Spectral Theory Sum Rules, Meromorphic Herglotz Functions and Large Deviations

Barry Simon

IBM Professor of Mathematics and Theoretical Physics, Emeritus

California Institute of Technology

Pasadena, CA, U.S.A.

Overview

My goal in this talk is to explain two sum rules in the spectral theory of orthogonal polynomials, one of which I was involved with about 20 years ago. I will then describe the original method of proof exploiting methods of complex analysis. We'll see various functions that arise in that approach seem ad hoc and mysterious. Finally I'll explain a more recent approach of some probabilists that obtains the result using the method of large deviations for some standard random matrix models. This new approach will explain what the previously ad hoc functions are and take the mystery away. It will expose interesting new connections between random matrix theory and the spectral theory of orthogonal polynomials.

Szegő Recursion and Verblunsky Coefficients

We start with orthogonal polynomials on the unit circle, aka OPUC. Let $d\mu$ be a probability measure on $\partial\mathbb{D}$. Then, there are monic orthogonal polynomials, Φ_n, and recursion relations due to Szegő in 1939

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

where Φ_n^* is the polynomial obtained by conjugating and reversing the order of the coefficients. The $\{\alpha_n\}_{n=0}^{\infty}$, called *Verblunsky coefficients*, lie in $\mathbb{D}$ and there is a one-one correspondence, called the *Verblunsky map*, from measures of infinite support and sequences in $\mathbb{D}$. 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. $d\theta$. 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

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

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

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

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) 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. $\int \log(w(\theta))\frac{d\theta}{2\pi} > -\infty$. 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!

Jacobi Parameters

Next orthogonal polynomials on the real line, aka OPRL. One starts with a probability measure, μ, of compact support in $\mathbb{R}$ and forms the orthonormal polynomials, $\{p_n(x)\}_{n=0}^{\infty}$. They obey recursion relations

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

which sets up a one-one correspondence (which we'll call the *Jacobi map*) between such measures (with an infinity of points in their support) and sequences $\{a_n, b_n\}_{n=1}^{\infty}$ of bounded a's in $(0, \infty)$ and b's in $\mathbb{R}$ (called *Jacobi parameters*). If P_n are the monic OPs, then one has $\|P_n\| = a_1 \dots a_n$.

Jacobi Parameters

There is also a correspondence between point measures with finite support and suitable sets of finitely many Jacobi parameters. If there are n pure points, then P_n is 0 in $L^2(d\mu)$ so $a_n = 0$ and again there are $2n-1$ Jacobi parameters – n b's and $n-1$ a's.

For later purposes, I need some details on one approach to going from measures to Jacobi parameters. The more usual method than the one I want to discuss just forms the OPs and looks at the recursion parameters. Instead, consider the *once stripped Jacobi parameters*, i.e. $\{a_{j+1}, b_{j+1}\}_{j=1}^{\infty}$ obtained by dropping the first row and column of the Jacobi matrix. For any non–trivial probability measure of compact support, let $m(z) = \int d\mu(x)/(x-z)$ and let m_1 be the spectral measure for the once stripped problem.

Jacobi Parameters

Then one can prove that

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

Using that $m_1(z) = -z^{-1} + \mathrm{O}(z^{-2})$, one sees that one can 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. Conversely, one can go from Jacobi parameters to m (and so μ) by summing a continued fraction expansion (that goes back to Jacobi, Chebyshev and Markov). One can also get the just mentioned Taylor coefficients as polynomials in the Jacobi parameters.

One consequence of this is that the poles of m_1 (i.e. the pure points of μ_1) are precisely the zeros of m.

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

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

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 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!

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 $d=1$, $p=1$)

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

Killip–Simon Sum Rule

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

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.

