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 4: GNR Proof of Sum Rules
GNR Proof of 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


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) and suppose we have a sequence of probability measures on $X$ with an LDP.
Gamboa, Nagel and Rouault had the following lovely idea. Let $X$ be the set of probability measures on $\partial D$ or on $\mathbb{R}$ (with some song and dance to handle measures which don’t have compact support) 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.
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) 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
Gamboa, Nagel and Rouault had the following lovely idea. Let $X$ be the set of probability measures on $\partial D$ or on $\mathbb{R}$ (with some song and dance to handle measures which don’t have compact support) 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 \textit{and as rate functions, these functions are automatically non-negative}!!!!!
Gamboa, Nagel and Rouault had the following lovely idea. Let $X$ be the set of probability measures on $\partial D$ or on $\mathbb{R}$ (with some song and dance to handle measures which don’t have compact support) 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.
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.
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.
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.
GNR Approach

SGō Coefficient
Side

SGō Measure
Side

Killip Simon via
LDP

Further
Developments

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,
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.
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.
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.
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.
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.
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.
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.
To be explicit about the random matrix models:
To be explicit about the random matrix models:

- the Szegő–Verblunsky sum rule comes from CUE, aka Circular Unitary Ensemble, the family on the spectral measures induced by Haar measure on $\mathbb{U}(n)$.
To be explicit about the random matrix models:

- the Szegő–Verblunsky sum rule comes from CUE, aka Circular Unitary Ensemble, the family on the spectral measures induced by Haar measure on $U(n)$.

- the Killip–Simon sum rules comes from GUE, aka Gaussian Unitary Ensemble, the measure on random $n \times n$ self–adjoint matrices has $\{\text{Re} M_{ij}^{(n)}\}_{1 \leq i \leq j \leq n}$ and $\{\text{Im} M_{ij}^{(n)}\}_{1 \leq i < j \leq n}$ Gaussian iid with mean zero and $E([M_{ii}^{(n)}]^2) = n^{-1}$. 

To be explicit about the random matrix models:

- The Szegő–Verblunsky sum rule comes from CUE, aka Circular Unitary Ensemble, the family on the spectral measures induced by Haar measure on $\mathbb{U}(n)$.

- The Killip–Simon sum rules come from GUE, aka Gaussian Unitary Ensemble, the measure on random $n \times n$ self-adjoint matrices has \( \{\text{Re} M_{ij}^{(n)}\}_{1 \leq i \leq j \leq n} \) and \( \{\text{Im} M_{ij}^{(n)}\}_{1 \leq i < j \leq n} \) Gaussian iid with mean zero and \( \mathbb{E}([M_{ii}^{(n)}]^2) = n^{-1} \).

(GNR use GOE rather than GUE but that only means our sum rules are twice theirs).
To be explicit about the random matrix models:

- the Szegő–Verblunsky sum rule comes from CUE, aka Circular Unitary Ensemble, the family on the spectral measures induced by Haar measure on $\mathbb{U}(n)$.
- the Killip–Simon sum rules comes from GUE, aka Gaussian Unitary Ensemble, the measure on random $n \times n$ self–adjoint matrices has $\{\Re M_{ij}^{(n)}\}_{1 \leq i \leq j \leq n}$ and $\{\Im M_{ij}^{(n)}\}_{1 \leq i < j \leq n}$ Gaussian iid with mean zero and $\mathbb{E}([M_{ii}^{(n)}]^2) = n^{-1}$.

(GNR use GOE rather than GUE but that only means our sum rules are twice theirs). Note the curious fact that on the support of the measures $\mathbb{P}_n$ (which is easily seen to be the measures with at most $n$ pure points (only)), we have that $I = \infty$ because there is no a.c. part.
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof.
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let

\[
Y_\infty = \mathbb{D}^\infty, \quad Y_n = \left( \prod_{j=0}^{n-2} \mathbb{D} \right) \times \partial \mathbb{D}, \quad Y = Y_\infty \cup \bigcup_{n=1}^{\infty} Y_n
\]

This topology is such that the map from probability measures to \(Y\) is a homeomorphism.
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let

\[ Y_\infty = \mathbb{D}^\infty \quad Y_n = \left( \prod_{j=0}^{n-2} \mathbb{D} \right) \times \partial \mathbb{D} \quad Y = Y_\infty \cup \bigcup_{n=1}^{\infty} Y_n \]

The topology is metrizable with convergence given by

\[ \alpha^{(n)} \to \alpha^{(\infty)} \text{ with } \alpha^{(\infty)} \in Y_\infty \iff \alpha_j^{(n)} \to \alpha_j^{(\infty)} \text{ for all } j \]
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let

\[
Y_\infty = \mathbb{D}^\infty \quad Y_n = \left( \prod_{j=0}^{n-2} \mathbb{D} \right) \times \partial \mathbb{D} \quad Y = Y_\infty \cup \bigcup_{n=1}^{\infty} Y_n
\]

The topology is metrizable with convergence given by

\[
\alpha^{(n)} \to \alpha^{(\infty)} \text{ with } \alpha^{(\infty)} \in Y_\infty \iff \alpha_j^{(n)} \to \alpha_j^{(\infty)} \text{ for all } j
\]

and if \( \alpha^{(\infty)} \in Y_m \), then for eventually, \( \alpha^{(n)} \in Y_\infty \cup \bigcup_{n=m}^{\infty} Y_n \) and \( \alpha_j^{(n)} \to \alpha_j^{(\infty)} \), \( j = 0, \ldots, m - 1 \).
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let

\[ Y_\infty = \mathbb{D}^\infty \quad Y_n = \left( \prod_{j=0}^{n-2} \mathbb{D} \right) \times \partial \mathbb{D} \quad Y = Y_\infty \cup \bigcup_{n=1}^{\infty} Y_n \]

The topology is metrizable with convergence given by

\[ \alpha^{(n)} \to \alpha^{(\infty)} \text{ with } \alpha^{(\infty)} \in Y_\infty \iff \alpha_j^{(n)} \to \alpha_j^{(\infty)} \text{ for all } j \]

and if \( \alpha^{(\infty)} \in Y_m \), then for eventually, \( \alpha^{(n)} \in Y_\infty \cup (\bigcup_{n=m}^{\infty} Y_n) \) and \( \alpha_j^{(n)} \to \alpha_j^{(\infty)}, j = 0, \ldots, m-1 \).

This topology is such that the map from probability measures to \( Y \) is a homeomorphism.
Topology of VCs

In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let

\[
Y_\infty = \mathbb{D}^\infty \quad Y_n = \left( \prod_{j=0}^{n-2} \mathbb{D} \right) \times \partial \mathbb{D} \quad Y = Y_\infty \cup \bigcup_{n=1}^{\infty} Y_n
\]

The topology is metrizable with convergence given by \( \alpha^{(n)} \to \alpha^{(\infty)} \) with \( \alpha^{(\infty)} \in Y_\infty \iff \alpha_j^{(n)} \to \alpha_j^{(\infty)} \) for all \( j \) and if \( \alpha^{(\infty)} \in Y_m \), then for eventually, \( \alpha^{(n)} \in Y_\infty \cup \left( \bigcup_{n=m}^{\infty} Y_n \right) \) and \( \alpha_j^{(n)} \to \alpha_j^{(\infty)} \), \( j = 0, \ldots, m - 1 \). This topology is such that the map from probability measures to \( Y \) is a homeomorphism.

Let \( X = \overline{\mathbb{D}}^\infty \).
In the rest of the lectures, we’ll describe the CUE proof in some detail and then sketch the GUE proof. We begin by describing the set of Verblunsky coefficients and the topology on it. Let

\[
Y_\infty = \mathbb{D}^\infty \quad Y_n = \left( \prod_{j=0}^{n-2} \mathbb{D} \right) \times \partial \mathbb{D} \quad Y = Y_\infty \cup \bigcup_{n=1}^{\infty} Y_n
\]

The topology is metrizable with convergence given by

\[
\alpha^{(n)} \rightarrow \alpha^{(\infty)} \text{ with } \alpha^{(\infty)} \in Y_\infty \iff \alpha_j^{(n)} \rightarrow \alpha_j^{(\infty)} \text{ for all } j
\]

and if \( \alpha^{(\infty)} \in Y_m \), then for eventually, \( \alpha^{(n)} \in Y_\infty \cup \left( \bigcup_{n=m}^{\infty} Y_n \right) \) and \( \alpha_j^{(n)} \rightarrow \alpha_j^{(\infty)}, j = 0, \ldots, m - 1 \).

This topology is such that the map from probability measures to \( Y \) is a homeomorphism.

Let \( X = \overline{\mathbb{D}}^\infty \). Then the map \( H : X \rightarrow Y \) by dropping all \( \alpha_j \) after the first one in \( \partial \mathbb{D} \) is continuous.
Let $P_N$ be the measure on $X$ given by the Killip–Nenciu formula on the first $N$ factors and a point mass at 0 on the remaining coordinates.
Let $\mathbb{P}_N$ by the measure on $X$ given by the Killip–Nenciu formula on the first $N$ factors and a point mass at 0 on the remaining coordinates. Let $X_j$ be $\overline{D}^j$ and $\pi_j : X \to X_j$ projection onto the first $j$ coordinates.
Computation of $I$ on $Y$

Let $\mathbb{P}_N$ by the measure on $X$ given by the Killip–Nenciu formula on the first $N$ factors and a point mass at 0 on the remaining coordinates. Let $X_j$ be $\mathbb{D}^j$ and $\pi_j : X \to X_j$ projection onto the first $j$ coordinates. By our result on and LDP for measures of the form $F(x)e^{-NG(x)}d\nu x$, we see that $\pi_j^*(\mathbb{P}_N)$ obeys and LDP with speed $N$ and rate $I_j(\{\alpha_k\}_{k=0}^{j-1}) = -\sum_{k=0}^{j-1} \log(1 - |\alpha_k|^2)$. 

