Statistics C206 Lecture Notes
Eigenvalues of random matrices, Spring 2025

Professor: Vadim Gorin
Scribe: Zixun Huang, Yinghan Liu

Contents

1 Jan 21st

1.1 Origins, historic motivations, and sample results

In this course, we will primarily focus on random matrices and their eigenvalues. To provide motivation, we will begin by exploring the history of random matrix theory.

The history of random matrices can be divided into three significant waves. The first wave occurred during the **1920s to 1930s**, which in one side is related to the representation theory of Lie groups. From group theory, any compact group can induce a canonical probability measure, which is called Haar measure (formalization of uniform measure).

We can discuss a central example: consider the unitary group $G = \mathscr{U}(N)$, which consists of $N \times N$ unitary matrices u. The eigenvalues of u, denoted as $z_1, \ldots, z_N$, are N points on the unit circle. This is because for $ue_i = z_i e_i$, we have $|z_i|^2 e_i = z_i^* z_i e_i = u^* u e_i = e_i$, and thus $|z_i|^2 = 1$.

Moreover, H. Weyl developed the following theorem, which describes the density of the N eigenvalues of a randomly sampled matrix from $\mathscr{U}(N)$.

Theorem 1.1. *Let u be a uniformly random element of $\mathscr{U}(N)$. Then its eigenvalues are distributed with density*

$$\rho(z_1, \ldots, z_N) = \prod_{i<j} |z_j - z_i|^2.$$

In parallel, people derive random matrix from multi-dimensional stastics. Let $X = [x_{ij}]$ be $N \times T$ matrix of real data and define sample covariance matrix as $M = XX^*$ (we call M in that way because $\frac{1}{T}M$ is an estimator for the true covariance of N dim vector represented by one column). The main results are from Hsu, who proved something more general:

Theorem 1.2. *Suppose $x_{ij} \overset{iid}{\sim} \mathscr{N}(0,1)$, $T \geq N$. The density of N real eigenvalues of M is*

$$\rho(\lambda_1, \ldots, \lambda_N) = \prod_{i<j} |\lambda_j - \lambda_i| \prod_{i=1}^{N} \left(\lambda_i^{\frac{T-N-1}{2}} e^{-\frac{\lambda_i}{2}} \right).$$

As a conclusion, the first era focuses on explicit N computations.

Next, we arrive at the **1950s**, a period when the eigenvalues of large random matrices began to serve as a universal model for various point processes. This work contributed to Eugene Wigner receiving the Nobel Prize in Physics in 1963. The originates of this work are in quantum mechanics, where observations are closely linked to the eigenvalues of operators. One basic idea here is to approximate operators in infinite dimension spaces by large dimensional random matrices.

Consider the simplest matrix GOE (Gaussian Orthogonal Ensemble), which is defined as $M = \frac{1}{2}(X + X^*)$ where X is $N \times N$ with entries $\overset{\text{iid}}{\sim} \mathscr{N}(0,1)$. We also consider GUE (Guassian Unitary Ensemble) and GSE (Gaussian Symplectic Ensemble) in this course.

The main results of this era is Gaudin-Mehta Distribution, which involved spacing between two neighboured eigenvalues $\lambda_i - \lambda_{i+1}$. The GM distribution is related in a very range region, including neutron resonance spectroscopy, Dirichlet Boundary Value Problem, the imaginary part of Riemann zeta function and even bus interval distribution!

In the last 20 years, although many proofs of universal approximation of GM distribution have been provided, the underlying conceptual reasons remain not fully understood.

The third wave began in the **last 25 years** and continues to the present day, which focuses on the largest and smallest eigenvalues. The main results of this wave include the Tracy-Widom distributions, which have become widely recognized and influential.

Definition 1.3 (Tracy-Widom distribution)**.** *TW_1 and TW_2 are limits (after proper centering and rescaling) for the laws of largest eigenvalues in GOE and GUE, respectively.*

They are also widespread in 1) random matrix theory and 2) combinatorial probability.

Consider an intuitive example, where we take a permutation of numbers: $1, 2, ..., n \to i_1, i_2, ..., i_n$. Denote $l(\sigma)$ as the length of longest increasing subsequence, and we have the theorem

Theorem 1.4. *Baik–Deift–Johansson theorem Let l_n be the length of LIS in uniformly random permutation of $\{1, 2, ..., n\}$, then*

$$n^{-1/6}(l_n - 2\sqrt{n}) \xrightarrow{d} TW_2.$$

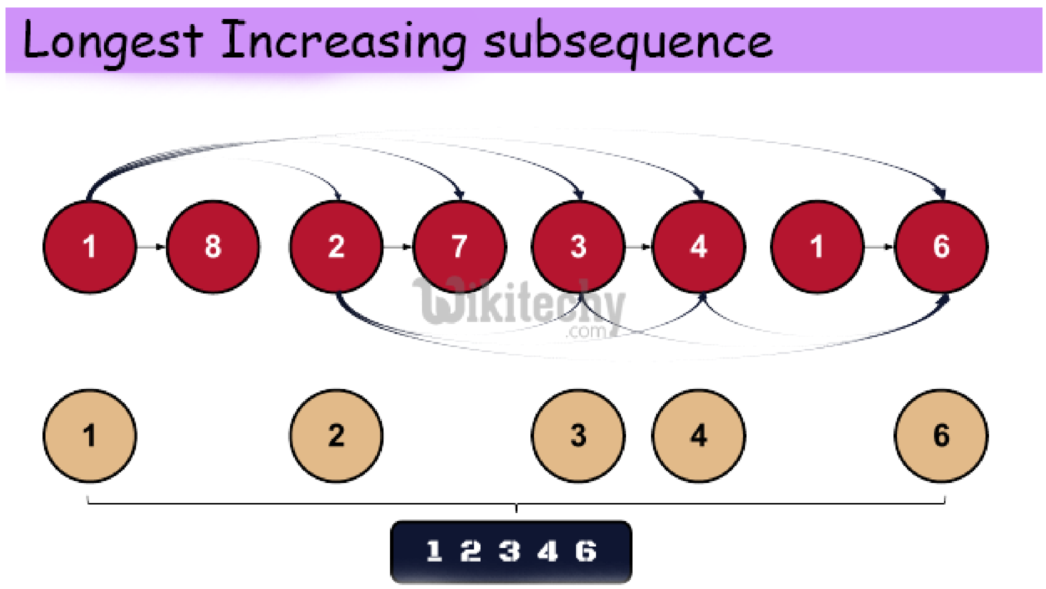


Figure 1: Largest increasing subsequence

Another application is the KPZ universality class for interface growth models, which played a role in the awarding of the 2021 Nobel Prize in Physics. Theoretical tools, such as replicas, are used to demonstrate that the large-time fluctuations of growing interfaces converge to the Tracy-Widom distribution TW_2, which depends on geometry of the system.

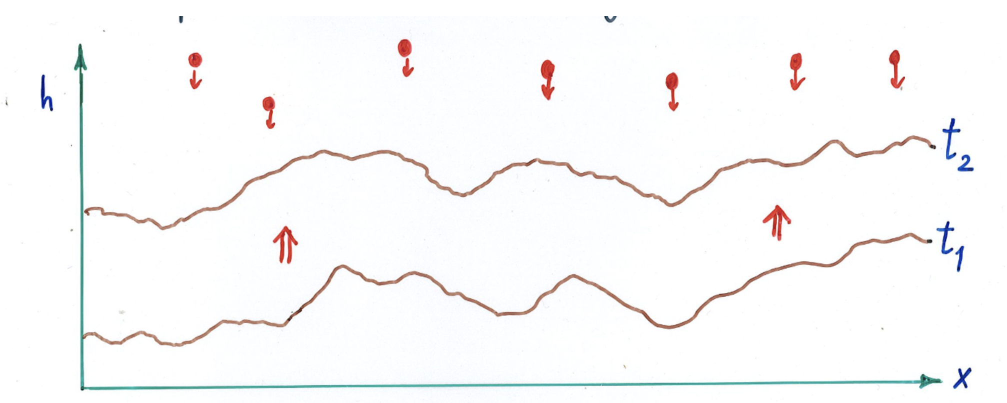


Figure 2: KPZ growth model

In recent years, with the advancement of artificial intelligence and machine learning, two notable applications have emerged. The first one is "Signal + Noise problem", which is formulated as:

$$C = A + B,$$

where C is the observed matrix, A is a deterministic matrix which is always low rank and B is a random matrix, e.g. GOE or GUE.

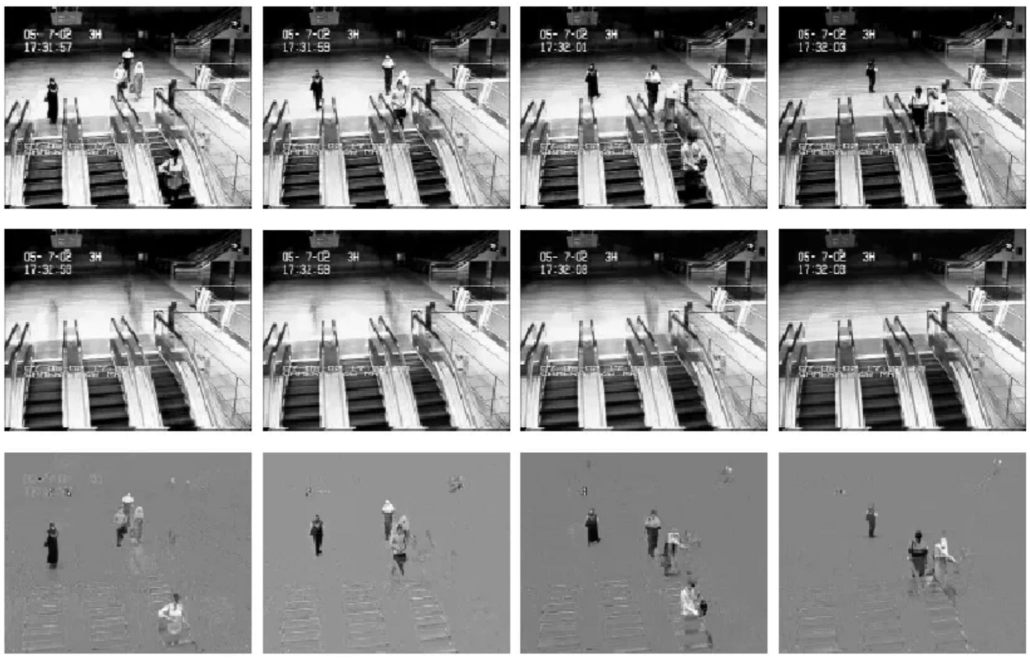


Figure 3: Background Subtraction from an Image

Determining what A is by examining C is a fundamental problem in sparse recovery, and the largest eigenvalue of C plays a significant role in this process.

Another example involves solving a system of linear equations:

$$AX = Y,$$

where A is an $N \times T$ matrix, X is an unknown $T \times 1$ vector and Y is a known $N \times 1$ vector. An important question is how sensitive is the solution to small perturbations of Y, which is related to the condition number $\kappa = \frac{\sigma_{\max}}{\sigma_{\min}}$.

Moreover, in these years large language model has developed a lot and they have many matrix multiplications inside. A natural question is: can we analysis them using random matrices?

In this class, we will give an overview of main types of random matrix behaviors and some tools to figure them out.

1.2 First computation: density of eigenvalues in classical ensembles of random matrices (GOE, GUE, GSE, etc)

Let X be $N \times N$ matrix with i.i.d. entries:

$$\begin{cases} a)\mathcal{N}(0,1) \\ b)\mathcal{N}(0,1) + i\mathcal{N}(0,1) \\ c)\mathcal{N}(0,1) + i\mathcal{N}(0,1) + j\mathcal{N}(0,1) + k\mathcal{N}(0,1) \end{cases}$$

Set $M = \frac{1}{2}(X + X^*)$, and we have the main theorem in today's lecture.

Theorem 1.5. *Let the eigenvalues of M be $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N$, and they have density*

$$\frac{1}{Z} \prod_{i<j} |\lambda_j - \lambda_i|^\beta \cdot \prod_{i=1}^{N} \exp(-\frac{\lambda_i^2}{2}) \,\mathrm{d}\lambda_1 ... \,\mathrm{d}\lambda_N.$$

where $\beta = 1, 2, 4$ for a), b), c) and

$$Z = \frac{(2\pi)^{N/2}}{N!} \prod_{j=0}^{N} \frac{\Gamma(1 + (j+1)\beta/2)}{\Gamma(1 + \beta/2)}$$

is the partition function (normalization factor).

Proof. We prove $\beta = 1$ here and leave $\beta = 2$ to the homework.
<u>**Step 1**</u> Claim density of $M = \frac{1}{2}(X + X^*) \sim \exp(-\frac{1}{2}\operatorname{tr}(M^2))$.
Indeed,

$$\operatorname{tr}(M^2) = \sum_{i,j} m_{ij}^2 = \sum_{i=1}^{N} \underbrace{x_{ii}^2}_{\mathscr{N}(0,1)} + \frac{1}{2} \sum_{i<j} \underbrace{(x_{ij} + x_{ji})^2}_{\mathscr{N}(0,2)}$$

implies that the density of $M \propto \exp(-\frac{1}{2}\operatorname{tr}(M^2))$

<u>**Step 2**</u> We can calculate that $\exp(-\frac{1}{2}\operatorname{tr}(M^2)) = \prod_{i=1}^{N} \exp(-\frac{\lambda_i^2}{2})$.
We derive it immediately by diagonalizing the the matrix M.

<u>**Step 3**</u> Each symmetric matrix is determined by its eigenvalues and eigenvectors, i.e. there exists an almost bijection π

$$\pi : \quad \underbrace{\mathscr{W}_N}_{\lambda_1<\lambda_2<...<\lambda_N} \quad \times \quad \underbrace{O(N)}_{\text{Orthogonal Bases}} \quad \to \quad \underbrace{\mathscr{H}_N}_{\text{Symmetric Matrix}}$$

We say "almost" because indeed the map is not injective: we can multiply the eigenvalue by ± 1. In other words, if the eigenvalues are unique, $\pi^{-1}(M)$ has exactly 2^N elements.

Now we give the key proposition of the proof.

Proposition 1.6. *Consider the map $\pi : (\Lambda, O) \mapsto B$, where $\pi((\Lambda, O)) = O\Lambda O^*$. Then the Jocobian of the map is $\prod_{i<j} |\lambda_j - \lambda_i|$.*

Remark 1.7. *It is an important technique to transform an intractable density calculation into a tractable one and a Jocobian.*

Proof. It is sufficient to calculate by taking $O = \mathrm{Id}$, as when we transform O to $A \cdot O$ (where A is orthogonal), the **uniform measure** on $O(N)$ and the Lebesgue measure on $\mathscr{H}(N)$ remain unchanged.

Then we take $O = \exp(B) = \mathrm{Id} + B + ...$ and we can deduce that $B + B^* = 0$ as $OO^* = \mathrm{Id}$. Then the map π can be written as

$$((\lambda_1,...,\lambda_N),\exp(B)) \mapsto \exp(B)\begin{pmatrix}\lambda_1 & 0 & 0 & \cdots & 0\\ 0 & \lambda_2 & 0 & \cdots & 0\\ 0 & 0 & \lambda_3 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & \lambda_n\end{pmatrix}\exp(-B)$$

$$\mapsto (\mathrm{Id}+B)\begin{pmatrix}\lambda_1 & 0 & 0 & \cdots & 0\\ 0 & \lambda_2 & 0 & \cdots & 0\\ 0 & 0 & \lambda_3 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & \lambda_n\end{pmatrix}(Id-B)+o(B)$$

$$\mapsto \begin{pmatrix}\lambda_1 & 0 & 0 & \cdots & 0\\ 0 & \lambda_2 & 0 & \cdots & 0\\ 0 & 0 & \lambda_3 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

$$+\begin{pmatrix}0 & b_{12}(\lambda_2-\lambda_1) & b_{13}(\lambda_3-\lambda_1) & \cdots & b_{1n}(\lambda_n-\lambda_1)\\ b_{21}(\lambda_1-\lambda_2) & 0 & b_{23}(\lambda_3-\lambda_2) & \cdots & b_{2n}(\lambda_n-\lambda_2)\\ b_{31}(\lambda_1-\lambda_3) & b_{32}(\lambda_2-\lambda_3) & 0 & \cdots & b_{3n}(\lambda_n-\lambda_3)\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ b_{n1}(\lambda_1-\lambda_n) & b_{n2}(\lambda_2-\lambda_n) & b_{n3}(\lambda_3-\lambda_n) & \cdots & 0\end{pmatrix}+o(B)$$

As B has only $\frac{n(n-1)}{2}$ free parameter, so the Jacobian of the map is $\prod_{i<j}|\lambda_j-\lambda_i|$. □

Step 4 We now calculate Z:

$$Z=\int_{\mathbb{R}^N}\prod_{i<j}|\lambda_j-\lambda_i|^\beta\cdot\prod_{i=1}^N\exp(-\frac{\lambda_i^2}{2})\,\mathrm{d}\lambda_1...\,\mathrm{d}\lambda_N=\frac{(2\pi)^{N/2}}{N!}\prod_{j=0}^N\frac{\Gamma(1+(j+1)\beta/2)}{\Gamma(1+\beta/2)}$$

□

Now we give some other ensembles. The first one is density of rectangular matrix.

Definition 1.8 (Laguerre or Wishart ensemble). *Take $N\times M$ matrix X with $N<M$ and assume it has singular value decomposition $X=U\,\mathrm{diag}\{s_1,...,s_N\}V^T$*

Theorem 1.9. *Let X be $N\times M$ with i.i.d. real/complex/quarternion Gaussian matrix elements, then the density of $\lambda_i=s_i^2$*

$$\rho(\lambda_1,...,\lambda_N)\propto\prod_{i<j}|\lambda_j-\lambda_i|^\beta\cdot\prod_{i=1}^N\exp(-\frac{\lambda_i}{2})\cdot\lambda_i^{\frac{\beta}{2}(M-N+1)-1}\,\mathrm{d}\lambda_1...\,\mathrm{d}\lambda_N.$$

Remark 1.10. *This density is also known as multivariate Γ distribution.*

Another ensemble involves subspace projection. Imagine that we have two rectangular arrays $X:N\times T$ and $Y:K\times T$, where $N\le K\le T$. P_X is a projection on N-dimensional subspace in T-dimensional space spanned by N rows of X and P_Y is defined similarly. Squared canonical correlations c_i are N non-zero eigenvalues P_XP_Y ($c_i^2=\cos^2\theta_i$).

Theorem 1.11. *Assume X and Y are independent with i.i.d. Gaussian real/complex/quarternion elements. The eigenvalues of $P_X P_Y$ have density*

$$\rho(\lambda_1, ..., \lambda_N) \propto \prod_{i<j} |\lambda_j - \lambda_i|^\beta \cdot \prod_{i=1}^{N} \lambda_i^{\frac{\beta}{2}(K-N+1)-1} (1-\lambda_i)^{\frac{\beta}{2}(T-N-K+1)-1} \, \mathrm{d}\lambda_1 ... \, \mathrm{d}\lambda_N,$$

where we assume $N \leq K \leq T$ and $N + K \leq T$, $0 \leq \lambda_i \leq 1$.

Remark 1.12. *This density is also known as multivariate β distribution.*

The general form of the eigenvalues in such ensembles is

$$\underbrace{\prod_{i<j} |\lambda_j - \lambda_i|^\beta}_{\text{logorithm pairwise interaction}} \quad \underbrace{\prod_{i=1}^{N} V(\lambda_i)}_{\text{potential}}.$$

whose name comes from log-gas or β-ensemble.

2 Jan 28th

Happy Chinese New Year!

2.1 Law of Large Numbers through tridiagonal matrices

Today our goal is to prove the law of large numbers for any $\beta > 0$, which is known as the semi-circle law. Our main tools are triangle matrix and moment method.

Theorem 2.1. *Consider the real symmetric tridiagonal matrix*

$$T_\beta = \begin{pmatrix} \mathcal{N}(0,1) & \frac{1}{\sqrt{2}}\chi_{\beta(n-1)} & \cdots & 0 \\ \frac{1}{\sqrt{2}}\chi_{\beta(n-1)} & \mathcal{N}(0,1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{N}(0,1) \end{pmatrix}$$

where the second diagonal entries are $\chi_{\beta(n-1)}, \chi_{\beta(n-2)}, \dots \chi_\beta$. $\chi_k = \sqrt{\chi_k^2} = \sqrt{\sum_{i=1}^k \mathcal{N}(0,1)^2}$, *whose density is*

$$\frac{1}{2^{\frac{k}{2}-1}\Gamma(\frac{k}{2})} x^{k-1} e^{-\frac{x^2}{2}}.$$

Then the eigenvalues of T_β have the same distribution of $G\beta E$.

Proof. We present for $\beta = 1$. By linear algebra, we can choose an orthogonal matrix U_1, and transform the matrix M to $U_1 M U_1^\top$, whose the first row is

$$(x_{11}, \sqrt{\sum_{k=2}^n x_{ik}^2}, 0, ..., 0).$$

Note that the eigenvalues do not change after the transformation.

We can inductively do the same operation on the rest sub-matrix. $\square$

Corollary 2.2. *For $\beta = 1, 2, 4$, the law of eigenvalues* $\sim \prod_{i<j}(x_i - x_j)^\beta \prod_{i=1}^N \exp(-\frac{x_i^2}{2})$.

Theorem 2.3. *In fact, the corollary is true for any $\beta > 0$.*

Now we begin to prove the semi-circle law. First we give the key result of trace calculation.

Theorem 2.4. *For each $k > 0, \beta > 0$, we have*

$$\frac{1}{N^{\frac{k}{2}+1}} \operatorname{tr}(T_\beta^k) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{i=1}^N X_i^k \xrightarrow{N\to\infty, \text{in probability}} \begin{cases} 0 & , \quad k \text{ is odd;} \\ (\frac{\beta}{2})^k \operatorname{Cat}_{\frac{k}{2}} & , \quad k \text{ is even.} \end{cases}$$

where Cat_k *is the k-the catalan number defined as*

$$\operatorname{Cat}_k = \frac{1}{k+1}\binom{2k}{k}.$$

Proof. By SLLN, we have

$$\lim_{N\to\infty} \frac{\chi_{\beta(N-1)}}{\sqrt{2N}} = \sqrt{\frac{\beta}{2}},$$

$$\lim_{N\to\infty} \frac{\chi_{\beta(N-\alpha N)}}{\sqrt{2N}} = \sqrt{\frac{\beta(1-\alpha)}{2}}.$$

So by the definition of the trace, we have

$$\operatorname{tr}\left(\frac{T_\beta}{\sqrt{N}}\right)^k = \sum_{m=1}^{N} \sum_{\substack{\text{k-step paths} \\ \text{from } m \text{ to } m}} \left(\text{some } \frac{\chi_{\beta(N-i)}}{\sqrt{2N}}\right)\left(\text{some } \frac{\mathcal{N}(0,1)}{\sqrt{N}}\right)$$

$$= \sum_{m=1}^{N} (\#\text{k-step paths from } m \text{ to } m)\sqrt{\frac{\beta}{2}\left(\frac{N-m}{N}\right)^k} + o(N).$$

The second equation follows from that if there exists "some $\frac{\mathcal{N}(0,1)}{\sqrt{N}}$", the term vanishes. When $2 \nmid k$, #k-step paths from m to m is zero. When $2 \mid k$, #k-step paths from m to $m = \binom{k}{\frac{k}{2}}$. Approximate the summation by integration and we have

$$\frac{1}{N}\operatorname{tr}\left(\frac{T_\beta}{\sqrt{N}}\right)^k \to \binom{k}{\frac{k}{2}}\left(\frac{\beta}{2}\right)^{\frac{k}{2}}\int_0^1 x^{\frac{k}{2}}\,\mathrm{d}x = \operatorname{Cat}_{\frac{k}{2}}\left(\frac{\beta}{2}\right)^{\frac{k}{2}}$$

□

Now we introduce the semi-circle distribution.

Definition 2.5 (Wiegner semi-circle distribution)**.** *The density of the distribution is*

$$\mu(x) = \frac{1}{2\pi}\sqrt{4-x^2}.$$

We use moment method to recover the proof and give the definition.

Theorem 2.6. *Let m_k are moments of T_β and we have*

$$m_u = \begin{cases} \frac{1}{\frac{u}{2}+1}\binom{u}{\frac{u}{2}} & , \quad u \text{ is even.} \end{cases}$$

Theorem 2.7. *m_k are moments of semi-circle law and*

$$m_k = \int_{-2}^{2} \mu(x) x^k \,\mathrm{d}x.$$

Proof. (Method I) We just calculate

$$m_k = \int \frac{1}{\sqrt{2\pi}}\sqrt{4-x^2}x^k\,\mathrm{d}x.$$

Let $x = 2\cos\theta$, and by inductive calculation, we derive the results. □

However, we are not satisfied for the calculation is not so intuitive. We then introduce another proof.

Proof. (Method II) Introduce generating function:

$$G(z) = \sum_{k=0}^{\infty} m_k z^{-k-1},$$

$m_0 = 1$. By classical results, we have

$$\sum_{n=0}^{\infty} \operatorname{Cat}_n x^n = \frac{1-\sqrt{1-4x}}{2x} =: C(x).$$

Now we can derive the expression of $G(z)$ using $C(x)$.

$$G(z) = \sum_{k=0}^{\infty} m_k z^{-k-1} = \sum_{l=0}^{\infty} \mathrm{Cat}_l \, z^{-2l-1} = \frac{z - \sqrt{z^2 - 4}}{2}.$$

Next we introduce two propositions.

Proposition 2.8.

$$G(z) = \int \frac{\mu(x)}{z - x} \, \mathrm{d}x.$$

Proof. By Taylor expansion, we have

$$\begin{aligned}
\int \frac{\mu(x)}{z - x} \, \mathrm{d}x &= \frac{1}{z} \int \frac{\mu(x)}{1 - \frac{x}{z}} \, \mathrm{d}x \\
&= \frac{1}{z} \int \mu(x) \sum_{k=0}^{\infty} \left(\frac{x}{z} \right)^k \\
&= \sum_{k=0}^{\infty} m_k z^{-k-1} \\
&= G(z)
\end{aligned}$$

□

Proposition 2.9.

$$\mu(x_0) = -\frac{1}{\pi} \lim_{y_0 \to 0^+} \Im(G(x_0 + iy_0)).$$

Remark 2.10. *We can just add a perturbation onto the imagine axis and obtain the information of the point.*

Proof. By proposition 2.8, we have

$$\begin{aligned}
-\frac{1}{\pi} \Im(G(x_0 + iy_0)) &= -\frac{1}{\pi} \int \Im \frac{1}{x_0 + iy_0 - x} \mu(x) \, \mathrm{d}x && (1) \\
&= -\frac{1}{\pi} \int \Im \frac{x_0 - x - iy_0}{(x - x_0)^2 + y_0^2} \mu(x) \, \mathrm{d}x && (2) \\
&= \int \frac{1}{\pi} \frac{y_0}{(x - x_0)^2 + y_0^2} \mu(x) \, \mathrm{d}x. && (3)
\end{aligned}$$

Notice that $\frac{1}{\pi} \frac{y_0}{(x - x_0)^2 + y_0^2}$ is a "good" kernel, and thus as $y_0 \to 0$

$$\int \frac{1}{\pi} \frac{y_0}{(x - x_0)^2 + y_0^2} \mu(x) \, \mathrm{d}x \to \mu(x_0).$$

□

Return to the Theorem 2.7, by proposition 2.9

$$\mu(x_0) = -\frac{1}{\pi} \lim_{y \to 0^+} \Im \left(\frac{1}{2} \left[(x + iy) - \sqrt{(x + iy)^2 - 4)} \right] \right) = \begin{cases} 0 & , \quad |x| > 2; \\ \frac{\sqrt{4 - x^2}}{2\pi} & , \quad |x| \leq 2. \end{cases}$$

□

Now we are prepared to formally state and prove the semi-circle law.

