Chapter 3 from AGZ: Hermite Polynomials, Spacings and Limit Distributions for the Gaussian Ensembles

February 6, 2018

1 Summary of main results: spacing distributions in the bulk and edge of the spectrum for the Gaussian ensembles

We recall that the N eigenvalues of the GOE/GUE/GSE are spread out on an interval of width roughly equal to $4\sqrt{N}$ and hence the spacing between the eigenvalues ought to be of order $1/\sqrt{N}$

1.1 Limit result for the GUE

Using the determinental structure of the eigenvalues $\{\lambda_1^N, \ldots, \lambda_N^N\}$ of the GUE, we will develope in sections 3.2-3.4 we will prove the following:

Theorem. *(3.1.1.) (Gaudin-Mehta) For any compact set $A \subset \mathbb{R}$:*

$$\lim_{N\to\infty} \mathbf{P}\left(\sqrt{N}\lambda_1^N, \sqrt{N}\lambda_2^N, \ldots, \sqrt{N}\lambda_N^N \notin A\right) = 1+\sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \int_A \cdots \int_A \det_{i,j=1}^{k} K_{sine}(x_i, x_j)\, dx_1\, dx_2 \ldots dx_k$$

where:

$$K_{sine}(x,y) = \frac{1}{\pi} \frac{\sin(x-y)}{x-y}$$

and is understood to be $1/\pi$ if $x = y$ (i.e. determined by continuity)

As a consequence of this theorem, we will show that the theory of integrable systems applies and we will get the following result for *eigenvalues in the bulk*:

Theorem. *(3.1.2) (Jimbo-Miwa-Mori-Sato) One has:*

$$\lim_{N\to\infty} \mathbf{P}\left[\sqrt{N}\lambda_1^N \ldots \sqrt{N}\lambda_N^N \notin \left(-\frac{t}{2}, \frac{t}{2}\right)\right] = 1 - F(t)$$

with:

$$1 - F(t) = \exp\left(\int_0^t \frac{\sigma(x)}{x} dx\right)$$

With σ the solution of a Painleve V formula and has asymptotics:

$$\sigma(t) = -\frac{t}{\pi} - \frac{t^2}{\pi^2} - \frac{t^3}{\pi^3} + O(t^4) \text{ as } t \downarrow 0$$

We will also develop results for *eigenvalues at the edge of the spectrum* (i.e. the top eigenvalue) Before we can say what this is, we will make some notation.

Definition. (3.1.3.) The **Airy function** is defined by the formula:

$$\mathrm{Ai}(x) = \frac{1}{2\pi i}\int_C \exp\left(\frac{1}{3}z^3 - xz\right) \mathrm{d}z$$

where C is the contour in the $z-$plane consititing of two rays, one from the direction $e^{-\pi i/3}$ coming from infinity to the origin and one from the origin to $e^{\pi i/3}$ to ∞.

The **Airy kernal** is defined by:

$$K_{Airy}(x,y) = A(x,y) := \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y}$$

with the value of $x = y$ determined by continuity.

Remark. By differentiating under the integral sign, we can check that the Airy function $\mathrm{Ai}(x)$ satisfies the differential equation $y = \mathrm{Ai}(x)$:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - xy = 0$$

We will look at further properties of the airy function later in the chapter.

The fundemental result for the eigenvalues of the GUE at the edge of the spectrum is:

Theorem. *(3.1.4.) [The top eigevalue of the GUE] For all $-\infty < t \le t' \le \infty$ we have:*

$$\lim_{N\to\infty} \mathbf{P}\left[N^{2/3}\left(\frac{\lambda_i^N}{\sqrt{N}} - 2\right) \notin [t,t'], i = 1\ldots, N\right] = 1 + \sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_t^{t'}\int_t^{t'}\cdots\int_t^{t'} \det_{i,j=1}^{k} A(x_i,x_j)dx_1dx_2\ldots dx_k$$

With A the Airy kernal. In particular, the TOP eigevalue λ_N^N has:

$$\begin{aligned}\lim_{N\to\infty} \mathbf{P}\left[N^{2/3}\left(\frac{\lambda_N^N}{\sqrt{N}} - 2\right) \le t\right] &= 1 + \sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_t^{\infty}\int_t^{\infty}\cdots\int_t^{\infty} \det_{i,j=1}^{k} A(x_i,x_j)dx_1dx_2\ldots dx_n \\ &=: F_2(t)\end{aligned}$$

$F_2(t)$ is the celebrated Tracy-Widom distribution.

Remark. The statement of theorem 3.1.4 does not ensure that F_2 is a proper distribution function (For example if you apply the Helly selection theorem to a sequence of measures, you get convergence to something for a subsequence, of the form $\mathbf{P}(X_{n_k} \leq t) \to F(t)$ but you are not guarenteed that the limiting object is a proper probability distribution unless you know X_n is tight, or have some additional information that stops mass from leaking away. An example is $X_n \sim N(0, t)$ will converge to $F(t) = \frac{1}{2}$....or boxes $X_n = 1_{[n,n+1]}$ will converge to 0. This type of convregence is called VAGUE convergence....this is convergence under expectation for test functions with COMPACT support. Weak convergece we need it to work for any continuous function. Since arbitary continus function need not vanish at ∞, we need to be sure mass isn't leaking away)

The following theorem clarifies the situation! As it happens, F_2 is indeed a proper probability distribution and the convernce above is indeed weak convergnece.

Theorem. *(3.1.5.) (The Tracy Widom distribution) The function $F_2(\cdot)$ is a proper distribution function that admits the representation:*

$$F_2(t) = \exp\left(-\int_t^\infty (x-t)q(x)^2\,dx\right)$$

Where q satisfies:

$$q'' = tq + 2q^3$$

and will have $q(t) \sim Ai(t)$ as $t \to +\infty$.

Remark. The function $F_2(\cdot)$ is called the **Tracy-Widom** distribution and the differential equation its solves is the Painleve II equation.

1.2 Generalizations: Limit formulas for the GOE and GSE

I'm going to skip this for now and focus on the GUE. There is a nice analaogy using a parameter β which is 1 for the GOE, 2 for the GUE, and 4 for the GSE.

2 Hermite Polynomials and the GUE

In this section we will show why orthongonal polynomials arise naturally in the study of the law of the GUE. The relavent orthogonal polynomials in this study are the Hermite polynomials and the associated oscillator wave-functions. We introduce these here and use these tools to derive the Fredholm determinant representation for some GUE things.

2.1 The GUE and determinantal laws

We will show that the joint distribution of the eigenvalues for the GUE has a nice formula as a determinantal point process. (Its actually a determinental projection process). This will lead us to the representation of the eigenvalues in the bulk as a Fredhold determinant. (this will be done in Lemma 3.2.2. and Lemma 3.2.4.)

Before we start, lets clarify some notation we will use in this section.

Definition. We consider

the GUE matrix with complex Gaussian entries of unit variance. Specfically, if $\xi_{i,j}$and $\eta_{i,j}$ are sequences of iid $N(0,1)$ Gaussians, then our matrices look like:

$$\begin{bmatrix} (\xi_{1,1}) & \frac{1}{\sqrt{2}}\left(\xi_{1,2}+i\eta_{1,2}\right) & \frac{1}{\sqrt{2}}\left(\xi_{1,3}+i\eta_{1,3}\right) \\ * & (\xi_{2,2}) & \frac{1}{\sqrt{2}}\left(\xi_{2,3}+i\eta_{2,3}\right) \\ * & * & (\xi_{3,3}) \end{bmatrix}$$

The terms below the diagonal are mirrored to make the matrix Hermitian ($A_{ij}=\overline{A_{ji}}$). We will briefly recall that the GUE is special due to the fact that it is invariant under conjugation with unitary matrices (i.e. if X is distributed like a GUE and U is unitary, then $UXU^*$is distributed like a GUE too) and the probability density of a matrix in the GUE depends only on the eigenvalues of that matrix (specifically, $\rho(X=H)=C\exp\left(-\mathrm{Tr}(H^2)/2\right)$ (Recall that $\mathrm{Tr}(A^2)$ is the sum of the squares of the eigenvalues). Using these facts one can find an explicit formula for the joint distribution of the eigenvalues of the GUE, namely, :

$$\rho\left((\lambda_1,\ldots,\lambda_N)=(x_1,\ldots,x_n)\right)=C\mathbb{1}_{\{x_1\leq x_2\leq\ldots\leq x_N\}}\left|\Delta(x)\right|^2\prod_{i=1}^{N}\exp\left(-\frac{1}{2}x_i^2\right)$$

Where $\Delta(x)=\prod_{i<j}(x_j-x_i)$ is the Vandermonde determinant. The constant C can be calculated explicitly. [All of this stuff is done in detail in Chapter 2 of AGZ]

We will denote the law of the **unordered eigenvalues** by $\mathscr{P}_N^{(2)}$ (the 2 stands for $\beta=2$ connecting to the case for the GOE and GSE) (i.e. $\mathscr{P}_N^{(2)}$ is the measure on $\mathbb{R}^N$ that has density $C\left|\Delta(x)\right|^2\prod_{i=1}^{N}\exp\left(-\frac{1}{2}x_i^2\right)$...they can come in any order).

For $p\leq N$, we will denote the marginal of $\mathscr{P}_N^{(2)}$ onto p coordinates by $\mathscr{P}_{p,N}$ the **distribution of p unordred eigenvalues of the GUE**. I.e this is the measure on $\mathbb{R}^p$ that tells you the probability of finding p of the N eigenvalues somewhere. More expliecly, $\mathscr{P}_{p,N}$ (I'm going to drop the superscript (2)'s for convencience) one has for any $f\in C_b(\mathbb{R}^p)$ that:

$$\int f(\theta_1,\ldots,\theta_p)\mathrm{d}\mathscr{P}_{p,N}\left(\theta_1,\ldots,\theta_p\right)=\int f(\theta_1,\ldots,\theta_p)\mathrm{d}\mathscr{P}_N(\theta_1,\ldots,\theta_N)$$

(Recall that $\mathscr{P}_N$ is the law of the *unordered* eigenvalues). Clearly one also has:

$$\int f(\theta_1,\ldots,\theta_p)\mathrm{d}\mathscr{P}_{p,N}\left(\theta_1,\ldots,\theta_p\right) = \frac{(N-p)!}{N!}\sum_{\sigma\in S_{p,N}}\int f\left(\theta_{\sigma(1)},\ldots\theta_{\sigma(p)}\right)\mathrm{d}\mathscr{P}_N(\theta_1,\ldots,\theta_N)$$

$$= \frac{p!}{\binom{N}{p}} \sum_{\sigma \in S_{p,N}} \int f\left(\theta_{\sigma(1)}, \ldots \theta_{\sigma(p)}\right) \mathrm{d}\mathscr{P}_N(\theta_1, \ldots, \theta_N) \quad (1)$$

Where $S_{p,N}$ is the set of injective maps from $\{1, \ldots, p\}$ to $\{1, \ldots N\}$. This is just saying that to observe p eigenvalues in a particular configuration, you must observe a subset of the full N eigenvalues in that configuration, and then you have a $p!/\binom{N}{p}$ chance of those being the p eigenvalues you are interested in. (Everything is kind of hairy to write down here because of the whole unordered bit....it is nicest to think of them perhaps as being in order and then being randomly assigned a permutation π to mix up their labels in order to give us the unordered $\mathscr{P}$ measure)

We now introduce Hermite polynomials and the associated normalized (harmonic) oscillator wave-function. [You will recall these the most from your introductory quantum thoery class, they are the polynomials that give the eigenfunctions for the quantum harmonic oscillator, and they obey the "ladder operators" etc]

Definition. (3.2.1.) a) The n-th order **Hermite polynomial** $\mathcal{H}_n(x)$ is defined by:

$$\mathcal{H}_n(x) := (-1)^n \exp\left(\frac{x^2}{2}\right) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \exp\left(-\frac{x^2}{2}\right)$$

(This looks like it might not be a polynomial, but it is!)

b) The **n-th normalized oscillator wave-function** is the function:

$$\psi_n(x) = \frac{\exp\left(-\frac{x^2}{4}\right)}{\sqrt{\sqrt{2\pi} n!}} \mathcal{H}_n(x)$$

Remark. Some authors define the Hermite polynomials differently, with a factor of $\sqrt{2}$ difference in x.

Remark. The Hermite polynomials and wave functions have lots of nice properties. The two that we will use for the moment are:

1) The wave functions are orthogonal and orthonormal:

$$\langle \psi_k . \psi_l \rangle = \int \psi_k(x) \psi_l(x) \mathrm{d}x = \delta_{k,l}$$

2) $\mathcal{H}_n(x)$ is a polynomials of degree n with leading term x^n.

There is a whole subsection later on devoted to these properties and more for the Hermite polynomials. We now have all the background we need to state the determinantal structure for $\mathscr{P}_{p,N}$ that we are strivng towards:

Lemma. *(3.2.2.) For any $p \leq N$, the law $\mathscr{P}_{p,N}$ is absolutly continuous with respect to the Lebesgue measure with density:*

$$\rho_{p,N}\left(\theta_1, \ldots, \theta_p\right) = \frac{(N-p)!}{N!} \det_{k,l=1}^{p} K^{(N)}\left(\theta_k, \theta_l\right)$$

where:

$$K^{(N)}(x,y) = \sum_{k=0}^{N-1} \psi_k(x)\psi_k(y)$$

Remark. This is saying that $\mathscr{P}_N$ is a determinantal projection process onto the space spanned by $\{\psi_1, \psi_2, \ldots, \psi_N\}$in $L^2(\mathbb{R})$.