Given the map $H$ from the set of allowed Verblunsky coefficients and $X$, one notes that the Killip–Nenciu Theorem says that $\mathbb{P}(H)N$ is precisely the measure on VCs induced by Haar measure on $U(n)$.
Computation of $I$ on $Y$

Let $\mathbb{P}_N$ by the measure on $X$ given by the Killip–Nenciu formula on the first $N$ factors and a point mass at 0 on the remaining coordinates. Let $X_j$ be $\mathbb{D}^j$ and $\pi_j : X \rightarrow X_j$ projection onto the first $j$ coordinates. By our result on and LDP for measures of the form $F(x) e^{-NG(x)} d\nu x$, we see that $\pi_j^*(\mathbb{P}_N)$ obeys and LDP with speed $N$ and rate

$$I_j(\{\alpha_k\}_{k=0}^{j-1}) = - \sum_{k=0}^{j-1} \log(1 - |\alpha_k|^2).$$

It follows by the projective limit theorem that $\mathbb{P}_N$ has an LDP with speed $N$ and rate function

$$I(\{\alpha_k\}_{k=0}^{\infty}) = - \sum_{k=0}^{\infty} \log(1 - |\alpha_k|^2).$$
Let $\mathbb{P}_N$ by the measure on $X$ given by the Killip–Nenciu formula on the first $N$ factors and a point mass at 0 on the remaining coordinates. Let $X_j$ be $\mathbb{D}^j$ and $\pi_j : X \to X_j$ projection onto the first $j$ coordinates. By our result on and LDP for measures of the form $F(x)e^{-NG(x)}d\nu x$, we see that $\pi_j^*(\mathbb{P}_N)$ obeys and LDP with speed $N$ and rate

$$I_j(\{\alpha_k\}_{k=0}^{j-1}) = -\sum_{k=0}^{j-1} \log(1 - |\alpha_k|^2).$$

It follows by the projective limit theorem that $\mathbb{P}_N$ has an LDP with speed $N$ and rate function $I(\{\alpha_k\}_{k=0}^{\infty}) = -\sum_{k=0}^{\infty} \log(1 - |\alpha_k|^2)$. Given the map $H$ from the set of allowed Verblunsky coefficients and $X$, one notes that the Killip–Nenciu Theorem says that $\mathbb{P}_N^H$ is precisely the measure on VCs induced by Haar measure on $\mathbb{U}(n)$. 
Let $\mathbb{P}_N$ be the measure on $X$ given by the Killip–Nenciu formula on the first $N$ factors and a point mass at 0 on the remaining coordinates. Let $X_j$ be $\mathbb{D}^j$ and $\pi_j : X \to X_j$ projection onto the first $j$ coordinates. By our result on and LDP for measures of the form $F(x)e^{-NG(x)}d\nu x$, we see that $\pi_j^*(\mathbb{P}_N)$ obeys and LDP with speed $N$ and rate

$$I_j(\{\alpha_k\}_{k=0}^{j-1}) = -\sum_{k=0}^{j-1} \log(1 - |\alpha_k|^2).$$

It follows by the projective limit theorem that $\mathbb{P}_N$ has an LDP with speed $N$ and rate function $I(\{\alpha_k\}_{k=0}^{\infty}) = -\sum_{k=0}^{\infty} \log(1 - |\alpha_k|^2)$.

Given the map $H$ from the set of allowed Verblunsky coefficients and $X$, one notes that the Killip–Nenciu Theorem says that $\mathbb{P}_N(H)$ is precisely the measure on VCs induced by Haar measure on $\mathbb{U}(n)$. Applying the contraction principle, we see these measures obey an LDP with rate $I$ as above, one side of the Szegő–Verblunsky sum rule.
We begin our presentation of the calculation of the rate function on the measure side by specifying the distribution of spectral measures induced by $\text{CUE}(n)$ which we’ll also call $\text{CUE}(n)$. 

Let $\{e_j\}_{j=1}^n$ be the standard basis for $\mathbb{C}^n$. It is easy to see that for a.e. $U$, $e_1$ is a cyclic vector for $U$ so that $U$ and $e_1$ define a spectral measure $d\mu(\theta) = \sum_{j=1}^n \lambda_j \delta_{\lambda_j}^\theta$ on $\partial \mathcal{D}$, with precisely $n$ pure points (aka atoms) $\lambda_j = e^{i\theta_j}, j = 1, \ldots, n$. 

Letting $\{\phi_j\}_{j=1}^n$ be the orthonormal basis of eigenvectors of $U$, so that $U\phi_j = \lambda_j \phi_j$, we have $w_j = |\langle \phi_j, e_1 \rangle|^2$. Of course, since $\|e_1\| = 1$, $\sum_{j=1}^n w_j = 1$. 

Distribution of Haar distributed spectral measures

We begin our presentation of the calculation of the rate function on the measure side by specifying the distribution of spectral measures induced by CUE\((n)\) which we’ll also call CUE\((n)\). Let \(\{e_j\}_{j=1}^n\) be the standard basis for \(\mathbb{C}^n\). It is easy to see that for a.e. \(U\), \(e_1\) is a cyclic vector for \(U\) so that \(U\) and \(e_1\) define a spectral measure
Distribution of Haar distributed spectral measures

We begin our presentation of the calculation of the rate function on the measure side by specifying the distribution of spectral measures induced by CUE($n$) which we’ll also call CUE($n$). Let $\{e_j\}_{j=1}^n$ be the standard basis for $\mathbb{C}^n$. It is easy to see that for a.e. $U$, $e_1$ is a cyclic vector for $U$ so that $U$ and $e_1$ define a spectral measure

$$d\mu(\theta) = \sum_{j=1}^n w_j \delta_{\lambda_j}$$
Distribution of Haar distributed spectral measures

We begin our presentation of the calculation of the rate function on the measure side by specifying the distribution of spectral measures induced by CUE($n$) which we’ll also call CUE($n$). Let $\{e_j\}_{j=1}^n$ be the standard basis for $C^n$. It is easy to see that for a.e. $U$, $e_1$ is a cyclic vector for $U$ so that $U$ and $e_1$ define a spectral measure

$$d\mu(\theta) = \sum_{j=1}^{n} w_j \delta_{\lambda_j}$$

on $\partial D$, with precisely $n$ pure points (aka atoms) $\lambda_j = e^{i\theta_j}, j = 1, \ldots, n.$
We begin our presentation of the calculation of the rate function on the measure side by specifying the distribution of spectral measures induced by CUE($n$) which we’ll also call CUE($n$). Let $\{e_j\}_{j=1}^n$ be the standard basis for $\mathbb{C}^n$. It is easy to see that for a.e. $U$, $e_1$ is a cyclic vector for $U$ so that $U$ and $e_1$ define a spectral measure

$$d\mu(\theta) = \sum_{j=1}^n w_j \delta_{\lambda_j}$$

on $\partial\mathbb{D}$, with precisely $n$ pure points (aka atoms) $\lambda_j = e^{i\theta_j}, j = 1, \ldots, n$. Letting $\{\varphi_j\}_{j=1}^n$ be the orthonormal basis of eigenvectors of $U$, so that $U\varphi_j = \lambda_j \varphi_j$, we have $w_j = |\langle \varphi_j, e_1 \rangle|^2$. Of course, since $\|e_1\| = 1$, 

\[ \text{Distribution of Haar distributed spectral measures} \]
We begin our presentation of the calculation of the rate function on the measure side by specifying the distribution of spectral measures induced by CUE\( (n) \) which we’ll also call CUE\( (n) \). Let \( \{ e_j \}_{j=1}^n \) be the standard basis for \( \mathbb{C}^n \). It is easy to see that for a.e. \( U \), \( e_1 \) is a cyclic vector for \( U \) so that \( U \) and \( e_1 \) define a spectral measure

\[
d\mu(\theta) = \sum_{j=1}^{n} w_j \delta_{\lambda_j}
\]

on \( \partial \mathbb{D} \), with precisely \( n \) pure points (aka atoms) \( \lambda_j = e^{i\theta_j}, j = 1, \ldots, n \). Letting \( \{ \varphi_j \}_{j=1}^n \) be the orthonormal basis of eigenvectors of \( U \), so that \( U \varphi_j = \lambda_j \varphi_j \), we have \( w_j = |\langle \varphi_j, e_1 \rangle|^2 \). Of course, since \( \|e_1\| = 1 \),

\[
\sum_{j=1}^{n} w_j = 1
\]
Distribution of Haar distributed spectral measures

For \( \tilde{U} \) an arbitrary unitary, \( \tilde{U}U\tilde{U}^{-1} \) has the same eigenvalues as \( U \) and \( \langle \varphi_j(\tilde{U}U\tilde{U}^{-1}), e_1 \rangle = \langle \tilde{U} \varphi_j(U), e_1 \rangle \).
For $\tilde{U}$ an arbitrary unitary, $\tilde{U}UU\tilde{U}^{-1}$ has the same eigenvalues as $U$ and $\langle \varphi_j(\tilde{U}UU\tilde{U}^{-1}), e_1 \rangle = \langle U\varphi_j(U), e_1 \rangle$. Since $U \mapsto \tilde{U}UU\tilde{U}^{-1}$ leaves Haar measure invariant, we see that the distribution of the unit vector $(\langle \varphi_1(U), e_1 \rangle, \langle \varphi_2(U), e_1 \rangle, \ldots, \langle \varphi_n(U), e_1 \rangle) \in \mathbb{C}^n$ is invariant under unitary transformations,
Distribution of Haar distributed spectral measures

For $\tilde{U}$ an arbitrary unitary, $\tilde{U}U\tilde{U}^{-1}$ has the same eigenvalues as $U$ and $\langle \varphi_j(\tilde{U}U\tilde{U}^{-1}), e_1 \rangle = \langle \tilde{U}\varphi_j(U), e_1 \rangle$. Since $U \mapsto \tilde{U}U\tilde{U}^{-1}$ leaves Haar measure invariant, we see that the distribution of the unit vector $(\langle \varphi_1(U), e_1 \rangle, \langle \varphi_2(U), e_1 \rangle, \ldots, \langle \varphi_n(U), e_1 \rangle) \in \mathbb{C}^n$ is invariant under unitary transformations, which implies it is the Euclidean measure restricted to the sphere.
Distribution of Haar distributed spectral measures

For \( \tilde{U} \) an arbitrary unitary, \( \tilde{U}U\tilde{U}^{-1} \) has the same eigenvalues as \( U \) and \( \langle \varphi_j(\tilde{U}U\tilde{U}^{-1}), e_1 \rangle = \langle \tilde{U}\varphi_j(U), e_1 \rangle \).

Since \( U \mapsto \tilde{U}U\tilde{U}^{-1} \) leaves Haar measure invariant, we see that the distribution of the unit vector
\[
(\langle \varphi_1(U), e_1 \rangle, \langle \varphi_2(U), e_1 \rangle, \ldots, \langle \varphi_n(U), e_1 \rangle) \in \mathbb{C}^n
\]
is invariant under unitary transformations, which implies it is the Euclidean measure restricted to the sphere. By using the fact that that \( d^2z = \frac{1}{2}d\theta d(|z|^2) \) (which shows it is essential we work in \( \mathbb{C} \)), it is not hard to show that the squares of the components of a complex \( n \)-vector uniformly distributed on the sphere are uniformly distributed on the simplex.
Distribution of Haar distributed spectral measures

For \( \tilde{U} \) an arbitrary unitary, \( \tilde{U}UU\tilde{U}^{-1} \) has the same eigenvalues as \( U \) and 
\[
\langle \varphi_j(\tilde{U}UU\tilde{U}^{-1}), e_1 \rangle = \langle \tilde{U}\varphi_j(U), e_1 \rangle.
\]
Since \( U \mapsto \tilde{U}UU\tilde{U}^{-1} \) leaves Haar measure invariant, we see that the distribution of the unit vector 
\[
(\langle \varphi_1(U), e_1 \rangle, \langle \varphi_2(U), e_1 \rangle, \ldots, \langle \varphi_n(U), e_1 \rangle) \in \mathbb{C}^n
\]
is invariant under unitary transformations, which implies it is the Euclidean measure restricted to the sphere. By using the fact that that 
\[
d^2z = \frac{1}{2}d\theta d(|z|^2)
\]
(which shows it is essential we work in \( \mathbb{C} \)), it is not hard to show that the squares of the components of a complex \( n \)-vector uniformly distributed on the sphere are uniformly distributed on the simplex. Thus we get that the \( \{w_j\}_{j=1}^n \) are independent of the eigenvalues and have \( \mathbb{P}_n \)-distribution.
Distribution of Haar distributed spectral measures