Step–by–Step Sum Rule

I know of many proofs of Szegő's Theorem but until recently all proofs of the Killip–Simon sum rule were variants of our original proof which I want to describe some parts of. A key part is that it is required to hold in case both sides are infinite and it is hard to control infinite sums so we had the idea of building up the sums. Suppose that the sum rule holds for both μ and the once striped measure, μ_1. Then subtracting one sum from the other we get that

$$Q(\mu) - Q(\mu_1) + \sum_{j,\pm} \left[F(E_j^\pm) - F(E_j^{(1)\pm}) \right] = \tfrac{1}{4} b_1^2 + \tfrac{1}{2} G(a_1)$$

Because the eigenvalues $E_j^\pm$ and $E_j^{(1)\pm}$ interlace and F is monotone, the sum is of positive terms and always convergent (interlacing sums).

Step–by–Step Sum Rule

While the $\log(w)$ integral might divergence, one can show that a $\log(w/w_1)$ integral is always convergent. So this formula always makes sense as finite terms. We'll discuss the proof of this formula, called, for obvious reasons, a *step–by-step sum rule*.

If you iterate coefficient striping and assume the boundary term goes away, you get the full sum rule. What Killip and I found is so long as lots of terms were positive, one could get an always–valid full sum rule from a step–by–step sum rule. The proof is a somewhat subtle. Among other things it used the fact that the function Q is lower semi-continuous in μ which we discovered by noting it was a relative entropy. We were proud of this realization although shortly afterwards we discovered Verblubsky's papers and found he also proved and used semicontinuity (although without realizing he had an entropy) – in 1935!

Poisson–Jensen Formula

The step–by–step sum rule will involve a Poisson–Jensen formula whose classical form we recall. Define Blaschke factors, $b(z,w)$ to be z is $w=0$ and otherwise $-\frac{|w|(z-w)}{w(1-\bar{w}z)}$ Let f be analytic on the unit disk and in Nevanlinna class, i.e. $\sup_{0<r<1}\int \log_+(|f(re^{i\theta})|)\,d\theta<\infty$. If $\{z_j\}_{j=1}^N$ is a listing of the zeros of f, then $\sum_{j=1}^N(1-|z_j|)<\infty$ which implies that $B(z)=\prod_{j=1}^N b(z,z_j)$ converges to an analytic function vanishing exactly at the z_j. Suppose also that for some $p>1$, we have that $\log(f(z)/B(z))$ lies in H^p (which we'll call the "L^p–condition").

The famous theorem of Smirnov and Beurling says that for some $\omega\in\partial\mathbb{D}$, we have that (Poisson–Jensen formula)

$$f(z)=\omega B(z)\exp\left(\int\frac{e^{i\theta}+z}{e^{i\theta}-z}\log|f(e^{i\theta})|\frac{d\theta}{2\pi}\right)$$

Without the L^p–condition, there a singular inner part.

Meromorphic Herglotz Functions

By a *meromorphic Herglotz* function, we mean a function meromorphic on $\mathbb{D}$, real on $(-1,1)$ with $\operatorname{Im} z > 0 \Rightarrow \operatorname{Im} f(z) > 0$. It is easy to see that such functions have zeros and poles only on $(-1,1)$ and the zeros and poles are simple and interlace. If one looks at the product of Blaschke factors and their inverses for the zeros and poles in $(-r,r)$, it can be shown that they have a limit as $r \uparrow 1$ – an analog of alternating sums converging. Let's suppose $f(0) = 0$ and let $B(z)$ be the limiting product of zero and pole Blashcke factors other than the zero at 0.

Meromorphic Herglotz Functions

One can prove that in $\mathbb{D} \cap \mathbb{C}_+$, one has that $|\arg z B(z)| \le 2\pi$ so that $\arg(f(z)/zB(z))$ is bounded on $\mathbb{D}$. Since $\arg(g) = \mathrm{Im}(\log g)$, M. Riesz's Theorem implies that $\log(f(z)/zB(z))$ is in all H^p with $p < \infty$ so it obeys a Poisson–Jensen formula (with no singular inner part). Thus

$$f(z) = zB(z) \exp\left(\int \frac{e^{i\theta} + z}{e^{i\theta} - z} \log|f(e^{i\theta})| \frac{d\theta}{2\pi} \right)$$

Taking $\log$'s, one gets relations between Taylor coefficients of $\log(f(z)/z)$, certain sums involving $\log$s or powers of zeros and poles and integrals $\cos(n\theta) \log|f(e^{i\theta})|$.

