## Section 3 Invariant Ensembles

Consider the density function of an invariant ensemble as

$$\mathbb{P}(H)\,dH = \frac{1}{Z}\exp\left(-\frac{\beta}{2}N\,\mathrm{Tr}\,V(H)\right)dH$$

For GOE / GEU, $V(H) = \frac{1}{2}H^2$. In this case, this model is called a $\beta$-ensemble

$\beta = 1$ for GOE ; $\beta = 2$ for GUE ; $\beta = 4$ for GSE.

symplectic

( Gaussian quaternion ensemble )

For general $\beta$, it does not correspond to any Wigner ensembles.

① Probability density for $\beta$-ensemble

For invariant ensembles, their eigenvectors are uniformly distributed on the unit sphere. So people mainly focus on their eigenvalues. We will integrate out the eigenvectors in $\mathbb{P}(H)\,dH$ and show the following result.

**Theorem 3.1** The joint probability density of the eigenvalues of $H$ is given by

$$(*)\quad p_N(\lambda_1, \lambda_2, \cdots, \lambda_N) = \frac{1}{Z_N}\prod_{i<j}|\lambda_i - \lambda_j|^{\beta}\, e^{-\frac{\beta}{2}N\sum_{i=1}^{N}V(\lambda_i)}.$$

Rmk 1: Recall that $\Delta(\lambda) = \prod_{i<j}(\lambda_i - \lambda_j)$ is the Vandermonde determinant.

Rmk 2: In the above formula, we can neglect the ordering of the eigenvalues. This can be done by multiplying a $\frac{1}{n!}$ factor when calculating various probabilities.

~~Rmk 3: We can rewrite (*) as:~~ $\cancel{p_N(\lambda_1, \lambda_2, \cdots, \lambda_N) = \frac{1}{Z_N}\exp\left[\beta\log|\lambda_i - \lambda_j| - \frac{\beta}{2}N\right.}$

Rmk 3: When $|\lambda_i - \lambda_j| = 0$, we have $p_N(\lambda_1, \cdots, \lambda_N) = 0$. This indicates a repulsion between different eigenvalues.

Rmk 4: The eigenvalues are strongly correlated. It is useful to think of $p_N$ as a Gibbs measure on $N$ "particles" $\vec{\lambda} = (\lambda_1, \cdots, \lambda_N)$:

$$p_N(\vec{\lambda})\,d\vec{\lambda} = \exp[-\beta N \mathcal{H}(\vec{\lambda})]\,\frac{1}{Z_N},$$

$$\mathcal{H}(\vec{\lambda}) := \underbrace{\frac{1}{2}\sum_{i=1}^{N}V(\lambda_i)}_{\text{A confining potential.}} - \underbrace{\frac{1}{N}\sum_{i<j}\log|\lambda_i - \lambda_j|}_{\text{A logarithmic potential giving repulsions between particles.}}$$

[The two terms have the same order "$N$".]

This gives an important statistical mechanical system, called the "log-Coulomb gas".

To prove Thm 3.1, we need to show that integrating out the eigenvectors gives the Vandermonde determinant.

Let $H = \mathcal{U} D \mathcal{U}^*$ be an eigenvalue decomposition of $H$, where $D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_N)$ with $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$ and $\mathcal{U}$ is an orthogonal / unitary matrix.
We consider the unitary case. Then $\mathcal{U}$ is Haar distributed on $U(N)$.

From the unitary invariance of $\mathbb{P}(H)dH$, we see that after conditioning on $D$, the eigenmatrix $\mathcal{U}$ is drawn from the Haar measure on $U(n)$. In particular, $\mathcal{U}$ and $D$ can be taken to be independent:

$$\cancel{\mathbb{P}(H = \mathcal{U}_0 D_0 \mathcal{U}_0^* + dH) = \mathbb{P}(H = (V\mathcal{U}_0) D_0 (\mathcal{U}_0^* V^*) + dH)}$$

~~for any $V \in U(N)$~~

$$\mathbb{P}(\mathcal{U} \in B,\ D \in A) = \mathbb{P}(\mathcal{U} \in B \mid D \in A)\, \mathbb{P}(D \in A)$$
$$= \mathbb{P}(\mathcal{U} \in B)\, \mathbb{P}(D \in A).$$

<u>Proof of Theorem 3.1</u>: Fix a diagonal matrix $D_0 = \mathrm{diag}(\lambda_1^\circ, \cdots, \lambda_N^\circ)$, $\lambda_1^\circ \le \cdots \le \lambda_N^\circ$. Let $\varepsilon > 0$ be arbitrarily small. We now compute the probability that $H_N$ lies in an $\varepsilon$-ball around $D_0$ in the H-S norm (Frobenius norm):

$$\mathbb{P}(\|H_N - D_0\|_F \le \varepsilon)$$

(The Euclidean norm on the space of $n \times n$ matrices.)

I. The simpler direction: The probability density of $M_n$ is $\propto \exp(-\frac{\beta}{2} N \sum_i V(\lambda_i^\circ))$.
The volume of the $\varepsilon$-ball is $C_N \varepsilon^{N^2}$. So

(1) $$\mathbb{P}(\|H_N - D_0\|_F \le \varepsilon) = (C_N + \underbrace{o(1)}_{o_\varepsilon(1) \to 0 \text{ as } \varepsilon \to 0 \text{ when } N \text{ is fixed.}})\, \varepsilon^{N^2} \exp(-\tfrac{\beta}{2} N \textstyle\sum_i V(\lambda_i))$$

II. The harder direction: Let's calculate $\mathbb{P}(\|H_N - D_0\|_F \le \varepsilon)$ in a different way.

<u>Weyl's inequality</u>: Given two hermitian matrices $A$ and $B$ with eigenvalues
$$\lambda_1(A) \le \cdots \le \lambda_N(A), \quad \lambda_1(B) \le \cdots \le \lambda_N(B).$$
Then, $\forall i$, $\lambda_i(A) + \lambda_1(B) \le \lambda_i(A+B) \le \lambda_i(A) + \lambda_N(B)$.
In particular, if $B$ is a small perturbation with $\|B\| \le \varepsilon$, then
$$\lambda_i(A) - \varepsilon \le \lambda_i(A+B) \le \lambda_i(A) + \varepsilon.$$

Thus, for $H_N = \mathcal{U} D \mathcal{U}^*$, we have $\|D - D_0\| \le \varepsilon$ when $\|H_N - D_0\|_F \le \varepsilon$.