For $\tilde{U}$ an arbitrary unitary, $\tilde{UU}\tilde{U}^{-1}$ has the same eigenvalues as $U$ and $\langle \varphi_j(\tilde{UU}\tilde{U}^{-1}), e_1 \rangle = \langle \tilde{U}\varphi_j(U), e_1 \rangle$. Since $U \mapsto \tilde{UU}\tilde{U}^{-1}$ leaves Haar measure invariant, we see that the distribution of the unit vector 
$$(\langle \varphi_1(U), e_1 \rangle, \langle \varphi_2(U), e_1 \rangle, \ldots, \langle \varphi_n(U), e_1 \rangle) \in \mathbb{C}^n$$
is invariant under unitary transformations, which implies it is the Euclidean measure restricted to the sphere. By using the fact that that $d^2z = \frac{1}{2}d\theta d(|z|^2)$ (which shows it is essential we work in $\mathbb{C}$), it is not hard to show that the squares of the components of a complex $n$–vector uniformly distributed on the sphere are uniformly distributed on the simplex. Thus we get that the $\{w_j\}_{j=1}^n$ are independent of the eigenvalues and have $\mathbb{P}_n$-distribution.

$$(n-1)!\chi_{\{\sum_{j=1}^{n-1} w_j \leq 1; w_j \geq 0\}}(w)dw_1 \ldots dw_{n-1}$$
The distribution of the eigenvalues is given by the celebrated Weyl integration formula which says that the distribution of the eigenvalues under Haar measure is
Weyl Integration Formula

The distribution of the eigenvalues is given by the celebrated Weyl integration formula which says that the distribution of the eigenvalues under Haar measure is

\[
\frac{1}{n!} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^2 \prod_{j=1}^{n} \frac{d\theta_j}{2\pi}
\]

\[
\Delta(\lambda_1, \ldots, \lambda_n) \equiv \prod_{i<j}(\lambda_i - \lambda_j)
\]
The distribution of the eigenvalues is given by the celebrated Weyl integration formula which says that the distribution of the eigenvalues under Haar measure is

$$\frac{1}{n!} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^2 \prod_{j=1}^{n} \frac{d\theta_j}{2\pi}$$

$$\Delta(\lambda_1, \ldots, \lambda_n) \equiv \prod_{i<j}(\lambda_i - \lambda_j)$$

For proofs of this formula from two different points of view, see Anderson et al Random Matrices book or my group representation book.
The distribution of the eigenvalues is given by the celebrated Weyl integration formula which says that the distribution of the eigenvalues under Haar measure is

\[ \frac{1}{n!} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^2 \prod_{j=1}^{n} \frac{d\theta_j}{2\pi} \]

\[ \Delta(\lambda_1, \ldots, \lambda_n) \equiv \prod_{i<j}(\lambda_i - \lambda_j) \]

For proofs of this formula from two different points of view, see Anderson et al Random Matrices book or my group representation book. Summarizing
The distribution of the eigenvalues is given by the celebrated Weyl integration formula which says that the distribution of the eigenvalues under Haar measure is

\[
\frac{1}{n!} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^2 \prod_{j=1}^{n} \frac{d\theta_j}{2\pi}
\]

\[
\Delta(\lambda_1, \ldots, \lambda_n) \equiv \prod_{i<j}(\lambda_i - \lambda_j)
\]

For proofs of this formula from two different points of view, see Anderson et al Random Matrices book or my group representation book. Summarizing

\[
d\mathbb{P}_n(\theta_1, \ldots, \theta_n, w_1, \ldots, w_n) = \frac{1}{n(2\pi)^n} \chi_{\{\sum_{j=1}^{n-1} w_j \leq 1; w_j \geq 0\}}(w)
\]

\[
|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n}|^2 d\theta_1 \ldots d\theta_n dw_1 \ldots dw_{n-1}
\]
LDP for the Empirical Measure

As a preliminary to computing the measure side rate, one needs to look at what spectral theorists call the density of states, OP workers the density of zeroes and probabilists the empirical measure, namely

\[ \mu(E) = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j} \]

where \( \lambda_j \) are the atoms of \( \mu \).

That is, we drop the weights from the spectral measure. \( P_n \) induces a distribution \( P_n(E) \) on point measures of the above form, essentially given by the Weyl Integration Formula.
LDP for the Empirical Measure

As a preliminary to computing the measure side rate, one needs to look at what spectral theorists call the density of states, OP workers the density of zeroes and probabilists the empirical measure, namely

\[ \mu^{(E)} = \frac{1}{n} \sum_{j=1}^{n} \delta \lambda_j \]
LDP for the Empirical Measure

As a preliminary to computing the measure side rate, one needs to look at what spectral theorists call the density of states, OP workers the density of zeroes and probabilists the empirical measure, namely

\[ \mu^{(E)} = \frac{1}{n} \sum_{j=1}^{n} \delta \lambda_j \]

where \( \lambda_j \) are the atoms of \( \mu \).
LDP for the Empirical Measure

As a preliminary to computing the measure side rate, one needs to look at what spectral theorists call the density of states, OP workers the density of zeroes and probabilists the empirical measure, namely

$$\mu^{(E)} = \frac{1}{n} \sum_{j=1}^{n} \delta \lambda_j$$

where $\lambda_j$ are the atoms of $\mu$. That is, we drop the weights from the spectral measure.
As a preliminary to computing the measure side rate, one needs to look at what spectral theorists call the density of states, OP workers the density of zeroes and probabilists the empirical measure, namely

\[ \mu(E) = \frac{1}{n} \sum_{j=1}^{n} \delta\lambda_j \]

where \( \lambda_j \) are the atoms of \( \mu \). That is, we drop the weights from the spectral measure.

\( \mathbb{P}_n \) induces a distribution \( \mathbb{P}_n(E) \) on point measures of the above form, essentially given by the Weyl Integration Formula.
One has the following result of Ben Arous and Guionnet – their results discuss GUE, not CUE – the analog for CUE uses the same ideas and is even simpler:
One has the following result of Ben Arous and Guionnet – their results discuss GUE, not CUE – the analog for CUE uses the same ideas and is even simpler:

**BAG Theorem** $\mathbb{P}_n^{(E)}$ obeys a LDP with speed $n^2$ and good rate function

\[ I(\mu) = -\int \log(|z-w|) d\mu(z) d\mu(w) \]

Remark. In the formula for $I$, $z$ and $w$ lie in the unit circle and $|z-w|$ is a two dimensional distance. This is a 2D Coulomb energy. There is a close connection between this result and Johansson’s proof of the Strong Szegő Theorem.
LDP for the Empirical Measure

One has the following result of Ben Arous and Guionnet – their results discuss GUE, not CUE – the analog for CUE uses the same ideas and is even simpler:

**BAG Theorem** $\mathbb{P}_n^{(E)}$ obeys a LDP with speed $n^2$ and good rate function

$$I(\mu) = -\int \log(|z - w|) \, d\mu(z) \, d\mu(w)$$

Remark. In the formula for $I(\mu)$, $z$ and $w$ lie in the unit circle and $|z - w|$ is a two dimensional distance. This is a 2D Coulomb energy. There is a close connection between this result and Johansson’s proof of the Strong Szegő Theorem.
One has the following result of Ben Arous and Guionnet – their results discuss GUE, not CUE – the analog for CUE uses the same ideas and is even simpler:

**BAG Theorem** $\mathbb{P}_n^{(E)}$ obeys a LDP with speed $n^2$ and good rate function

$$I(\mu) = -\int \log(|z - w|) \, d\mu(z) \, d\mu(w)$$

**Remark.** In the formula for $I$, $z$ and $w$ lie in the unit circle and $|z - w|$ is a two dimensional distance. This is a 2D Coulomb energy. There is a close connection between this result and Johansson’s proof of the Strong Szegő Theorem.
We will not give a formal proof of the BAG Theorem but instead indicate the basic intuition.
We will not give a formal proof of the BAG Theorem but instead indicate the basic intuition. For distinct $\lambda_i$s,
We will not give a formal proof of the BAG Theorem but instead indicate the basic intuition. For distinct $\lambda_i$s,

$$\prod_{i<j} |e^{i\theta_i} - e^{i\theta_j}|^2 = \exp\left(-n^2 J_n(\lambda_1, \ldots, \lambda_n)\right)$$

$$J_n(\lambda_1, \ldots, \lambda_n) = -\frac{2}{n^2} \sum_{i<j} \log(|\lambda_i - \lambda_j|)$$

$$= -\frac{1}{n^2} \sum_{i \neq j} \log(|\lambda_i - \lambda_j|)$$
LDP for the Empirical Measure

We will not give a formal proof of the BAG Theorem but instead indicate the basic intuition. For distinct $\lambda_i$s,

$$\prod_{i<j} |e^{i\theta_i} - e^{i\theta_j}|^2 = \exp\left(-n^2 J_n(\lambda_1, \ldots, \lambda_n)\right)$$

$$J_n(\lambda_1, \ldots, \lambda_n) = -\frac{2}{n^2} \sum_{i<j} \log(|\lambda_i - \lambda_j|)$$

$$= -\frac{1}{n^2} \sum_{i\neq j} \log(|\lambda_i - \lambda_j|)$$

If $\mu^E$ is an $n$–point measure near $\mu$ and the $\lambda$ have reasonable local spacing,
LDP for the Empirical Measure

We will not give a formal proof of the BAG Theorem but instead indicate the basic intuition. For distinct $\lambda_i$s,

$$\prod_{i<j} |e^{i\theta_i} - e^{i\theta_j}|^2 = \exp \left(-n^2 J_n(\lambda_1, \ldots, \lambda_n) \right)$$

$$J_n(\lambda_1, \ldots, \lambda_n) = -\frac{2}{n^2} \sum_{i<j} \log(|\lambda_i - \lambda_j|)$$

$$= -\frac{1}{n^2} \sum_{i \neq j} \log(|\lambda_i - \lambda_j|)$$

If $\mu^{(E)}$ is an $n$–point measure near $\mu$ and the $\lambda$ have reasonable local spacing, the final sum, which is a discrete Coulomb energy should be near the integral which gives a continuum Coulomb energy.
Slightly Simplified Problem

The weights and eigenvalues are independent. We’ll consider a **fixed** triangular array of eigenvalues

\[ \{ \lambda^{(n)}_{\ell} \}_{1 \leq \ell \leq n; n=1,...} \]  

where we suppose that
The weights and eigenvalues are independent. We’ll consider a **fixed** triangular array of eigenvalues \( \{\lambda^{(n)}_\ell\} \) \( 1 \leq \ell \leq n; n=1,... \) where we suppose that

\[
\frac{1}{n} \sum_{\ell=1}^{n} \delta \lambda^{(n)}_\ell \to \frac{d\theta}{2\pi}
\]

weakly.
The weights and eigenvalues are independent. We’ll consider a **fixed** triangular array of eigenvalues \( \{\lambda^{(n)}_{\ell}\} \) \( 1 \leq \ell \leq n; \ n = 1,... \) where we suppose that

\[
\frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda^{(n)}_{\ell}} \to \frac{d\theta}{2\pi}
\]

weakly. We distribute weights uniformly on the simplex and look at
Slightly Simplified Problem

