Closest Spacing of Eigenvalues

Jade P. Vinson

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MATHEMATICS

JUNE, 2001
Abstract

We consider several variants of the following problem: pick an $n \times n$ matrix from some unitary ensemble of random matrices. In an interval containing many eigenvalues, what is the closest spacing between two eigenvalues? We are interested in the correct scaling for this random variable as $n \to \infty$, and the limiting distribution of the rescaled random variable.

One can predict the distribution of the minimum spacing heuristically by assuming that the consecutive spacings are independent random variables, chosen from the consecutive spacing distribution for unitary ensembles. The consecutive spacings are not independent. However, in all cases studied, the heuristics predict asymptotically the correct scaling and distribution of the minimum spacing.

Using the method of moments, we show that the number of eigenvalue pairs in a given interval and closer than a small distance $\gamma$, is approximately a Poisson random variable with mean as predicted by heuristics. Varying $\gamma$, we obtain the result for minimum spacing.

In the most concrete special case, our Main Theorem is this:

**Theorem 0.1 (Main Theorem).** Choose a random $n \times n$ unitary matrix. Let $Z_n$ be the closest spacing between any two eigenvalues. Fix $\beta > 0$. Then as $n \to \infty$,

$$\Pr \left( Z_n \left( \frac{n^4}{72\pi} \right)^{\frac{1}{3}} > \beta \right) \to e^{-\beta^3}.$$

The correct scaling for the closest eigenvalue spacing is $n^{-1}(n|I_n|)^{-\frac{1}{3}}$. In this case, $n^{-1}$ is comparable to the mean spacing and $n|I_n|$ is comparable to the expected number of eigenvalues in the interval being considered. On the other hand, if the eigenvalues were independent and distributed according to a Poisson process, then the closest spacing would scale like $n^{-1}(n|I_n|)^{-1}$, which is much closer. Thus we have another confirmation of what is often said: “The eigenvalues of a unitary matrix repel each other”.

Our results also apply to the Gaussian unitary ensemble (GUE) and, with restrictions, to a universal class of unitary ensembles (UUE) studied by Deift, Kreicherbauer, McLaughlin, Venakides, and Zhou. In each of these cases, the expected number of eigenvalues in the interval must be large, and the interval must be contained inside the bulk distribution of the eigenvalues.
Acknowledgements

I would like to thank my advisor, Peter Sarnak, for suggesting the problem of minimum spacing and an approach using the method of moments. Peter Sarnak always shared his advice on mathematical problems, and exercised good judgement with regard to his role in my studies. At first he played an active role, suggesting topics to think about and articles to read, but during the last year his involvement tapered off rapidly as soon as the minimum spacing problem looked promising. I think that this is the correct balance.

Financial support was provided by the National Defense Science and Engineering Graduate Fellowship Program.

I find that writing a sincere Acknowledgement is an extremely difficult task, and it would take days of reflection to do it well. Therefore instead of attempting to thank the many people who have influenced my mathematical life, I will limit the scope of this Acknowledgement to the dissertation itself.

I would like to thank my family for encouraging me to complete the Ph.D. in time for graduation in June. Assuming the final public oral in a week is successful, I will be able to relax for the first summer of my adult life. I would also like to thank my colleagues at the Institute for Defense Analyses for encouraging me to finish my Ph.D. now rather than later.

The dissertation was completed with little time to spare. I would like to thank Peter Sarnak and Yakov Sinai for reading it and making suggestions. I would also like to thank Scott Kenney, the department manager, for walking me through the formalities required to schedule a final public oral.
1. Introduction

We consider several variants of the following problem: pick an $n \times n$ matrix from some unitary ensemble of random matrices. In an interval containing many eigenvalues, what is the closest spacing between two eigenvalues? We are interested in the correct scaling for this random variable as $n \to \infty$, and the limiting distribution of the rescaled random variable.

One can predict the distribution of the minimum spacing heuristically by assuming that the consecutive spacings are independent random variables, chosen from the consecutive spacing distribution for unitary ensembles. The consecutive spacings are not independent. However, in all cases studied, the heuristics predict asymptotically the correct scaling and distribution of the minimum spacing.

Using the method of moments, we show that the number of eigenvalue pairs in a given interval and closer than a small distance $\gamma$, is approximately a Poisson random variable with mean as predicted by heuristics. Varying $\gamma$, we obtain the result for minimum spacing.

Our results apply to the circular unitary ensemble (CUE), the Gaussian unitary ensemble (GUE), and with restrictions to a universal class of unitary ensembles (UUE) studied by Deift, Kreicherbauer, McLaughlin, Venakides, and Zhou. In each of these cases, the expected number of eigenvalues in the interval must be large, and the interval must be contained inside the bulk distribution of the eigenvalues.

1.1. Statement of the Main Theorem. The Main Theorem is easiest to state in the case of the circular unitary ensemble (CUE), which is the compact group $U_n$ with Haar measure.

**Theorem 1.1.** For each $n$, let $I_n \subset [0, 2\pi)$ be an interval so that $n|I_n| \to \infty$ as $n \to \infty$.

Choose a matrix randomly from $CUE_n$ and let $Z_n$ be the closest spacing between any two eigenvalues whose average is in $I_n$. Fix $\beta > 0$. Then as $n \to \infty$,

$$\Pr \left( Z_n \left( \frac{n^4 |I_n|}{144\pi^2} \right) ^{\frac{1}{3}} > \beta \right) \to e^{-\beta^3}.$$

Of particular interest is the case when $I_n = [0, 2\pi)$ for all $n$:

**Theorem 1.2** (Main Theorem, CUE Version). Choose a matrix randomly from $CUE_n$ and let $Z_n$ be the closest spacing between any two eigenvalues. Fix $\beta > 0$. Then as $n \to \infty$,

$$\Pr \left( Z_n \left( \frac{n^4}{72\pi} \right) ^{\frac{1}{4}} > \beta \right) \to e^{-\beta^3}.$$
Given Theorem 1.2 and its Corollary, observe that the correct scaling for the closest eigenvalue spacing is \( n^{-1}(n|I_n|)^{-\frac{1}{3}} \). In this case, \( n^{-1} \) is comparable to the mean spacing and \( n|I_n| \) is comparable to the expected number of eigenvalues in the interval being considered. On the other hand, if the eigenvalues were independent and distributed according to a Poisson process, then the closest spacing would scale like \( n^{-1}(n|I_n|)^{-1} \), which is much closer. Thus we have another confirmation of what is often said: “The eigenvalues of a unitary matrix repel each other”.

The universal unitary ensembles (UUE) are not completely standard. Given a real analytic potential \( V(x) \) with sufficient growth at infinity, \( UUE_n \) is the ensemble of Hermitian matrices with the following joint probability density function (j.p.d.f.) for the matrix entries:

\[
\text{(const)} \ e^{-n \sum \lambda_j} dM
\]

Using Weyl integration (Appendix 13), the j.p.d.f. for the eigenvalues is

\[
\text{(const)} \ \prod_{i<j} (\lambda_i - \lambda_j)^2 e^{-n \sum \lambda_i} d\Lambda
\]

One could recover the Gaussian Unitary Ensemble (GUE) as a special case of UUE by choosing the potential \( V(x) = x^2 \) and rescaling properly. The word “unitary” in this context confuses many newcomers to the field of random matrices. The GUE is a probability distribution on the set of Hermitian matrices which is invariant under conjugation by any unitary matrix.

Our main result, in the case of universal ensembles, is the following:

**Theorem 1.3** (Main Theorem, universal version). Let \( V(x) \) be a real analytic potential which is regular and whose equilibrium measure \( \Psi(x)dx \) is supported on a single interval \([a, b]\). Fix \( \epsilon > 0 \). For each \( n \), let \( I_n \subset [a + \epsilon, b - \epsilon] \) be an interval contained in the bulk distribution of eigenvalues, so that \( n|I_n| \to \infty \) as \( n \to \infty \).

Choose a matrix randomly from \( UUE_n \) and let \( Z_n \) be the closest spacing between any two eigenvalues whose average is in \( I_n \). Fix \( \beta > 0 \). Then as \( n \to \infty \),

\[
\Pr \left( Z_n \left( \frac{\pi^2 n^4}{9} \int_{I_n} \Psi(x)^4 dx \right)^{\frac{1}{2}} > \beta \right) \to e^{-\beta^3}.
\]

**1.2. Overall Strategy.** We analyze the random variable \( Z_n \) indirectly. Let \( G_{\gamma,I_n} \) count the number of eigenvalue pairs whose average is in \( I_n \) and whose separation is at most \( \gamma \). For example, if \( \lambda_1 < \lambda_2 < \lambda_3 < \lambda_4 < \lambda_1 + \gamma \), then this contributes 6 pairs to \( G_{\gamma,I_n} \). For each \( \gamma \) we analyze the
random variable $G_{\gamma,I,n}$ in detail, but what interests us most is the probability $\Pr(G_{\gamma,I,n} = 0)$:

\[
\begin{align*}
(Z_n > \gamma) & \iff (G_{\gamma,I,n} = 0) \\
\Pr(Z_n > \gamma) & = \Pr(G_{\gamma,I,n} = 0)
\end{align*}
\]

Since $G_{\gamma,I,n}$ is an integer valued random variable, $\Pr(G_{\gamma,I,n} = 0)$ is accessible from the moments of $G_{\gamma,I,n}$. The following Theorem says that the moments of $G_{\gamma,I,n}$ are approximately the moments of a Poisson distribution.

**Theorem 1.4** (Moment Estimation Theorem, universal version). Let $V(x)$ be a real analytic potential which is regular and whose equilibrium measure $\Psi(x)dx$ is supported on a single interval $[a,b]$. Fix $\epsilon > 0$. For each $n$, let $I_n \subset [a + \epsilon, b - \epsilon]$ be an interval contained in the bulk distribution of eigenvalues.

For each $n$, let $\gamma_n > 0$. Let $G_{\gamma,I,n}$ be the random variable which counts the number of UUE $n$ eigenvalues whose average is in $I_n$ and whose difference is at most $\gamma_n$. Let $G_\mu$ be the Poisson distribution with mean

\[
\mu_n = \frac{\pi^2 n^4}{9} \int_{I_n} \Psi(x)^4 dx.
\]

Then for all $k \geq 1$, as $n \to \infty$,

\[
E(G_{k,I,n}^k) = E(G_\mu^k) \left(1 + O \left(n^{-1} + \gamma^2 n^2 + (\gamma^2 n^2)^3 + (n |I|^{-2}) \right) \right).
\]

The constant implied by $O$ depends only on $k$, the potential $V(x)$, and $\epsilon$.

Proving the Moment Estimation Theorem is the major task of this paper. First, let us assume the Moment Estimation Theorem and use it to prove the Main Theorem.

**1.3. Proof that the Moment Estimation Theorem implies the Main Theorem.** In applying the Moment Estimation Theorem, we get to choose $\gamma_n$. We choose:

\[
\gamma_n = \beta \left( \frac{\pi^2 n^4}{9} \int_{I_n} \Psi(x)^4 dx \right)^{-\frac{1}{4}} \\
\propto n^{-\frac{1}{3} \frac{|I_n|^{-\frac{1}{3}}}}.
\]

With this choice of $\gamma_n$, $\mu_n = \beta^3$ for all $n$. Applying the Moment Estimation Theorem,

\[
E(G_{\gamma,I,n}) = E(G_\beta^k) \left(1 + O \left(n |I_n|^{-\frac{3}{2}} \right) \right).
\]

By assumption, $(n |I_n|) \to \infty$ as $n \to \infty$, so the moments of $G_{\gamma,I,n}$ converge to those of $G_\beta$. Thus,

\[
\Pr(G_{\gamma,I,n} = 0) \to \Pr(G_\beta = 0) = e^{-\beta^3} \quad \text{as} \quad n \to \infty.
\]
We retrace our steps back to the minimum spacing $Z_n$. The random variable $Z_n$ is greater than $\gamma_n$ if and only if $G_{\gamma,I,n} = 0$. Substituting in the chosen value of $\gamma_n$, this is equivalent to

$$Z_n \left( \frac{\pi^2 n^4}{9} \int_{I_n} \Psi(x)^4 \, dx \right)^{\frac{1}{4}} > \beta,$$

completing the proof.

### 1.4. Outline of Paper.

In the remainder of this paper we prove the Moment Estimation Theorem. We first prove the case of CUE because this case is by far the easiest. The novel calculations for CUE are

- The combinatorial methods used enumerate the various contributions to the moment $E(G_{\gamma,I,n})$. Each contribution is expressed as an integral using the method of Gaudin.
- Calculation of the main contributions using asymptotics for the projection kernel $K_n(x,y)$ and its derivatives.
- Bounds for the undesired contributions, which are implied by bounds for $K_n(x,y)$ and its derivatives.

For CUE, the projection kernel is very easy to work with:

$$K_n(x,y) = \frac{1}{2\pi} \frac{e^{in(x-y)} - 1}{e^{i(x-y)} - 1}.$$ 

See Appendix [14.4](#).

We prove the UUE version of the Moment Estimation Theorem using exactly the same techniques. Parts of the proof are identical to the CUE case, and are omitted. The new feature the UUE case is that the projection kernel $K_n(x,y)$ is less easy to work with:

$$K_n(x,y) = \sum_{j=0}^{n} \eta_j(x) \eta_j(y) = \frac{(\text{const}) \eta_n(x) \eta_{n-1}(y) - \eta_n(y) \eta_{n-1}(x)}{x - y}.$$ 

where $\phi_j(x)$ is the $(j)$th normalized orthogonal polynomial with respect to the weight $e^{-nV(x)}$ and $\eta_j(x) = e^{-\frac{2}{n}V(x)} \phi_j(x)$.

The asymptotics for the orthogonal polynomials $\phi_{n-1}$ and $\phi_n$ were obtained recently by Deift, Kreicherbauser, McLaughlin, Venakides and Zhou. Based closely on [DKM+97](#) and [Dei99](#), we outline the derivation of the leading order asymptotics. By a very minor modification of these techniques, we obtain the leading order asymptotics for the derivatives of $\eta_{n-1}$ and $\eta_n$—which, as one would expect, are the derivatives of the leading order asymptotics.
The case of the Gaussian unitary ensemble is of special interest. The analogs of the Main Theorem and the Moment Estimation Theorem could be obtained by rescaling the GUE and applying the universal results for the potential $V(x) = x^2$. We present an alternate proof of the GUE Theorems, using Plancherel-Rotach asymptotics for Hermite polynomials in place of the more general Deift-Zhou asymptotics for orthogonal polynomials.

Given our results about the closest spacing for eigenvalues in an interval, it is natural to ask about the distribution of the maximum spacing between consecutive eigenvalues in some interval. We have not studied this problem in great detail, but believe it is more difficult than the minimum spacing problem. At least we have not been able to deal with it yet. In section 15, we outline one possible approach to studying the maximum spacing.

We call the reader’s attention to [Edc92] and to related papers available on Edelman’s web page. In these paper Edelman finds the correct scaling and distribution of the condition number

$$\sqrt{\sum |\lambda_j|^2 / |\lambda_1|},$$

where $\lambda_1$ is the complex eigenvalue with smallest absolute value. The condition number is an indicator of the difficulty of inverting a matrix numerically. Edelman considers the ensemble of matrices whose real or complex entries are chosen independently from Gaussians. Since the numerator is very tightly distributed because of the law of large numbers, Edelman’s result concerns the distribution of the smallest eigenvalue for a matrix in this ensemble.

2. Case 1: The Circular Unitary Ensemble (CUE)

We now state and prove the Moment Estimation Theorem for the circular unitary ensemble. The Moment Estimation Theorem for the CUE case implies the Main Theorem for the CUE case, which was already stated in the introduction.

**Theorem 2.1** (Moment Estimation Theorem, CUE version). For each $n$, let $\gamma_n > 0$ and $I_n \subset [0, 2\pi)$.

Let $G_{\gamma,I,n}$ be the random variable which counts the number of CUE eigenvalues whose average is in $I_n$ and whose difference is at most $\gamma_n$. Let $G_\mu$ be the Poisson distribution with mean

$$\mu_n = |I_n|\gamma_n^3 n^4 / 144\pi^2.$$

Then for all $k \geq 1$, as $n \to \infty$,

$$E(G_{\gamma,I,n}^k) = E(G_\mu^k) \left(1 + O\left(n^{-1} + \gamma^2 n^2 + (\gamma^2 n^2)^3 + (n|I|)^{-\frac{5}{2}}\right)\right).$$

The constant implied by $O$ depends only on $k$. 


Before proving this Theorem, we will show that it agrees with heuristic predictions.

2.1. Heuristic Prediction Based on the Consecutive Spacing Distribution. We make the simplifying assumption that the consecutive spacings are independent random variables, all chosen from the consecutive spacing distribution for unitary ensembles. This assumption of independence is false, since adjacent consecutive spacings are anti-correlated. See p. 111 of [Meh91] for a contour plot of the joint probability density function of two adjacent consecutive spacings. Observe that short spacings tend to be followed by longer ones and vice-versa. However, we will see that the independence assumption does lead to correct predictions for the minimum spacing.

For the unitary ensembles, the consecutive spacing distribution vanishes to order two at the origin, so that very short consecutive spacings are unlikely and the eigenvalues are said to “repel.” See [Meh91] Chapter 5 and Appendix 13. For a power series expansion of this density function at the origin, the first few terms are:

\[ p(x) = \frac{\pi^2}{3} x^2 - \frac{2\pi^4}{45} x^4 + \frac{\pi^6}{315} x^6 - \ldots \]

When a matrix is chosen from \( U_n \), the mean spacing of the eigenvalues is the constant \( \frac{2\pi}{n} \), independent of \( \theta \). Thus a separation of \( \gamma \) is equal to \( \left( \frac{2\pi}{n} \right) \) times the mean spacing. When \( \left( \frac{2\pi}{n} \gamma \right) \) is small, the probability of any one of these spacings being less than \( \gamma \) is about \( \left( \frac{\pi^2}{9} \right) \left( \frac{2\pi}{n} \gamma \right)^3 \). The number of consecutive spacings is about \( \frac{n|I_n|}{2\pi} \), so the expected number of consecutive spacings less than \( \gamma \) is about

\[ \mu = \frac{n|I_n|}{2\pi} \left( \frac{\pi^2}{9} \right) \left( \frac{n\gamma}{2\pi} \right)^3 = \frac{|I_n|n^4\gamma^3}{144\pi^2} \]

Since \( G_{\gamma,I,n} \) is a sum of many independent unlikely events, it is approximately Poisson with mean \( \mu \).

3. The First Moment of \( G_{\gamma,I,n} \)

Recall that \( G_{\gamma,I,n} \) is the number of GUE-n eigenvalue pairs whose average is in \( I_n \) and whose difference is less than \( \gamma \). Thus \( G_{\gamma,I,n} \) is the symmetrization to \( n \) variables of a function of 2 variables:

\[ g(u, t) = \begin{cases} \frac{1}{2} & \text{if } |u - t| < \gamma \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \frac{u + t}{2} \in I \]

\[ G_{\gamma,I,n}(t_1, t_2, \ldots, t_n) = \sum_{i \neq j} g(t_i, t_j). \]
Using Theorem 14.5, the expected value of $G_{\gamma,I,n}$ is

$$E(G_{\gamma,I,n}) = \int \int_{[0,2\pi]^2} g(u,t) \begin{bmatrix} K_n(u,u) & K_n(u,t) \\ K_n(t,u) & K_n(t,t) \end{bmatrix} dudt$$

$$= \frac{1}{2} \int \int_{\Omega} \begin{bmatrix} K_n(u,u) & K_n(u,t) \\ K_n(t,u) & K_n(t,t) \end{bmatrix} dudt.$$

In the above formula, the region $\Omega$ is the set where $g_2(t_1,t_2) = \frac{1}{2}$:

$$\Omega = \left\{ (u,t) \text{ s.t. } \frac{u+t}{2} \in I, |u-t| < \gamma \right\}.$$

The projection kernel $K_n(u,t)$ for $U_n$ is

$$K_n(u,t) = \frac{1}{2\pi} e^{i(u-t)} - 1 = K_n(u-t)$$

Since $(t-u)$ is small in the region $\Omega$, we change variables and expand the determinant as a power series. We let $x = \frac{u+t}{2}$, $y = \frac{u-t}{2}$, and $dxdy = \frac{1}{2}dudt$.

$$E(G_{\gamma,I,n}) = \int \int_{I_n} \int_{-\gamma}^{\gamma} \left[ \frac{K_n(0)}{K_n(-2y)} \frac{K_n(2y)}{K_n(0)} \right] dudt$$

$$= \int \int_{I_n} \int_{-\gamma}^{\gamma} 4y^2 \left( (K'(0))^2 - K(0)K''(0) \right) dydx,$$

where $c_j$ is an upper bound for the $(j)$th derivative of $K_n$. We apply the following Lemma:

**Lemma 3.1.** Consider the kernel $\frac{1}{2\pi} e^{inx} - 1$. For $x > x_0 \geq \frac{1}{n} (\text{mod } 2\pi)$, the $(k)$th derivative of this kernel is

$$O\left( \frac{1}{x_0 n^k} \right)$$

Without the restriction on $x$, the derivative is $O(n^{k+1})$.

**Proof.** Consider taking $k$ derivatives symbolically. After taking $0 \leq j \leq k$ of these derivatives, the result is a sum of terms of the form:

$$\text{monomial}(n, e^{inx}, e^{ix}) \left( e^{ix} - 1 \right)^t.$$

At each step, the worst that can happen is that either the exponent in the denominator increases by 1, or the power of $n$ in the numerator increases by 1. Since we assumed that $x_0 \geq \frac{1}{n}$, the worse of these two outcomes is multiplying by $n$. The result follows by induction, and observing that the number of terms is finite, depending only on $k$. 
For the bound without the restriction on \( t \), express the kernel as a sum of complex exponentials and differentiate termwise.

Returning to our expression for \( E(F_{m,n}) \),

\[
E(F_{m,n}) = \int_{I_n} \int_{-\frac{1}{2}\gamma}^{\frac{1}{2}\gamma} 4y^2 ((K'(0))^2 - K(0)K''(0)) + O(y^4n^6 + \cdots + y^8n^{10})dydx
\]

\[
= \frac{1}{4\pi^2} \int_{I_n} \int_{-\frac{1}{2}\gamma}^{\frac{1}{2}\gamma} 4y^2 \left(-\frac{n^4}{4} + \frac{n^4}{3} + O(n^3)\right) dydx + O\left(|I|\gamma^5n^6 + \gamma^9n^{10}\right)
\]

\[
= \frac{|I_n|\gamma^3n^4}{144\pi^2} \left(1 + O\left(\frac{1}{n} + \gamma^2n^2 + (\gamma^2n^2)^3\right)\right)
\]

This agrees with the heuristic predictions in Subsection 2.1.