~~Moreover,~~ $\cancel{\mathcal{U} D \mathcal{U}^* = D_0 + O(\varepsilon) = D + O(\varepsilon) \Rightarrow \mathcal{U}D - D\mathcal{U} = O(\varepsilon)}$

~~Now, we make the ansatz~~

$$\cancel{D = D_0 + \varepsilon E, \quad \mathcal{U} = \exp(\varepsilon S)}$$

~~where $E$ is a diagonal matrix and $S$ is a skew-adjoint matrix.~~

~~[Note that $\mathcal{U}\mathcal{U}^* = \exp(\varepsilon S)\exp(\varepsilon S^*) = \exp(\varepsilon S)\exp(-\varepsilon S) = I$]~~

~~Note that $S \mapsto \exp(\varepsilon S)$ has a non-degenerate Jacobian in the small $\varepsilon$-ball, so the inverse function thm applies to specify $S$ from $\mathcal{U}$. (Formally, $S = \frac{1}{\varepsilon}\log \mathcal{U}$.)~~

~~WLOG, suppose all $\lambda_i^0$ are different (othe this holds a.s.).~~

Thus, we make the ansatz: $D = D_0 + \varepsilon E$, $E$ is diagonal bounded.

Note, the eigenvalues of $D_0$ are non-zero and non-degenerate almost surely. Moreover, we have $\quad \mathcal{U} D \mathcal{U}^* = D_0 + O(\varepsilon) \Rightarrow \mathcal{U} D_0 \mathcal{U}^* = D_0 + O(\varepsilon) \Rightarrow \mathcal{U} D_0 - D_0 \mathcal{U} = O(\varepsilon)$. $(*)$

The only unitary matrices that commute with $D_0$ are diagonal unitary matrices

$$R = (r_1, r_2, \cdots, r_N), \quad |r_i|^2 = 1.$$

From $(*)$, we can make the ansatz: $\mathcal{U} = \exp(\varepsilon S) R$,

where $S$ is a bounded skew-adjoint ($S^* = -S$) matrix with zero diagonal. (The diagonal parts can be included into $R$.)

We can check that $\mathcal{U}\mathcal{U}^* = \exp(\varepsilon S)\exp(\varepsilon S^*) = 1$,

$\mathcal{U} D_0 - D_0 \mathcal{U} = \exp(\varepsilon S) R D_0 - D_0 \exp(\varepsilon S) R = R D_0 - D_0 R + O(\varepsilon) = O(\varepsilon)$.

It is easy to check that $(R, S) \mapsto \exp(\varepsilon S) R$ has a non-degenerate Jacobian, so with the inverse function thm, we can uniquely determine $R$ and $S$ from $\mathcal{U}$ in the ~~sa~~ small $\varepsilon$-ball around $D_0$.

Now with the ansatz: $D = D_0 + \varepsilon E$, $\mathcal{U} = \exp(\varepsilon S) R$, we get

$$\begin{aligned} H_N = \mathcal{U} D \mathcal{U}^* &= \exp(\varepsilon S) R (D_0 + \varepsilon E) R^* \exp(-\varepsilon S) \\ &= D_0 + \varepsilon E + \varepsilon S R D_0 R^* - \varepsilon R D_0 R^* S + O(\varepsilon^2) \\ &= D_0 + \varepsilon E + \varepsilon (S D_0 - D_0 S) + O(\varepsilon^2). \end{aligned}$$

Thus, $\mathbb{P}(\|H_N - D_0\|_F \leq \varepsilon) = \mathbb{P}(\|E + (S D_0 - D_0 S)\| \leq 1 + O(\varepsilon))$. $(+)$

Since $\mathcal{U}$ is Haar distributed on $\mathcal{U}(N)$, $S$ is locally distributed as ~~$C_N \varepsilon^{N^2-N}$~~ $C_N' \varepsilon^{N^2 - N}$ times the Lebesgue measure on the space of skew-adjoint matrices with 0 diagonal. ~~Be~~ On the other hand, $E$ is distributed as $(\rho_N(\vec{\lambda}^0) + o(1))\varepsilon^N$ times the Lebesgue measure on the space of diagonal matrices. Thus we can calculate $(+)$ as:

$$C_N' \varepsilon^{N^2} (\rho_N(\vec{\lambda}^0) + o(1)) \iint_{\|E + S D_0 - D_0 S\|_F \leq 1 + O(\varepsilon)} dE\, dS$$

Consider the map $S \mapsto S D_0 - D_0 S$, where $(S D_0 - D_0 S)_{ij} = (\lambda_j^0 - \lambda_i^0) S_{ij}$.

In other words, the map dilates the $(i,j)$th entry of $S$ by $\lambda_j^0 - \lambda_i^0$. Hence, the Jacobian of this map is $\displaystyle\prod_{i \neq j} |\lambda_j^0 - \lambda_i^0| = \prod_{1 \leq i < j \leq N} |\lambda_j^0 - \lambda_i^0|^2$.