Proof. By the explicit calculation of the density for $\mathscr{P}_N$, and our discussion about how to take the marginal, we have already an explicity formula for the density namely:

$$\rho_{p,N}(\theta_1, \ldots, \theta_p) = C_{p,N} \int |\Delta(x)|^2 \prod_{i=1}^{p} \exp\left(-\frac{\theta_i^2}{2}\right) \prod_{i=p+1}^{N} \exp\left(-\frac{\zeta_i^2}{2}\right) \prod_{i=p+1}^{N} \mathrm{d}\zeta_i$$

(Here the vector x that appears in the Vandermonde determinant has its first p components equal to θ_i and its last components equal to ζ_i, the variables being integrated out) We now have to manipulate this formula to make it look like a determinant.

Fortunatly a Vandermonde determinant is already a determinant, so we have:

$$\Delta(x) = \prod_{1 \le i < j \le N} (x_j - x_i) = \det_{i,j=1}^{N} x_i^{j-1}$$

Since the Hermite polynomials,$\mathcal{H}_k$, are all monic with leading term x^k, we claim that $\det_{i,j=1}^N x_i^{j-1} = \det_{i,j=1}^N \mathcal{H}_{j-1}(x_i)$. Indeed, the contribution from the non-monic terms can be shown to be zero, since there are many dependent rows/columns in the matrix. [Remark: you might ask yourself why we are using the Hermite polynomials and not any other polynomials here. Any monic polynomials will do at this stage, but the Hermite polynomials play nice with the kernal $\exp\left(-\frac{x^2}{2}\right)$ which appears here and that is still to come]

Let's now consider the case $p = N$ first. We will write $\rho_{N,N}$ as ρ_N for shorthand. We have:

$$\begin{aligned}
\rho_N\left(\theta_1, \ldots, \theta_N\right) &= C_{N,N} \left|\Delta(\theta)\right|^2 \prod_{i=1}^{N} \exp\left(-\frac{\theta_i^2}{2}\right) \\
&= C_{N,N} \left(\det_{i,j=1}^{N} \mathcal{H}_{j-1}(\theta_i)\right)^2 \prod_{i=1}^{N} \exp\left(-\frac{\theta_i^2}{2}\right) \\
&= \tilde{C}_{N,N} \left(\det_{i,j=1}^{N} \psi_{j-1}(\theta_i)\right)^2
\end{aligned}$$

Here we have just used the definition ofψ_k in terms of $\mathcal{H}_k$, and absorbed the constant prefactors into the factor $\tilde{C}_{N,N}$ out front. Now we use the fact that $\det(AB) = \det(A)\det(B)$ and $\det(B^*) = \det(B)$. If we put $A = [\psi_{j-1}(\theta_i)]_{i,j=1}^N$ and $B = \left[\overline{\psi_{i-1}}(\theta_j)\right]_{i,j=1}^N$ so that $A = B^*$, then the matrix AB is exactly the matrix $AB = [K(\theta_i, \theta_j)]_{i,j=1}^N$, so we can replace the determinant squared that appears above by:

$$\rho_N(\theta_1, \ldots, \theta_N) = \tilde{C}_{N,N}\left(\det_{i,j=1}^{N} K^{(N)}(\theta_i, \theta_j)\right)$$

We now introduce a lemma that is handy for this kind of calculation and will be used later on [The proof's not over yet.....I know there is a square on the right of the page over there....just ignore that!) □

Lemma. *(3.2.3.) For any square-integrable $f_1, \ldots, f_n$ and $g_1, \ldots, g_n$ on the real line we have:*

$$\begin{aligned}\frac{1}{n!}\int\int\cdots\int \det_{i,j=1}^{n}\left(\sum_{k=1}^{n} f_k(x_i)g_k(x_j)\right)\prod_{i=1}^{n} dx_i &= \frac{1}{n!}\int\int\cdots\int \det_{i,j=1}^{n} f_i(x_j) \det_{i,.j=1}^{n} g_i(x_j)\prod_{i=1}^{n} dx_i \\ &= \det_{i,j=1}^{n}\int f_i(x)g_j(x)dx\end{aligned}$$

Proof. The first equality is the trick with $\det(AB) = \det(A)\det(B)$ and $\det(B^*) = \det(B)$ and defining A, B with $B^* = A$ that we just did. To get the next inequality, do a permutation expansion for both determinants:

$$\begin{aligned}\int\int\cdots\int \det_{i,j=1}^{n} f_i(x_j) \det_{i,.j=1}^{n} g_i(x_j)\prod_{i=1}^{n} \mathrm{d}x_i &= \sum_{\sigma,\tau\in S_n}\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)\int\cdots\int\prod_{i=1}^{n} f_{\sigma(i)}(x_i)g_{\tau(i)}(x_i)\prod_{i=1}^{n}\mathrm{d}x_i \\ &= \sum_{\sigma,\tau\in S_n}\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)\prod_{i=1}^{n}\int f_{\sigma(i)}(x)g_{\tau(i)}(x)\mathrm{d}x \\ &= n!\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)\prod_{i=1}^{n}\int f_i(x)g_{\sigma(i)}(x)\mathrm{d}x \\ &= n!\det_{i,j=1}^{n}\int f_i(x)g_j(x)\mathrm{d}x\end{aligned}$$

The only tricky step is to regroup the permutations; if you wanted to do this with more steps you could sum over σ first then over τ and then notice that summing over $\tau \in S_n$ is the same as summing over $\tau\sigma^{-1} \in S_n$ or something, and then you coudl "factor out" the σ so that each summand is identical. Ok, anyways, then you do permutation expansion the oppositie direction to get the inequality and thats the lemma. □

Proof. [Contining the proof of Lemma 3.2.2.] Put $f_i = g_i = \psi_{i-1}$ and $n = N$ in Lemma 3.2.3 that we just proved, and we get that:

$$\begin{aligned}\int \det_{i,j=1}^{N} K^{(N)}(\theta_i, \theta_j) \quad \mathrm{d}\theta &= \int \det_{i.j=1}^{N}\sum \psi_k(\theta_i)\psi_k(\theta_j)\mathrm{d}\theta \\ &= N!\det_{i,j=1}^{n}\int \psi_i(x)\psi_j(x)\mathrm{d}x\end{aligned}$$

$$= \quad N! \det(Id)$$
$$= \quad N!$$

Since the ψ's are orthonormal. This shows that $\tilde{C}_{N,N}$, the normalizing constant in $\rho_N(\theta_1, \ldots, \theta_N) = \tilde{C}_{N,N}\left(\det_{i,j=1}^N K^{(N)}(\theta_i, \theta_j)\right)$ is $\tilde{C}_{N,N} = \frac{1}{N!}$. Undoing the constants (they just came from normalizing the ψ_k's) we have $C_{N,N}^{-1} = N! \prod_{k=0}^{N-1}\left(\sqrt{2\pi}k!\right)$. This completes the proof in the case that $p = N$.

To get the result in the case $p \leq N$ you can go the route I suggested in the remark before the lemma (which I think is correct, but am not 100% sure) or you can procede as we did at first in the case $p = N$ to get to:

$$\rho_{p,N}(\theta_1, \ldots, \theta_p) = \tilde{C}_{p,N} \int \left(\det_{i,j=1}^N \psi_{j-1}(x_i)\right)^2 \prod_{i=p+1}^N \mathrm{d}\zeta_i$$

From here you can do a permutation expansion, and many of the terms will die by the orthogonality of the ψ_j's . You then get to a sum for which you can apply the Cauchy-Binet formula and get the result, and then you have to integrate out (and then using the Lemma 3.2.3 as we did above) to get to $\tilde{C}_{p,N} = \frac{(N-p)!}{N!}$. Its important that the polynomials are orthogonal with resepct to the $\exp\left(-\frac{x^2}{2}\right)$ weight here for this to simplify down nicely. This kind of result going from the DPP to the marginal of the DPP is true in general for projection processes like this one. □

Remark. So we needed the polynomials to be monic with leading term x^n so we could put them in as the Vandermonde determinant, and we needed them to be orhtogonal with respect to the weight $\exp(-x^2/2)$ so that the resulting Kernal would be that of a projection determinantal point process and everything simplifies nicely.

Lemma. *(3.2.4.) For any measurable subset* A *of* $\mathbb{R}$*:*

$$\mathbf{P}_N\left(\bigcap_{i=1}^N \{\lambda_i \in A\}\right) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^c}, \ldots, \int_{A^c} \det_{i,j=1}^k K^{(N)}(x_i, x_j) \prod_{i=1}^k dx_i$$

Proof. By our determinatnal formula for the density (Lemma 3.2.3) and our $\int \det$ to $\det \int$ lemma (Lemma 3.2.3) and orthogonality we have that:

$$\begin{aligned}
\mathbf{P}_N\left(\bigcap_{i=1}^N \{\lambda_i \in A\}\right) &= \int_A \cdots \int_A \frac{1}{N!}\left(\det_{i,j=1}^N K^{(N)}(\theta_i, \theta_j)\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n \\
&= \det_{i,j=1}^N \int_A \psi_i(x)\psi_j(x)\mathrm{d}x \\
&= \det_{i,j=1}^N \left(\delta_{ij} - \int_{A^c} \psi_i(x)\psi_j(x)\mathrm{d}x\right)
\end{aligned}$$

From here, we invoke the expansion for determinants of the form $= \det\left[I + hA_{ij}\right] = 1 + \sum_{k=1}^{N} h^k \sum_{0\le v_1<\ldots<v_k\le N-1} \det_{i,j=1}^{k}\left(A_{v_i v_j}\right)$ when $h = -1$. (See the proposition after the lemma). So we get too:

$$\mathbf{P}_N\left(\bigcap_{i=1}^{N}\{\lambda_i \in A\}\right) = 1 + \sum_{k=1}^{N}(-1)^k \sum_{0\le v_1<\ldots<v_k\le N-1} \det_{i,j=1}^{k}\left(\int_{A^c} \psi_{v_i}(x)\psi_{v_j}(x)\right)\prod_{i=1}^{k} \mathrm{d}x_i$$

Using the det $\int$ to $\int$det lemma the other way now, and the standard $\det(AB) = \det(A)\det(B)$ trick we get:

$$\mathbf{P}_N\left(\bigcap_{i=1}^{N}\{\lambda_i \in A\}\right) = 1 + \sum_{k=1}^{N}\frac{(-1)^k}{k!}\int_{A^c}\int_{A^c}\cdots\int_{A^c}\sum_{0\le v_1<\ldots<v_k\le N-1}\left(\det_{i,j=1}^{k}\psi_{v_i}(x_j)\right)^2\prod_{i=1}^{k}\mathrm{d}x_i$$

By the Cauchy Binet theorem now, [In our case we use the following "version" of the theorem: Let A be a $p \times N$ matrix and let $C = AA^{\star}$(this is a $p \times p$ matrix), then $\det C = \sum_{K\in\mathcal{K}_{p,N}} \det A_K \det A_K^{\star}$where $\mathcal{K}_{p,N}$ is the set of all p element subsets of $\{1, \ldots, N\}$ and A_K is the $p \times p$ matrix which is obtained from A by keeping only the columns in K. This is exactly the set up we have here with $A_{i,j} = \psi_{j-1}(\theta_i)$.] Have:

$$\mathbf{P}_N\left(\bigcap_{i=1}^{N}\{\lambda_i \in A\}\right) = 1 + \sum_{k=1}^{N}\frac{(-1)^k}{k!}\int_{A^c}\int_{A^c}\cdots\int_{A^c}\det_{i,j=1}^{k} K^{(N)}(x_i, x_j)\prod_{i=1}^{k}\mathrm{d}x_i$$

Since the summand is 0 for $k \ge N$ (since the matrix $\left[K^{(N)}(x_i x_j)\right]_{i,j=1}^{k}$ is a $k \times k$ matrix here and it arises as the product of a $k \times N$ matrix with a $N \times k$ matrix (namely $[\psi_i(x_j)]_{i=1..N,j=1,,k}$ we see that the matrix cannot have rank more than N, so we might as well write it as an infinite sum:

$$\mathbf{P}_N\left(\bigcap_{i=1}^{N}\{\lambda_i \in A\}\right) = 1 + \sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_{A^c}\int_{A^c}\cdots\int_{A^c}\det_{i,j=1}^{k} K^{(N)}(x_i, x_j)\prod_{i=1}^{k}\mathrm{d}x_i$$

□

Remark. See also the notes from Johannson on "Determinantal Point Processes and Random Matrices" for a slightly more general look at this.