4. The Higher Moments of \( G_{\gamma,I,n} \)

The \((k)\)th power of \( G_{\gamma,I,n}(t_1, t_2, \ldots, t_n) \) may be written

\[
G_{\gamma,I,n}^k = \left(\sum_{i \neq j} g(t_i, t_j)\right)^k
\]

\[
= \sum_{i_1 \neq j_1; i_2 \neq j_2; \ldots; i_k \neq j_k} g(t_{i_1}, t_{j_1})g(t_{i_2}, t_{j_2})\cdots g(t_{i_k}, t_{j_k}),
\]

where

\[
g(u, t) = \begin{cases} 
\frac{1}{2} & \text{if } |u - t| < \gamma \\
\frac{u + t}{2} & \text{if } \frac{u + t}{2} \in I \\
0 & \text{otherwise}
\end{cases}
\]

The indexing set for the sum on the right is not of the form required for Theorem 14.5. However it may be written as a disjoint union of indexing sets of the proper form.

4.1. The Combinatorics of Collapses. Let us, for example, partition the indexing set \( i_1 \neq j_1; i_2 \neq j_2; i_3 \neq j_3 \) into smaller indexing sets of the proper form. We write this indexing set in shorthand as

\[
(i_1|j_1) \wedge (i_2|j_2) \wedge (i_3|j_3),
\]
and expand as follows. Except for the subtle distinction between “|” and “∀”, each step below should be clear.

\[(i_1|j_1) \land (i_2|j_2) \land (i_3|j_3)\]
\[= (i_1|j_1) ((i_2|j_2 \tilde{\pi}_3|j_3) \lor (i_2, i_3|j_2|j_3) \lor (i_2, i_3|j_2|j_3) \lor (i_2, i_3|j_3|j_2) \lor (i_2, j_3|j_2) \lor (i_2, j_3|j_2))\]
\[= ((i_1|j_1) (i_2|j_2 \tilde{\pi}_3|j_3) \lor (i_1|j_1, j_2 \tilde{\pi}_3|j_3) \lor (i_1|j_1, j_2|\tilde{\pi}_3|j_3))\]
\[= ((i_1|j_1, i_2|j_2 \tilde{\pi}_3|j_3) \lor (i_1|j_1, i_3|j_2|j_3) \lor (i_1|j_1, i_3|j_2|j_3) \lor (i_1|j_1, i_3|j_2|j_3))\]
\[= ((i_1|j_1, i_3|j_3|j_3) \lor (i_1|j_1, j_3|j_3|j_3) \lor (i_1, i_2|j_1, i_3|j_3|j_3) \lor (i_1, i_3|j_1, i_2|j_3|j_3) \lor (i_1, i_3|j_1, i_2|j_3|j_3))\]
\[= ((i_1, i_2|j_1, i_3|j_3) \lor (i_1, i_2|j_1, i_3|j_3) \lor (i_1, i_3|j_1, j_2|j_3) \lor (i_1, i_3|j_1, j_2|j_3))\]
\[= ((i_1, i_3|j_3|j_3|j_3) \lor (i_1, j_3|j_3|j_3|j_3) \lor (i_1, j_3|j_3|j_3|j_3) \lor (i_1, j_3|j_3|j_3|j_3))\]

Each term above is an indexing set of the form required for Theorem 14.5. We call such indexing sets “collapses.” Let us take one of the more complicated collapses above and explain it’s meaning:

\[(i_1|j_1, i_3|j_3 \tilde{\pi}_2|j_2)\]

This means that \(j_1 = i_3\), but otherwise the indices \(i_1, j_1, i_2, j_2, i_3, j_3\) are distinct. There are 53 indexing sets in the full expansion, but we only list 28 of them above. The symbols “|” and “∀” both separate the \(i\)’s and \(j\)’s into equivalence classes. The symbol \(\tilde{\pi}\) plays an additional role which will not be clear until later in this section. If \(i_l\) and \(j_l\) are in different equivalence classes, then we say that these two equivalence classes fall into the same cluster. The clusters are separated by \(\tilde{\pi}\).

Every cluster contains at least two equivalence classes. If any cluster in a collapse contains more than two equivalence classes, we call the collapse a mixed collapse. If every cluster contains exactly two equivalence classes we call it a clean collapse. We will show that the mixed collapses make a negligible contribution to \(E(G^k_n, I_n)\).

There are 42 mixed collapses for \(k = 3\), which we do not list explicitly.

The 11 clean collapses for \(k = 3\) are:

\[
[(i_1|j_1 \tilde{\pi}_2|j_2 \tilde{\pi}_3|j_3)] \iff (1|2|3)
\]
\[
[(i_1, i_2|j_1, j_2 \tilde{\pi}_3|j_3)] \iff (1|2|3)
\]
\[
[(i_1, i_3|j_1, j_3 \tilde{\pi}_2|j_2)] \iff (1|2|3)
\]
\[
[(i_1|j_1 \tilde{\pi}_2|j_2, j_3|j_2)] \iff (1|2|3)
\]
\[
[(i_1|j_1 \tilde{\pi}_2|j_2, i_3|j_3)] \iff (1|2|3)
\]
\[
[(i_1|j_1 \tilde{\pi}_2|j_2, i_3|j_3)] \iff (1|2|3)
\]
\[
[(i_1|j_1 \tilde{\pi}_2|j_2, i_3|j_3)] \iff (1|2|3)
\]

In the list above, we have collected those indexing sets together so that each collection corresponds to one way of partitioning \(\{1, 2, 3\}\).
The following Lemma is evident, based on consideration of the expansion of \((i_1|j_1) \land (i_2|j_2) \land (i_3|j_3)\) into collapses which we considered above.

**Lemma 4.1.** When the indexing set \((i_1|j_1) \land (i_2|j_2) \land \cdots \land (i_k|j_k)\) is expanded into collapses, every cluster contains at least two equivalence classes. Clean collapses with \(l\) clusters arise from a partition of the integers \(\{1, 2, \ldots, k\}\) into \(l\) nonempty subsets. Each of \(l\) subsets of \(\{1, 2, \ldots, k\}\) leads to a cluster of equivalence classes of \(\{i_1, j_1, i_2, j_2, \ldots, i_k, j_k\}\). Each partition of \(\{1, 2, \ldots, k\}\) into \(l\) equivalence classes corresponds to \(2^{k-1}\) different clean collapses with \(l\) clusters.

We will see that clean collapses make the dominant contribution to \(E(G_k^\gamma, I, n)\). For each clean collapse, the “block diagonal” is the dominant contribution. Using a combinatorial analysis, we’ll see that the sum over all clean collapse block diagonal terms is approximately equal to \(E(G_k^\mu)\), in agreement with Theorem 2.1.

**4.2. Estimating the clean collapses.** As an example, we select one of the clean collapses from \(G_3^\gamma, I, n\), use Theorem 14.5 to express its contribution to \(E(G_3^\gamma, I, n)\) as an integral, and estimate that integral. We select:

\((i_1, j_2|j_1, i_2, j_3|j_3)\).

Making the substitutions \(j_2 = i_1\) and \(i_2 = j_1\), this indexing set makes the following contribution to \(E(G_3^\gamma, I, n)\):

\[
\sum_{i_1, j_1, i_3, j_3 \text{ distinct}} g(t_{i_1}, t_{j_1}) g(t_{j_1}, t_{i_1}) g(t_{i_3}, t_{j_3})
\]

Since \(g\) takes only the values 0 or \(\frac{1}{2}\) and is symmetric, \(g(t_{i_1}, t_{j_1}) g(t_{j_1}, t_{i_1}) = \frac{1}{2} g(t_{i_1}, t_{j_1})\). The contribution to \(E(G_3^\gamma, I, n)\) simplifies to:

\[
\frac{1}{2} \sum_{i_1, j_1, i_3, j_3 \text{ distinct}} g(t_{i_1}, t_{j_1}) g(t_{i_3}, t_{j_3})
\]

This is the symmetrization of a function of four variables, so Theorem 14.5 expresses the expected value of this sum as an integral involving four variables:
For a moment, let us take for granted that the block-diagonal terms of the determinant, the ones which make a significant contribution to the integral are the four terms on the block diagonal. Their contribution is:

\[
\frac{1}{2} \int_{\Omega} \begin{bmatrix}
K_n(u_1, u_1) & K_n(u_1, u_1) & 0 & 0 \\
0 & K_n(t_1, t_1) & K_n(u_3, u_3) & 0 \\
K_n(t_3, u_1) & 0 & K_n(t_3, u_3) & 0 \\
0 & 0 & K_n(t_3, u_3) & K_n(t_3, t_3)
\end{bmatrix} dudt_1du_3dt_3
\]

\[
\frac{1}{2} \int_{\Omega} \begin{bmatrix}
K_n(u_1, u_1) & K_n(u_1, t_1) & 0 & 0 \\
0 & K_n(t_1, t_1) & K_n(u_3, u_3) & 0 \\
K_n(t_3, u_1) & 0 & K_n(t_3, u_3) & 0 \\
0 & 0 & K_n(t_3, u_3) & K_n(t_3, t_3)
\end{bmatrix} dt_1du_3dt_3
\]

Notice that \( g(t_{i_1}, t_{j_1})g(t_{i_3}, t_{j_3}) \) takes only the values 0 and \( \frac{1}{4} \). Instead of including \( g(t_{i_1}, t_{j_1})g(t_{i_3}, t_{j_3}) \) in the integrand, we incorporate the constant \( \frac{1}{4} \) and restrict the region of integration:

\[
\Omega = \Omega_{I,(i_1,j_2|j_1,i_2|j_3)} = \left\{ (u_1, t_1, u_3, t_3) \mid \begin{array}{c}
\frac{u_1 + t_1}{2} \in I \\
|u_1 - t_1| < \gamma \\
\frac{u_3 + t_3}{2} \in I \\
|u_3 - t_3| < \gamma
\end{array} \right\}
\]

We see for the first time the meaning of clustering the equivalence classes. The variables \( u_1, t_1, u_3, t_3 \) each represent one of the equivalence classes in \((i_1, j_2|j_1, i_2|j_3)\). The equivalence classes corresponding to \( u_1 \) and \( t_1 \) are in the same cluster and so the region of integration \( \Omega \) carries the restriction that \(|u_1 - t_1|\) is small. Similarly, \(|u_3 - t_3|\) is small because the equivalence classes \( i_3 \) and \( j_3 \) are in the same cluster.

### 4.3. The main contribution from the clean collapses.

Of the 24 = 4! terms in the determinant, the ones which make a significant contribution to the integral are the four terms on the block diagonal. Their contribution is:

\[
\frac{1}{2} \left( \int_{\Omega} \begin{bmatrix}
K_n(u, u) & K_n(u, t) \\
K_n(t, u) & K_n(t, t)
\end{bmatrix} dt \right)^2
\]

\[
= \frac{1}{2} E(G_{\gamma,I,n})^2
\]

\[
= \frac{1}{2} \left( \frac{|I_n|^4}{144\pi^2} \right)^2 \left( 1 + O \left( \frac{1}{n} + n^2 \gamma^2 + (n^2 \gamma^2)^3 \right) \right)
\]

For a moment, let us take for granted that the block-diagonal terms of the clean collapses are the dominant contributions to \( E(G_{\gamma,I,n}^k) \). Using identical
techniques to the ones above, the block diagonal terms of a clean collapse of \((i_1|j_1) \wedge (1_2|j_2) \wedge \cdots \wedge (i_k|j_k)\) into \(l\) clusters contributes

\[
\left(\frac{1}{2}\right)^{k-l} E(G_{\gamma,I,n})^l.
\]

According to Lemma 4.1, for every partition of \(\{1,2,\ldots,k\}\) into \(l\) equivalence classes there correspond \(2^{k-l}\) clean collapses with \(l\) clusters. Thus each partition of \(\{1,2,\ldots,k\}\) into \(l\) equivalence classes contributes \(E(G_{\gamma,I,n})^l\) to the expected value \(E(G_{\gamma,I,n}^k)\).

In the following Subsection, we show that each partition of \(\{1,2,\ldots,k\}\) contributes \(\mu^l\) to the moment \(E(G_{\gamma,I,n}^k)\). Taking for granted that the block diagonal terms of clean collapses are the main contribution and expecting other error terms to sneak in, we have recovered the conclusions of Theorem 2.1.

4.4. Moments of the Poisson Distribution. The moments of a Poisson distribution are given by polynomials in the first moment with non-negative integer coefficients:

\[
E(G_{\gamma,I,n}^k) = \mu^k
\]

\[
E(G_{\gamma,I,n}^{k+1}) = \mu E(G_{\gamma,I,n}^k) = \mu^2 + \mu^3
\]

\[
E(G_{\gamma,I,n}^{k+2}) = \mu^2 E(G_{\gamma,I,n}^k) = \mu^3 + 3\mu^2 + \mu^4
\]

\[
E(G_{\gamma,I,n}^{k+3}) = \mu^3 E(G_{\gamma,I,n}^k) = \mu^4 + 7\mu^3 + 6\mu^4 + \mu^5
\]

\[
E(G_{\gamma,I,n}^{k+i}) = \sum_{j=1}^{i} a_{ij} \mu^j
\]

**Lemma 4.2.** The \(a_{i,k}\) satisfy the following recurrence relation and initial conditions:

\[
a_{1,1} = 1
\]

\[
a_{1,k} = 0 \quad \text{for } k \neq 1
\]

\[
a_{i,k} = ka_{i-1,k} + a_{i-1,k-1} \quad \text{for } i > 1.
\]

The \(a_{i,j}\) have a combinatorial description: \(a_{i,k}\) is the number of ways of partitioning \(\{1,2,3,\ldots,i\}\) into \(k\) nonempty subsets.

**Proof.** Define the \(a_{i,k}\) by the stated recurrence relation and initial conditions.

Suppose that \(X\) is Poisson with mean \(\mu\). Then

\[
Pr(X = k) = \frac{\mu^k}{k!} e^{-\mu}.
\]

Using the above formula one sees that

\[
\mu^j = E(X(X-1)(X-1)\ldots(X-j+1))
\]
Let the coefficients $b^{ji}$ be the result of expanding the above in terms of moments:

$$\mu^j = \sum_i b^{ji} E(G_{\gamma,I,n}^i)$$

The $b^{ji}$ are specified by the following recursion and initial conditions:

$$b^{1,1} = 1$$
$$b^{1,i} = 0 \quad \text{for } j \neq 1$$
$$b^{j,i} = b^{j-1,i-1} - (j - 1)b^{j-1,i}$$

The following is true for $k = 1, j \geq 1$, and for $k \geq 1, j = 1$:

$$\sum_i b^{k,i}a_{i,j} = \delta^k_j$$

Assume for the purposes of induction that the above is true for all smaller values of $j$ and $k$. The following symbolic manipulations, which use the Einstein summation convention, comprise the inductive step.

$$b^{j,i}a_{i,k} = (b^{j-1,i-1} - (j - 1)b^{j-1,i}) a_{i,k} (ka_{i-1,k} + a_{i-1,k-1})$$
$$= b^{j-1,i-1} (ka_{i-1,k} + a_{i-1,k-1}) - (j - 1)b^{j-1,i}a_{i,k}$$
$$= b^{j-1,i-1}a_{i-1,k-1} + kb^{j-1,i-1}a_{i-1,k} - (j - 1)\delta^j_{k-1}$$
$$= \delta^j_{k-1} + k\delta^j_{k-1} - (j - 1)\delta^j_{k-1}$$
$$= \delta^j_k$$

We have now shown that

$$\sum_{i=1}^k b^{k,i}a_{i,j} = \delta(j = k)$$

for all $k \geq 1$. This equation uniquely determines the $a_{i,j}$. This equation is satisfied if and only if

$$\mu^k = \sum b^{k,i}m_i$$
$$= E(X(X-1)(X-1)\ldots(X-k+1)),$$

where $m_i = \sum_{j=1}^i a_{ij} \mu^j$. □

5. The Errors Effecting Higher Moments of $G_{\gamma,I,n}$

There are two types of contributions to $E(G_{\gamma,I,n}^k)$ which we have not yet considered. We have yet to consider the off block diagonal terms of the clean collapses, and the mixed collapses. Since the block diagonal terms
of clean collapses contribute approximately \( E(G^n_k) \), which is what we want, these remaining terms are “errors”. Our main tool in bounding these errors will be Lemma 3.1.

5.1. Controlling the Off Block Diagonal Terms. We now estimate the off block diagonal terms. For this, we perform row and column operations on the matrix which leave the determinant unchanged. Specifically, we subtract each odd column from the following even column, and subtract each odd row from the following even row. In addition to leaving the determinant unchanged, this operation does not effect which terms in the determinant are on the block diagonal and which are off block diagonal.

$$
\frac{1}{8} \int_{\Omega} \begin{bmatrix}
\frac{n}{2\pi} & K_n(u_1, t_1) & K_n(u_1, u_3) & K_n(u_1, t_3) \\
K_n(t_1, u_1) & \frac{n}{2\pi} & K_n(t_1, u_3) & K_n(t_1, t_3) \\
K_n(u_3, u_1) & K_n(u_3, t_1) & \frac{n}{2\pi} & K_n(u_3, t_3) \\
K_n(t_3, u_1) & K_n(t_3, t_1) & K_n(t_3, u_3) & \frac{n}{2\pi}
\end{bmatrix} du_1 dt_1 du_3 dt_3
$$

$$
= \frac{1}{8} \int_{\Omega} \begin{bmatrix}
\frac{n}{2\pi} K_n(t_1, u_1) - \frac{n}{2\pi} K_n(u_3, u_1) & 2n/\pi - K_n(u_1, t_1) - K_n(t_1, u_1) & K_n(u_1, t_3) - K_n(u_1, u_3) & K_n(u_1, t_3) - K_n(u_1, t_3) - K(t_3, t_1) - K(t_3, u_1) \\
k(t_1, u_3) - k(t_3, u_1) & k(t_3, t_1) + k(u_3, u_1) - K(u_3, t_1) - K(u_3, t_3) & K_n(u_3, t_3) - \frac{n}{2\pi} - K_n(u_3, t_3) - K_n(t_3, u_3)
\end{bmatrix} du_1 \ldots dt_3
$$

To estimate the individual entries in the matrix in the integrand, we use Lemma 3.1 in conjunction with a Taylor series expansion of \( K_n \).

5.2. Dividing \( \Omega \) into Two Regions. We first divide the region \( \Omega \) into two regions. In the first region \( \Omega_\alpha \), at least one pair of the variables \( u_1 \) and \( u_3 \) are closer than \( \alpha \), where \( \alpha >> (\frac{1}{n}) \). In the latter region \( \Omega_\bar{\alpha} \), no pair of variables among \( u_1 \) and \( u_3 \) are closer than \( \alpha \):

\[
\Omega_\alpha = \begin{cases}
(u_1, t_1, u_3, t_3) & \frac{u_1 + t_1}{2} \in I \\
& |u_1 - t_1| < \gamma \\
& \frac{u_1 + t_1}{2} \in I \\
& |u_3 - t_3| < \gamma \\
& |u_1 - u_3| < \alpha
\end{cases}
\]

\[
\Omega_\bar{\alpha} = \begin{cases}
(u_1, t_1, u_3, t_3) & \frac{u_1 + t_1}{2} \in I \\
& |u_1 - t_1| < \gamma \\
& \frac{u_1 + t_1}{2} \in I \\
& |u_3 - t_3| < \gamma \\
& |u_1 - u_3| > \alpha
\end{cases}
\]
The size of the region $\Omega_\alpha$ is $O(|\alpha\gamma^2|)$. Using Lemma 3.1 in conjunction with the power series expansions of $K_n$, the entries in the matrix are:

$$
\frac{1}{8} \int_{\Omega_\alpha} \begin{bmatrix}
O(n) & O(n^2\gamma) & O(n) & O(n^2\gamma) \\
O(n^2\gamma) & O(n^3\gamma^2) & O(n^2\gamma) & O(n^3\gamma^2) \\
O(n) & O(n^2\gamma) & O(n) & O(n^2\gamma) \\
O(n^2\gamma) & O(n^3\gamma^2) & O(n^2\gamma) & O(n^3\gamma^2)
\end{bmatrix} \, du_1 \, dt_1 \, du_3 \, dt_3
$$

Each term in the determinant, and hence the determinant itself, is $O(n^5\gamma^4)$. Multiplying this by the size of the region of integration yields the total contribution of $O(n^8|I|\alpha\gamma^6)$ of off block diagonal terms in the region $\Omega_\alpha$.

The size of the region $\Omega_{\bar{\alpha}}$ is $O(|I|^2\gamma^2)$. Again using Lemma 3.1, the terms in the matrix are:

$$
\frac{1}{8} \int_{\Omega_{\bar{\alpha}}} \begin{bmatrix}
O(n) & O(n^2\gamma) & O\left(\frac{1}{\alpha}\right) & O\left(\frac{2n^3}{\alpha}\right) \\
O(n^2\gamma) & O(n^3\gamma^2) & O\left(\frac{2n^3}{\alpha}\right) & O\left(\frac{3n^3}{\alpha}\right) \\
O\left(\frac{1}{\alpha}\right) & O\left(\frac{2n^3}{\alpha}\right) & O(n) & O(n^2\gamma) \\
O\left(\frac{2n^3}{\alpha}\right) & O\left(\frac{3n^3}{\alpha}\right) & O(n^2\gamma) & O(n^3\gamma^2)
\end{bmatrix} \, du_1 \, dt_1 \, du_3 \, dt_3
$$

The off block terms in the determinant above are $O\left(\frac{n^{6\gamma^4}}{\alpha^2}\right)$. Multiplying this by the size of the region of integration yields $O\left(\frac{n^{6\gamma^4}|I|^2}{\alpha^2}\right)$.