Applying the change of variable, we get

$$\iint_{\|E + S D_0 - D_0 S\|_F \leq 1 + O(\varepsilon)} dE\, dS = \frac{1}{\prod_{1 \leq i < j \leq N} |\lambda_j^0 - \lambda_i^0|^2} \iint_{\|E + S\|_F \leq 1 + O(\varepsilon)} dE\, dS$$

$$= \frac{C_N''}{\prod_{1 \leq i < j \leq N} |\lambda_j^0 - \lambda_i^0|^2}.$$

In sum, we have found that

$$\mathbb{P}(\|H_N - D_0\|_F \le \varepsilon) = (C_N + o(1))\,\varepsilon^{N^2} \exp\left[-\frac{\beta}{2} N \sum_i V(\lambda_i^\circ)\right]$$

$$= \frac{C_N'''}{\prod_{1\le i<j\le N} |\lambda_j^\circ - \lambda_i^\circ|^2} \left[\rho_N(\vec{\lambda}^\circ) + o(1)\right] \varepsilon^{N^2}$$

$$\Rightarrow \rho_N(\vec{\lambda}^\circ) = \frac{1}{Z_N} \prod_{1\le i<j\le N} |\lambda_j^\circ - \lambda_i^\circ|^2 \exp\left[-\frac{\beta}{2} N \sum_{i=1}^N V(\lambda_i^\circ)\right].$$ □

② Mean field approximation of semicircle law

$$\rho_N(\vec{\lambda}) = \frac{1}{Z_N} \exp\left[-\beta N \mathcal{H}(\vec{\lambda})\right], \quad \mathcal{H}(\vec{\lambda}) = \frac{1}{2}\sum_{i=1}^N V(\lambda_i) - \frac{1}{N}\sum_{i<j} \log|\lambda_i - \lambda_j|.$$

Intuitively, it is plausible that the spectrum should concentrate around the maximum of $\rho_N(\vec{\lambda})$ as $N \to +\infty$. It is equivalent to study the minimum of $\mathcal{H}(\vec{\lambda})$.

Heuristically, we make a mean-field approximation, i.e., the ESD $\frac{1}{N}\sum_i \delta_{\lambda_i}$ can be approximated by a continuous probability measure $\rho(x)\,dx$ (→ this is reasonable for GOE/GUE at least with $\rho = \rho_{sc}$.)

Then, we expect that $\mathcal{H}(\vec{\lambda})$ is approximately given by

~~$\mathcal{H}(\rho) = \frac{1}{2}\int_{\mathbb{R}} V(x)\,dx$~~

$$\mathcal{H}(\rho) = \frac{1}{2}\int_{\mathbb{R}} V(x)\rho(x)\,dx + \int_{\mathbb{R}}\int_{\mathbb{R}} \rho(x)\rho(y) \log\frac{1}{|x-y|}\,dx\,dy.$$

$\rho$ should minimize this functional. Now, we derive the Euler-Lagrange equation for $\mathcal{H}(\rho)$: consider $\rho + \delta\rho$, where $\rho$ is the minimizer. We should have

$$\frac{\delta \mathcal{H}(\rho)}{\delta\rho} = 0, \quad \text{because } \mathcal{H}(\rho + \delta\rho) \ge \mathcal{H}(\rho) \ \forall \delta\rho.$$

We have
$$\delta\mathcal{H}(\rho) = \frac{1}{2}\int_{\mathbb{R}} V(x)\,\delta\rho(x)\,dx + \int_{\mathbb{R}}\int_{\mathbb{R}} (\delta\rho(x))\rho(y)\log\frac{1}{|x-y|}\,dx\,dy$$
$$+ \int_{\mathbb{R}}\int_{\mathbb{R}} \rho(x)(\delta\rho(y))\log\frac{1}{|x-y|}\,dx\,dy$$

$$\Rightarrow \frac{1}{2}\int_{\mathbb{R}} V(x)\,\delta\rho(x)\,dx + 2\int_{\mathbb{R}}\left(\int_{\mathbb{R}} \rho(y)\log\frac{1}{|x-y|}\,dy\right)\delta\rho(x)\,dx = 0$$

Since $\delta\rho$ is arbitrary, with $\int_{\mathbb{R}} \delta\rho(x)\,dx = 0$ (since $\rho + \delta\rho$ should be a prob. density), we ought to have $\frac{1}{2}V(x) + 2\int_{\mathbb{R}} \rho(y)\log\frac{1}{|x-y|}\,dy = C$.

This is the E-L equation for $\rho$.

Solving the E-L equation, we get a density $\rho(x)$, which should be the limiting ESD of $\mathbb{P}(H)\,dH = \frac{1}{Z_N}\exp[-\frac{\beta}{2}N\,\mathrm{Tr}\,V(H)]\,dH$.

Let's solve it for GUE with $V(x)=x^2$. We need to solve

$$\frac{1}{2}x^2 + 2\int_{\mathbb{R}} \rho(y)\log\frac{1}{|x-y|}dy = C \Rightarrow x + 2\int_{\mathbb{R}}\frac{\rho(y)}{y-x}dy = 0 \quad (*)$$

Here, $\int_{\mathbb{R}}\frac{\rho(y)}{y-x}dy$ ~~should~~ should be formally understood as a principal value

$$\mathrm{p.v.}\int_{\mathbb{R}}\frac{\rho(y)}{y-x}dy := \lim_{\varepsilon\to 0}\int_{|y-x|\geq\varepsilon}\frac{\rho(y)}{x-y}dy = \lim_{\eta\to 0^+}\mathrm{Re}\int_{\mathbb{R}}\frac{\rho(y)}{y-(x+i\eta)}dy$$

$$= \lim_{\eta\to 0^+}\mathrm{Re}\,S_\rho(x+i\eta) =: \mathrm{Re}\,S_\rho(x+i0^+)$$

On the other hand, recall $\rho(x) = \frac{1}{\pi}\mathrm{Im}\,S_\rho(x+i0^+)$.