The weights and eigenvalues are independent. We’ll consider a fixed triangular array of eigenvalues \( \{\lambda^{(n)}_\ell\}_{1 \leq \ell \leq n; n=1,...} \) where we suppose that

\[
\frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda^{(n)}_\ell} \rightarrow \frac{d\theta}{2\pi}
\]

weakly. We distribute weights uniformly on the simplex and look at

\[
\{w_\ell\}_{\ell=1}^{n} \mapsto \sum_{\ell=1}^{n} w_\ell \delta_{\lambda^{(n)}_\ell} \equiv \mu_n(w_\ell)
\]
The weights and eigenvalues are independent. We’ll consider a fixed triangular array of eigenvalues 
\( \{ \lambda_\ell^{(n)} \}_{1 \leq \ell \leq n; \ n=1,...} \) where we suppose that 
\[
\frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda_\ell^{(n)}} \rightarrow \frac{d\theta}{2\pi}
\]
weakly. We distribute weights uniformly on the simplex and look at 
\[
\{ w_\ell \}_{\ell=1}^{n} \mapsto \sum_{\ell=1}^{n} w_\ell \delta_{\lambda_\ell^{(n)}} \equiv \mu_n(w_\ell)
\]
This gives a distribution, \( \mathbb{P}_n^{(\lambda)} \), on measures
Slightly Simplified Problem

The weights and eigenvalues are independent. We’ll consider a fixed triangular array of eigenvalues
\( \{\lambda^{(n)}_\ell\} \}_{1 \leq \ell \leq n; n=1,...} \) where we suppose that
\[
\frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda^{(n)}_\ell} \to \frac{d\theta}{2\pi}
\]
weakly. We distribute weights uniformly on the simplex and look at
\[
\{w_\ell\}_{\ell=1}^{n} \mapsto \sum_{\ell=1}^{n} w_\ell \delta_{\lambda^{(n)}_\ell} \equiv \mu_n(w_\ell)
\]
This gives a distribution, \( \mathbb{P}_n^{(\lambda)} \), on measures and we’ll prove these measures obey a LDP with speed \( n \) and rate function \( H\left(\frac{d\theta}{2\pi}, \mu\right) \), the KL divergence.
The weights and eigenvalues are independent. We’ll consider a fixed triangular array of eigenvalues
\[ \{ \lambda^{(n)}_{\ell} \}_{1 \leq \ell \leq n; n=1,...} \] where we suppose that
\[ \frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda^{(n)}_{\ell}} \to \frac{d\theta}{2\pi} \]
weakly. We distribute weights uniformly on the simplex and look at
\[ \{ w_{\ell} \}_{\ell=1}^{n} \mapsto \sum_{\ell=1}^{n} w_{\ell} \delta_{\lambda^{(n)}_{\ell}} \equiv \mu_{n}(w_{\ell}) \]
This gives a distribution, \( \mathbb{P}_{n}^{(\lambda)} \), on measures and we’ll prove these measures obey a LDP with speed \( n \) and rate function \( H(\frac{d\theta}{2\pi}, \mu) \), the KL divergence. A full analysis depends on proving for each \( \epsilon > 0, j \) and \( k = 1, \ldots, 2^{j} \), the probability that
\[ \left| \frac{2^{j}}{n} \#(\ell \mid \lambda^{(n)}_{\ell} \in I_{k}^{(j)}) - 1 \right| \geq \epsilon \]
(with
\[ I_{k}^{(j)} \equiv \{ e^{2\pi i \theta} \mid \frac{k-1}{2^{j}} \leq \theta < \frac{k}{2^{j}} \}) \] goes to zero faster than exponentially in \( n \).
Slightly Simplified Problem

The weights and eigenvalues are independent. We’ll consider a **fixed** triangular array of eigenvalues
\[ \{\lambda_{\ell}^{(n)}\}_{1 \leq \ell \leq n; n=1,...} \] where we suppose that
\[ \frac{1}{n} \sum_{\ell=1}^{n} \delta_{\lambda_{\ell}^{(n)}} \rightarrow \frac{d\theta}{2\pi} \]
weakly. We distribute weights uniformly on the simplex and look at
\[ \{w_{\ell}\}_{\ell=1}^{n} \mapsto \sum_{\ell=1}^{n} w_{\ell} \delta_{\lambda_{\ell}^{(n)}} \equiv \mu_{n}(w_{\ell}) \]
This gives a distribution, \( P_{n}(\lambda) \), on measures and we’ll prove these measures obey a LDP with speed \( n \) and rate function \( H(\frac{d\theta}{2\pi}, \mu) \), the KL divergence. A full analysis depends on proving for each \( \epsilon > 0, j \) and \( k = 1, \ldots, 2^j \), the probability that
\[ \left| \frac{2^j}{n} \#(\ell \mid \lambda_{\ell}^{(n)} \in I_{k}^{(j)}) - 1 \right| \geq \epsilon \] (with
\[ I_{k}^{(j)} \equiv \{e^{2\pi i \theta} \mid \frac{k-1}{2^j} \leq \theta < \frac{k}{2^j}\} \]) goes to zero faster than exponentially in \( n \). This depends on the BAG Theorem.
The proof will be to use projective limits with the maps
\[ \pi_j : \mathcal{M}_{+1}(\partial \mathbb{D}) \to \mathbb{R}^{2^j} \text{ given by } \mu \mapsto \mu(I_{k}^{(j)}). \]
The proof will be to use projective limits with the maps $\pi_j : \mathcal{M}_{+,1}(\partial \mathbb{D}) \to \mathbb{R}^{2^j}$ given by $\mu \mapsto \mu(I_k^{(j)})$. We’ll get a LDP for the projections using our LDP for sums of exponential random variables and control the sup of the projected rate functions by a general continuity result.
The proof will be to use projective limits with the maps \( \pi_j : \mathcal{M}_{+,1}(\partial \mathbb{D}) \to \mathbb{R}^{2^j} \) given by \( \mu \mapsto \mu(I_k^{(j)}) \). We’ll get a LDP for the projections using our LDP for sums of exponential random variables and control the \( \sup \) of the projected rate functions by a general continuity result. It is this last fact that will show singular parts of the measure only change the rate by their impact on the total weight of the a.c. part.
The proof will be to use projective limits with the maps \( \pi_j : \mathcal{M}_{+1}(\partial \mathbb{D}) \to \mathbb{R}^{2^j} \) given by \( \mu \mapsto \mu(I_k^{(j)}) \). We’ll get a LDP for the projections using our LDP for sums of exponential random variables and control the \( \sup \) of the projected rate functions by a general continuity result. It is this last fact that will show singular parts of the measure only change the rate by their impact on the total weight of the a.c. part.

For each \( j = 1, \ldots \) and \( k = 1, \ldots, 2^j \), let \( I_k^{(j)} \) be given as above and \( \pi_j(\mu) \) the measure with constant a.c. weight on each \( I_k^{(j)} \) which gives the same weight to each \( I_k^{(j)} \) as \( \mu \).
LDP for Projected Haar Distribution

The proof will be to use projective limits with the maps
\[ \pi_j : \mathcal{M}_{+1}(\partial \mathbb{D}) \to \mathbb{R}^{2^j} \] given by \( \mu \mapsto \mu(I^{(j)}_k) \). We’ll get a LDP for the projections using our LDP for sums of exponential random variables and control the \( \sup \) of the projected rate functions by a general continuity result. It is this last fact that will show singular parts of the measure only change the rate by their impact on the total weight of the a.c. part.

For each \( j = 1, \ldots \) and \( k = 1, \ldots, 2^j \), let \( I^{(j)}_k \) be given as above and \( \pi_j(\mu) \) the measure with constant a.c. weight on each \( I^{(j)}_k \) which gives the same weight to each \( I^{(j)}_k \) as \( \mu \). This is exactly the setup we described in Lecture 3 for an example of projective limits.
LDP for Projected Haar Distribution

Given \( \{w_\ell\}_{\ell=1}^n \), let \( \tilde{\mu}_n^j(w_\ell) \) be the measure on \( \partial \mathbb{D} \) with constant a.c. weight on each \( I_{k}^{(j)} \) so that
Given \( \{w_\ell\}_{\ell=1}^n \), let \( \tilde{\mu}_n^j(w_\ell) \) be the measure on \( \partial \mathbb{D} \) with constant a.c. weight on each \( I_{k}^{(j)} \) so that

\[
\tilde{\mu}_n^j(I_{k}^{(j)}) = \sum_{\lambda_{\ell}^{(n)} \in I_{k}^{(j)}} w_\ell
\]

Thus we have that \( \pi_j(\mu_n(w_\ell)) = \tilde{\mu}_n^j(w_\ell) \).

The \( w_j \) are almost independent except for the bothersome normalization condition. We will deal this by noting that if \( \{W_j\}_{n=1}^j \) are iidrv with exponential distribution, then \( w_j = \frac{W_j}{\sum_{k=1}^n W_k} \) are distributed uniformly on a simplex.
LDP for Projected Haar Distribution

Given \( \{w_\ell\}_{\ell=1}^n \), let \( \tilde{\mu}_n^j(w_\ell) \) be the measure on \( \partial \mathbb{D} \) with constant a.c. weight on each \( I_{k}^{(j)} \) so that

\[
\tilde{\mu}_n^j(I_{k}^{(j)}) = \sum_{\lambda^{(n)}_\ell \in I_{k}^{(j)}} w_\ell
\]

Thus we have that \( \pi_j(\mu_n(w_\ell)) = \tilde{\mu}_n^j(w_\ell) \).
Given \( \{w_{\ell}\}_{\ell=1}^{n} \), let \( \tilde{\mu}_{n}^{j}(w_{\ell}) \) be the measure on \( \partial \mathbb{D} \) with constant a.c. weight on each \( I_{k}^{(j)} \) so that

\[
\tilde{\mu}_{n}^{j}(I_{k}^{(j)}) = \sum_{\lambda_{\ell}^{(n)} \in I_{k}^{(j)}} w_{\ell}
\]

Thus we have that \( \pi_{j}(\mu_{n}(w_{\ell})) = \tilde{\mu}_{n}^{j}(w_{\ell}) \). The \( w_{j} \) are almost independent except for the bothersome normalization condition.
Given \( \{w_\ell\}_{\ell=1}^n \), let \( \tilde{\mu}_n^j(w_\ell) \) be the measure on \( \partial \mathbb{D} \) with constant a.c. weight on each \( I_k^{(j)} \) so that
\[
\tilde{\mu}_n^j(I_k^{(j)}) = \sum_{\lambda_\ell^{(n)} \in I_k^{(j)}} w_\ell
\]
Thus we have that \( \pi_j(\mu_n(w_\ell)) = \tilde{\mu}_n^j(w_\ell) \). The \( w_j \) are almost independent except for the bothersome normalization condition. We will deal this by noting that if \( \{W_j\}_{j=1}^n \) are iidrv with exponential distribution, then \( w_j = W_j / \sum_{k=1}^n W_k \) are distributed uniformly on a simplex.
We will be able to prove a LDP for subsums of $W$’s and then use the contraction principle to pass to $w$’s.
LDP for Projected Haar Distribution

We will be able to prove a LDP for subsums of $W$’s and then use the contraction principle to pass to $w$’s.

