes the gems so interesting is that they allow arbitrary singular parts of the measures so long as the Szegő condition holds, i.e. $\int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$. If $\sum_{j=0}^{\infty}|\alpha_j| < \infty$, one can show that there is a scattering theory and strong asymptotic completeness holds in that there is only a.c. spectrum. The VS sum rules implies in going from ℓ^1 to ℓ^2 Verblunsky coefficients, one can have arbitrary mixed spectral types.

In the late 1990's unaware of the OPUC literature, my research group was studying $1D$ Schrodinger operators, $-\frac{d^2}{dx^2} + V(x)$ and the difference between L^1 and L^2 conditions.

Szegő's Theorem: Szegő-Verblunsky sum rule

What makes the gems so interesting is that they allow arbitrary singular parts of the measures so long as the Szegő condition holds, i.e. $\int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$. If $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, one can show that there is a scattering theory and strong asymptotic completeness holds in that there is only a.c. spectrum. The VS sum rules implies in going from ℓ^1 to ℓ^2 Verblunsky coefficients, one can have arbitrary mixed spectral types.

In the late 1990's unaware of the OPUC literature, my research group was studying $1D$ Schrodinger operators, $-\frac{d^2}{dx^2} + V(x)$ and the difference between L^1 and L^2 conditions. Deift–Killip had proven there was a.c. spectrum for L^2 and showing there were examples with mixed spectrum was one of the problems in my list at the 2000 ICMP.

Szegő's Theorem: Szegő-Verblunsky sum rule

What makes the gems so interesting is that they allow arbitrary singular parts of the measures so long as the Szegő condition holds, i.e. $\int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$. If $\sum_{j=0}^{\infty}|\alpha_j| < \infty$, one can show that there is a scattering theory and strong asymptotic completeness holds in that there is only a.c. spectrum. The VS sum rules implies in going from ℓ^1 to ℓ^2 Verblunsky coefficients, one can have arbitrary mixed spectral types.

In the late 1990's unaware of the OPUC literature, my research group was studying $1D$ Schrodinger operators, $-\frac{d^2}{dx^2} + V(x)$ and the difference between L^1 and L^2 conditions. Deift–Killip had proven there was a.c. spectrum for L^2 and showing there were examples with mixed spectrum was one of the problems in my list at the 2000 ICMP. Little did I know that an analogous problem had been solved in 1935!

Szegő Condition

Here is one version of Szegő's Theorem for OPRL.

Szegő Condition

Here is one version of Szegő's Theorem for OPRL. The map $z \mapsto z + z^{-1}$ maps $\partial\mathbb{D}$ to $[-2, 2]$ (via $e^{i\theta} \mapsto 2\cos\theta$)

Szegő Condition

Here is one version of Szegő's Theorem for OPRL. The map $z \mapsto z + z^{-1}$ maps $\partial\mathbb{D}$ to $[-2, 2]$ (via $e^{i\theta} \mapsto 2\cos\theta$) and so measures, μ, on $[-2, 2]$ to measures, ρ, on $\partial\mathbb{D}$ which are symmetric under complex conjugation (since the above map is 2 to 1 except at ± 1).

Szegő Condition

Here is one version of Szegő's Theorem for OPRL. The map $z \mapsto z + z^{-1}$ maps $\partial\mathbb{D}$ to $[-2, 2]$ (via $e^{i\theta} \mapsto 2\cos\theta$) and so measures, μ, on $[-2, 2]$ to measures, ρ, on $\partial\mathbb{D}$ which are symmetric under complex conjugation (since the above map is 2 to 1 except at ± 1). In 1922, Szegő found a relation of the OPRL for μ to the OPUC for ρ and this allowed later authors to prove a version of Szegő's theorem for $d\mu = w(x)\,dx + d\mu_s$ (with $s(x) = (4 - x^2)^{-1/2}$):