Then the equation $(*)$ reads: $\quad x + 2\mathrm{Re}\,S_\rho(x+i0^+) = 0$

$$\Rightarrow x\,\mathrm{Im}\,S_\rho(x+i0^+) + 2[\mathrm{Re}\,S_\rho(x+i0^+)][\mathrm{Im}\,S_\rho(x+i0^+)] = 0$$

$$\Rightarrow \mathrm{Im}[x S_\rho(x+i0^+) + S_\rho^2(x+i0^+)] = 0$$

Define function $f(z) = zS_\rho(z) + S_\rho^2(z)$. The above function means we can make an analytic continuation from $z\in\mathbb{C}_+ \to \mathbb{C}$ by making a reflection along $\mathbb{R}$:

$$f(z) = \overline{f(\bar{z})} \quad \text{for } z\in\mathbb{C}_-.$$

On the other hand, since $S(z) = \frac{-(1+o(1))}{z}$ as $z\to\infty$, $f\to -1$ as $z\to+\infty$.

By Liouville's theorem, $f$ is a constant and

$$f(z) \equiv \lim_{z\to+\infty}[zS_\rho(z) + S_\rho^2(z)] = -1$$

This gives the familiar self-consistent equation:

$$S_\rho^2(z) + zS_\rho(z) + 1 = 0 \Rightarrow S_\rho(z) = m_{sc}(z) \Rightarrow \rho = \rho_{sc}.$$

The above arguments can be made rigorous for general $\beta$ ensembles as long as we impose some conditions on $V$ (in particular, need certain condition on its growth at $\infty$).

③ Bulk eigenvalue distribution of GUE and orthogonal polynomials

Recall that the Vandermont determinant $\Delta(\vec{\lambda})$ can be expressed (up to a $\pm$ sign) as a determinant of an $n\times n$ matrix

$$M = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

[We take $n\equiv N$ in this section.]

Then, $|\Delta_n(\vec{\lambda})|^2$ can be expressed as $\det(MM^*)$, where

$$(MM^*)_{ij} = \sum_{k=0}^{n-1}\lambda_i^k\lambda_j^k, \quad 1\leq i\leq j\leq n.$$

Through ~~too~~ column operations of $M$, we observe the following fact: if $P_0(\lambda), P_1(\lambda), \cdots, P_{n-1}(\lambda)$ are any sequence of polynomials, in which $P_i(x)$ has degree $i$ and the degree $i$ term is exactly $x^i$, then $\Delta(\lambda_1, \cdots, \lambda_n)$ is equal to the det of the matrix $(P_{j-1}(\lambda_i))_{1\le i,j\le n}$.
Hence, $|\Delta(\vec{\lambda})|^2$ is equal to the det of the following matrix

$$\left( \sum_{k=0}^{n-1} P_k(\lambda_i) P_k(\lambda_j) \right)_{1\le i,j\le n}$$

Then, we rewrite the prob. density for the eigenvalues of GUE as

We resort to the original scaling with $\mathbb{E}|h_{ij}|^2=1$

$$p(\vec{\lambda}) \propto \det\left( \sum_{k=0}^{n-1} P_k(\lambda_i) e^{-\frac{\lambda_i^2}{4}} P_k(\lambda_j) e^{-\frac{\lambda_j^2}{4}} \right)_{1\le i,j\le n}.$$

One particular class of polynomials $\{P_i(\lambda)\}$ is of particular importance to us, that is, the "orthogonal polynomials" with respect to the Gaussian measure.
These are the famous "Hermite polynomials". Define the inner product

$$\langle f, g\rangle := \int_{\mathbb{R}} f(x) g(x) e^{-\frac{x^2}{2}} dx.$$

Let $P_0(x) = 1$. Then we define $P_i(x)$ through the Gram-Schmidt process, i.e., given $P_0(x), P_1(x), \cdots, P_{n-1}(x)$ we define $P_n(x)$ (of degree $n$ ~~and has degree n term as $x^n$~~) such that $\langle P_n, P_i\rangle = 0$ for $0\le i\le n-1$. Furthermore, we can normalize $P_n(x)$ so that $\langle P_n, P_n\rangle = 1$. Then we get a sequence of orthonormal polynomials with respect to the Gaussian measure: $\langle P_i, P_j\rangle = \delta_{ij}$.
Then, define the kernel

$$K_n(x,y) := \sum_{k=0}^{n-1} P_k(x) e^{-\frac{x^2}{4}} P_k(y) e^{-\frac{y^2}{4}}.$$

This kernel defines an orthogonal projection $\Pi_{V_n}$ in $L^2(\mathbb{R})$ to the span $V_n$ of $\{1, x, \cdots, x^n\}$. In addition, we have

$$p_n(\vec{\lambda}) \propto \det(K_n(\lambda_i, \lambda_j))_{1\le i,j\le n}.$$

<u>Lemma 3.2.</u> For any $k\ge 0$, we have

$$\int_{\mathbb{R}} \det[K_n(\lambda_i, \lambda_j)]_{1\le i,j\le k+1} \, d\lambda_{k+1} = (n-k) \det(K_n(\lambda_i, \lambda_j))_{1\le i,j\le k}.$$

<u>Rmk</u>: Here, we have neglected the ordering of eigenvalues.

<u>Pf</u>: We observe that

① $$\int_{\mathbb{R}} K_n(x,x)\,dx = \sum_{k=0}^{n-1} \int P_k(x)^2 e^{-\frac{x^2}{2}} dx = n.$$

② $$\int_{\mathbb{R}} K_n(x,z) K_n(z,y)\,dz = \int_{\mathbb{R}} \sum_{k,k'=0}^{n-1} P_k(x) e^{-\frac{x^2}{4}} \left[P_k(z) e^{-\frac{z^2}{4}} P_{k'}(z) e^{-\frac{z^2}{4}}\right] P_{k'}(y) e^{-\frac{y^2}{4}} dz$$
$$= \sum_{k=0}^{n-1} P_k(x) e^{-\frac{x^2}{4}} P_k(y) e^{-\frac{y^2}{4}} = K_n(x,y).$$