Now we wish to choose $\alpha$ to minimize the total error from the off block diagonal terms. We use the method of dominant balance to choose $\alpha = \left(\frac{|I|}{n^2}\right)^{\frac{1}{3}}$, so that the two contributions from $\Omega_\alpha$ and $\Omega_{\bar{\alpha}}$ are comparable. The resulting contribution from the off block diagonal terms is

$$
O\left(\left(\frac{1}{|I|n}\right)^{\frac{2}{3}} \left(n^4|I|\gamma^3\right)^2\right).
$$

Let us generalize the above example to other clean collapses. Consider a clean collapse of $(i_1,j_1) \land (i_2,j_2) \land \cdots \land (i_k,j_k)$ into $l$ clusters. Using the same techniques as above, the contribution from $\Omega_\alpha$ is $O(n^{4l}|I|^{l-1}\alpha\gamma^{3l})$. The contribution from $\Omega_{\bar{\alpha}}$ is $O\left(\left(\frac{1}{\alpha n^3}\right)^2 \left(|I|^4\gamma^3\right)^l\right)$. Using the method of dominant balance, we again choose $\alpha = \left(\frac{|I|}{n^2}\right)^{\frac{1}{3}}$. The total contribution of the off-block-diagonal terms to a clean collapse with $l$ clusters is

$$
O\left(\left(\frac{1}{|I|n}\right)^{\frac{2}{3}} \left(n^4|I|\gamma^3\right)^l\right).
$$

### 5.3. Controlling the Mixed Collapses.

Let us consider one of the mixed collapses from $G_3^{ij,n}$:

$$(i_1,j_2,i_3|j_1|i_2,j_3).$$
The contribution to $E(G^3_{\gamma, I, n})$ from this mixed collapse may be written as an integral:

$$
E \left( \sum_{i_1, j_1, j_2 \text{ distinct}} g(t_{i_1}, t_{j_1})g(t_{i_2}, t_{i_1})g(t_{i_2}, t_{j_2}) \right)
$$

$$
= \int_{I^3} g(u_1, t_1)g(u_2, u_1)g(u_1, u_2)
\times \left[ \begin{array}{ccc}
K_n(u_1, u_1) & K_n(u_1, t_1) & K_n(u_1, u_2) \\
K_n(t_1, u_1) & K_n(t_1, t_1) & K_n(t_1, u_2) \\
K_n(u_2, u_1) & K_n(u_2, t_1) & K_n(u_2, u_2)
\end{array} \right] du_1 dt_1 du_2
$$

$$
= \frac{1}{8} \int_{\Omega} \left[ \begin{array}{ccc}
K_n(u_1, u_1) & K_n(u_1, t_1) & K_n(u_1, u_2) \\
K_n(t_1, u_1) & K_n(t_1, t_1) & K_n(t_1, u_2) \\
K_n(u_2, u_1) & K_n(u_2, t_1) & K_n(u_2, u_2)
\end{array} \right] du_1 dt_1 du_2,
$$

where $\Omega$ is the region:

$$
\Omega = \Omega_{I,(i_1,i_2,i_3|i_1|i_2,j_3)} = \left\{ (u_1, t_1, u_2) \left| \begin{array}{c}
\frac{u_1 + t_1}{2} \in I \\
\frac{|u_1 - t_1|}{2} \in I \\
\frac{|u_1 + u_2|}{2} \in I \\
\frac{|u_1 - u_2|}{2} \in \gamma
\end{array} \right. \right\}.
$$

As before, we perform row and column operations. For each cluster, one of the equivalence classes is selected; in this case the equivalence class represented by the variable $u_1$. The $u_1$ column is then subtracted from the columns corresponding to the other equivalence classes in its cluster. Similarly, the $u_1$ row is subtracted from the others equivalence classes in its cluster. We then use Lemma 3.1 to estimate the terms in the matrix.

$$
\frac{1}{8} \int_{\Omega} \left[ \begin{array}{ccc}
K_n(u_1, u_1) & K_n(u_1, t_1) & K_n(u_1, u_2) \\
K(t_1, u_1) - K(u_1, u_1) & K(t_1, t_1) + K(u_1, u_1) - K(t_1, t_1) - K(t_1, u_1) \\
K(u_2, u_1) - K(u_1, u_1) & K(u_2, t_1) + K(u_1, u_1) - K(u_1, t_1) - K(u_2, u_1)
\end{array} \right] du_1 dt_1 du_2
$$

$$
= \frac{1}{8} \int_{\Omega} \left[ \begin{array}{ccc}
O(n) & O(n^2) & O(n^2) \\
O(n^2) & O(n^3) & O(n^3) \\
O(n^2) & O(n^3) & O(n^3)
\end{array} \right] du_1 dt_1 du_2.
$$

The size of each term in the determinant is $O(n^7 \gamma^4)$. The size of the region of integration is $O(|I| \gamma^2)$, so the total contribution from this mixed collapse is $O(|I| n^{7} \gamma^{6})$.

To generalize, consider a mixed collapse of $(i_1 | j_1) \wedge (i_2 | j_2) \wedge \cdots \wedge (i_k | j_k)$ with $l_1$ equivalence classes and $l_2$ clusters. The size of the region of integration is
\( \mathcal{O}(|I|^{l_2}) \). The size of each term in the determinant is \( \mathcal{O}(n^{l_1}(n\gamma)^{2(l_1-l_2)}) \).
The total contribution is
\[
\mathcal{O}(|I|^{l_2}n^{-3l_2}n^{3l_1-2l_2}) = \mathcal{O}\left(\frac{\mu^{l_1-l_2}}{(n|I|)l_1-2l_2}\right),
\]
where
\[
\mu = \frac{|I| n^4\gamma^3}{2\pi^2 72\pi}.
\]

The mixed collapses of \((i_1|j_1) \land (1_2|j_2) \land \cdots \land (i_k|j_k)\) have the constraints \(2l_2 < l_1 < 2k\). Thus \(l_1 - 2l_2 \geq 1\). Since \(l_1 \geq 2\),
\[
l_1 - l_2 = \frac{1}{2}l_1 + \frac{1}{2}(l_1 - 2l_2) \geq \frac{3}{2} \geq 1.
\]
Considering the expansion of \((i_1|j_1) \land (1_2|j_2) \land \cdots \land (i_k|j_k)\), one sees that \(l_1 - l_2 \leq k\). Since
\[
1 \leq l_1 - l_2 \leq k,
\]
the main contribution to \(E(G_{\gamma,I,n})\) contains contributions comparable to \(\mu^{l_1-l_2}\). Relative to the main term, the contribution from this, or any, mixed collapse is
\[
\mathcal{O}\left(\frac{1}{n|I|}\right).
\]

6. Review of Case 1: The Circular Unitary Ensemble

There are several sources of error in the estimation of \(E(G_{\gamma,I,n}^k)\). We summarize these sources in a table. In this table, \(\mu = \frac{|I| n^4\gamma^3}{144\pi^2}\).

| Description | Contribution | Relative Size |
|-------------|--------------|--------------|
| Main term: | \(\sum_{j=1}^{k} a_{k,j}\mu^j\) | 1 |
| Approximation \(E(G_{\gamma,I,n}) \approx \mu\) used to estimate main term | \(\mu^l \mathcal{O}\left(n^{-1} + n^2\gamma^2 + (n^2\gamma^2)^3\right)\) | \(\mathcal{O}\left(n^{-1} + n^2\gamma^2 + (n^2\gamma^2)^3\right)\) |
| Clean collapse with \(l\) clusters, off-block diagonal terms \(\Omega_\alpha\) | \(\mathcal{O}\left(n^{3l}l^{-1}\alpha\gamma^{3l}\right)\) | \(\mathcal{O}\left(\frac{\alpha}{7}\right)\) |
| Clean collapse with \(l\) clusters, off-block diagonal terms \(\Omega_\delta\) | \(\mathcal{O}\left(\left(\frac{1}{\alpha n}\right)^2 (In^4\gamma^3)^l\right)\) | \(\mathcal{O}\left(\frac{1}{\alpha^2 n^2}\right)\) |
| Clean collapse with \(l\) clusters, total of off-block diagonal terms, choosing \(\alpha = \left(\frac{l}{In}\right)^{\frac{1}{2}}\) | \(\mathcal{O}\left(\left(\frac{1}{n}\right)^{\frac{2}{3}} (In^4\gamma^3)^l\right)\) | \(\mathcal{O}\left(\frac{1}{n^2}\right)\) |
| Mixed collapse, \(l_1\) equivalence classes, and \(l_2\) clusters | \(\mathcal{O}\left(\frac{\mu^{l_1-l_2}}{n|I|}\right)\) | \(\mathcal{O}\left(\frac{1}{n|I|}\right)\) |
7. Case 2: Universal Unitary Ensembles

Let $V(x)$ be a potential which is real analytic on $\mathcal{R}$ and has sufficient growth at $\infty$:

$$\lim_{|x| \to \infty} \frac{V(x)}{\log(x^2 + 1)} = \infty.$$ 

The universal unitary ensemble $UUE_n$, with potential $V(x)$, is the set of $n \times n$ Hermitian matrices $M = (m_{ij})$ with joint probability density function

$$(\text{const}) \ e^{-n \sum V(\lambda_j)} dM$$

The joint probability density function for the eigenvalues is obtained from the joint probability density function for the matrix entries using Weyl integration. See Appendix [13].

$$(\text{const}) \prod_{i<j} (\lambda_i - \lambda_j)^2 e^{-n \sum V(\lambda_i)} d\Lambda$$

As discussed in Section [11.1] we assume that the potential is regular and that the equilibrium measure $\Psi(x)dx$ is supported on a single interval $[a, b]$.

7.1. Heuristic Prediction for the Universal Ensemble. As in the case of CUE, we again make the simplifying but false assumption that the consecutive spacings are independent random variables. Thus $G_{\gamma,I,n}$, as a sum of many unlikely events, will be approximately Poisson. We now estimate its mean.

At a point $x \in I$, the local density of eigenvalues is given by the diagonal of the projection kernel:

$$\text{density} = K_n(x, x).$$

At the point $x$, a separation of $\gamma$ is equal to $\gamma K_n(x, x)$ times the local mean spacing. Thus the probability of any one of these spacings being less than $\gamma$ is approximately

$$\left( \frac{\pi^2}{9} \right) (\gamma K_n(x, x))^3.$$

In the region $[x, x + dx]$, the number of consecutive spacings is approximately $K_n(x, x)dx$. Integrating over $x \in I$, the expected value of $G_{\gamma,I,n}$ is approximately:

$$E(G_{\gamma,I,n}) \approx \frac{\pi^2 \gamma^3}{9} \int_I K_n(x, x)^4 dx.$$

For $x \in [a + \epsilon, b - \epsilon]$, $K_n(x, x) = n \Psi(x)(1 + O(n^{-1}))$. See [DKM'97]. Thus, the heuristic prediction agrees with the Moment Estimation Theorem.
As in the case of CUE, \(G_{\gamma,I,n}\) is the symmetrization of a function of 2 variables:

\[
g_2(x,y) = \begin{cases} 
\frac{1}{2} & \text{if } |x-y| < \gamma \quad \left(\frac{x+y}{2}\right) \in I_n \\
0 & \text{otherwise}
\end{cases}
\]

\(G_{\gamma,I,n}(t_1,t_2,\ldots,t_n) = \sum_{i\neq j} g_2(t_i,t_j).\)

Using the method of Gaudin (see Appendix 14), the expected value of \(G_{\gamma,I,n}\) is

\[
E(G_{\gamma,I,n}) = \int \int_{\mathbb{R}^2} g_2(u,t) \begin{bmatrix} K_n(u,u) & K_n(u,t) \\
K_n(t,u) & K_n(t,t) \end{bmatrix} \, dudt
\]

\[
= \frac{1}{2} \int \int_{\Omega} \begin{bmatrix} K_n(u,u) & K_n(u,t) \\
K_n(t,u) & K_n(t,t) \end{bmatrix} \, dudt.
\]

In the above formula, the region \(\Omega\) is the set where \(g_2(t_1,t_2) = \frac{1}{2}:\)

\[
\Omega = \left\{(u,t) \quad \text{s.t.} \quad \frac{u+t}{2} \in I, \quad |t-u| < \gamma \right\}.
\]

The projection kernel \(K_n(x,y)\) is defined in Section 14.

Compared to the CUE\(_0\) case, it is now slightly more difficult to estimate \(E(G_{\gamma,I,n})\) because the kernel \(K_n(x,y)\) no longer depends solely on the difference \((y-x)\).

The region \(\Omega\) is narrow in the \((t-u)\) direction. This suggests expanding the integrand in a Taylor series. Let \(x = \frac{u+t}{2}\) and \(y = \frac{u-t}{2}\). Then \(dxdy = \frac{1}{2}dudt\). In terms of the new variables \(x\) and \(y\), the first moment is:

\[
E(G_{\gamma,I,n}) = \int \int_{I} \left( y^2 \left( KK^{\gamma^2} - KK^{\gamma^2} - \frac{2}{5} K^{\gamma^2} \right) \right) \, dxdy,
\]
where

\[
K^{\triangledown}(x, y) = \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) K(x, y)
\]

\[
K^{\triangledown\triangledown}(x, y) = \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) K(x, y)
\]

\[
K^{\searrow\searrow}(x, y) = \left( \frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right) \left( \frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right) K(x, y)
\]

and \( c_j \) is the maximum over \( \Omega \) of any \((j)\)th partial derivative of \( K_n(x, y) \).

Using the Lemma 12.2, the integral of \( O\left(y^4 c_0 c_4 \right) \) over \( \Omega \) is \( O\left(|I_n| \gamma^5 n^6 \right) \), the integral of \( O\left(y^5 c_1 c_4 \right) \) is \( O\left(|I_n| \gamma^6 n^7 \right) \), and so on, and the integral of \( O\left(y^8 c_4 c_4 \right) \) is \( O\left(|I_n| \gamma^9 n^{10} \right) \). Dropping the redundant error terms,

\[
E(G_{\gamma, I, n}) = \int_I \frac{\gamma^3}{12} \left( K K^{\triangledown\triangledown} - K K^{\searrow\searrow} - (K^{\triangledown})^2 \right) dx + O\left(|I_n| \gamma^5 n^6 + |I_n| \gamma^9 n^{10} \right).
\]

We first evaluate the derivatives \( K^{\triangledown} \), \( K^{\triangledown\triangledown} \), and \( K^{\searrow\searrow} \) symbolically as a function of \( x \) and \( y \). Let

\[
[i, j, k] = b_{n-1} \frac{\eta^{(i)}_n(x) \eta^{(j)}_{n-1}(y) - \eta^{(j)}_n(y) \eta^{(i)}_{n-1}(x)}{(x - y)^k}.
\]

where \( \eta_j(x) = \phi_j(x) e^{-\frac{n}{2} V(x)} \), and \( \phi_j \) is the \((j)\)th normalized orthogonal polynomial with respect to \( e^{-n V(x)} dx \). In terms of this new notation, we
have the following expressions for $K(x, y)$ and its derivatives:

\[
K = [0, 0, 1] \\
\frac{\partial}{\partial x} K = [1, 0, 1] - [0, 0, 2] \\
\frac{\partial}{\partial y} K = [0, 1, 1] + [0, 0, 2] \\
\frac{\partial^2}{\partial x^2} K = [2, 0, 1] - 2[1, 0, 2] + 2[0, 0, 3] \\
\frac{\partial^2}{\partial x \partial y} K = [1, 1, 1] - [0, 1, 2] + [1, 0, 2] - 2[0, 0, 3] \\
\frac{\partial^2}{\partial y^2} K = [0, 2, 1] + 2[0, 1, 2] + 2[0, 0, 3] \\
K^\nearrow = \left( \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right) K \\
= [1, 0, 1] + [0, 1, 1] \\
K^{\searrow\nearrow} = \left( \frac{\partial^2}{\partial x^2} + 2 \frac{\partial^2}{\partial x \partial y} + \frac{\partial^2}{\partial y^2} \right) K \\
= [2, 0, 1] + 2[1, 1, 1] + [0, 2, 1] \\
K^{\nwarrow\nearrow} = \left( \frac{\partial^2}{\partial x^2} + 2 \frac{\partial^2}{\partial x \partial y} + \frac{\partial^2}{\partial y^2} \right) K \\
= ([2, 0, 1] - 2[1, 1, 1] + [0, 2, 1]) + (-4[1, 0, 2] + 4[0, 1, 2]) + 8[0, 0, 3].
\]

We wish to evaluate the above quantities $K^\nearrow$, $K^{\searrow\nearrow}$, and $K^{\nwarrow\nearrow}$ along the diagonal $x = y$. At first we are taken aback by the presence of $(x - y)$ in the denominator, but then recall that $K_n(x, y)$ is smooth everywhere, including the diagonal. We determine the limiting values along the diagonal by use of
L'Hopital's rule, which in the case of \( K^{\nearrow\searrow} \) must be applied three times:

\[
\begin{align*}
K^{\nearrow}(x, x) & = [200] + [110] \\
K^{\nearrow\nearrow}(x, x) & = [300] + [120] + 2[210] \\
K^{\searrow\searrow}(x, x) & = (x - y)^2([2, 0, 3] - 2[1, 1, 3] + [0, 2, 3]) + (x - y)(-4[1, 0, 3] + 4[0, 1, 3]) + 8[0, 0, 3] \\
& = \frac{1}{3}([3, 0, 0] - 2[2, 1, 0] + [1, 2, 2]) \\
& +\frac{1}{3}((x - y)(-2[2, 0, 2] + 2[0, 2, 2]) + (4[1, 0, 2] + 4[0, 1, 2])) \\
& = \frac{1}{3}([3, 0, 0] - 2[2, 1, 0] + [1, 2, 0]) \\
& +\frac{1}{6}((x - y)(-2[3, 0, 1] + 2[1, 2, 1]) \\
& +(-2[2, 0, 1] + 2[0, 2, 1]) \\
& +4[2, 0, 1] + 4[1, 1, 1]) \\
& = \left(\frac{2}{3}[2, 1, 0] + \frac{2}{3}[1, 2, 0]\right) + \left(\frac{1}{3}[2, 0, 1] + \frac{2}{3}[1, 1, 1] + \frac{1}{3}[0, 2, 1]\right) \\
& = \left(\frac{2}{3}[2, 1, 0] + \frac{2}{3}[1, 2, 0]\right) + \left(\frac{1}{3}[3, 0, 0] + \frac{2}{3}[2, 1, 0] + \frac{1}{3}[1, 2, 0]\right) \\
& = \frac{1}{3}[3, 0, 0] + [1, 2, 0].
\end{align*}
\]

and substitute the results into the integrand:

\[
\left(KK^{\nearrow\nearrow} - KK^{\searrow\searrow} - (K^{\nearrow})^2\right) \\
= \left([1, 0, 0]([3, 0, 0] + [2, 1, 0]) - [1, 0, 0]\left(\frac{1}{3}[3, 0, 0] - [2, 1, 0]\right) - [2, 0, 0]^2\right) \\
= \left(\frac{2}{3}[1, 0, 0][3, 0, 0] + 2[1, 0, 0][2, 1, 0] - [2, 0, 0]^2\right).
\]
We derive estimates for \([3,0,0]\) and then state without proof the analogous estimates for \([1,0,0], [2,1,0], \text{and} [2,0,0].\)

\[
[3,0,0] = \left(\frac{b-a}{4} + O\left(\frac{1}{n}\right)\right) \left(\eta_n^{(3)}(x)\eta_{n-1}(x) - \eta_n(x)\eta_{n-1}^{(3)}(x)\right)
\]

\[
= \left(\frac{b-a}{4}\right) (n\pi\Psi(x))^3 \frac{2}{(b-a)\pi} \left[\Re(-ie^{i\theta(x)}u(x))\Re(e^{i\theta(x)}v(x)) - \Re(e^{i\theta(x)}u(x))\Re(-ie^{i\theta(x)}v(x)) + O\left(\frac{1}{n}\right)\right]
\]

\[
= \frac{(n\pi\Psi(x))^3}{2\pi} |u(x)||v(x)| \left[\sin(\theta + \alpha)\cos(\theta + \beta) - \cos(\theta + \alpha)\sin(\theta + \beta) + O\left(\frac{1}{n}\right)\right]
\]

\[
= \frac{(n\pi\Psi(x))^3}{2\pi} \left(\sqrt{\frac{b-x}{x-a}} + \sqrt{\frac{x-a}{b-x}}\right) (\sin(\alpha - \beta) + O(n^{-1}))
\]

\[
= \frac{(n\pi\Psi(x))^3}{2\pi} \left(\sqrt{\frac{b-x}{x-a}} + \sqrt{\frac{x-a}{b-x}}\right) \left(\Im \left(\frac{u(x)}{v(x)}\right) + O(n^{-1})\right)
\]

\[
= \frac{(n\pi\Psi(x))^3}{\pi} (1 + O(n^{-1})).
\]

where

\[
\theta(x) = n\pi\int_x^b \Psi(s)ds
\]

\[
u(x) = e^{i\frac{\pi}{4}} \left(\frac{b-x}{x-a}\right)^{\frac{1}{4}} + e^{-i\frac{\pi}{4}} \left(\frac{x-a}{b-x}\right)^{\frac{1}{4}}
\]

\[
v(x) = e^{-i\frac{\pi}{4}} \left(\frac{b-x}{x-a}\right)^{\frac{1}{4}} + e^{i\frac{\pi}{4}} \left(\frac{x-a}{b-x}\right)^{\frac{1}{4}}
\]

\[
\alpha(x) = \arg(u(x))
\]

\[
\beta(x) = \arg(v(x)).
\]

Using the same techniques,

\[
[2,1,0] = \frac{(n\pi\Psi(x))^3}{\pi} (-1 + O(n^{-1}))
\]

\[
[2,0,0] = \frac{(n\pi\Psi(x))^2}{\pi} (0 + O(n^{-1}))
\]

\[
[1,0,0] = \frac{(n\pi\Psi(x))}{\pi} (-1 + O(n^{-1})).
\]
Returning to our expression for $E(G_{\gamma,I,n})$,
\[
E(G_{\gamma,I,n}) = O(|I|^5 n^6 + |I|^9 n^{10}) + \int \frac{\gamma^3}{12} \left( KK^{\wedge,\wedge} - KK^{\vee,\vee} - (K^{\wedge})^2 \right) dx
\]
\[
= O(|I|^5 n^6, |I|^9 n^{10}) + \int \frac{\gamma^3}{12} \left( \frac{2}{3} (-1)(1) + 2(-1)(-1) - 0^2 + O(n^{-1}) \right) dx
\]
\[
= O(|I|^5 n^6 + |I|^9 n^{10}) + \int \frac{\pi^2 \gamma^3 n^4 \Psi(x)^4}{9} (1 + O(n^{-1})) dx
\]
\[
= \frac{\pi^2 \gamma^3 n^4}{9} \int \Psi(x)^4 dx \left( 1 + O \left( n^{-1} + \gamma^2 n^2 + (\gamma^2 n^2)^3 \right) \right).
\]

These calculations confirm that, for the first moment $E(G_{\gamma,I,n})$, the conclusions of Theorem 1.4 are correct. We now estimate the higher moments.