So let $\widetilde{P}_n(j)$ be the measure on $\mathbb{R}^{2^j}$ but where now the $w_\ell$ are replaced by iid exponential random variables, $W_\ell$. Thus, $\widetilde{P}_n(j)$ is the probability measure for the $\mathbb{R}^{2^j}$-valued random variable given by
We will be able to prove a LDP for subsums of $W$’s and then use the contraction principle to pass to $w$’s.

So let $\widetilde{P}_n(j)$ be the measure on $\mathbb{R}^{2^j}$ but where now the $w_\ell$ are replaced by iid exponential random variables, $W_\ell$. Thus, $\widetilde{P}_n(j)$ is the probability measure for the $\mathbb{R}^{2^j}$-valued random variable given by

$$\beta_k^n = \sum_{\chi^{(n)}_\ell \in I^{(j)}_k} W_\ell$$
LDP for Projected Haar Distribution

We will be able to prove a LDP for subsums of $W$’s and then use the contraction principle to pass to $w$’s.

So let $\widetilde{P}_n^{(j)}$ be the measure on $\mathbb{R}^{2^j}$ but where now the $w_\ell$ are replaced by iid exponential random variables, $W_\ell$. Thus, $\widetilde{P}_n^{(j)}$ is the probability measure for the $\mathbb{R}^{2^j}$-valued random variable given by

$$\beta_k^n = \sum_{\chi^{(n)}_\ell \in I_k^{(j)}} W_\ell$$

Fix $j$ and take $n \to \infty$. By our analysis of sums of exponential iidrvs, $\widetilde{P}_n^{(j)}$ obeys a LDP with speed $n$ and rate function at the point $\vec{\beta} \equiv \{\beta_\ell\}_{\ell=1}^{2^j} \in \mathbb{R}^{2^j}$
We will be able to prove a LDP for subsums of $W$’s and then use the contraction principle to pass to $w$’s.

So let $\tilde{P}^{(j)}_n$ be the measure on $\mathbb{R}^{2^j}$ but where now the $w_\ell$ are replaced by iid exponential random variables, $W_\ell$. Thus, $\tilde{P}^{(j)}_n$ is the probability measure for the $\mathbb{R}^{2^j}$-valued random variable given by

$$\beta^n_k = \sum_{\chi^{(n)}_\ell \in I^{(j)}_k} W_\ell$$

Fix $j$ and take $n \to \infty$. By our analysis of sums of exponential iidrvs, $\tilde{P}^{(j)}_n$ obeys a LDP with speed $n$ and rate function at the point $\vec{\beta} \equiv \{\beta_\ell\}^{2^j}_{\ell=1} \in \mathbb{R}^{2^j}$

$$\varphi(\vec{\beta}) = \sum_{\ell=1}^{2^j} \left[ (\beta_\ell - 2^{-j}) - 2^{-j} \log(2^j \beta_\ell) \right]$$
Recall that given two probability measures $\mu$ and $\nu$ on the same space, their KL divergence, $H(\mu|\nu)$, is given by the negative of a log integral.
Recall that given two probability measures $\mu$ and $\nu$ on the same space, their KL divergence, $H(\mu|\nu)$, is given by the negative of a log integral. Write $\beta_\ell = \beta s_\ell$ with $\beta = \sum_{q=1}^{2^j} \beta_q$ so that $\vec{s}$ lies in a $2^j$-simplex. Write $\mu_{\vec{s}}$ for the probability measure giving uniform weight $s_k$ to $I_k^{(j)}$ and let $\nu$ be normalized Lebesgue measure on the circle (i.e. $\mu_{\vec{s}}$ for the $\vec{s}$ with equal components, $2^{-j}$). Then $\varphi$ can be rewritten:
Recall that given two probability measures $\mu$ and $\nu$ on the same space, their KL divergence, $H(\mu|\nu)$, is given by the negative of a log integral. Write $\beta_\ell = \beta s_\ell$ with $\beta = \sum_{q=1}^{2^j} \beta_q$ so that $\vec{s}$ lies in a $2^j$-simplex. Write $\mu_{\vec{s}}$ for the probability measure giving uniform weight $s_k$ to $I_k^{(j)}$ and let $\nu$ be normalized Lebesgue measure on the circle (i.e. $\mu_{\vec{s}}$ for the $\vec{s}$ with equal components, $2^{-j}$). Then $\varphi$ can be rewritten:

$$\varphi(\vec{\beta}) = \beta - 1 - \log(\beta) + H(\nu|\mu_{\vec{s}})$$
LDP for Projected Haar Distribution

Recall that given two probability measures $\mu$ and $\nu$ on the same space, their KL divergence, $H(\mu|\nu)$, is given by the negative of a log integral. Write $\beta_\ell = \beta s_\ell$ with $\beta = \sum_{q=1}^{2^j} \beta_q$ so that $\bar{s}$ lies in a $2^j$-simplex. Write $\mu_{\bar{s}}$ for the probability measure giving uniform weight $s_k$ to $I_{k}^{(j)}$ and let $\nu$ be normalized Lebesgue measure on the circle (i.e. $\mu_{\bar{s}}$ for the $\bar{s}$ with equal components, $2^{-j}$). Then $\varphi$ can be rewritten:

$$\varphi(\vec{\beta}) = \beta - 1 - \log(\beta) + H(\nu|\mu_{\vec{s}})$$

Note this is the sum of a function of $\beta$ only and a function of the $s$’s only.
Recall that given two probability measures $\mu$ and $\nu$ on the same space, their KL divergence, $H(\mu|\nu)$, is given by the negative of a log integral. Write $\beta_\ell = \beta s_\ell$ with $\beta = \sum_{q=1}^{2^j} \beta_q$ so that $\vec{s}$ lies in a $2^j$-simplex. Write $\mu_{\vec{s}}$ for the probability measure giving uniform weight $s_k$ to $I_k^{(j)}$ and let $\nu$ be normalized Lebesgue measure on the circle (i.e. $\mu_{\vec{s}}$ for the $\vec{s}$ with equal components, $2^{-j}$). Then $\varphi$ can be rewritten:

$$\varphi(\vec{\beta}) = \beta - 1 - \log(\beta) + H(\nu|\mu_{\vec{s}})$$

Note this is the sum of a function of $\beta$ only and a function of the $s$'s only. This is a consequence of the fact that for independent exponential random variables, $\sum_{k=1}^{N} X_k$ is independent of $\{X_j/\sum_{k=1}^{N} X_k\}_{j=1}^{N}$. It makes the use of the contraction principle (which, in general, is already simple), extremely simple.
For fixed λ’s, let $P_n^{(j)} = \pi_j^* \left( P_n^{(\lambda)} \right)$. This is just the contraction of $\tilde{P}_n^{(j)}$ under the map $G(\tilde{\beta}) \equiv \tilde{\beta}/\beta$ from $\mathbb{R}^{2^j}$ to the $2^j$–simplex. By the contraction principle and
LDP for Projected Haar Distribution

For fixed $\lambda$’s, let $\mathbb{P}_n^{(j)} = \pi_j^* \left( \mathbb{P}_n^{(\lambda)} \right)$. This is just the contraction of $\widetilde{\mathbb{P}}_n^{(j)}$ under the map $G(\vec{\beta}) \equiv \vec{\beta}/\beta$ from $\mathbb{R}^{2^j}$ to the $2^j$–simplex. By the contraction principle and

$$\inf_{\beta > 0} [\beta - 1 - \log(\beta)] = 0$$
For fixed $\lambda$’s, let $P_n^{(j)} = \pi^*_j \left( P_n^{(\lambda)} \right)$. This is just the contraction of $\tilde{P}_n^{(j)}$ under the map $G(\vec{\beta}) \equiv \vec{\beta} / \beta$ from $\mathbb{R}^{2j}$ to the $2^j$–simplex. By the contraction principle and
\[
\inf_{\beta > 0} [\beta - 1 - \log(\beta)] = 0
\]
(as it must as the rate function, for averages of exponentials),
For fixed $\lambda$'s, let $\mathbb{P}_n^{(j)} = \pi^* \left( \mathbb{P}_n^{(\lambda)} \right)$. This is just the contraction of $\tilde{\mathbb{P}}_n^{(j)}$ under the map $G(\beta) \equiv \beta / \beta$ from $\mathbb{R}^{2j}$ to the $2^j$–simplex. By the contraction principle and
\[ \inf_{\beta > 0} [\beta - 1 - \log(\beta)] = 0 \]
(as it must as the rate function, for averages of exponentials), we see that for each fixed $j$, $\mathbb{P}_n^{(j)}$ obeys a LDP with speed $n$ and rate function $H(\nu | \mu_S)$. Given the projection theorem, the following completes the proof that the measure theory rate function is $H(\nu | \mu)$. 
For fixed $\lambda$’s, let $P^{(j)}_n = \pi^*_j \left( P^{(\lambda)}_n \right)$. This is just the contraction of $\tilde{P}_n^{(j)}$ under the map $G(\vec{\beta}) \equiv \vec{\beta}/\beta$ from $\mathbb{R}^{2^j}$ to the $2^j$–simplex. By the contraction principle and

$$\inf_{\beta > 0} [\beta - 1 - \log(\beta)] = 0$$

(as it must as the rate function, for averages of exponentials), we see that for each fixed $j$, $P_n^{(j)}$ obeys a LDP with speed $n$ and rate function $H(\nu | \mu_\Sigma)$. Given the projection theorem, the following completes the proof that the measure theory rate function is $H(\nu | \mu)$.

**Key Fact.** Let $\mu$ be an arbitrary probability measure on $\partial \mathbb{D}$ and $\nu = \frac{d\theta}{2\pi}$. Then
LDP for Projected Haar Distribution

For fixed $\lambda$’s, let $\mathbb{P}_n^{(j)} = \pi_j^* \left( \mathbb{P}_n^{(\lambda)} \right)$. This is just the contraction of $\tilde{\mathbb{P}}_n^{(j)}$ under the map $G(\vec{\beta}) \equiv \vec{\beta}/\beta$ from $\mathbb{R}^{2^j}$ to the $2^j$–simplex. By the contraction principle and

$$\inf_{\beta > 0} [\beta - 1 - \log(\beta)] = 0$$

(as it must as the rate function, for averages of exponentials), we see that for each fixed $j$, $\mathbb{P}_n^{(j)}$ obeys a LDP with speed $n$ and rate function $H(\nu|\mu_\mathcal{S})$.

Given the projection theorem, the following completes the proof that the measure theory rate function is $H(\nu|\mu)$.