When $k=0$, ① completes the proof. Now suppose the result has been proved for $k \geq 0$. Then, we consider the $k+1$ case. We apply the cofactor expansion to the $(k+1)$-th row of $\det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq k+1}$. Then,

$$\det(K(\lambda_i,\lambda_j))_{1\leq i,j\leq k+1} = \det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq k} K_n(\lambda_{k+1},\lambda_{k+1}) + \sum_{l=1}^{k} (-1)^{k+1-l} K_n(\lambda_l,\lambda_{k+1}) \det(K_n(\lambda_i,\lambda_j))_{1\leq i\leq k;\ 1\leq j\leq k+1,\ j\neq l}$$

$$\int \det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq k} K_n(\lambda_{k+1},\lambda_{k+1})\, d\lambda_{k+1} = n \det(\lambda_i,\lambda_j)_{1\leq i,j\leq k}.$$

For the sum, we use ② to get that

$$\int K_n(\lambda_l,\lambda_{k+1}) \det(K_n(\lambda_i,\lambda_j))_{1\leq i\leq k;\ 1\leq j\leq k+1,\ j\neq l}\, d\lambda_{k+1}$$

$$= \det\left( (K_n(\lambda_i,\lambda_j))_{1\leq i\leq k;\ 1\leq j\leq k,\ j\neq l},\ (K_n(\lambda_i,\lambda_l))_{1\leq i\leq k} \right)$$

After row exchange, we get $(-1)^{k-l} \det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq k}$. □

Iterating the formula, we get $\int_{\mathbb{R}^n} \det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq n}\, d\lambda_1\cdots d\lambda_n = n!$.

On the other hand, neglecting the ordering eigenvalues, if we extend $p_n(\vec{\lambda})$ from the simplex $\{\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_n\}$ to $\mathbb{R}^n$, its integral is $n!$. Since $\det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq n}$ is symmetric in $\lambda_1,\cdots,\lambda_n$, we thus get the Gaudin–Mehta formula

$$p_n(\lambda_1,\cdots,\lambda_n) = \det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq n} \text{ on } \{\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_n\}.$$

Moreover, the above lemma 3.2 shows that $p_k(\lambda_1,\cdots,\lambda_k) = \det(K_n(\lambda_i,\lambda_j))_{1\leq i,j\leq k}$ is the $k$-point correlation function:

$$\int_{\mathbb{R}^k} p_k(\lambda_1,\cdots,\lambda_k) F(\lambda_1,\cdots,\lambda_k)\, d\lambda_1\cdots d\lambda_k$$

$$= \mathbb{E} \sum_{\substack{1\leq i_1,\cdots,i_k\leq n \\ \text{distinct}}} F(\lambda_{i_1},\cdots,\lambda_{i_k}),$$

$F(x_1,\cdots,x_k)$ supported on $\{x_1\leq x_2\leq\cdots\leq x_k\}$

When $k=1$, we get $\mathbb{E}\mu_{H_n}(dx) = \frac{1}{n} K_n(x,x)\,dx$, or

$$\mathbb{E}\mu_{\frac{1}{\sqrt{n}}H_n} = \frac{1}{\sqrt{n}} K_n(\sqrt{n}x,\sqrt{n}x)\,dx.$$

When $k=2$, we get $p_2(\lambda_1,\lambda_2) = \det\begin{pmatrix} K(\lambda_1,\lambda_1) & K(\lambda_1,\lambda_2) \\ K(\lambda_2,\lambda_1) & K(\lambda_2,\lambda_2) \end{pmatrix}.$

Recall that $P_i(x)$ is orthogonal to $x^j$ for $0 \le j < i$. This implies that $xP_i(x)$ is orthogonal to $x^j$ for $0 \le j < i-1$. On the other hand, $xP_i(x)$ has degree $i+1$, so $xP_i(x)$ must lie in the span of $\{P_{i-1}, P_i, P_{i+1}\}$. This gives that

$$xP_i = a P_{i-1} + b P_i + c P_{i+1}, \text{ where } c \neq 0.$$

Rearrange, we can write that

$$P_{i+1}(x) = (a_i x + b_i) P_i - c_i P_{i-1}, \quad a_i \neq 0.$$

Note $\int_{\mathbb{R}} x P_i(x) P_{i+1}(x) e^{-\frac{x^2}{2}} dx = \frac{1}{a_i}$, $\int_{\mathbb{R}} x P_i(x) P_{i-1}(x) dx = \frac{c_i}{a_i}$.

$$\Rightarrow \frac{1}{a_{i-1}} = \frac{c_i}{a_i} \Rightarrow c_i = \frac{a_i}{a_{i-1}}, \quad \text{with } a_{-1} = \infty.$$

Consider $P_{i+1}(x) P_i(y) - P_i(x) P_{i+1}(y) = a_i (x-y) P_i(x) P_i(y) - \frac{a_i}{a_{i-1}} \left( P_{i-1}(x) P_i(y) - P_{i-1}(y) P_i(x) \right).$

$$\Rightarrow P_i(x) P_i(y) = \frac{P_{i+1}(x) P_i(y) - P_i(x) P_{i+1}(y)}{a_i (x-y)} - \frac{P_i(x) P_{i-1}(y) - P_{i-1}(x) P_i(y)}{a_{i-1}(x-y)}.$$

Summing over them gives the Christoffel-Darboux formula:

$$K_n(x,y) := \sum_{i=0}^{n-1} P_i(x) P_i(y) e^{-\frac{x^2+y^2}{4}} = \frac{P_n(x) P_{n-1}(y) - P_{n-1}(x) P_n(y)}{a_{n-1}(x-y)} e^{-\frac{x^2+y^2}{4}}.$$

To understand the asymptotic behavior of $K_n(x,y)$ as $n \to +\infty$, need to understand the behavior $P_n(x)$ as $n \to +\infty$.

---

Note that the above discussion is very general and not restricted to Hermite polynomials. It only requires a system of orthogonal polynomials.