Theorem 2.11. *Let $\lambda_1 < ... < \lambda_N$ be eigenvalues of $G\beta E$. Set $x_i = \lambda_i\sqrt{\frac{2}{\beta N}}$ and let μ_N be their empirical measure: $\mu_N = \frac{1}{\sqrt{N}}\sum_{i=1}^N \delta_{x_i}$. Then*

$$\lim_{N\to\infty} \mu_N = \frac{1}{2\pi}\sqrt{4-x^2}\mathbb{1}_{|x|\le 2} = \mu_{circle}(x)$$

weakly in probability, which means $\forall$ bounded continuous function $f(x)$

$$\lim_{N\to\infty} \int_{\mathbb{R}} f(x)\mu_N(x) = \int_{\mathbb{R}} f(x)\mu(x)\,\mathrm{d}x.$$

Proof. **Step 1** The results hold for $f(x) = x^k$.

Step 2 The results hold for any polynomials.

Step 3 Take $L > 2$, the results hold for $f(x) = \mathbb{1}_{|x|>L}x^k$. We have

$$\begin{aligned}
\left|\int f(x)\mu_N(\mathrm{d}x)\right| &= \left|\int x^k\mu_N(\mathrm{d}x)\right| \\
&\le L^{-2m}\int_{|x|\ge L} x^{2k+2m}\mu_N(\mathrm{d}x) \\
&\to L^{-2m}\int x^{2k+2m}\mu(\mathrm{d}x) \\
&\le L^{-2m}2^{2k+2m} = 2^{2k}\left(\frac{2}{L}\right)^m \to 0.
\end{aligned}$$

Step 4 By step 3, we can restrict the support set of f on a compact set, i.e. $[-4,4]$. Apply the Weierstrass theorem and we get the proof. $\square$

At last, we give three generalizations of the semi-circle law. Next theorem tells us the gaussian assumption of the semi-circle law is not neccessary.

Theorem 2.12. *Let $z_{ij}, i < j$ be i.i.d. random variables with finite moments. $\mathbb{E}z_{ij} = 0$ and $\mathbb{E}z_{ij}^2 = \frac{1}{2}$. Let Y_i be i.i.d. with finite moments. Then the semi-circle law still holds for the matrix with entries Y_i and z_{ij}.*

Theorem 2.13 (Marchenko–Pastur Law)**.** *Consider Laguerre β-ensemble $N \le M$, and its density is*

$$\prod_{i<j}(\lambda_j-\lambda_i)^\beta\prod_{i=1}^N \lambda_i^{\frac{\beta}{2}}e^{-\frac{\lambda_i}{2}}.$$

Let $x_i = \frac{\lambda_i}{\beta N}$, then when $M, N \to \infty$ with thermodynamics condition $\frac{M}{N} \to C \ge 1$. Then $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i} \xrightarrow{N\to\infty}$ distribution on $\mathbb{R}_{>0}$ with density $\frac{(\lambda_+-x)(x-\lambda_-)}{x}\mathbb{1}_{\lambda_-<x<\lambda_+}$, where $\lambda_\pm = \lambda_\pm(C)$.

Theorem 2.14 (Wachter Law)**.** *Consider Jacobi β-ensemble with $N \le K \le T$ and density function $\prod_{i<j}\prod_{i=1}^N \lambda_i^{\frac{\beta}{2}(K-N+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(T-N-K+1)-1}$. Assume $N, K, T \to \infty$ with thermodynamics condition $\frac{K}{T} \to C_1$ amd $\frac{N}{T} \to C_2 \ll C_1$, $C_1 + C_2 < 1$. Then $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i} \to$ distribution on $\mathbb{R}_{>0}$ with density $\frac{\sqrt{(\lambda_+-x)(x-\lambda_-)}}{x(1-x)}\mathbb{1}_{\lambda_-<x<\lambda_+}$, where $\lambda_\pm = \lambda_\pm(C_1, C_2)$.*

3 Feb 4th

3.1 Point process

In the last lecture, we discuss macroscopic behavior of eigenvalues, and today we are going to focus on the microscopic behavior of eigenvalues of local limits. We view eigenvalues as N points on the label line, and the limits here can be split into two classes

- Bulk limit: limit somewhere in the middle, so that practicles extend in both directions from a reference point.
- Edge limit: focus on the largest/smallest eigenvalue, so that the limit is a semiinfinite configuration: $x_1 < x_2 < ...$ or $x_1 > x_2 > x_3 > ....$

The question is how to describe random infinite point configurations. And the difficulty lies in that there is no good underlying configurations.

Assume $\mathscr{X}$ is a statespace with topology structure, and we can give the definition of the configuration.

Definition 3.1 (Point configuration and window)**.** *Define a locally finite subset of $\mathscr{X}$ as point configuration and Conf($\mathscr{X}$) be the set of all the point configuration. Define window as compact subset of X. The window induce a measure*

$$N_A(X) = \#\ \textit{points of } A \textit{ in } X.$$

We state that the configuration has intrinsic Borel structure, which means Conf(x) is the minimal $\sigma-$ algebra which makes all N_A measurable functions. We can ask questions about probability which can be inferred from random variables N_A. For instance, $\mathbb{P}$ (No particles in [a, b]).

Example 3.2. *(Bernoulli point process) Let $X = \mathbb{Z}$ and for each $a \in \mathbb{Z}$, we place a particle there with probabilities $0 < p < 1$. Take $A \subset \mathbb{Z}$ containing n elements, then*

$$\mathbb{P}(N_A = k) = p^k(1-p)^{n-k}\binom{n}{k}, \quad \textit{for } 0 \leq k \leq n.$$

N_A and N_B are independent if $A \bigcap B = \emptyset$.

Example 3.3. *(Poisson point process) Let $X = \mathbb{R}$ and $\lambda \in \mathbb{R}$, for $A \subset \mathbb{R}$, we have $N_A \sim \mathrm{Poi}(\lambda|A|)$, which means*

$$\mathbb{P}(N_A = k) = e^{-\lambda|A|}\frac{(\lambda|A|)^k}{k!}, \quad \textit{for } k = 0, 1, 2, ...$$

For disjoint $A_1, ..., A_m$, we have $N_{A_1}, ..., N_{A_m}$ are independent.

We can derive Poisson point process by taking $p \to 0$ in Bernoulli point process and let $\mathbb{Z} \to m\mathbb{Z}$. Here we give an intuitive derivation (choose λ to be the probability)

$$\begin{aligned}
\mathbb{P}(N_A = k) &= \binom{(b-a)m}{k}\left(\frac{\lambda}{m}\right)^k\left(1-\frac{\lambda}{m}\right)^{(b-a)m-k} \\
&= \frac{1}{k!}e^{-k}\frac{((b-a)m)^{(b-a)m}}{((b-a)m-k)^{(b-a)m-k}}\left(\frac{\lambda}{m}\right)^k\left(1-\frac{\lambda}{m}\right)^{(b-a)m-k} \\
&= \frac{e^{-k}}{k!}(\lambda(b-a))^k\left(\left(1-\frac{\lambda}{m}\right)\left(1+\frac{k}{(b-a)m-k}\right)\right)^{(b-a)m-k} \\
&= e^{-k}\frac{(\lambda(b-a))^k}{k!}e^{-(b-a)\lambda}e^k \\
&= e^{-(b-a)\lambda}\frac{(\lambda(b-a))^k}{k!}
\end{aligned}$$

By the construction of Lebesgue measure, we derive the results for arbitrary A.

Now we introduce a tool for describing point processes – "correlation function".

Definition 3.4 (Correlation function)**.** *Assume $\mathscr{X}$ is descrete and X is a point process of $\mathscr{X}$. The n-th correlation function ρ_n is a function of n distinct variables.*

$$\rho_n(x_1, ..., x_n) = \mathbb{P}(x_1 \in X, ..., x_n \in X).$$

Example 3.5 (Bernoulli)**.**

$$\rho_n(x_1, ..., x_n) = p^n.$$

Proposition 3.6. *For discrete X, the sequence of functions $\rho_1, \rho_2, ...$ uniquely describe the law of X.*

Proof. The law of X is the joint law of all N_A, and also equals to all probabilities of the sort

$$\mathbb{P}\begin{pmatrix} a_1 \in X, & a_2 \in X, & ..., & a_n \in X \\ b_1 \notin X, & b_2 \notin X, & ..., & b_n \notin X \end{pmatrix}.$$

By inclusive-exclusive formula, we can get the results. □

Definition 3.7. *The n-th correlation measure ρ_n is a symmetric measure on $\mathscr{X}^n$ such that for any compactly supported bounded measurable $f : \mathscr{X}^n \to \mathbb{R}$, we have*

$$\int_{\mathbb{R}^n} f \rho_n(\mathrm{d}x_1, ..., \mathrm{d}x_n) = \mathbb{E}_x \left[\sum_{x_1,...,x_n \ distinct} f(x_{i_1}, ..., x_{i_k}) \right].$$

Proposition 3.8. *If $\mathscr{X}$ is discrete, then this is the same definition as before.*

Proof. Both sides are linear in f, so we can take

$$f(x_1, ..., x_n) = \mathbb{1}_{x_1=a_1,...,x_n=a_n}.$$

and the theorem follows. □

Remark 3.9. *Often we take $\mathscr{X} = \mathbb{R}$ and μ is the Lebesgue measure so that $\rho_n(x_1, ..., x_n)\,\mathrm{d}x_1 ... \mathrm{d}x_n \approx \mathbb{P}$ (there are particles in $[x_1, x_1 + \mathrm{d}x_1] \bigcup ... \bigcup [x_n, x_n + \mathrm{d}x_n]$).*

In Homework 2, we can prove for poisson process of intensity λ, $\rho_n(x_1, ..., x_n)$ w.r.t Lebesgue measure.

Proposition 3.10. *For compact $A \subset \mathscr{X}$, we have*

$$\mathbb{E}\left[N_A(N_A - 1)...(N_A - n + 1)\right] = \int_{A^n} \rho_n(\mathrm{d}x_1, ..., \mathrm{d}x_n).$$

Proof. Take $f = \mathbb{1}_{A_1}(x)...\mathbb{1}_{A_n}(x)$ and we get the proof. □

Then we state a theorem without proof.

Theorem 3.11. *Under mild growth conditions, correlation measures exist and unique determine the law of the point process.*

Notice that in proposition 3.10, ρ_n are linked to moments of N_A, so we need conditions similar to the ones in "Moment problem for random variables". We give the following proposition.

Proposition 3.12. *Take a point process formed by N particles $X_1, ..., X_N$ with joint probability density $\mathbb{P}(\mathrm{d}x_1, ..., \mathrm{d}x_N)$ assumed to symmetric w.r.t. permutations of $x_1, ..., x_N$.*

$$\rho_n = \begin{cases} 0 & , \quad \textit{for } n > N; \\ \frac{N!}{(N-n)!} \int_{x_{n+1},...,x_N \in \mathscr{X}} \rho_N(\mathrm{d}x_1, ..., \mathrm{d}x_N) & , \quad \textit{for } n \leq N. \end{cases}$$

Proof. When $n > N$, we can't choose n distinct particles from $[N]$; and for $n \leq N$,

$$\mathbb{E}_X \left[\sum f(x_{i_1}, ..., x_{i_n}) \right] = \int_{\mathscr{X}^N} \sum_{x_{i_1},...,x_{i_n} \text{ distinct}} f(x_{i_1}, ..., x_{i_n}) \rho_N(\mathrm{d}x_1, ..., \mathrm{d}x_N).$$

By symmetry, we derive the proof. □

3.2 Correlation kernel

Now we introduce an important class.

Definition 3.13 (Determinantal)**.** *A point process X is determinantal if there $\exists$ $k(x, y)$ - correlation kernel on $\mathscr{X} \times \mathscr{X}$ such that correlation functions with respect to some reference measure μ have the form*

$$\rho_N(x_1, ..., x_N) = \det[k(x_i, x_j)]_{i,j=1}^N$$

Remark 3.14.

1. *The order of x_i does not matter.*
2. *Replacement $k(x, y)$ with $\frac{f(x)}{f(y)} k(x, y)$ leads to the same correlation functions.*
3. *Reduction of complexity: all ρ_n are encoded in a single function $k(x, y)$.*

Example 3.15 (Poisson process)**.** *For Poisson process $\rho_n = \lambda^n$ is determinantal with $k(x, y) = \lambda\delta(x = y)$*

Next, we introduce another important class.

Definition 3.16. *(Biorthogonal ensemble) Consider space $\mathscr{X}$ with reference measure μ. An N-point biorthogonal ensemble is a probability measure on $\{x_1, ..., x_N\} \subset \mathscr{X}$ of the form*

$$\rho_N(\mathrm{d}x_1, ..., \mathrm{d}x_N) = C_N \det[\phi_i(x_j)]_{i,j=1}^N \det[\psi_i(x_j)]_{i,j=1}^N \mu(\mathrm{d}x_1)...\mu(\mathrm{d}x_N).$$

for constant $C_N > 0$ and functions ϕ_i, ψ_j such that $\int_{\mathscr{X}} \phi_i(x)\psi_j(x) < \infty$ for all $i, j = 1, ..., N$.

Example 3.17. *Consider measure on $\mathbb{R}^N$ of density $\sim \prod_{i<j}(x_i - x_j)^2 \prod_{i=1}^N V(x_i)$. Then choose $\phi_i(x) =$ arbitrary degree $(i - 1)$ polynomials, and similar for ψ_j. Choose $\mu(\mathrm{d}x) = V(x)\,\mathrm{d}x$. Then, by Vandermonde's determinant, we have*

$$\det[\phi_i(x_j)]_{i,j=1}^N = \prod_{i<j}(\lambda_j - \lambda_i).$$

Remark 3.18. *The derivation of Vandermonde's determinant is as follows: The degree of the polynomials is $\frac{(N-1)N}{2}$, and $\lambda_i - \lambda_j$ is a factor of the determinant.*

Theorem 3.19. *Biorthogonal ensemble is a determinantal point process with*

$$K(x, y) = \sum_{i,j=1}^N \phi_i(x)\psi_j(y)[G^{-\top}]_{ij}$$

Remark 3.20. *If ϕ_i and ψ_j are biorthogonal, which means*

$$\int \phi_i \psi_j \mu(\mathrm{d}x) = d_i \delta(i = j),$$

then $G^{-\top}$ is easy to calculate.

Proof. <u>**Step 1**</u> We first calculate C_N:

$$\begin{aligned}
C_N^{-1} &= \int_{\mathscr{X}^N} \det[\phi_i(x_j)]_{i,j=1}^N \det[\psi_i(x_j)]_{i,j=1}^N \mu(\mathrm{d}x_1)...\mu(\mathrm{d}x_N) \\
&= \sum_{\sigma,\tau} (-1)^{\sigma\tau} \prod_{i=1}^N \phi_{\sigma_i}(x_i)\psi_{\tau_i}(x_i)\mu(\mathrm{d}x_i) \\
&= N! \det(G).
\end{aligned}$$

<u>**Step 2**</u> Notice that

$$\det G \det[\phi_i(x_j)]_{i,j=1}^N \det[\psi_i(x_j)]_{i,j=1}^N = 1$$

We have

$$\rho_N(x_1, ..., x_N) = \int_{\mathscr{X}^N} \det[\phi_i(x_j)]_{i,j=1}^N \det[\psi_i(x_j)]_{i,j=1}^N \mu(\mathrm{d}x_1)...\mu(\mathrm{d}x_N).$$

<u>**Step 3**</u> We can choose

$$K(x_i, x_j) = (\phi(x_i))(G^{-\top})(\psi(x_j))^\top.$$

□

Corollary 3.21 (GUE). *For $N-$particle GUE eigenvalues of density $\sim \prod_{i<j}(x_i - x_j)^2 \prod_{i=1}^N \exp(-\frac{x_i^2}{2})$, it is a determinantal point process with correlation kernel*

$$k^N(x, y) = \sum_{k=0}^N \frac{H_k(x)H_k(y)}{\langle H_k, H_k \rangle} \cdot e^{-\frac{y^2}{2}}$$

w.r.t Lebesgue measure, where $H_n(x)$ is Hermite polynomial.

Remark 3.22. *It seems the $k^N * (x, y)$ is not symmetric, but we can multiply an $\frac{f(x)}{f(y)}$ and transform $e^{-\frac{y^2}{2}}$ into $e^{-\frac{x^2+y^2}{4}}$.*

3.3 Properties of Hermite polynomial

In this section, we take a review of Hermite polynomials.

Definition 3.23 (Hermite polynomial). *$H_n(x)$ is defined as*

- $H_n(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$ *for* $n = 0, 1, 2, ...$
- $\int_{\mathbb{R}} H_n(x)H_m(x)e^{-\frac{x^2}{2}} = \sqrt{2\pi} n! \delta(n = m)$.

Proposition 3.24. *We have*

$$\begin{aligned}
H_n(x) &= (-1)^n e^{\frac{x^2}{2}} \frac{\partial^n}{\partial x^n} e^{-\frac{x^2}{2}} \\
&= \frac{n!}{2\pi i}(-1)^n e^{\frac{x^2}{2}} \oint_x \frac{e^{-\frac{z^2}{2}}}{(z - x)^{n+1}} \,\mathrm{d}z \\
&= \frac{n!}{2\pi i} \oint_O \frac{e^{-\frac{z^2}{2} + zx}}{z^{n+1}} \,\mathrm{d}z
\end{aligned}$$

Proof. The first line follows from integration by parts. The second line is residue formula and the third kine is Integration by substitution. □

Corollary 3.25.

$$H_{n+1}(x) - xH_n(x) + nH_{n-1}(x) = 0.$$

Proof.

$$\text{LHS} = \frac{n!}{2\pi i}\oint_O \frac{e^{-\frac{z^2}{2}+zx}}{z^{n+1}}[\frac{n+1}{z} - x + z]\,\mathrm{d}z = \frac{n!}{2\pi i}\oint_O \mathrm{d}\left(\frac{e^{-\frac{z^2}{2}+zx}}{z^{n+1}}\right) = 0.$$

□

Proposition 3.26. *The sum in det of $k^N(x,y)$ telescopes to*

$$\sum_{k=0}^{N-1}\frac{H_k(x)H_k(y)}{\langle H_k, H_k\rangle} = \frac{1}{\langle H_{N-1}, H_{N-1}\rangle}\frac{H_N(x)H_{N-1}(y) - H_{N-1}(x)H_N(y)}{x-y}.$$

Proof. Induction. Use Corollary 3.25. □

4 Feb 11th

4.1 Steepest Descent

Today we use the Hermite polynomials and relative kernel for asymptotic analysis. The main result is as follows.

Theorem 4.1. *Take $s \in (-2,2)$, then*

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} K_N^{GUE}(s\sqrt{N} + \frac{x}{\sqrt{N}}, s\sqrt{N} + \frac{y}{\sqrt{N}}) = K_{\sin}^{\rho}(x,y) e^{\frac{s}{2}(x-y)},$$

where

$$K_{\sin}^{\rho}(x,y) = \begin{cases} \frac{\sin[\pi\rho(x-y)]}{\pi(x-y)} & , \ x \neq y; \\ \rho & , \ x = y; \end{cases}$$

and

$$\rho = \frac{1}{\sqrt{2\pi}} \sqrt{4 - s^2}.$$

Remark 4.2. *If we take $x = y$, we have*

$$K(x,x) \to \rho,$$

which recovers the semi-circle law.

Remark 4.3. *Eigenvalues live on $[-2\sqrt{N}, 2\sqrt{N}]$, and we expect spacing of order $\frac{1}{\sqrt{N}}$ (the behaviour of eigenvalues align with $\frac{x}{\sqrt{N}}$). We call the limit process at $\beta = 2$ sine process of intensity ρ.*

Proof. (for $s = 0$) First, we recall some important properties of Hermite polynomials H_n:

- $H_n(x) = x^n + a_{n-1}x^{n-1} + ...$
- $\int_{\mathbb{R}} H_n(x) H_m(x) = \sqrt{2\pi}\ n!\ \delta_{n=m}$
- (By induction) For even n, $H_n(x) = n!\ (-1)^{\frac{n}{2}}\ \sum_{m=0}^{\frac{n}{2}} \frac{(-1)^m}{(2m)!(\frac{n}{2}-m)! 2^{\frac{n}{2}-m}} x^{2m}$.
- (By induction) For odd n, $H_n(x) = n!\ (-1)^{\frac{n-1}{2}}\ \sum_{m=0}^{\frac{n-1}{2}} \frac{(-1)^m}{(2m+1)!(\frac{n-1}{2}-m)! 2^{\frac{n-1}{2}-m}} x^{2m+1}$.

For even n,

$$H_n(\frac{x}{\sqrt{n}}) = \frac{n!\ 2^{-\frac{n}{2}} (-1)^{-\frac{n}{2}}}{\left(\frac{n}{2}\right)!} \sum_{m=0}^{\frac{n}{2}} \frac{\left(\frac{n}{2}\right) ... \left(\frac{n}{2} - m + 1\right)}{\left(\frac{n}{2}\right)^m} \frac{x^{2m}}{(2m)!} (-1)^m \to \frac{n!\ 2^{-\frac{n}{2}} (-1)^{-\frac{n}{2}}}{\left(\frac{n}{2}\right)!} \cos x.$$

And for odd n,

$$H_n(\frac{x}{\sqrt{n}}) = \frac{n!\ 2^{-\frac{n-1}{2}} (-1)^{-\frac{n-1}{2}}}{\left(\frac{n-1}{2}\right)! \sqrt{n}} \sum_{m=0}^{\frac{n-1}{2}} \frac{\left(\frac{n-1}{2}\right) ... \left(\frac{n-1}{2} - m + 1\right)}{\left(\frac{n}{2}\right)^m} \frac{x^{2m+1}}{(2m+1)!} (-1)^m \to \frac{n!\ 2^{-\frac{n-1}{2}} (-1)^{-\frac{n-1}{2}}}{\left(\frac{n-1}{2}\right)! \sqrt{n}} \sin x.$$

And we have

$$\begin{aligned} K_N(\frac{x}{\sqrt{N}}, \frac{y}{\sqrt{N}}) &= \frac{\sqrt{N}}{N!\sqrt{2\pi}} \frac{H_N(\frac{x}{\sqrt{N}}) H_{N-1}(\frac{y}{\sqrt{N}}) - H_{N-1}(\frac{x}{\sqrt{N}}) H_N(\frac{y}{\sqrt{N}})}{x - y} \\ &= \frac{2^{-N+1}}{N!\sqrt{2\pi}} \frac{N!(N-1)!(-1)}{\left(\frac{N}{2}\right)! \left(\frac{N-2}{2}\right)!} \frac{\cos x \sin y - \sin x \cos y}{x - y} \\ &\to \frac{\sin(x-y)}{x - y}. \end{aligned}$$

□

for general s.

Proposition 4.4. *We have*

- $$H_n(x) = \frac{1}{i\sqrt{2\pi}} \int_{-i\infty}^{i\infty} w^n e^{\frac{(x-w)^2}{2}} \, \mathrm{d}w.$$

- $$H_n(ix) = \frac{i^n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^n e^{-\frac{(t-x)^2}{2}} \, \mathrm{d}t.$$

Proof. We apply the properties of characteristic function of $\mathscr{N}(0,1)$.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}+itx} \, \mathrm{d}t = e^{-\frac{x^2}{2}}$$

Take derivations for n times at each side

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} [\frac{\partial^n}{\partial x^n} e^{-\frac{x^2}{2}}] = \frac{1}{\sqrt{2\pi}} (-i)^n \int_{-\infty}^{\infty} t^n e^{-\frac{t^2}{2}+itx+\frac{x^2}{2}} \, \mathrm{d}t.$$

Change t into it and we get the proof. □

Now we can apply Proposition 4.4 to our problem.

Corollary 4.5. *$K_N(x,y)$ has a double contour integral form*

$$\frac{1}{(2\pi i)^2} \oiint \frac{w^N e^{\frac{w^2}{2}-wy}}{z^N e^{\frac{z^2}{2}-zx}} \frac{\mathrm{d}w \, \mathrm{d}z}{w-z}$$

Proof. By $\frac{H_k(x)}{k!} = \frac{1}{2\pi i} \oint \frac{e^{-\frac{z^2}{2}+zx}}{z^{N+1}} \, \mathrm{d}z$ and Proposition 4.4,

$$K_N(x,y) = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{N-1} \frac{H_k(x)}{k!} H_k(y) e^{-\frac{y^2}{2}} = \frac{1}{(2\pi i)^2} \oiint \frac{w^N e^{\frac{w^2}{2}-wy}}{z^N e^{\frac{z^2}{2}-zx}} \sum_{k=0}^{N-1} \frac{w^k}{z^{k+1}} \, \mathrm{d}w \, \mathrm{d}z.$$

Rearrange the contour: $|z| < |w|$ and we have

$$\sum_{k=0}^{N-1} \frac{w^k}{z^{k+1}} = \frac{1}{w-z}.$$

□

Return to our problem, we want to analyze

$$\frac{1}{\sqrt{N}} K_N(s\sqrt{N} + \frac{x}{\sqrt{N}}, s\sqrt{N} + \frac{y}{\sqrt{N}}) = \frac{1}{(2\pi i)^2} \oiint \frac{\exp(N(\log w + \frac{w^2}{2} - ws))}{\exp(N(\log z + \frac{z^2}{2} - zs))} \exp(zx - wy) \frac{\mathrm{d}w \, \mathrm{d}z}{w-z}.$$

Now we need to introduce a strong tool to calculate the asymptotic property of the integral called "steepest descent" or "saddle point method".

Remark 4.6. *The method is highly related to the Landau's calculation in statistical physics – use Taylor's expansion to analyze the phase transition. First, we warm up by exploring two interesting and familiar examples.*

Example 4.7 (Stiring's formula). *Here we are going to prove*

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + o_n(1)).$$

Proof.

$$n! = \Gamma(n+1) = \int_0^\infty x^n e^{-x}\,\mathrm{d}x = n^{n+1}\int_0^\infty \exp(-n(y - \log y))\,\mathrm{d}y.$$

The minimum of $y - \log y$ is $y = 1$, and notice that

$$y - \log y = 1 - \frac{1}{2}(y-1)^2 + O((y-1)^3).$$

By setting $y = 1 + \frac{z}{\sqrt{n}}$, we have

$$n! = n^{n+1}\frac{1}{\sqrt{n}}\int \exp(-n - \frac{z^2}{2} + O(\frac{1}{\sqrt{n}}))\,\mathrm{d}z.$$

which derives the result

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + o_n(1)).$$

□

Example 4.8 (CLT of binomial distribution). *Here we are going to prove*

$$\binom{n}{k} = \frac{1 + o_n(1)}{\sqrt{2\pi n}}\frac{1}{\alpha^{k+1/2}(1-\alpha)^{n-k+1/2}}$$

where $\alpha = \frac{k}{n}$. The formula can be directly derived from the CLT, but here we provide a new proof by steepest descent.