Case Step–by-Step Sum Rules

Recall that $m(z) = \int d\mu(x)/(x-z)$. It defines a Herglotz function on $\mathbb{C}_+$, real on $\mathbb{R}$. Thus $M(z) = -m(z+z^{-1})$ is what we called a meromorphic Herglotz function. Its poles are the eigenvalues of J under the inverse image of the map $z \mapsto z + z^{-1}$ and its zeros are the same for J_1. The Taylor coefficients of $\log M(z)$ about zero are related to those of $m(z)$ at infinity and so polynomials in the Jacobi parameters.

The above procedure thus yields a relation between polynomials of Jacobi parameters, the difference of functions of the eigenvalues of J and J_1 and integral of $\log|M(e^{i\theta})|$. Because $m(z)^{-1} = b_1 - z - a_1^2 m_1(z)$, one finds that $|M(e^{i\theta})|^{-2}\mathrm{Im}\, M(e^{i\theta}) = a_1^2 \mathrm{Im}\, M_1(e^{i\theta})$ so the $\log$ integral is a $\log$ of ratios of w and w_1.

P_2 Sum Rule

What results is a step–by–step sum rule which if iterated with boundary terms dropped yields the formal sum rules stated by Case (although, unlike Case, Killip and I had explicit formulae for the polynomials in the Jacobi parameters). These C_n step–by–step sum rules, especially C_0 have turned out to be useful in spectral theory, but to get a gem, one needs positivity and Killip and I found that none of the Case rules had the required positivity.

However, we discovered that $C_0 + \frac{1}{2}C_2$ had the required positivity. We had no explanation of why this was so but observed it. We called this the P_2 sum rule (P for positive) and it is now known as the Killip–Simon sum rule. The rather complicated functions F and G just arose by taking the functions from the Case sum rule and combining them.

Mysteries

While the gem one gets from the P_2 sum rule is simple and elegant, the proof has lots of mysteries:

1. Why are there any positive combinations?
2. It is easy to understand the $(4 - x^2)^{-1/2}\, dx$ of the Szegő condition. It is $d\theta$ under $x = \cos(\theta)$. Equivalently, it is the potential theoretic equilibrium measure for $[-2, 2]$ but where the heck does the $(4 - x^2)^{1/2}\, dx$ come from?
3. What does the function
$$G(a) = a^2 - 1 - \log(a^2)$$
mean?
4. What does the function
$$F(E) = \tfrac{1}{4}[\beta^2 - \beta^{-2} - \log \beta^4]; \quad E = \beta + \beta^{-1}$$
mean?

The LD Framework

Large deviations go back to Laplace. The modern theory was initiated by Cramér in the 1930's and made into a powerful machine by Donsker–Varadhan and Freidlin–Wentzel and then Varadhan alone (work for which he got the Abel prize). Two standard texts are Deuschel–Stroock and Dembo–Zeitouni.

We consider a sequence of probability measures, $\{\mu_n\}_{n=1}^{\infty}$, on a space, X. Naively, one has a Large Deviation Principle (aka LDP) if the μ_n–probability that x is near x_0 is $\mathrm{O}(e^{-nI(x_0)})$. To be mathematically precise, one supposes that X is a Polish space (aka complete metric space), allows multiplicative factors other than n and so speaks of the *speed*, a_n, *rate function*, $I : X \to [0, \infty]$ and requires that:

The LD Framework

1. I is lower semicontinuous
2. For all closed sets $F \subset X$
   $$\limsup_{n\to\infty} \frac{1}{a_n} \log \mu_n(F) \le -\inf_{x\in F} I(x)$$
3. For all open sets $U \subset X$
   $$\liminf_{n\to\infty} \frac{1}{a_n} \log \mu_n(U) \ge -\inf_{x\in U} I(x)$$