Szegő Condition

Here is one version of Szegő's Theorem for OPRL. The map $z \mapsto z + z^{-1}$ maps $\partial\mathbb{D}$ to $[-2, 2]$ (via $e^{i\theta} \mapsto 2\cos\theta$) and so measures, μ, on $[-2, 2]$ to measures, ρ, on $\partial\mathbb{D}$ which are symmetric under complex conjugation (since the above map is 2 to 1 except at ± 1). In 1922, Szegő found a relation of the OPRL for μ to the OPUC for ρ and this allowed later authors to prove a version of Szegő's theorem for $d\mu = w(x)\,dx + d\mu_s$ (with $s(x) = (4 - x^2)^{-1/2}$):

$$\liminf_{n\to\infty} \prod_{j=1}^{n} a_j = \sqrt{2}\exp\left(\int_{-2}^{2} \log|\pi s(x) w(x)| s(x) \frac{dx}{4\pi}\right)$$

Szegő Condition

Here is one version of Szegő's Theorem for OPRL. The map $z \mapsto z + z^{-1}$ maps $\partial\mathbb{D}$ to $[-2, 2]$ (via $e^{i\theta} \mapsto 2\cos\theta$) and so measures, μ, on $[-2, 2]$ to measures, ρ, on $\partial\mathbb{D}$ which are symmetric under complex conjugation (since the above map is 2 to 1 except at ± 1). In 1922, Szegő found a relation of the OPRL for μ to the OPUC for ρ and this allowed later authors to prove a version of Szegő's theorem for $d\mu = w(x)\,dx + d\mu_s$ (with $s(x) = (4 - x^2)^{-1/2}$):

$$\liminf_{n\to\infty} \prod_{j=1}^{n} a_j = \sqrt{2}\exp\left(\int_{-2}^{2} \log|\pi s(x) w(x)| s(x) \frac{dx}{4\pi}\right)$$

The condition for the finiteness of the integral is called the *Szegő condition*:

Szegő Condition

Here is one version of Szegő's Theorem for OPRL. The map $z \mapsto z + z^{-1}$ maps $\partial\mathbb{D}$ to $[-2,2]$ (via $e^{i\theta} \mapsto 2\cos\theta$) and so measures, μ, on $[-2,2]$ to measures, ρ, on $\partial\mathbb{D}$ which are symmetric under complex conjugation (since the above map is 2 to 1 except at ± 1). In 1922, Szegő found a relation of the OPRL for μ to the OPUC for ρ and this allowed later authors to prove a version of Szegő's theorem for $d\mu = w(x)\,dx + d\mu_s$ (with $s(x) = (4-x^2)^{-1/2}$):

$$\liminf_{n\to\infty} \prod_{j=1}^{n} a_j = \sqrt{2}\exp\left(\int_{-2}^{2} \log|\pi s(x) w(x)|\, s(x) \frac{dx}{4\pi}\right)$$

The condition for the finiteness of the integral is called the *Szegő condition*:

$$\int_{-2}^{2} \log|w(x)|(4-x^2)^{-1/2}\,dx > -\infty$$

Szegő Condition

This doesn't yield a gem because

Szegő Condition

This doesn't yield a gem because

$$\inf_n \prod_{j=1}^{n} a_j > -\infty \iff \int_{-2}^{2} \log|w(x)|(4-x^2)^{-1/2}\,dx$$

Szegő Condition

This doesn't yield a gem because

$$\inf_n \prod_{j=1}^{n} a_j > -\infty \iff \int_{-2}^{2} \log |w(x)| (4 - x^2)^{-1/2} \, dx$$

only holds under the a priori condition that μ is supported inside $[-2, 2]$

Szegő Condition

This doesn't yield a gem because

$$\inf_n \prod_{j=1}^{n} a_j > -\infty \iff \int_{-2}^{2} \log|w(x)|(4-x^2)^{-1/2}\,dx$$

only holds under the a priori condition that μ is supported inside $[-2,2]$ and this is not simply expressible in terms of the Jacobi parameters;

Szegő Condition

This doesn't yield a gem because

$$\inf_n \prod_{j=1}^{n} a_j > -\infty \iff \int_{-2}^{2} \log|w(x)|(4-x^2)^{-1/2}\,dx$$

only holds under the a priori condition that μ is supported inside $[-2,2]$ and this is not simply expressible in terms of the Jacobi parameters; for example, it doesn't only depend on the parameters near ∞ and can be changed by modifying a single a or b.