The genius idea is to observe that

$$\binom{n}{k} = \oint_O \frac{(1+z)^n}{z^{k+1}}\,\mathrm{d}z = \oint_O \exp(-n(\alpha\log z - \log(1+z)))\,\mathrm{d}z.$$

The critical point $z_c = \frac{\alpha}{1-\alpha}$, and we have $f(z) = f(z_c) + \frac{f''(z_c)}{2}(z - z_c)^2 + O((z-z_c)^3)$. Notice that from complex analysis, we have

Proposition 4.9. *There exists some contour around O, passing through the critical point z_c, such that $\Re(f)$ is maximized at z_c.*

Take the contour as above, we give

$$\begin{aligned}\binom{n}{k} &= \frac{1+o(1)}{2\pi i}\frac{(1+z_c)^n}{z_c^{k+1}}\int \exp(\frac{1}{2}\frac{(1-\alpha)^3}{\alpha}(z-z_c)^2)\,\mathrm{d}z \\ &= \frac{1+o_n(1)}{\sqrt{2\pi n}}\frac{1}{\alpha^{k+1/2}(1-\alpha)^{n-k+1/2}}.\end{aligned}$$

Return to our problem, let $f(z) = \log z + \frac{z^2}{2} - zs$, and we want to analyze

$$\frac{1}{(2\pi i)^2}\oiint \frac{\exp(N(f(w) - f(z)))\exp(zx - wy)}{w - z}\,\mathrm{d}w\,\mathrm{d}z.$$

To calculate the integral, we design a new contour, which satisfies

$$\Re(f(w) - f(z)) \le 0,$$

the equality holds if and only if $z = z_c$. By the steepest descent, we localize $w = z_c + \frac{\tilde{w}}{\sqrt{N}}$ and $z = z_c + \frac{\tilde{z}}{\sqrt{N}}$, where $z_c = \frac{s \pm i\sqrt{4-s^2}}{2}$. The integrand decays to 0 at the order of $\frac{1}{\sqrt{N}}$.

$$\begin{aligned}
\oiint &\to 0 + \underbrace{\int_{\bar{z}_c}^{z_c} \exp(z(x-y))\,\mathrm{d}z}_{\text{residual at } w=z} \\
&= \frac{1}{\pi} \frac{\exp(z_c(x-y)) - \exp(\bar{z}_c(x-y))}{(x-y)2i} \\
&= e^{\Re(z_c(x-y))} \frac{\sin[\Im(z_c(x-y))]}{\pi(x-y)} \\
&= e^{\frac{s}{2}(x-y)} \frac{\sin \pi\rho(x-y)}{\pi(x-y)}.
\end{aligned}$$

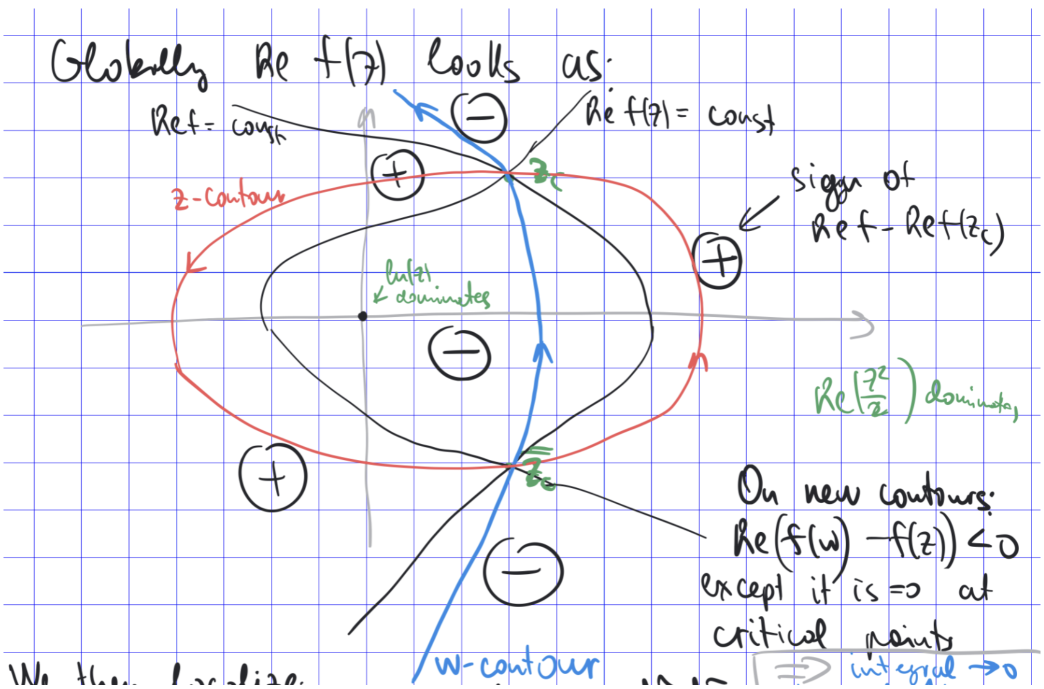


Figure 4: Choose specific contours for integral

$\square$

Remark 4.10. *For $|s| > 2$, we can show that double contour integrand decays exponential fast in N.*

Corollary 4.11. *For $|s| < 2$, the point process $\{\sqrt{N}(\lambda_i - s\sqrt{N})\}$ converges in distribution as $N \to \infty$ to the sine process of intensity $\rho = \frac{1}{2\pi}\sqrt{4-s^2}$.*

Proof. (proof sketch)

- There exists some DPP with kernel $K_{\sin}^{\rho}$.
- Express distributions of N_A and deduce convergence from Theorem 1.

□

Next we consider $|s| = 2$. We give an intuitive understanding first. Consider the semi-circle law and let $\frac{y_{N-k}}{\sqrt{N}} \approx y_{N-k}$. We have

$$\int_{y_{N-k}}^{2} \frac{1}{\sqrt{2\pi}} \sqrt{4 - x^2}\, \mathrm{d}x = \frac{k}{N}.$$

As the integral can be approximated as $(2 - y_{N-k})^{3/2}$ and we have $\lambda_N - \lambda_{N-1} \approx N^{-\frac{1}{6}}$.

Another intuitive understanding comes from the spacing. The spacing of $\sqrt{N}\lambda_N$ has the order $N^{\frac{1}{3}}$. Now, we introduce the theorem as belows.

Theorem 4.12.

$$\begin{aligned}\lim_{N\to\infty} N^{-\frac{1}{6}} K_N^{GUE}(2\sqrt{N} + xN^{-\frac{1}{6}}, 2\sqrt{N} + yN^{-\frac{1}{6}}) e^{(y-x)N^{\frac{1}{3}}} &= K_{Airy}(x, y) \\ &:= \frac{A_i(x)A_i'(y) - A_i'(x)A_i(y)}{x - y},\end{aligned}$$

where

$$A_i''(x) = xA_i(x) \quad \textit{and} \quad A_i(x) \to 0 \quad \textit{as} \quad x \to \infty.$$

Proof. (proof sketch) The idea is similar to the proof of $|s| < 2$. The only difference is that Taylor's expansion here is to the 3-rd order and we need to use $N^{\frac{1}{3}}$ to balance the N in steepest descent method. □

5 Feb 18th

5.1 Tracy-Widom, Gaudin-Mehta

In today's lecture, we focus on

- Largest eigenvalue.
- Gaps/spacings between eigenvalues.

Theorem 5.1 (Gap probability). *Let X be a DPP on $\mathscr{X}$ with correlation kernel $K(x,y)$ w.r.t measure μ. Then*

$$\mathbb{P}(\textit{there are no particles in } A) = \det\,[\mathrm{Id} - K]_{L_2(A,\mu)}$$

Remark 5.2. *If K has finite rank, i.e. $K(x,y) = \sum_{i=1}^m \phi_i(x)\psi_i(y)$, the expansion can be reduced to $m \times m$ determinant*

$$\det(I - UV) = \det(I - VU).$$

Proof. (Discrete X and finite A) By inclusive-exclusive theory

$$\begin{aligned}\mathbb{P}(\text{there are no particles in A}) &= 1 - \sum_a \mathbb{P}(\text{no particles at } a) + \frac{1}{2!}\sum_{a,b}\mathbb{P}(\text{no particles at } a,b) + ...\\ &= 1 - \sum_a \rho_1(a) + \frac{1}{2!}\sum_{a,b}\rho_2(a,b) + ...\end{aligned}$$

As $\rho_m(x_1,...,x_m) = \det K(x_i,x_j)_{1\le i,j\le m}$, by Fredholm expansion, we get the result. □

Remark 5.3. *We generalize the result*

$$\mathbb{E}\prod_{x\in X}(1-\phi(x)) \le \mathbb{E}\det(\mathrm{Id} - \phi K)$$

Definition 5.4. *Let $a_1 \ge a_2 \ge ...$ be point of the Airy$_2$ point process, then the law of a_1 is called Tracy-Widom Distribution ($TW_2, TW_{GUE}, F_2, F_{GUE}$).*

Proposition 5.5. *The distribution of TW_2 is*

$$\mathbb{P}(TW_2 \le t) = \det\,[Id - K_{Airy}]_{L_2(t,\infty)}.$$

Proof. We have

$$\mathbb{P}(TW_2 \le t) = \mathbb{P}(\text{there are no particles in } (t,+\infty)).$$

□

By some numerical results, the TW_2-distribution has some properties

- $\mathbb{E}\,[TW_2] \approx -1.77$
- $\mathrm{Var}\,(TW_2) \approx 0.81$
- $\mathbb{P}(TW_2 \ge s) \asymp \exp(-\frac{4}{3}s^{3/2})$
- $\mathbb{P}(TW_2 \le s) \asymp \exp(-\frac{1}{12}s^3)$

The rigorous analysis is as follows.

Theorem 5.6. *We have*

$$\mathbb{P}(TW_2 \le t) = \exp(-\int_t^\infty (x-t)^2 q(x)^2 \, \mathrm{d}x),$$

where

$$q''(x) = xq(x) + 2q(x)^3.$$

Remark 5.7. *We have $q(x) \approx A_i(x)$ as $x \to \infty$.*

By analyzing $\mathbb{P}(TW_2 \le t)$ and $\mathbb{P}(N^{-1/6}(\lambda_N - 2\sqrt{N}) \le t)$, we can derive

Theorem 5.8.

$$N^{-1/6}(\lambda_N - 2\sqrt{N}) \xrightarrow{d} TW_2.$$

Next, we care about the the second question: the spacing between two eigenvalues.

Definition 5.9. *For a shift invariant point process on $\mathbb{R}$, define its spacing as*

$$S_\rho = \textit{law of first positive point conditioning on a point at origin.}$$

Remark 5.10. *It is different from the law of the distance S'_ρ between the first positive and the first negative point. Usually $S_\rho > S'_\rho$.*

Theorem 5.11. *For a shift-invariant point process on $\mathbb{R}$ of density ρ_1 and assuming all ingredients exist and are smooth, we have*

$$\frac{\partial^2}{\partial x^2}\mathbb{P}(\textit{no articles in } (0,x)) = \rho_1 \cdot \mathbb{P}_{S_\rho}(x)$$

Proof. (Heuristic derivation in a discrete version) For a shift-invariant point process on $\mathbb{Z}$. The left hand is

$$\mathbb{P}(\text{there are particles at } l, \text{ but no articles at } 1, ...l-1)$$
$$-\mathbb{P}(\text{there are particles at } l, \text{ but no articles at } 0, ...l-1),$$

which also equals to

$$\mathbb{P}(\text{there are particles at } 0, l, \text{ but no articles at } 1, ...l-1) = \rho_1 \cdot \mathbb{P}_{S_\rho}(x)$$

□

Remark 5.12. *The probability for spacing in the bulk of GUE becomes*

$$\frac{\partial^2}{\partial x^2} \det\, (I - K_{sin})_{L^2(0,x)}.$$

The result is an application of the theorem in the last lecture.

robertsweeneyblanco.github.io/Computational_Random_Matrix_Theory/Eigenvalues/Zeta_and_GUE/Zeta_and_GUE.html

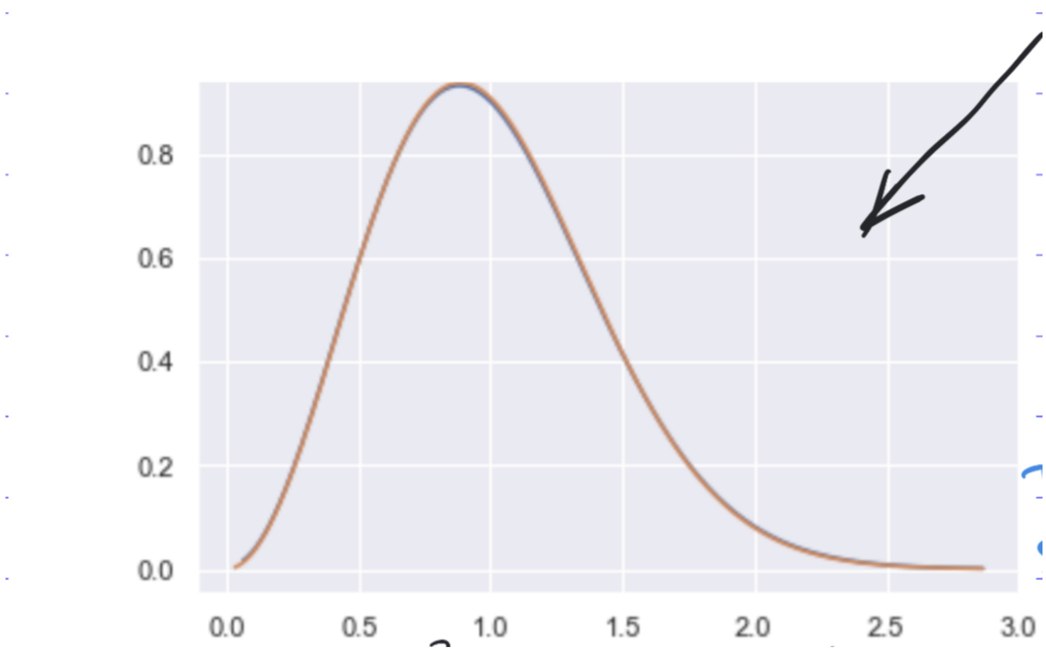


Figure 5: The blue curve represents $\frac{\partial^2}{\partial x^2}\det(\mathrm{Id} - K_{sin})$, and the orange curve represents $\frac{32s^2}{\pi^2}e^{-\frac{4s^2}{\pi}}$.

5.2 General beta

Now our question is how to generalize $\beta = 2$ to other values?

In 1960's, for $\beta = 1, 4$, there is a parallel but more complicated theory based on Pfaffian Point Process. (A simple understanding for Pfaffian of A is $\sqrt{\det(A)}$.)

In 2000's, there emerges theory which covers all $\beta > 0$ in a uniform way. It is very different from previous method, for no correlation functions are known for general β and one needs to proceed differently with different β.

Theorem 5.13. *Consider rescaled $G\beta E$ of density*

$$\prod_{i<j}(\lambda_j - \lambda_i)^\beta \prod_{i=1}^{N} \exp\left(-\frac{\beta\lambda_i^2}{4}\right).$$

*Then **for fix i***

$$\lim_{N\to\infty} N^{1/6}(\lambda_{N-i} - 2\sqrt{N}) \stackrel{d}{=} a_i,$$

where $\lambda_1 < \ldots < \lambda_N$ are eigenvalues of $G\beta E$ and $a_1 > a_2 > \ldots$ is $Airy_\beta$ process defined as the set of eigenvalues of the stochastic Airy Operator $SAO_\beta = \frac{\partial^2}{\partial x^2} - x + \frac{2}{\sqrt{\beta}}W'(x)$ acting on $L^2(\mathbb{R}_{\geq 0})$ and vanishing at $x = 0$.

Remark 5.14. *The advantages are the formula and the dependence of β are simple. The disadvantage is that we need to make sense of these eigenvalues. We give three approaches as follows, which are all based on the integral by parts.*

Remark 5.15. *This note is a fantastic material to understand the SAO_β: note.*

Approach 5.16 (Rauirez-Rider-Virag; Bloemendal-Virag)**.** *Idea: Use variational characterization of eigenvalues and eigenfunctions of self adjoint operators*

Consider Hilbert space L^* of functions of $[0,+\infty]$ with $f(0)=0$ and

$$||f||_*^2 = \int_0^\infty \left[(f'(x)^2 + (1+x)f^2(x))\right] dx < \infty$$

Then L^* is completion of smooth, compactly supported functions by this form.

Define a quadratic form

$$\begin{aligned} H(f,g) = \langle f, SAO_\beta g\rangle &= \int_o^\infty f\cdot\left(\frac{\partial^2}{\partial x^2}g - xg + \frac{2}{\sqrt{\beta}}w'g\right)dx \\ &= -\int_0^\infty f'(x)g'(x)dx + \int_0^\infty \left(\frac{x^2}{2} - \frac{2}{\sqrt{\beta}}w\right)(fg)'dx \end{aligned}$$

The last expression does not have w' and is a well-defined r.v. Then (a_i, f_i) is maxinum of $H(f,f)$ over $f\in L^*, ||f||=1$.

Approach 5.17 (Bloemendol, UToromto Phd Thesis)**.** *Transform the function by*

$$\tilde{f} = f\cdot e^{\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}$$

Then

$$\begin{aligned} \frac{\partial^2}{\partial x^2}f &= \frac{\partial^2}{\partial x^2}\left(\tilde{f}\cdot e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\right) \\ &= \frac{\partial}{\partial x}\left(\tilde{f}\cdot e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\cdot\left(-\frac{2}{\sqrt{\beta}}w(y)\right) + \tilde{f}'\cdot e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\right) \\ &= -\frac{2}{\sqrt{\beta}}w'(x)\cdot e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\cdot\tilde{f} + \frac{4}{\beta}\cdot w^2(x)\cdot e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\cdot\tilde{f} \\ &\quad - \frac{2}{\sqrt{\beta}}w(x)\cdot e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\cdot\tilde{f}' + e^{-\int_0^x \frac{2}{\sqrt{\beta}}w(y)\,\mathrm{d}y}\cdot\tilde{f}'' \end{aligned}$$

So

$$SAO_\beta f = \lambda f \Leftrightarrow \left(\frac{\partial^2}{\partial x^2} - \frac{2}{\sqrt{\beta}}w(x)\frac{\partial}{\partial x} + \frac{4}{\beta}w^2(x) - x\right)\tilde{f} = \lambda\tilde{f}$$

which can be solved by the classical theory of S-L operators.

Approach 5.18 (Gorin-Shkolnikov)**.** *Using Feynman Kac's Formula for SAO_β, we have*

$$K(x,y;T) = \frac{1}{\sqrt{2\pi T}}\exp\left(-\frac{(x-y)^2)}{2T}\right)\mathbb{E}_{B^{x\to y}}[\mathbb{1}_{B\geq 0}\exp\left(-\frac{1}{2}\int_0^T B(t)\,\mathrm{d}t + \frac{1}{\sqrt{\beta}}\int_0^\infty L_a(B)\,\mathrm{d}W(a)\right)]$$

where $B^{x\to y}$ is the Brownian bridge from $x\to y$ in times $t\in[0,T]$. W is a Brownian motion independent of B.

Then we give a heuristics for Theorem 5.13. From the second lecture, we have

$$T_\beta = \begin{pmatrix} \mathcal{N}(0,\frac{2}{\beta}) & \frac{\chi_{\beta(n-1)}}{\sqrt{\beta}} & \cdots & 0 \\ \frac{\chi_{\beta(n-1)}}{\sqrt{\beta}} & \mathcal{N}(0,\frac{2}{\beta}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{N}(0,\frac{2}{\beta}) \end{pmatrix}$$

We treat N as $N \times N$ matrix acting linearly on $f(\frac{1}{N^{1/3}}), (\frac{2}{N^{1/3}}), ..., (\frac{N}{N^{1/3}})$, and thus $N^{1/6}(T - 2\sqrt{N})$ becomes an operator. Notice that

$$\frac{\chi_{\beta(N-K)}}{\sqrt{\beta}} = \sqrt{N-K} + \mathscr{N}\left(\frac{1}{2\beta}\right) \approx \sqrt{N} - \frac{1}{2}\sqrt{N}\frac{K}{N} + \mathscr{N}\left(\frac{1}{2\beta}\right).$$

$$\begin{aligned} f(x) \to & N^{1/6}\left(f(x)\mathscr{N}\left(0, \frac{2}{\beta}\right) + [f(x+N^{-1/3}) + f(x-N^{-1/3})]\mathscr{N}\left(0, \frac{1}{2\beta}\right)\right) \\ & + N^{2/3}(f(x+N^{-1/3}) - 2f(x) + f(x-N^{-1/3})) \\ & - \frac{1}{2}x(f(x+N^{-1/3}) + f(x-N^{-1/3})) \\ \approx & N^{1/6}\left(f(x)\mathscr{N}\left(0, \frac{4}{\beta}\right)\right) + f''(x) - xf(x). \end{aligned}$$

Remark 5.19. *$N^{1/6}$ is the correct scale if we try to calculate*

$$N^{-1/3}\sum_{i=1}^{N} f\left(\frac{i}{N^{1/3}}\right) = \Omega(1)$$

for compact support f.

In order to make this rigorous, start from one of the three approaches, do same opperator for tridiagonal matrix and pass to the limit in the result.

6 Feb 25th

6.1 Corners and Bessel functions

In this lecture, we focus on the general question: How eigenvalues play with algebraic operations?

- Cutting Corners. (Lecture 6)
- Addition $A, B \mapsto A + B$. (Lecture 8+9)
- Multiplication $A, B \mapsto AB$. (No time)

First, we give the definition of "Orbital measure".

Definition 6.1. *Take $\lambda = (\lambda_1, ..., \lambda_N) \in \mathbb{R}$,*

$$Orbit(\lambda) = \textit{all matrices with eigenvalues } \lambda \textit{ equipped with uniform measure.}$$

Now we need to make sense of this uniform measure.

Approach 6.2. *This is a smooth compact manifold embedded into the Euclidean space of all symmetric matrices ($\langle X, Y \rangle = \mathrm{tr}(XY)$). And there is a well-defined metric and volumn form on $Orbit(\lambda)$, which can be renormalized to have volumn 1.*

Example 6.3. *Let $\beta = 1$, $N = 2$,*

$$X = \begin{pmatrix} a & \frac{b}{\sqrt{2}} \\ \frac{b}{\sqrt{2}} & c \end{pmatrix},$$

whose eigenvalues $= (\lambda_1, \lambda_2)$. By the properties of trace and determinant,

$$\begin{cases} a + c & = \lambda_1 + \lambda_2 \\ ac - \frac{b^2}{2} & = \lambda_1 \lambda_2 \end{cases}.$$

Change the coordinates to $(\frac{a+c}{\sqrt{2}}, \frac{a-c}{\sqrt{2}}, b)$, and we have

$$\begin{cases} \frac{a+c}{\sqrt{2}} & = \frac{\lambda_1+\lambda_2}{\sqrt{2}} \\ (\frac{a-c}{\sqrt{2}})^2 + b^2 & = \frac{(\lambda_1+\lambda_2)^2}{2} \end{cases}.$$

So (a, b, c) solving this forms a circle, which has a natural uniform measure.

Approach 6.4. *Matrix = Eigenvalues + Eigenvectors.*

$$Orbit(\lambda) = \textit{image of } O(N) \textit{ or } U(N) \textit{ under the map,}$$

and Haar measure on $O(N)$ and $U(N)$ directs to the orbit measure.

Example 6.5. *$O(2)$ has two components*

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \textit{ and } \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}.$$

WLOG, we choose the first rotation matrix.

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$
$$= \begin{pmatrix} \lambda_1 \cos^2\phi + \lambda_2 \sin^2\phi & (\lambda_1 - \lambda_2)\sin\phi\cos\phi \\ (\lambda_1 - \lambda_2)\sin\phi\cos\phi & \lambda_1 \sin^2\phi + \lambda_2 \cos^2\phi \end{pmatrix}$$

Here we can choose

$$\begin{cases} b & = \frac{\lambda_1-\lambda_2}{\sqrt{2}} \sin 2\phi \\ \frac{a-c}{\sqrt{2}} & = \frac{\lambda_1-\lambda_2}{\sqrt{2}} \cos 2\phi \end{cases},$$

which gives the same circle and the same uniform measure induced from $\phi \in [0, 2\pi]$.

Theorem 6.6. *Let $\mathscr{M}$ be a random element of Orbit(λ) and let $\mu_1 \le ... \le \mu_{N-1}$ be eigenvalues of its $(N-1)\times(N-1)$ corner. Then for $\beta = 1, 2, 4$,*

- $\lambda_1 \le \mu_1 \le \lambda_2 \le ... \le \mu_{N-1} \le \mu_N$ *almost surely.*
- *The density of μ w.r.t Lebesgue measure is*

$$\frac{\Gamma(N\frac{\beta}{2})}{(\Gamma(\frac{\beta}{2}))^N} \prod_{i<j}(\mu_j-\mu_i) \prod_{i=1}^{N-1}\prod_{j=1}^{N}(\mu_i-\lambda_j)^{\frac{\beta}{2}-1} \prod_{i<j}(\lambda_j-\lambda_i)^{1-\beta}.$$

Our main technique is the lemma as follows.

Lemma 6.7. *The law of $\mu_1, ...\mu_{N-1}$ is the same as $N-1$ roots of equation*

$$\sum_{i=1}^{N} \frac{\xi_i}{z-\lambda_i} = 0,$$

where $\xi_i \overset{iid}{\sim} \chi^2_\beta$.

Proof. We have $\mu_1, ..., \mu_{N-1}$ are $N-1$ roots of

$$\det \begin{pmatrix} & & & 0 \\ & U\,\mathrm{diag}(\Lambda)\,U^* - zI_N & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Multiplies $\begin{pmatrix} U^* & 0 \\ 0 & 1 \end{pmatrix}$ on the left and $\begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix}$ on the right, we get

$$\begin{aligned} &\det \begin{pmatrix} & & & 0 \\ & U\,\mathrm{diag}(\Lambda)\,U^* - zI_N & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\ =&\det \begin{pmatrix} \lambda_1 - z & 0 & \cdots & 0 & u_1^* \\ 0 & \lambda_2 - z & \cdots & 0 & u_2^* \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N - z & u_N^* \\ u_1 & u_2 & \cdots & u_N & 0 \end{pmatrix} = -\sum_{i=1}^{N} \frac{u_i u_i^*}{\lambda_i - z} \prod (\lambda_i - z). \end{aligned}$$

Actually, $(u_1 \dots, u_N)$ is a row of uniformly random orthogonal (unitary) matrix, which is equivalent to the uniformly random unit vertor in $\mathbb{R}^N$ $(\mathbb{C}^N)$, which is

$$\left(\frac{v_1}{\sqrt{|v_1|^2 + \cdots + |v_N|^2}}, \frac{v_2}{\sqrt{|v_1|^2 + \cdots + |v_N|^2}}, \dots, \frac{v_N}{\sqrt{|v_1|^2 + \cdots + |v_N|^2}} \right)$$

where $(v_1, \ldots, v_N)$ is i.i.d. Guassian vector. Notice that $v_i v_i^* \sim \chi^2_\beta$ ($u_i u_i^* = \frac{v_i v_i^*}{\sum_{i=1}^N |v_i|^2} = \frac{\xi_i}{\sum_{i=1}^N \xi_i}$), we find $\mu_1, \ldots, \mu_N$ solve

$$-\frac{\prod(\lambda_i - z)}{\sum_{i=1}^N \xi_i}\left(\sum_{i=1}^N \frac{\xi_i}{\lambda_i - z}\right) = 0$$

□

Now we are ready to prove the theorem.