Hermite polynomials are defined as

$$H_k(x) := (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2} \quad \text{with leading coefficient } 1.$$

Hermite functions are defined as

$$\psi_k(x) = \underbrace{\frac{H_k(x)}{(2\pi)^{1/4}\sqrt{k!}}}_{\to P_k(x)} e^{-x^2/4}, \quad \int \psi_k(x) \psi_\ell(x) dx = \delta_{k\ell}.$$

Hermite functions are eigenfunctions of the Hamiltonian for the Harmonic oscillator in quantum mechanics. They have the celebrated asymptotics:

$$\psi_{2m}(x) = \frac{(-1)^m}{n^{1/4}\sqrt{\pi}} \cos(\sqrt{n}\, x) + \bar{o}(n^{-\frac{1}{4}}),$$

$$\psi_{2m+1}(x) = \frac{(-1)^m}{n^{1/4}\sqrt{\pi}} \sin(\sqrt{n}\, x) + \bar{o}(n^{-\frac{1}{4}}), \quad \text{as } n \to +\infty \text{ and } |2m-n| = O(1).$$

They are uniform over $|x| \le C n^{-\frac{1}{2}}$. In addition, $a_{n-1} \sim \frac{1}{\sqrt{n}}$.

Hence, we get $\displaystyle K_n(x,y) \approx \sqrt{n}\, \frac{\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)}{x-y}$

$$\approx \frac{1}{\pi} \frac{\sin(\sqrt{n}x)\cos(\sqrt{n}y) - \sin(\sqrt{n}y)\cos(\sqrt{n}x)}{x-y}$$

$$= \frac{1}{\pi} \frac{\sin(\sqrt{n}(x-y))}{x-y}$$

We consider the scaling $\frac{y}{\sqrt{n}\rho_{sc}(0)}$, $\rho_{sc}(0) = \frac{1}{\pi}$. Then

$$K_n\left(\frac{y_1}{\sqrt{n}\rho_{sc}(0)}, \frac{y_2}{\sqrt{n}\rho_{sc}(0)}\right) \approx \sqrt{n}\rho_{sc}(0) \frac{\sin[\pi(y_1-y_2)]}{\pi(y_1-y_2)}.$$

This gives the Dyson sine kernel around 0. In general, we have

$$\cancel{\frac{1}{\sqrt{n}\rho_{sc}(x)} K_n\left(\frac{y_1}{\sqrt{n}\rho_{sc}(x)}, \frac{y_2}{\sqrt{n}\rho_{sc}(x)}\right)} \to S(y_1-y_2) := \frac{\sin(\pi(y_1-y_2))}{\pi(y_1-y_2)} \quad \text{for } x\in(-2,2).$$

$$\frac{1}{\sqrt{n}\rho_{sc}(x)} K_n\left(\sqrt{n}x + \frac{y_1}{\sqrt{n}\rho_{sc}(x)}, \sqrt{n}x + \frac{y_2}{\sqrt{n}\rho_{sc}(x)}\right)$$

We resort to the scaling $H_n \to \frac{1}{\sqrt{n}}H_n$, where the typical eigengap is of order $\frac{1}{n}$. Then,

$$\frac{1}{[\rho_{sc}(E)]^n} p_k\left(E + \frac{x_1}{n\rho_{sc}(E)}, \cdots, E + \frac{x_k}{n\rho_{sc}(E)}\right) \to \det\left(S(x_i - x_j)\right)_{1\le i,j\le n}.$$

This means, for any $E\in(-2,2)$ and bounded $F$,

$$\mathbb{E} \sum_{1\le i_1<i_2<\cdots<i_k\le n} F\left(n\rho_{sc}(E)(\lambda_{i_1}-E), \cdots, n\rho_{sc}(E)(\lambda_{i_k}-E)\right)$$

$$\to \int_{y_1\le y_2\le\cdots\le y_k} F(y_1,y_2,\cdots,y_k) \det\left(S(y_i-y_j)\right)_{1\le i,j\le k} dy_1\cdots dy_k.$$

In particular, using the 2-point correlation function, one can derive that "Wigner surmise": given a bulk energy $E$, let $i$ s.t. $i = N\left(\int_{-2}^{E} \rho_{sc}(x)dx + \bar{o}(1)\right)$. Then

$N\rho_{sc}(E)(\lambda_{i+1}-\lambda_i)$ converges in law to $p_W(x) = \frac{32x^2}{\pi^2} e^{-4x^2/\pi}$ in the Hermitian case.

Later, we will show the "bulk universality", i.e., the Dyson sine kernel and the Wigner surmise also occur for general Wigner matrices.

---

At edge, a different scaling has to be used. Using the definition of Hermite functions, we can check the following identity:

$$\psi_N'(x) = -\frac{x}{2}\psi_N(x) + \sqrt{N}\psi_{N-1}(x).$$

With this identity, we can check that

$$K_N(x,y) = \sqrt{N}\, \frac{\psi_N(x)\psi_{N-1}(y) - \psi_N(y)\psi_{N-1}(x)}{x-y} = \frac{\psi_N(x)\psi_N'(y) - \psi_N(y)\psi_N'(x)}{x-y} - \frac{1}{2}\psi_N(x)\psi_N(y).$$

The Plancherel-Rotach edge asymptotics for $\psi_N$ gives that

$$N^{\frac{1}{12}} \psi_N\left(2\sqrt{N} + \frac{u}{N^{1/6}}\right) \to Ai(u), \quad \text{where } Ai(u) \text{ is the Airy function:}$$

$$Ai(u) = \frac{1}{\pi}\int_0^\infty \cos\left(\frac{1}{3}t^3 + ut\right)dt$$

(It is called Airy function because it is a solution to the Airy equation $y'' - xy = 0$.)

Then, we see that

$$\frac{1}{N^{1/6}} K_N\left(2\sqrt{N} + \frac{u}{N^{1/6}}, 2\sqrt{N} + \frac{v}{N^{1/6}}\right) \to A(u,v),$$

$$A(u,v) := \frac{Ai(u)Ai'(v) - Ai'(u)Ai(v)}{u-v}$$ the Airy kernel.