Proposition. *Let A_{ij} be an $N \times N$ matrix. Have:*

$$\begin{aligned}\det\left[I + hA_{ij}\right] &= 1 + h\sum_{i} A_{ii} + \frac{h^2}{2}\sum_{i,j}\det\begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj}\end{bmatrix} + \ldots \\ &= 1 + \sum_{k=1}^{N} h^k \sum_{0\le v_1<\ldots<v_k\le N-1}\det_{i,j=1}^{k}\left(A_{v_i v_j}\right)\end{aligned}$$

Proof. Let us call $D(h) := \det[I + hA_{ij}]$. $D(h)$ is a polynomial of degree at most N in the variable h, $D(h) = \sum_0^N a_m h^m$, so it suffices to find the coefficients a_m. Using derivatives, we have that:

$$a_m = \frac{1}{m!}\left(\frac{\mathrm{d}}{\mathrm{d}h}\right)^m D(h)\bigg|_{h=0}$$

Label the columns of $I + hA_{ij}$ as $C_j(h)$. Notice that each column $C_j(h)$ has components which are linear in h and also that $C_j(0) = e_j$. Now, think of the determinant as being a linear function of all the columns. Since the derivative is multilinear as a function of the columns C_j, we have the following differentiation rule:

$$\frac{\mathrm{d}}{\mathrm{d}h}\det[C_1(h), C_2(h), \ldots, C_N(h)] = \sum_{k=1}^{N}\det\left[C_1(h), \ldots, \frac{\mathrm{d}}{\mathrm{d}h}C_k(h), \ldots, C_N(h)\right]$$

(Here a derivative $\frac{\mathrm{d}}{\mathrm{d}h}C_i$ is the vector that we get by taking component-wise derivatives. A skeptical reader could prove this result from the definition of the derivative and using induction, along with the usual "add and subtract" trick that comes up in this type of derivative argument. Hint: Induction hypothesis for $l \leq N$ is that: $\lim_{\Delta\to 0}\frac{1}{\Delta}\left(\det[C_1(h+\Delta), , \ldots, C_l(h+\Delta), C_{l+1}(h), \ldots, C_N(h)] - \det[C_1(h), \ldots, C_N(h)]\right)$ $\sum_{k=1}^{l}\det\left[C_1(h), \ldots, \frac{\mathrm{d}}{\mathrm{d}h}C_k(h), \ldots, C_N(h)\right])$ □

Using this rule repeatedly, gives that:

$$\left(\frac{\mathrm{d}}{\mathrm{d}h}\right)^m \det[C_1(h), C_2(h), \ldots, C_N(h)] = \sum_{k_1,\ldots,k_m=1}^{N}\det\left[C_1(h), \ldots, \frac{\mathrm{d}}{\mathrm{d}h}C_{k_1}(h), \ldots \frac{\mathrm{d}}{\mathrm{d}h}C_{k_m}, \ldots C_N(h)\right]$$

In our case, since every component $C_i(h)$ is a linear function of h, taking two derivatives of any column C_k would give a zero-column, and then the determinant from that term would vanish leaving no contribution. For this reason, we only need to consider k_i all distinct. In our effort to evaluate a_m now, we evalute the above at $h = 0$. For columns with no derivative we have $C_k(0) = e_k$, and for columns with a derivative we have that $\frac{\mathrm{d}}{\mathrm{d}h}C_k(0) = A_{\cdot k}$ is the column from A. Hence we have:

$$\begin{aligned} m!a_m &= \sum_{k_1\ldots k_m}\det[e_1, \ldots, A_{\cdot k_1}, \ldots A_{\cdot k_m}, \ldots, e_N] \\ &= \sum_{k_1\ldots k_m}\det\left[A_{k_i k_j}\right]_{i,j=1}^{m} \end{aligned}$$

If we sort the indeces k_i so that $k_1 < \ldots < k_m$ then we pick up a factor of $m!$ which exactly cancels out the $m!$ on the LHS. Plugging back into $D(h) = \sum h^k a_k$ completes the result.

2.2 Properties of the Hermite Polynomials and Oscillator Wave-Functions

Definition. We define the inner produc $\langle f, g\rangle_{\mathscr{G}} = \int_{\mathbb{R}} f(x)g(x)\exp\left(-x^2/2\right) \mathrm{d}x$

Lemma. *(3.2.5.) Properties of the Hermite polynomials*

1. $\mathcal{H}_0(x) = 1, \mathcal{H}_1(x) = x$ and $\mathcal{H}_{n+1}(x) = x\mathcal{H}_n(x) - \mathcal{H}'_n(x)$
2. $\mathcal{H}_n(x)$ is a *monic* polynomial of degree n
3. $\mathcal{H}_n(x)$ is even when n is even and odd when n is odd.
4. $\left\langle x, \mathcal{H}_n^2\right\rangle_{\mathscr{G}} = 0$
5. $\langle \mathcal{H}_k, \mathcal{H}_l\rangle_{\mathscr{G}} = \sqrt{2\pi}k!\delta_{kl}$
6. $\langle f, \mathcal{H}_n\rangle_{\mathscr{G}} = 0$ for all polynomials $f(x)$ of degree $< n$
7. $x\mathcal{H}_n(x) = \mathcal{H}_{n+1}(x) + n\mathcal{H}_{n-1}(x)$ for $n \geq 1$
8. $\mathcal{H}'_n(x) = n\mathcal{H}_{n-1}(x)$ [Corr: $\mathcal{H}_n^{(k)} = n(n-1)\dots(n-k+1)\mathcal{H}_{n-k}(x)$]
9. $\mathcal{H}''_n(x) - x\mathcal{H}'_n(x) + n\mathcal{H}_n(x) = 0$
10. For $x \neq y$: $\sum_{k=0}^{n-1} \frac{\mathcal{H}_k(x)\mathcal{H}_k(y)}{k!} = \frac{(\mathcal{H}_n(x)\mathcal{H}_{n-1}(y) - \mathcal{H}_{n-1}(x)\mathcal{H}_n(y))}{(n-1)!(x-y)}$

Proposition. *$\mathcal{H}_n(x+t)$ can be written in terms of $\{\mathcal{H}_k(x)\}$ by the following combinatorial formula:*

$$\mathcal{H}_n(x+t) = \sum_{k=0}^{n} \binom{n}{k} \mathcal{H}_k(x) t^{n-k}$$

Proof. These are polynomials, so they are anayltic. We use $\mathcal{H}_n^{(k)} = n(n-1)\dots(n-k+1)\mathcal{H}_{n-k}(x)$ to explictly compute the derivatives. By Taylor's theorem:

$$\begin{aligned}
\mathcal{H}_n(x+t) &= \sum_{k=0}^{N} \frac{t^k}{k!}\mathcal{H}_n^{(k)}(x) \\
&= \sum_{k=0}^{N} \frac{t^k}{k!} n(n-1)\dots(n-k+1)\mathcal{H}_{n-k}(x) \\
&= \sum_{k=0}^{N} t^k \binom{n}{k} \mathcal{H}_{n-k}(x) \\
&= \sum_{k=0}^{N} t^{n-k} \binom{n}{n-k} \mathcal{H}_k(x) \\
&= \sum_{k=0}^{N} t^{n-k} \binom{n}{k} \mathcal{H}_k(x)
\end{aligned}$$

As desired. $\square$

Definition. Lets have the definition of the oscillator wave functions here again:

$$\psi_n(x) = \frac{\exp\left(-\frac{x^2}{4}\right)}{\sqrt{\sqrt{2\pi}n!}}\mathcal{H}_n(x)$$

Notice that since the $\mathcal{H}_n$'s are polynomials, the osciallator wave functions go to 0 as $x \to \pm\infty$ like a polynomial times $\exp\left(-\frac{x^2}{2}\right)$. The same will hold for any of its derivatives too (it will turn out that we can write ψ_n' in terms of ψ_n andψ_{n-1})

Lemma. *(3.2.7.) Properties of the Oscillator wave function (these just follow from the properties of the Hermite polynomials and the definition of the oscillator wave function in terms of those)*

1. $\int \psi_k \psi_l = \delta_{kl}$
2. $x\psi_n = \sqrt{n+1}\psi_{n+1} + \sqrt{n}\psi_{n-1}$
3. $K^{(n)}(x,y) = \sum_{k=0}^{n-1} \psi_k(x)\psi_k(y) = \sqrt{n}\left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right)/(x-y)$
4. $\psi_n'(x) = -\frac{x}{2}\psi_n(x) + \sqrt{n}\psi_{n-1}(x)$
5. $\psi_n'' + \left(n + \frac{1}{2} - \frac{x^2}{2}\right)\psi_n(x) = 0$

3 The semicircle law revisited

Recall that we proved the semicircle law for *Wigner* matrices by using combinatorial methods to compute the limits as $N \to \infty$ of the moments. In this section we will compute *explicitly* as a function of N the moments for the GUE (which is a special type of Wigner matrix....it is a Gaussian Wigner matrix)

Let $X \in \mathscr{H}_N$ be a random Hermitian GUE matrix with eigenvalues $\lambda_1^N \leq \ldots \leq \lambda_N^N$ and let:

$$L_N = \left(\delta_{\lambda_1^N/\sqrt{N}} + \ldots + \delta_{\lambda_N^N/\sqrt{N}}\right)/N$$

Be the scaled empirical distribution functions for the matrix $X_N/\sqrt{N}$. Let $\bar{L}_N$ be the average empirical spectral distribution, that is $\langle \bar{L}_N, f\rangle = \mathbf{E}\langle L_N, f\rangle$. We will derive a recursion for the moments of $\bar{L}_N$ and estimate the order of fluctuation of the renormalized maximum eigenvalue $\lambda_N^N/\sqrt{N}$ above the spectrum edge.

3.1 Calculation of moments of $\bar{L}_N$

Lemma. *(3.3.1) For any $s \in \mathbb{R}$ and any $N \in \mathbb{N}$:*

$$\begin{aligned} \left\langle \bar{L}_N, e^{s\cdot} \right\rangle &= e^{s^2/2N} \sum_{k=0}^{N-1} \frac{1}{k+1} \binom{2k}{k} \frac{(N-1)\dots(N-k)}{N^k} \frac{s^{2k}}{(2k)!} \\ &= e^{s^2/2N} \sum_{k=0}^{N-1} C_k \frac{(N-1)\dots(N-k)}{N^k} \frac{s^{2k}}{(2k)!} \end{aligned}$$

(I.e. we have an explicit formula for the moment generating function of $\overline{L}_N$ involving Catalan numbers)

Proof. By the determinantal formula for the spectrum of the GUE, we have the one -point correlation formula that $\rho_{1,N}(x) = \frac{1}{N} K^{(N)}(x,x)$, so then we have that for any ϕ:

$$\begin{aligned} \left\langle \bar{L}_N, \phi \right\rangle &= \mathbf{E}\left(\langle L_N, \phi\rangle\right) \\ &= \mathbf{E}\left(\frac{1}{N}\phi\left(\lambda_1^N/\sqrt{N}\right) + \dots + \frac{1}{N}\phi\left(\lambda_N^N/\sqrt{N}\right) \right) \\ &= N\mathbf{E}\left(\frac{1}{N}\phi\left(\lambda^N/\sqrt{N}\right)\right) \text{ (take them unoredered)} \\ &= N\frac{1}{N}\int_{-\infty}^{\infty} \phi\left(\frac{x}{\sqrt{N}}\right)\frac{1}{N}K^{(N)}(x,x)\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \phi(x) \frac{K^{(N)}\left(\sqrt{N}x, \sqrt{N}x\right)}{\sqrt{N}} \mathrm{d}x \text{ (change of variable)} \end{aligned}$$

This shows that $\bar{L}_N$ is absolutely continuous with respect to the Lebesgue measure with density $\frac{K^{(N)}\left(\sqrt{N}x,\sqrt{N}x\right)}{\sqrt{N}}$. For this reason we are interested in $K(x,x)$. Fortunatly, this has a simple for in terms of the oscillator wave functions. We will compute $\frac{\mathrm{d}}{\mathrm{d}x}K(x,x)$ and then integrate by parts to the formula for the moment generatring function we want. Now, by the identity $\sum_{k=0}^{n-1} \psi_k(x)\psi_k(y) = \sqrt{n}\left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right)/(x-y)$ we have that:

$$\frac{K^{(n)}(x,y)}{\sqrt{n}} = \frac{\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)}{x-y}$$

and hence by L'Hopital have:

$$\frac{K^{(n)}(x,x)}{\sqrt{n}} = \psi_n'(x)\psi_{n-1}(x) - \psi_{n-1}'(x)\psi_n(x)$$

(One observation y ou can make at this point is that $K^{(n)}(x,x) \to 0$ like $\exp\left(-x^2/2\right)$ as $x \to \pm\infty$, just as the oscillator functions do) Therefore, using

$\psi_n'' + \left(n + \frac{1}{2} - \frac{x^2}{2}\right)\psi_n(x) = 0$, we get to:

$$\begin{aligned}\frac{\frac{\mathrm{d}}{\mathrm{d}x}K^{(n)}(x,x)}{\sqrt{n}} &= \psi_n''(x)\psi_{n-1}(x) - \psi_{n-1}''(x)\psi_n(x) \\ &= \ldots \\ &= -\psi_n(x)\psi_{n-1}(x)\end{aligned}$$

We can use this, and the fact that $\frac{K^{(N)}(\sqrt{N}x,\sqrt{N}x)}{\sqrt{N}}$ is our density function to calculate the moment generating function by integration by parts:

$$\begin{aligned}\left\langle \bar{L}_N, \exp(s\cdot)\right\rangle &= \int_{-\infty}^{\infty} \exp(sx)\frac{K^{(N)}\left(\sqrt{N}x,\sqrt{N}x\right)}{\sqrt{N}}\mathrm{d}x \\ &= \frac{1}{N}\int_{-\infty}^{\infty}\exp\left(\frac{sx}{\sqrt{N}}\right)K^{(N)}(x,x)\mathrm{d}x \\ &= \frac{1}{N}\left[\frac{\sqrt{N}}{s}\exp\left(\frac{sx}{\sqrt{N}}\right)K^{(N)}(x,x)\right]_{-\infty}^{\infty} + \frac{1}{N}\int_{-\infty}^{\infty}\frac{\sqrt{N}}{s}\exp\left(\frac{sx}{\sqrt{N}}\right)\frac{\mathrm{d}}{\mathrm{d}x}K^{(N)}(x,x)\mathrm{d}x\end{aligned}$$