Killip–Simon Theorem

In 2001 (published 2003), Killip and I proved the following gem which we regard as an OPRL analog of the Verblunsky–Szegő gem where $\{E_j^\pm\}_{j=1}^{N_\pm}$ are the eigenvalues outside $[-2,2]$ (with + above 2 and - below -2):

Killip–Simon Theorem

In 2001 (published 2003), Killip and I proved the following gem which we regard as an OPRL analog of the Verblunsky–Szegő gem where $\{E_j^{\pm}\}_{j=1}^{N_{\pm}}$ are the eigenvalues outside $[-2,2]$ (with + above 2 and - below -2):

Killip-Simon Theorem If $d\mu = w(x)dx + d\mu_s$ is a measure of compact support on $\mathbb{R}$ and $\{a_n, b_n\}_{n=1}^{\infty}$ its Jacobi parameters, then

$$\sum_{j=1}^{\infty} |a_j - 1|^2 + b_j^2 < \infty$$

Killip–Simon Theorem

In 2001 (published 2003), Killip and I proved the following gem which we regard as an OPRL analog of the Verblunsky–Szegő gem where $\{E_j^{\pm}\}_{j=1}^{N_\pm}$ are the eigenvalues outside $[-2,2]$ (with + above 2 and - below -2):

Killip-Simon Theorem If $d\mu = w(x)dx + d\mu_s$ is a measure of compact support on $\mathbb{R}$ and $\{a_n, b_n\}_{n=1}^{\infty}$ its Jacobi parameters, then

$$\sum_{j=1}^{\infty} |a_j - 1|^2 + b_j^2 < \infty$$

if and only if the essential support of μ is $[-2,2]$ and

$$\int_{-2}^{2} \log(w(x))\sqrt{4-x^2}\,dx > -\infty \qquad \sum_{j,\pm} (|E_j^{\pm}| - 2)^{3/2} < \infty$$

Killip–Simon Theorem

This result on Jacobi Hilbert-Schmidt perturbations of the free Jacobi matrix should be compared with a celebrated theorem of von-Neumann that any bounded self-adjoint operator has a Hilbert-Schmidt perturbation with only dense point spectrum!

Killip–Simon Theorem

This result on Jacobi Hilbert-Schmidt perturbations of the free Jacobi matrix should be compared with a celebrated theorem of von-Neumann that any bounded self-adjoint operator has a Hilbert-Schmidt perturbation with only dense point spectrum!

We called $\int_{-2}^{2} \log(w(x))\sqrt{4-x^2}\,dx > -\infty$ *the quasi-Szegő condition* since the square root appeared to the $+1/2$ power rather than the $-1/2$ in the Szegő condition.

Killip–Simon Theorem

This result on Jacobi Hilbert-Schmidt perturbations of the free Jacobi matrix should be compared with a celebrated theorem of von-Neumann that any bounded self-adjoint operator has a Hilbert-Schmidt perturbation with only dense point spectrum!

We called $\int_{-2}^{2} \log(w(x))\sqrt{4-x^2}\,dx > -\infty$ *the quasi-Szegő condition* since the square root appeared to the $+1/2$ power rather than the $-1/2$ in the Szegő condition. We called $\sum_{j,\pm}(|E_j^{\pm}|-2)^{3/2} < \infty$ a *Lieb-Thirring condition* since it is the discrete analog of the celebrated inequality from which Lieb and Thirring proved stability of matter,

Killip–Simon Theorem

This result on Jacobi Hilbert-Schmidt perturbations of the free Jacobi matrix should be compared with a celebrated theorem of von-Neumann that any bounded self-adjoint operator has a Hilbert-Schmidt perturbation with only dense point spectrum!