9. Higher Moments in the Universal Case

As in the case of a random unitary matrix, we express $E(G_{\gamma,I,n})$ as a sum of several terms, with each term corresponding to a collapse of
\[
(i_1|j_1) \land (i_2|j_2) \land (i_3|j_3).
\]

Using the method of Gaudin, each term is expressed as an integral involving only a few variables. As before, the main contribution will come from the block-diagonal terms of the clean collapses:
\[
\text{(clean collapse, block diagonal)} = \sum_{j=1}^{k} a_{k,j} E(G_{\gamma,I,n})^j
\]
\[
= \sum_{j=1}^{k} a_{k,j} \mu^j \left( 1 + O(n^{-1} \gamma^2 n) \right),
\]
where
\[
\mu = \frac{\pi^2 \gamma^3 n^4}{9} \int \Psi(t)^4 dt.
\]

The errors in this approximation have the same sources as in the CUE case:

- the approximation $E(G_{\gamma,I,n}) = \mu$ used in estimating the main term.
- clean collapse, off-block diagonal terms.
- mixed collapses.

The bounds for these errors are obtained using exactly the same reasoning as in the CUE case. The table from the CUE case applies here with only minor changes. The only differences are that
• The value of $\mu$ has changed.
• We now use Lemma 10.1 instead of Lemma 3.1 to bound $K_n(x, y)$ and its derivatives.

10. Controlling the Derivatives of $K_n(x, y)$ in the Universal Case

For bounding the error terms for higher moments in the UUE case, we require estimates for $K_n(x, y)$ and its derivatives. In order to apply the Christoffel-Darboux formula from Lemma 14.2, we approximate the coefficient $b_{n-1}$. We refer the reader to [DKM+97] and state the result and error term without proof. This approximation is valid when the potential $V(x)$ is regular and the equilibrium measure is supported on a single interval $[a, b]$.

\[
b_{n-1} = \int_{-\infty}^{\infty} x \phi_n(x) \phi_{n-1}(x) e^{-nV(x)} dx
\]
\[
= \int_{-\infty}^{\infty} x \eta_n(x) \eta_{n-1}(x) dx
\]
\[
= \int_{a}^{b} x \eta_n(x) \eta_{n-1}(x) dx + O\left(\frac{1}{n}\right)
\]
\[
= \int_{a}^{b} \frac{2}{(b-a)\pi} \left( -\frac{1}{2} \left( \frac{b-x}{x-a} \right)^{\frac{1}{2}} + \frac{1}{2} \left( \frac{x-a}{b-x} \right)^{\frac{1}{2}} \right) dx + O\left(\frac{1}{n}\right)
\]
\[
= \frac{(b-a)}{4} + O\left(\frac{1}{n}\right).
\]

The projection kernel is

\[
K_n(x, y) = b_{n-1} \frac{\eta_n(x) \eta_{n-1}(y) - \eta_n(y) \eta_{n-1}(x)}{x-y}.
\]

**Lemma 10.1.** Suppose that the potential $V(x)$ is regular and that it’s equilibrium measure $\Psi(x)dx$ is supported on the interval $[a, b]$. Fix $\epsilon > 0$ and let $R$ be the region $[a + \epsilon, b - \epsilon] \times [a + \epsilon, b - \epsilon]$. For $(x, y) \in R$, a mixed partial derivative of $K_n(x, y)$ of total degree $k \geq 0$ is $O\left(\frac{1}{n^{k+1}}\right)$. When $|x - y| \geq \frac{1}{n}$, we have the stronger bound $O\left(\frac{1}{|x-y|^{k+1}}\right)$.

**Proof.** Take $k$ partial derivatives of $K_n(x, y)$ symbolically. The result is a finite number of terms of the form

\[
b_{n-1}c_{k_1,k_2,k_3} \frac{\eta_n^{(k_1)}(x) \eta_{n-1}^{(k_2)}(y) - \eta_n^{(k_2)}(y) \eta_{n-1}^{(k_1)}(x)}{(x-y)^{1+k_3}},
\]

where $k_1 + k_2 + k_3 = k$, the coefficients $c_{k_1,k_2,k_3}$ are integers which do not depend on $n$, and a superscript in parenthesis indicates multiple differentiation.
We divide the region $R$ into two regions:
\[
R_{\text{diag}} = \left\{ (x, y) \in R \mid |x - y| < \frac{1}{n} \right\}
\]
\[
R_{\text{bulk}} = \left\{ (x, y) \in R \mid |x - y| \geq \frac{1}{n} \right\}
\]

In the region $R_{\text{bulk}}$, $\eta^{(k_1)}_n(x)$ and $\eta^{(k_1)}_{n-1}(x)$ are of size $O(n^{k_1})$; $\eta^{(k_2)}_n(y)$ and $\eta^{(k_2)}_{n-1}(y)$ are of size $O(n^{k_2})$; and $\frac{1}{(x-y)^{1+k_3}}$ is of size $O\left(\frac{1}{|x-y|^{1+k_3}}\right)$. The total contribution in the bulk region is then
\[
O\left(\frac{1}{|x-y|^{n^k}}\right).
\]

The region $R_{\text{diag}}$ is more subtle. Consider the term
\[
\frac{\eta^{(k_1)}_n(x)\eta^{(k_2)}_{n-1}(y) - \eta^{(k_2)}_n(y)\eta^{(k_1)}_{n-1}(x)}{(x-y)^{1+k_3}},
\]
We treat both $\eta^{(k_1)}_n(x)\eta^{(k_2)}_{n-1}(y)$ and $\eta^{(k_2)}_n(y)\eta^{(k_1)}_{n-1}(x)$ the same way, so let us discuss $\eta^{(k_1)}_n(x)\eta^{(k_2)}_{n-1}(y)$. We expand $\eta^{(k_2)}_{n-1}(y)$ in a Taylor series centered at $x$.
\[
\eta^{(k_2)}_{n-1}(y) = \eta^{(k_2)}_{n-1}(x) + \cdots + \frac{(y-x)^{k_3}}{k_3!}\eta^{(k_2+k_3)}_{n-1}(x) + \frac{(y-x)^{k_3+1}}{(k_3+1)!}\eta^{(k_2+k_3+1)}_{n-1}(\tilde{x}),
\]
for some $\tilde{x}$ between $x$ and $y$. Selecting the remainder term yields
\[
\frac{\eta^{(k_1)}_n(x)(y-x)^{k_3+1}}{(k_3+1)!}\eta^{(k_2+k_3+1)}_{n-1}(\tilde{x})\]
\[
= O(n^{k_1+k_2+k_3+1}) = O(n^{1+k}),
\]
which is within the desired bound. Selecting one of the other terms in the expansion of $\eta^{(k_2)}_{n-1}(y)$ yields
\[
\frac{\eta^{(k_1)}_n(x)(y-x)^j}{(y-x)^{1+k_3}}\eta^{(k_2+k_3)}_{n-1}(x)
\]
\[
= \frac{\eta^{(k_1)}_n(x)\eta^{(k_2+j)}_{n-1}(x)}{(y-x)^{1+k_3-j}},
\]
for some $0 \leq j \leq k_3$. Observe that this term is of the form
\[
e^{-nV(x)} \frac{p(x)}{(y-x)^{1+k_3-j}},
\]
where $p(x)$ is a polynomial in $x$, and the exponent $1 + k_3 - j$ in the denominator is positive. Upon adding all such terms in the Taylor expansion, for
every term in the symbolic differentiation of \( K_n(x, y) \), the result is
\[
\sum_{j=0}^{k_3} e^{-nV(x)} \frac{p_j(x)}{(y-x)^{1+k_3-j}}.
\]
Since \( K_n(x, y) = \sum_{i=0}^{n-1} \eta_i(x)\eta_i(y) \), it is smooth along the diagonal \( x = y \) and \( p_j(x) \) is identically zero for \( 0 \leq j \leq k_3 \).

11. Asymptotics for Orthogonal Polynomials with General Weights

We outline the derivation of the leading order asymptotics for the \( \eta_n(x) \) and \( \eta_{n-1}(x) \), where
\[
\eta_j(x) = \frac{\phi_j(x)}{e^{\frac{1}{2}V(x)}}
\]
and \( \phi_j(x) \) is the \( (j) \)th normalized orthogonal polynomial with respect to the measure \( e^{-nV(x)} \).

Our derivation follows the exposition of [DKM+97] and [Dei99] very closely, and omits the proofs of several facts which are proven in these articles. The strongest results in this direction are proven in [DKM+99a] and [DKM+99b]. The only departure from their presentation comes when we obtain bounds for the derivatives of \( L(z) \), and then use these bounds to derive the leading order asymptotics for the derivatives of \( \eta_n(x) \) and \( \eta_{n-1}(x) \).

The leading order asymptotics of the derivatives of \( \eta_n \) and \( \eta_{n-1} \) turn out to be the derivatives of the leading order asymptotics. Of course one expects this, but it is not something which can be taken for granted.

11.1. Facts About the Equilibrium Measure of a Potential. Let \( V(x) \) be a potential which is real analytic on \( \mathcal{R} \) and has sufficient growth at \( \infty \):
\[
\lim_{|x| \to \infty} \frac{V(x)}{\log(x^2 + 1)} = \infty
\]
This growth condition guarantees, in particular, that all the moments of \( e^{-nV(x)} \) are finite.

Given a probability measure \( \mu \) on \( \mathcal{R} \), let \( I^V \) be the energy functional:
\[
I^V(\mu) = \int_{\mathcal{R}} \int_{\mathcal{R}} \log |t-s|^{-1} d\mu(t) d\mu(s) + \int_{\mathcal{R}} V(t) d\mu(t)
\]

**Theorem 11.1.** Let \( V \) and \( I^V \) be as above. Then there exists a unique probability measure \( \mu = \mu^V \) such that
\[
E^V = \inf_{\mu \in M^1} I^V(\mu) = I^V(\mu^V),
\]
where the infimum is over all probability measures on $\mathcal{R}$. The equilibrium measure $\mu^V$ is compactly supported.  

Let the support of $\mu^V$ be the following disjoint union of intervals:

$$\text{Supp}(\mu) = \bigcup_{j=1}^{l} [a_j, b_j].$$

Now if

- $V(x)$ is real analytic
- $V(x)$ grows fast enough at infinity: $\lim_{|x| \to \infty} \frac{V(x)}{\log(x^2+1)} = \infty$,

then the equilibrium measure is $\Psi(x)dx$, where

$$\Psi(x) = R_+(x) \frac{1}{2} h(x)$$

$$R(x) = -\prod_{j=1}^{l} (z - a_j)(z - b_j),$$

and $h(x)$ is real analytic on $\mathcal{R}$. See [DKM98].

Since $\Psi$ is continuous one can use the calculus of variations to derive the Euler-Lagrange equations for the equilibrium measure:

**Theorem 11.2.** There is a constant $l \in \mathcal{R}$ such that the equilibrium measure $\mu^V$ satisfies the following conditions.

- $2 \int \log |x - y|^{-1} d\mu^V(y) + V(x) \geq l$ for all $x \in \mathcal{R}$.
- $2 \int \log |x - y|^{-1} d\mu^V(y) + V(x) = l$ for $x$ in the support of $\Psi$.

Conversely, if a compactly supported measure $\mu$ satisfies the above conditions for some $l$, then it is the equilibrium measure $\mu^V$.

If $\psi(x) > 0$ except at the endpoints of $J$, and we have strict inequality in the above theorem for $x \not\in J$, then we say that the potential $V(x)$ is regular. Otherwise we call the potential singular.

For simplicity, we consider only the case when $V(x)$ is regular and the support of $\psi$ is a single interval $J = [a, b]$. This is always the case when $V(x)$ is convex.

### 11.2. Uniqueness for $2x2$ Riemann Hilbert Problem.

Let $\Sigma$ a contour in $\mathcal{C}$, and $\Sigma_0$ be the same contour excluding the points of intersection. Suppose that $\nu(z) : \Sigma_0 \to GL(2, \mathcal{C})$, such that $\nu$ is smooth, bounded, and approaches $Id$ rapidly on the unbounded components of $\Sigma_0$. 

Lemma 11.3. Suppose that \( m : \mathbb{C} \setminus \Sigma \to GL(2, \mathbb{C}) \) satisfies the Riemann-Hilbert problem \((\Sigma, \nu)\) if

1. \( m \) is analytic in \( \mathbb{C} \setminus \Sigma \).
2. \( m_+(z) = m_-(z)\nu(z) \) for \( z \in \Sigma_0 \).
3. \( m(z) \to Id \) as \( z \to \infty \).

Assume further that \( \det(\nu) = 1 \). Then the solution of the Riemann-Hilbert problem is unique, if it exists.

Proof. We sketch the proof from p. 194-198 of [Dei99]. First, because \( \det(\nu) = 1 \), \( \det(m) \) is analytic across the contour. So \( \det(m) \) is an analytic function and, because of the behavior of \( m \) at infinity, \( \det(m) \) must be the constant 1. In particular \( m \) is always invertible.

Now suppose that \( m \) and \( \tilde{m} \) are two solutions. Then, using an algebraic trick specific to \( 2 \times 2 \) matrices with determinant 1, it turns out that \( H = \tilde{m}m^{-1} \) satisfies a Riemann-Hilbert problem with the same contour but \( \nu = Id \). In other words \( H \) is analytic, approaches \( Id \) at infinity, hence is identically equal to \( Id \), establishing uniqueness. \( \square \)

11.3. Expressing Orthogonal Polynomials as the Solution of Riemann Hilbert Problems. Let \( \pi_j(z) \) be the monic orthogonal polynomials with respect to the weight \( e^{-nV(x)}dx \).

Theorem 11.4 (Fokas, Its, Kitaev). Let \( Y(z) \) be the \( 2 \times 2 \) matrix-valued function satisfying the following RHP:

1. \( Y(z) \) is analytic in \( \mathbb{C} \setminus \mathbb{R} \).
2. \( Y(z) = (1 + O(z^{-1})) \begin{pmatrix} z^q & 0 \\ 0 & z^{-q} \end{pmatrix} \) as \( z \to \infty \), \( z \in \mathbb{C} \setminus \mathbb{R} \).
3. \( Y_+(z) = Y_-(z) \begin{pmatrix} 1 & e^{-nV(z)} \\ 0 & 1 \end{pmatrix} \) for \( z \in \mathbb{R} \).

Then \( Y \) encodes information about \( \pi_q(z) \), the \( (q) \)th monic orthogonal polynomial with respect to the measure \( e^{-nV(x)}dx \):

\[
Y(z) = \begin{pmatrix} \pi_q(z) & \int_{\mathbb{R}} \frac{\pi_q(s)e^{-nV(s)}ds}{s-z} \frac{1}{2\pi i} \\ \gamma_{q-1}\pi_{q-1}(z) & \gamma_{q-1}\int_{\mathbb{R}} \frac{\pi_{q-1}(s)e^{-nV(s)}ds}{s-z} \frac{1}{2\pi i} \end{pmatrix},
\]

where the constant \( \gamma_{q-1} \) is

\[
\gamma_{q-1} = \frac{-2\pi i}{\int_{\mathbb{R}} \pi_{q-1}^2(s)e^{-nV(s)}ds}.
\]
The $2 \times 2$ jump matrix in this Theorem has determinant 1, so the solution of the Riemann-Hilbert problem is unique if it exists. To prove this Theorem, one verifies that the given solution satisfies the Riemann-Hilbert problem.

11.4. Removing the Bulk of the Oscillatory Behavior. For $\Im(z) \neq 0$, let

$$g(z) = \int \log(z - s)\Psi(s)ds.$$ 

We choose $\log(z)$ to have a branch cut along the negative real axis, and to be real on the positive real axis.

Observe that $g(z)$ is analytic in the region $\mathcal{C} \setminus (-\infty, b_l]$. For real $z$, let $g_+(z)$ be the limiting value from above the branch cut, and $g_-(z)$ the limiting value from below. Then

$$g_\pm(z) = \int_a^b \log|z - s|\Psi(s)\,ds \pm \pi i \int_z^b \Psi(s)\,ds$$

$$g_+(z) + g_-(z) = 2 \int \log|z - s|\Psi(s)\,ds$$

$$g_+(z) - g_-(z) = 2\pi i \int_z^b \Psi(s)\,ds$$

Observe that because $\Psi(s)\,ds$ is a probability measure, $g_+(z) - g_-(z) = 2\pi i$ for $z < a$. Thus $e^{ng(z)}$ is analytic in $\mathcal{C} \setminus [a, b]$.

For $z \in \mathcal{C} \setminus \mathcal{R}$, let

$$m^{(1)}(z) = \begin{pmatrix} e^{\frac{ng(z)}{2}} & 1 \\ 0 & e^{-\frac{ng(z)}{2}} \end{pmatrix} \begin{pmatrix} Y(z)e^{-ng(z)} & e^{-\frac{ng(z)}{2}} \\ 0 & e^{\frac{ng(z)}{2}} \end{pmatrix} \begin{pmatrix} e^{\frac{ng(z)}{2}} & 1 \\ 0 & e^{-\frac{ng(z)}{2}} \end{pmatrix}$$

Then $m^{(1)}$ satisfies a simpler Riemann-Hilbert problem:

1. $m^{(1)}(z)$ is analytic in $\mathcal{C} \setminus \mathcal{R}$
2. $m_+(z) = m_+(z)\nu^{(1)}(z)$ for $z \in \mathcal{R}$
3. $m^{(1)}(z) = I + O(z^{-1})$ as $z \to \infty$,

where

$$\nu^{(1)}(z) = \begin{pmatrix} e^{\frac{ng(z)}{2} + ng_-(z)} & 1 \\ 0 & e^{-\frac{ng(z)}{2} - ng_-(z)} \end{pmatrix} \begin{pmatrix} 1 & e^{-nV(z)} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^{-ng_+(z) + \frac{nV(z)}{2}} & 1 \\ 0 & e^{ng_+(z) + \frac{nV(z)}{2}} \end{pmatrix}$$

$$= \begin{pmatrix} e^{n(g_-(z) - g_+(z))} & e^n(g_-(z) + g_+(z) - V(z) + \frac{1}{2}) \\ 0 & e^{n(g_+(z) - g_-(z))} \end{pmatrix}.$$
The upper right entry in the jump matrix $\nu$ is $e^{n(g_{+}+g_{-}-V+l)}$. We analyze the exponent in detail.

$$g_{+}(x) + g_{-}(x) - V(x) + l = 2 \int_{-\infty}^{\infty} \log(|x-s|) \Psi(s) ds - V(x) + l$$

Recall that the equilibrium measure $\Psi$ satisfies the Euler-Lagrange equations. Thus

$$g_{+}(x) + g_{-}(x) - V(x) + l =
\begin{cases}
< 0 & \text{for } x < a \\
= 0 & \text{for } a \leq x \leq b \\
< 0 & \text{for } x > b
\end{cases}$$

Because of our assumption that the potential $V$ is regular, the inequalities above are sharp.

A further simplification to the jump matrix is this. For $x < a$ and $x > b$,

$$e^{n(g_{+}(x) - g_{-}(x))} = e^{n(g_{+}(x) - g_{-}(x))} = 1.$$  

To summarize, the jump matrix defined on $\mathcal{R}$ has the following properties in the three regions:

$$\nu^{(1)}(x) =
\begin{cases}
\begin{pmatrix}
1 & e^{n(g_{+}(x) + g_{+}(x) - V(x) + l)} \\
0 & 1
\end{pmatrix} & \text{for } x < a \\
\begin{pmatrix}
e^{n(g_{+}(x) - g_{-}(x))} & 1 \\
0 & e^{n(g_{+}(x) - g_{-}(x))}
\end{pmatrix} & \text{for } a \leq x \leq b \\
\begin{pmatrix}
1 & e^{n(g_{+}(x) + g_{+}(x) - V(x) + l)} \\
0 & 1
\end{pmatrix} & \text{for } x > b
\end{cases}$$

11.5. Analytic Continuation. The difference $g_{+}(x) - g_{-}(x)$ has been defined on the real axis as the difference between $g(z)$ on opposite sides of a branch cut. Recall from the previous Subsection that $g_{+}(x) - g_{-}(x)$ has an integral representation for $x \in \mathcal{R}$. Call this difference $G(x)$.

$$G(x) = 2 \pi i \int_{x}^{b} \Psi(s) ds.$$  

Recall that because $V(x)$ is real analytic, $\Psi(z)$ is an analytic function times $\sqrt{(z-a)(b-z)}$:

$$\Psi(z) = h(z) \sqrt{(z-a)(b-z)}.$$  

The function $h(z)$ is analytic in a neighborhood of $\mathcal{R}$, and real on the real axis. We choose $\Psi(z)$ to be positive on the interval $[a,b]$ and have branch cuts along $(-\infty,a]$ and $[b,\infty)$. Because of its integral representation, $G(z)$ is also analytic in this region.
11.6. Factoring the Jump Matrix. Along \((a, b)\), \(G'(x) = -2\pi i \Psi(x)\). Thus there is a complex neighborhood \(U\) of \((a, b)\) such that \(\Re(G) > 0\) in \(U \cap \mathbb{C}^+\), and \(\Re(G) < 0\) in \(U \cap \mathbb{C}^-\). Because \(G\) behaves as \(z^{3/2}\) at the endpoints \(a\) and \(b\), \(U\) cannot form an angle greater than \(\pi/3\), with respect to the real axis, at either endpoint of \([a, b]\). See Figure 1.