The first term vanishes because we know that $K^{(N)}(x,x) \to 0$ like a polynomial timesexp$\left(\frac{-x^2}{2}\right)$. The second term we know since we just calculated $\frac{\mathrm{d}}{\mathrm{d}x}K^{(n)}(x,x)$:

$$\begin{aligned}\left\langle \bar{L}_N, \exp(s\cdot)\right\rangle &= \frac{1}{s}\int_{-\infty}^{\infty}\exp\left(\frac{sx}{\sqrt{N}}\right)\psi_N(x)\psi_{N-1}(x)\mathrm{d}x \\ &= \frac{1}{s\sqrt{2\pi}n!}\int_{-\infty}^{\infty}\exp\left(\frac{sx}{\sqrt{N}}\right)\mathcal{H}_N(x)\mathcal{H}_{N-1}(x)\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x\end{aligned}$$

This is a nice formula! We will now work a bit to get it more explitcitly. If it wasn't for the $\exp\left(\frac{sx}{\sqrt{N}}\right)$ we would be in buisness here since the Hermite polynomials are orthogonal w.r.t. the weight $\exp\left(-\frac{x^2}{2}\right)$. To deal with that term, we can do a change of variable and complete the square (this reminds me of caluclating the moment generating function for a Gaussian). The cost of this is that we end up with terms like $\mathcal{H}_n(x+t)$. Fortunatly we can handle this by the identity $\mathcal{H}_n(x+t) = \sum_{k=0}^{N} t^{n-k}\binom{n}{k}\mathcal{H}_k(x)$ which we got by doing a Taylor expansion and using the fact that derivatives of $\mathcal{H}_n$ are other hermite poly's. (AGZ do this identitiy here, but I bundled it up in t "properties of Hermite polynomials" section). For convenience, let

$$S_t^n : = \int_{-\infty}^{\infty}\exp(tx)\psi_n(x)\psi_{n-1}(x)\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}n!}\int_{-\infty}^{\infty} \exp(tx)\,\mathcal{H}_n(x)\mathcal{H}_{n-1}(x)\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x$$

(So that $\langle \bar{L}_N, \exp(s\cdot)\rangle = \frac{1}{s}S^N_{s/\sqrt{N}}$ is what we are actually interested in). We compute now:

$$\begin{aligned}
S_t^n &= \frac{1}{\sqrt{2\pi}n!}\int_{-\infty}^{\infty} \mathcal{H}_{n-1}(x)\mathcal{H}_{n-1}(x)\exp\left(-\frac{x^2}{2}+tx\right)\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}n!}\exp\left(\frac{t^2}{2}\right)\int_{-\infty}^{\infty} \mathcal{H}_{n-1}(x)\mathcal{H}_{n-1}(x)\exp\left(-\frac{(x-t)^2}{2}\right)\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}n!}\exp\left(\frac{t^2}{2}\right)\int_{-\infty}^{\infty} \mathcal{H}_{n-1}(y+t)\mathcal{H}_{n-1}(y+t)\exp\left(-\frac{y^2}{2}\right)\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}n!}\exp\left(\frac{t^2}{2}\right)\int_{-\infty}^{\infty} \sum_{i=1}^{n}\sum_{j=1}^{n-1}\binom{n}{i}\binom{n-1}{j}\mathcal{H}_i(x)\mathcal{H}_j(x)t^{n-i}t^{n-1-j}\exp\left(-\frac{y^2}{2}\right)\mathrm{d}y
\end{aligned}$$

Now by the orthogonality relation $\langle \mathcal{H}_k, \mathcal{H}_l\rangle_{\mathscr{G}} = \sqrt{2\pi}k!\delta_{kl}$ the terms where $i \neq j$ vanish and we remain with:

$$S_t^n = \exp\left(\frac{t^2}{2}\right)\sum_{k=0}^{n-1}\frac{k!}{n!}\binom{n}{k}\binom{n-1}{k}t^{2n-1-2k}$$

From here, some manipulation with binomial coefficients in needed to get to the form we had originally. □

Lemma. *(2.1.6. For Gaussian Wigner Matrices) For every k, the k-th moment of the mean empirical distribution function converges to the k-th moment of the semicircle law, namely $m_{2k} = C_k$ and $m_{2k+1} = 0$. I.e:*

$$\begin{aligned}
\langle \bar{L}_N, x^{2k}\rangle &\to C_k \text{ as } N \to \infty \\
\langle \bar{L}_N, x^{2k+1}\rangle &\to 0 \text{ as } N \to \infty
\end{aligned}$$

Proof. We have the explicit formula for the generating function:

$$\langle \bar{L}_N, e^{s\cdot}\rangle = e^{s^2/2N}\sum_{k=0}^{N-1} C_k\frac{(N-1)\dots(N-k)}{N^k}\frac{s^{2k}}{(2k)!}$$

From the RHS we see that the moment generating function is infinetly differentiable, so we have moments of all orders. Since everything is nice here, we can Taylor expand both sides in s, and we extract the $2k-$th coeffcient, we have:

$$\left\langle \bar{L}_N, x^{2k} \right\rangle \frac{1}{(2k)!} = C_k \frac{(N-1)\dots(N-k)}{N^k} \frac{1}{(2k)!} + \text{finitely many terms coming from the expansion of } e^{s^2}$$
$$\to C_k \frac{1}{(2k!)} + 0 \text{ as } N \to \infty$$

The same kind of dealio works for k odd. □

3.2 The Harer-Zagier recursion and Ledoux's Argument

In this section λ_N^N denotes the maximal eigenvalue of the GUE. Our goal in this section is to prove that:

Lemma. *(3.3.2) (Ledoux's Bound) There exist positive constants c' and C' such that:*

$$\mathbf{P}\left(\frac{\lambda_N^N}{2\sqrt{N}} \geq \exp\left(\epsilon N^{-2/3}\right)\right) \leq C' \exp\left(-c'\epsilon\right)$$

Remark. Rouglhy speaking this says that the fluctuations of the top rescaled top eignevalue $\tilde{\lambda}_N^N = \lambda_N^N/2\sqrt{N} - 1$ above 0 are of order of magnitude $N^{-2/3}$. This is an a-priori indication that we might have convergence in distribution as in the Tracy Widom law. In fact, this lemma is part of what makes the theory work.

By using the method of moments, you can prove this kind of thing for a general Wigner matrix. In this section we re only looking at the GUE though.

Definition. Let $b_k^{(N)}$ be such that:

$$\left\langle \bar{L}_N, e^{s\cdot} \right\rangle = \sum_{k=0}^{\infty} b_k^{(N)} C_k \frac{s^{2k}}{(2k)!}$$

i.e. $\left\langle \bar{L}_N, x^{2k} \right\rangle = \frac{b_k^{(N)}}{k+1}\binom{2k}{k} = b_k^{(N)} C_k$. Compare this to the formula $e^{s^2/2N} \sum_{k=0}^{N-1} C_k \frac{(N-1)\dots(N-k)}{N^k} \frac{s^{2k}}{(2k)!}$ we proved in the last section. We know for example that $b_k^{(N)} \to 1$ for all k as $N \to \infty$.

Lemma. *(3.3.3.) (Harer-Zagier Recursions) For any integer numbers k and N we have:*

$$b_{k+1}^{(N)} = b_k^{(N)} + \frac{k(k+1)}{4N^2} b_{k-1}^{(N)}$$

Proof. This proof basically comes from reconcling the two formulas, $e^{s^2/2N} \sum_{k=0}^{N-1} C_k \frac{(N-1)\dots(N-k)}{N^k} \frac{s^{2k}}{(2k)!}$ and $\sum_{k=0}^{\infty} b_k^{(N)} C_k \frac{s^{2k}}{(2k)!}$. You can see the recurrsion in question by seeing what differential equation the generating function of the coefficienst satisfy. □

Lemma. *(3.3.2) (Ledoux's Bound) There exist positive constants c' and C' such that:*

$$\mathbf{P}\left(\frac{\lambda_N^N}{2\sqrt{N}} \geq \exp\left(\epsilon N^{-2/3}\right)\right) \leq C' \exp\left(-c'\epsilon\right)$$

Proof. [Idea: The Harer-Zagier recursion gives us a recursion for the moments of $\bar{L}_n$ which controls very tightly how they can grow. This upper bounds the moments of the eigenvalues of course because for any positive function ϕ:

$$\begin{aligned}\langle \bar{L}_n, \phi \rangle &= \frac{1}{N}\sum_{j=1}^{N} \mathbf{E}\left[\phi\left(\frac{\lambda_j^N}{\sqrt{N}}\right)\right] \\ &\geq \frac{1}{N}\mathbf{E}\left[\phi\left(\frac{\lambda_N^N}{\sqrt{N}}\right)\right]\end{aligned}$$

With $\phi(x) = x^{2k}$, this is:

$$b_k^{(N)} C_k = \langle \bar{L}_n, x^{2k} \rangle \geq \frac{1}{N^{k+1}} \mathbf{E}\left[\left(\lambda_N^N\right)^{2k}\right]$$

we get Cheb type moment bound to control the probability.]

The H-Z recursion gives us immediatly the inequalites that:

$$0 \leq b_k^{(N)} \leq b_{k+1}^{(N)} = b_k^{(N)} + \frac{k(k+1)}{4N^2} b_{k-1}^{(N)} \leq \left(1 + \frac{k(k+1)}{4N^2}\right) b_k^{(N)}$$

Hence:

$$\begin{aligned}b_k^{(N)} &\leq \prod_{i=1}^{k}\left(1 + \frac{i(i+1)}{4N^2}\right) b_1^{(N)} \\ &= \exp\left(\sum_{i=1}^{k} \log\left(1 + \frac{i(i+1)}{4N^2}\right) + \log\left(b_1^{(N)}\right)\right) \\ &\leq \exp\left(\sum_{i=1}^{k}\left(\frac{i(i+1)}{4N^2}\right) + \log\left(b_1^{(N)}\right)\right) \\ &\leq \exp\left(c\frac{k^3}{N^2}\right)\end{aligned}$$

For some constant $c > 0$. We now just estimate $\mathbf{P}\left(\frac{\lambda_N^N}{2\sqrt{N}} \geq \exp\left(\epsilon N^{-2/3}\right)\right)$ with a Cheb moment bound:

$$\begin{aligned}\mathbf{P}\left(\frac{\lambda_N^N}{2\sqrt{N}} \geq \exp\left(\epsilon N^{-2/3}\right)\right) &\leq \left(\frac{1}{2\sqrt{N}\exp\left(\epsilon N^{-2/3}\right)}\right)^{2k} \mathbf{E}\left[\left(\lambda_N^N\right)^{2k}\right] \\ &= \left(\frac{1}{2\sqrt{N}\exp\left(\epsilon N^{-2/3}\right)}\right)^{2k} b_k^{(N)} C_k N^{k+1} \\ &= \exp\left(-2\epsilon k N^{-2/3}\right) \frac{N}{4^k} C_k b_k^{(N)}\end{aligned}$$

By stirling's approximation, we know that the Catalan numbers grow like $C_k \sim \frac{4^k}{k^{3/2}\sqrt{\pi}}$. So then we have some constant C for which $C_k \leq C\frac{4^k}{k^{3/2}\sqrt{\pi}}$ for all

k, and putting this in as well as $b_k^{(N)} \leq \exp\left(c\frac{k^3}{N^2}\right)$ we get:

$$\begin{aligned}\mathbf{P}\left(\frac{\lambda_N^N}{2\sqrt{N}} \geq \exp\left(\epsilon N^{-2/3}\right)\right) &\leq \exp\left(-2\epsilon k N^{-2/3}\right)\frac{CN}{k^{3/2}}\exp\left(c\frac{k^3}{N^2}\right) \\ &= CNk^{-3/2}\exp\left(-2\epsilon N^{-2/3}k + ck^3/N^2\right)\end{aligned}$$

This bound holds for any k. Taking $k \simeq N^{2/3}$ both terms in the exponential are like $N^{-2/3}k \simeq ck^3/N^2 \simeq O(1)$ the coeffient out front is also $Nk^{-3/2} \simeq O(1)$, so this gives us exactly the estimate we want. □

3.3 Working with the n-point correlation function (Exercise 3.3.4.)

So far, we have calculated the determinantal form for mean empircal spectral distribution and then used that to get the formula for the moments then the HZ bound and then Ledoux's argument. These all actually came from the one point correlation function, namley $\rho_{1,N}(x) = \frac{1}{N}K(x,x)$. In this section we will use the full N-point correlation function to show that:

$$\left\langle L_N, x^k\right\rangle - \left\langle \bar{L}_N, x^k\right\rangle \to 0 \text{ in probability}$$

(Recall: $\left\langle L_N, x^k\right\rangle$ is a random variable, $\left\langle \bar{L}_N, x^k\right\rangle$ is a number). $\left\langle L_N, x^k\right\rangle$ is given by some crazy integral over the full n point correlation function.)