Resort to the scaling $H_N \to \frac{1}{\sqrt{N}} H_N$, we have

$$N^{k/3} p_k\left(2 + \frac{\alpha_1}{N^{2/3}}, 2 + \frac{\alpha_2}{N^{2/3}}, \cdots, 2 + \frac{\alpha_k}{N^{2/3}}\right) \to \det\left(A(\alpha_i, \alpha_j)\right)_{1\le i,j\le k}.$$

When $k=2$, we have

$$\sum_{j\neq k} \mathbb{E}\, F\left(N^{2/3}(\lambda_j - 2), N^{2/3}(\lambda_k - 2)\right) \to \int_{\mathbb{R}^2} d\alpha_1 d\alpha_2\, F(\alpha_1, \alpha_2) \det\left(A(\alpha_i,\alpha_j)\right)_{1\le i,j\le 2}.$$

A much harder problem is to derive the limiting distribution of $N^{2/3}(\lambda_N - 2)$.
Tracy and Widom (1993, 1994) show that it converges to a limiting distribution, which is now referred to as the Tracy-Widom law: $F_\beta(x)$, $\beta = 1, 2, 4$.
In particular,

$$F_2(s) = \exp\left(-\int_s^\infty (x-s)q^2(x)dx\right),$$ where $q(x)$ is the solution to Painlevé equation II: $q''(s) = sq(s) + 2q(s)^3$, $q(s) \sim Ai(s)$ as $s \to \infty$.

"Edge universality" also holds, i.e., $N^{2/3}(\lambda_N - 2)$ also converges to the TW law as $N \to +\infty$ for general Wigner matrices.

$\boxed{\text{Section 4}}$ Rigidity of eigenvalues

Thm 4.1 (Local semicircle law) Let $H_N$ be a real Wigner matrix with independent entries (up to $H=H^*$) with mean 0 and variance $\frac{1}{N}$. Its entries have finite moments up to any order, i.e., $\forall k \in \mathbb{N}$, $\exists C_k > 0$ so that

$$\max_{i,j} \mathbb{E} |\sqrt{N} h_{ij}|^k \le C_k .$$

Then, uniformly in $|E| \le \varepsilon^{-1}$ and $\eta \ge N^{-1+\varepsilon}$ for a small constant $\varepsilon > 0$, we have:

(1) (Entrywise local law)

$$\max_{i,j} | G_{ij}(z) - \delta_{ij} m_{sc}(z) | \prec \Psi(z), \qquad \Psi(z) := \sqrt{\frac{\operatorname{Im} m_{sc}(z)}{N\eta}} + \frac{1}{N\eta} .$$

(2) (Averaged local law) $|m_N(z) - m_{sc}(z)| \prec \frac{1}{N\eta}$, $m_N = \frac{1}{N} \operatorname{Tr} G(z)$.

Outside of the spectrum (i.e., $[-2,2]$), we have the stronger averaged local law:

(3) (Averaged local law)

$$| m_N(z) - m_{sc}(z) | \prec \frac{1}{N(\kappa_E + \eta)} + \frac{1}{(N\eta)^2 \sqrt{\kappa_E + \eta}}$$

uniformly in $\{z : 2 \le |E| \le \varepsilon^{-1}, N^{-1+\varepsilon} \le \eta \le 10, N\eta\sqrt{\kappa_E + \eta} \ge N^{\varepsilon}\}$ for any small const $\varepsilon > 0$.
Here $\kappa_E$ is defined as the distance to the edges: $\kappa_E := \min\{|E-2|, |E+2|\}$.

We will prove this theorem later. Before that, we consider some of it applications. In this section, we use it to establish the "rigidity of eigenvalues".

Thm 4.2: Given the local laws, we have that

$$|\lambda_j - \gamma_j| \prec (j, N-j+1)^{-\frac{1}{3}} N^{-\frac{2}{3}} \quad \forall\ 1 \le j \le N,$$

where $\gamma_j$ is defined as

$$\int_{-2}^{\gamma_j} \rho_{sc}(x)\, dx = \frac{j}{N} . \quad \left( \text{Recall that } \int_{-\infty}^{\lambda_j} d\mu_{H_N}(x) = \frac{j}{N} . \right)$$

Rmk: Note that this is much stronger than the semicircle law we proved before. Previously, we have only shown that the Kolmogorov distance between $\mu_{H_N}$ and $\mu_{sc}$ is $o(1)$ with prob. $1-o(1)$. Thm 4.2 indeed implies that

$$\sup_x \left| \int_{-\infty}^{x} \rho_{sc}(t)\, dt - \int_{-\infty}^{x} d\mu_{H_N}(x) \right| \prec \frac{1}{N} .$$

This indicates a strong correlation between eigenvalues. For independent point process, the typical order of fluctuation should be $N^{-\frac{1}{2}}$.

① We first show an upper / lower bound on the largest / smallest eigenvalue.

**Prop**: Given the local law, for any constants $\varepsilon, D > 0$, we have that

$$\mathbb{P}\left( \lambda_N \geq 2 + N^{-\frac{2}{3}+\varepsilon} \right) \leq N^{-D} \quad \text{for large enough } N.$$

**Pf**: Let $\eta = N^{-\frac{2}{3}}$ and choose $E = 2 + \kappa$ for some $\kappa \geq N^{-\frac{2}{3}+\varepsilon} \gg N^{\frac{\varepsilon}{2}}\eta$. From the local law, $\displaystyle |m_N(z) - m_{sc}(z)| \prec \frac{1}{N(\kappa_E+\eta)} + \frac{1}{(N\eta)^2\sqrt{\kappa_E+\eta}}$, we get

$$|\operatorname{Im} m_N(z) - \operatorname{Im} m_{sc}(z)| \leq \frac{N^{\varepsilon/2}}{N\kappa} \ll \frac{1}{N\eta} \quad \text{w.h.p.}$$

On the other hand, if there is an eigenvalue $\lambda$ with $|\lambda - E| \leq \eta$, we have

$$\operatorname{Im} m_N(z) = \frac{1}{N}\sum_k \frac{\eta}{(\lambda_k - E)^2 + \eta^2} \geq \frac{1}{2N\eta}.$$

For $m_{sc}(z) = \dfrac{-z + \sqrt{z^2-4}}{2}$, we can check that $\displaystyle \operatorname{Im} m_{sc}(z) \lesssim \frac{\eta}{\sqrt{\kappa+\eta}} \ll \frac{1}{N\eta}$. (see the lemma below)

Hence, $\displaystyle \operatorname{Im} m_N(z) \leq \operatorname{Im} m_{sc}(z) + |\operatorname{Im} m_N(z) - \operatorname{Im} m_{sc}(z)| \ll \frac{1}{N\eta}$ w.h.p.