We called $\int_{-2}^{2} \log(w(x))\sqrt{4-x^2}\,dx > -\infty$ *the quasi-Szegő condition* since the square root appeared to the $+1/2$ power rather than the $-1/2$ in the Szegő condition. We called $\sum_{j,\pm}(|E_j^{\pm}|-2)^{3/2} < \infty$ a *Lieb-Thirring condition* since it is the discrete analog of the celebrated inequality from which Lieb and Thirring proved stability of matter, viz for $-\Delta + V(x)$ on $L^2(\mathbb{R}^d)$ (for $d=1$, $p=1$)

$$\sum |E_n|^p \le C \int |V(x)|^{p+d/2}\,dx$$

Killip–Simon Sum Rule

The gem comes from a sum rule. Let
$Q(\mu) = \frac{1}{2\pi} \int_0^{2\pi} \log\left(\frac{\sin(\theta)}{\operatorname{Im} m(2\cos(\theta))}\right) \sin^2(\theta) d\theta,$

Killip–Simon Sum Rule

The gem comes from a sum rule. Let
$Q(\mu) = \frac{1}{2\pi} \int_0^{2\pi} \log\left(\frac{\sin(\theta)}{\operatorname{Im} m(2\cos(\theta))}\right) \sin^2(\theta) d\theta$,
$G(a) = a^2 - 1 - \log(a^2)$ and

Killip–Simon Sum Rule

The gem comes from a sum rule. Let
$Q(\mu) = \frac{1}{2\pi} \int_0^{2\pi} \log\left(\frac{\sin(\theta)}{\operatorname{Im} m(2\cos(\theta))}\right) \sin^2(\theta) d\theta$,
$G(a) = a^2 - 1 - \log(a^2)$ and

$$F(E) \equiv \frac{1}{4}[\beta^2 - \beta^{-2} - \log(\beta^4)] \qquad E = \beta + \beta^{-1} \qquad |\beta| > 1$$

Killip–Simon Sum Rule

The gem comes from a sum rule. Let
$Q(\mu) = \frac{1}{2\pi} \int_0^{2\pi} \log\left(\frac{\sin(\theta)}{\operatorname{Im} m(2\cos(\theta))}\right) \sin^2(\theta) d\theta$,
$G(a) = a^2 - 1 - \log(a^2)$ and

$$F(E) \equiv \tfrac{1}{4}[\beta^2 - \beta^{-2} - \log(\beta^4)] \qquad E = \beta + \beta^{-1} \qquad |\beta| > 1$$

Then the Killip-Simon sum rule says

$$Q(\mu) + \sum_{j,\pm} F(E_j^{\pm}) = \sum_{n=1}^{\infty} \tfrac{1}{4} b_n^2 + \tfrac{1}{2} G(a_n)$$

As with the Szegő–Verblunsky sum rule, an important point is that it always holds although both sides may be $+\infty$.

Killip-Simon Sum Rule

The gem comes from the fact that $F \geq 0$, vanishes exactly at $E = \pm 2$ and is $\mathrm{O}((|E| - 2)^{3/2})$ there and that $G \geq 0$, vanishes exactly at $a = 1$ and is $\mathrm{O}((a-1)^2)$ there.

Killip-Simon Sum Rule

The gem comes from the fact that $F \geq 0$, vanishes exactly at $E = \pm 2$ and is $\mathrm{O}((|E| - 2)^{3/2})$ there and that $G \geq 0$, vanishes exactly at $a = 1$ and is $\mathrm{O}((a-1)^2)$ there.

The positivity of the terms is essential to be sure that there aren't cancelations.

Killip-Simon Sum Rule

The gem comes from the fact that $F \geq 0$, vanishes exactly at $E = \pm 2$ and is $\mathrm{O}((|E|-2)^{3/2})$ there and that $G \geq 0$, vanishes exactly at $a = 1$ and is $\mathrm{O}((a-1)^2)$ there.

The positivity of the terms is essential to be sure that there aren't cancelations. Case had an infinite number of sum rules that he stated (without indication of when they hold nor rigorous proof), coming from terms of a suitable Taylor series, but none was positive.

Killip-Simon Sum Rule

The gem comes from the fact that $F \geq 0$, vanishes exactly at $E = \pm 2$ and is $\mathrm{O}((|E|-2)^{3/2})$ there and that $G \geq 0$, vanishes exactly at $a = 1$ and is $\mathrm{O}((a-1)^2)$ there.