We will actually first show that $\lim_{N\to\infty}\left(\mathbf{E}\left\langle L_n, x^k\right\rangle^2 - \left\langle \bar{L}_N, x^k\right\rangle^2\right) = 0$, and then use this to deduce the result. The fact that $\left\langle L_N, x^k\right\rangle - \left\langle \bar{L}_N, x^k\right\rangle \to 0$ in probability and the fact that $\left\langle \bar{L}_N, x^{2k}\right\rangle \to C_k$ where the two facts we used in chapter 2, along with a "moment method" argument to prove Wigner's theorem, so if you like we have proven Wigner's theorem for the GUE using the determinantal form.

Lemma. $\int K^{(n)}(x,t)K^{(n)}(t,y)\,dt = K^{(n)}(x,y)$

Proof. This just comes from the definition $K^{(n)}(x,y) = \sum_{i=1}^n \psi_i(x)\psi_i(y)$ and the fact the $\psi's$ are orthonormal. If you expand out:

$$K^{(n)}(x,t)K^{(n)}(t,y) = \left(\sum_{i=1}^n \psi_i(x)\psi_i(t)\right)\left(\sum_{j=1}^n \psi_j(t)\psi_j(y)\right)$$

Any terms where $i \neq j$ will die when integrated out, and any terms $i = j$ reduce to $\psi_i(x)\psi_i(y)$. □

Lemma. $\lim_{N\to\infty}\left(\mathbf{E}\left\langle L_n, x^k\right\rangle^2 - \left\langle \bar{L}_N, x^k\right\rangle^2\right) = 0$

Proof. From the 1-point correlation formula, we have $\left\langle \bar{L}_N, x^k \right\rangle = \frac{1}{N} \int \left(\frac{x}{\sqrt{N}} \right)^k K^{(N)}(x,x)\mathrm{d}x = \frac{1}{N^{k/2+1}} \int x^k K^{(N)}(x,x)\mathrm{d}x$. Hence:

$$\begin{aligned} \left\langle \bar{L}_N, x^k \right\rangle^2 &= \frac{1}{N^{k+2}} \left(\int x^k K^{(N)}(x,x)\mathrm{d}x \right)^2 \\ &= \frac{1}{N^{k+2}} \int \int x^k y^k K^{(N)}(x,x) K^{(N)}(y,y)\mathrm{d}x\mathrm{d}y \end{aligned}$$

On the other hand, consider that:

$$\begin{aligned} \langle L_N, x_k \rangle^2 &= \left(\frac{1}{N} \left(\frac{\lambda_1^N}{\sqrt{N}} \right)^k + \ldots + \frac{1}{N} \left(\frac{\lambda_N^N}{\sqrt{N}} \right)^k \right)^2 \\ &= \frac{1}{N^{k+2}} \left(\sum_{i=1}^{N} \left(\lambda_i^N \right)^k \right)^2 \\ &= \frac{1}{N^{k+2}} \left[\sum_{i=1}^{N} \left(\lambda_i^N \right)^{2k} + 2 \sum_{1 \le i < j \le N} \left(\lambda_i^N \right)^k \left(\lambda_j^N \right)^k \right] \end{aligned}$$

So we can evaluate the $\mathbf{E}$ of this using the one and two point correlation functions:

$$\mathbf{E}\left[\left\langle L_N, x^k \right\rangle^2 \right] = \frac{1}{N^{k+2}} \frac{1}{N} \int x^{2k} K^{(N)}(x,x)\mathrm{d}x + \frac{1}{N^{k+2}} \frac{2}{N(N-1)} \sum_{1 \le i \le j \le N} \int \int x^k y^k \det \left[\begin{array}{c} K^{(N)}(x,x) \\ K^{(N)}(x,y) \end{array} \right.$$

Since all the terms in the last sum are the same, we can simply multiply by the number of terms. As fortune has it, this cancels out the coefficient of $\frac{2}{N(N-1)}$ in the front! Remain with:

$$\begin{aligned} \mathbf{E}\left[\left\langle L_N, x^k \right\rangle^2 \right] &= \frac{1}{N^{k+2}} \frac{1}{N} \int x^{2k} K^{(N)}(x,x)\mathrm{d}x + \frac{1}{N^{k+2}} \int \int x^k y^k \det \left[\begin{array}{cc} K^{(N)}(x,x) & K^{(N)}(x,y) \\ K^{(N)}(x,y) & K^{(N)}(y,y) \end{array} \right] \mathrm{d}x\mathrm{d}y \\ &= \frac{1}{N^{k+2}} \frac{1}{N} \int x^{2k} K^{(N)}(x,x)\mathrm{d}x \\ &\quad + \frac{1}{N^{k+2}} \frac{1}{N(N-1)} \int \int x^k y^k K^{(N)}(x,x) K^{(N)}(y,y)\mathrm{d}x\mathrm{d}y \\ &\quad - \frac{1}{N^{k+2}} \frac{1}{N(N-1)} \int \int x^k y^k K^{(N)}(x,y) K^{(N)}(y,x)\mathrm{d}x\mathrm{d}y \end{aligned}$$

The middle term is precisely the expression we had for $\left\langle \bar{L}_N, x^k \right\rangle^2$! We can make the first term look a bit more like the last term by writing $K^{(N)}(x,x) = \int K^{(N)}(x,y) K^{(N)}(y,x)\mathrm{d}x$. We remain with:

$$\mathbf{E}\left[\left\langle L_N, x^k \right\rangle^2 \right] - \left\langle \bar{L}_N, x^k \right\rangle^2 = \frac{1}{N^{k+2}} \int \int \left(x^{2k} - x^k y^k \right) K^{(N)}(x,y) K^{(N)}(y,x)\mathrm{d}x\mathrm{d}y$$

So it suffices to check that this thing on the RHS goes to zero. Lets give a name, $I_k^{(N)} := RHS$ of the above. Here's the strategy to deal with this guy. We need some identies, firstly:

$$K^{(n)}(x,y) = \sum_{k=0}^{n-1} \psi_k(x)\psi_k(y) = \sqrt{n}\left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right)/(x-y)$$

Reduces us to looking at the integral:

$$\frac{1}{N^{k+3/2}} \int\int \frac{x^{2k} - x^k y^k}{x-y} \left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right) K^{(N)}(y,x)\mathrm{d}x\mathrm{d}y$$

(The lower case n's should be upper case from here on in..ooops!) Now we will use the three series recursion $x\psi_n = \sqrt{n+1}\psi_{n+1} + \sqrt{n}\psi_{n-1}$ to try and simplify the integrand:

$$\frac{x^{2k} - x^k y^k}{x-y} \left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right)$$

If we can reduce this down to somethign like $\sum c_k f_k(x) g_k(y)$, where f,and g are Hermite polynomials, then the whole double integral appearing above will be split up into a big sum of products of integrals $\sum\sum c_k \left(\int f(x)\psi_k(x)\mathrm{d}x\right)\left(\int g(y)\psi_k(y)\mathrm{d}y\right)$ which we'll be able to deal with by orthogonality. Since $K^{(N)}(x,y)$ only has terms up to $\psi_{N-1}(x)\psi_{N-1}(y)$ we dont have to keep track of anything ψ_l for $l \geq n$. We will write "0" for these ommited terms

Let's do it! Start by expanding:

$$\begin{aligned}
\frac{x^{2k} - x^k y^k}{x-y} \left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right) &= x^k\left(\sum_{i=1}^{k} y^{k-i}x^{i-1}\right)\left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right) \\
&= \sum_{i=1}^{k} x^{k+i-1}y^{k-i}\psi_n(x)\psi_{n-1}(y) \\
&\quad - \sum_{i=1}^{k} x^{k+i-1}y^{k-i}\psi_{n-1}(x)\psi_n(y)
\end{aligned}$$

So then putting this into $I_k^{(N)}$ and using $K^{(N)}(x,y) = \sum_{j=0}^{N-1}\psi_j(x)\psi_j(y)$we get to:

$$\begin{aligned}
I_k^{(N)} &= \frac{1}{N^{k+3/2}} \int\int \frac{x^{2k} - x^k y^k}{x-y} \left(\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)\right) K^{(N)}(y,x)\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{N^{k+3/2}} \sum_{j=0}^{N-1}\sum_{i=1}^{k} \left(\int x^{k+i-1}\psi_n(x)\psi_j(x)\mathrm{d}x\right)\left(\int y^{k-i}\psi_{n-1}(y)\psi_j(y)\mathrm{d}x\right) \\
&\quad - \frac{1}{N^{k+3/2}} \sum_{j=0}^{N-1}\sum_{i=1}^{k} \left(\int x^{k+i-1}\psi_{n-1}(x)\psi_j(x)\mathrm{d}x\right)\left(\int y^{k-i}\psi_n(y)\psi_j(y)\mathrm{d}x\right)
\end{aligned}$$

From here I think you can calculate $\int x^k \psi_n \psi_m \mathrm{d}x$ using the 3 term recurrance $x\psi_n = \sqrt{n+1}\psi_{n+1} + \sqrt{n}\psi_{n-1}$ to reduce the power k down to 0. Instead of doing this exactly, I'll just do an estimate. At each step you double the number of terms, and you get a factor of at most $\sqrt{n+1}$ in front, and the resulting integral at the last step is either 0 or 1. So doing this k times gives us the upper bound $\int x^k \psi_n \psi_m \mathrm{d}x \leq (2\sqrt{n+1})^k \cdot 1 \simeq n^{k/2} 2^k$ Putting this estimate in, I have (I'll do it up to constants and put t):

$$\begin{aligned} I_k^{(N)} &\leq \frac{1}{N^{k+3/2}} \sum_{j=0}^{N-1} \sum_{i=1}^{k} \sqrt{n}^{k+i-1} \sqrt{n}^{k-i} \\ &\leq \frac{1}{N^{k+3/2}} N \cdot N^{k-1/2} \\ = \tfrac{1}{N} &\to 0 \end{aligned}$$

□

4 Quick Introduction to Fredholm Determinants

4.1 The setting, fundemental estimates and defintion of the Fredholm determinant

Let X be a locally compact Polish space (a complete seperable metric space) with $\mathcal{B}_X$ its Borel sigma algebra. Let ν be a complex-valued measure on $(X, \mathcal{B}_X)$ such that:

$$\|\nu\|_1 = \int_X |\nu(\mathrm{d}x)| < \infty$$

(In many applications $X = \mathbb{R}$ and ν is a scalar multiple of the Lebesgue measure on a bounded interval)

Definition. (3.4.1.) A **kernal** is a Borel measurable, complex-valued function $K(x, y)$ defined on $X \times Y$ such that:

$$\|K\| := \sup_{x,y \in X \times X} |K(x, y)| < \infty$$

The **trace** is:

$$\mathrm{Tr}(K) = \int K(x, x)\mathrm{d}\nu(x)$$

Given two kernals, we define their composition by:

$$(K \star L)(x, y) = \int K(x, z)L(z, y)\mathrm{d}\nu(z)$$

These are well defined for $\|\nu\|_1 < \infty$ and $\|K\| < \infty$. By Fubini (its ok since the things are bounded), we have that $\mathrm{Tr}(K \star L) = \mathrm{Tr}(L \star K)$ and $(K \star L) \star M = K \star (L \star M)$

Remark. Watch out becasue the trace looks on the diagonal which is a null set, so for instance we could have $K = K'$a.e. and yet $\mathrm{Tr}(K) \neq \mathrm{Tr}(K')$.

Lemma. *(3.4.2.) Fix $n > 0$ for any kernal F we have:*

$$\left| \det_{i,j=1}^{n} F(x_i, y_j) \right| \leq n^{n/2} \|F\|^n$$

two kernals $F(x,y)$ and $G(x,y)$ we have:

$$\left| \det_{i,j=1}^{n} F(x_i, x_j) - \det_{i,j=1}^{n} G(x_i, y_j) \right| \leq n^{1+n/2} \|F - G\| \max\left(\|F\|, \|G\|\right)^{n-1}$$

Proof. I'm going to skip the details of the proof. The big finish is to use Hadamard's iinequality that for a matrix A whose columns are v_i we have $|\det(A)| \leq \prod \|v_i\|_{L^2}$ (this makes sense if you think about the determinant as the "volume" of a box). Our matrices have entries no larger than $\|F\|$, so each column has $\|v_i\|_{L^2} \leq \sqrt{n \|F\|^2} = n^{1/2} \|F\|$ and so $\left|\det_{i,j=1}^{n} F(x_i, y_j)\right| \leq n^{n/2} \|F\|^n$ follows. The other inequality is a bit harder, but the idea is to make an auxilliary kernals $H_i^{(k)}$ so that $\det_{i,j=1}^{n} F(x_i, x_j) - \det_{i,j=1}^{n} G(x_i, y_j) = \sum_{k=1}^{n} \det_{i,j=1}^{n} H_i^{(k)}(x_i, x_j)$ becomes a telescoping sum and applying the Hadmarad inequality to these guys gives the inequality. □

Definition. For a kernal K put:

$$\Delta_n = \Delta_n(K, \nu) = \int \ldots \int \det_{i,j=1}^{n} K(\xi_i, \xi_j) \, \mathrm{d}\nu(\xi_1) \ldots \mathrm{d}\nu(\xi_n)$$

and define $\Delta_0 = 1$. The above estimate on $\det_{i,j=1}^{n} K(\xi_i, \xi_j)$ shows that the integral is well defined and moreover we have the estimae that $|\Delta_n| \leq \|\nu\|_1^n \|K\|^n n^{n/2}$

Definition. (3.4.3.) The **Fredholm determinant** associated with the kernal K is defined as:

$$\Delta(K) = \Delta(K, \nu) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \Delta_n(K, \nu)$$