**Key Fact.** Let $\mu$ be an arbitrary probability measure on $\partial \mathbb{D}$ and $\nu = \frac{d\theta}{2\pi}$. Then

$$\lim_{k \to \infty} H(\pi_j(\nu)|\pi_j(\mu)) = H(\nu|\mu)$$
Before turning to the proof of the Key Fact, a quick remark:
\[ \pi_j(\nu) = \nu \text{ for this } \nu. \]
Before turning to the proof of the Key Fact, a quick remark: $\pi_j(\nu) = \nu$ for this $\nu$. We write it this way because with a slight change in the proof, it holds for any $\nu$ (and $\mu$).
Before turning to the proof of the Key Fact, a quick remark: \( \pi_j(\nu) = \nu \) for this \( \nu \). We write it this way because with a slight change in the proof, it holds for any \( \nu \) (and \( \mu \)). This extended version is needed for the Killip–Simon theorem and other cases where the limiting empirical measure is not unweighted Lebesgue measure.
Before turning to the proof of the Key Fact, a quick remark: $\pi_j(\nu) = \nu$ for this $\nu$. We write it this way because with a slight change in the proof, it holds for any $\nu$ (and $\mu$). This extended version is needed for the Killip–Simon theorem and other cases where the limiting empirical measure is not unweighted Lebesgue measure.

We’ll prove the limit result in two parts.
Before turning to the proof of the Key Fact, a quick remark: $\pi_j(\nu) = \nu$ for this $\nu$. We write it this way because with a slight change in the proof, it holds for any $\nu$ (and $\mu$). This extended version is needed for the Killip–Simon theorem and other cases where the limiting empirical measure is not unweighted Lebesgue measure.

We’ll prove the limit result in two parts. We’ll prove a general upper bound: $H(\pi_j(\nu)|\pi_j(\mu)) \leq H(\nu|\mu)$. 
Before turning to the proof of the Key Fact, a quick remark: \( \pi_j(\nu) = \nu \) for this \( \nu \). We write it this way because with a slight change in the proof, it holds for any \( \nu \) (and \( \mu \)). This extended version is needed for the Killip–Simon theorem and other cases where the limiting empirical measure is not unweighted Lebesgue measure.

We’ll prove the limit result in two parts. We’ll prove a general upper bound: \( H(\pi_j(\nu)|\pi_j(\mu)) \leq H(\nu|\mu) \). (By slightly expanding the argument, one sees that \( H(\pi_j(\nu)|\pi_j(\mu)) \) is monotone increasing in \( j \).)
Before turning to the proof of the Key Fact, a quick remark: $\pi_j(\nu) = \nu$ for this $\nu$. We write it this way because with a slight change in the proof, it holds for any $\nu$ (and $\mu$). This extended version is needed for the Killip–Simon theorem and other cases where the limiting empirical measure is not unweighted Lebesgue measure.

We’ll prove the limit result in two parts. We’ll prove a general upper bound: $H(\pi_j(\nu)|\pi_j(\mu)) \leq H(\nu|\mu)$. (By slightly expanding the argument, one sees that $H(\pi_j(\nu)|\pi_j(\mu))$ is monotone increasing in $j$.)

The other direction – that $H(\nu|\mu) \leq \lim\inf H(\pi_j(\nu)|\pi_j(\mu))$ comes from weak convergence, $\lim \pi_j(\eta) = \eta$ (for any probability measure $\eta$) and the lower semi–continuity.
To get the upper bound, note that by convexity of $y \mapsto -\log y$ and Jensen’s inequality, for any positive function $h$ and probability measure $d\eta(y)$, we have that
Limit Theorem for KL Divergences

To get the upper bound, note that by convexity of $y \mapsto -\log y$ and Jensen’s inequality, for any positive function $h$ and probability measure $d\eta(y)$, we have that

$$-\int \log h(y) \, d\eta(y) \geq -\log \left( \int h(y) \, d\eta(y) \right)$$
Limit Theorem for KL Divergences

To get the upper bound, note that by convexity of $y \mapsto -\log y$ and Jensen’s inequality, for any positive function $h$ and probability measure $d\eta(y)$, we have that

$$-\int \log h(y) \, d\eta(y) \geq -\log \left( \int h(y) \, d\eta(y) \right)$$

In just the same way that this implies that $H(\nu|\mu) \geq 0$, it implies that
To get the upper bound, note that by convexity of $y \mapsto -\log y$ and Jensen’s inequality, for any positive function $h$ and probability measure $d\eta(y)$, we have that

$$- \int \log h(y) \, d\eta(y) \geq - \log \left( \int h(y) \, d\eta(y) \right)$$

In just the same way that this implies that $H(\nu|\mu) \geq 0$, it implies that

$$- \int_{I^{(j)}_k} \log(w(\theta)) \, 2^j \frac{d\theta}{2\pi} \geq - \log \left( 2^j \mu(I^{(j)}_k) \right)$$
Limit Theorem for KL Divergences

To get the upper bound, note that by convexity of $y \mapsto -\log y$ and Jensen’s inequality, for any positive function $h$ and probability measure $d\eta(y)$, we have that

$$-\int \log h(y) \, d\eta(y) \geq -\log \left( \int h(y) \, d\eta(y) \right)$$

In just the same way that this implies that $H(\nu|\mu) \geq 0$, it implies that

$$-\int_{I_k^{(j)}} \log(w(\theta)) \frac{2^j \, d\theta}{2\pi} \geq -\log \left( 2^j \mu(I_k^{(j)}) \right)$$

Summing this yields the upper bound.
Killip Simon via LDP

The large deviation proof of the Killip–Simon sum rule is similar to the one I just presented for Szegő–Verblunsky sum rule with some changes and additions which we briefly describe.
The large deviation proof of the Killip–Simon sum rule is similar to the one I just presented for Szegő–Verblunsky sum rule with some changes and additions which we briefly describe.

1 One uses GUE instead of CUE. Thus the measure on random $n \times n$ self–adjoint matrices has

\[
\{\text{Re} M_{ij}^{(n)} \}_{1 \leq i \leq j \leq n} \quad \text{and} \quad \{\text{Im} M_{ij}^{(n)} \}_{1 \leq i < j \leq n}
\]

Gaussian iid with mean zero and $\mathbb{E}( [M_{ii}^{(n)}]^2 ) = n^{-1}$.
The eigenvalue distribution has $\lambda_j \in \mathbb{R}$ with distribution

$$\left[ \prod_{i<j} |\lambda_i - \lambda_j|^2 \right] e^{-n \sum_{j=1}^{n} \lambda_j^2}$$

(4.1)
The eigenvalue distribution has $\lambda_j \in \mathbb{R}$ with distribution

$$\prod_{i<j} |\lambda_i - \lambda_j|^2 e^{-n \sum_{j=1}^{n} \lambda_j^2} \quad (4.1)$$

so the empirical measure converges to the equilibrium measure in a quadratic external field, i.e. the minimizer for $-\int \log |x - y| \, d\mu(x) \, d\mu(y) + 2 \int x^2 \, d\mu(x)$. 
The eigenvalue distribution has $\lambda_j \in \mathbb{R}$ with distribution

$$
\left[ \prod_{i<j} |\lambda_i - \lambda_j|^2 \right] e^{-n \sum_{j=1}^{n} \lambda_j^2}
$$

so the empirical measure converges to the equilibrium measure in a quadratic external field, i.e. the minimizer for $-\int \log |x-y| d\mu(x) d\mu(y) + 2 \int x^2 d\mu(x)$. It is well–known that this minimizer is the semicircle law $d\nu_0(x) \equiv \pi^{-1} (1-x^2)^{1/2} \chi_{[-1,1]}(x) dx$. 

\[ (4.1) \]
The eigenvalue distribution has $\lambda_j \in \mathbb{R}$ with distribution

$$\prod_{i<j} |\lambda_i - \lambda_j|^2 e^{-n \sum_{j=1}^{n} \lambda_j^2}$$

(4.1)

so the empirical measure converges to the equilibrium measure in a quadratic external field, i.e. the minimizer for $-\int \log |x - y| \, d\mu(x) \, d\mu(y) + 2 \int x^2 \, d\mu(x)$. It is well–known that this minimizer is the semicircle law $d\nu_0(x) \equiv \pi^{-1} (1 - x^2)^{1/2} \chi_{[-1,1]}(x) \, dx$. To agree with the Killip–Simon notation, one rescales the matrix so the support is $[-2, 2]$. 
The empirical measure converges to $\nu_0$. 
The empirical measure converges to $\nu_0$. By mimicking the argument above, the contribution of the part of the spectral measure on $[-2, 2]$ is just $H(\nu_0|\mu)$. Thus the weight in the Killip–Simon quasi–Szegő integral is exactly the Wigner semicircle weight.
As we’ve seen, a single point in the measure, if the point is in the bulk, involves the increase of $H(\nu|\mu)$ due to the weight having a smaller integral.
As we’ve seen, a single point in the measure, if the point is in the bulk, involves the increase of $H(\nu|\mu)$ due to the weight having a smaller integral. But if the point is outside $[-2, 2]$, there is a contribution due to the location, $\lambda_0$, of the eigenvalue.
As we’ve seen, a single point in the measure, if the point is in the bulk, involves the increase of $H(\nu|\mu)$ due to the weight having a smaller integral. But if the point is outside $[-2, 2]$, there is a contribution due to the location, $\lambda_0$, of the eigenvalue. By looking at the log of the part of the weight depending on $\lambda_0$, one sees that the decrease in the eigenvalue density involves $\lambda_0$ interacting with $n$ eigenvalues.
As we’ve seen, a single point in the measure, if the point is in the bulk, involves the increase of $H(\nu|\mu)$ due to the weight having a smaller integral. But if the point is outside $[-2, 2]$, there is a contribution due to the location, $\lambda_0$, of the eigenvalue. By looking at the log of the part of the weight depending on $\lambda_0$, one sees that the decrease in the eigenvalue density involves $\lambda_0$ interacting with $n$ eigenvalues. The decrease is approximately $\exp(-nF(\lambda_0))$ where $F$ is the potential in the quadratic external field in the equilibrium measure (this idea is due to Ben Arous, Dembo and Guionnet).
As we’ve seen, a single point in the measure, if the point is in the bulk, involves the increase of \( H(\nu|\mu) \) due to the weight having a smaller integral. But if the point is outside \([-2, 2]\), there is a contribution due to the location, \( \lambda_0 \), of the eigenvalue. By looking at the log of the part of the weight depending on \( \lambda_0 \), one sees that the decrease in the eigenvalue density involves \( \lambda_0 \) interacting with \( n \) eigenvalues. The decrease is approximately \( \exp(-nF(\lambda_0)) \) where \( F \) is the potential in the quadratic external field in the equilibrium measure (this idea is due to Ben Arous, Dembo and Guionnet). It is known that this function is the same as the Killip–Simon \( F \).
For finitely many eigenvalues outside \([-2, 2]\) you just get the sums of single costs since the interaction between eigenvalues is \(O(1)\), not \(O(n)\).
For finitely many eigenvalues outside $[-2, 2]$ you just get the sums of single costs since the interaction between eigenvalues is $O(1)$, not $O(n)$. Handling infinitely many eigenvalues converging to $\pm 2$ requires a careful use of projective limits.
For finitely many eigenvalues outside $[-2, 2]$ you just get the sums of single costs since the interaction between eigenvalues is $O(1)$, not $O(n)$. Handling infinitely many eigenvalues converging to $\pm 2$ requires a careful use of projective limits.

For the coefficient side, Killip–Nenciu is replaced by earlier results of Dumitriu–Edelman (whose work motivated Killip and Nenciu) who found the distribution of Jacobi parameters for GUE and GOE.