In the interval \([a, b]\), we factor the jump matrix as

\[
\nu = \begin{pmatrix} 1 & 0 \\ e^{nG(z)} & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ e^{-nG(z)} & 1 \end{pmatrix}
\]

\[
\nu = \nu_{(1)}^+ \nu_{(1)}^- \nu_{(1)}^0
\]

As illustrated in Figure 1, we deform the original contour \(\Sigma^{(1)}\) to obtain a Riemann Hilbert problem on a new contour \(\Sigma^{(2)}\). It is possible to translate between a solution of one Riemann-Hilbert problem and a solution of the other. Except for \(z\) in the region enclosed by the lenses, \(m^{(2)}(z) = m^{(1)}(z)\).

Inside the upper region,

\[
m^{(2)}(z) = m^{(1)}(z) \left( \nu_{(1)}^+(z) \right)^{-1}.
\]

For \(z\) inside the lower region,

\[
m^{(2)}(z) = m^{(1)}(z)^{\nu_{(1)}^-}(z).
\]

The jump functions on the new contour are:
\[ \nu^{(2)}(z) = \begin{cases} 
\begin{pmatrix} 1 & e^{n(g_-(z) + g_+(z) - V(z) + l)} \\
0 & 1 
\end{pmatrix} & \text{for } z < a \\
\begin{pmatrix} 1 & 0 \\
e^{nG(z)} & 1 
\end{pmatrix} & \text{for } z \text{ on lower lens} \\
\begin{pmatrix} 0 & 1 \\
-1 & 0 
\end{pmatrix} & \text{for } a \leq z \leq b \\
\begin{pmatrix} 1 & 0 \\
e^{-nG(z)} & 1 
\end{pmatrix} & \text{for } z \text{ on upper lens} \\
\begin{pmatrix} 1 & e^{n(g_-(z) + g_+(z) - V(z) + l)} \\
0 & 1 
\end{pmatrix} & \text{for } z > b 
\end{cases} \]

All of the exponents appearing in these jump functions are negative. As \( n \to \infty \), these terms disappear. The limiting jump functions are:

\[ \nu^{(\infty)}(z) = \begin{cases} 
\begin{pmatrix} 1 & 0 \\
0 & 1 
\end{pmatrix} & \text{for } z < a \\
\begin{pmatrix} 1 & 0 \\
0 & 1 
\end{pmatrix} & \text{for } z \text{ on lower lip} \\
\begin{pmatrix} 0 & 1 \\
-1 & 0 
\end{pmatrix} & \text{for } a \leq z \leq b \\
\begin{pmatrix} 1 & 0 \\
0 & 1 
\end{pmatrix} & \text{for } z \text{ on upper lip} \\
\begin{pmatrix} 1 & 0 \\
0 & 1 
\end{pmatrix} & \text{for } z > b 
\end{cases} \]

Let \( m^{(\infty)} \) be the solution of this limiting Riemann-Hilbert problem. The convergence \( \nu^{(2)} \to \nu^{(\infty)} \) as \( n \to \infty \) is not uniform and occurs more slowly in neighborhoods of \( a \) and \( b \). Thus we cannot automatically conclude that \( m_2 \to m_\infty \) as \( n \to \infty \).

Because we assumed the potential \( V(x) \) is regular, there are no other areas of slow convergence.

**11.7. Solution of the Limiting Riemann Hilbert Problem.** To find the solution \( m_\infty \), we perform a change of basis to diagonalize the jump function. This replaces the \( 2 \times 2 \) Riemann-Hilbert problem for \( m_\infty \) with a pair of scalar Riemann-Hilbert problems.

For \( a < x < b \), the diagonalization of the jump function is:

\[
\begin{pmatrix} 0 & 1 \\
-1 & 0 
\end{pmatrix} = \begin{pmatrix} 1 & 1 \\
i & -i 
\end{pmatrix} \begin{pmatrix} i & 0 \\
0 & -i 
\end{pmatrix} \begin{pmatrix} 1 & 1 \\
i & -i 
\end{pmatrix}^{-1}
\]
The transformed Riemann-Hilbert problem is:

\[ \hat{m}(z) = \left( \begin{array}{cc} 1 & 1 \\ i & -i \end{array} \right)^{-1} m_\infty(z) \left( \begin{array}{cc} 1 & 1 \\ i & -i \end{array} \right) \]

\[ \hat{\nu}(z) = \begin{cases} 
(1, 0) & \text{for } x < a \\
(i, 0) & \text{for } a \leq x \leq b \\
(1, 0) & \text{for } x > b 
\end{cases} \]

The solution of this pair of Riemann-Hilbert problems is

\[ \hat{m}(z) = \left( \begin{array}{cc} z - b \sqrt{2} & 1 \\ -1 & z - a \sqrt{2} \end{array} \right) \]

\[ \beta = \frac{z - b \sqrt{2}}{z - a \sqrt{2}}, \]

where \( \beta \) is analytic on \( \mathbb{C} \setminus [a, b] \) and \( \beta(z) \to 1 \) as \( n \to \infty \).

Changing back to the original basis we recover the solution \( m_\infty(z) \):

\[ m_\infty(z) = \left( \begin{array}{cc} 1 & 1 \\ i & -i \end{array} \right) \hat{m}(z) \left( \begin{array}{cc} 1 & 1 \\ i & -i \end{array} \right)^{-1} \]

\[ = \left( \begin{array}{cc} 1 & 1 \\ i & -i \end{array} \right) \hat{m}(z) \left( \begin{array}{cc} 1 & 1 \\ i & -i \end{array} \right)^{-1} \]

\[ = \left( \begin{array}{cc} z - b \sqrt{2} \beta - \beta^{-1} & \beta + \beta^{-1} \\ -1 & z - a \sqrt{2} \beta - \beta^{-1} \end{array} \right). \]

11.8. Parametrices at the Endpoints. The convergence \( \nu^{(2)} \to \nu^{(\infty)} \) is not uniform at the endpoints \( a \) and \( b \). Thus the approximation \( m^{(2)}(z) \approx m^{(\infty)}(z) \) is inappropriate near \( a \) and \( b \).

Let \( O_a \) and \( O_b \) be neighborhoods of \( a \) and \( b \), which can be chosen as small as desired. Inside the neighborhood \( O_a \), one constructs a parametrix \( m_a \) which satisfies the following Riemann-Hilbert problem:

- \( m^{(a)} \) is analytic on \( O_a \setminus \Sigma^{(2)} \)
- \( m_+^{(a)}(z) = m_+^{(a)}(z) \nu^{(2)}(z) \) for \( z \in \Sigma^{(2)} \cap O_a \)
- \( m_+^{(a)}(z) = m^{(\infty)}(z) \left( 1 + \mathcal{O}\left( \frac{1}{n} \right) \right) \) for \( z \in \partial O_a \).

In other words, \( m_a \) satisfies the Riemann-Hilbert problem \( (\Sigma^{(2)}, \nu^{(2)}) \) exactly in a neighborhood of \( a \), and matches \( m^{(\infty)} \) to within \( \mathcal{O}(n^{-1}) \). Similarly, one constructs a parametrix \( m^{(b)} \) in \( O_b \). We will not discuss the construction of
We now patch together an approximate solution of the Riemann-Hilbert problem for $m^{(2)}$:

$$m^{(p)} = \begin{cases} 
m^{(\infty)} & \text{for } z \in C \setminus O_a \setminus O_b \\
m^{(a)} & \text{for } z \in O_a \\
m^{(b)} & \text{for } z \in O_b
\end{cases}$$

The actual solution $m^{(2)}$ will be recovered as a perturbation of the approximate solution. Let $L(z) = m^{(2)}(z)(m^{(p)}(z))^{-1}$, so that $m^{(2)}$ can be recovered from $m^{(p)}$ and $L$. Then $L$ approaches $Id$ at $\infty$ and has a jump function as indicated in Figure 2.

Observe that the jump function for $L(z)$ is in both $L^2(\Sigma)$ and $L^\infty(\Sigma)$.

### 11.9. The Solution $L(z)$ and its Properties.

Since $m^{(2)}(z) = L(z)m^{(p)}(z)$, the quantity $L(z) - Id$ plays the role of an error term in the approximation $m^{(2)}(z) = m^{(p)}(z)$. In order to bound these errors, we make use of Theorem 11.3, which appears in [Dei99], near p. 219.

Suppose that for a given Riemann-Hilbert problem $(\Sigma, \nu)$, the jump function $\nu$ factors as $\nu = b_-^{-1}b_+$, where the $b_\pm$ are bounded and invertible. Let

$$b_\pm = Id \pm \omega_\pm \quad \text{and} \quad \omega = \omega_+ + \omega_-.$$ 

For $f \in L^2(\Sigma)$, let $C_\pm$ be the limits of the Cauchy operators:

$$(Cf)(z) = \frac{1}{2\pi i} \int_{\Sigma} \frac{f(s)}{s - z} \, ds \quad \text{for } z \in CC \setminus \Sigma$$

$$(C_\pm f)(z) = \lim_{z' \to z} Cf(z') \quad \text{for } z \in \Sigma_0$$

Let $C_\omega f = C_+(f\omega_-) + C_-(f\omega_+)$. Assume that $\omega_\pm$ are in $L^2(\Sigma)$, so that $C_\omega Id \in L^2(\Sigma)$. 

---

**Figure 2.** The Riemann-Hilbert problem for $L$. 

These parametrices, which comprise a large part of the work of [DKM+99a] and [DKM+99b].
Theorem 11.5. Suppose that $I - C_\omega$ is invertible on $L^2(\Sigma)$ and let $\mu \in I + L^2(\Sigma)$ be the unique solution of

$$(I - C_\omega)\mu = Id$$

or, more properly,

$$(I - C_\omega)(\mu - Id) = C_\omega Id \in L^2(\Sigma).$$

Then

$$m(z) = I + (C(\mu\omega))(z) \quad \text{for} \quad z \in C \setminus \Sigma$$

is the solution of the Riemann Hilbert problem.

We now depart from [Dei99] in order to bound $L(z)$ and its derivatives.

In our case, we may choose $b_- = Id$, $b_+ = \nu$, and $\omega_+ = \nu - Id$. Then $C_\omega f = C_-(f\omega_+)$. For smooth contours, the $C_\pm$ are bounded on $L^2(\Sigma)$ with constants depending only on $\Sigma$. Since $\omega_+ \in L^\infty(\Sigma)$,

$$||C_\omega f||_{L^2(\Sigma)} \leq \frac{c||\omega_+||_{L^2(\Sigma)}}{||\omega_+||_{L^\infty(\Sigma)}}||f||_{L^2(\Sigma)}.$$

Since $||\omega_+||_{L^\infty(\Sigma)} = O\left(\frac{1}{n}\right)$, we eventually have $||C_\omega||_{L^2 \to L^2} < \frac{1}{2}$. Once this occurs, $(I - C_\omega)$ is invertible with $L^2 \to L^2$ norm at most 2. This allows us to bound $(\mu - Id)$ in $L^2(\Sigma)$:

$$||\mu - I||_{L^2(\Sigma)} \leq ||(I - C_\omega)C_\omega I||_{L^2(\Sigma)} \leq 2||C_\omega I||_{L^2(\Sigma)} \leq 2c||\omega_+||_{L^2(\Sigma)} = O\left(\frac{1}{n}\right).$$

Finally we have an expression for $L(z)$:

$$L(z) = Id + \frac{1}{2\pi i} \int_\Sigma \frac{\mu\omega}{s - z} \, ds.$$
For $z$ not within distance $d$ of the contour $\Sigma$,

$$|L(z) - Id| = \left| \frac{1}{2\pi i} \int_\Sigma \frac{\mu \omega}{s-z} ds \right|$$

$$\leq \frac{1}{2\pi d} ||\mu \omega||_{L^1(\Sigma)}$$

$$= O\left( \frac{1}{nd} \right).$$

For $|z - \Sigma| \geq d$, we can also control the derivatives of $L(z)$:

$$\left| \frac{\partial^k}{\partial z^k} L(z) \right| = \left| \frac{k!}{2\pi i} \int_\Sigma \frac{\mu \omega}{(s-z)^{k+1}} ds \right|$$

$$\leq \frac{k!}{2\pi d^{k+1}} ||\mu \omega||_{L^1(\Sigma)}$$

$$= O\left( \frac{1}{nd^{k+1}} \right).$$

11.10. Retracing Our Steps. We now retrace our steps in order to obtain asymptotics for orthogonal polynomials, on the real axis for $a < z < b$.

First, we choose a proper subinterval of $[a+\epsilon, b-\epsilon] \subset [a, b]$. Next, review the above steps to make sure that the neighborhoods $O_a$ and $O_b$ have radius strictly less than $\epsilon$. Let $d$ be the distance between the contour $\Sigma$ and the interval $[a+\epsilon, b-\epsilon]$.

For $z \in [a+\epsilon, b-\epsilon]$, we obtain the asymptotics for $Y(z)$ by approaching the real axis from above. Recall that $Y_{11}(z) = \pi_n(z)$ and $Y_{21}(z) = -i(\text{const})\pi_{n-1}(z)$ are the quantities that interest us, and are unaffected by the jump function along the real axis.

$$Y(z) = \begin{pmatrix} e^{-\frac{ni^2}{2}} & 0 & 0 & e^{\frac{ni}{2}} \\ 0 & e^{\frac{ni}{2}} & 0 & e^{-\frac{ni}{2}} \end{pmatrix} L(z)m^{(\infty)}(z) \begin{pmatrix} 1 & 0 \\ e^{-nG(z)} & 1 \end{pmatrix}$$

$$\times \begin{pmatrix} e^{\frac{ni}{2}} & 0 & 0 & e^{-\frac{ni}{2}} \\ 0 & e^{-\frac{ni}{2}} & 0 & e^{\frac{ni}{2}} \end{pmatrix} e^{ng_+(z)}$$

$$= \begin{pmatrix} e^{-\frac{ni}{2}} L_{11}(z) & e^{-\frac{ni}{2}} L_{12}(z) & e^{-\frac{ni}{2}} L_{21}(z) & e^{-\frac{ni}{2}} L_{22}(z) \end{pmatrix} \begin{pmatrix} \beta + \beta^{-1} & \beta - \beta^{-1} & 0 \\ \beta - \beta^{-1} & \beta + \beta^{-1} & 0 \\ 0 & 0 & \beta + \beta^{-1} \end{pmatrix}$$

$$\times \begin{pmatrix} e^{\frac{ni}{2} + ng_+(z)} & e^{\frac{ni}{2} + ng_+(z)} & e^{\frac{ni}{2} + ng_+(z)} & e^{\frac{ni}{2} + ng_+(z)} \end{pmatrix}$$
The entry $Y_{11}(z)$ in the above product is

$$
Y_{11}(z) = e^{ng_{+}(z)} \left( L_{11}(z) \frac{\beta + \beta^{-1}}{2} - L_{12}(z) \frac{\beta - \beta^{-1}}{2i} \right) \\
+ e^{ng_{-}(z)-nG(z)} \left( L_{11}(z) \frac{\beta - \beta^{-1}}{2i} + L_{12}(z) \frac{\beta + \beta^{-1}}{2} \right) \\
= \Re \left[ e^{ng_{+}(z)} \left( e^{i\pi} (L_{11}(z) + iL_{12}(z)) \left( \frac{b - z}{z - a} \right)^{\frac{1}{4}} \\
+ e^{-i\pi} (L_{11}(z) - iL_{12}(z)) \left( \frac{z - a}{b - z} \right)^{\frac{1}{4}} \right) \right] \\
= e^{n \int_{-a}^{b} \log |z-s| \Psi(s) ds} \times \\
\Re \left[ e^{ni\pi} \int_{-a}^{b} \Psi(s) ds \left( e^{i\pi} (L_{11}(z) + iL_{12}(z)) \left( \frac{b - z}{z - a} \right)^{\frac{1}{4}} \\
+ e^{-i\pi} (L_{11}(z) - iL_{12}(z)) \left( \frac{z - a}{b - z} \right)^{\frac{1}{4}} \right) \right].
$$

The entry $Y_{21}(z)$ in the above product is:

$$
= e^{n+l+n \int_{-a}^{b} \log |z-s| \Psi(s) ds} \times \\
i \times \Im \left[ e^{ni\pi} \int_{-a}^{b} \Psi(s) ds \left( e^{i\pi} (L_{22}(z) - iL_{21}(z)) \left( \frac{b - z}{z - a} \right)^{\frac{1}{4}} \\
+ e^{-i\pi} (L_{22}(z) + iL_{21}(z)) \left( \frac{z - a}{b - z} \right)^{\frac{1}{4}} \right) \right].
$$

In the Christoffel-Darboux formula, it is easier to deal with normalized orthogonal polynomials $\phi_j(z)$ than with monic orthogonal polynomials $\pi_j(z)$:

$$
\phi_j(x) = \frac{\pi_j(x)}{\left( \int_{-\infty}^{\infty} (\pi_j(x))^2 e^{-nV(x)} dx \right)^{\frac{1}{2}}}
$$

When $j = n$, we approximate heuristically the normalizing constant in the denominator, refer the reader to [DKM+97] for the precise result, and
state without proof the error term in our approximation:

\[
\int_{-\infty}^{\infty} (\pi_n(x))^2 e^{-nV(x)} \, dx \\
= \left(1 + O\left(\frac{1}{n}\right)\right) \int_{a}^{b} (\pi_n(x))^2 e^{-nV(x)} \, dx \\
= \left(1 + O\left(\frac{1}{n}\right)\right) \times \int_{a}^{b} e^{2n \int_{x}^{b} \log|x-s| \Psi(s) \, ds - nV(x)} \\
\times \mathfrak{R} \left[ \text{oscillatory} \cdot \sqrt{\left(\frac{b-x}{x-a}\right)^{\frac{1}{2}} + \left(\frac{x-a}{b-x}\right)^{\frac{1}{2}}} \right]^2 \, dx \\
= \left(1 + O\left(\frac{1}{n}\right)\right) e^{-nl} \int_{a}^{b} \frac{1}{2} \left(\frac{b-x}{x-a}\right)^{\frac{1}{2}} + \left(\frac{x-a}{b-x}\right)^{\frac{1}{2}} \right) \, dx \\
= e^{-nl} \frac{(b-a)\pi}{2} \left(1 + O\left(\frac{1}{n}\right)\right).
\]

Even more convenient than the normalized orthogonal polynomials, for later computations, are

\[
\eta_j(x) = \frac{\phi_j(x)}{e^\frac{n\pi}{2}V(x)}.
\]

The advantage of using \(\eta_j(x)\) is that they are orthonormal with respect to Lebesgue measure on \(\mathcal{R}\).

The entry \(Y_{21}(z)\) is a negative imaginary constant times \(\pi_{n-1}(z)\). Using techniques similar to those above, we approximate normalize this polynomial, obtaining an approximation for \(\phi_{n-1}(x)\).

Using our knowledge of \(Y_{11}(z), Y_{21}(z)\), the normalizing constants, and \(L(z)\), we obtain the leading order asymptotics for \(\eta_n(x), \eta_{n-1}(x)\), and their derivatives.

**Theorem 11.6.** Let \(\epsilon > 0\), and \(k\) a nonnegative integer. For \(z \in [a+\epsilon, b-\epsilon]\), the leading order asymptotics of

\[
\frac{\partial^k}{\partial z^k} \eta_n(z) \quad \text{and} \quad \frac{\partial^k}{\partial z^k} \eta_{n-1}(z)
\]

as \(n \to \infty\) are, respectively,

\[
(n\pi \Psi(z))^k \sqrt{\frac{2}{(b-a)\pi}} \mathfrak{R} \left[ i^k e^{n\pi \int_{s}^{b} \Psi(s) \, ds} \left(e^{-\frac{n\pi}{z-a}} \left(\frac{b-z}{z-a}\right)^{\frac{1}{4}} + e^{-i\frac{n\pi}{z-b}} \left(\frac{z-a}{b-z}\right)^{\frac{1}{4}} \right) + O\left(\frac{1}{n}\right) \right]
\]
\[
\left( n \pi \Psi(z) \right)^k \sqrt{\frac{2}{(b-a)\pi}} \quad \Re \left[ i^k e^{n \pi i} \int_a^b \psi(s) ds \left( e^{-i \frac{3\pi}{4}} \left( \frac{b-z}{z-a} \right)^{\frac{1}{4}} + e^{-i \frac{\pi}{4}} \left( \frac{z-a}{b-z} \right)^{\frac{1}{4}} \right) + O \left( \frac{1}{n} \right) \right].
\]

The constants implied by O depend only on \( k \), the potential \( V(x) \), and \( \epsilon \).

12. Special Case: The Gaussian Unitary Ensemble

The ensemble \( GUE_n \) is the set of \( n \times n \) Hermitian matrices with joint probability density function (j.p.d.f.)

\[
\left( \pi^{-\frac{m^2}{2} \frac{n(n+1)}{2}} \right) \exp \left( -\sum_i m_{i,i}^2 - 2 \sum_{i<j} |m_{i,j}|^2 \right) dM \\
= C_n e^{-Tr(M^2)} dM
\]

In other words, the entries on the diagonal are real, and are chosen independently from the Gaussian distribution with variance 1. In the upper triangle, the real and imaginary parts of each entry are chosen independently, all from the Gaussian distribution with variance \( \frac{1}{2} \).

The j.p.d.f. for the eigenvalues is obtained from the j.p.d.f. for the matrix entries using Weyl integration:

\[
\left( \prod_{j=0}^{n-1} \left( \frac{2j}{j!} \right) \pi^{-n/2} \frac{1}{n!} \right) \prod_{i<j} (\lambda_i - \lambda_j)^2 e^{-\sum \lambda_i^2}
\]

At first glance the Gaussian unitary ensemble seems not to be an instance of the universal unitary ensemble, because the j.p.d.f. contains the term \( e^{-\sum \lambda_i^2} \) instead of \( e^{-n \sum \lambda_i^2} \). Since the potential \( V(x) = x^2 \) is homogeneous, this difference is eliminated by rescaling. Thus, the Moment Estimation Theorem we present for GUE is really a special case of the universal version of the Theorem.

Instead of just using the universal Theorem, we prove the Moment Estimation Theorem in the GUE case by using Plancherel-Rotach asymptotics for Hermite polynomials instead of the more general asymptotics by Deift, Kriecherbauer, McLaughlin, Venakides, and Zhou. The standard reference for Plancherel-Rotach asymptotics is [Sze75].

We present this alternative proof because many readers will be familiar with Plancherel-Rotach asymptotics but not the Deift-Zhou asymptotics.
For such readers, the GUE version of the Moment Estimation Theorem is a good warm-up for the universal case.