Remark. (3.4.4.) Here is some motivation for calling $\Delta(K)$ a determinant. Let $f_1(x), \ldots, f_N(x)$,$g_1(x), \ldots, g_N(x)$ be given. Put:

$$K(x,y) = \sum_{i=1}^{N} f_i(x) g_i(y)$$

Assume that $K(x,y)$ is a kernal (it suffices that the $f's$and $g's$ are all bounded. By using the expansion we saw earleir, $\det\left[I + hA_{ij}\right] = 1 + \sum_{k=1}^{N} h^k \sum_{0 \leq v_1 < \ldots < v_k \leq N-1} \det_{i,j=1}^{k} \big(A$

with $h=-1$, and using the $\int$ det to det $\int$ lemma, $\frac{1}{n!}\int\int\ldots\int \det_{i,j=1}^{n}\left(\sum_{k=1}^{n} f_k(x_i)g_k(x_j)\right)\prod_{i=1}^{n}\mathrm{d}x_i = \det_{i,j=1}^{n}\int f_i(x)g_j(x)\mathrm{d}x$ we have that:

$$\begin{aligned}\det_{i,.j=1}^{N}\left(\delta_{ij}-\int f_i(x)g_j(x)\mathrm{d}\nu(x)\right) &= 1+\sum_{k=1}^{N}(-1)^k\sum_{0\le v_1<\ldots<v_k\le N-1}\det_{i,j=1}^{k}\left(\int f_{v_i}(x)g_{v_j}(x)\mathrm{d}\nu(x)\right)\\ &= 1+\sum_{k=1}^{N}\frac{(-1)^k}{k!}\sum_{0\le v_1<\ldots<v_k\le N-1}\int\int\ldots\int\det_{i,j=1}^{k}\left(\sum_{i=1}^{k}f_{v_i}(x_i)g_{v_j}(x_j)\right)\\ &= 1+\sum_{k=1}^{N}\frac{(-1)^k}{k!}\int\int\ldots\int\sum_{0\le v_1<\ldots<v_k\le N-1}\det_{i,j=1}^{k}\left(\sum_{i=1}^{k}f_{v_i}(x_i)g_{v_j}(x_j)\right)\end{aligned}$$

Focus attention on the k-th term of the sum now. We know use the following version of the Cauchy Binet formula: Let A be a $k\times N$ matrix and let $C=AA^\star$(this is a $k\times k$ matrix), then $\det C=\sum_{K\in\mathcal{K}_{p,N}}\det A_K\det A_K^\star$where $\mathcal{K}_{k,N}$ is the set of all k element subsets of $\{1,\ldots,N\}$ and A_K is the $k\times k$ matrix which is obtained from A by keeping only the columns in K. If we take A to be the $k\times N$ matrix $A=[f_i(x_j)]_{\,i=1..N,j=1..k}$ then $C=\left[\sum_{l=1}^{N}f_l(x_i)g_l(x_j)\right]_{i,j=1}^{k}=[K(x_i,x_j)]_{i,j=1}^{k}$ is exaclth $C=AA^*$. The matrix $A_KA_K^*$ is exactly $\left[\sum_{i=1}^{k}f_{v_i}(x_i)g_{v_j}(x_j)\right]_{i,j=1}^{k}$ so the Cauchy Binet theorem exactly applies and we have:

$$\begin{aligned}\det_{i,.j=1}^{N}\left(\delta_{ij}-\int f_i(x)g_j(x)\mathrm{d}\nu(x)\right) &= 1+\sum_{k=1}^{N}\frac{(-1)^k}{k!}\int\int\ldots\int\det_{i,j=1}^{k}\left(\sum_{l=1}^{N}f_l(x_i)g_l(x_j)\right)\prod_{i=1}^{n}\mathrm{d}x_i\\ &= 1+\sum_{k=1}^{N}\frac{(-1)^k}{k!}\Delta_k\\ &= \Delta\end{aligned}$$

It is ok that we only sum to N in this case, since $\Delta_k=0$ for $k\ge N$ as the matrix $C=AA^*$ cannot be of full rank.

Example. (Determinantal Projection Process)

Let's see what this means for a DPP. A determinental point process is a point process whose k point correlation function is given by:

$$\rho_k(x_1,\ldots,x_k)=\det\left(K(x_i,x_j)\right)_{1\le i,j\le k}$$

Where the k point correlation function is defined so that for disjoiint sets $D_1,D_2,\ldots,D_n$ we have:

$$\mathbf{E}\left[\prod_{i=1}^{k}\chi(D_i)\right]=\int_{D_1}\ldots\int_{D_k}\rho_k\left(x_1,\ldots,x_k\right)\mathrm{d}\mu(x_1)\ldots\mathrm{d}\mu(x_k)$$

i.e. $\rho_k(x_1,\ldots,x_k)$ is the density of finding particles at the points $x_1,\ldots,x_k$. To handle overlapping sets, $B=\overbrace{D_1\times\ldots\times D_1}^{k_1}\times\ldots\times\overbrace{D_r\times\ldots\times D_r}^{k_r}$ where

$D_1, \ldots D_r$ are disjoint sets, and $\sum_{i=1}^{r} k_i = k$ then this becomes:

$$\mathbf{E}\left[\prod_{i=1}^{r}\binom{\chi(D_i)}{k_i}k_i!\right] = \int_B \rho_k\left(x_1, \ldots, x_k\right) \mathrm{d}\mu(x_1)\ldots d\mu(x_k)$$

($\binom{x}{k}k!$ is the falling factoiral $(x)_k = x(x-1)\ldots(x-k+1)$)If we choose $B = \overbrace{\Lambda \times \Lambda \times \ldots \times \Lambda}^{k}$ then $\chi(\Lambda)$ is simply the total number of points (which is a random variable), so this says that:

$$\begin{aligned}\mathbf{E}\left[\binom{\chi(\Lambda)}{k}k!\right] = \mathbf{E}\left[(\chi(\Lambda))_k\right] &== \int_{\Lambda\times\Lambda\times\ldots\times\Lambda} \rho_k\left(x_1, \ldots, x_k\right) \mathrm{d}\mu(x_1)\ldots d\mu(x_k) \\ &= \Delta_k\end{aligned}$$

So Δ_k has a very concrete interpratation for DPP's.

For example, a poisson process, (which is a DPP with a trivial kernal $K(x,x) = \delta_{x,x}$) is characterized by $\mathbf{E}\left[(\chi(\Lambda))_k\right] = \Delta_k = \lambda^k$. Hence $\Delta = 1 + \sum_{k=1}^{N} \frac{(-1)^k}{k!}\Delta_k = \exp\left(-\lambda\right)$. As it happens, $\exp(-\lambda) = \mathbf{P}$ of zero points, I'm not sure if this is related.

Another example: Suppose that $\chi(\Lambda) = N$ almost surely. Then this is saying that there are exactly N points. In this case then $\Delta_k = 0$ for $k > N$. So we have a finte sum for Δ, namely:

$$\begin{aligned}\Delta &= 1 + \sum_{k=1}^{N}\frac{(-1)^k}{k!}\Delta_k \\ &= 1 + \sum_{k=1}^{N}\frac{(-1)^k}{k!}\binom{N}{k}k! \\ &= 1 + \sum_{k=1}^{N}(-1)^k\binom{N}{k} \\ &= 0\end{aligned}$$

This is 0 for example by the binomial expansion for $(1-1)^N$. I think this proof actually shows that if the number of points is BOUNDED then $\Delta = 0$ too, for we can right $\Delta = \mathbf{E}(\cdots) = \sum_k \mathbf{E}\left(\cdots|\chi(\Lambda) = k\right)\mathbf{P}\left(\chi(\Lambda) = k\right)$ and it is zero on each piece. (It doesn't work for an infinite sum though because Δ_k could be very large on the set where

An example of a DPP with exactly N points is a projection process. For a projection process, where the kernal is made of some orthogonal set of $L^2(\Lambda)$ functions, say $\psi_1, \ldots, \psi_n$ and the kernal is given by $\sum_{i=1}^{n}\psi_i(x)\psi_i(y)$, the other formula for Δ in this case is: $\Delta = \det_{i,j=1}^{N}\left(\delta_{ij} - \int \psi_i\psi_j \mathrm{d}x\right) = \det_{i.j=1}^{n}\left[\delta_{ij} - \delta_{ij}\right] = 0$, so it lines up!

This gives me the feeling that Δ is measuring something about how much randomness there is in the number of points....ok back to AGZ now.

From looking at the paper on DPP's and Random Matrices by Johansson, I see that Δ is like a GAP probability. What I've been calculating above is the probability that there are no particles anywhere. For a projection process or a Poisson process, its not surprising that this is 0.

Lemma. *(3.4.5.) For any two kernals $K(x,y)$ and $L(x,y)$ we have:*

$$|\Delta(K) - \Delta(L)| \leq \left(\sum_{n=1}^{\infty} \frac{n^{1+n/2} \|\nu\|_1^n \max\left(\|K\|, \|L\|\right)^{n-1}}{n!} \right) \|K - L\|$$

Proof. Sum the estimate we had from Lemma 3.4.2. □

Corollary. *If K is fixed and we have L varying in such a way that $\|K - L\| \to 0$ then it follows that $\Delta(L) \to \Delta(K)$.*

Remark. In practice, we make it so that $\Delta(L)$ is some sort of relavent probability thing we care about, we show that $L \to K$ and thus we know that $\Delta(L) \to \Delta(K)$.

4.2 Definition of the Fredhold Adjugant, Fredholm resolvent and a fundemental identity.

Define a shorthand:

$$K\begin{pmatrix} x_1 \dots x_n \\ y_1 \dots y_n \end{pmatrix} := \det_{i,j=1}^{n} K(x_i, y_j)$$

Set:

$$H_n(x,y) = \int \dots \int K\begin{pmatrix} x\ \xi_1\ \dots\ \xi_n \\ y\ \xi_1\ \dots\ \xi_n \end{pmatrix} \mathrm{d}\nu(\xi_1)\dots\mathrm{d}\nu(\xi_n)$$

and:

$$H_0(x,y) = K(x,y)$$

We have from Lemma 3.4.2. that:

$$|H_n(x,y)| \leq \|K\|^{n+1} \|\nu\|_1^n (n+1)^{(n+1)/2}$$

Definition. (3.4.6.) The **Fredholm adjugant** of the Kernal $K(x,y)$ is the function:

$$H(x,y) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} H_n(x,y)$$

If $\Delta(K) \neq 0$ we define the **resolvent** of the Kernal $K(x,y)$ as the function:

$$R(x,y) = \frac{H(x,y)}{\Delta(K)}$$

By the estimate on $|H_n(x,y)|$, the series that defines $H(x,y)$ above converges absolutly and uniofrmly on $X \times X$. Therefore H is well defined and it is continuous on X.

Lemma. *(3.4.7.) (The fundemental identity) Let $H(x,y)$ be the Fredholm adjugant of the kernal $K(x,y)$ then:*

$$\begin{aligned}\int K(x,z)H(z,y)d\nu(z) &= H(x,y) - \Delta(K)\cdot K(x,y)\\ &= \int H(x,z)Z(z,y)d\nu(z)\end{aligned}$$

i.e.:

$$K \star H = H - \Delta(K)\cdot K = H \star K$$

Remark. If $\Delta(K) \neq 0$ so that R makes sense, then this says that:

$$K \star R = R - K = R \star K$$

Which one could write as the operator identity:

$$1 + R = (1-K)^{-1}$$

Proof. I'm going to skip this for now. □

Corollary. *(3.4.9.) (i) For all $n \geq 0$:*

$$\frac{(-1)^n}{n!}H_n(x,y) = \sum_{k=0}^{n}\frac{(-1)^k}{k!}\Delta_k \cdot \left(\underbrace{K \star \ldots \star K}_{n+1-k}\right)(x,y)$$

ii) Furthermore:

$$\frac{(-1)^n}{n!}\Delta_{n+1} = \sum_{k=0}^{n}\frac{(-1)^k}{k!}\Delta_k \cdot Tr\left(\underbrace{K \star \ldots \star K}_{n+1-k}\right)$$

In partiuclar the sequence of numbers $Tr(K)$, $Tr(K \star K)$, $Tr(K \star K \star K)$ uniquely determines the Fredholm determinant $\Delta(K)$

I'm going to skip a section now thats useful for the GOE and GSE.

5 Gap Probabilites at 0 and the Proof of Theorem 3.1.1.

Let $X_n \in \mathscr{H}_N^{(2)}$ be a random Hermitian matrix from the GUE with eigenvalues $\lambda_1^N \leq \ldots \leq \lambda_N^N$. We initate in this section the study of the spacings between the eigenvaleues of X_N. We focus on those eigencalues that lie near 0. i.e. we wish to calculate something like:

$$\lim_{N\to\infty}\mathbf{P}\left[\sqrt{N}\lambda_1^N, \ldots, \sqrt{N}\lambda_N^N \notin (-t/2, t/2)\right]$$

This has a chance of being some non-degenerate probability because the N random varaibles $\sqrt{N}\lambda_1^N, \ldots \sqrt{N}\lambda_N^N$ are spread out over an interval of length rouhgly $4N$.