One of the simplest but also most powerful results is that of Cramér: Let $\{X_j\}_{j=1}^\infty$ be iidrv with individual expectation $\mathbb{E}$. Let μ_n be the distribution on $\mathbb{R}$ of $\frac{1}{n}\sum_{j=1}^n X_j$. Then an LDP holds with speed n and rate function

$$I(x) = \sup_\theta \left[\theta x - \log\left(\mathbb{E}(e^{\theta X})\right)\right]$$

Cramér Example

Let X be an exponential random variable, i.e. with density $\chi_{[0,\infty)}(x)e^{-x}\,dx$. Then

$$\log\left(\mathbb{E}(e^{\theta X})\right) = \begin{cases} -\log(1-\theta), & \text{if } \theta < 1 \\ \infty, & \text{if } \theta \geq 1 \end{cases}$$

For $x \leq 0$, taking $\theta \to -\infty$ in $\theta x - \log\left(\mathbb{E}(e^{\theta X})\right)$, we see that $I(x) = \infty$. If $x > 0$, the θ derivative of $\theta x - \log\left(\mathbb{E}(e^{\theta X})\right)$ vanishes at $\theta = 1 - x^{-1}$ at which point $\theta x - \log\left(\mathbb{E}(e^{\theta X})\right)$ has the value $x - 1 - \log(x)$. Thus

$$I(x) = \begin{cases} x - 1 - \log(x), & \text{if } x > 0 \\ \infty, & \text{if } x \leq 0 \end{cases}$$

Notice that $G(a) = I(a^2)$, which we'll see is no coincidence!!!

LDP and Sum Rules

Gamboa, Nagel and Rouault had the following lovely idea. Let X be the set of probability measures on $\partial\mathbb{D}$ or on $\mathbb{R}$ (with some song and dance to handle measures which don't have compact support — I'll henceforth suppress this phrase) and suppose we have a sequence of probability measures on X with an LDP. The Verblunsky and Jacobi maps are continuous to sequences of Verblunsky coefficients or Jacobi parameters and so one has an LDP on sequence space. But the rate functions are clearly the same, so we have the equality of a function of the spectral measures and of a function of the parameters *and as rate functions, these functions are automatically non-negative*!!!!! We thus have a way to generate positive sum rules and demanding they be finite gives us a gem.

LDP and Sum Rules

GNR had the further idea that the measures on the spectral measures should come from random matrix measures with a cyclic vector in the limit as the matrix dimension goes to infinity.

Of course, the issue becomes to effectively compute the rate function on both sides and alas, we haven't yet found a magic way to do these calculations in a general context.

The reception of the GNR paper illustrates the dangers of working in between two disparate areas. They wrote the paper in a way that only experts on large deviations could understand it, but such experts didn't understand the spectral theory context.

LDP and Sum Rules

Jonathan Breuer and I couldn't understand the paper so we consulted Ofer Zeitouni who said he'd looked quickly at the paper and there didn't seem to be much new there! In fact, the calculations of rate functions on the two sides wasn't so far from prior calculations of rate functions. What was new was the realization that because a rate function could be computed in two ways, one is able to prove interesting equalities. So they had some troubles getting published what I regard as one of the more interesting recent papers in spectral theory. In the end, Jonathan, Ofer and I used their methods to study higher order sum rules and we also wrote a pedagogic translation of their paper accessible to spectral theorists.

CUE: Measure Side

Circular Unitary Ensemble, aka CUE, is just another name for Haar Measure in $U(n)$, the n x n unitary matrices, for varying n. Any fixed vector is cyclic with probability one, so the corresponding spectral measures have the form $\sum_{j=1}^{n} w_j \delta_{\theta_j}$ where $\lambda_j \equiv e^{i\theta_j}$ are the eigenvalues. Haar measure induces a measure on measures which is supported on the n-point measures.