**Theorem 12.1** (Moment Estimation Theorem, GUE version). Let $\epsilon > 0$. For each $n$, let $I_n \subset (-\sqrt{\frac{2}{\pi}} - \epsilon \sqrt{n}, \sqrt{\frac{2}{\pi}} - \epsilon \sqrt{n})$ be an interval, such that $\sqrt{n}|I_n| \to \infty$ as $n \to \infty$. Let $G_{\gamma,I,n}$ be the random variable which counts the number of GUE$_n$ eigenvalues whose average is in $I_n$ and whose difference is at most $\gamma n$. Let $G_{\mu}$ be the Poisson distribution with mean

$$\mu_n = \frac{\pi^2 \gamma^3}{9} \int I K_n(x,x)^4 dx.$$ 

Then for all $k \geq 1$, 

$$E(G_{\gamma,I,n}^k) = E(G_{\mu}^k) \left(1 + O\left(n^{-1}, n\gamma^2, (n^{\frac{1}{2}}|I|)^{-\frac{2}{3}}\right)\right).$$

The constant implied by $O$ depends only on $k$ and $\epsilon$.

The novel aspects of the proof of the GUE Theorem are the estimation of the first moment using Plancherel-Rotach asymptotics, the different scaling, and a slightly different proof of the bounds for $K_n(X,y)$ and its derivatives. We only treat these aspects of the GUE proof.

### 12.1. The First Moment in the GUE Case.

As in the CUE and UE cases, one uses the method of Gaudin to express the expected value of $G_{\gamma,I,n}$ as a two-dimensional integral.

$$E(G_{\gamma,I,n}) = \int \int_{\mathbb{R}^2} g_2(u,t) \left[ K_n(u,u) K_n(u,t) \right. \
\left. K_n(t,u) K_n(t,t) \right] dudt$$

$$= \frac{1}{2} \int \int_{\Omega} \left[ K_n(u,u) K_n(u,t) \right. \
\left. K_n(t,u) K_n(t,t) \right] dudt,$$

where $\Omega$ is the region

$$\Omega = \left\{(u,t) \text{ s.t. } \frac{u+t}{2} \in I, |t-u| < \gamma \right\}.$$ 

We state without proof an expression for the projection kernel:

$$K_n(x,y) = \frac{e^{-\frac{(x^2+y^2)}{2}}}{2^n(n-1)!\sqrt{\pi}} \left( H_n(x)H_{n-1}(y) - H_{n-1}(x)H_n(y) \right),$$

where $H_n(x)$ is the $(n)$th Hermite polynomial. This expression is an instance of the Christoffel-Darboux formula, which we encountered in the UUE case. The only challenge in the above formula is to derive the correct normalizing constant.

As compared to the $U_n$ case, it is now slightly more difficult to estimate $E(G_{\gamma,I,n})$ because the kernel $K_n(x,y)$ no longer depends solely on the difference $(y-x)$. 
As in the case of UUE (Subsection 8), we change variables and expand the integrand in a Taylor series:

The region Ω is narrow in the \((t-u)\) direction. This suggests expanding the integrand in a Taylor series. Let \(x = \tfrac{u+t}{2}\) and \(y = \tfrac{u-t}{2}\). Then \(dxdy = \tfrac{1}{2}dudt\). In terms of the new variables \(x\) and \(y\), the first moment is:

\[
E(G_{\gamma,I,n}) = \int_I \int_{-\frac{\gamma}{2}}^{\frac{\gamma}{2}} \left( y^2 \left( KK'\gamma - KK'\gamma - K'\gamma^2 \right) + O(y^4c_0c_4 + y^8c_4c_4) \right) dxdy,
\]

where and \(c_j\) is the maximum over \(\Omega\) of any \((j)\)th partial derivative of \(K_n(x,y)\). Using the Lemma 12.2 instead of Lemma 10.1, the integral of \(O(y^4c_0c_4)\) and \(O(y^5c_4c_4)\) over \(\Omega\) are \(O(|I|\gamma^5n^3)\) and \(O(|I|\gamma^9n^5)\) respectively.

We explicitly expand the integrand in terms of Hermite polynomials. After collecting similar terms,

\[
E(G_{\gamma,I,n}) = O \left( |I|\gamma^5n^3 + |I|\gamma^9n^5 \right) + \int_I \frac{\gamma^3}{12} \left( \frac{e^{-2t^2}}{2^{2n}(n-1)!2\pi} \right) \times \left[ \begin{array}{c} 2(H_{n-1}^{(1)}H_n^{(1)} - H_nH_{n-1}^{(1)})(H_n^{(2)} - H_n^{(2)}) \\ + \frac{2}{3}(H_{n-1}^{(3)} - H_nH_{n-1}^{(3)})(H_n^{(3)} - H_n^{(3)}) \\ - (H_nH_{n-1}^{(3)} - H_n^{(3)})^2 \end{array} \right] dt.
\]

All Hermite polynomials or their derivatives are evaluated at \(t\); this is omitted to save space. In the next section, we will use the Plancherel Rotach asymptotics to derive an approximation to the integrand valid as \(n \to \infty\).

### 12.2. Application of Plancherel Rotach Asymptotics to the First Moment.

Let’s use Plancherel Rotach to estimate \(E(G_{\gamma,I,n})\). Unlike other systems of orthogonal polynomials (such as we will encounter in the case of universal unitary ensembles), the Hermite polynomials can be differentiated easily. For fixed \(k\) as \(n \to \infty\):

\[
H_{n-k}^{(1)}(x) = 2(n-k)H_{n-k-1}(x) = 2nH_{n-k-1}(x)(1 + O(n^{-1})�).
\]

This shortcut for differentiating Hermite polynomials allows us to simplify the integrand from the most recent formula for \(E(G_{\gamma,I,n})\):

\[
E(G_{\gamma,I,n}) = O \left( |I|\gamma^5n^3 + |I|\gamma^9n^5 \right) + \int_I \frac{4\gamma^3}{3} \left( \frac{e^{-2t^2}}{2^{2n}(n-1)!2\pi} \right) n^4 \left( 1 + O(n^{-1}) \right) \times \left[ \begin{array}{c} 2(H_{n-1}^2 - H_nH_{n-2})(H_{n-2}^2 - H_{n-1}H_n) \\ + \frac{2}{3}(H_{n-3}H_{n-1}^3 - H_nH_{n-3})(H_{n-1}^3 - H_nH_{n-1}) \\ - (H_nH_{n-3} - H_{n-1}H_{n-2})^2 \end{array} \right] dt.
\]

Notice that we factored the \((1 + O(n^{-1}))\) outside of the brackets in order to save space. Technically this is dangerous because of the possibility that the
terms inside brackets interfere destructively, but we will see that destructive interference does not occur in our case.

We use the Plancherel Rotach asymptotics in Theorem \ref{thm:plancherel_rotach} to write the Hermite polynomials as an envelope times a phase function and estimate each of the above groupings. For the first grouping, the result is:

\[
(H_{n-1}H_{n-1} - H_nH_{n-2}) = e^{t^2}2^n(n-1)^{n-1}e^{1-n}\left(1 - \frac{t^2}{2n}\right)^{-\frac{1}{2}}
\times \left[\Re(\Phi(n-1))^2 - \Re(\Phi(t,n))\Re(\Phi(t,n-2)) + O(n^{-1})\right].
\]

We will now use Lemma \ref{lem:phase_shift}. Let

\[
u = \Phi(x,n)
\]
\[
v = \left(x - i\sqrt{2n - x^2}\right)\sqrt{2n}
\]

The in terms of these variables,

\[
\Phi(x,n) = u
\]
\[
\Phi(x,n-1) = uv
\]
\[
\Phi(x,n-2) = uv^2
\]

Using the identity \(\Re[uv]^2 - \Re[u]\Re[vu] = \Im[v]^2\) for unimodular \(u\) and \(v\), we have

\[
(H_{n-1}^2H_{n-2}) = e^{t^2}2^n(n-1)^{n-1}e^{1-n}\left(1 - \frac{t^2}{2n}\right)^{-\frac{1}{2}} \left[1 + O(n^{-1})\right].
\]

Applying the same techniques to other groupings,

\[
(H_{n-1}^2 - H_nH_{n-2}) = e^{t^2}2^n(n-1)^{n-1}e^{1-n}s^{-1}(s^2)^{-1}\left[1 + O(n^{-1})\right]
\]
\[
(H_{n-2}^2 - H_{n-1}H_{n-3}) = e^{t^2}2^{n-1}(n-2)^{n-2}e^{2-n}s^{-1}(s^2)^{-1}\left[1 + O(n^{-1})\right]
\]
\[
(H_{n-1}H_{n-3} - H_nH_{n-4}) = e^{t^2}2^{n-1}(n-2)^{n-2}e^{2-n}s^{-1}(3s^2 - 4s^4)^{-1}\left[1 + O(n^{-1})\right]
\]
\[
(H_{n-3}H_{n-3} - H_{n-1}H_{n-2}) = e^{t^2}2^{n-1}n^2(n-3)^{n-2}e^{2-n}s^{-1}(-2cs^2)^{-1}\left[1 + O(n^{-1})\right]
\]

where

\[
s = \sqrt{1 - \frac{t^2}{2n}}
\]
\[
c = \frac{t}{\sqrt{2n}}.
\]
Incorporating these newly derived asymptotics into the recent expression for 
\(E(G_{\gamma,I,n})\) yields:

\[
E(G_{\gamma,I,n}) = O(I^{\gamma}n^3 + |I|^{\gamma}n^5) + \int_I \frac{4\gamma^3n^2}{3}s^{-2}\left(1 + O(n^{-1})\right) \left[ \frac{2(s^2)(s^2)^2}{3}(s^2 - 4s^4) - 2cs^2 \right] \, dt
\]

Recall the Wigner Semicircle law from Appendix 12.5 which says that

\[
K_n(t,t) = \sqrt{\frac{2n}{\pi}} \sqrt{1 - \frac{t^2}{2n}(1 + O(n^{-1}))}.
\]

We therefore see that:

\[
E(G_{\gamma,I,n}) = O(I^{\gamma}n^3 + |I|^{\gamma}n^5) + \frac{\pi^{2\gamma^3}}{9} \int_I K_n(t,t)^4(1 + O(n^{-1})) dt.
\]

This agrees with our heuristic prediction.

### 12.3. Controlling the Derivatives of \(K_n(x,y)\) in the GUE Case.

The error terms in our estimates of the moments of \(E(G^k_{\gamma,I,n})\) are bounded in terms of derivatives of \(K_n(x,y)\). We now bound these derivatives. Unlike the projection kernels of universal ensembles to come, we can now exploit a shortcut when differentiating the orthogonal polynomials. In this case the orthogonal polynomials are Hermite polynomials, and the shortcut is

\[
\frac{\partial}{\partial x}H_n(x) = 2nH_{n-1}(x).
\]

The following Theorem is the GUE version of Lemma 3.1.

**Lemma 12.2.** Fix \(\epsilon > 0\). For \((x,y)\) in the region

\[
R = \left[ -(1 - \epsilon)\sqrt{2n}, (1 - \epsilon)\sqrt{2n} \right] \times \left[ -(1 - \epsilon)\sqrt{2n}, (1 - \epsilon)\sqrt{2n} \right],
\]

a mixed partial derivative of \(K_n(x,y)\) of total degree \(k\) is \(O\left( n^{\frac{k+1}{2}} \right)\). When

\[
|x - y| \geq n^{-\frac{1}{2}},
\]

we have the stronger bound \(K_n(x,y) = O\left( \frac{1}{|x-y|^{k \frac{3}{2}}} \right)\). The constants implied by \(O\) depend only on \(k\) and \(\epsilon\).

**Proof.** Recall the closed form expression for \(K_n(x,y)\):

\[
K_n(x,y) = \frac{e^{-(x^2+y^2)/2}}{2^n(n-1)!\sqrt{\pi}} \frac{H_n(x)H_{n-1}(y) - H_{n-1}(x)H_n(y)}{x-y}
\]
Take $k$ partial derivatives of $K_n(x, y)$. The result is a sum of several terms of the form:

$$e^{-\frac{(x^2+y^2)}{2}} \frac{p_{k_0}(n)q_{k_1}(x, y)H_{n-k_2}(x)H_{n-k_3}(y)}{2^n(n-1)!\sqrt{\pi}} (x-y)^{k_4},$$

where $p_{k_0}(n)$ and $q_{k_1}(x, y)$ are integral polynomials of total degrees $k_0$ and $k_1$ respectively. Before taking any derivatives, i.e. $k = 0$, there are two terms, each with $(k_0, k_1, k_2, k_3) = (0, 0, 1, 1)$. Each new partial derivative can hit in one of four places, and increments the worst-case vector $(k_1, k_2+k_3, k_4)$.

Here are the four cases:

- It hits $e^{-\frac{(x^2+y^2)}{2}}$. The increment is $(0, 0, 1, 0)$.
- It hits $q$. The increment is $(0, -1, 0, 0)$.
- It hits a Hermite polynomial. The increment is $(1, 0, 1, 0)$.
- It hits the denominator. The increment is $(0, 0, 0, 1)$.

We divide the region $R$ into two regions:

\begin{align*}
R_{\text{diag}} &= \left\{(x, y) \in R \mid |x-y| < \frac{1}{\sqrt{n}} \right\} \\
R_{\text{bulk}} &= \left\{(x, y) \in R \mid |x-y| \geq \frac{1}{\sqrt{n}} \right\}
\end{align*}

Our estimates are straightforward in the bulk region $R_{\text{bulk}}$. Consider any of the finitely many terms resulting from taking $k$ partial derivatives of $K_n(x, y)$:

\begin{align*}
\frac{e^{-\frac{(x^2+y^2)}{2}}}{2^n(n-1)!\sqrt{\pi}} p_{k_0}(n)q_{k_1}(x, y)H_{n-k_2}(x)H_{n-k_3}(y) (x-y)^{k_4}
&= O \left( \frac{1}{|x-y|^{n-\frac{k_1-k_2-k_3+k_4}{2}}} \right)
&= O \left( \frac{1}{|x-y|^{n-\frac{k_1}{2}}} \right)
&= O \left( n^{\frac{1+k}{2}} \right)
\end{align*}

Estimates for the diagonal region $R_{\text{diag}}$ are more subtle.

Consider the term

$$\frac{e^{-\frac{(x^2+y^2)}{2}}}{2^n(n-1)!\sqrt{\pi}} p_{k_0}(n)q_{k_1}(x, y)H_{n-k_2}(x)H_{n-k_3}(y) (x-y)^{k_4}. $$
We expand $H_{n-k}(y)$ and $q_k(x, y)$ as a Taylor series centered at $x$:

$$q_k(x, y) = q_k(x, x) + (y - x)(\partial_2 q_k(x, x)) + \ldots$$

$$+ \frac{(y - x)^{k_1 - 1}}{(k_1 - 1)!}(\partial_2^{k_1 - 1} q_k(x, x)) + \frac{(y - x)^{k_4}}{k_4!}(\partial_2^{k_4} q_k(x, x))$$

$$H_{n-k}(y) = H_{n-k}(x) + (y - x)H'_{n-k}(x) + \ldots$$

$$+ \frac{(y - x)^{k_4 - 1}}{(k_4 - 1)!}H'_{n-k}(x) + \frac{(y - x)^{k_4}}{k_4!}H^{k_4}_{n-k}(x),$$

for some $x_1 \in (x, y)$ and $x_3 \in (x, y)$. Here’s what happens if we select the remainder term from the expansion of $H_{n-k}(y)$, and a lower order term, $0 \leq j_1 < k_4$, in the expansion of $q_k(x, y)$:

$$e^{-\frac{(x^2 + y^2)}{2}} \frac{p_{k_0}(n) (\frac{(y - x)^{j_1}}{j_1!}(\partial_2^{j_1} q_k(x, x)))}{2^n(n - 1)!\sqrt{\pi}} \frac{H_{n-k}(x) (\frac{(y - x)^{k_4}}{k_4!}H^{k_4}_{n-k}(x))}{(x - y)^{k_4}}$$

$$= O\left(n^{\frac{1 + 2k_0 + k_1 - j_1 - k_2 - k_4 + k_4}{2}}\right) = O\left(n^{\frac{1 + k - j_1}{2}}\right)$$

Here’s what happens when we take a lower order term, $0 \leq j_3 < k_4$, from the expansion of $H_{n-k}(y)$, and the remainder term for the expansion of $q_k(x, y)$:

$$e^{-\frac{(x^2 + y^2)}{2}} \frac{p_{k_0}(n) (\frac{(y - x)^{k_4}}{k_4!}(\partial_2^{k_4} q_k(x, x)))}{2^n(n - 1)!\sqrt{\pi}} \frac{H_{n-k}(x) (\frac{(y - x)^{j_3}}{j_3!}H^{j_3}_{n-k}(x))}{(x - y)^{j_3}}$$

$$= O\left(n^{\frac{1 + 2k_0 + k_1 - j_1 - k_4 - k_3 + j_1}{2}}\right) = O\left(n^{\frac{1 + k - j_3}{2}}\right)$$

A term with lower order terms $j_1$ and $j_3$ with $j_1 + j_3 \geq k_4$ also make a small contribution,

$$O\left(n^{\frac{1 + 2k_0 + k_1 - j_1 - k_4 - k_3 + j_1}{2}}\right) = O\left(n^{\frac{1 + k - j_1}{2}}\right).$$

Most subtle of all are the remaining terms

$$e^{-\frac{(x^2 + y^2)}{2}} \frac{p_{k_0}(n) (\frac{(y - x)^{j_1}}{j_1!}(\partial_2^{j_1} q_k(x, x)))}{2^n(n - 1)!\sqrt{\pi}} \frac{H_{n-k}(x) (\frac{(y - x)^{j_3}}{j_3!}H^{j_3}_{n-k}(x))}{(x - y)^{j_3}}$$

where $j_1 + j_3 < k_4$. Adding all of these terms together, we obtain, for each $1 \leq l \leq k_4$, a polynomial in $x$ times

$$e^{-\frac{x^2 + y^2}{2}} \frac{(x - y)^l}{(x - y)^l}.$$

Since $K_n(x, y)$ is smooth, there are no singularities on the diagonal, and all these polynomials in $x$ must vanish identically. Thus all the terms with $j_1 + j_3 < k_4$ cancel out against each other.
12.4. Derivation of Plancherel-Rotach Asymptotics. We will use the method of steepest descent (see [BO78]) to prove the Plancherel Rotach asymptotics for Hermite polynomials.

**Theorem 12.3** (Plancherel-Rotach). Fix $\epsilon > 0$, and let $(x, n)$ be a pair for which $|x| < (\sqrt{2} - \epsilon)\sqrt{n}$. Then:

$$H_n(x) = e^{x^2/2} 2^{n+1/2} n^2 e^{-n x^2/2n} \left(1 - \frac{x^2}{2n}\right)^{-\frac{1}{4}} \left[\text{Re}(\Phi(x, n)) + O(n^{-1})\right],$$

where the phase $\Phi(x, n)$ is:

$$\Phi(x, n) = e^{i\pi/4} \left(\frac{x}{\sqrt{2n}} + i\sqrt{1 - \frac{x^2}{2n}}\right) - \frac{i}{2} x \sqrt{2n-x^2}.$$

The constant implied by $O$ depends only on $\epsilon$.

**Proof.** Begin with the well-known integral representation for Hermite polynomials:

$$H_n(z) = \frac{2^n}{\sqrt{\pi}} \int_{-\infty}^{\infty} (z + it)^n e^{-t^2} dt.$$

One may verify this integral representation by checking that $H_0(x)$ and $H_1(x)$ are correct, and then verifying the three term recurrence $H_n(t) = 2tH_{n-1}(t) - 2(n-1)H_{n-2}(t)$ for Hermite polynomials.

First change variables so that the locations of the saddle points will remain constant as $n \to \infty$. Let $w = \frac{x}{\sqrt{n}}$ and $u = \frac{t}{\sqrt{n}}$. The integral representation becomes:

$$H_n(\sqrt{n}w) = \frac{2^n}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{n(\log(\sqrt{n}w+iu) - u^2)} du.$$

For fixed $w \in (-\sqrt{2} - \epsilon, \sqrt{2} - \epsilon)$ let us apply stationary phase to the above integral in the $n \to \infty$ limit. The saddle points occur when $\frac{i}{w + iu} - 2u = 0$, or

$$u_\pm = \frac{1}{2} \left(\frac{i}{2} \pm \sqrt{2 - w^2}\right).$$

The second derivatives of the phase function at these points are

$$f_\pm = 2 \left((w^2 - 2) \mp i w \sqrt{2 - w^2}\right).$$

The absolute value of $f_\pm$ is:

$$2\sqrt{4 - 2w^2}.$$

The slope of the stationary phase contour at the saddle point $u_+$ is

$$\sqrt{\frac{-f_+}{|f_+|^2}}$$

with the sign chosen to have positive real part.
The value of the integrand is $v_{\pm}$, where

$$v_{\pm} = \left( \frac{1}{2}w \pm \frac{1}{2}i\sqrt{2 - w^2} \right) e^{-\frac{1}{4}(iw\pm\sqrt{2-w^2})^2}.$$ 

The integral representation for $H_n(\sqrt{nw})$ has the real axis as its contour of integration. We change this contour to a stationary phase contour, that is $\Im(\log(w+iu)-u^2) = const$, which passes through the two saddle points. The stationary phase contour begins at $-\infty$, hits the left saddle point, crosses over to the right saddle point, then goes to $+\infty$, and overall looks a bit like a Gaussian. Reducing to small neighborhoods $S_-$ and $S_+$ of the saddle points produces only exponentially small errors.