Proof. Set $\varphi_i = \frac{\xi_i}{\sum_{i=1}^N \xi_i}$ ($= u_i u_i^*$). Then $\varphi \geq 0, \sum_{i=1}^N \varphi_i = 1$. The joint density of $\varphi_1, \ldots, \varphi_N$ is given by the Dirichlet distribution of density

$$p(x_1, \ldots, x_N) = \frac{\Gamma(N\frac{\beta}{2})}{\Gamma(\frac{\beta}{2})^N} x_1^{\frac{\beta}{2}-1} \cdots x_N^{\frac{\beta}{2}-1} dx_1 \cdots dx_{N-1}$$

Because ξ_i are all Γ-distribution with density

$$\frac{1}{2^{\beta/2}} \frac{1}{\Gamma(\frac{\beta}{2})} x^{\frac{\beta}{2}-1} e^{-\frac{x}{2}}, \ x > 0$$

and then integrate them on the condition $x_1 + \cdots + x_N$. By the Lemma, we multiply $\prod_{i=1}^N (z - \lambda_i)$ with $A = \sum_{i=1}^N \frac{\xi_i}{\lambda_i - z}$ gives a polynomial equations of degree $N-1$, so it has at most $N-1$ real roots, then A also has at most $N-1$ real roots.

- When $z = \lambda_i + \varepsilon$ (small enough), $A > 0$
- When $z = \lambda_{i+1} - \varepsilon$, $A < 0$

Hence, there is a root in each $[\lambda_i, \lambda_{i+1}]$ denoted by μ_i, $i = 1, \ldots, N-1$. Thus we porve the first conclusion. To calculate the density of $\mu_1, \ldots, \mu_{N-1}$, we need to compute the Jacobian of the map: $(\varphi_1, \ldots, \varphi_N) \to (\mu_1, \ldots, \mu_{N-1})$

$$\begin{aligned}\prod_{i=1}^N (z - \lambda_i) \sum_{i=1}^N \frac{\xi_i}{\lambda_i - z} &= \prod_{j=1}^{N-1} (z - \mu_j) \\ \sum_{i=1}^N \frac{\xi_i}{\lambda_i - z} &= \frac{\prod_{j=1}^{N-1} (z - \mu_j)}{\prod_{i=1}^N (z - \lambda_i)}\end{aligned}$$

Multiply by $z - \lambda_a$, set λ_a to get

$$\varphi_a = \frac{\prod_{j=1}^{N-1} (\lambda_a - \mu_j)}{\prod_{i \neq a} (\lambda_a - \lambda_i)}, \ a = 1, \ldots, N$$

Thus

$$\frac{\partial \varphi_a}{\partial \mu_b} = \frac{\prod_{j \neq b} (\lambda_a - \mu_j)}{\prod_{i \neq a} (\lambda_a - \lambda_i)} = \frac{\prod_{j=1}^{N-1} (\lambda_a - \mu_j)}{\prod_{i \neq a} (\lambda_a - \lambda_i)} \cdot \frac{1}{\mu_b - \lambda_a}$$

Then the Jacobian is

$$\det \left[\frac{\partial \varphi_a}{\partial \mu_b}\right]_{a,b=1}^{N-1} = \frac{\prod_{i,j=1}^{N-1} (\lambda_i - \mu_j)}{\prod_{j<i<N} (\lambda_j - \lambda_i)^2 \prod_{i=1}^N (\lambda_N - \lambda_i)} \cdot \det \left[\frac{1}{\mu_b - \lambda_a}\right]_{a,b=1}^{N-1}$$

While Cauchy determinant tells us that

$$\det \left[\frac{1}{\mu_b - \lambda_a}\right]_{a,b=1}^{N-1} = \frac{\prod_{i<j<N}(\mu_i - \mu_j)\prod_{i<j<N}(\lambda_j - \lambda_i)}{\prod_{i,j}(\mu_i - \lambda_j)}$$

which gives the Jacobian

$$\frac{\prod_{i<j<N}(\mu_j - \mu_i)}{\prod_{i<j\le N}(\lambda_j - \lambda_i)}$$

Then we plug the

$$\varphi_a = \frac{\prod_{j=1}^{N-1}(\lambda_a - \mu_j)}{\prod_{i\neq a}(\lambda_a - \lambda_i)},\ a = 1, \dots, N$$

into $p(x_1, \dots, x_N)$ and use the expression of Jacobian, which gives the 2nd conclusion. □

Corollary 6.8. *Take Orbit(λ), $N \times N$ matrix and consider eigenvalues of all principal corners. Let $x_i^K (1 \le i \le K \le N)$ be i-th eigenvalue of $K \times K$ corner. Then:*

- *Eigenvalues form an interlacing triangular array, i.e.*

$$x_i^{K+1} \le x_i^K \le x_{i+1}^{K+1} \text{ and } x_i^N = \lambda_i.$$

- *The joint density of x_i^K is*

$$\propto \prod_{K=1}^{N-1}\prod_{i<j\le K}(x_j^K - x_i^K)^{2-\beta}\prod_{a=1}^{K-1}\prod_{b=1}^{K}(x_a^{K-1} - x_b^K)^{\beta/2-1}$$

Proof. By iteration of theorem 1. □

Remark 6.9. *For $\beta = 1, 2, 4$, this is a theorem. For other $\beta > 0$, this is a definition of "β corners process with top row λ".*

For $\beta = 2$, all factors disappear and you end up with uniform measure on interlacing array.

Example 6.10 (N=2)**.**

- *For $\beta = 2$, x is uniform between λ_1 and λ_2.*
- *For $\beta = 1$, x has density $(x - \lambda_1)^{-1/2}(\lambda_2 - x)^{-1/2}$.*

Theorem 6.11. $\lim_{\beta\to\infty}$ *β-corners process with top row λ = determistic (a_i^K), such that*

$$\prod_{i=1}^{K}(z - a_i^K) = \frac{K!}{N!}\left(\frac{\partial}{\partial z}\right)^{N-K}\prod_{i=1}^{N}(z - \lambda_i).$$

Proof. We only prove the case $K = 1$. By Lemma 6.7, we have

$$\left(\frac{\partial}{\partial z}\right)\prod_{i=1}^{N}(z - \lambda_i) = \prod_{i=1}^{N}(z - \lambda_i)\sum_{i=1}^{N}\frac{1}{z - \lambda_i} = \lim_{\beta\to\infty}\prod_{i=1}^{N}(z - \lambda_i)\sum_{i=1}^{N}\frac{\chi^2_{N\beta}/N\beta}{z - \lambda_i}.$$

□

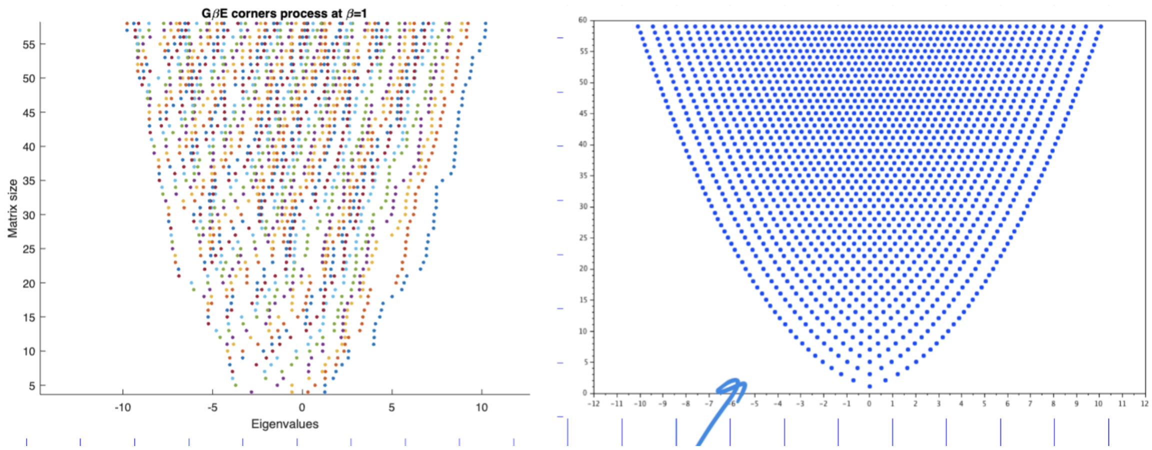


Figure 6: The left figure demonstrates the sample GOE ($N = 60$) and we plot the eigenvalues of the corner. The right figure demonstrates the roots of Hermite polynomials.

Now we introduce tools for corners processes: Multivariate Bessel functions. The key idea is similar to the characteristic function, i.e

$$A \mapsto F_A(Z) = \mathbb{E}[\exp(\mathrm{tr}(AZ))].$$

Observation 6.12. *If the law of A is invariant under orthogonal/unitary conjugation, then $F_A(Z)$ depends on Z only through eigenvalues of Z.*

Example 6.13. *Let A be GOE/GUE,*

$$\mathbb{E}[\exp(\mathrm{tr}(AZ))] = \mathbb{E}\left[\exp\left(\sum_{i=1}^{N} a_{ii} z_i\right)\right] = \exp\left(\sum_{i=1}^{N} \frac{z_i^2}{2}\right).$$

Definition 6.14 (The Mutivariate Bessel function)**.** *For real λ and $\beta = 1, 2, 4$,*

$$B_{\lambda_1,\dots,\lambda_N}(z_1, \dots, z_N; \beta) = \mathbb{E}_{A\sim Orbit(\lambda)}[\exp(\mathrm{tr}(AZ))],$$

where $z_1, \dots, z_N$ is the eigenvalues of Z. The trace only depends on the eigenvalues of Z.

Remark 6.15. *We recall the classic Bessel functions, which is defined as*

$$J_\alpha(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\ \Gamma(n+\alpha+1)} \left(\frac{x}{2}\right)^{2n+\alpha}.$$

We have $J_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi}} \frac{\sin x}{\sqrt{x}}$ We have the iterative equation:

$$x^2 J_\alpha'' + x J_\alpha' + (x^2 - \alpha^2) J_\alpha = 0.$$

We have the theorem as follows.

Theorem 6.16.

$$\int e^{i\langle a,u\rangle}\, \mathrm{d}u = \left(\frac{2}{x}\right)^\alpha \Gamma(\alpha+1) J_\alpha(x),$$

where $\alpha = \frac{n}{2} - 1$ and $x = \|a\|$.

Remark 6.17. *On the analogy of this, we take*

$$\begin{aligned} S^{n-1} &\leftrightarrow Orbit(\lambda) \\ x &\leftrightarrow \textit{eigenvalues } \lambda_1 < \dots < \lambda_n. \end{aligned}$$

We now give the RMT version.

Theorem 6.18. *At $\beta = 2$, we have*

$$B_{\lambda_1,\dots,\lambda_N}(z_1,\dots,z_N,2) = \prod_{K=1}^{N-1} K! \cdot \frac{\det[(\exp(\lambda_i z_j)]_{i,j=1}^N}{\prod_{i<j}(\lambda_i - \lambda_j)(z_i - z_j)}.$$

Let's give a restatement. By Harish-Chandra integral, we have

$$\int_{U(N)} \exp(U\Lambda U^* Z) = \prod_{K=1}^{N-1} K! \cdot \frac{\det[(\exp(\lambda_i z_j)]_{i,j=1}^N}{\prod_{i<j}(\lambda_i - \lambda_j)(z_i - z_j)},$$

where $\Lambda = \mathrm{diag}\{\lambda_1,\dots,\lambda_N\}$ and $Z = \mathrm{diag}\{z_1,\dots,z_N\}$.

Proof. Almost the same as Theorem 6.6. □

Remark 6.19. *What about other values of β?*

- *At $N = 1$, $B_x(x;\beta) = e^{\lambda x}$ for all β.*
- *For $N \geq 2$, there is a similar series expansion of $B_\lambda(z_1,\dots,z_N;\beta)$ in power series in $z_1,\dots z_N$.*

Theorem 6.20 (Definition)**.** *Let $\{x_i^K\}(1 \leq i \leq K \leq N)$ be β corners process with top row λ. Then*

$$B_{\lambda_1,\dots,\lambda_N}(z_1,\dots,z_N;\beta) = \mathbb{E}\left[\exp\left(\sum_{K=1}^{N} z_K \left(\sum_{i=1}^{K} x_i^K - \sum_{j=1}^{K-1} x_j^{K-1}\right)\right)\right],$$

where x_i^K is defined in Corollary 6.8, and we take expectation for $z \sim \mathcal{N}(0, I_N)$ and matrix $\sim$ $Orbit(\lambda)$. For $\beta = 1, 2, 4$, it is a theorem. While for other β, it is a definition.

7 Mar 4th

7.1 Asymptotics of corners

In this lecture, we explore the limit of corners as $N \to \infty$. There are several meanings of the limit behavior:

- How K grows with N (fixed or obey thermodynamic conditions).
- Macroscopic limits (semicircle law) or microscopic limits (bulk or edge).

First, we consider the case K is fixed. The main result is as follows.

Theorem 7.1. *Set $\beta = 1, 2, 4$, $\lambda = (\lambda_1, ..., \lambda_N)$ depends on N. Define $m(\lambda) = \frac{1}{N}\sum_{i=1}^N \lambda_i$, $V(\lambda) = \frac{1}{N}\sum_{i=1}^N \lambda_i^2 - m(\lambda)^2$ and $T(\lambda) = \frac{1}{N}\sum_{i=1}^N [\lambda_i - m(\lambda)]^3$. Assume*

$$\limsup_{N\to\infty} \frac{T(\lambda)^{1/3}}{V(\lambda)^{1/2}} \leq C,$$

then

$$\lim_{N\to\infty} \frac{A_K(\lambda) - m(\lambda)\,\mathrm{Id}}{V(\lambda)^{1/2}} \sqrt{\frac{\beta N}{2}} = G\beta E.$$

Example 7.2 (General)**.** *Consider*

$$\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i} \to \mu \text{ and } |\lambda_i| < C.$$

Example 7.3 (Particular)**.** *Consider $\lambda_1 = \lambda_2 = \cdots = \lambda_{N/2} = 0$ and $\lambda_{N/2+1} = \cdots = \lambda_N = 1$, $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. We have $m(\lambda) = \frac{1}{2}$, $V(\lambda) = \frac{1}{4}$ and $T(\lambda) = \frac{1}{8}$.*

Next we give a proof for $\beta = 1$ and $K = 1, 2$.

Proof. We have

$$A^\lambda = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*,$$

We have

$$A^\lambda_{[1,1]} = \sum_{i=1}^N \lambda_i u_i^2 \overset{d}{=} \sum_{i=1}^N \lambda_i \frac{\xi_i^2}{\sum_{j=1}^N \xi_j^2},$$

and

$$\frac{A^\lambda_{[1,1]} - m(\lambda)}{\sqrt{V(\lambda)}} = \frac{1}{\sum_{j=1}^N \xi_j^2} \sum_{i=1}^N \frac{\lambda_i - m(\lambda)}{\sqrt{V(\lambda)}} \xi_i^2.$$

Now we compute the $1, 2, 3$ order of moments.

$$\mathbb{E}\left[\sum_{i=1}^N \frac{\lambda_i - m(\lambda)}{\sqrt{V(\lambda)}} \xi_i^2\right] = \frac{\sum_{i=1}^N (\lambda_i - m(\lambda))}{\sqrt{V(\lambda)}} = 0,$$

$$\mathbb{E}\left[\sum_{i=1}^N \frac{(\lambda_i - m(\lambda))^2}{V(\lambda)} (\xi_i^2 - 1)^2\right] = N \cdot \mathbb{E}(\xi_i^2 - 1)^2 = 2N,$$

$$\mathbb{E}\left[\sum\left|\frac{\lambda_i - m(\lambda)}{\sqrt{V(\lambda)}}\right|^3 \left|\xi_i^2 - 1\right|^3\right] \leq \text{const} \cdot N.$$

By Lindeberg's or Lyapunov's CLT, we have

$$\frac{A^{\lambda}_{[1,1]} - m(\lambda)}{\sqrt{V(\lambda)}} \approx \frac{1}{N}\sqrt{2N}\mathscr{N}(0,1).$$

We can multiply $\sqrt{\frac{N}{2}}$ to get the result.

On the other hand, for $K = 2$, we can apply the Gram-Schimidt orthogonalization to get an orthogonal matrix. We have the two projected unit vector is

$$(u_1, u_2) \stackrel{d}{=} (\frac{\xi}{\|\xi\|_2}, \frac{\eta}{\|\eta\|_2}).$$

Applied CLT, the diagonal element obeys $\mathscr{N}(0,1)$ as proved. And the non-diagonal element is $\mathscr{N}(0,\frac{1}{2})$ as $\mathbb{E}\ [\xi_i^2\eta_i^2] = 1 = \frac{1}{2}\mathbb{E}\ [(\xi_i^2 - 1)^2]$. □

Corollary 7.4. *In the setting of Theorem 7.1, the first K rows of $\beta-$corners process $\{x_i^j\}_{1\leq i\leq j\leq K}$ converge to $G\beta E$ corners process.*

Here we give a intuitive check for the corollary. Take λ to be $\frac{1}{\sqrt{N}}GUE$-eigenvalues. $m(\lambda) \to 0$, $V(\lambda) \to 1$ and $T(\lambda)$ bounded. Then

$$\lim_{N\to\infty} GUE_{[K]}(N) = GUE_{[K]},$$

which is obviously true.

Now we consider K grows with N, and we introduce the theorem as follows.

Theorem 7.5. *Suppose λ depends on N in such a way that $\frac{1}{N}\delta_{\lambda_i} \to \mu_1$ and $\sup_i |\lambda_i| < C$. Let $x_1, ..., x_K$ be eigenvalues of $K \times K$ corner of $Orbit(\lambda)$ and assume $\frac{K}{N} \to \alpha$. Then*

$$\frac{1}{K}\sum_{i=1}^{K} \delta_{x_i} \to \mu^{\alpha} \text{ (weakly, in probability).}$$

We only care about the existence for now.

Remark 7.6. *The statement should be true for all β, but can be found only for $\beta = 1, 2, 4$.*

We here give a proof sketch for $\beta = 1$.

Proof. Let X be a $N \times N$ corner. We start with the following expression:

$$\frac{1}{K}\sum_{i=1}^{K}(x_i)^m = \frac{1}{K}\text{Trace}\,(X^m) = \text{sum of finite products of elements of } X$$

We want to apply a version of LLN or CLT to this large sum, using representation through i.i.d. Gaussians like in Theorem 7.1. Let's only do it for $m = 1$:

$$\frac{1}{K}\text{Trace}(X) = \frac{1}{K}\sum_{j=1}^{K}\sum_{i=1}^{N}\lambda_i(u_i^{(j)})^2$$

where $u_i^{(j)}$ is the matrix element of the orthogonal matrix U. By $\sum_{i=1}^N (u_i^{(j)})^2 = 1$ and symmetry, we have

$$\mathbb{E}\left((u_i^{(j)})^2\right) = \frac{1}{N}$$

Thus, we get:

$$\frac{1}{K}\text{Trace}(X) \approx \frac{1}{K} \cdot K \cdot \frac{\sum_{i=1}^N \lambda_i}{N} = \frac{\sum_{i=1}^N \lambda_i}{N}$$

This implies that the first moment of μ^α is the same as for μ^1:

$$\int x\, d\mu^\alpha(x) = \int x\, d\mu^1(x).$$

□

The question now is how to compute μ^α.

Example 7.7. *Consider $\mu^1 = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. We give the plot of the plot of μ^α as follows.*

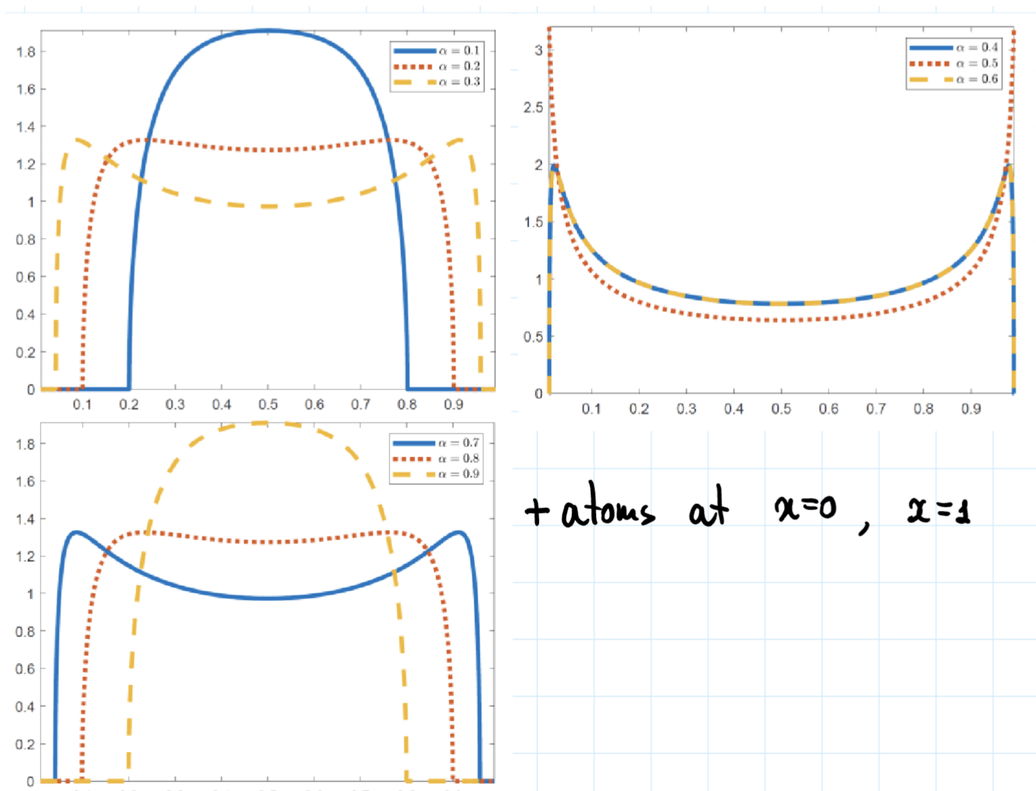


Figure 7: The behavior of μ^α for different α.

Now we give a new encoding of measures by functions of complex variable.

Definition 7.8 (Voivulescu, R-transform)**.** *For a compactly supported measure μ, recall $G_\mu(z) = \int \frac{1}{z-x}\mu(\mathrm{d}x) = \frac{1}{z} + \frac{1}{z^2}m_1 + \frac{1}{z^3}m_2 + \cdots$. Now we define*

$$R_\mu(z) = G^{-1}(z) - \frac{1}{z} : 0 \mapsto \textit{finity}.$$

Our theorem describes μ^α.

Theorem 7.9. *We have*

$$R_{\tilde{\mu}^\alpha}(z) = \frac{1}{\alpha} R_{\mu^1}(z),$$

where $\tilde{\mu}^\alpha[A] = \mu^\alpha[\alpha A]$.

Example 7.10 (Semicircle Law)**.** *We have* $G(z) = \frac{1}{2}(z - \sqrt{z^2 - 4})$ *and we can derive that* $G^{-1}(z) = z + \frac{1}{z}$, *which implies*

$$R(z) = z.$$

We have $\tilde{\mu}^\alpha$ *is the same semicircle but stretched. The stretch coefficient is like*

$$G^\alpha = \frac{y - \sqrt{y^2 - 4\alpha}}{2\alpha}.$$

Now we give a proof of Theorem 7.9 for $\beta = 2$. First, we give some lemmas.

Lemma 7.11. *For Bessel functions*

$$\begin{aligned} B_{\lambda_1,\dots,\lambda_n}(z_1, ..., z_n; \beta = 2) &= \mathbb{E} \exp \left[\mathit{Trace} \left(A \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \right) \right] \\ &= \prod_{i=1}^{N-1} [i!] \frac{\det[\exp(\lambda_i z_j)]_{i,j=1}^N}{\prod_{i<j}[(\lambda_i - \lambda_j)(z_i - z_j)]}. \end{aligned}$$

We have

$$\mathbb{E}_{x_1,\dots,x_k}[B_{x_1,\dots,x_k}(z_1, ..., z_k)] = B_{\lambda_1,\dots,\lambda_N}(z_1, ...z_k, 0^{N-k}).$$

Proof. Plug $z_{k+1} = ... = z_n = 0$ into the expression. ☐

Then we claim a key theorem as follows.

Theorem 7.12. *Suppose that* $\frac{1}{N}\delta_{\lambda_i} \to \mu$ *and* $\sup_i |\lambda_i| < C$. *Then*

$$\frac{1}{N} \log B_{\lambda_1,\dots,\lambda_N}(Nz, 0^{N-1}) \to \int_0^z R_\mu(u)\, \mathrm{d}u.$$

If the theorem holds we can prove the Theorem 7.9 directly. (The difference lies in plugging Nz or Kz.) To prove Theorem 7.12, we need another lemma.

Lemma 7.13. *For* $\beta = 2$, *we have*

$$B_{\lambda_1,\dots,\lambda_N}(z, 0^{N-1}) = \frac{(N-1)!}{z^{N-1}} \frac{1}{2\pi i} \oint \exp(vz) \prod_{i=1}^N \frac{1}{v - \lambda_i}\, \mathrm{d}v$$

Proof. By decomposition of the determinant

$$\det[\exp(\lambda_i z_j)] = \sum_{e=1}^N \exp(\lambda_e z_1)(-1)^{e-1} \det[\exp(\lambda_i z_j)]_{i \neq e}$$

When $z_2 = \ldots = z_N = 0$, using $B_{\lambda_2,\dots,\lambda_N}(0, ..., 0) = 1$:

$$B_{\lambda_1,\dots,\lambda_n}(z, 0^{N-1}) = \frac{(N-1)!}{z^{N-1}} \sum_{l=1}^N \exp(\lambda_l z) \frac{1}{\prod_{i \neq l}(\lambda_l - \lambda_i)}$$

This corresponds to the residue expansion of the integral. ☐

Now we give the whole proof of Theorem 7.12 using steepest descent method.