As is well–known, the λ's and w's are independent of each other, the w's are uniformly distributed on the simplex $\{\mathbf{w} | \sum_{j=1}^{n} w_j = 1\}$ and by the Weyl integration formula, the θ's have distribution

$$\frac{1}{n!} \prod_{1 \le j < k \le n} |e^{i\theta_j} - e^{i\theta_k}|^2 \prod_{j=1}^{n} \frac{d\theta_j}{2\pi}$$

CUE: Measure Side

The first step in the analysis of the measure side is to analyze what probabilists call the *empirical measure* and physicists *the density of states*, namely the random measure $\frac{1}{n}\sum_{j=1}^{n}\delta_{\theta_j}$. This also defines a family of measures on measures and, in 1997, Ben Arous and Guionnet made the important discovery that this (or rather an analog on the real line with a confining potential) has an LDP with speed n^2 (note the square) and rate function the 2D Coulomb energy $-\int \log|x-y|\,d\mu(x)\,d\mu(y)$.

This is easy to understand. The Weyl distribution can be viewed as a discrete two dimensional Coulomb gas in the canonical ensemble (2D because $|x-y|^{-2}$ is the exponential of $-2\log|x-y|$). The $n\to\infty$ limit is a high density limit and due to repulsion, there is a strong tendency towards equal spacing.

CUE: Measure Side

To get a significant difference from equal spacing, one has $\mathrm{O}(n^2)$ smaller distances and so the speed is n^2. The optimal spacing will still be locally equal and the discrete Coulomb energy will converge to the continuum.

The fact that n^2 is much larger than n implies that for a measure to have finite rate at speed n, it has to have points close to uniformly distributed and the large deviations comes entirely from the lack of a uniform weight.

The weights are close to independent (except for the normalization they are) – a slick way to see this is to note if Y_j are positive exponentially distributed iidrv, then $w_j = Y_j / \sum_{j=1}^{n} Y_j$. This allows one (using Cramér's theorem on small blocks) to prove an LDP for the spectral measure with speed n and rate function the Szegő integral $-\int \log(w(\theta))\frac{d\theta}{2\pi}$.

CUE: Coefficient Side

In 2004, Killip and Nenciu wrote down the distribution of $\{\alpha_j\}_{j=0}^{n-1}$ induced by restricting Haar measure as we are. The α's are independent with α_{n-1} (which lies on $\partial\mathbb{D}$) uniformly distributed on $\partial\mathbb{D}$ and for $j = 0 \ldots n-2$, α_j has density on $\mathbb{D}$

$$\frac{n-j-1}{\pi}(1-|z|^2)^{n-j-2}\, d^2z$$

which says that α_j is distributed as the first complex component of a unit vector in $\mathbb{C}^{n-j}$.

$\prod \rho_j^2$ appears to the nth power so the rate function is $-\sum_{j=1}^{\infty} \log(1-|\alpha_j|^2)$. In this calculation, one makes use of the theory of LDP projective limits to handle the technicalities of going from finite to infinite support. So, voilá, a new proof of Szegő's Theorem!!!!!

GUE

The Gaussian Unitary Ensemble, aka GUE, is the probability measure on $n \times n$ self adjoint matrices so that $\{a_{ii}\}_{i=1,\dots,n}, \{\mathrm{Re}(a_{ij})\}_{1\le i<j\le n}$ and $\{\mathrm{Im}(a_{ij})\}_{1\le i<j\le n}$ are independent identically distributed Gaussian random variables of mean zero and suitable, n–dependent variance.

The argument for GUE, normalized so the limiting density is the semicircle law on $[-2,2]$, is similar to that for CUE. Instead of results of Killip-Nenciu for the distribution of α's, one has earlier results of Dumitriu and Edelman for the Jacobi parameters. The calculation is made easier by the independence of the Jacobi parameters (which leads to sums of terms that depend only on a single a or b).

GUE

One needs to make some additional arguments going back to Ben Arous-Dembo-Guionnet to deal with eigenvalues outside the essential support.