The contributions from the two neighborhoods are complex conjugates of each other, so we focus on $S_+$:
\[
H_n(\sqrt{n}w) = 2\Re \left[ \frac{2^n n^{n+1}}{\sqrt{\pi}} \int_{S^+} e^{n(\log(w+iw)-w^2)} du \left( 1 + O(e^{-c\sqrt{n}}) \right) \right]
\]
\[
= 2\Re \left[ \frac{2^n n^{n+1}}{\sqrt{\pi}} n \sqrt{\frac{f_+}{|f_+|}} \sqrt{\frac{2\pi}{n|f_+|}} \left( 1 + O(n^{-1}) \right) \right]
\]
\[
= 2 \frac{n^{n+1}}{n} e^{\frac{1}{2} n(w^2-1)} \left( 1 - \frac{w^2}{2} \right)^{-\frac{1}{4}}
\times
\Re \left[ \left( \frac{w}{\sqrt{2}} + \frac{i\sqrt{2-w^2}}{\sqrt{2}} \right)^n e^{-\frac{1}{2} iwn\sqrt{2-w^2}} \sqrt{\frac{2-w^2}{\sqrt{2}}} + \frac{iw}{\sqrt{2}} \right].
\]

Now recall that \( w = \frac{z}{\sqrt{n}} \) and notice that two of the terms in the phase function can be combined:

\[
H_n(z) = 2 \frac{n^{n+1}}{n} e^{\frac{z^2}{2n}} e^{\frac{z^2}{2}} \left( 1 - \frac{z^2}{2n} \right)^{\frac{1}{4}}
\times
\Re \left[ e^{\frac{z}{\sqrt{2}}} \left( \frac{z}{\sqrt{2n}} + i \sqrt{1 - \frac{z^2}{2n}} \right)^{n-\frac{1}{2}} e^{-\frac{1}{2} i\pi \sqrt{2n-z^2}} + O(n^{-1}) \right]
\]

□

We’ll often use Plancherel-Rotach by comparing \( H_n(x) \) for consecutive values of \( n \). The following simple Lemma describes the way in which \( \Phi \) varies with \( n \):

**Lemma 12.4.** Fix \( \epsilon > 0 \). Then for pairs \((x,n)\) for which \(|x| < (\sqrt{2} - \epsilon)\sqrt{n}\),

\[
\frac{\Phi(x, n+1)}{\Phi(x, n)} = \left( \frac{x + i\sqrt{2n-x^2}}{\sqrt{2n}} \right) + O(n^{-1})
\]
Proof.

\[
\frac{\Phi(x, n+1)}{\Phi(x, n)} = \left( \frac{z}{\sqrt{2n+2}} + i \sqrt{1 - \frac{z^2}{2n+2}} \right)^{n-\frac{1}{2}} e^{-\frac{1}{2}iz \sqrt{2n+2-z^2+\frac{1}{2}z \sqrt{2n-z^2}}} \\
= \left( 1 + i \frac{z}{\sqrt{2n+2}} \sqrt{1 - \frac{z^2}{2n+2}} - i \frac{z}{\sqrt{2n+2}} \sqrt{1 - \frac{z^2}{2n+2}} + O(n^{-2}) \right)^n \\
\times \left( \frac{z}{\sqrt{2n+2}} + i \sqrt{1 - \frac{z^2}{2n+2}} \right) e^{-\frac{1}{2}iz \sqrt{2n-z^2}} + O(n^{-1}) \\
= \left( \frac{z}{\sqrt{2n+2}} + i \sqrt{1 - \frac{z^2}{2n+2}} \right) + O(n^{-1}) \\
\]

\[\square\]

12.5. Derivation of Wigner Semicircle Law using Plancherel
Rotach Asymptotics. We use Plancherel Rotach asymptotics, Theorem 12.3, to prove the Wigner semicircle law. This is similar to our estimation of \(E(G, I, n)\) using Plancherel-Rotach asymptotics.

**Lemma 12.5.** For fixed \(\varepsilon > 0\) and \(x \in \left[ -(1 - \varepsilon)\sqrt{2n}, (1 - \varepsilon)\sqrt{2n} \right]\), the diagonal of the kernel \(K_n(x, y)\) satisfies the following asymptotics as \(n \to \infty\):

\[
K_n(x, x) = \frac{\sqrt{2n}}{\pi} \sqrt{1 - \frac{x^2}{2n}} (1 + O(n^{-1})).
\]

**Proof.** Since \(K_n(x, y)\) is smooth everywhere, it is continuous on the diagonal \(x = y\). By continuity,

\[
K_n(x, x) = \lim_{h \to 0} K_n(x + h, x) \\
= \lim_{h \to 0} \frac{e^{-x^2}}{2^n(n-1)! \sqrt{\pi}} \left( \frac{H_n(x + h) H_{n-1}(x) - H_{n-1}(x + h) H_n(x)}{h} \right) \\
= \frac{e^{-x^2}}{2^n(n-1)! \sqrt{\pi}} \left( H_n'(x) H_{n-1}(x) - H_{n-1}'(x) H_n(x) \right) \\
= \frac{e^{-x^2}}{2^n(n-1)! \sqrt{\pi}} \left( 2n H_{n-1}(x)^2 - 2(n-1) H_{n-2}(x) H_n(x) \right) \\
= \frac{\sqrt{2n}}{\pi} \left( 1 - \frac{x^2}{2n} \right)^{-\frac{1}{2}} \left( \Re(\Phi(x, n-1))^2 - \Re(\Phi(x, n)) \Re(\Phi(x, n-2)) + O \left( \frac{1}{n} \right) \right)
\]
We will now use Lemma 12.4. Let

\[ u = \Phi(x, n) \]
\[ v = \left( x - i\sqrt{2n - x^2} \right) / \sqrt{2n} \].

The in terms of these variables,
\[ \Phi(x, n) = u \]
\[ \Phi(x, n - 1) = uv \]
\[ \Phi(x, n - 2) = uv^2 \]

Using the identity \( \Re[uv]^2 - \Re[u]\Re[uv^2] = \Im[v]^2 \) for unimodular \( u \) and \( v \), we have
\[
K_n(x, x) = \frac{\sqrt{2n}}{\pi} \left( 1 - \frac{x^2}{2n} \right)^{-\frac{1}{2}} \left( \Re(uv)^2 - \Re(u)\Re(uv^2) + O\left( \frac{1}{n} \right) \right)
\]
\[
= \frac{\sqrt{2n}}{\pi} \left( 1 - \frac{x^2}{2n} \right)^{-\frac{1}{2}} \left( 1 + O\left( \frac{1}{n} \right) \right)
\]

\[\Box\]

13. Appendix: Weyl integration

The following discussion of Weyl integration is from [Meh91], p.62-63.

Lemma 13.1. Suppose that an ensemble of Hermitian matrices has the following joint probability density function for the matrix entries:

\[ C_n e^{-n \sum V(\theta_i)} dM, \]

where \( V(x) \) is a potential, the \( \theta_i \) are eigenvalues, and \( C_n \) is a normalization constant. Then the joint probability density function for the eigenvalues is

\[ \tilde{C}_n \prod_{i < j} (\theta_j - \theta_i)^2 e^{-n \sum V(\theta_i)} d\Theta, \]

where \( \tilde{C}_n \) is a new normalization constant.

Let \( U \) be a unitary matrix and \( \Theta \) be a diagonal matrix so that

\[ H = U\Theta U^*. \]

Except on a set of measure zero, the eigenvalues will be distinct. Assume without loss of generality that the diagonal entries of \( \Theta \) are in ascending order. We also assume without loss of generality that the first nonzero entry in each column of \( U \) is positive real. With these conventions, the \( \Theta \) and \( U \)
are uniquely determined by $H$, and $H$ is of course uniquely determined by $U$ and $\Theta$. Let $p_\mu$ be $n(n-1)$ real variables which specify a unitary matrix.

$$J(\theta, p) = \det \left[ \frac{\partial \left( H_{11}^{(0)}, \ldots, H_{nn}^{(0)}, H_{12}^{(0)}, H_{12}^{(1)}, \ldots \right)}{\partial \left( \theta_1, \theta_2, \ldots, \theta_n, p_1, p_2, \ldots, p_{n(n-1)} \right)} \right]$$

If one takes $J(\theta, p)$ above, and integrates out the variables $p$, then multiplying by the joint probability density function for matrix entries, one has the joint probability density function for the eigenvalues themselves. It will turn out that $J(\theta, p)$ is $\prod_{i<j}(\theta_j - \theta_i)^2$ times a function of $p$, so that all that remains is a function of $\theta$ times a constant.

We will multiply the above $n^2 \times n^2$ matrix for $J(\theta, p)$ by another matrix of the same dimensions depending only on $U$, and the determinant of the result will be $\prod_{i<j}(\theta_j - \theta_i)^2$ times a function depending only on $U$.

Now $UU^* = 1$, so for each variable $p_\mu$,

$$S(\mu) = U^* \frac{\partial U}{\partial p_\mu} = \frac{\partial U^*}{\partial p_\mu} U$$

is conjugate Hermitian. Then

$$U^* \frac{\partial H}{\partial p_\mu} U = S(\mu) \Theta - \Theta S(\mu),$$

or

$$\sum_{j,k} U_{\alpha,j}^* \frac{\partial H_{j,k}}{\partial p_\mu} U_{k,\beta} = S(\mu)_{\alpha,\beta} (\theta_\beta - \theta_\alpha)$$

$$= \sum_j \frac{\partial H_{j,j}}{\partial p_\mu} \left[ U_{\alpha,j}^* U_{j,\beta} \right] + \sum_{j<k} \Re \left( \frac{\partial H_{j,k}}{\partial p_\mu} \right) \left[ U_{\alpha,j}^* U_{k,\beta} + U_{\alpha,k}^* U_{j,\beta} \right]$$

$$+ i \sum_{j<k} \Im \left( \frac{\partial H_{j,k}}{\partial p_\mu} \right) \left[ U_{\alpha,j}^* U_{k,\beta} - U_{\alpha,k}^* U_{j,\beta} \right]$$

Similarly,

$$\sum_{j,k} U_{\alpha,j}^* \frac{\partial H_{j,k}}{\partial \theta_\gamma} U_{k,\beta} = \delta_{\alpha,\beta,\gamma}.$$
\[
\begin{bmatrix}
\frac{\partial H^{(0)}_{ij}}{\partial \theta_j} & \frac{\partial H^{(0)}_{ik}}{\partial \theta_j} & \frac{\partial H^{(1)}_{jk}}{\partial \theta_j} \\
\frac{\partial H^{(0)}_{ik}}{\partial p_\mu} & \frac{\partial H^{(0)}_{ik}}{\partial p_\mu} & \frac{\partial H^{(1)}_{jk}}{\partial p_\mu}
\end{bmatrix}
\times
\begin{bmatrix}
[U^*_{\alpha,j} U_{j,\alpha}] & \Re[U^*_{\alpha,j} U_{j,\alpha}] & \Im[U^*_{\alpha,j} U_{j,\alpha}]
\\
[U^*_{\alpha,j} U_{k,\beta} + U^*_{\alpha,k} U_{j,\beta}] & \Re[U^*_{\alpha,j} U_{k,\beta} + U^*_{\alpha,k} U_{j,\beta}] & \Im[U^*_{\alpha,j} U_{k,\beta} + U^*_{\alpha,k} U_{j,\beta}]
\\
[U^*_{\alpha,j} U_{k,\beta} - U^*_{\alpha,k} U_{j,\beta}] & -\Im[U^*_{\alpha,j} U_{k,\beta} - U^*_{\alpha,k} U_{j,\beta}] & \Re[U^*_{\alpha,j} U_{k,\beta} - U^*_{\alpha,k} U_{j,\beta}]
\end{bmatrix}
\]

\[= \begin{bmatrix} I_0 & 0 & 0 \\
0 & \Re(S^\mu_{\alpha,\beta})(\theta_\beta - \theta_\alpha) & \Im(S^\mu_{\alpha,\beta})(\theta_\beta - \theta_\alpha) \end{bmatrix}\]

### 14. Appendix: The Method of Gaudin

Suppose that \(F_{m,n}(t_1, t_2, \ldots, t_n)\) is a symmetric function of \(n\) variables, which only depends on the variables \(m\) at a time. Applying \(F_{m,n}\) to the eigenvalues of an \(n \times n\) matrix chosen randomly from some ensemble makes \(F_{m,n}\) a random variable. We express the expected value of \(F_{m,n}\) as an \(m\)-dimensional integral.

#### 14.1. The Universal Unitary Ensemble

Let \(V(x)\) be a real-analytic potential function growing sufficiently rapidly at infinity. Then for the ensemble \(UUE_n\), the j.p.d.f. of the eigenvalues is

\[
\kappa_{1,V,n} \prod_{0 \leq j < k < n} (t_j - t_k)^2 e^{-n \sum V(t_j)} dT,
\]

where \(\kappa_{1,V,n}\) is a constant depending only on \(V(x)\) and \(n\). See Appendix 13.

This leads to an expression for the expected value of \(F_{m,n}\).

\[
E(F_{m,n}) = \kappa_{1,V,n} \int F_{m,n}(t_1, \ldots, t_n) \times \prod_{0 \leq j < k < n} (t_j - t_k)^2 e^{-n \sum V(t_j)} dT
\]

**Theorem 14.1.** Let \(m \geq 1\) and \(F\) a function of \(m\) variables. Define the symmetrization of \(F\) to \(n\) variables as:

\[
F_{m,n}(t_1, t_2, \ldots, t_n) = \sum_{i_1, i_2, \ldots, i_m \text{ distinct}} F(t_{i_1}, t_{i_2}, \ldots, t_{i_m}).
\]

Then the expected value of \(F_{m,n}\), applied to the eigenvalues of a random unitary matrix from \(UUE_n\), can be expressed as an integral in \(m\) variables:

\[
E(F_{m,n}) = \int F(t_1, t_2, \ldots, t_m) \det_{m \times m} [K_n(t_j, t_k)] dt_1 \ldots dt_m.
\]

The projection kernel \(K_n(x, y)\) appearing inside the determinant will be defined below.
14.2. Definition of the Projection Kernel $K_n(x, y)$. We form the sequence of normalized orthogonal polynomials by performing Gram-Schmidt orthonormalization on the sequence $1, x, x^2, x^3, \ldots$, with respect to the inner product

$$(f, g) = \int_{-\infty}^{\infty} f(x)g(x)e^{-nV(x)}dx.$$ 

We always choose the sign of the leading coefficient of the polynomials to be positive. We call the resulting sequence $\phi_{n,j}(x)$ the normalized orthogonal polynomials with respect to the measure $e^{-nV(x)}dx$.

For each polynomial we have a choice of sign – we choose so that the leading coefficient of each polynomial is positive. Let the resulting sequence of polynomials be $\phi_j$. Thus,

$$\int_{-\infty}^{\infty} \phi_j(x)\phi_k(x)e^{-nV(x)}dx = \delta_{j,k}.$$ 

An alternative normalization is to choose the leading coefficient of each polynomial to be 1. This would result in monic orthogonal polynomials. We will use $\phi_{n,j}$ to denote normalized orthogonal polynomials and $\pi_{n,j}$ to denote monic orthogonal polynomials. When the context is clear, we will omit the subscript $n$.

The most familiar examples of orthogonal polynomials are Hermite polynomials. Hermite polynomials are orthogonal polynomials with respect to $e^{-x^2}dx$. They are normalized so that the leading term is $2^j x^j$.

Lemma 14.2 (Christoffel-Darboux + three-term recurrence relation). Using the above notation, let

$$a_k = \int x\phi_k^2(x)e^{-nV(x)}dx \quad \text{for} \quad k \geq 0$$

$$b_k = \int x\phi_{k+1}(x)\phi_k(x)e^{-nV(x)}dx \quad \text{for} \quad k \geq 0.$$ 

Then

$$x\phi_0(x) = b_0\phi_1(x) + a_0\phi_0(x)$$

$$x\phi_k(x) = b_k\phi_{k+1}(x) + a_k\phi_k(x) + b_{k-1}\phi_{k-1}(x) \quad \text{for} \quad k \geq 1.$$ 

and

$$\sum_{j=0}^{n-1} \phi_j(x)\phi_j(y) = b_{n-1} \frac{\phi_n(x)\phi_{n-1}(y) - \phi_n(y)\phi_{n-1}(x)}{x-y}.$$ 

Proof. To prove the three term recurrence relation, observe that $x\phi_k(x)$ is a polynomial of degree $(k + 1)$, so it is orthogonal to all polynomials of degree higher than $(k + 1)$. It is also orthogonal to polynomials of degree $j$
less than \((k - 1)\):

\[
\int (x\phi_k(x))\phi_j(x)d\mu(x) = \int \phi_k(x)(x\phi_j(x))d\mu(x) = 0.
\]

The second integral above is zero because \((x\phi_j)\) is a polynomial of degree at most \((k - 1)\), and \(\phi_k(x)\) is orthogonal to polynomials of such low degree. The Christoffel-Darboux formula itself is proven by induction. The inductive step is the following:

\[
(x - y) \left( b_k \frac{\phi_{k+1}(x)\phi_k(y) - \phi_{k+1}(y)\phi_k(x)}{x - y} - b_{k-1} \frac{\phi_k(x)\phi_{k-1}(y) - \phi_k(y)\phi_{k-1}(x)}{x - y} \right)
\]

\[
= (b_k \phi_{k+1}(x) + b_{k-1} \phi_{k-1}(x)) \phi_k(y) - \phi_k(x) (b_k \phi_{k+1}(y) + b_{k-1} \phi_{k-1}(y))
\]

\[
= (b_k \phi_{k+1}(x) + b_{k-1} \phi_{k-1}(x) + a_k \phi_k(x)) \phi_k(y) - \phi_k(x) (b_k \phi_{k+1}(y) + b_{k-1} \phi_{k-1}(y) + a_k \phi_k(y))
\]

\[
= x\phi_k(x)\phi_k(y) - \phi_k(x)y\phi_k(y)
\]

\[
= (x - y)\phi_k(x)\phi_k(y).
\]

\[\Box\]

For computations elsewhere, we find it more convenient to use \(\eta_{n,j}(x)\), where

\[
\eta_{n,j}(x) = \frac{\phi_{n,j}(x)}{e^{\frac{1}{2}V(x)}}.
\]

The advantage of using \(\eta_{n,j}(x)\) is that they are orthonormal with respect to Lebesgue measure on \(\mathbb{R}\). Again, we will omit the subscript \(n\) when the context is clear. Using this notation, let

\[
K_n(x, y) = \sum_{j=0}^{n-1} \eta_j(x)\eta_j(y)
\]

\[
= b_{n-1} \frac{\eta_n(x)\eta_{n-1}(y) - \eta_n(y)\eta_{n-1}(x)}{x - y}.
\]

14.3. Proof of Theorem 14.1. We will use the following Lemma to express \(E(F_{m,n})\) as the integral of an \(n \times n\) determinant:

Lemma 14.3. Let \(K_n(x, y)\) be the projection kernel above. Then:

\[
\prod_{0 \leq j < k < n} (t_j - t_k)^2 e^{-n\sum V(t_j)} = \kappa_{2,V,n}^2 \det_{n \times n}[K_n(t_j, t_k)],
\]

where \(\kappa_{2,V,n}\) depends only on the potential \(V(x)\) and \(n\).
Proof. We recognize the component \(\prod_{0 \leq j < k < n} (t_j - t_k)\) as a Vandermonde determinant:

\[
\prod_{1 \leq j < k \leq n} (t_j - t_k) = \det \begin{bmatrix}
1 & t_1 \ldots t_1^{n-1} \\
\vdots & \vdots & \vdots \\
1 & t_n \ldots t_n^{n-1}
\end{bmatrix}
\]

Performing column operations we replace the monomials \(t_k^j\) with orthogonal polynomials \(\pi_{n,k}(t_j)\). To replace the monic orthogonal polynomials \(\pi_{n,k}\) with the normalized orthogonal polynomials \(\varphi_{n,k}\), we divide by the leading coefficients. We next replace \(\varphi_j\) with \(\eta_j\) in order to absorb a factor of \(e^{-\frac{n}{2} \sum V(t)}\):

\[
\prod_{1 \leq j < k \leq n} (t_j - t_k) e^{-\frac{n}{2} \sum V(t)}
\]

\[
= \det \begin{bmatrix}
\pi_0(t_1) & \pi_1(t_1) \ldots \pi_{n-1}(t_1) \\
\vdots & \vdots \vdots \\
\pi_0(t_n) & \pi_1(t_n) \ldots \pi_{n-1}(t_N)
\end{bmatrix} e^{-\frac{n}{2} \sum V(t)}
\]

\[
= \kappa_{2,V,n} \det \begin{bmatrix}
\phi_0(t_1) & \phi_1(t_1) \ldots \phi_{n-1}(t_1) \\
\vdots & \vdots \vdots \\
\phi_0(t_n) & \phi_1(t_n) \ldots \phi_{n-1}(t_N)
\end{bmatrix} e^{-\frac{n}{2} \sum V(t)}
\]

\[
= \kappa_{2,V,n} \det \begin{bmatrix}
\eta_0(t_1) & \eta_1(t_1) \ldots \eta_{n-1}(t_1) \\
\vdots & \vdots \vdots \\
\eta_0(t_n) & \eta_1(t_n) \ldots \eta_{n-1}(t_N)
\end{bmatrix}
\]

\[
= \kappa_{2,V,n} \det[M].
\]

Squaring the above formula,

\[
\prod_{1 \leq j < k \leq n} (t_j - t_k)^2 e^{-\frac{n}{2} \sum V(t)}
\]

\[
= \kappa_{2,V,n}^2 \det(M)^2
\]

\[
= \kappa_{2,V,n}^2 \det(MM^T)^2
\]

\[
= \kappa_{2,V,n}^2 \det[K_{n}(t_j, t_k)]
\]

\[\square\]

We now use the results of Lemma 14.3 to express \(E(F_{m,n})\) as the integral of an \(n \times n\) determinant:

\[
E(F_{m,n}) = \frac{\kappa_{1,V,n}}{\kappa_{2,V,n}^2} \int_{F_{m,n}(t_1, \ldots, t_n)} \det[K_{n}(t_j, t_k)] dt_1 dt_2 \ldots dt_n
\]
Unfortunately, this is an $n$ dimensional integral instead of $m$ dimensional. We use the next Lemma to “integrate out” the extra dimensions one at a time.