For finitely many eigenvalues outside \([-2, 2]\) you just get the sums of single costs since the interaction between eigenvalues is \(O(1)\), not \(O(n)\). Handling infinitely many eigenvalues converging to \(\pm 2\) requires a careful use of projective limits.

For the coefficient side, Killip–Nenciu is replaced by earlier results of Dumitriu–Edelman (whose work motivated Killip and Nenciu) who found the distribution of Jacobi parameters for GUE and GOE. The \(\{b_j\}_{j=1}^n\) are Gaussian (with \(O(n)\) widths leading to the \(b_j^2\) term in the Killip–Simon sum rule).
For finitely many eigenvalues outside $[-2, 2]$ you just get the sums of single costs since the interaction between eigenvalues is $O(1)$, not $O(n)$. Handling infinitely many eigenvalues converging to $\pm 2$ requires a careful use of projective limits.

For the coefficient side, Killip–Nenciu is replaced by earlier results of Dumitriu–Edelman (whose work motivated Killip and Nenciu) who found the distribution of Jacobi parameters for GUE and GOE. The $\{b_j\}_{j=1}^n$ are Gaussian (with $O(n)$ widths leading to the $b_j^2$ term in the Killip–Simon sum rule). The $\{a_j^2\}_{j=1}^{n-1}$ are gamma distributed, essentially behaving like sums of exponential random variables and so we get the $G(a_j)$ terms.
For finitely many eigenvalues outside $[-2, 2]$ you just get the sums of single costs since the interaction between eigenvalues is $O(1)$, not $O(n)$. Handling infinitely many eigenvalues converging to ±2 requires a careful use of projective limits.

For the coefficient side, Killip–Nenciu is replaced by earlier results of Dumitriu–Edelman (whose work motivated Killip and Nenciu) who found the distribution of Jacobi parameters for GUE and GOE. The $\{b_j\}_{j=1}^n$ are Gaussian (with $O(n)$ widths leading to the $b_j^2$ term in the Killip–Simon sum rule). The $\{a_j^2\}_{j=1}^{n-1}$ are gamma distributed, essentially behaving like sums of exponential random variables and so we get the $G(a_j)$ terms. Thus $G$ occurs in the sum rule as the rate function for suitable gamma distributions.
There is a technical issue involving the equality of the two sides of the sum rule that we want to discuss, addressed in related ways by Gamboa-Rouault and by BSZ.
There is a technical issue involving the equality of the two sides of the sum rule that we want to discuss, addressed in related ways by Gamboa-Rouault and by BSZ. The natural setting for the LDP for measures is the space, \( X' \), of all probability measures on \( \mathbb{R} \), and for Jacobi parameters the Polish space \( Y' \equiv [\mathbb{R} \times (0, \infty)]^\infty \) with finite sequences added to it.
There is a technical issue involving the equality of the two sides of the sum rule that we want to discuss, addressed in related ways by Gamboa-Rouault and by BSZ. The natural setting for the LDP for measures is the space, $X'$, of all probability measures on $\mathbb{R}$, and for Jacobi parameters the Polish space $Y' \equiv [\mathbb{R} \times (0, \infty)]^\infty$ with finite sequences added to it. The issue is that the inverse Jacobi map isn’t defined for all measures but only those with all moments finite and, in general, this inverse map is many–to–one in certain cases where the measure has unbounded support.
There is a technical issue involving the equality of the two sides of the sum rule that we want to discuss, addressed in related ways by Gamboa-Rouault and by BSZ. The natural setting for the LDP for measures is the space, $X'$, of all probability measures on $\mathbb{R}$, and for Jacobi parameters the Polish space $Y' \equiv [\mathbb{R} \times (0, \infty)]^\infty$ with finite sequences added to it. The issue is that the inverse Jacobi map isn’t defined for all measures but only those with all moments finite and, in general, this inverse map is many–to–one in certain cases where the measure has unbounded support. BSZ handle this by restricting to measures supported in $[-k - 2, k + 2]$
There is a technical issue involving the equality of the two sides of the sum rule that we want to discuss, addressed in related ways by Gamboa-Rouault and by BSZ. The natural setting for the LDP for measures is the space, $X'$, of all probability measures on $\mathbb{R}$, and for Jacobi parameters the Polish space $Y' \equiv [\mathbb{R} \times (0, \infty)]^\infty$ with finite sequences added to it. The issue is that the inverse Jacobi map isn’t defined for all measures but only those with all moments finite and, in general, this inverse map is many–to–one in certain cases where the measure has unbounded support. BSZ handle this by restricting to measures supported in $[-k - 2, k + 2]$ and its image under the Jacobi map.
Mysteries Solved

We can now solve the mysteries:

1. Why are there any positive combinations?
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.
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?**
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;
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.
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?
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.
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**

   \[
   F(E) = \frac{1}{4} [\beta^2 - \beta^{-2} - \log \beta^4]; \quad E = \beta + \beta^{-1}
   \]

   **mean?**
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**

   \[
   F(E) = \frac{1}{4} [\beta^2 - \beta^{-2} - \log \beta^4]; \quad E = \beta + \beta^{-1}
   \]

   **mean?** This is the Coulomb potential of the Wigner semi–circle distribution plus a quadratic external field.
Higher Order Sum Rules

In OPUC1, I found a sum rule involving
\[ -\int (1 - \cos(\theta)) \log(w(\theta)) \, \frac{d\theta}{2\pi} \] on the measure side and made a conjecture concerning

\[ \int \log(w(\theta)) \, d\eta(\theta) \] where
\[ d\eta(\theta) = \frac{Z}{\prod_{j=1}^{\infty} (1 - \cos(\theta - \theta_j))^m_j} \, d\theta \] where \( Z \) is a normalization factor to make \( d\eta \) into a probability measure.

There developed a huge literature on these so-called higher order sum rules for OPUC and OPRL including papers by Denissov, Golinskii, Kupin, Laptev et al., Lukic and Nazarov et al.
In OPUC1, I found a sum rule involving
\[- \int (1 - \cos(\theta)) \log(w(\theta)) \frac{d\theta}{2\pi} \] on the measure side and made a conjecture concerning
\[- \int \log(w(\theta)) \, d\eta(\theta) \]
Higher Order Sum Rules

In OPUC1, I found a sum rule involving
\[- \int (1 - \cos(\theta)) \log(w(\theta)) \frac{d\theta}{2\pi}\]
on the measure side and made a conjecture concerning
\[- \int \log(w(\theta)) \, d\eta(\theta)\]

where
\[d\eta(\theta) = Z^{-1} \prod_{j=1}^{k} (1 - \cos(\theta - \theta_j))^{m_j} \, d\theta\]
Higher Order Sum Rules

In OPUC1, I found a sum rule involving
\[- \int (1 - \cos(\theta)) \log(w(\theta)) \frac{d\theta}{2\pi}\] on the measure side and made a conjecture concerning
\[- \int \log(w(\theta)) \, d\eta(\theta)\]

where
\[d\eta(\theta) = Z^{-1} \prod_{j=1}^{k} (1 - \cos(\theta - \theta_j))^{m_j} d\theta\]

where $Z$ is a normalization factor to make $d\eta$ into a probability measure.
Higher Order Sum Rules

In OPUC1, I found a sum rule involving
\[- \int (1 - \cos(\theta)) \log(w(\theta)) \frac{d\theta}{2\pi}\] on the measure side and made a conjecture concerning

\[- \int \log(w(\theta)) \, d\eta(\theta)\]

where

\[d\eta(\theta) = Z^{-1} \prod_{j=1}^{k} (1 - \cos(\theta - \theta_j))^{m_j} \, d\theta\]

where \(Z\) is a normalization factor to make \(d\eta\) into a probability measure. There developed a huge literature on these so called higher order sum rules for OPUC and OPRL including papers by Denissov, Golinskii, Kupin, Laptev et al, Lukic and Nazarov et al.
Higher Order Sum Rules

The key to understanding such sum rules (for OPUC) in the context of large deviations is to replace Haar measure, $d\mathbb{P}_N$, by

$$Z^{-1}N\exp\left(-\sum_{j=1}^{N} V(\lambda_j)\right)d\mathbb{P}_N$$

where $V$ is a function on $\partial D$ and $\{\lambda_j\}_{j=1}^{N}$ are the eigenvalues.
Higher Order Sum Rules

The key to understanding such sum rules (for OPUC) in the context of large deviations is to replace Haar measure, $d\mathbb{P}_N$, by

$$Z_N^{-1} \exp \left[ -N \sum_{j=1}^{N} V(\lambda_j) \right] d\mathbb{P}_N$$
Higher Order Sum Rules

The key to understanding such sum rules (for OPUC) in the context of large deviations is to replace Haar measure, $d\mathbb{P}_N$, by

$$Z_N^{-1} \exp \left[ -N \sum_{j=1}^{N} V(\lambda_j) \right] d\mathbb{P}_N$$

where $V$ is a function on $\partial \mathbb{D}$ and $\{\lambda_j\}_{j=1}^{N}$ are the eigenvalues.
Higher Order Sum Rules

The key to understanding such sum rules (for OPUC) in the context of large deviations is to replace Haar measure, $d\mathbb{P}_N$, by

$$Z_N^{-1} \exp \left[ -N \sum_{j=1}^N V(\lambda_j) \right] d\mathbb{P}_N$$

where $V$ is a function on $\partial \mathbb{D}$ and $\{\lambda_j\}_{j=1}^N$ are the eigenvalues. It is well known in the random matrix literature that when $V$ is nice enough, we will get $d\eta$ as the empirical measure if
Higher Order Sum Rules

The key to understanding such sum rules (for OPUC) in the context of large deviations is to replace Haar measure, $d\mathbb{P}_N$, by

$$Z_N^{-1} \exp \left[ -N \sum_{j=1}^{N} V(\lambda_j) \right] d\mathbb{P}_N$$

where $V$ is a function on $\partial \mathbb{D}$ and $\{\lambda_j\}_{j=1}^{N}$ are the eigenvalues. It is well known in the random matrix literature that when $V$ is nice enough, we will get $d\eta$ as the empirical measure if

$$V(e^{i\theta}) = 2 \int \log |e^{i\theta} - e^{i\psi}| d\eta(\psi)$$
In a forthcoming paper BSZ study this when $d\eta$ is given as above.
In a forthcoming paper BSZ study this when $d\eta$ is given as above. In the cases we study, $V(e^{i\theta})$ is a finite linear combination of $\cos(m\theta)$. 
In a forthcoming paper BSZ study this when $d\eta$ is given as above. In the cases we study, $V(e^{i\theta})$ is a finite linear combination of $\cos(m\theta)$. In terms of $U$, if $e^{i\theta_j}$ are the eigenvalues, $\sum_{j=1}^{n} \cos(m\theta_j) = \text{Re}(\text{Tr}(U^m))$ which one can write in terms of Verblunsky coefficients using the CMV (or the GGT) representation of $U$. 
In a forthcoming paper BSZ study this when $d\eta$ is given as above. In the cases we study, $V(e^{i\theta})$ is a finite linear combination of $\cos(m\theta)$. In terms of $U$, if $e^{i\theta_j}$ are the eigenvalues, $\sum_{j=1}^n \cos(m\theta_j) = \text{Re}(\text{Tr}(U^m))$ which one can write in terms of Verblunsky coefficients using the CMV (or the GGT) representation of $U$. We obtain a large deviations proof of the $(1 - \cos(\theta))$ sum rule of Simon and the gems of Simon–Zlatoš.
In a forthcoming paper BSZ study this when $d\eta$ is given as above. In the cases we study, $V(e^{i\theta})$ is a finite linear combination of $\cos(m\theta)$. In terms of $U$, if $e^{i\theta_j}$ are the eigenvalues, $\sum_{j=1}^{n} \cos(m\theta_j) = \text{Re}(\text{Tr}(U^m))$ which one can write in terms of Verblunsky coefficients using the CMV (or the GGT) representation of $U$. We obtain a large deviations proof of the $(1 - \cos(\theta))$ sum rule of Simon and the gems of Simon–Zlatoš. In addition, we prove a partial special case of a conjecture of Lukic that replaces a wrong conjecture of Simon, providing evidence for Lukic’s conjecture.
Higher Order Sum Rules