Proof. By the lemma, we have

$$B_{\lambda_1,\dots,\lambda_n}(Nz, 0^{N-1}) = \frac{(N-1)!}{N^{N-1}} \frac{1}{z^N} \frac{1}{2\pi i} \oint \exp \underbrace{\left(N \left(vz - \frac{1}{N} \sum \log(v - \lambda_i) \right) \right)}_{NF(v)} dv$$

Using Stirling's approximation

$$(N-1)! \approx \sqrt{2\pi N} \left(\frac{N}{e} \right)^{N-1}$$

we have

$$F(v) = vz - \frac{1}{N} \sum \log(v - \lambda_i).$$

Solve the equation for the critical point

$$0 = F'(v) = z - \frac{1}{N} \sum \frac{1}{v - \lambda_i},$$

and we get

$$v_c \approx G_\mu^{(-1)}(z).$$

Omitting existence of the contour deformation:

$$\frac{1}{N} \log B_{\lambda_1,\dots,\lambda_n}(Nz, 0^{N-1}) = -1 - \log z + F(v_c) + \bar{o}(1).$$

This simplifies to:

$$\begin{aligned} &= -1 - \log z + zG^{(-1)}(z) - \int_q^{G^{-1}(z)} G(v)dv + \frac{1}{N} \sum_{i=1}^N \log(q - \lambda_i) + o(1) \\ &= \int_0^z \left(G^{-1}(u) - \frac{1}{u} \right) \mathrm{d}u + o(1). \end{aligned}$$

Here, q can be anything, for example $q \to \infty$ is a good choice. To show coincidence, we show that they both tend to 0 at $z = 0$ and that their derivatives coincide.

$$\partial_z \text{LHS} = -\frac{1}{z} + G^{(-1)}(z) + z\partial_z G^{(-1)}(z) - \partial_z G^{(-1)}(z) G(G^{(-1)}(z)) = G^{(-1)}(z) - \frac{1}{z} = \partial_z \text{RHS}$$

□

At last, we turn back to the example $\mu^1 = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ and demonstrate the relationship to the Wachter's Law.

Example 7.14. *We have* $G(z) = \frac{1}{2}(\frac{1}{z} + \frac{1}{z-1})$ *and*

$$R(z) = \frac{z - 1 + \sqrt{z^2 + 1}}{2z}.$$

Rescaling by α*, we can calculate* $G^{-1}_{\tilde{\mu}^\alpha}(z)$ *and*

$$\tilde{\mu}^\alpha = C \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{x(\alpha^{-1} - x)}.$$

Remark 7.15. *We saw it*

- *When* $\beta = \infty$*, it is HW2. (By Theorem 6.11)*
- *In lecture 2, it is a limit for Jacobi ensemble = eigenvalues of two projectors.*

8 Mar 11th

8.1 Additions of matrices and free convolution

The main question for this section is what the eigenvalue of $C = A + B$ should be. First, we give some examples.

Example 8.1 (N=1). *Simply γ (eigenvalue of C) $= \alpha$ (eigenvalue of A) $+ \beta$ (eigenvalue of B)*

Example 8.2 (N=2). *By $\mathrm{Tr}(C) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$, we have equality*

$$\gamma_1 + \gamma_2 = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) \tag{$*$}$$

By some spectral inequality, we can derive that

$$\gamma_1 \le \alpha_1 + \beta_1 \tag{I}$$

$$\gamma_2 \ge \alpha_1 + \beta_2 \tag{II}$$

$$\gamma_2 \ge \alpha_2 + \beta_1 \tag{III}$$

Theorem 8.3. *The triplet $(\alpha_1 \le \alpha_2), (\beta_1 \le \beta_2), (\gamma_1 \le \gamma_2)$ satisfies $(*), \mathrm{(I)}, \mathrm{(II)}, \mathrm{(III)}$ iff there exists Hermitian A, B, C with such spectra and $A + B = C$.*

The answer for the question is that there are always one equality $\sum_{k=1}^N \gamma_k = \sum_{k=1}^N \alpha_k + \sum_{k=1}^N \beta_k$ and several inequalities.

Theorem 8.4. *There exists Hermitian A, B, C with $A + B + C = 0$ iff there exists honeycomb with boundaries parametrized by spectra of A, B, C.*

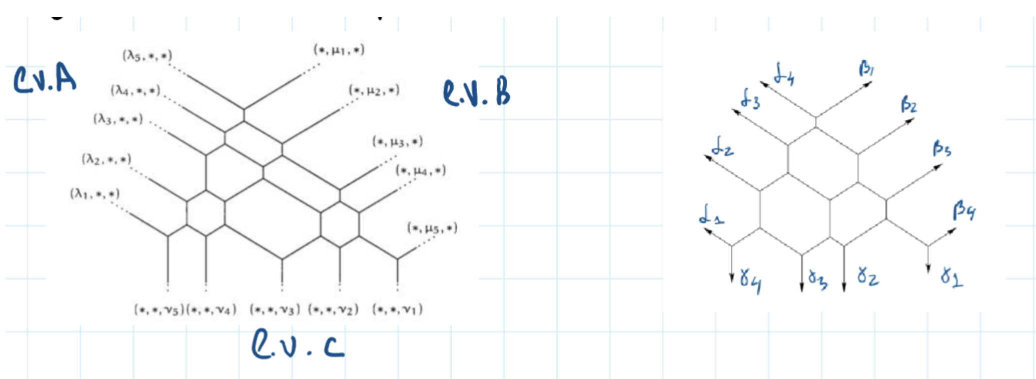


Figure 8: Honey comb

Probabilistic point of view We see A, B as (1) r.v. (2) With uniformly random eigenvectors (3) Independent

Then we have 2 questions:

1. What's the law of eigenvectors of $C = A + B$?
2. What about $N \to \infty$

Theorem 8.5. *Suppose that $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_N$ and $\beta_1 \le \beta_2 \le \cdots \le \beta_N$ depending on N in such way that*

$$\sup_{i,N} |\alpha_i| + |\beta_i| \le C$$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{\alpha_i} = \mu^A$$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{\beta_i} = \mu^B$$

Let A, B be matrices from $\operatorname{Orbit}(\alpha), \operatorname{Orbit}(\beta)$ and independent. Let $C = A+B$ having eigenvalues $\gamma_1 \le \gamma_2 \le \cdots \le \gamma_N$. Then

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{\gamma_i} = \mu^A \boxplus \mu^B \quad \textit{weakly and in probability}$$

*where $\mu^A \boxplus \mu^B$ is called the **free convolution** of μ^A and μ^B*

Proof. (Sketch of the proof.) Claim that

$$\operatorname{diag}\{\alpha_1, \dots, \alpha_N\} + U \operatorname{diag}\{\beta_1, \dots, \beta_N\} U^* = D$$

where U is random orthogonal/unitary and D has the same law of e.v. as $\gamma_1, \dots, \gamma_N$.

Then we compute

$$\frac{1}{N}\operatorname{Tr}(D^k) = \frac{1}{N}(\gamma_1^k + \cdots + \gamma_N^k)$$

$k = 1$

$$\operatorname{Tr}(D) = \sum_{i=1}^{N} \alpha_i + \sum_{i,j=1}^{N} u_{ij}^2 \beta_j = \sum_{i=1}^{N} \alpha_i + \sum_{i=1}^{N} \beta_i$$

So

$$\frac{1}{N}\operatorname{Tr}(D) \to \int x\mu^A(\mathrm{d}x) + \int x\mu^B(\mathrm{d}x)$$

$k = 2$

$$\begin{aligned}
\operatorname{Tr}(D^2) &= \operatorname{Tr}(\operatorname{diag}\{\alpha_1^2, \dots, \alpha_{\mathrm{N}}^2\}) + \operatorname{Tr}(U \operatorname{diag}\{\beta_1^2, \dots, \beta_{\mathrm{N}}^2\}\, U^*) \\
&\quad + 2\operatorname{Tr}(\operatorname{diag}\{\alpha_1, \dots, \alpha_N\} U \operatorname{diag}\{\beta_1, \dots, \beta_N\} U^*) \\
&= \sum_{i=1}^{N} \alpha_i^2 + \sum_{i=1}^{N} \beta_i^2 + 2\sum_{i,j=1}^{N} u_{ij}^2 \alpha_i \beta_j
\end{aligned}$$

From last lecture we know $u_{ij}, u_{i'j'}$ is very close to be independent. Applying LLN, we have

$$\begin{aligned}
\frac{1}{N}\operatorname{Tr}(D^2) &= \frac{1}{N}\sum_{i=1}^{N} \alpha_i^2 + \frac{1}{N}\sum_{i=1}^{N} \beta_i^2 + 2\left(\frac{1}{N}\sum_{i=1}^{N} \alpha_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} \beta_i\right) \\
&\to \int x^2\mu^A(\mathrm{d}x) + \int x^2\mu^B(\mathrm{d}x) + 2\int x\mu^A(\mathrm{d}x)\int x\mu^B(\mathrm{d}x).
\end{aligned}$$

Intuitively, we can derive that

$$\operatorname{Var}(\mu^A \boxplus \mu^B) = \operatorname{Var}(\mu^A) + \operatorname{Var}(\mu^B).$$

For larger K, the result is similar but more complicated. $\square$

Theorem 8.6 (how to compute $\mu^A \boxplus \mu^B$)**.**

$$R_{\mu^A \boxplus \mu^B}(z) = R_{\mu^A}(z) + R_{\mu^B}(z)$$

where $R_\mu(z) = G_\mu^{-1}(z) - \frac{1}{z}$ *and* $G(z) = \int \frac{1}{z-x}\mu(\mathrm{d}x) = \sum_{k=1}^{\infty} m_k z^{-k-1}$ *(*m_k *is the k-th moment of* μ*).*

Lemma 8.7. *For* $\beta = 1, 2, 4$*,*

$$\mathbb{E}[B_{\gamma_1,\dots\gamma_N}(z_1,\dots,z_N)] = B_{\alpha_1,\dots\alpha_N}(z_1,\dots,z_N) \cdot B_{\beta_1,\dots\beta_N}(z_1,\dots,z_N) \qquad (**)$$

Proof.

$$\begin{aligned} \mathrm{RHS} &= \mathbb{E}[\exp(\mathrm{Tr}(A\,\mathrm{diag}\{\mathrm{z}_1,\dots \mathrm{z}_\mathrm{N}\}))] \cdot \mathbb{E}[\exp(\mathrm{Tr}(B\,\mathrm{diag}\{\mathrm{z}_1,\dots \mathrm{z}_\mathrm{N}\}))] \\ &= \mathbb{E}[\exp(\mathrm{Tr}(C\,\mathrm{diag}\{\mathrm{z}_1,\dots \mathrm{z}_\mathrm{N}\}))] = \mathbb{E}_{\gamma_1,\dots,\gamma_N}\mathbb{E}[\exp(\mathrm{Tr}(H\,\mathrm{diag}\{\mathrm{z}_1,\dots \mathrm{z}_\mathrm{N}\}))] = \mathrm{LHS} \end{aligned}$$

where $H \sim \mathrm{Orbit}(\gamma_1,\dots,\gamma_N)$. □

Remark 8.8. *(**) is also the definition of* β*-dependent (not just 1,2,4) operation addition.*

Proof. (Proof for $\beta = 2$ of Theorem 8.6.) Set $z_1 = Nz$, $z_2 = \cdots = z_N = 0$ in Lemma 8.7. Take logarithm, divide by N and let N go to ∞,

$$\int_0^z R_{\mu^A \boxplus \mu^B}(u)\,\mathrm{d}u = \int_0^z R_{\mu^A}(u)\,\mathrm{d}u + \int_0^z R_{\mu^B}(u)\,\mathrm{d}u.$$

Take derivation and get the result. □

How do you think about free convolution? Analogy with classical convolution $\mu^1 * \mu^2$.

Definition 8.9. *Classical Convolution*

1. *Take* ξ^1 *to be* μ^1*-distributed and* ξ^2 *to be* μ^2*-distributed, then* $\mu^1 * \mu^2$ *is the distribution of* $\xi^1 + \xi^2$*.*
2. *If* $\mu^1(x), \mu^2(x)$ *are corresponding densities, then* $\mu^1 * \mu^2(x) = \int_{-\infty}^{\infty} \mu^1(y)\mu^2(x-y)\,\mathrm{d}y$*.*
3. $$\int x^n(\mu^1 + \mu^2)(\mathrm{d}x) = \sum_{k=1}^{n} \binom{n}{k} \int x^k \mu^1(\mathrm{d}x) \int x^{n-k}\mu^2(\mathrm{d}x).$$
   $$\mathbb{E}[\xi^1 + \xi^2]^n = \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}[\xi^1]^k \mathbb{E}[\xi^2]^{n-k}.$$
4. $$\log \int e^{itx}(\mu^1 + \mu^2)(\mathrm{d}x) = \log \int e^{itx}\mu^1(\mathrm{d}x) + \log \int e^{itx}\mu^2(\mathrm{d}x) \quad \textit{for all } t \textit{ in a small neighbourhood of } 0.$$
   $$\mathbb{E}[e^{it(\xi^1+\xi^2)}] = \mathbb{E}[e^{it\xi^1}]\mathbb{E}[e^{it\xi^2}].$$

Definition 8.10. *Free Convolution*

1. $\xi^1 + \xi^2 \approx$ *Our first definition of* $\mu^A \boxplus \mu^B$ *as a limit of addition of independent matrices.*
2. *The second definition does not exist.*

3. A version of the third definition can be developed.

4. $\log(\cdots) = \log(\cdots) + \log(\cdots) \approx R_{\mu^A \boxplus \mu^B} = R_{\mu^A} = R_{\mu^B}$.

Theorem 8.11. *Define classical cumulants ($K_n = K_n^{classical}$) of a prob measure μ through (assuming all moments exists)*

$$\log \int e^{itx} \mu(\mathrm{d}x) = K_1(it) + K_2 \frac{(it)^2}{2} + K_3 \frac{(it)^3}{3!} + \cdots$$

Then

1. K_n *is a polynomial in moments* m_n *of* μ

$$K_n = m_n + (\textit{homogenuous polynomial in } m_1, \ldots, m_{n-1} \textit{ of degree } n, \textit{ if } \deg(m_p) = p).$$

2.
$$m_n = \sum_{\pi=(B_1,\ldots,B_l)} \prod_{B\in\pi} K_{|B|}$$

where π *is set partition of* $\{1, \ldots, n\}$.

3.
$$K_n(\mu^1 * \mu^2) = K_n(\mu^1) + K_n(\mu^2).$$

Example 8.12. *For* $\mathscr{N}(m, \sigma^2)$, $K_1 = m, K_2 = \sigma^2, K_l = 0$ $(l \geq 3)$

$$\log \int e^{itx} \mu(\mathrm{d}x) = \log\left(\exp\left(itm - \frac{t^2\sigma^2}{2}\right)\right) = itm - \frac{t^2\sigma^2}{2}.$$

For $\lambda \cdot \mathrm{Poisson}(\gamma)$

$$\log \int e^{itx} \mu(\mathrm{d}x) = \gamma(e^{itx} - 1),$$

implying that

$$K_N = \gamma \lambda^N.$$

Proof. (Sketch of proof of Theorem 8.11.) Obviously *2.* ⇒ *1. 3.* is the statement of Definition 1.9.4. For *2.*,

$$\sum_{n\geq 1} \frac{K_n}{n!} z^n = \log\left(\sum_{j\geq 0} \frac{m_j}{j!} z^j\right)$$

$$\text{Take } \frac{\partial}{\partial z}: \quad \sum_{n\geq 1} \frac{K_n}{(n-1)!} z^{n-1} = \frac{\sum_{j\geq 1} \frac{m_j}{(j-1)!} z^{j-1}}{\sum_{j\geq 0} \frac{m_j}{j!} z^j}.$$

Multiply by denominator, compare coefficient of z^l, yields

$$\frac{m_{l+1}}{l!} = \frac{K_{l+1}}{l!} + \sum_{j=1}^{l} \frac{m_j}{j!} \frac{K_{l+1-j}}{(l+1-j)!}.$$

which is equivalent to *2.* (Exercise!) □

Theorem 8.13 (Analogy for free convolution)**.** *Definr free cumulants* $K_n = K_n^{free}$ *of* μ *through*

$$G_\mu^{-1}(z) - \frac{1}{z} = K_1 + K_2 z + K_3 z^2 + \cdots$$

$$G(z) = \sum_{k=1}^{\infty} m_k z^{-k-1} \quad \textit{(existence of all } m_n \textit{ is sufficed for existence of all } K_n\textit{).}$$

Then

1. K_n *is a polynomial in moments* m_n *of the form*

$$K_n = m_n + (\textit{homogenuous polynomial in } m_1, \dots, m_{n-1} \textit{ of degree } n, \textit{ if } \deg(m_p) = p).$$

2. $$m_n = \sum_{\pi=(B_1,\dots,B_l)} \prod_{B \in \pi} K_{|B|}$$

   where π *is non-crossing set partition of* $\{1, \dots, n\}$.

3. $$K_n(\mu^1 \boxplus \mu^2) = K_n(\mu^1) + K_n(\mu^2).$$

Non-crossing set partitions:

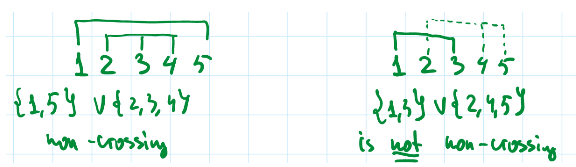


Figure 9: Explanation of Non-crossing

Example 8.14. *Semicircle law is the analogy of* $\mathcal{N}(0,1)$ *and Marchenllo-Pastur law is the analogy of Poisson distribution.*

Proof. (Sketch of the proof of Theorem 8.13.) Obviously *2.* ⇒ *1.* *3.* is the same as Theorem 8.6.

For *2.*,

$$R(z) = K_1 + K_2 z + \cdots \qquad G(z) = \frac{1}{z} + \frac{m_1}{z^2} + \cdots$$

$$(R(z) + \frac{1}{z}) \circ G(z) = z$$

$$R(G(z)) + \frac{1}{G(z)} = z \qquad 1 + G(z)P(G(z)) = zG(z)$$

$$1 + \sum_{n \geq 1} (G(z))^n K_n = \sum_{n \geq 0} m_n z^{-n}$$

Evaluating the coefficient of z^{-n}, we get

$$m_n = K_n + \cdots.$$

□

Corollary 8.15. *Recall* $\tilde{\mu}^\alpha$ *from Lecture 7. For each* $n = 1, 2, \dots$, $\tilde{\mu}^\alpha = \boxplus^n \mu^1$, *which means*

Cutting corners = addition in large matrix limit.

Proof. Comparing with Thm 7.9, both sides have R-transform $nR_{\mu^1}(z)$. □

Remark 8.16. *Lecture notes on free probability by Roland Speicher is recommended.*

9 Mar 18th

9.1 Signal plus noise

Today we continue to discussing

$$A + B = C,$$

where here A has low rank as $N \to \infty$ and B has i.i.d. matrix elements. There are many variants in statistics literature, e.g.

- Self-adjoint version:

$$A = \sum_{i=1}^{K} \alpha_i v_i v_i^*,$$

  and B is GβE.

- Rectangle version:

$$A = \sum_{i=1}^{K} \alpha_i v_i u_i^*,$$

  and the elements of B are i.i.d. Gaussians.

Remark 9.1. I *and* II *are like semicircle law and M-P law. We discuss* I *here.*

Consider the case $K = 1$,

$$C = \alpha v v^* + \sqrt{\frac{2}{\beta}} G\beta E.$$

A signal-noise setup requires us to recover α and v from C. In the setting of previous lecture, $A \approx 0$ and we expect C obeys semicircle law.

Theorem 9.2 (Spiked Random Matrices)**.** *Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ are N eigenvalues of*

$$C = \sqrt{N} a v v^* + \sqrt{\frac{2}{\beta}} G\beta E,$$

where $\|v\| = 1$ Then there exists a_{crit} such that

1. *If $a > a_{crit} = 1$, then*

$$\lim_{N\to\infty} \frac{\lambda_N}{\sqrt{N}} = a + \frac{1}{a} > 2.$$

2. *If $a < a_{crit}$, then*

$$\lim_{N\to\infty} \frac{\lambda_N}{\sqrt{N}} = 2.$$

By the theorem, we can easily get

Corollary 9.3.

$$\hat{a} = \frac{1}{2}\left(\frac{\lambda_N}{\sqrt{N}} + \sqrt{\frac{\lambda_N^2}{N} - 4}\right) \xrightarrow{N\to\infty} a.$$

Remark 9.4. *The corollary is only applicable if you see a spike.*

Theorem 9.5 (Recover v)**.** *In the setting of Theorem 9.2. Let $\hat{v}$ denote the eigenvector corresponding to λ_N and let ϕ be the angle between v and $\hat{v}$,*

$$\lim_{N\to\infty} \sin\phi = \frac{1}{a} \wedge 1.$$

Proof. **Step 1** a, λ_N and ϕ are unchanged under orthogonal/unitary transformations of matrix C. We do two transformations:

- Rotate v ti be the first basis vector $(1, 0^{N-1})$.
- Rotate in the orthogonal complement of $(1, 0^{N-1})$, so that the $(N-1)\times(N-1)$ bottom random corner of GβE becomes diagonal, while the law of the first row and column of GβE preserved.

Step 2 Now we brought C into

$$\hat{C} = \begin{pmatrix} a\delta_N + \mathcal{N}(0, \frac{2}{\beta}) & \xi_2 & \xi_3 & \cdots & \xi_N \\ \bar{\xi}_2 & \mu_2 & & & \\ \bar{\xi}_3 & & \mu_3 & & \\ \vdots & & & \ddots & \\ \bar{\xi}_N & & & & \mu_N \end{pmatrix}$$

and $\mu_2, \cdots, \mu_N$ be the eigenvalues of $\sqrt{\frac{2}{\beta}} G\beta E$ and $\xi_2, \cdots, \xi_N \overset{i.i.d}{\sim} \sqrt{\frac{2}{\beta}}$ corresponding Normal distribution. We find an eigenvector $(x_1, \cdots, x_N)$ of $\tilde{C}$ with eigenvalues λ

$$\begin{cases} x_1\left(a\sqrt{N} + \mathcal{N}\left(0, \frac{2}{\beta}\right)\right) + \sum_{i=1}^N \xi_i x_i = \lambda x_1 \\ x_1\bar{\xi}_2 + \mu_2 x_2 = \lambda x_2 \Rightarrow x_2 = \frac{x_1\bar{\xi}_2}{\lambda-\mu_2} \\ \vdots \\ x_1\bar{\xi}_N + \mu_N x_N = \lambda x_N \Rightarrow x_N = \frac{x_1\bar{\xi}_N}{\lambda-\mu_N} \end{cases} \Longrightarrow \begin{cases} x_2 = \frac{\bar{\xi}_2}{\lambda-\mu_2} \\ \vdots \\ x_N = \frac{\bar{\xi}_N}{\lambda-\mu_N} \end{cases}$$

Plug the $2 \sim N$ equations into the first line and we get

$$x_1\left[a\sqrt{N} + \mathcal{N}\left(0, \frac{2}{\beta}\right) - \lambda + \sum_{i=1}^N \frac{\xi_i\bar{\xi}_i}{\lambda - \mu_i}\right] = 0. \tag{$*$}$$

By interpolating of eigenvalues, C has 1 eigenvalue larger than μ_N. Only this eigenvalue has chance to become larger than $2\sqrt{N}$ as $\frac{\mu_N}{\sqrt{N}} \to 2$. Denote $y = \frac{\lambda_N}{\sqrt{N}}$ and investigate $(*)$ for $y > 2$.

$$a + \frac{\mathcal{N}\left(0, \frac{2}{\beta}\right)}{\sqrt{N}} - y + \frac{1}{N}\sum_{i=2}^N \frac{\xi_i\bar{\xi}_i}{y - \frac{\mu_i}{\sqrt{N}}} = 0. \tag{$**$}$$

By LLN, we have

$$\frac{1}{N}\sum_{i=2}^N \frac{\xi_i\bar{\xi}_i}{y - \frac{\mu_i}{\sqrt{N}}} \approx \frac{1}{N}\sum_{i=2}^N \frac{1}{y - \frac{\mu_i}{\sqrt{N}}} \to G(y) = \frac{1}{2}(y - \sqrt{y^2-4}).$$

Then we have

$$a - y + \frac{1}{2}(y - \sqrt{y^2-4}) = 0 \Longrightarrow y = a + \frac{1}{a},$$

and this proves Theorem 9.2.

Step 3 For Theorem 9.5, we notice that the eigenvalues are

$$(1, 0, \cdots, 0) \text{ and } \left(1, \frac{\bar{\xi}_2}{\lambda - \mu_2}, \cdots, \frac{\bar{\xi}_N}{\lambda - \mu_N}\right),$$

and we have

$$\cos^2\phi = \frac{1}{1+\sum_{i=2}^N \frac{\xi_i\bar{\xi}_i}{(\lambda-\mu_i)^2}} \to \frac{1}{1+\frac{1}{2}\frac{y}{\sqrt{y^2-4}}-\frac{1}{2}} = \frac{2\sqrt{y^2-4}}{y+\sqrt{y^2-4}}.$$

As $y = a+\frac{1}{a}$, we have

$$\sin^2\phi \to \frac{1}{a^2}.$$

□

Next, we discuss the fluctuation. Notice that without spike, the fluctuation of GβE is $N^{-1/2}\cdot N^{1/3} = N^{-1/6}$. In the spiked case the fluctuation is $N^{-1/2}\cdot N^{1/2} = \text{Const}$ – much larger.