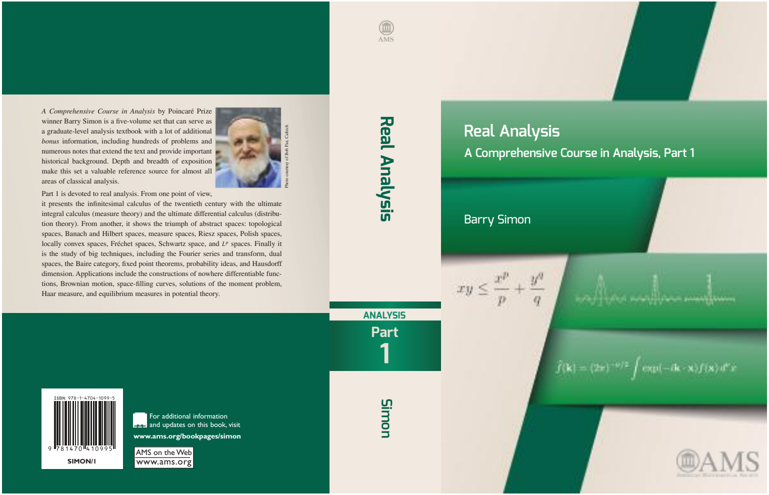
What results is a new proof of the Killip-Simon sum rule.

Mysteries Solved

We can now solve the mysteries:

1. *Why are there any positive combinations?* This is the basic GNR theory of positive sum rules.
2. *It is easy to understand the* $(4-x^2)^{-1/2}\,dx$ *of the Szegő condition but where the heck does the* $(4-x^2)^{1/2}\,dx$ *come from?* This is the Wigner semi–circle law; essentially the measure is the potential theory equilibrium measure in quadratic external field.
3. *What does the function*
$$G(a) = a^2 - 1 - \log(a^2)$$
mean? As we've seen, this is the rate function for square roots of sums of exponential RVs.
4. *What does the function*
$$F(E) = \tfrac{1}{4}[\beta^2 - \beta^{-2} - \log\beta^4]; \quad E = \beta + \beta^{-1}$$
mean? This is the Coulomb potential of the Wigner semi–circle distribution plus a quadratic external field.