**Lemma 14.4.** Let $K_n(x, y) = \sum_{j=0}^{n-1} \eta_j(x)\eta_j(y)$ be the projection of $n$ functions orthonormal with respect to Lebesgue measure. Then integrating

$$\det_{l \times l}[K_n(t_j, t_k)]$$

over the last variable $t_l$ is the same as multiplying by $(n - l + 1)$. That is,

$$\int_{-\infty}^{\infty} \det_{l \times l}[K_n(t_j, t_k)]dt_l = (n - l + 1) \det_{(l-1) \times (l-1)}[K_n(t_j, t_k)].$$

**Proof.** Expand by minors along the last column. The last entry in this column contributes

$$\frac{\det_{(l-1) \times (l-1)}[K_n(t_j, t_k)]}{\det_{(l-1) \times (l-1)}[K_n(t_j, t_k)]} \left( \int K_n(t_l, t_l)dt_l \right) = n \frac{\det_{(l-1) \times (l-1)}[K_n(t_j, t_k)]}{\det_{(l-1) \times (l-1)}[K_n(t_j, t_k)]}.$$

Now for the $(m - 1)$ other minors. Consider the minor by expanding along $K_n(t_j, t_l)$, where $1 \leq j \leq (l - 1)$:

$$(-1)^{l-j} \int \det \begin{bmatrix} K_n(t_1, t_1) & K_n(t_1, t_2) & \ldots & K_n(t_1, t_{l-1}) \\ & & & \\ & & & \\ K_n(t_{j-1}, t_1) & K_n(t_{j-1}, t_2) & \ldots & K_n(t_{j-1}, t_{l-1}) \\ K_n(t_{j+1}, t_1) & K_n(t_{j+1}, t_2) & \ldots & K_n(t_{j+1}, t_{l-1}) \\ & \vdots & \ddots & \vdots \\ K_n(t_l, t_1) & K_n(t_l, t_2) & \ldots & K_n(t_l, t_{l-1}) \end{bmatrix} K_n(t_j, t_l)dt_l$$

Each of the $(l - 1)!$ terms in the determinant has exactly one term in row $l$, say $K_n(t_l, t_l)$, but otherwise has no dependence of $t_l$. To integrate this term against $K_n(t_j, t_l)dt_l$, observe that $K_n(x, y)$ is a projection kernel:

$$\int K_n(x, y)K_n(y, z)dy = K_n(x, z).$$

The resulting contribution is:

$$(-1)^{l-j} \det \begin{bmatrix} K_n(t_1, t_1) & K_n(t_1, t_2) & \ldots & K_n(t_1, t_{l-1}) \\ & & & \\ & & & \\ K_n(t_{j-1}, t_1) & K_n(t_{j-1}, t_2) & \ldots & \ldots & K_n(t_{j-1}, t_{l-1}) \\ K_n(t_{j+1}, t_1) & K_n(t_{j+1}, t_2) & \ldots & \ldots & K_n(t_{j+1}, t_{l-1}) \\ & \vdots & \ddots & \ddots & \vdots \\ K_n(t_j, t_1) & K_n(t_j, t_2) & \ldots & \ldots & K_n(t_j, t_{l-1}) \\ & & & & \\ K_n(t_{l-1}, t_1) & K_n(t_{l-1}, t_2) & \ldots & \ldots & K_n(t_{l-1}, t_{l-1}) \end{bmatrix}$$

$$= - \det \begin{bmatrix} K_n(t_1, t_1) & K_n(t_1, t_2) & \ldots & K_n(t_1, t_{l-1}) \\ & & & \\ & & & \\ K_n(t_{j-1}, t_1) & K_n(t_{j-1}, t_2) & \ldots & \ldots & K_n(t_{j-1}, t_{l-1}) \\ & & & & \\ & & \vdots & \ddots & \vdots \\ K_n(t_j, t_1) & K_n(t_j, t_2) & \ldots & \ldots & K_n(t_j, t_{l-1}) \\ & & & & \\ K_n(t_{l-1}, t_1) & K_n(t_{l-1}, t_2) & \ldots & \ldots & K_n(t_{l-1}, t_{l-1}) \end{bmatrix}.$$
Adding the contribution from the last minor to the other \((m-1)\) minors yields the conclusion of the Lemma.

\[\square\]

Using Lemmas 14.3 and 14.4 allows an induction on the following statement for \(m \leq l \leq n\):

\[E(F_{m,n}) = \frac{\kappa_{1,V,n}}{\kappa_{2,V,n}} \frac{(n-l)!}{(n-m)!} \int F(t_1, t_2, \ldots, t_m) \det_{l \times l} [K_n(t_j, t_k)] dt_1 dt_2 \ldots dt_l\]

To prove the base case \(l = n\), observe that \(F_{m,n}\) is a sum of \(\frac{n!}{(n-m)!}\) terms. Each of these terms contributes as much to \(E(F_{m,n})\) as the term \(F(t_1, t_2, \ldots, t_m)\) does. Specializing to the case \(l = m\), we obtain:

\[E(F_{m,n}) = \frac{\kappa_{1,V,n}}{\kappa_{2,V,n}} \int F(t_1, t_2, \ldots, t_m) \det_{m \times m} [K_n(t_j, t_k)] dt_1 \ldots dt_m.\]

This proves Theorem 14.1 except for the presence of a multiplicative constant depending only on \(V\) and \(n\). To find this constant, let \(F(x) \equiv 1\) be a function of one variable. Then

\[E(F_{m,n}) = \sum_{1 \leq j \leq n} 1 = n\]

and

\[\int K_n(x,x)dx = n,\]

establishing that the multiplicative constant is 1.

### 14.4. The Circular Unitary Ensemble.

Applying the method of Gaudin to the circular unitary ensemble (CUE) yields the following Theorem, which differs from Theorem 14.1 only in the definition of the projection kernel \(K_n(x,y)\).

**Theorem 14.5** (Method of Gaudin, CUE version). Let \(m \geq 1\) and \(F\) a function of \(n\) variables. Define the symmetrization of \(F\) to \(n\) variables as:

\[F_{m,n}(t_1, t_2, \ldots, t_n) = \sum_{i_1, i_2, \ldots, i_m \text{ distinct}} F(t_{i_1}, t_{i_2}, \ldots, t_{i_m}).\]

Then the expected value of \(F_{m,n}\), applied to the eigenvalues of a random unitary matrix from \(U_n\), can be expressed as an integral in \(m\) variables:

\[E(F_{m,n}) = \int F(t_1, t_2, \ldots, t_m) \det_{1 \leq j, k \leq m} [K_n(t_j, t_k)] dt_1 \ldots dt_m,\]

where the projection kernel is

\[K_n(x,y) = \frac{1}{2\pi} \frac{e^{in(x-y)} - 1}{e^{i(x-y)} - 1}.\]
Proof. For the CUE, the j.p.d.f. for the eigenvalues is
\[
\kappa_3, n \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2
\]  
\[
= \kappa_3, n \prod_{j<k} (e^{i\theta_j} - e^{i\theta_k}) \prod_{j<k} (e^{-i\theta_j} - e^{-i\theta_k})
\]  
\[
= \kappa_3, n \begin{pmatrix}
1 & e^{i\theta_1} & \cdots & e^{i(n-1)\theta_1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & e^{i\theta_n} & \cdots & e^{i(n-1)\theta_n}
\end{pmatrix}
\begin{pmatrix}
1 & \cdots & 1 \\
e^{-i\theta_1} & \cdots & e^{-i\theta_n} \\
e^{-i(n-1)\theta_1} & \cdots & e^{-i(n-1)\theta_n}
\end{pmatrix}
\]  
\[
= \kappa_{4,n} \det[K_n(\theta_j, \theta_k)],
\]  
where
\[
K_n(\theta_1, \theta_2) = \frac{1}{2\pi} \frac{e^{in(\theta_1-\theta_2)} - 1}{e^{i\theta_1} - e^{i\theta_2}}.
\]  

Proceeding as in the proof of Theorem 14.1, one must verify that $K_n$ is a projection kernel:
\[
\int K_n(\theta_1, \theta_2)K_n(\theta_2, \theta_3)dy = K_n(\theta_1, \theta_3).
\]  

It is this requirement which dictates the choice of normalization constant for $K_n$. The remainder of the proof is identical to the proof of Theorem 14.1. □

15. Appendix: One Possible Approach to the Maximum Spacing Problem

In this paper, we considered the following question: given an interval $I$ and an ensemble of random matrices, what is the minimum spacing between two consecutive eigenvalues in the interval $I$?

It is natural to ask the same question about maximum spacing: what is the maximum spacing between two consecutive eigenvalues in the interval $I$?

We suspect that maximum spacing is a more difficult question than minimum spacing; at least we have not been able to deal with it yet. We begin with heuristic predictions based on the tail probabilities of the GUE consecutive spacing distribution. As far as we can tell, even these tail probabilities have not been proven rigorously.

In a region of unit mean spacing, in an appropriate limit, the probability that an interval $I$ is free of eigenvalues is the Fredholm determinant of the integral operator
\[
(Id - A_I \psi)(x) = \psi(x) - \int I K(x, y)\psi(y)dy,
\]
where $K(x, y)$ is the sine kernel

$$K(x, y) = \frac{\sin(x - y)}{\pi(x - y)}.$$ 

In the above statement and formula, we could replace $I$ by a union $J$ of intervals. See [Wid95] Let $F(J)$ be the value of this Fredholm determinant.

For a single interval, $F(I)$ depends only on the length of the interval $s = |I|$. Let $E_2(0; s)$ be this probability. Then the consecutive spacing distribution is

$$p_2(0; s) = \frac{\partial^2}{\partial s^2} E_2(0; s).$$

See [Meh91], Chapter 5 and Appendix 13. In the minimum spacing problem, we used the Taylor expansions of $p_2(0; s)$ centered at $s = 0$:

$$E_2(0; 2) = 1 - s + \frac{\pi^2 s^4}{36} - \frac{\pi^4 s^6}{675} + \frac{\pi^6 s^8}{17640} + \ldots$$

$$p_2(0; s) = \frac{\pi^2 s^2}{3} - \frac{2\pi^4 s^4}{45} + \frac{\pi^6 s^6}{315} + \ldots$$

For maximum spacing, we require the asymptotics of $p_2(0; s)$ for large $s$. Dyson [Dys76] derived the following expansion for large $s$:

$$\log(E_2(0; s)) = -\frac{s^2}{8} - \frac{1}{4} \log(s) + 3\zeta'(-1) + \frac{1}{3} \log(2) + o(1).$$

Unfortunately Dyson’s methods did not yield rigorous asymptotics. The first rigorous asymptotics are more limited, and were obtained only recently by Widom [Wid94]:

$$\frac{\partial}{\partial s} \log(E_2(0; s)) = \frac{s}{4} + O(1)$$

$$\implies \log(E_2(0; s)) = -\frac{s^2}{8} + O(s)$$

Since we are only doing heuristics, we choose to use Dyson’s expansion. Truncating this expansion and differentiating twice we obtain asymptotics for $p_2(0; s)$ for large $s$:

$$\int_s^\infty p_s(0; t) dt = \left( \frac{s^3}{4} + 2 \frac{s^3}{4} \right) e^{-\frac{s^2}{8} + 3\zeta'(-1) + \frac{1}{3} \log(2)}$$

$$p_2(0; s) = \left( \frac{s^7}{16} - 2 \frac{s^7}{16} \right) e^{-\frac{s^2}{8} + 3\zeta'(-1) + \frac{1}{3} \log(2)}$$

$$= \frac{s^7}{16} e^{-\frac{s^2}{8} + 3\zeta'(-1) + \frac{1}{3} \log(2)}$$

As in the minimum spacing problem, we assume that the consecutive spacings are independent random variables, aware that the assumption is false but convenient. Recycling notation from the minimum spacing problem, let $G_{\gamma, I, n}$ be the number of consecutive spacings, of a random matrix from
an ensemble of $n \times n$ matrices, in $I_n$, which are larger than $\gamma$. Let $Z_n$ be the maximum spacing itself. Then, assuming a sum of independent unlikely events, $G_{\gamma,I,n}$ will be approximately Poisson with mean

$$\mu = \int F_n(x,x) \frac{\left(\gamma K_n(x,x)\right)^2}{4} e^{-\frac{\gamma K_n(x,x)}{8} + \frac{3\zeta(-1)}{4} + \frac{1}{2} \log(2)} \, dx$$

Then $\Pr(Z_n < \gamma) = \Pr(G_{\gamma,I,n} = 0) \approx e^{-\mu}$ is the probability that the maximum spacing is less than $\gamma$. We specialize to the case of constant eigenvalue density 1, so that

$$\mu = |I| \frac{\gamma^2}{16} e^{-\frac{\gamma^2}{8} + \frac{3\zeta(-1)}{4} + \frac{1}{2} \log(2)}.$$

Our proposed strategy for recovering $\Pr(Z_n < \gamma)$ is less direct than in the minimum spacing problem. Let $S_{\gamma,I,n}$ be the set

$$S_{\gamma,I,n} = \left\{ x \mid \left[x, x + \gamma\right] \subset I \text{ and } \left[x, x + \gamma\right] \text{ contains no eigenvalues.} \right\} = \left\{ x \mid x \in \tilde{I} \text{ and } \left[x, x + \gamma\right] \text{ contains no eigenvalues.} \right\}$$

Let $X_\gamma = |S_{\gamma,I,n}|$ be the size of the set $S_{\gamma,I,n}$. We wish to estimate $\Pr(X_\gamma = 0)$, which we bound using the following inequalities:

$$\Pr(X_\gamma - \epsilon \leq \epsilon) \leq \Pr(X_\gamma = 0) \leq \Pr(X_\gamma \leq \epsilon).$$

It may be possible to estimate $\Pr(X_\gamma \leq \epsilon)$ and $\Pr(X_\gamma - \epsilon \leq \epsilon)$ using estimates for the moments of $X_\gamma$ and $X_\gamma - \epsilon$, respectively.

The $(k)$th moment of $X_\gamma$ is

$$\int \prod_{j=1}^k F([x_1, x_1 + \gamma] \cup \cdots \cup [x_k, x_k + \gamma]) \, dx_1 \, dx_2 \cdots \, dx_k.$$ We expect that the dominant contributions to this integral occur when the $x_j$ are “clustered together” according to some partition of $\{1, 2, \ldots, k\}$. For example, consider the contribution to $E(X_\gamma^4)$ corresponding to the partition $(13|2|4)$. The region $\Omega_{(13|2|4)} \subset \tilde{I}^4$ corresponding to this partition is

$$\Omega_{(13|2|4)} \subset \tilde{I}^4 = \left\{ (x_1, x_2, x_3, x_4) \in \tilde{I}^4 \mid \begin{array}{c} |x_1 - x_3| \text{ small} \\ |x_1 - x_2|, |x_1 - x_4| \text{ large} \end{array} \right\}$$

Since $|x_1 - x_3|$ is small, $[x_1, x_1 + \gamma] \cup [x_3, x_3 + \gamma]$ is a single interval of length $\gamma + |x_1 - x_3|$. Thus $F([x_1, x_1 + \gamma] \cup [x_3, x_3 + \gamma]) = E_2(0; \gamma + |x_1 - x_3|)$, which is easily calculated and compared to $E_2(0; \gamma)$.

When intervals $I_1$ and $I_2$ are well separated, it should be the case that

$$F(I_1 \cup I_2) \approx F(I_1) \cdot F(I_2),$$
and similarly for more than two intervals. We call this *approximate splitting*. Since $|x_1 - x_2|$ and $|x_1 - x_4|$ are large, we have three well-separated intervals. Thus,

$$F ([x_1, x_1 + \gamma] \cup [x_2, x_2 + \gamma] \cup [x_3, x_3 + \gamma] \cup [x_4, x_4 + \gamma])$$

$$\approx F ([x_1, x_1 + \gamma] \cup [x_3, x_3 + \gamma]) \cdot F ([x_2, x_2 + \gamma]) \cdot F ([x_4, x_4 + \gamma])$$

$$\approx E_2(0; \gamma + |x_1 - x_3|) \times E_2(0; \gamma) \times E_2(0; \gamma).$$

The integral over $\Omega_{(13|24)}$ can then be approximated by a product of three more easily evaluated integrals.

We call the readers attention to [Wid95]. It seems that these asymptotics are for unions of intervals like $[0, n] \cup [2n, 3n]$. That is, the ratios of length and separation remain fixed and $x \to \infty$. For approximate splitting, however, we require asymptotics where the intervals are of fixed length and they are moved far apart.

Before attempting to show that the remainder of $\tilde{I}_k$ makes a negligible contribution to $E(X_k^F)$, one would first have to quantify “large” and “small” in the definition of $\Omega_{(13|24)}$ and the in the regions corresponding to other partitions. The definition of “small” would depend on how rapidly $E_2(0; \gamma + \epsilon)$ decays. The definition of “large” would depend on the details of the approximate splitting conjecture.

Here is one approach to proving the approximate splitting

$$F(I_1 \cup I_2) \approx F(I_1) \cdot F(I_2).$$

Let $\alpha_{1,j}$ be the eigenvalues of $A_{I_1 \cup I_2}$, $\alpha_{1,j}$ the eigenvalues of $A_{I_1}$, and $\alpha_{2,k}$ the eigenvalues of $A_{I_2}$. We must show that

$$\prod_i (1 - \alpha_{12,i}) \approx \prod_j (1 - \alpha_{1,j}) \prod_k (1 - \alpha_{2,k}).$$

Let us take for granted that all of the $\alpha$ are positive and less than 1; perhaps this is something that could be proven.

We expect that for the large $\alpha_{1,j}$ and $\alpha_{2,k}$ there correspond $\alpha_{12,i}$. If this were true for all $\alpha$, it would prove approximate separation. Regarding the 1−1 correspondence, there is a subtlety which we illustrate in Figure 4.

It seems plausible that the large $\alpha_{1,j}$ are well separated from other $\alpha_{1,j}$, and that this could be proven. Similarly for the $\alpha_{2,k}$. However, there is no reason why one of the $\alpha_{1,j}$ cannot be close to one of the $\alpha_{2,k}$. If we only have assurances that one eigenvalue $\alpha_{12,i}$ is nearby, then we are missing one. We therefore need a multiple eigenvalue version of Weyl’s Criterion:

**Weyl’s Criterion:** Let $T$ be a self adjoint operator on a Hilbert space and $\psi$ and $\lambda$ such that $|T\psi - \lambda \psi| \leq \epsilon |\psi|$, where $\lambda$ is real. Then the intersection of the spectrum of $T$ with the interval $[\lambda - \epsilon, \lambda + \epsilon]$ is nonempty.
Figure 4. The correspondence between eigenvalues of $A_{I_1}$, $A_{I_2}$, and $A_{I_1 \cup I_2}$. When the circles are disjoint, there is no danger of confusion and we have a one-to-one correspondence of eigenvalues.

Embarrassingly, the author has been unable to find a reference for Weyl’s Criterion above, or the following statement. We do not know whether the following statement is true as stated or whether or not it is known.

**Statement:** For $\epsilon > 0$ sufficiently small, there is a constant $C_\epsilon$ so that the following is true. Suppose that $T$ is a self-adjoint operator on a Hilbert space $\mathcal{H}$. Suppose $\psi_1, \psi_2 \in \mathcal{H}$ and $\lambda \in \mathcal{R}$ are such that:

$$
|T\psi_1 - \lambda \psi_1| \leq \epsilon |\psi_1|
$$

$$
|T\psi_2 - \lambda \psi_2| \leq \epsilon |\psi_2|
$$

$$
|(\psi_1, \psi_2)| \leq \epsilon |\psi_1| \cdot |\psi_2|
$$

Then the interval $[\lambda - C_\epsilon, \lambda + C_\epsilon]$ contains at least two eigenvalues of $T$. The constants $C_\epsilon$ may be chosen so that $C_\epsilon \to 0$ as $\epsilon \to 0$.

To apply the multiple eigenvalue statement to the current situation, $T = A_{I_1 \cup I_2}$, and $\lambda$ be the average value of two nearby eigenvalues $\alpha_{1,j}$ and $\alpha_{2,k}$. Let $\psi_1$ and $\psi_2$ be the corresponding eigenfunctions for $A_{I_1}$ and $A_{I_2}$. Then $\psi_1$ should be localized around $I_1$ and $\psi_2$ should be localized around $I_2$. Hence $\psi_1$ and $\psi_2$ are nearly orthogonal, and also $A_{I_2} \psi_1$ and $A_{I_1} \psi_2$ are small, satisfying the hypotheses of the multiple eigenvalue statement.

Thus for large $\alpha$, there is a one to one correspondence between eigenvalues of either $A_{I_1}$ or $A_{I_2}$ and eigenvalues of $A_{I_1 \cup I_2}$. For small $\alpha$, however, this correspondence may no longer hold. Fortunately the small $\alpha$ cannot contribute much to any of the products $\prod (1 - \alpha)$. The reason is that $A_{I_1 \cup I_2}$ has trace $|I_1| + |I_2|$ and also has, we hope, all positive eigenvalues. The large eigenvalues $\alpha_{12,i}$ have accounted for all but a small amount of this sum, and so the product $\prod (1 - \alpha_{12,i})$ over the remaining $\alpha$ is close to one. Similarly for the other two products.
References

[BO78] Carl M. Bender and Steven A. Orszag. *Advanced mathematical methods for scientists and engineers*. McGraw-Hill Book Co., New York, 1978. International Series in Pure and Applied Mathematics.

[Dei99] P. A. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*. New York University Courant Institute of Mathematical Sciences, New York, 1999.

[DKM+97] P. Deift, T. Kriecherbauer, K. T-R McLaughlin, S. Venakides, and X. Zhou. Asymptotics for polynomials orthogonal with respect to varying exponential weights. *Internat. Math. Res. Notices*, (16):759–782, 1997.

[DKM98] P. Deift, T. Kriecherbauer, and K. T.-R. McLaughlin. New results on the equilibrium measure for logarithmic potentials in the presence of an external field. *J. Approx. Theory*, 95(3):388–475, 1998.

[DKM+99a] P. Deift, T. Kriecherbauer, K. T-R McLaughlin, S. Venakides, and X. Zhou. Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.*, 52(12):1491–1552, 1999.

[DKM+99b] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.*, 52(11):1335–1425, 1999.

[Dys76] Freeman J. Dyson. Fredholm determinants and inverse scattering problems. *Comm. Math. Phys.*, 47(2):171–183, 1976.

[Ede92] Alan Edelman. On the distribution of a scaled condition number. *Math. Comp.*, 58(197):185–190, 1992.

[Meh91] Madan Lal Mehta. *Random matrices*. Academic Press Inc., Boston, MA, second edition, 1991.

[Sze75] Gábor Szegő. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[Wid94] Harold Widom. The asymptotics of a continuous analogue of orthogonal polynomials. *J. Approx. Theory*, 77(1):51–64, 1994.

[Wid95] Harold Widom. Asymptotics for the Fredholm determinant of the sine kernel on a union of intervals. *Comm. Math. Phys.*, 171(1):159–180, 1995.