Theorem 9.6. *Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ are N eigenvalues of*

$$C = \sqrt{N}avv^* + \sqrt{\frac{2}{\beta}}G\beta E,$$

where $\|v\| = 1$. Assuem $a > 1$, then

$$\lim_{N\to\infty}\left(\lambda_N - \left(a+\frac{1}{a}\right)\sqrt{N}\right) = \mathscr{N}\left(0, \frac{2}{\beta}\left(1-\frac{1}{a^2}\right)\right).$$

Proof.

$$a\sqrt{N} + \mathscr{N}\left(0,\frac{2}{\beta}\right) - \lambda + \sum\frac{\xi_i\bar{\xi}_i}{\lambda-\mu_i} = 0$$

Set $\lambda = \sqrt{N}(a+\frac{1}{a}) + \Delta\lambda$ in the equation

$$\begin{aligned}
0 &= -\frac{\sqrt{N}}{a} + \mathscr{N}\left(0,\frac{2}{\beta}\right) - \Delta\lambda + \frac{1}{\sqrt{N}}\sum_{i=2}^N \frac{\xi_i\bar{\xi}_i}{(a+\frac{1}{a}) - \frac{\mu_i}{\sqrt{N}} - \frac{\Delta\lambda}{\sqrt{N}}} \\
&= -\frac{\sqrt{N}}{a} + \mathscr{N}\left(0,\frac{2}{\beta}\right) - \Delta\lambda + \frac{1}{\sqrt{N}}\sum_{i=2}^N \frac{\xi_i\bar{\xi}_i}{(a+\frac{1}{a}) - \frac{\mu_i}{\sqrt{N}}} - \frac{\Delta\lambda}{N}\sum_{i=2}^N \frac{\xi_i\bar{\xi}_i}{((a+\frac{1}{a}) - \frac{\mu_i}{\sqrt{N}})^2} + O(N^{-\frac{1}{2}})
\end{aligned}$$

$$\xi_i\bar{\xi}_i\frac{1}{\beta}\chi^2_\beta \Rightarrow \mathbb{E}[\xi_i\bar{\xi}_i] = 1, \quad \text{Var}(\xi_i\bar{\xi}_i) = \frac{1}{\beta^2}\cdot 2\beta = \frac{2}{\beta}$$

Applying CLT,

$$\frac{1}{\sqrt{N}}\sum_{i=2}^N \frac{\frac{1}{\beta}\chi^2_\beta}{(a+\frac{1}{a}) - \frac{\mu_i}{N}} = \sqrt{N}G\left(a+\frac{1}{a}\right) + \mathscr{N}\left(0,\frac{2}{\beta}\left(-G'\left(a+\frac{1}{a}\right)\right)\right) + O(N^{-\frac{1}{2}})$$

where

$$G\left(a+\frac{1}{a}\right) = \frac{1}{a}, \quad G'\left(a+\frac{1}{a}\right) = -\frac{1}{a^2+1}$$

Thus we have

$$\mathscr{N}\left(0,\frac{2}{\beta}\right) + \mathscr{N}\left(0,\frac{2}{\beta}\left(-G'\left(a+\frac{1}{a}\right)\right)\right) + \Delta\lambda\left(-1+G'\left(a+\frac{1}{a}\right)\right) + O(N^{-\frac{1}{2}}) = 0$$

$$\Delta\lambda = \mathscr{N}\left(0,\frac{2}{\beta}\left(1-\frac{1}{a^2}\right)\right)$$

□

Remark 9.7. *In summary,*

- *for $a < 1$, λ_N has fluctuations $N^{-1/6} TW_\beta$;*
- *for $a > 1$, λ_N has fluctuations $\mathscr{N}\left(0, \frac{2}{\beta}\left(1 - \frac{1}{a^2}\right)\right)$.*

The change of behavior is called BBP phase transition.

Now we discuss what is happening exactly at $a = 1$. Let us derive some Heuristics for critical scaling of a,

$$2\sqrt{N} + N^{-1/6} TW_\beta = \left(a + \frac{1}{a}\right)\sqrt{N} + \mathscr{N}\left(0, \frac{2}{\beta}\left(1 - \frac{1}{a^2}\right)\right).$$

We get the right scale of a is $a = 1 + \theta N^{-1/3}$. We have the theorem as follows.

Theorem 9.8. *Assume $a = 1 + \theta N^{-1/3}$, $\theta \in \mathbb{R}$. Let λ_N be the largest eigenvalue of $C = \sqrt{N} a v v^* + \sqrt{\frac{2}{\beta}} G\beta E$. Then*

$$\lim_{N\to\infty} N^{1/6}\left(\lambda_N - 2\sqrt{N}\right) = F_{\beta,\theta}.$$

Remark 9.9. *$F_{\beta,\theta}$ interpolates between Tracy-Widom distribution ($\theta = -\infty$) and Gaussian ($\theta = +\infty$)*

We do not prove it, but give several approaches instead.

- *Deterministic Point Process for $A + GUE$ at $\beta = 2$ (which will discussed in the next lecture).*
- *Tridiagonal matrices for arbitrary $\beta > 0$. This approach is restricted to $K = 1$ spike, because tridiagonalization only works for $K = 1$.*

At the last of this lecture, we talk about the multi-spiked case. We choose a_i as eigenvalues and v_i as eigenvectors.

Theorem 9.10. *Suppose $a_1 > a_2 > \cdots > a_K > 1$. Then for each $i = 1, 2, \cdots, K$, we have*

$$\lim_{N\to\infty}\left(\lambda_{N-i+1} - \sqrt{N}\left(a_i + \frac{1}{a_i}\right)\right) = \mathscr{N}\left(0, \frac{2}{\beta}\left(1 - \frac{1}{a_i^2}\right)\right)$$

and

$$\lim_{N\to\infty} \sin\phi_i = \frac{1}{a_i}.$$

Remark 9.11. *It should be independent over i, but haven't found in any literature.*

Proof. Similar to the case $K = 1$. □

Now imagine that $a_1 = \cdots = a_K = a > 1$,

$$C = \sqrt{N} a \sum_{i=1}^{K} v_i v_i^* + \sqrt{\frac{2}{\beta}} G\beta E.$$

Theorem 9.12. *The asymptotics in this case becomes*

$$\left(\lambda_N - \left(a + \frac{1}{a}\right)\sqrt{N}, \cdots, \lambda_{N-K+1} - \left(a + \frac{1}{a}\right)\sqrt{N}\right) \to \sqrt{\frac{2}{\beta}\left(1 - \frac{1}{a^2}\right)} \quad G\beta E \text{ eigenvalues.}$$

Remark 9.13. *In practical, you see K spikes close to each other and then need to be careful with interpreting eigenvectors.*

10 Api 1st

10.1 Dyson Browian Motion

The task is to add time to random matrices with motivation that

- Physical objects come with evolutions, so we want RW to become a part of some dynamics.
- When you add matrices $C = A + B$. Sum at n terms can become time.
- Connection to 2d stat mechanics and Markov chain.
- An important formula proving universality theorems for RW ensembles.

Consider the case $N = 1$. GOE/GUE/GSE with $\mathscr{N}(0, 1)$ random variable.
Consider Brownian motion $B(t)$ started from $B(0) = a, t \geq 0$

- $B(s) - B(t) \sim \mathscr{N}(0, s - t)$.
- Continuous curve.
- Independent increments.

Matrix Brownian motion $X(t)$ which is $N \times N$ matrix filled with i.i.d. Brownian motions: $B(t) + i\tilde{B}(t)$ for $\beta = 2$, where $X(0)$ can be arbitrary deterministic. Set $U(t) = \frac{1}{2}(X(t) + X^*(t))$

Lemma 10.1. *Set $U(0) = A$. Then $U(t)$ has the law of $A + \sqrt{t}G\beta E$ for a fixed $t \geq 0$ (But $t-$dependence is different!)*

Proof. For $\beta = 2$,

$$\frac{X(t) - X(0)}{\sqrt{t}} = \text{matrix of i.i.d. } \mathscr{N}(0, 1) + i\mathscr{N}(0, 1)$$

$$U(t) = U(0) + \sqrt{t}\,\frac{1}{2}\left(\frac{X(t) - X(0)}{\sqrt{t}} + \frac{X^*(t) - X^*(0)}{\sqrt{t}}\right) = A + \sqrt{t}G\beta E$$

□

Theorem 10.2. *Let $\lambda(t) = (\lambda_1(t) \leq \lambda_2(t) \leq \cdots \leq \lambda_N(t))$ be eigenvalues of $U(t)$. Then $(\lambda(t))_{t\geq 0}$ is a Markov process.*

(Given present, future and past are indep = all info about $(\lambda(t))_{t\leq T}$) useful for predicting $(\lambda(T))_{t>T}$ is given by $\lambda(T)$

Remark 10.3. *$(\lambda^N, \lambda^{N-1})(t)$ is also Markov, but $(\lambda^N, \lambda^{N-1}, \lambda^{N-2})(t)$ is not. λ^N is e.v. of $N \times N$ matrix, and λ^{N-1} is for $N-1 \times N-1$ submatrix. Projection of a Markov rocess is rarely a arkov Process*

Proof. Study $\lambda(s)_{S>T}$ conditional on $\lambda(t)_{t<T}$. Write $U(t) = U(T) + (U(t) - U(T))$ (2 terms are indep by def of BW)

Conjugate with orthogonal/unitary U to diagonalize $U(T)$.

$$UU(t)U^* = \mathrm{diag}(\lambda_1(T), \ldots, \lambda_N(T)) + U(U(t) - U(T))U^*$$

LHS leads to the same e.v. $\lambda_1(t), \ldots \lambda_N(t)$. The second term in RHS is again a BM indep of $(\lambda(S))_{S\leq T}$. So no dependence on $\lambda(S)_{S<T}$ remained. □

How to describe an $N-$dim Markov process $\lambda_1(t) \leq \cdots \leq \lambda_N(t)$?

I By transition prob

II As a solution to a Stochastic Differential equation

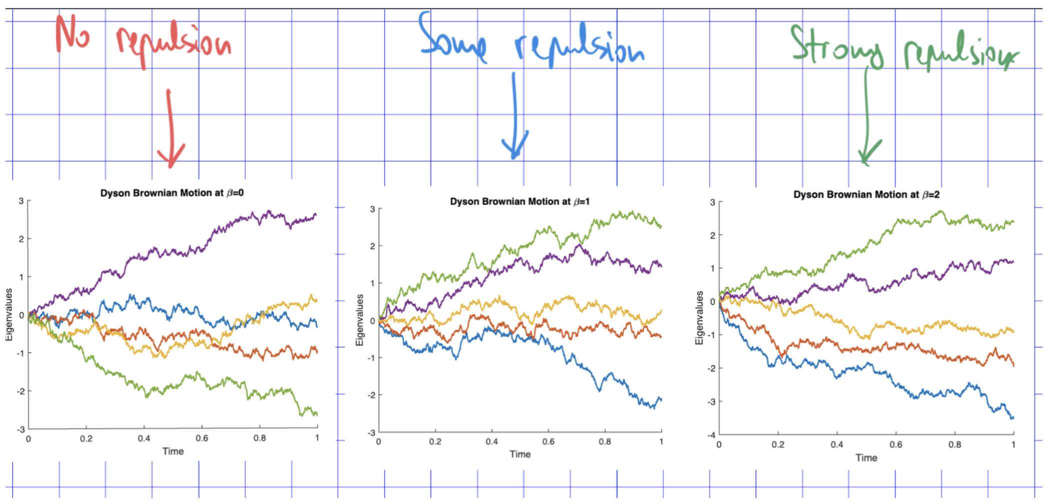
Start from II

Theorem 10.4. *$\lambda(t)$ solves an SDE*

$$\mathrm{d}\lambda_i(t) = \frac{\beta}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{\lambda_i(t)-\lambda_j(t)} + \mathrm{d}W_i(t), \quad 1\le i\le N \tag{$*$}$$

where $W_1,\dots W_N$ are indep Brownian motions. This solution is called Dyson Brownian Motion.

Remark 10.5. *$(*)$ makes sense for all $\beta>0$. $\frac{\beta}{2}$ represents the strength of repulsion, and it's harder to make sense for small β.*



Sketch of the proof. Deal with $\beta=1$. As in the previous thm

$$\begin{aligned}
\lambda(T+\Delta t)_{\Delta t\ge 0} &= \text{e.v.}\left[\operatorname{diag}(\lambda_1(t),\dots,\lambda_N(t)) + \text{Matrix at time }\Delta t\right]\\
&= \text{e.v.}\begin{pmatrix}
\lambda_1(t)+B_1(\Delta t) & \frac{1}{\sqrt{2}}B_{12}(\Delta t) & \cdots & \frac{1}{\sqrt{2}}B_{1n}(\Delta t)\\
\frac{1}{\sqrt{2}}B_{12}(\Delta t) & \lambda_2(t)+B_2(\Delta t) & \cdots & \frac{1}{\sqrt{2}}B_{2n}(\Delta t)\\
\cdots & \cdots & \cdots & \cdots\\
\frac{1}{\sqrt{2}}B_{1n}(\Delta t) & \frac{1}{\sqrt{2}}B_{2n}(\Delta t) & \cdots & \lambda_n(t)+B_n(\Delta t)
\end{pmatrix}
\end{aligned}$$

Assume $\lambda=\lambda_i(t)+\Delta\lambda$ is the eignevalue of $U(t)$, we have

$$\begin{aligned}
0=\det(U(t)) &= \prod_{m=1}^{N}(\lambda_m(t)-\lambda_i(t)+B_m(\Delta t)-\Delta\lambda)\\
&\quad + \sum_{j\neq i}\prod_{m\neq j,i}(\lambda_m(t)-\lambda_i(t)+B_m(\Delta t)-\Delta\lambda)\cdot\left(\frac{1}{2}B_{ij}(\Delta t)\right)^2 + o(\Delta t)
\end{aligned}$$

$$\Longrightarrow \Delta\lambda = \Delta t\cdot\frac{1}{2}\sum_{j\neq i}\frac{1}{\lambda_i(t)-\lambda_j(t)} + B_i(\Delta t) + \frac{1}{2}\sum_{j\neq i}\frac{B_{ij}^2(\Delta t)-\Delta t}{\lambda_i(t)-\lambda_j(t)} + o(\Delta t).$$

In this matrix, $N-1$ diagonal elements are of const order, and the rest is very small as $\Delta t\to 0$. We expect $\Delta\lambda\sim O(\Delta t)$ Scaling $\Delta t\to 0$, we get the desired SDE. $\square$

Proposition 10.6. *For each $\beta > 0$, the SDE $(*)$ has a (unique) solution, such that at each fixed $t > 0$, $(\lambda_1(t), \dots, \lambda_N(t))$ has distribution of density*

$$\sim \prod_{i<j} |\lambda_i - \lambda_j|^\beta \cdot \exp(-\frac{1}{2t}\sum_{i=1}^N \lambda_i^2) \qquad (**)$$

where $\lambda_1(0) = \cdots = \lambda_N(0) = 0$

For $\beta = 1, 2, 4$, this follows from **Lemma** 10.1 and **Theorem** 10.4 For other $\beta > 0$, we need to

- make sense of SDE and its solutions.
- make a formal computation checking that the density $(**)$ is preserved under $(*)$.

E.g. in $N = 1$ case. $\mathrm{d}\lambda = \mathrm{d}W(t)$ is a Markov process and you want to show that densities

$$\rho(t, \lambda) = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{1}{2t}\lambda^2)$$

are preserved. For that you need a generator of a Markov process which is $\frac{1}{2}\frac{\partial^2}{\partial\lambda^2}$.

Here, we need to check that

$$\frac{\partial}{\partial t}p(t, \lambda) = \frac{1}{2}\frac{\partial^2}{\partial\lambda^2}p(t, \lambda).$$

For $\beta = 2$, one can go much further.

Theorem 10.7. *For $\beta = 2$, DBM has the following transition probabilities*

$$P(\lambda(t) = \boldsymbol{x} \mid \lambda(0) = \boldsymbol{a}) = N! \left(\frac{1}{\sqrt{2\pi t}}\right)^N \prod_{i<j} \frac{x_i - x_j}{a_i - a_j} \det\left[\exp\left(-\frac{(x_i - a_j)^2}{2t}\right)\right]_{i,j=1}^N$$

Proof. $\lambda(t)$ are e.v. of $X = A + \sqrt{t}\,GUE$. The law of X has density$\sim \exp(-\frac{\mathrm{Tr}(X-A)^2}{2t})$

A is deterministic, however, for any unitary U, UAU^* will lead to the same e.v. of A and X. Hence, the density

$$\begin{aligned}
&\int_{U\in U(\lambda)} \exp\left(-\frac{\mathrm{Tr}(X - UAU^*)^2}{2t}\right) \mathrm{d}U \\
=& \int_U \exp\left(-\frac{\mathrm{Tr}\, X^2}{2t} + \frac{\mathrm{Tr}(XUAU^*)}{t} - \frac{\mathrm{Tr}(UAU^*)^2}{2t}\right) \mathrm{d}U
\end{aligned}$$

Integral over U is now HCTZ formula from Theorem 6.18. Hence, the density of X is

$$\sim \exp\left(-\frac{\sum x_i^2}{2t} - \frac{\sum a_i^2}{2t}\right) \cdot \det\left[\exp\left(\frac{x_i a_j}{t}\right)\right] \prod_{i<j} \frac{1}{(x_i - x_j)(a_i - a_j)}$$

To get the density of its eigenvalues, we use the Jacobian from Lecture 1 and need to multiply with $\prod_{i<j}(x_i - x_j)^2$. ☐

Theorem 10.8. *Let $x_1, \dots, x_N$ be coordinates at time t for $\beta = 2$ DBM started at $(a_1, \dots, a_N)$ at time $t = 0$. Then $\{x_i\}$ form a determinental pt process with kernel*

$$K(x, y) = \frac{1}{(2\pi i)^2 t} \oint\oint \frac{\exp(\frac{w^2 - 2yw}{2t})}{\exp(\frac{z^2 - 2xz}{2t})} \prod_{i=1}^N \frac{w - a_i}{z - a_i} \frac{\mathrm{d}w\, \mathrm{d}z}{w - z}$$

Remark 10.9.
- *If $a_1 = \cdots = a_N = 0$ and $t = 1$, we reproduce the correlation functions in Lecture 5.*
- *Can be used for study of BBP phase transition for $\beta = 2$ by setting $(a_1, \cdots, a_N) = (\sqrt{N}\alpha_1, \cdots, \sqrt{N}\alpha_K, 0^{N-K})$.*
- *University of RM statistics (next lecture).*

Lemma 10.10. *The density of $(x_1, \ldots, x_N)$ can be expressed as*

$$\lim_{s\to\infty}\left(\frac{1}{z}\det[P_t(a_i \to x_j)]_{i,j=1}^N \det[P_t(x_j \to K-1)]_{j,K=1}^N\right)$$

where $P_t(x \to y)$ is the transition prob of BM $\frac{1}{\sqrt{2\pi t}}\exp(-\frac{(x-y)^2}{2t})$

Proof. We match with Theorem 10.8 by computing the 2nd det under the limit. (1st one is already there)

$$\begin{aligned}
&\det\left[\exp\left(-\frac{x_j^2}{2s} - \frac{x_j(K-1)}{s} - \frac{(K-1)^2}{2s}\right)\right]_{j,K=1}^N \\
\sim &\exp\left(-\frac{\sum x_j^2}{2s}\right)\det\left[\exp\left(\frac{x_j(K-1)}{s}\right)\right] = \exp\left(-\frac{\sum x_j^2}{2s}\right)\prod_{i<j}\left(e^{\frac{x_i}{s}} - e^{\frac{x_j}{s}}\right) \\
\sim &\prod(x_i - x_j) \qquad \text{as } s \to \infty
\end{aligned}$$

□

Proof sketch of ***Theorem 10.8****.* Using the result of Biorthogonal ensemble, we can get

$$K(x,y) = \lim_{s\to\infty}\sum_{i,k=1}^N P_t(a_i \to x)P_s(x \to k-1)[G^{-\top}]_{ik}.$$

□

Theorem 10.11. *For $\beta = 2$, DBM, $\lambda(t)$, started from $\lambda(0) = (a_1, \cdots, a_N)$, coincides in law with N independent BM started from $\lambda(0) = (a_1, \cdots, a_N)$ and condition on never intersect.*

Remark 10.12. *Two independent BM started from arbitrary $a_1 \leq a_2$ almost surely intersect.*

Definition 10.13. *Choose $a_1 < \cdots a_N$. For each $T > 0$, let $\lambda^T(t)$ be BM conditioned on no intersections until time T and $B_i(T) = i - 1$.*

Lemma 10.14. *λ^T is a Markov process with translation density*

$$\mathbb{P}(\lambda(t+\Delta t) = x \mid \lambda(t) = y) = \frac{\det[P_{\Delta t}(y \to x)]\det[P_{T-t-\Delta t}(x \to (0,1,\cdots,N-1))]}{\det[P_{T-t}(y \to (0,1,\cdots,N-1))]}.$$

11 Apr 8th

11.1 Universality

Definition 11.1. *Hermitioan i.i.d. Wigner $W = (w_{ij})$ is a $N \times N$ complex Hermitian matrix, such that $\operatorname{Re} w_{ij}, \operatorname{Im} w_{ij}, i < j$ and w_{ii} are indep with*

$$\operatorname{Re} w_{ij} \quad i.i.d. \quad \mathbb{E}[\operatorname{Re} w_{ij}] = 0 \quad \mathbb{E}[\operatorname{Re} w_{ij}]^2 = \frac{1}{2}$$

$$\operatorname{Im} w_{ij} \quad i.i.d. \quad \mathbb{E}[\operatorname{Im} w_{ij}] = 0 \quad \mathbb{E}[\operatorname{Im} w_{ij}]^2 = \frac{1}{2}$$

$$w_{ii} \quad i.i.d. \quad \mathbb{E}[w_{ij}] = 0 \quad \mathbb{E}[w_{ij}]^2 = \sigma^2$$

We address 3 regimes for GUE

- Global: semicircle law
- Bulk limits: local lim by sine process
- Edge lim: Airy/TW lim

Under mild restrictions all 3 extend to general Wigner matrices. However our methods, based on DPP fail. Tnstead we use comparison methods. Treat general Wigner case as a deformation of GUE case and argue that nothing changes. (GUE is still improtant)

3 approaches

- Lindeberg swapping method
- Moment's method
- Dyson BM

Theorem 11.2 (universality of semicircle)**.** *Hermitian Wigner matrix W, the empirical dist of the e.v. satisfies*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{\frac{\lambda_i}{\sqrt{N}}} = \textit{semicircle law of density } \frac{1}{2\pi}\sqrt{4-x^2}$$

We only prove a weaker result

Proposition 11.3. *Assume $\exists\ \mathbb{E}|w_{ij}|^2 < \infty, \mathbb{E}|w_{ii}|^3 < \infty$. Then*

$$\mathbb{E}\left[\int \frac{1}{z-x}\left(\frac{1}{N}\sum_{i=1}^{N} \delta_{\frac{\lambda_i}{\sqrt{N}}}\right)\right] = \mathbb{E}[G_N(z)] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} \frac{1}{z-\frac{\lambda_i}{\sqrt{N}}}\right]$$
$$\xrightarrow{N\to\infty} \int \frac{1}{z-x}\left(\frac{1}{2\pi}\sqrt{4-x^2}\right) = \frac{1}{2}(z-\sqrt{z^2-4}) = G(z)$$

As we saw in Lecture 2, $\int \frac{1}{z-x}\mu(\mathrm{d}x)$ uniquely determines μ. Hence, this fixes semicircle. For Theorem 11.2, we additionally need $\operatorname{Var}(\frac{1}{N}\sum_{i=1}^{N} \frac{1}{z-\frac{\lambda_i}{\sqrt{N}}}) \to 0$ for concentration.

Strategy: For $W = GUE$, i.e. Gaussian W_{ij}, we already know this from Lecture 2, we replace matrix elements one by one from GUE to W_{ij}, control the change of $\mathbb{E}[G_N(z)]$ and show that after all steps, the total change is $O(N^{-1/2})$. Hence, GUE limit = general Wigner limit.

History: Approach intended by Lindeberg in order to prove CLT. It is trivial, because $\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \xi_i \sim \mathscr{N}(0,1)$. Then replace $\mathscr{N}(0,1)$ to desired distribution one by one for each ξ_i.

Implementation: Introduce resolvent of A

$$r_A(z) = \frac{1}{A - z \cdot \mathrm{Id}}$$

Note that

$$G_N(z) = -\frac{1}{N} \operatorname{Tr} r_{\frac{W}{\sqrt{N}}}(z)$$

Lemma 11.4 (Resolvent identity)**.**

$$r_B(z) = r_A(z) + r_A(z) + (A - B) r_B(z)$$

$$r_B(z) = r_A(z) + \sum_{k=1}^{m} (r_A(z)(A-B))^k r_A(z) + (r_A(z)(A-B))^{k+1} r_B(z)$$

Proof. We need

$$\frac{1}{B-Z} = \frac{1}{A-Z} + \frac{1}{A-Z}(A-B)\frac{1}{B-Z}$$

which is obviouslu correct. 2nd is the iteration of the 1st. □

Proof of Prop. Only prove for off-diagonal elemets. Start with GUE, replace matrix elements by w_{ij} in order (left to right and up to down).

Define M^{ij}: Hermitian matrix in which all elements up to (i,j) were replaced by $\frac{w_{ij}}{\sqrt{N}}$; $(i,j),(j,i)$ elements are 0; and all elements after (i,j) are $\frac{g_{ij}}{\sqrt{N}}$ (g_{ij} is elements of GUE). Similarly, define A^{ij} except at $(i,j) \to \frac{g_{ij}}{\sqrt{N}}$, and at $(j,i) \to \frac{w_{ij}}{\sqrt{N}}$.

We need

$$\sum_{i \le j} \mathbb{E}\left[\frac{1}{N} \operatorname{Tr}\left(r_{A^{ij}}(z) - r_{B^{ij}}(z)\right)\right] = O\left(\frac{1}{\sqrt{N}}\right) \quad \text{for all } z \text{ with } \operatorname{Im} z \neq 0$$

Equivalently,

$$\sum_{i \le j} \mathbb{E}\left[\operatorname{Tr}\left(r_{A^{ij}}(z) - r_{B^{ij}}(z)\right)\right] = O\left(N^{-\frac{3}{2}}\right) \quad \text{for all } z \text{ with } \operatorname{Im} z \neq 0$$

We get rid of i, j, z in the notation and write

$$r_A = r_M + r_M(M-A)r_M + r_M(M-A)r_M(M-A)r_M + [r_M(M-A)]^3 r_A$$

$$r_B = r_M + r_M(M-B)r_M + r_M(M-B)r_M(M-B)r_M + [r_M(M-B)]^3 r_B$$

$$\mathbb{E}[r_A - r_B] = 0 + 0 + 0 + \mathbb{E}\left[\right][r_M(M-A)]^3 r_A\big] - \mathbb{E}\left[\right][r_M(M-B)]^3 r_B\big] .$$

So we only need

$$\mathbb{E} \operatorname{Tr}\left[[r_M(M-A)]^3 r_A - [r_M(M-B)]^3 r_B\right] = O\left(N^{-\frac{3}{2}}\right)$$

Lemma 11.5. *Let Y be a matrix of rank 1 or 2, then*

$$|\operatorname{Tr} Y| \le 2||Y|| = 2\sqrt{\lambda_{max}(YY^*)}$$

Proof. Decomposition

$$Y = UDV \quad 0 \le d_1 \le d_2 = ||Y||$$

$$|\operatorname{Tr} Y| = |\sum_i (u_{i1} d_1 v_{1i} + u_{i2} d_2 v_{2i})| \le |d_1| + |d_2| \le 2||Y||$$

□

We use the lemma and note

$$||r_M|| \le \frac{1}{|\operatorname{Im} z|}$$

because r_M is a normal operator with e.v. $\frac{1}{z-x}$ for real x. Hence we have

$$||[r_M(M-A)]^3 r_A|| \le \frac{1}{|\operatorname{Im} z|^4}||M-A||^3 = O\left(N^{-\frac{3}{2}}\right)$$

For diagonal elements replacement, $\frac{g_{ii}}{\sqrt{N}} \to \frac{w_{ii}}{\sqrt{N}}$, in $\mathbb{E}[r_A - r_B]$, the 3rd term is not 0. But it is small enough, so that because there are only N diagonal elements, the overall bound still works. $\square$

How far can this go?

- We can move from $\mathbb{E}G_N(z)$ to the dist of $G_N(z)$ by computing $\mathbb{E}[f(G_N(z))]$: $G_N(z)$ changes just a little bit, hence for smooth f, by Taylor expansion $f(G_N(z))$ also change just a little bit.
- you can take z approaching real axis as $N \to \infty$. This eventually gives access to local statistics. But need to match more moment to GUE.

Next we come to spikes.