The positivity of the terms is essential to be sure that there aren't cancelations. Case had an infinite number of sum rules that he stated (without indication of when they hold nor rigorous proof), coming from terms of a suitable Taylor series, but none was positive. What Killip-Simon realized is that $C_0 + \frac{1}{2}C_2$ had only positive terms

Killip-Simon Sum Rule

The gem comes from the fact that $F \geq 0$, vanishes exactly at $E = \pm 2$ and is $\mathrm{O}((|E| - 2)^{3/2})$ there and that $G \geq 0$, vanishes exactly at $a = 1$ and is $\mathrm{O}((a-1)^2)$ there.

The positivity of the terms is essential to be sure that there aren't cancelations. Case had an infinite number of sum rules that he stated (without indication of when they hold nor rigorous proof), coming from terms of a suitable Taylor series, but none was positive. What Killip-Simon realized is that $C_0 + \frac{1}{2}C_2$ had only positive terms although it was mysterious why this sum is positive and so unclear how to generate positive sum rules.

Killip-Simon Sum Rule

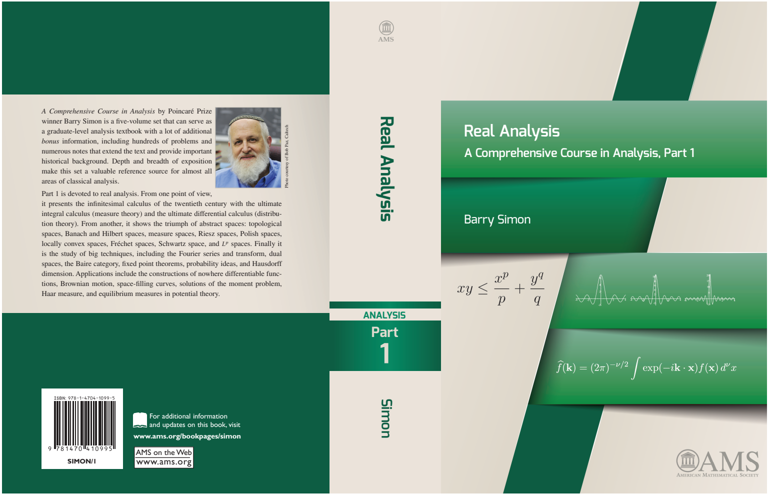
The gem comes from the fact that $F \geq 0$, vanishes exactly at $E = \pm 2$ and is $\mathrm{O}((|E| - 2)^{3/2})$ there and that $G \geq 0$, vanishes exactly at $a = 1$ and is $\mathrm{O}((a-1)^2)$ there.

The positivity of the terms is essential to be sure that there aren't cancelations. Case had an infinite number of sum rules that he stated (without indication of when they hold nor rigorous proof), coming from terms of a suitable Taylor series, but none was positive. What Killip-Simon realized is that $C_0 + \frac{1}{2}C_2$ had only positive terms although it was mysterious why this sum is positive and so unclear how to generate positive sum rules.

As in the OPUC case, this sum rule implies the existence of Hilbert–Schmidt perturbations with mixed spectrum.