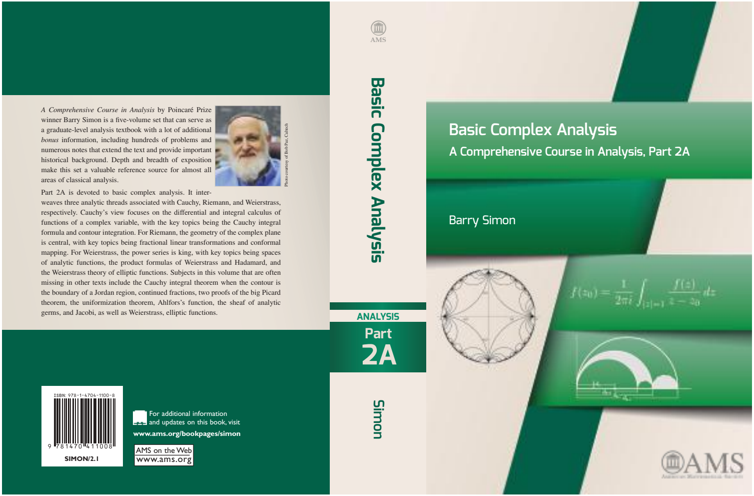
And Now a Word from Our Sponsor



Google *simon comprehensive course*

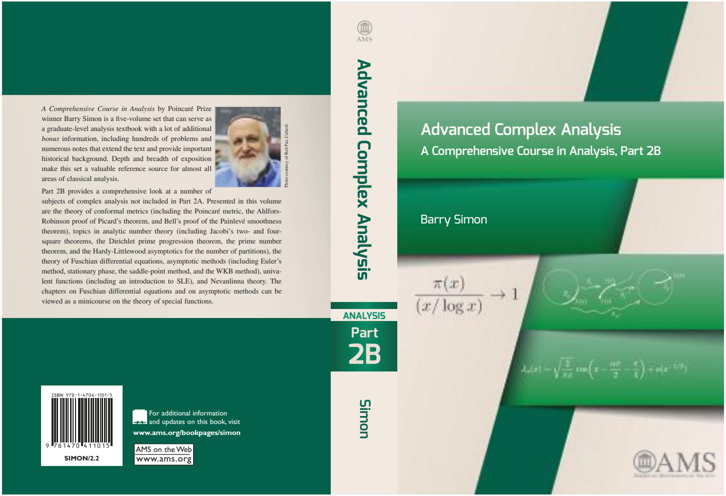
And Now a Word from Our Sponsor

- OPUC
- Szegő–Verblunsky Sum Rule
- OPRL
- Killip–Simon Sum Rule
- Meromorphic Herglotz Functions
- MHT Proof
- Mysteries
- Large Deviations
- GNR Strategy
- CUE and Szegő
- **GUE and Killip–Simon**



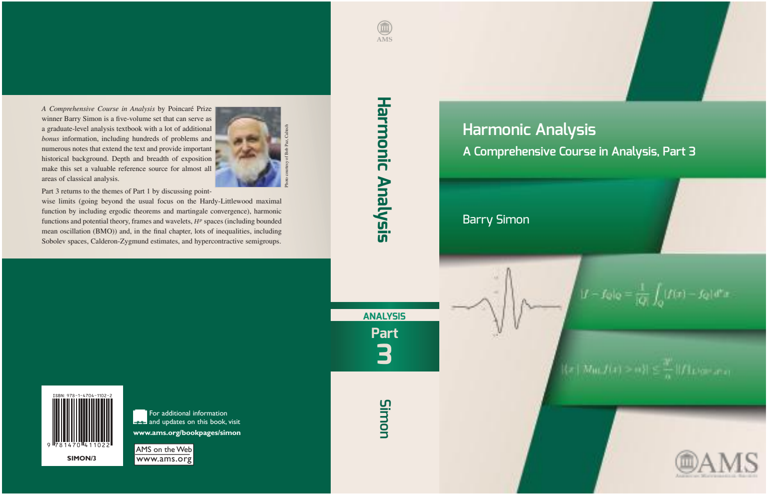
Google *simon comprehensive course*

And Now a Word from Our Sponsor



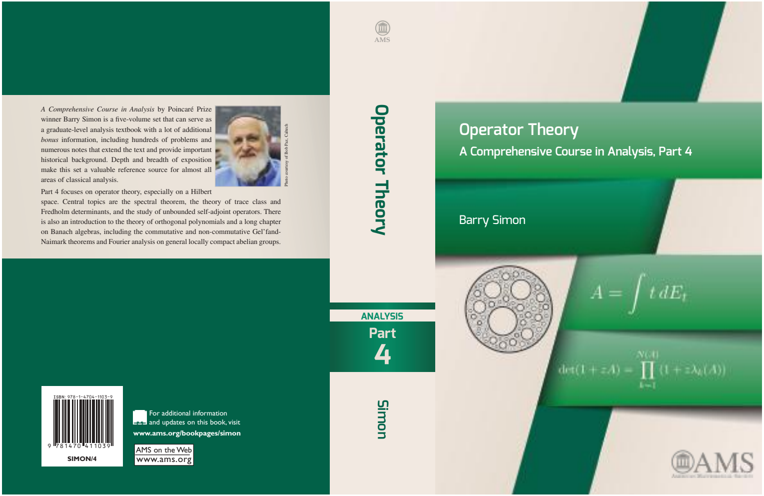
Google *simon comprehensive course*

And Now a Word from Our Sponsor



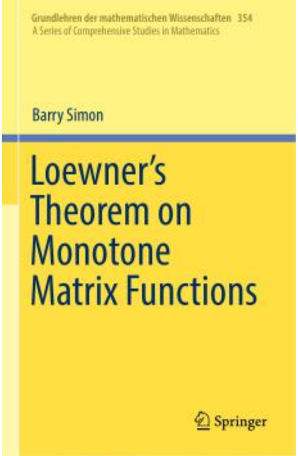
Google *simon comprehensive course*

And Now a Word from Our Sponsor



Google *simon comprehensive course*

And Now a Word from Our Sponsor



And tada, the latest book