Theorem 11.6 (Bai-Yin theorem)**.** *Assume $\mathbb{E}|w_{ij}|^4 < \infty$ and let $N \times N$ Hermitian Wigner matrix W. Then*

$$\lim_{N\to\infty} \frac{\lambda_N}{\sqrt{N}} = 2 \quad \textit{in prob and a.s.}$$

It is known that if the 4th moment is infinite, then

$$\limsup_{N\to\infty} \left(\frac{\lambda_N}{\sqrt{N}}\right) > 2 \quad \text{a.s.}$$

Sketch of the proof. Assume that w_{ij} are bounded, $|w_{ij}| < c$ a.s. Claim that suppose $K = K(N)$ grows with N, then

$$\mathbb{E}\operatorname{Tr}\left(\frac{W_N}{2\sqrt{N}}\right)^K = \begin{cases} \frac{2\sqrt{2}N}{\sqrt{\pi}}K^{-\frac{3}{2}}(1+o(1)) & ,\text{even } k; \\ \text{small order} & ,\text{odd } k; \end{cases}$$

The claim remains true as long as $K = o(N^{\frac{2}{3}})$, but complexity of the proof grows as K becomes larger.

For even K,

$$\mathbb{E}\operatorname{Tr}\left(\frac{W_N}{2\sqrt{N}}\right)^K = \mathbb{E}\sum_{i=1}^N \left(\frac{\lambda_i}{2\sqrt{N}}\right)^K \ge \mathbb{P}\left(\frac{\lambda_N}{2\sqrt{N}} \ge 1+\epsilon\right)(1+\epsilon)^K.$$

So, if $K = N^\delta$ with $\delta > 0$, then the equation implies $\mathbb{P}\left(\frac{\lambda_N}{2\sqrt{N}} \ge 1+\epsilon\right) \to 0$. $\sum_N \mathbb{P}\left(\frac{\lambda_N}{2\sqrt{N}} \ge 1+\epsilon\right) < \infty \implies$ the a.s. convergence.

For the proof of the claim, we use moments

$$\mathbb{E}\operatorname{Tr}(w_N)^K = \sum_{i_1,\dots,i_K} \mathbb{E}W_{i_1,i_2}W_{i_2,i_3}\cdots W_{i_K i_1} \qquad (**)$$

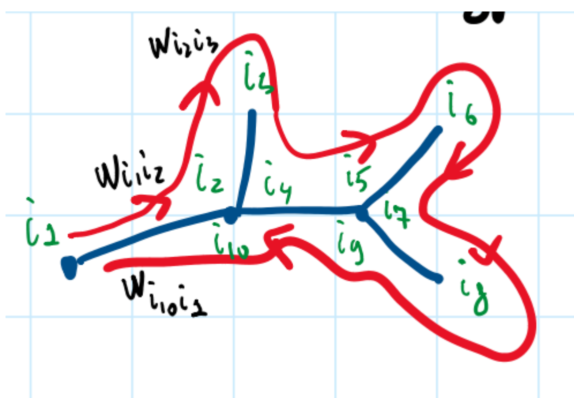
Key obsevation: If W_{ij} is present in the product, but there is no other w_{ij}/w_{ji}, then $\mathbb{E} = 0$ because $\mathbb{E}w_{ij} = 0$ and indep.

Moreover, $\mathbb{E}(w_{ij})^2 = 0$ as well. Getting twice of the same w_{ij} also gives the $\mathbb{E} = 0$.

Lemma 11.7. *For the leading contribution to the same (**), we only need to consider* $i_1, \dots, i_K$ *s.t.*

- w_{ii} *is not present.*
- *For each* w_{ij}*, there is a single matching* w_{ji} *and usen additional appearance of* w_{ij} *or* w_{ji}*.*

Lemma 11.7 is a hard part, whose complexity grows with K. And we omit it. The terms of Lemma 11.7 can be encoded by trees.



Indices on the same vertex coincide:

$$\mathbb{E}[W_{i_1 i_2} W_{i_2 i_3} \cdots W_{i_N i_1}] = 1.$$

It remains to count the trees: $N^{\frac{K}{2}+1}$ planar rooted trees on $\frac{K}{2}+1$ vertices $= N^{\frac{K}{2}+1} \frac{1}{\frac{K}{2}+1}\binom{K}{\frac{K}{2}}$. Hence

$$\mathbb{E}\,\mathrm{Tr}(W_N)^K = N^{\frac{K}{2}+1} \frac{4^{\frac{K}{2}}}{(\frac{K}{2})^{3/2}\sqrt{\pi}}(1+o(1)).$$

$\square$

Remark 11.8.

- *Under the same conditions as Theorem 11.6 in fact*

$$N^{-1/6}(\lambda_N - 2\sqrt{N}) \to TW_2 \quad \textit{as } N \to \infty.$$

- *Under additional assumptions (e.g.* $|w_{ij}| < C$*), one can prove* TW_2 *limit by high moments*

$$\mathrm{Tr}(W_N)^K, K = \tau N^{2/3}.$$

In the bulk, we do not even need the 4th moment.

Theorem 11.9. *Suppose that* $\mathbb{E}|w_{ij}|^{2+\delta} < \infty$ *for some* $\delta > 0$*. Let* $\lambda_1 \le \cdots \le \lambda_N$ *be eigenvalues of matrix* W_N *take* $-2 \le s \le 2$*. Then*

$$\left\{\sqrt{N}(\lambda_i - s\sqrt{N})\right\} \to \textit{sine process of intensity } \frac{1}{2\pi}\sqrt{4-s^2}.$$

Remark 11.10.

- $\mathbb{E}|w_{ij}|^{2+\delta}$ *is believed not to be necessary,* $\mathbb{E}|w_{ij}|^2$ *should be sufficient.*
- $\mathbb{E}|w_{ij}| < \infty$ *is in fact sufficient if we change the scaling.*
- *Even* $\mathbb{E}|w_{ij}|^{\delta} < \infty$ *is sufficient in a part of the spectrum.*

Proof. We prove Gaussian divisible case.

$$W_{ij} = V_{ij} + \sqrt{t} g_{ij}.$$

We assume $\mathbb{E}|V_{ij}|^4 < \infty$, so that we can use Theorem 11.6. We need $\mathbb{E}|V_{ij}|^2 = 1-t$ for $\mathbb{E}|W_{ij}|^2 = 1$ and t can be arbitrary small.

Note $W = V + \sqrt{t} GUE$, V has eigenvalues $(a_1, \cdots, a_N)$. Recall Theorem 10.8

Theorem 11.11. *Let $x_1, \ldots, x_N$ be coordinates at time t for $\beta = 2$ DBM started at $(a_1, \ldots, a_N)$ at time $t = 0$. Then $\{x_i\}$ form a determinental pt process with kernel*

$$K(x,y) = \frac{1}{(2\pi i)^2 t} \oiint \frac{\exp(\frac{w^2 - 2yw}{2t})}{\exp(\frac{z^2-2xz}{2t})} \prod_{i=1}^{N} \frac{w - a_i}{z - a_i} \frac{\mathrm{d}w \, \mathrm{d}z}{w - z}$$

Eigenvalues of W_N form a DPP with an increasing kernel, conditioning on $(a_1, \cdots, a_N) =$ e.v. of V.

Strategy: Analyze conditionally on (a_i), hope that only limits of (a_i) which we know enter into the answers.

Steepest descent: Let

$$x \to \sqrt{N} s + \frac{\Delta x}{\sqrt{N}}$$
$$y \to \sqrt{N} s + \frac{\Delta y}{\sqrt{N}}$$
$$w \to \sqrt{N} w$$
$$z \to \sqrt{N} z$$

and the kernel becomes

$$K(x,y) = \frac{\sqrt{N}}{(2\pi i)^2 z} \oiint \frac{\exp(N(\frac{w^2}{2t} - \frac{2sw}{2t} + \frac{1}{N}\sum_i \log(w - \frac{a_i}{\sqrt{N}})))}{\exp(N(\frac{z^2}{2t} - \frac{2sz}{2t} + \frac{1}{N}\sum_i \log(z - \frac{a_i}{\sqrt{N}})))} \cdot \frac{\Delta y z - \Delta x w}{t} \frac{\mathrm{d}w \, \mathrm{d}z}{w - z}.$$

We investigate

$$F(w) = \frac{w^2}{2t} - \frac{2sw}{2t} + \frac{1}{N} \sum_i \log\left(w - \frac{a_i}{\sqrt{N}} \right).$$

By the semisircle law,

$$0 = F'(w) = \frac{w}{t} - \frac{s}{t} + \frac{1}{N} \sum_i \frac{1}{w - \frac{a_i}{\sqrt{N}}} \approx \frac{w}{t} - \frac{s}{t} + \frac{1}{2(1-t)} \left(w - \sqrt{w^2 - 4(1-t)} \right).$$

$$\implies w_c = \frac{s(2-t) \pm t\sqrt{s^2 - 4}}{2}.$$

When $|s| < 2$, there are two complex conjugate critical points. Deform the contours to pass through them. Similar analysis... (Additional step to be careful about: (a_i)-dependent steepest descent contours). □

Remark 11.12. *Further developments using DBM methods:*

- *Make t smaller and smaller (until it is so small and Gaussian divisibility condition no longer needed).*
- *Develop alternative analysis of DBM not using DPP (and so extending beyond $\beta = 2$).*

12 Apr 15th

12.1 CLT and GFF

Recall Semicircle Law: $\lambda_1 \leq \cdots \leq \lambda_N$ are eigenvalues of $\sqrt{\frac{2}{\beta}}G\beta E$.

Theorem 12.1. *Let f be analytic in a small neighboring of $[-2,2]$. Let $\lambda_1 \leq \cdots \leq \lambda_N$ be e.v. of $\sqrt{\frac{2}{\beta}}G\beta E, \beta > 0$. Then*

$$\sum_{i=1}^{N} f\left(\frac{\lambda_i}{\sqrt{N}}\right) - N \int f(x) \frac{1}{2\pi}\sqrt{4-x^2}$$

converges to dist to a Gaussian r.v. ξ_f jointly other several $f = f_1, \dots, f_K$ with

$$\mathbb{E}[\xi_f] = \left(\frac{2}{\beta} - 1\right) m(f), \quad \mathrm{Cov}(\xi_f, \xi_g) = \frac{2}{\beta} C(f,g)$$

$$m(f) = \frac{1}{4}(f(2) + f(-2)) - \frac{1}{2\pi}\int_{-2}^{2} \frac{f(x)}{\sqrt{4-x^2}}\,\mathrm{d}x$$

$$C(f,g) = \frac{1}{4\pi}\int_{-2}^{2}\int_{-2}^{2} \frac{(f(x)-f(y))(g(x)-g(y))}{(x-y)^2} \frac{4-xy}{\sqrt{4-x^2}\sqrt{4-y^2}}\,\mathrm{d}x\,\mathrm{d}y.$$

or

$$C\left(\frac{1}{z-x}, \frac{1}{w-x}\right) = -\frac{1}{2(z-w)^2}\left(1 - \frac{zw-4}{\sqrt{z^2-4}\sqrt{w^2-4}}\right) \quad \textit{for } z, w \textit{ out of } [-2,2]$$

where $m(f)$ and $C(f,g)$ do not depend on β.

- In contrast to edge, β-dependence is simple.
- Gaussian limits, no new distributions.
- $m(f), C(f,g) \to$ research direction.
- For other ensembles, similar β-dependence, we can give similar $m(f)$ and $C(f,g)$.
- Optimal results: for all f with $C(f,f) < \infty$ is enough. But for $f(x) = I_{x<0}$, the scaling is wrong, $\mathrm{Var}(\sum_i f(\frac{\lambda_i}{\sqrt{N}}))$ grows logarithmically.

Proof. The main idea is fancy moments method.

Question: Let $\eta_1, \dots, \eta_m$ be jointly Gaussian r.v. with $\mathbb{E}[\eta_i] = 0$ and $\mathbb{E}[\eta_i\eta_j] = \sigma_{ij}$, then what is $\mathbb{E}[\eta_1 \cdots \eta_m]$?

Lemma 12.2 (Wick's formula)**.**

$$\mathbb{E}[\eta_1, \dots, \eta_m] = \sum_{\text{perfectmatchings}} \prod_{(i,j)\in\text{mathching}} \sigma_{ij}$$

Proof. By Laplacian's transform

$$\mathbb{E}[e^{t_1\eta_1 + \cdots + t_m\eta_m}] = \exp\left(\frac{1}{2}\sum_{i,j=1}^{m} t_i t_j \sigma_{ij}\right).$$

Hence, we take derivative and get the expectation

$$\mathbb{E}[\eta_1 \cdots \eta_m] = \frac{\partial^m}{\partial t_1 \cdots \partial t_m} \left[\exp\left(\frac{1}{2} \sum_{i,j=1}^{m} t_i t_j \sigma_{ij} \right) \right] \Big|_{t_1 = \cdots = t_m = 0}.$$

Equivalently, this is the coefficient of $t_1 \cdots t_m$ in Taylor's expansion.

$$= \sum_{ij} 2^{-\frac{m}{2}} \prod_{ij} \sigma_{ij}.$$

Notice that each perfect match appears $2^{\frac{m}{2}}$ times, we prove the Lemma 12.2 $\square$

Takeaway: Moments can be reconstructed by applying a differential operator to Laplace transform.

We do the same for GUE.

Lemma 12.3. *Introduce an operator*

$$\mathscr{D}_a = \prod_{i,j} (z_i - z_j)^{-1} \left(\sum_{i=1}^{N} T_{a,i} \right) \prod_{i,j} (z_i - z_j)$$

$$T_{a,i} f(z_1, \ldots, z_N) = f(z_1, \ldots, z_{i-1}, z_i + a, z_{i+1}, \ldots, z_N)$$

Then for $\lambda_1, \cdots, \lambda_N$ eigenvalues of GUE, we have

$$\underbrace{\mathbb{E} \prod_{k=1}^{M} \left[\sum_{i=1}^{N} e^{a_k \lambda_i} \right]}_{\text{Moments of linear statistic for } f(\lambda) = e^{a\lambda}} = \underbrace{\mathscr{D}_{a_m} \cdots \mathscr{D}_{a_1}}_{\text{They all commute}} \exp\left(\sum_{i=1}^{N} \frac{z_i^2}{2} \right) \Big|_{z_1 = \cdots = z_N = 0}.$$

Proof. From Lecture 6,

$$\mathbb{E} \exp(\mathrm{Tr}(\mathrm{GUE} \cdot \mathrm{Z})) = \mathbb{E} B_{\lambda_1, \ldots, \lambda_N}(z_1, \ldots, z_N) = \exp\left(\sum_{i=1}^{N} \frac{\lambda_i^2}{2} \right)$$

where $Z = \mathrm{diag}(z_1, \ldots, z_N)$ and $\lambda_1, \ldots, \lambda_N$ denote the e.v. of GUE.

Then act with $\mathscr{D}$

$$B_{\lambda_1, \ldots, \lambda_N}(z_1, \ldots, z_N) = \prod_{K=1}^{N} (K!) \cdot \frac{\det[\exp(\lambda_i z_j)]}{\prod_{i,j} (\lambda_i - \lambda_j)(z_i - z_j)}.$$

$$\mathscr{D} = \prod_{i,j} (z_i - z_j)^{-1} \left(\sum_{i=1}^{N} T_{a,i} \right) \prod_{i,j} (z_i - z_j).$$

So B is eigenfunction of $\mathscr{D}$ with eigenvalue $\sum_{i=1}^{N} \exp(a\lambda_i)$. Hence,

$$\mathbb{E} \left[\prod_{k=1}^{m} \left(\sum_{i=1}^{N} e^{a_k \lambda_i} \right) \right] B_{\lambda_1 \cdots \lambda_N}(z_1, \cdots, z_N) = \mathscr{D}_{a_m} \cdots \mathscr{D}_{a_1} \exp\left(\sum_{i=1}^{N} \frac{\lambda_i^2}{2} \right).$$

Plug $z_1, \ldots, z_N = 0$ and notice $B(0, \cdots, 0) = 1$. $\square$

Lemma 12.4.

$$\mathscr{D} f(z_1)\cdots f(z_N) = f(z_1)\cdots f(z_N)\frac{a^{-1}}{2\pi i}\oint_{\{z_1,\dots,z_N\}}\left[\prod_{j=1}^{N}\frac{v+a-z_j}{v-z_j}\right]\frac{f(v+a)}{f(v)}$$

The contour encloses $z_1,\dots,z_N$ with no singularities of f.

Proof.

$$\begin{aligned}\mathscr{D} f(z_1)\cdots f(z_N) &= \sum_{i=1}^{N}\left[\prod_{j\neq i}\frac{z_i+a-z_j}{z_i-z_j}\right]\frac{f(z_i+a)}{f(z_i)}\cdot f(z_1)\cdots f(z_N)\\ &= \frac{a^{-1}}{2\pi i}\oint_{\{z_1,\dots,z_N\}}\prod_{i=1}^{N}\frac{v+a-z_j}{v-z_j}\cdot\frac{f(v+a)}{f(v)}\,\mathrm{d}v f(z_1)\cdots f(z_N)\end{aligned}$$

□

Corollary 12.5.

$$\begin{aligned}\mathbb{E}\left[\prod_{K=1}^{m}\sum_{i=1}^{N}e^{a_K\lambda_i}\right] =&\frac{(a_1\cdots a_m)^{-1}}{(2\pi i)^m}\oint\sum_{k=1}^{m}\left[\frac{(v_k+a_k)^N}{v_k^N}\exp\left(\frac{a_k^2}{2}+a_kv_k\right)\right]\\ &\prod_{k<l}\frac{v_k-v_l+a_k-a_l}{v_k-v_l-a_l}\cdot\frac{v_k-v_l}{v_k-v_l+a_k}\,\mathrm{d}v_1\cdots\mathrm{d}v_k.\end{aligned}$$

Proof. We compute $\mathscr{D}_{a_m}\cdots\mathscr{D}_{a_1}\exp(\frac{z_i^2}{2})\cdots\exp(\frac{z_N^2}{2})$. By sequentially applying Lemma 12.4 and then setting $z_1=\cdots=z_N=0$ at the end □

Now we prove Theorem 1 for $\beta=2, f(\frac{\lambda}{\sqrt{N}})=\exp(a\cdot\frac{\lambda}{\sqrt{N}})$
Step 1: Expectation. By Cor with $m=1$

$$\mathbb{E}\left[\sum\exp\left(a\cdot\frac{\lambda_i}{\sqrt{N}}\right)\right] = \frac{(\frac{a}{\sqrt{n}})^{-1}}{2\pi i}\oint_{\{0\}}\left(\frac{v+\frac{a}{\sqrt{N}}}{v}\right)^N\exp\left(\frac{a^2}{2N}+\frac{a}{\sqrt{N}}v\right)\mathrm{d}v$$

No steepest descent needed. Set $v=u\sqrt{N}$ to get

$$\begin{aligned}&\frac{N}{a}\frac{1}{2\pi i}\oint\exp\left(N\log\left(1+\frac{a}{Nu}\right)+au+\frac{2a^2}{N}\right)\\ =&\frac{N}{a}\frac{1}{2\pi i}\oint\exp\left(a\left(u+\frac{1}{u}\right)+\frac{a^2}{2N}\left(1-\frac{1}{u^2}\right)+O(N^{-2})\right)\mathrm{d}u\\ =&\frac{N}{a}\frac{1}{2\pi i}\oint\exp\left(a\left(u+\frac{1}{u}\right)\right)\mathrm{d}u+\frac{a}{4\pi i}\oint\exp\left(a\left(u+\frac{1}{u}\right)\right)\left(1-\frac{1}{u^2}\right)\mathrm{d}u+O(N^{-1}).\end{aligned}$$

We only need to prove

$$\frac{N}{a}\frac{1}{2\pi i}\oint\exp\left(a\left(u+\frac{1}{u}\right)\right)\mathrm{d}u = N\int_{-2}^{2}\exp(ax)\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x.$$

We transform the contour to the unit circle and change variables $x = u + \frac{1}{u}$ $(u = \frac{1}{2}(x \pm i\sqrt{4-x^2}), \mathrm{d}u = \frac{1}{2}(1 \pm i\frac{-x}{\sqrt{4-x^2}})\,\mathrm{d}x)$

$$\begin{aligned}\frac{a^{-1}}{2\pi i}\oint \exp\left(a\left(u+\frac{1}{u}\right)\right)\mathrm{d}u &= \frac{a^{-1}}{2\pi i}\left(\int_{-2}^{2} e^{ax}\frac{1}{2}(1+i\frac{-x}{\sqrt{4-x^2}})\,\mathrm{d}x + \int_{-2}^{2} e^{ax}\frac{1}{2}(1-i\frac{-x}{\sqrt{4-x^2}})\,\mathrm{d}x\right)\\ &= \frac{a^{-1}}{2\pi}\int_{-2}^{2} e^{ax}\frac{x}{\sqrt{4-x^2}}\,\mathrm{d}x \overset{\text{by parts}}{=} \frac{a^{-1}}{2\pi}\int_{-2}^{2}(ae^{ax})\sqrt{4-x^2}\,\mathrm{d}x\end{aligned}$$

Conclusion:

$$\mathbb{E}\left[\sum_{i=1}^{N}\exp\left(a\frac{\lambda_i}{\sqrt{N}}\right)\right] = N\int_{-2}^{2} e^{ax}\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x + O\left(N^{-1}\right)$$

Step 2: Variance.

$$\begin{aligned}&\mathbb{E}\left[\sum_{i=1}^{N}\exp\left(a_1\frac{\lambda_i}{\sqrt{N}}\right)\sum_{i=1}^{N}\exp\left(a_2\frac{\lambda_i}{\sqrt{N}}\right)\right] - \mathbb{E}\left[\sum_{i=1}^{N}\exp\left(a_1\frac{\lambda_i}{\sqrt{N}}\right)\right]\mathbb{E}\left[\sum_{i=1}^{N}\exp\left(a_2\frac{\lambda_i}{\sqrt{N}}\right)\right]\\ =&N\frac{(a_1a_2)^{-1}}{(2\pi i)^2}\oiint\prod_{k=1}^{2}\left(\frac{v_k+\frac{a_k}{\sqrt{N}}}{v_k}\right)^N\exp\left(\frac{a_k^2}{2N}+\frac{a_kv_k}{N}\right)\cdot\left[\frac{v_1-v_2+\frac{a_1}{\sqrt{N}}-\frac{a_2}{\sqrt{N}}}{(v_1-v_2-\frac{a_2}{\sqrt{N}})(v_1-v_2+\frac{a_1}{\sqrt{N}})}-1\right]\mathrm{d}v_1\,\mathrm{d}v_2.\\ \approx&N^2\frac{(a_1a_2)^{-1}}{2\pi i^2}\oiint\exp\left(a_1\left(v_1+\frac{1}{v_1}\right)+a_2\left(v_2+\frac{1}{v_2}\right)\right)\left[\frac{1+\frac{a_1-a_2}{N(u_1-u_2)}}{(1-\frac{a_2}{N(u_1-u_2)})(1+\frac{a_1}{N(u_1-u_2)})}\right]\mathrm{d}u_1\,\mathrm{d}u_2.\\ =&N^2\frac{(a_1a_2)^{-1}}{(2\pi i)^2}\oiint\exp\left(a_1\left(v_1+\frac{1}{v_1}\right)+a_2\left(v_2+\frac{1}{v_2}\right)\right)\left[\frac{a_1a_2}{N^2(u_1-u_2)^2}\right]\mathrm{d}u_1\,\mathrm{d}u_2.\end{aligned}$$

Conclusion: Variance $\to \frac{1}{(2\pi i)^2}\oiint\exp\left(a_1\left(v_1+\frac{1}{v_1}\right)+a_2\left(v_2+\frac{1}{v_2}\right)\right)\frac{\mathrm{d}u_1\,\mathrm{d}u_2}{(u_1-u_2)^2}$.

Match this with formula in Theorem 1. Hint: Either directly with the 1st formula for $C(f,g)$, or

$$\sum_{i=1}^{N} f(\lambda_i) = \frac{1}{2\pi i}\oint_{\text{around all }\lambda_i} f(z)\sum_{i=1}^{N}\frac{1}{z-\lambda_i}\,\mathrm{d}z$$

and use it to compute with $C(\frac{1}{z-x},\frac{1}{w-x})$ in Theorems.

Step 3: Gaussianity. We need to show that

$$\mathbb{E}\left[\prod_{k=1}^{m}\left(\sum_{i=1}^{N}\exp\left(a_k\frac{\lambda_i}{\sqrt{N}}\right)-\mathbb{E}\sum_{i=1}^{N}\exp\left(a_k\frac{\lambda_i}{\sqrt{N}}\right)\right)\right] \to \text{expressions of the Wick's formula in Lemma 12.2.}$$

All of them have exactly the same $\prod_{k=1}^{m}$ part, but cross term part $\prod_{k<l}$ varies. As in Steps 1,2, we change variables $u_k = \sqrt{N}v_k$ and use

$$\text{crossterm} = 1 + \frac{a_ka_l}{N^2(u_k-u_l)^2} + O\left(N^{-3}\right)$$

The summation over 2^m integrals leads to cancellation at parts involving 1. The next term $\to$ perfect matching. $\square$

So far

- $\mathrm{Var}(\sum f(\frac{\lambda_i}{\sqrt{N}})) = o(1)$
- $\sum f(\frac{\lambda_i}{\sqrt{N}})$−asymptotically Gaussian

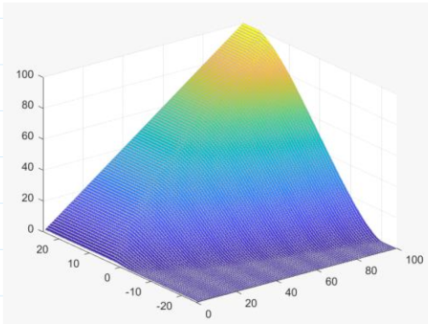
- $\mathrm{Cov}(\sum f(\frac{\lambda_i}{\sqrt{N}}, \sum g(\frac{\lambda_i}{\sqrt{N}})))$ = one of three complicated formulas.