GNR have a paper that discusses in some detail the case $V(\theta) = \cos(\theta)$ where the random matrix model has been studied by Gross–Witten whose names GNR apply to the model.
GNR have a paper that discusses in some detail the case $V(\theta) = \cos(\theta)$ where the random matrix model has been studied by Gross–Witten whose names GNR apply to the model. They note that formally the large deviations argument leads to a sum rule but for technical reasons, they aren’t able to provide a proof.
GNR have a paper that discusses in some detail the case $V(\theta) = \cos(\theta)$ where the random matrix model has been studied by Gross–Witten whose names GNR apply to the model. They note that formally the large deviations argument leads to a sum rule but for technical reasons, they aren’t able to provide a proof. By using some results from the theory of OPUC, we do prove sum rules in this and the other cases.
There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC.
There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC. In a recent preprint, GNR obtain a sum rule and gem for
\[ e = \{ e^{i\theta} \mid \alpha \leq \theta \leq 2\pi - \alpha \} \text{ for } 0 < \alpha < \pi. \]
Finite Gap OPUC

There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC. In a recent preprint, GNR obtain a sum rule and gem for 
\[ e = \{ e^{i\theta} \mid \alpha \leq \theta \leq 2\pi - \alpha \} \] for \( 0 < \alpha < \pi \). For real \( \alpha \), the Verblunsky side has the expected \( \sum |\alpha_j - a|^2 \) form but for general \( \alpha \), it has the form \( \sum |\gamma_j - a|^2 \) where \( \gamma_j \) is a non–local function of the \( \alpha \)'s.
There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC. In a recent preprint, GNR obtain a sum rule and gem for 
\( e = \{ e^{i\theta} \mid \alpha \leq \theta \leq 2\pi - \alpha \} \) for \( 0 < \alpha < \pi \). For real \( \alpha \), the Verblunsky side has the expected \( \sum |\alpha_j - a|^2 \) form but for general \( \alpha \), it has the form \( \sum |\gamma_j - a|^2 \) where \( \gamma_j \) is a non–local function of the \( \alpha \)’s. In particular, it is not clear if the finiteness of their Verblunsky side only depends on the behavior near \( j = \infty \).
There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC. In a recent preprint, GNR obtain a sum rule and gem for
\[ e = \{ e^{i\theta} \mid \alpha \leq \theta \leq 2\pi - \alpha \} \text{ for } 0 < \alpha < \pi. \]
For real \( \alpha \), the Verblunsky side has the expected \( \sum |\alpha_j - a|^2 \) form but for general \( \alpha \), it has the form \( \sum |\gamma_j - a|^2 \) where \( \gamma_j \) is a non–local function of the \( \alpha \)'s. In particular, it is not clear if the finiteness of their Verblunsky side only depends on the behavior near \( j = \infty \). At least for the real case, it would be interesting to get the sum rule via the Poisson–Jensen methods used in the original Killip–Simon proof.
There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC. In a recent preprint, GNR obtain a sum rule and gem for $e = \{e^{i\theta} | \alpha \leq \theta \leq 2\pi - \alpha\}$ for $0 < \alpha < \pi$. For real $\alpha$, the Verblunsky side has the expected $\sum |\alpha_j - a|^2$ form but for general $\alpha$, it has the form $\sum |\gamma_j - a|^2$ where $\gamma_j$ is a non–local function of the $\alpha$’s. In particular, it is not clear if the finiteness of their Verblunsky side only depends on the behavior near $j = \infty$. At least for the real case, it would be interesting to get the sum rule via the Poisson–Jensen methods used in the original Killip–Simon proof. It would also be interesting to understand the $\gamma_j$’s in a more conventional setting.
There has been very little work on Killip–Simon type theorems for finite gap sets in OPUC. In a recent preprint, GNR obtain a sum rule and gem for $e = \{e^{i\theta} \mid \alpha \leq \theta \leq 2\pi - \alpha\}$ for $0 < \alpha < \pi$. For real $\alpha$, the Verblunsky side has the expected $\sum |\alpha_j - a|^2$ form but for general $\alpha$, it has the form $\sum |\gamma_j - a|^2$ where $\gamma_j$ is a non–local function of the $\alpha$’s. In particular, it is not clear if the finiteness of their Verblunsky side only depends on the behavior near $j = \infty$. At least for the real case, it would be interesting to get the sum rule via the Poisson–Jensen methods used in the original Killip–Simon proof. It would also be interesting to understand the $\gamma_j$’s in a more conventional setting.

Understanding perturbations of periodic and the more general finite gap OPUC remains open.
Finally, we note that Killip–Simon have proven a sum rule and gem for half–line Schrödinger operators when
\[ V \in L^2((0, \infty); dx). \]
Finally, we note that Killip–Simon have proven a sum rule and gem for half–line Schrödinger operators when $V \in L^2((0, \infty); dx)$. It would be very interesting to find a large deviation proof of this result.
Finally, we note that Killip–Simon have proven a sum rule and gem for half–line Schrödinger operators when $V \in L^2((0, \infty); dx)$. It would be very interesting to find a large deviation proof of this result. In particular, what is the analog of random matrix models for the study of Schrödinger operators?
And Now a Word from Our Sponsor
A Comprehensive Course in Analysis by Poincaré Prize winner Barry Simon is a five-volume set that can serve as a graduate-level analysis textbook with a lot of additional bonus information, including hundreds of problems and numerous notes that extend the text and provide important historical background. Depth and breadth of exposition make this set a valuable reference source for almost all areas of classical analysis.

Part 1 is devoted to real analysis. From one point of view, it presents the infinitesimal calculus of the twentieth century with the ultimate integral calculus (measure theory) and the ultimate differential calculus (distribution theory). From another, it shows the triumph of abstract spaces: topological spaces, Banach and Hilbert spaces, measure spaces, Riesz spaces, Polish spaces, locally convex spaces, Fréchet spaces, Schwartz space, and $L^p$ spaces. Finally, it is the study of big techniques, including the Fourier series and transform, dual spaces, the Baire category, fixed point theorems, probability ideas, and Hausdorff dimension. Applications include the constructions of nowhere differentiable functions, Brownian motion, space-filling curves, solutions of the moment problem, Haar measure, and equilibrium measures in potential theory.
A Comprehensive Course in Analysis by Poincaré Prize winner Barry Simon is a five-volume set that can serve as a graduate-level analysis textbook with a lot of additional bonus information, including hundreds of problems and numerous notes that extend the text and provide important historical background. Depth and breadth of exposition make this set a valuable reference source for almost all areas of classical analysis. Part 2A is devoted to basic complex analysis. It interweaves three analytic threads associated with Cauchy, Riemann, and Weierstrass, respectively. Cauchy’s view focuses on the differential and integral calculus of functions of a complex variable, with the key topics being the Cauchy integral formula and contour integration. For Riemann, the geometry of the complex plane is central, with key topics being fractional linear transformations and conformal mapping. For Weierstrass, the power series is king, with key topics being spaces of analytic functions, the product formulas of Weierstrass and Hadamard, and the Weierstrass theory of elliptic functions. Subjects in this volume that are often missing in other texts include the Cauchy integral theorem when the contour is the boundary of a Jordan region, continued fractions, two proofs of the big Picard theorem, the uniformization theorem, Ahlfors’s function, the sheaf of analytic germs, and Jacob, as well as Weierstrass, elliptic functions.
A Comprehensive Course in Analysis by Poincaré Prize winner Barry Simon is a five-volume set that can serve as a graduate-level analysis textbook with a lot of additional bonus information, including hundreds of problems and numerous notes that extend the text and provide important historical background. Depth and breadth of exposition make this set a valuable reference source for almost all areas of classical analysis.

Part 2B provides a comprehensive look at a number of subjects of complex analysis not included in Part 2A. Presented in this volume are the theory of conformal metrics (including the Poincaré metric, the Ahlfors-Robinson proof of Picard’s theorem, and Bell’s proof of the Painlevé smoothness theorem), topics in analytic number theory (including Jacobi’s two- and four-square theorems, the Dirichlet prime progression theorem, the prime number theorem, and the Hardy-Littlewood asymptotics for the number of partitions), the theory of Fuchsian differential equations, asymptotic methods (including Euler’s method, stationary phase, the saddle-point method, and the WKB method), univalent functions (including an introduction to SLE), and Nevanlinna theory. The chapters on Fuchsian differential equations and on asymptotic methods can be viewed as a minicourse on the theory of special functions.
A Comprehensive Course in Analysis by Poincaré Prize winner Barry Simon is a five-volume set that can serve as a graduate-level analysis textbook with a lot of additional bonus information, including hundreds of problems and numerous notes that extend the text and provide important historical background. Depth and breadth of exposition make this set a valuable reference source for almost all areas of classical analysis.

Part 3 returns to the themes of Part 1 by discussing pointwise limits (going beyond the usual focus on the Hardy-Littlewood maximal function by including ergodic theorems and martingale convergence), harmonic functions and potential theory, frames and wavelets, $H^p$ spaces (including bounded mean oscillation (BMO)) and, in the final chapter, lots of inequalities, including Sobolev spaces, Calderon-Zygmund estimates, and hypercontractive semigroups.
A Comprehensive Course in Analysis by Poincaré Prize winner Barry Simon is a five-volume set that can serve as a graduate-level analysis textbook with a lot of additional bonus information, including hundreds of problems and numerous notes that extend the text and provide important historical background. Depth and breadth of exposition make this set a valuable reference source for almost all areas of classical analysis.

Part 4 focuses on operator theory, especially on a Hilbert space. Central topics are the spectral theorem, the theory of trace class and Fredholm determinants, and the study of unbounded self-adjoint operators. There is also an introduction to the theory of orthogonal polynomials and a long chapter on Banach algebras, including the commutative and non-commutative Gel'fand-Naimark theorems and Fourier analysis on general locally compact abelian groups.

For additional information and updates on this book, visit www.ams.org/bookpages/simon

AMS on the Web www.ams.org


Operator Theory
A Comprehensive Course in Analysis, Part 4

Barry Simon

A = \int t \, dE_t

\det(1 + zA) = \prod_{k=1}^{N(A)} (1 + z\lambda_k(A))

Google simon comprehensive course