As in (3.2.4.) set:

$$K^{(n)}(x,y) = \sum_{k=0}^{n-1} \psi_k(x)\psi_k(y) = \sqrt{n}\frac{\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)}{x-y}$$

and set:

$$S^{(n)}(x,y) = \frac{1}{\sqrt{n}} K^{(n)}\left(\frac{x}{\sqrt{n}}, \frac{y}{\sqrt{n}}\right)$$

A crucial step in the proof is the following convergence lemma:

Lemma. *(3.5.1.) :*

$$\lim_{n\to\infty} S^{(n)}(x,y) = \frac{1}{\pi}\frac{\sin(x-y)}{x-y}$$

and the convergence is uniform on each bounded subset of the x, y plane.

Proof. We will deffer the proof for now and prove it a bit later. □

Theorem. *(3.1.1.) (Gaudin-Mehta) For any compact set $A \subset \mathbb{R}$:*

$$\lim_{N\to\infty} \mathbf{P}\left(\sqrt{N}\lambda_1^N, \sqrt{N}\lambda_2^N, \ldots, \sqrt{N}\lambda_N^N \notin A\right) = 1+\sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \int_A \ldots \int_A \det_{i,j=1}^{k} K_{sine}(x_i, x_j)\, dx_1\, dx_2 \ldots dx_k$$

where:

$$K_{sine}(x,y) = \frac{1}{\pi}\frac{\sin(x-y)}{x-y}$$

and is understood to be $1/\pi$ if $x = y$ (i.e. determined by continuity)

Proof. By lemma 3.2.4, we have that $\mathbf{P}_N\left(\bigcap_{i=1}^N \{\lambda_i \in A\}\right) = 1+\sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^c}, \ldots, \int_{A^c} \det_{i,j=1}^k K^{(N)}(x_i, x$
Hence:

$$\begin{aligned}
\mathbf{P}_N\left(\sqrt{N}\lambda_1^N, \sqrt{N}\lambda_2^N, \ldots, \sqrt{N}\lambda_N^N \notin A\right) &= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\sqrt{n}^{-1}A}, \ldots, \int_{\sqrt{n}^{-1}A} \det_{i,j=1}^{k} K^{(N)}(x_i, x_j) \prod_{i=1}^{k} \mathrm{d}x_i \\
&= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A, \ldots, \int_A \det_{i,j=1}^{k} S^{(N)}(x_i, x_j) \prod_{i=1}^{k} \mathrm{d}x_i \\
&= \Delta_n(S^N, \nu 1_A) \\
&\to \Delta_n\left(K_{sine}, \nu 1_A\right) \\
&\to 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \int_A \ldots \int_A \det_{i,j=1}^{k} K_{sine}(x_i, x_j) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_k
\end{aligned}$$

The last convergence follows since $S^{(N)} \to K_{sine}$ uniformly on compact sets (this is the lemma 3.5.1. we have deffered for now) and since Fredholm determinants respect convergence like this (because of the estimate $|\Delta(K) - \Delta(L)| \leq (\sum_{n=1}^{\infty} *stuff*) \|K-L\|$, this is lemma 3.4.5.) □

5.1 The Method of Laplace

Laplace's method deals with asymptoics as $s \to \infty$ of integrals of the form:

$$\int f(x)^s g(x) \mathrm{d}x$$

We will be concerend with the situation in which the function f posses a global maximum at some point a. You can see that $f(a)$ will then dominate the integral as $s \to \infty$, so only the value $g(a)$ should count asymptotically.

Definition. To make this rigourous we need some extra conditions. Let $f : \mathbb{R} \to \mathbb{R}_+$ be given and for some constant a and some positive constants s_0, K, L, M let $\mathscr{G} = \mathscr{G}(a, \epsilon_0, s_0, f(\cdot), K, L, M)$ be the class of measurable functions $g : \mathbb{R} \to \mathbb{R}$ satsifying:

i) $|g(a)| \leq K$
(i.e. $g(a)$ is bounded)
ii) $\sup_{0<|x-a|\leq\epsilon_0} \left|\frac{g(x)-g(a)}{x-a}\right| \leq L$
(i.e. g is Lipshitz with constant L in a n'h'd of a)
iii) $\int f(x)^{s_0} |g(x)| \mathrm{d}x \leq M$
The theorem is that:

Theorem. *(3.5.3.) (Laplace) Let $f : \mathbb{R} \to \mathbb{R}_+$ be a function such that for some $a \in \mathbb{R}$ and some positive constants ϵ_0, c the following hold:*
a) $f(x) \leq f(x')$ if $a - \epsilon_0 \leq x \leq x' \leq a$ or $a \leq x' \leq x \leq a + \epsilon_0$
(i.e. f is increasing in $[a - \epsilon_0, a]$ and is decreasing in $[a, a + \epsilon_0]$. This means that f must have a maximum at $x = a$)
b) For all $\epsilon < \epsilon_0$ $\sup_{|x-a|>\epsilon} f(x) \leq f(a) - c\epsilon^2$
(i.e. the maximum at $x = a$ is like a "quadratic maximum")
c) $f(x)$ has two continuous derivatives in the interval $(a - 2\epsilon_0, a + 2\epsilon_0)$
d) $f''(a) < 0$
Then for any function $g \in \mathscr{G}(a, \epsilon_0, s_0, f(\cdot), K, L, M)$ we have:

$$\lim_{s\to\infty} \sqrt{s} f(a)^{-s} \int f(x)^s g(x) dx = \sqrt{\frac{2\pi f(a)}{|f''(a)|}} g(a)$$

Remark. The intuition of the proof is as follows. We have that $(f(x)/f(a))^s$ becomes negligible everywhere except near $x = a$ and here we know that it peaks like a bell curve.

Remark. Here is a heuristic proof of a related formula:

$$\exp(-n\ell(a)) \sqrt{n} \int \phi(x) \exp(n \cdot \ell(x)) \mathrm{d}x \quad \sim \quad \phi(a) \left[\frac{2\pi}{|\ell''(a)|}\right]^{1/2}$$

The proof goes by Taylor expanding around a:

$$\begin{aligned}\int \phi(x)\exp\left(n\cdot\ell(x)\right)\mathrm{d}x &\approx \int \phi(a)\exp\left(n\left(\ell(a)+(x-a)^2\frac{\ell''(a)}{2}\right)\right)\mathrm{d}x \\ &\approx \phi(a)\exp\left(n\cdot\ell(a)\right)\int_{-\infty}^{\infty}\exp\left(n\cdot u^2\frac{\ell''(a)}{2}\right)\frac{\mathrm{d}u}{\sqrt{n}} \\ &= \phi(a)\exp\left(n\cdot\ell(a)\right)\left[\frac{2\pi}{n\left|\ell''(a)\right|}\right]^{1/2}\end{aligned}$$

To translate to the AGZ version of the proof, put $\ell(x) = \log(f(x))$ and $\phi(x) = g(x)$. Notice also that $\ell'(x) = f'(x)/f(x)$ (so $\ell'(a) = 0$) and $\ell''(x) = \left(f''(x)f(x) - f'(x)^2\right)/f(x)^2$ at $x = a$, since $f'(a) = 0$ this gives $\ell''(a) = f''(a)/f(a)$ to finish the translation between the two.

Proof. We will split the integral up into the contribution from near $x = a$ and far away from $x = a$, where near/far means either $|x-a| < \epsilon(s)$ or $|x-a| > \epsilon(s)$ where $\epsilon(s)$ is a parameter depending on s which we will fix precisely later. We will then show that the contribution from far away terms $\to 0$ as $s \to \infty$ while the contribution from terms near a will tend to the RHS of the claimed equality. (It will turn out that any $\epsilon(s)$ satisfying $\epsilon(s) < \epsilon_0$ and $\epsilon(s) \to 0$ as $s \to \infty$ and $\sqrt{s}\epsilon(s) \to \infty$ as $s \to \infty$ will do)

Write:

$$\int f(x)^s g(x)\mathrm{d}x = g(a)I_1 + I_2 + I_3$$

Where I_1, I_2, I_3 are integrals that depend on s:

$$\begin{aligned}I_1 &= \int_{|x-a|\le\epsilon(s)} f(x)^s\mathrm{d}x \\ I_2 &= \int_{|x-a|\le\epsilon(s)} f(x)^s\left(g(x)-g(a)\right)\mathrm{d}x \\ I_3 &= \int_{|x-a|>\epsilon(s)} f(x)^s g(x)\mathrm{d}x\end{aligned}$$

[Demand #1 on $\epsilon(s)$: need $\epsilon(s) < \epsilon_0$ for every value of s so that we can use the Lipshitz property for g in this nhd and the fact that f is increasing/decreasing in this n'h'd of a.]

Claim 1: $\sqrt{s}f(a)^{-s}I_1 \to \sqrt{\frac{2\pi f(a)}{|f''(a)|}}$

As in the heuristic, it is more convenient to work with $\ell = \log(f)$ rather than f itself. (ℓ stands for the "l" in "log") We start by doing the integral version of the Taylor expansion for ℓ:

$$\ell(a+x) = \ell(a) + \ell'(a)x + \int_0^x u\cdot\ell''(x+a-u)\cdot\mathrm{d}u$$

$$= \ell(a) + \ell'(a)x + \int_0^x (x-t) \cdot \ell''(a+t) \cdot \mathrm{d}t$$

$$= \ell(a) + \ell'(a)x + x^2 \int_0^1 (1-r) \cdot \ell''(a+rx) \cdot \mathrm{d}r$$

Notice $\ell'(a) = 0$ since $f'(a) = 0$, so this term drops. Exponentiation both sides give

$$f(a+x) = f(a) \exp\left(x^2 \int_0^1 (1-r) \cdot \ell''(a+rx) \cdot \mathrm{d}r \right)$$

We can then do a change of variables on I_1 to get the result we want (this is a written exams type question!). It is convenient to denote $h(x) = \int_0^1 (1-r) \cdot \ell''(a+rx) \cdot \mathrm{d}r$ here.

$$\begin{aligned} I_1 &= \int_{|x-a| \le \epsilon(s)} f(x)^s \mathrm{d}x \\ &= \int_{-\epsilon(s)}^{\epsilon(s)} f(a+x)^s \mathrm{d}x \\ &= \int_{-\epsilon(s)}^{\epsilon(s)} f(a)^s \exp\left(sx^2 h(x)\right) \mathrm{d}x \\ &= \frac{f(a)^s}{\sqrt{s}} \cdot \int_{-\epsilon(s)\sqrt{s}}^{\epsilon(s)\sqrt{s}} \exp\left(t^2 h\left(\frac{t}{\sqrt{s}}\right)\right) \mathrm{d}t \text{ put } t = \sqrt{s}x \end{aligned}$$

[Demand #2 on ϵ_s, need $\sqrt{s}\epsilon(s) \to \infty$ as $s \to \infty$ so that this integrand becomes an integral on the whole real line.] The integrand converges pointwise as $s \to \infty$ to $\exp\left(t^2 h(0)\right)$ and the domain of integration converges to all of $(-\infty, \infty)$. By the dominated convergence theorem, we know that this integral is converging to the integral of $\exp\left(t^2 h(0)\right)$ on the whole real line then. Finally then:

$$\lim_{s \to \infty} \sqrt{s} f(a)^{-s} I_1 = \int_{-\infty}^{\infty} \exp\left(t^2 h(0)\right) \mathrm{d}t = \frac{\sqrt{\pi}}{\sqrt{|h(0)|}} = \sqrt{\frac{2\pi f(a)}{|f''(a)|}}$$

Where we have used the computation $h(0) = h(x) = \int_0^1 (1-r) \cdot \ell''(a) \cdot \mathrm{d}r = \int_0^1 (1-r) \frac{f''(a)}{f(a)} \mathrm{d}r = \frac{1}{2} \frac{f''(a)}{f(a)}$ and that this is < 0.

Claim 2: $\sqrt{s} f(a)^{-s} I_2 \to 0$ as $s \to \infty$.

Pf: This follows by the use of the Lipshitz constant L for g and the fact that the length of the interval is small. Have:

$$\begin{aligned}\left|\int\limits_{|x-a|\leq\epsilon(s)} f(x)^s\left(g(x)-g(a)\right)\mathrm{d}x\right| &\leq \int\limits_{|x-a|\leq\epsilon(s)} f(x)^s\,|x-a|\,L\mathrm{d}x\\ &\leq \epsilon(s)L\int\limits_{|x-a|\leq\epsilon(s)} f(x)^s\mathrm{d}x\\ &= \epsilon(s)L\cdot I_1\end{aligned}$$

Hence $\sqrt{s}f(a)^{-s}I_2 = \left(\sqrt{s}f(a)^{-s}I_1\right)L\epsilon(s)\to 0$ as long as we demand $\epsilon(s)\to 0$.