There exist a "formula-less" point of view of (3). Consider $G\beta E$-corners process $\lambda_i^K(1 \le i \le K)$ = eigenvalues of $K \times K$ corners. Introduce Height function

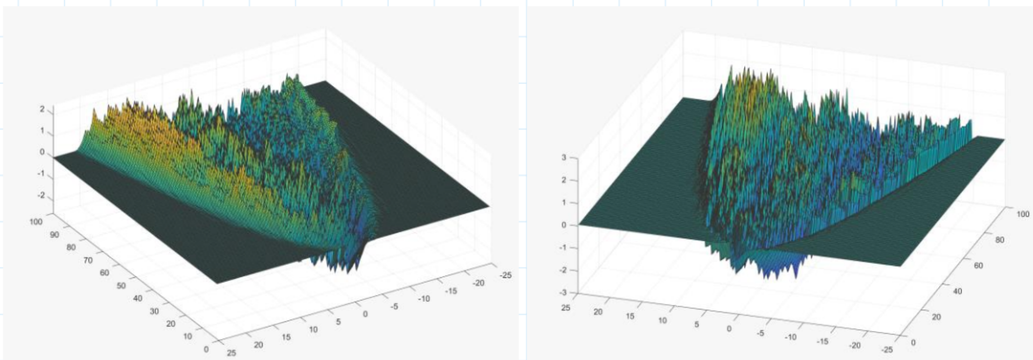
$$H(x, K) = \sum_{i=1}^{K} I(\lambda_i^K < x) = \#\left\{\text{eigenvalues of } K \times K \text{ corner which are} < \lambda\right\}.$$

- $H(x, K)$ is a random function or a random surface.
- As a linear statistic with non-smooth f, it should be more singular.
- Any other linear statistic is obtained by integration by parts:

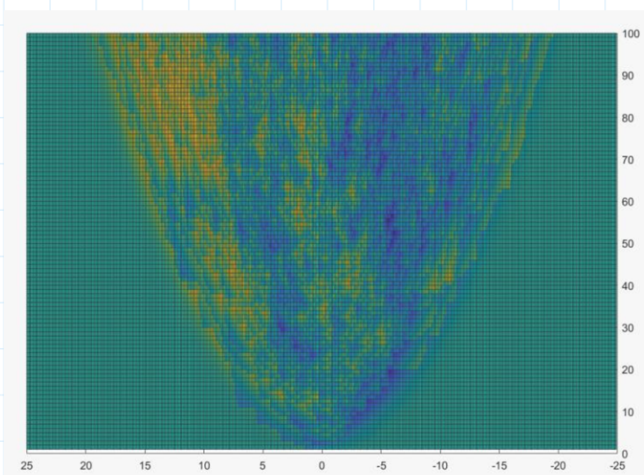
$$\begin{aligned} &-\int_{-\infty}^{\infty} f'(x)H(x, N)\,\mathrm{d}x = -\int f'(x)\sum I(\lambda_i^N < x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} f(x)(\sum I(\lambda_i^N < x))'\,\mathrm{d}x = \sum_{i=1}^{N} f(\lambda_i^N). \end{aligned}$$



Expected Height



H-EH in 3 projections



The 2D random field appearing in the pictures has a name "the Gaussian Free Field" or GFF.

Definition 12.6. *The GFF with Dirichlet boundary condition in the upper half plane* $\mathbb{H} = \{z \in \mathbb{C}|\ \operatorname{Im} z \geq 0\}$ *is a (generalized) Gaussian random field* $\bar{f}$ *on* $\mathbb{H}$ *with covariance given by*

$$\mathbb{E}(\bar{f}(z)\bar{f}(w)) = \frac{1}{2\pi} \ln \left| \frac{z-w}{z-\bar{w}} \right|$$

- *Usually, gaussian field is a random function* $\mathscr{F} : \mathbb{H} \to \mathbb{R}$, *s.t.* $\mathscr{F}(z_1), \ldots, \mathscr{F}(z_n)$ *are Gaussian covariance* (Δ).
- *Then* $\mathbb{E}|f(z)|^2 = +\infty \to$ *reflected by sharp peaks on pictures.*

- *Generalized means that we should not look at individual values, but instead consider pairings with test-measures*
$$\langle \mathscr{F}, \mu \rangle \approx \textit{heuristically} \iint \mathscr{F}\, \mathrm{d}\mu.$$

Formally, $\langle \mathscr{F}, \mu \rangle$ =mean 0 Gaussian r.v. with cov given by
$$\mathbb{E}\langle \mathscr{F}, \mu \rangle \langle \mathscr{F}, \nu \rangle = \iint_{\mathbb{H}} -\frac{1}{2\pi} \ln \left| \frac{z-w}{z-\bar{w}} \right| \mu(\mathrm{d}z)\nu(\mathrm{d}w).$$

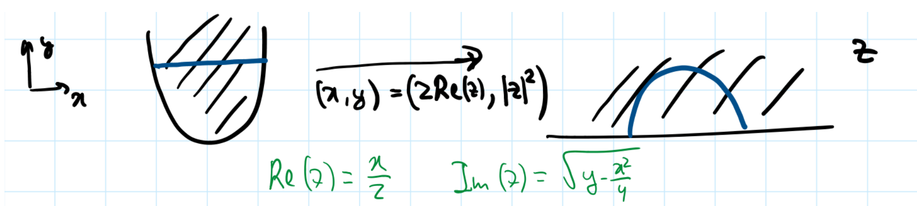
Example: If $z = x + iy$, $\mu(\mathrm{d}z) = \mu(z)\,\mathrm{d}x\,\mathrm{d}y$, then
$$\langle F, \mu \rangle = \iint_{\mathbb{H}} F(z)\mu(z)\,\mathrm{d}z.$$

The covariance $-\frac{1}{2\pi} \ln \left| \frac{z-w}{z-\bar{w}} \right| = C(z, w)$ satisfies

1. If z or w are real, $C(z, w) = 0$ (Dirichlet boundary condition)
2. $z = x + iy$
$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) C(z, w) = -\delta(z = w)$$
(δ function in functional analysis sense) This is Green's functions for the Laplace operator $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ with Dirichlet boundary conditions.
3. $\mathscr{F}$ is conformally invariant: If we change the variables $z \to z' = \frac{a+bz}{c+dz}, a, b, c, d \in \mathbb{R}$, preserving $\mathbb{H}$. It's called Mobius transformation, then $C(z, w)$ is unchanged.
4. $\mathscr{F}$ is a 2d analogue Brownian bridge.

Our GFF lives in $\mathbb{H}$, but e.v. live inside a parabola. Because the support of semicircle law is $[-2\sqrt{K}, 2\sqrt{K}]$.

Definition 12.7. *We introduce a bijection Ω : (interior of parabola) $\to \mathbb{H}$*



Theorem 12.8. *Let $H(x, k)$, $k = 1, 2, \cdots$ be random height function of $G\beta E$ corners process. Then, as $N \to \infty$*
$$\sqrt{\frac{\beta}{2}\pi} \left[H\left(x\sqrt{\frac{\beta N}{2}}, yN \right) - \mathbb{E}\left[H\left(x\sqrt{\frac{\beta N}{2}}, yN \right) \right] \right] \to \Omega - \textit{pullback of GFF in } \mathbb{H}$$

Remark 12.9. *Full proof can be found in General β-Jacobi Corners Process and the Gaussian Free Field.*

13 Apr 22rd

13.1 Random matrices and 2d statistical mechanics

Discrete eigenvalue models of random matrices naturally correspond to two-dimensional statistical mechanics models, such as lozenge tilings.

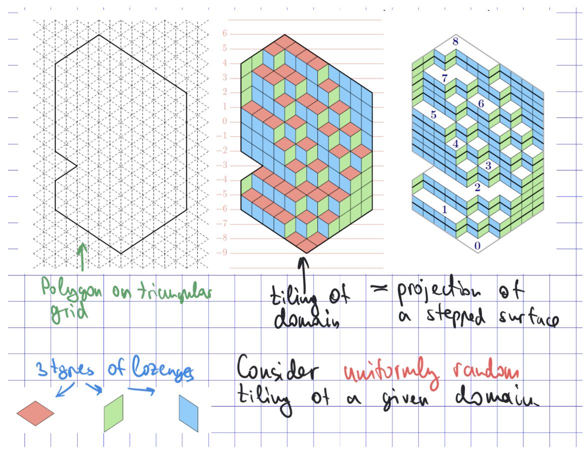


Figure 10: Lozenge Tilings

Motivation:

- Simplest model of stepped surfaces.
- Connects to other stat mech models like 3d Ising model, six vertex model, but tilings are simpler to study.
- Extremely interesting asymptotic.

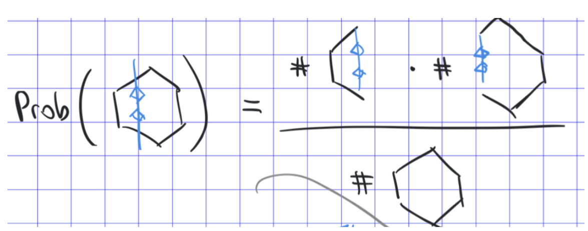
Theorem 13.1. *Consider $2A \times 2A \times 2A$ hexagon. For $N \le 2A$, consider a vertical section at distance N from the left. Then, there are N horizontal lozenges on the line at positions $y \in \{-A - \frac{N}{2}, \cdots, A + \frac{N}{2} - \frac{1}{2}\}$ and their distribution has weight:*

$$\frac{1}{Z}\prod_{i<j}(y_i - y_j)^2 \prod_{i=1}^{N} \frac{(3A - \frac{N}{2} - \frac{1}{2} - y_i)!(3A - \frac{N}{2} - \frac{1}{2} + y_i)!}{(A + \frac{N}{2} - \frac{1}{2} - y_i)!(A + \frac{N}{2} - \frac{1}{2} + y_i)!}$$

Remark 13.2. *Like GUE-eigenvalues, but:*

- $y_i \in \mathbb{Z}$ *or* $\mathbb{Z} + \frac{1}{2}$.
- *More complicated weight than* $e^{-\frac{x^2}{2}}$.

Proof. By some combination analysis, we get there are exact N horizontal lozenges.



By the definition of the Wikipedia page on Schur polynomials, we have

$$\text{The left part of numerator} = S_\lambda(1,\cdots,1) = \prod_{i<j}\left[\frac{(\lambda_j - j)-(\lambda_i - i)}{j-i}\right] = \prod_{i<j}\left[\frac{y_j - y_i}{j-i}\right].$$

The first equal sign comes from properties of Schur polynomials, and the second one comes from $\{\lambda_i - i\}_{i=1}^N = \{y_i - \frac{N}{2}\}_{i=1}^N$.

Similarly, we can complete the right part of nominator into a trapezoid and get

$$\text{The right part of numerator} = \prod_{i<j}\left[\frac{x_j - x_i}{j-i}\right],$$

where $\{x_i\} = \{-3A - \frac{N}{2} + \frac{1}{2}, \cdots, -A - \frac{N}{2} - \frac{1}{2}\} \cup \{A + \frac{N}{2} + \frac{1}{2}, \cdots, 3A + \frac{N}{2} - \frac{1}{2}\} \cup \{y_i\}$.

Combine them together, we get the result. □

Theorem 13.3. *Assume N is fixed, $A \to \infty$. Then*

$$\left\{\frac{y_i}{\sqrt{\frac{3}{4}A}}\right\} \to \textit{GUE - eigenvalues.}$$

Corollary 13.4. *The left corner of the graph corresponds to the Guassian corner process.*

Proof of Theorem 1.3. Analyse the factorial from thm 1, using Stirling formula ($K! = \sqrt{2\pi K}\left(\frac{K}{e}\right)^K$)

Notice that

$$(M-y)!(M+y)! \approx 2\pi M e^{-2M} \exp((M-y)\ln(M-y) + (M+y)\ln(M+y))$$

$$(M-y)\ln(M-y) + (M+y)\ln(M+y) = 2M\ln M + \frac{y^2}{M} + O\left(\frac{y^3}{M^2}\right)$$

so the weight becomes

$$w(y) = C\cdot\exp\left(\frac{y^2}{3A} - \frac{y^2}{A} + o(1)\right) = C\cdot\exp\left(-\frac{y^2}{2}\cdot\frac{4}{3A}\right)$$

Therefore,

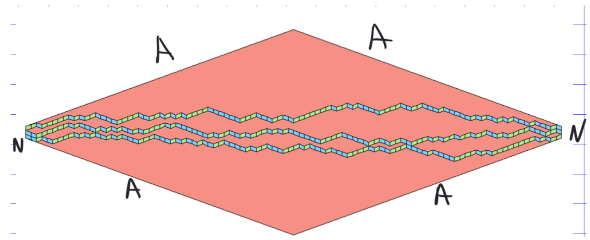
$$\mathrm{Prob}(y_1,\ldots,y_N) = C\cdot\prod_{i<j}\left(\frac{y_i}{\sqrt{\frac{3}{4}A}} - \frac{y_j}{\sqrt{\frac{3}{4}A}}\right)^2\prod_{i=1}^N \exp\left(-\frac{1}{2}\left(\frac{y_i}{\sqrt{\frac{3}{4}A}}\right)^2 + o(1)\right)$$

which is the GUE density. □

Theorem 13.5. *Keep track of the non-red Lozenge as $y_1(t) < y_2(t) < \cdots < y_N(t)$, then for any $\gamma < 1$,*

$$\lim_{A\to\infty}\left(\frac{y_i(tA^\gamma)}{A^{\frac{\gamma}{2}}}\right)_{i=1}^N = [\beta = 2]\ \textit{Dyson Brownian Motion } (t).$$

Sketch of the proof. By CLT, N non-intersecting random walks $\to$ N non-intersecting Brownian Motions.



Theorem 13.6. *Consider uniformly random tilings of $2A \times 2A \times 2A$ hexagon. Then as $A \to \infty$, inside the inscribed circle one sees all 3 types of lozenges, but outside, the configuration is "frozen", i.e. in each of 6 zones we see only 1 type of lozenges.*

Sketch of the proof. Let y_i be horizontal lozenges on $N-$th vertical. Assume $\frac{N}{A} \to x$, as $A \to \infty$ and study

$$\lim_{A\to\infty} \frac{1}{A} \sum_{i=1}^{N} \delta_{y_i/A} = \mu^x.$$

Notice that μ^x is not a probability measure for the total mass is x. We will not prove that the limit exists, but assuming that we find what μ^x should be. ☐

Proposition 13.7 (Nekrasov Equation; discrete loop equation, discrete Schinger-Dyson equation)**.** *Take a prob measure on N-tuples of integers $L \le y_1 < y_2 < \cdots < y_N \le R$ of the form*

$$\mathrm{Prob}(y_1, \ldots, y_N) = \frac{1}{Z} \prod_{i<j} (y_i - y_j)^2 \prod_{i=1}^{N} w(y_i)$$

with

- $w(L-1) = w(R+1) = 0$
- $\frac{w(y)}{w(y-1)} = \frac{\varphi^+(y)}{\varphi^-(y)}$ *where $\varphi^{\pm}(y)$ are holomorphic in a nbhd of $[L, R]$.*

Then

$$R(z) = \varphi^-(z) \mathbb{E} \prod_{i=1}^{N} \left[1 - \frac{1}{z - y_i}\right] + \varphi^+(z) \mathbb{E} \prod_{i=1}^{N} \left[1 + \frac{1}{z - y_i - 1}\right]$$

is also holomorphic in the same nbhd of $[L, R]$.

Proof. Examine the possible simple pole at $z = h$ by computing its residence

$$-\varphi^-(h) \sum_{i=1}^{N} \prod_{j \ne i} \left[1 - \frac{1}{h - y_j}\right] \mathrm{Prob}(y_j = h) + \varphi^+(h) \sum_{i=1}^{N} \prod_{j\ne i} \left[1 + \frac{1}{h - y_j - 1}\right] \mathrm{Prob}(y_j = h-1) \overset{?}{=} 0$$

Fix i, take 2 configurations y^+ and y^- s.t. $y_i^+ = h, y_i^- = h-1$ and are the same otherwise. They enter in the both sums

$$\frac{\varphi^-(h) \prod \left[1 - \frac{1}{h-y_j}\right] \mathrm{Prob}(y_i^+)}{\varphi^+(h) \prod \left[1 + \frac{1}{h-y_j-1}\right] \mathrm{Prob}(y_i^-)} = \frac{\varphi^-(h)}{\varphi^+(h)} \prod_{j\ne i} \left[\frac{h - y_j - 1}{h - y_j}\right]^2 \prod_{j \ne i} \left[\frac{h - y_j}{h - 1 - y_j}\right]^2 \cdot \frac{w(h)}{w(h-1)} = 1$$

Hence, the terms cancel out. ☐

Returning to the proof of the theorem, we first verify the two conditions stated in Proposition 13.7. We then observe that the R-function converges to a quadratic function, allowing us to explicitly identify its coefficients. Finally, we apply the Stieltjes inversion formula to determine the limiting distribution.

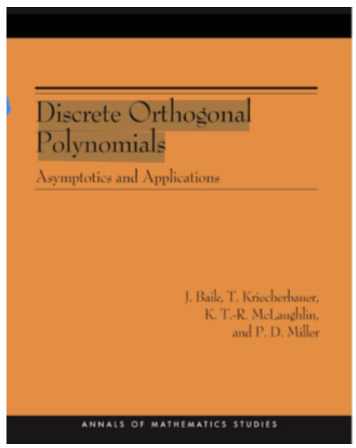
$\square$

Remark 13.8. *A general question is how the discrete frozen boundary approximates the circle?*

Theorem 13.9.

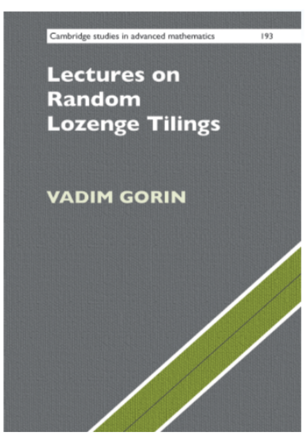
$$\frac{\textit{Discrete frozen - circle}}{cA^{1/3}} \overset{A\to\infty}{\to} \mathrm{TW}_2\,.$$

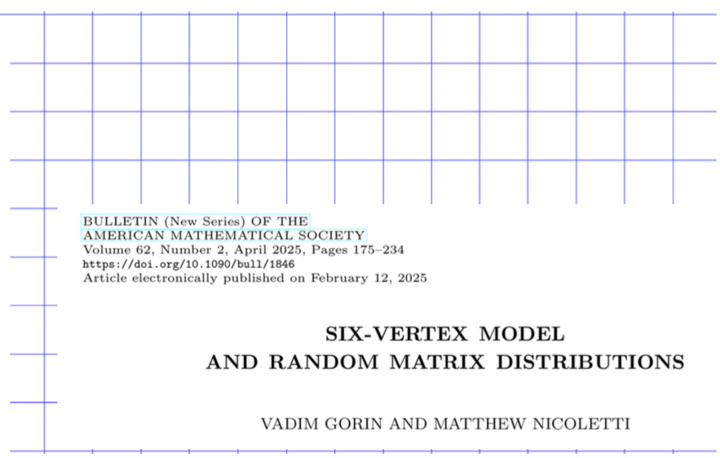
Proof. Using DPP, where double contour integrals can be found in and orthogonal polynomials can be found in



$\square$

Further reading.




BULLETIN (New Series) OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 62, Number 2, April 2025, Pages 175–234
https://doi.org/10.1090/bull/1846
Article electronically published on February 12, 2025

SIX-VERTEX MODEL AND RANDOM MATRIX DISTRIBUTIONS

VADIM GORIN AND MATTHEW NICOLETTI

14 Apr 29th

14.1 Random growth Models

Today we talk about Ulam's problem. Consider a uniformly random permutation of $\{1, \dots, n\}$, define the increasing subsequences, as the name means. And denote l_n to be the length of the largest increasing subsequence. So the question is how does l_n grow as $n \to \infty$?

The history: Ulan simulates by Monte Carlo and find $l_n \geq c\sqrt{n}$. Later in 1972, Hammersley proved that $l_n \sim c\sqrt{n}$. Then Vershik-Kerov / Logan-Shepp proved $l_n \sim 2\sqrt{n}$ in 1976. And Baik-Deitt-Johansson proved

$$l_n = 2\sqrt{n} + n^{1/6}\,\mathrm{TW}_2 + o(n^{1/6}).$$

Now we add a second dimension, and we consider point process in the quadrant:

1. $\mathbb{P}(\text{K points in set } A) = \exp(-\text{area}(A))[\text{area}(A)]^K/K!$

2. For disjoint $A_1, \cdots, A_m$, their particle counts are independent.

Proposition 14.1. *Let $L(x,t)$ = maximal number of points on monotone up-path (angles between $-\frac{\pi}{4}$ and $\frac{\pi}{4}$) from $(0,0)$ to (t,x).*

Then $L(x,t) \overset{d}{=} l_\rho$, where $\rho \sim Poisson\left(\frac{t^2-x^2}{2}\right)$ (sample $n = \rho$, and then sample l_n).

Proof. How many points are there between $(0,0)$ and (t,x)? This is ρ. Points are (y_i, z_i), order them by $y_1 < \cdots < y_n$. The question is what is the law of permutation given the order of z_i? Should be uniform, because $(y_i, z_i) \sim$i.i.d. in rectangle. Hence,

$$L(x,t) = l_n = l_\rho.$$

□

Now we have two points of view on random function $L(x,y)$. The first approach is random geometry: Let us discretize the quadrant. Put i.i.d. weights wedge on edges of the lattice. Set

$$L^w(x,t) = \min_{\text{Monotone lattice paths from (0,0) to (x,t)}} (w_{l_1} + w_{l_2} + \cdots + w_{l_K}).$$

- Liquid percolates $(0,0) \to (x,t)$, known as first passage percolation.
- Discrete time $(0,0) \to (x,t)$. This is the fastest time to reach it.

If we multiply all the w_i by -1, and we replace minimize problem by maximize problem, which is called last passage percolation.

Proposition 14.2. *Suppose $w \sim_{i.i.d.}$ Bernoulli$(\frac{1}{2N^2})$, then last passage time $(0,0) \to (Nx, Nt)$ will converge to $L(x,t)$ as $N \to \infty$.*

Proof. The set of edges where $w = 1$ becomes Poisson Point Process as $N \to \infty$. The remaining edges are 0. As $N \to \infty$, sum of w becomes the maximum number of points you take along the path. $\square$

Another point of view is to think about $L(x,t)$ as a random function at time t.

Lemma 14.3. *$L(x,t)$ is monotone in t.*

Proof. Any path from $(0,0)$ to (x,t) can be extended to $(x, t+\Delta t)$. $\square$

$L(x,t)$ is a random growth of an interface. What are the rules?

- $L(x,0) = 0$
- Point at (x,t) create an island, which grows linearly.
- Islands merge into one when they collide.

It is called Polynuclear Growth Model (PGM).

Then we want to understand why is asymptotics given by RM-object. We do it through Robinson-Schensted correspondence.

Definition 14.4 (Partition)**.** *Partition of n = representation of n as a sum of positive integers: $n = \lambda_1 + \lambda_2 + \cdots$.*

Definition 14.5 (YD)**.** *Young diagram = drawing of λ as a collection of boxes.*

Definition 14.6 (SYT)**.** *Standard Young Tableau of shape λ = filling of boxes with numbers $1, \dots, n$ in increasing order along rows and columns.*

Definition 14.7 (RS-correspondence)**.** *RS-correspondence: $\sigma \in S_n \to$ pairs of SYT of same shape $\lambda \vdash n$.*

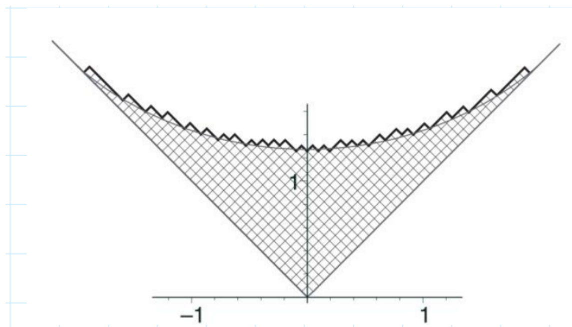
We grow SYT by adding boxes to YD one by one $\sigma = \sigma(1) \cdots \sigma(n)$.

Theorem 14.8. *Given $\sigma \in S_n$, construct a pair of SYT.*

1. *The outcome of the algorithm*
2. *The tableau which records how the YD is growing*

Then this is a bijection between σ and pairs of the same shape λ. And λ_1 = length of largest increasing subsequence in σ.

Corollary 14.9. *Set* $\dim \lambda = SYT$ *of shape* λ*. Then* $l_n \overset{d}{=} \lambda_1$*, where* $\lambda \vdash n$ *is a random partition distributed with weights* $\mathrm{Prob}_n(\lambda) = \frac{\dim^2 \lambda}{n!}$.



Theorem 14.10. $\lambda = (\lambda_1 \geq \cdots \geq \lambda_K \geq 0), \lambda \vdash n$*. Then*

$$\dim \lambda = \frac{n!}{\prod_{i=1}^{K} (\lambda_i + K - i)!} \prod_{1 \leq i < j \leq K} ((\lambda_i - i) - (\lambda_j - j))$$

Remark 14.11. K *can be arbitrary large, as long as* $\lambda_{K+1} = 0$ *(This formula is stable).*

Proof. Check recurrence.

$$\dim \lambda = \sum_{\lambda = \mu + \square} \dim \mu.$$

□

Corollary 14.12. *Plancherel measure is*

$$\mathrm{Prob}_n(\lambda) = n! \prod_{1 \leq i < j \leq K} ((\lambda_i - i) - (\lambda_j - j))^2 \prod_{i=1}^{K} \frac{1}{[(\lambda_i + K - i)!]^2}$$

Like a RM dist at $\beta = 2$, but

1. $\sum_{i=1}^{K} \lambda_i = n \to$ fixed (but Trace of matrix was not fixed previously)
2. K is not fixed (but for matrix, K was dimension, fixed)

Because of (1) Plancherel is not a DPP. But there is a remedy.

Definition 14.13 (Poissonized Plancherel measure)**.** *For* $\tau > 0$*, the Poissonized Plancherel measure* $\mathrm{PP}(\tau)$ *is a probability measure on all Young diagrams* λ *(i.e.* n *is not fixed) such that*

$$\mathrm{Prob}(\lambda) = e^{-\tau} \frac{\tau^n}{n!} \left(\frac{\dim^2 \lambda}{n!} \right) = e^{-\tau} \tau^n \dim^2 \lambda \left(\frac{1}{n!} \right)^2.$$

Remark 14.14. *Sample* n *as Poisson(*τ*), then sample* $\lambda \vdash n$*. It is the same mechanism as* $l_n \to L(x, y)$ *in the first half.*

Theorem 14.15. *Associate to* $\mathrm{PP}(\tau)$ *random* λ *an infinite particle configuration* $\{\lambda_i - i + \frac{1}{2}\}_{i=1}^{\infty} \subset \mathbb{Z} + \frac{1}{2}$*. These points form a DPP with kernel*

$$K(x, y) = \frac{1}{2\pi i} \oiint \exp\left(\sqrt{\tau} \left(v - v^{-1} - w + w^{-1} \right) \right) v^{x-1} w^{y-1} \frac{\sqrt{vw}}{v - w} \, dv \, dw.$$

Idea of the proof. See Lectures on integrable probability. □

Theorem 14.16. *Let λ be $PP(\tau)$. Then as $\tau \to \infty$*

$$\left\{ \frac{\lambda_i - 2\sqrt{\tau}}{\tau^{\frac{1}{6}}} \right\}_{i=1}^{\infty} \to \textit{ Airy Point Process}$$

And

$$\left\{ \frac{\lambda_i - 2\sqrt{\tau}}{\tau^{\frac{1}{6}}} \right\}_{i=1}^{\infty} \to \ TW_2$$

Corollary 14.17. *As $\tau \to \infty$*

$$\frac{L(x,t) - 2\sqrt{\frac{t^2-x^2}{2}}}{(\frac{t^2-x^2}{2})^{\frac{1}{6}}} \to \mathrm{TW}_2$$

And

$$\frac{l_n - 2\sqrt{n}}{n^{\frac{1}{6}}} \to \mathrm{TW}_2$$

Proof. Combine all the results. □