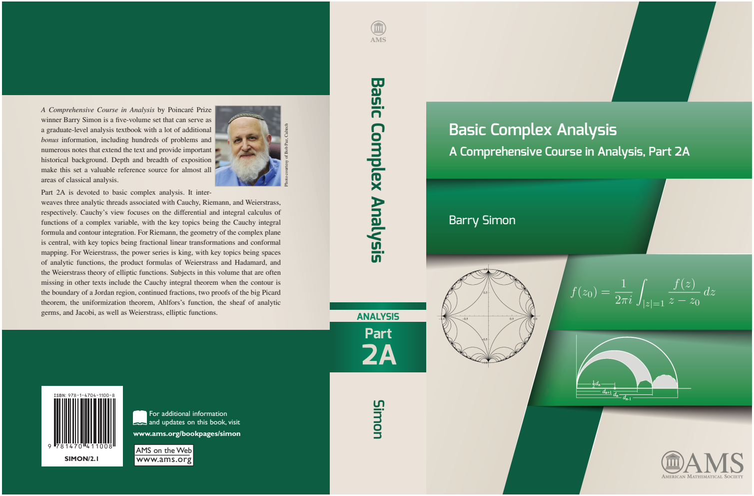
And Now a Word from Our Sponsor

And Now a Word from Our Sponsor



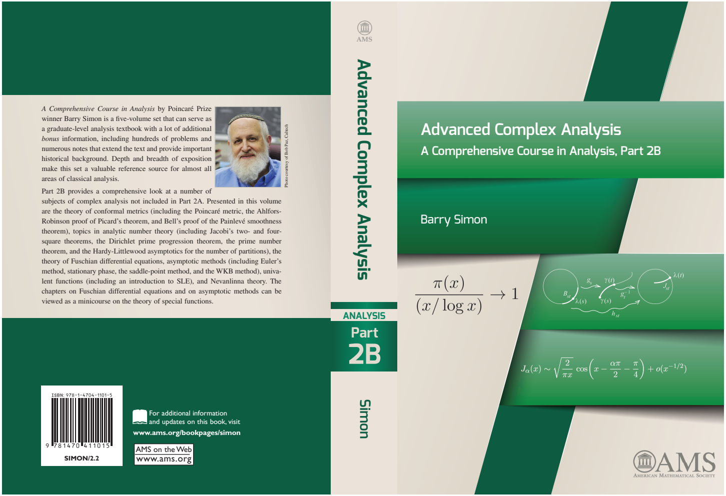
Google *simon comprehensive course*

And Now a Word from Our Sponsor



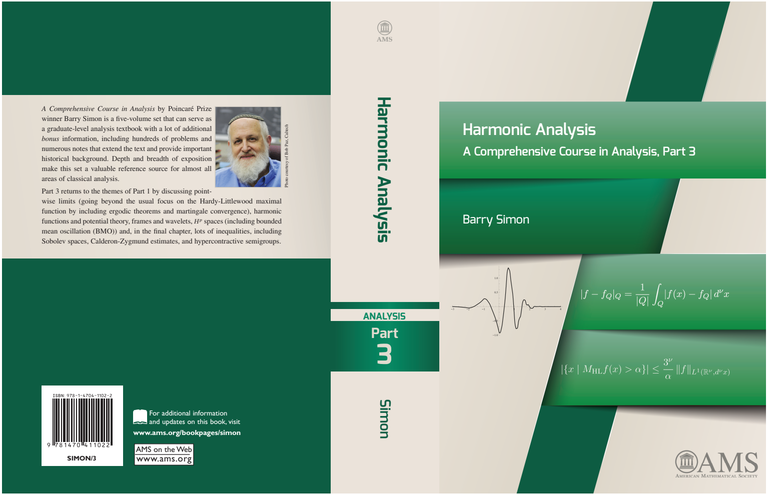
Google *simon comprehensive course*

And Now a Word from Our Sponsor



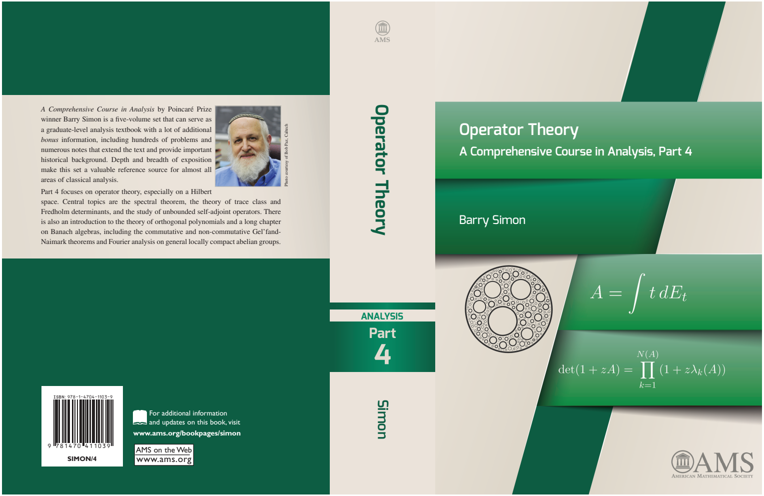
Google *simon comprehensive course*

And Now a Word from Our Sponsor



Google *simon comprehensive course*

And Now a Word from Our Sponsor



Google *simon comprehensive course*