Claim 3: $\sqrt{s}f(a)^{-s}I_3\to 0$ as $s\to\infty$

Pf: This is due to the fact that f drops off quadratically away from a. Have:

$$\begin{aligned}|I_3| &= \left|\int\limits_{|x-a|>\epsilon(s)} f(x)^s g(x)\mathrm{d}x\right|\\ &\leq \left|\int\limits_{|x-a|>\epsilon(s)} f(x)^{s-s_0}f(x)^{s_0}g(x)\mathrm{d}x\right|\\ &\leq \left|\int\limits_{|x-a|>\epsilon(s)} f(x)^{s_0}g(x)\mathrm{d}x\right|\sup_{|x-a|>\epsilon(s)}|f(x)|^{s-s_0}\\ &\leq M\left(f(a)-c\epsilon(s)^2\right)^{s-s_0}\end{aligned}$$

Hence:

$$\begin{aligned}\sqrt{s}f(a)^{-s}I_3 &\leq M\sqrt{s}f(a)^{-s}\left(f(a)-c\epsilon(s)^2\right)^{s-s_0}\\ &= M\sqrt{s}f(a)^{-s_0}\left(1-\frac{c\epsilon(s)^2}{f(a)}\right)^{s-s_0}\end{aligned}$$

And since $\epsilon(s)^2 s\to\infty$ this goes to "$e^{-\infty}$" $=0$ as $s\to\infty$. (In general $(1-\alpha(x))^{\beta(s)}\to e^{-\alpha(x)\beta(x)}$ as $x\to\infty$.) □

Exercise. Use Laplace's method with $a=1$ to prove Stirling's approximation:

$$\Gamma(s)=\int_0^\infty x^s e^{-x}\frac{\mathrm{d}x}{x}=s^s\int_0^\infty\left(xe^{-x}\right)^s\frac{\mathrm{d}x}{x}\sim\sqrt{2\pi}s^{s-1/2}e^{-s}$$

Proof. Put $g(x)=\frac{1}{x}$ and $f(x)=xe^{-x}$ so that $f(1)=e^{-1}$, $f'(x)=-xe^{-x}+e^{-x}$,$f''(x)=xe^{-x}-e^{-x}-e^{x}$,$f''(1)=-1$ then Laplace's method tells us that:

$$\lim_{s\to\infty}s^{1/2}\cdot(e^{-1})^{-s}\cdot\int\left(xe^{-x}\right)^s\left(\frac{1}{x}\right)\mathrm{d}x=\sqrt{2\pi\cdot\frac{1}{1}}\cdot 1$$

Or in other words:

$$s^s \int_0^\infty \left(xe^{-x}\right)^s \frac{\mathrm{d}x}{x} \sim \sqrt{2\pi} s^{s-1/2} e^{-s}$$

□

5.2 Evaluation of the scaling limit: Proof of Lemma 3.5.1.

Let:

$$\Psi_v(t) = n^{1/4}\psi_v\left(\frac{t}{\sqrt{n}}\right)$$

Where v is a quantity whose difference from n is fixed (e.g. $v = n, n-1$ or $n-2$) The main asymptotic result we need is that:

Lemma. *(3.5.4.) Uniformly of for t in a fixed bounded interval:*

$$\lim_{n\to\infty}\left|\Psi_v(t) - \frac{1}{\sqrt{\pi}}\cos\left(t - \frac{\pi v}{2}\right)\right| = 0$$

Proof. Recall the Fourier transform identity:

$$e^{-x^2/2} = \frac{1}{\sqrt{2\pi}}\int e^{-\xi^2/2}e^{-i\xi x}\mathrm{d}\xi$$

Hence the n-th Hermite polynomial is:

$$\mathcal{H}_n(x)e^{-x^2/2} = (-1)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(e^{-x^2/2}\right) = \frac{1}{\sqrt{2\pi}}\int (i\xi)^n e^{-\xi^2/2-i\xi x}\mathrm{d}\xi$$

So we get:

$$\begin{aligned}\Psi_v(t) &= \frac{i^v e^{t^2/4n}n^{1/4}}{(2\pi)^{3/4}\sqrt{v!}}\int \xi^v e^{-\xi^2/2-i\xi t/\sqrt{n}}\mathrm{d}\xi \\ &= \dots \\ &= C_{v,n}e^{n/2}\int_{-\infty}^{\infty} f(\xi)^n g_t(\xi)\mathrm{d}\xi\end{aligned}$$

Where $f(x) = xe^{-x^2/2}1_{x\geq 0}$ and $g(x) = g_t(x) = \cos(xt - \frac{\pi v}{2})x^{v-n}$. So from here we can apply Lapalce's method to get the result. □

With this method in hand:

Lemma. *(3.5.1.):*

$$\lim_{n\to\infty} S^{(n)}(x,y) = \frac{1}{\pi}\frac{\sin(x-y)}{x-y}$$

and the convergence is uniform on each bounded subset of the x, y plane.

Proof. Recall that:

$$S^{(n)}(x,y) = \sqrt{n}\frac{\psi_n\left(\frac{x}{\sqrt{n}}\right)\psi_{n-1}\left(\frac{y}{\sqrt{n}}\right) - \psi_{n-1}\left(\frac{x}{\sqrt{n}}\right)\psi_n\left(\frac{y}{\sqrt{n}}\right)}{x-y}$$

We now rewrite this in such a way as to make the $x-y$ under the rug. The trick is:

$$\begin{aligned}\frac{f(x)g(y)-f(y)g(x)}{x-y} &= \left(\frac{f(x)-f(y)}{x-y}\right)g(y) + f(y)\left(\frac{g(y)-g(x)}{x-y}\right) \\ &= g(y)\int_0^1 f'(tx+(1-t)y)\mathrm{d}t - f(y)\int_0^1 g'(tx+(1-t)y)\mathrm{d}t\end{aligned}$$

So we get:

$$\begin{aligned}S^{(n)}(x,y) &= \psi_{n-1}\left(\frac{y}{\sqrt{n}}\right)\int_0^1 \psi_n'\left(t\frac{x}{\sqrt{n}}+(1-t)\frac{y}{\sqrt{n}}\right)\mathrm{d}t \\ &\quad -\psi_n\left(\frac{y}{\sqrt{n}}\right)\int_0^1 \psi_{n-1}'\left(t\frac{x}{\sqrt{n}}+(1-t)\frac{y}{\sqrt{n}}\right)\mathrm{d}t \\ &= \psi_{n-1}\left(\frac{y}{\sqrt{n}}\right)\int_0^1 \sqrt{n}\psi_{n-1}(z) - \frac{z}{2}\psi_n(z)\Big|_{z=\left(t\frac{x}{\sqrt{n}}+(1-t)\frac{y}{\sqrt{n}}\right)}\mathrm{d}t \\ &\quad -\psi_n\left(\frac{y}{\sqrt{n}}\right)\int_0^1 \sqrt{n-1}\psi_{n-2}(z) - \frac{z}{2}\psi_{n-1}(z)\Big|_{z=\left(t\frac{x}{\sqrt{n}}+(1-t)\frac{y}{\sqrt{n}}\right)}\mathrm{d}t\end{aligned}$$

By the convergence of ψ_n we proved in Lemma 3.5.4. this converges to: (The terms with $z\psi_k$ die since $z = O(\frac{1}{\sqrt{n}})$ is very small. The other terms are exactly the right scaling for Lemma 3.5.4. by splitting up $\sqrt{n} = n^{1/4}n^{1/4}$. So we get:

$$\begin{aligned}S^{(n)}(x,y) &\sim \frac{1}{\pi}\left(\cos\left(y-\frac{\pi(n-1)}{2}\right)\right)\int_0^1 \cos\left(tx+(1-t)y-\frac{\pi(n-1)}{2}\right)\mathrm{d}t \\ &\quad -\frac{1}{\pi}\left(\cos\left(y-\frac{\pi n}{2}\right)\right)\int_0^1 \cos\left(tx+(1-t)y-\frac{\pi(n-2)}{2}\right)\mathrm{d}t \\ &= \frac{1}{\pi}\left(\sin\left(y-\frac{\pi n}{2}\right)\right)\int_0^1 \sin\left(tx+(1-t)y-\frac{\pi n}{2}\right)\mathrm{d}t \\ &\quad -\frac{1}{\pi}\left(\cos\left(y-\frac{\pi n}{2}\right)\right)\int_0^1 -\cos\left(tx+(1-t)y-\frac{\pi n}{2}\right)\mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{\pi}\frac{\cos(x-\frac{\pi n}{2})\sin\left(y-\frac{\pi n}{2}\right)-\cos(y-\frac{\pi n}{2})\sin\left(x-\frac{\pi n}{2}\right)}{x-y} \\
&= \frac{1}{\pi}\frac{\sin\left(\left(x-\frac{\pi n}{2}\right)-\left(y-\frac{\pi n}{2}\right)\right)}{x-y} \\
&= \frac{1}{\pi}\frac{\sin(x-y)}{x-y}
\end{aligned}
$$

(The last line follows by our identity for $\frac{f(x)g(y)-f(y)g(x)}{x-y}$ we had before) □

5.3 A complement: determinatal relations

Remark. This kind of thing is done in a bit more generality by Johansson in his survey "DPP and random matrices". I think I'm actually going to skip this for now.

6 Analysis of the sine-kernal

In this section the since kernal is analyzed and it is shown that it satisfies the Painleve differential equation and so on, I'm going to skip this for now.

7 Edge-Scaling: Proof of Theorem 3.1.4

Remark. (3.1.3.) Recall the **Airy function** is defined by the formula:

$$
\mathrm{Ai}(x) = \frac{1}{2\pi i}\int_C \exp\left(\frac{1}{3}z^3 - xz\right) \mathrm{d}z
$$

where C is the contour in the $z-$plane consititing of two rays, one from the direction $e^{-\pi i/3}$ coming from infinity to the origin and one from the origin to $e^{\pi i/3}$ to ∞.

The **Airy kernal** is defined by:

$$
K_{Airy}(x,y) = A(x,y) := \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y}
$$

with the value of $x = y$ determined by continuity.

Here is a preliminary estimate we need. The proof is deffered till later:

Lemma. *(3.7.1.) For any $x_0 \in \mathbb{R}$*

$$
\sup_{x,y\geq x_0} e^{x+y}\left|A(x,y)\right| < \infty
$$

7.1 Vague convergence of the largest eigenvalye: proof of Theorem 3.1.4.

Again, let $X_N \in \mathscr{H}_N^{(2)}$ be a random Hermitian matrix from the GUE with eigenvalues $\lambda_1^N \leq \ldots \leq \lambda_N^N$.

Theorem. *(3.1.4.) [The top eigevalue of the GUE] For all $-\infty < t \leq t' \leq \infty$ we have:*

$$\lim_{N\to\infty} \mathbf{P}\left[N^{2/3}\left(\frac{\lambda_i^N}{\sqrt{N}} - 2\right) \notin [t,t'], i = 1\ldots, N\right] = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_t^{t'}\int_t^{t'}\ldots\int_t^{t'} \det_{i,j=1}^{k} A(x_i,x_j)\,dx_1\,dx_2\ldots dx_k$$

With A the Airy kernal. In particular, the TOP eigevalue λ_N^N has:

$$\begin{aligned}\lim_{N\to\infty} \mathbf{P}\left[N^{2/3}\left(\frac{\lambda_N^N}{\sqrt{N}} - 2\right) \leq t\right] &= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_t^{\infty}\int_t^{\infty}\ldots\int_t^{\infty} \det_{i,j=1}^{k} A(x_i,x_j)\,dx_1\,dx_2\ldots dx_n \\ &=: F_2(t)\end{aligned}$$

$F_2(t)$ is the celebrated Tracy-Widom distribution.

Proof. As before, put:

$$K^{(n)}(x,y) = \sqrt{n}\frac{\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)}{x-y}$$

Define:

$$A^{(n)}(x,y) = \frac{1}{n^{1/6}} K^{(n)}\left(2\sqrt{n} + \frac{x}{n^{1/6}}, 2\sqrt{n} + \frac{y}{n^{1/6}}\right)$$

We will begin by extending $K^{(n)}$ and $A^{(n)}$ to the entire complex plane $\mathbb{C}$ by analyticity. Our goal will be to prove the convergence of $A^{(n)}$ to A on compacts sets of $\mathbb{C}$, which will also imply convergence of the derivatives. Recall that by part 4 of Lemma 3.2.7.:

$$K^{(n)}(x,y) = \frac{\psi_n(x)\psi_n'(y) - \psi_n(x)\psi_n'(x)}{x-y} - \frac{1}{2}\psi_n(x)\psi_n(y)$$

so that if we set:

$$\Psi_n(x) := n^{1/12}\psi_n(2\sqrt{n} + \frac{x}{n^{1/6}})$$

then:

$$A^{(n)}(x,y) = \frac{\Psi_n(x)\Psi_n'(y) - \Psi_n(y)\Psi_n'(x)}{x-y} - \frac{1}{2n^{1/3}}\Psi_n(x)\Psi_n(y)$$

The following lemma plays the role of Lemma 3.5.1 in the study of spacings in the bulk: (Rmk proof not over yet...ignore the square) □

Lemma. *(3.7.2.) Fix a number $C > 1$. Then:*

$$\lim_{n\to\infty} \sup_{|u|<C} |\Psi_n(u) - Ai(u)| = 0$$

Proof. We defer the rather long and technical proof to subsection 3.7.2. □

With this lemma in hand, we first notice that since Ψ_n and Ai are entire, this uniform convergence on compact subsets measnt atha $\Psi_n' \to \mathrm{Ai}'$ also uniformly on comapct subsets (e.g. by Cauchy integral fromulat). Hence

$$\begin{aligned} A^{(n)}(x,y) &= \frac{\Psi_n(x)\Psi_n'(y) - \Psi_n(y)\Psi_n'(x)}{x-y} - \frac{1}{2n^{1/3}}\Psi_n(x)\Psi_n(y) \\ &\to \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}(y)\mathrm{Ai}'(x)}{x-y} - 0 \\ &= A(x,y) \end{aligned}$$

Now by the estimate in lemma 3.4.5 we know that convergence of kernals implies convergence of the gap probabilites and we get the result.

7.2 Steepest descent: proof of Lemma 3.7.2.