We get a contradiction. $\square$

**Lemma** (Properties of $m_{sc}(z)$) $m_{sc}(z) = \dfrac{-z + \sqrt{z^2-4}}{2}$ is the unique solution to the self-consistent equation $\displaystyle m(z) + \frac{1}{m(z)} + z = 0$ with $\operatorname{Im} m(z) > 0$ for $\operatorname{Im} z > 0$. Furthermore, we have that:

(i) $|m_{sc}(z)| = |m_{sc}(z) + z|^{-1} \leq 1$.

(ii) For any constant $C > 0$, there exists a constant $\varepsilon > 0$ such that for $E \in [-C, C]$ and $\eta \in [0, C]$, we have

$$\varepsilon \leq |m_{sc}(z)| \leq 1 - \varepsilon\eta.$$

Furthermore, $|m_{sc}(E)| = 1$ for $E \in [-2, 2]$.

(iii) $|1 - m_{sc}^2(z)| \simeq \sqrt{\kappa_E + \eta}$, $\quad \kappa_E := ||E| - 2|$.

(iv)

$$\operatorname{Im} m_{sc}(z) \simeq \begin{cases} \sqrt{\kappa_E + \eta} & \text{if } |E| \leq 2 \\ \dfrac{\eta}{\sqrt{\kappa_E+\eta}} & \text{if } |E| > 2 \end{cases}.$$

Proof involves a direct calculation. The estimate (iv) is more general, and, unlike other estimates, does not pertain to the semicircle law. For $z = E + i\eta$,

$$\operatorname{Im} m_{sc}(z) = \operatorname{Im}\int_{-2}^{2} \frac{\rho_{sc}(x)\,dx}{x - (E+i\eta)} = \int_{-2}^{2} \frac{\eta}{(x-E)^2+\eta^2}\rho_{sc}(x)\,dx.$$

~~When $E > 2$, we have~~

$$\cancel{\operatorname{Im} m_{sc}(z) \simeq \int_0^2 \frac{\sqrt{2-x}}{(E-x)^2+\eta^2}dx + O(\eta) = \int_0^2 \frac{\sqrt{x}}{(\kappa+x)^2+\eta^2}dx + O(\eta)}$$

$\leq$

$*$ When $E \geq 2$, we have $\operatorname{Im} m_{sc}(z) \sim \eta \int_{2-c}^{2} \frac{\sqrt{2-x}}{(E-x)^2+\eta^2} dx + O(\eta)$

$$= \eta \int_0^c \frac{\sqrt{x}}{(x+\kappa)^2+\eta^2} dx + O(\eta),$$

where $\int_0^c \frac{\sqrt{x}}{(x+\kappa)^2+\eta^2} dx \sim \int_0^c \frac{\sqrt{x}}{(x+\kappa+\eta)^2} dx \sim \int_0^{\frac{c}{\kappa+\eta}} \frac{1}{\sqrt{\kappa+\eta}} \frac{\sqrt{y}}{y^2+1} dy \sim \frac{1}{\sqrt{\kappa+\eta}}$.

$*$ When $E \in [0,2)$, we have $\operatorname{Im} m_{sc}(z) \sim \underbrace{\eta \int_{E-c\kappa}^{2} \frac{\sqrt{2-x}}{(E-x)^2+\eta^2} dx}_{\mathrm{I}} + \underbrace{\eta \int_{-2}^{E-c\kappa} \frac{\sqrt{2-x}}{(E-x)^2+\eta^2} dx}_{\mathrm{II}}$.

For the second term, $\mathrm{II} = \eta \int_{(1+c)\kappa}^{4} \frac{\sqrt{y}}{(y-\kappa)^2+\eta^2} dy \sim \eta \int_{c\kappa}^{4-\kappa} \frac{\sqrt{y}}{y^2+\eta^2} dy \sim \sqrt{\eta} \int_{c\kappa/\eta}^{(4-\kappa)/\eta} \frac{\sqrt{y}}{y^2+1} dy$

$$\sim \sqrt{\eta} \wedge \frac{\eta}{\sqrt{\kappa}}.$$

For the first term, $\mathrm{I} = \eta \int_0^{(1+c)\kappa} \frac{\sqrt{y}}{(y-\kappa)^2+\eta^2} dy = \eta \int_{-\kappa}^{c\kappa} \frac{\sqrt{y+\kappa}}{y^2+\eta^2} dy = \sqrt{\eta} \int_{-\kappa/\eta}^{c\kappa/\eta} \frac{\sqrt{y+\kappa/\eta}}{y^2+1} dy$.

Note that $\mathrm{I} \sim \sqrt{\eta}$ if $\kappa \leq \eta$.

On the other hand, if $\kappa \geq \eta$, we have $\mathrm{I} = \sqrt{\eta} \int_{-c\kappa/\eta}^{c\kappa/\eta} \frac{\sqrt{y+\kappa/\eta}}{y^2+1} dy + O\left(\frac{\eta}{\sqrt{\kappa}}\right)$

$$\sim \sqrt{\kappa} \int_{-c\kappa/\eta}^{c\kappa/\eta} \frac{1}{y^2+1} dy + O\left(\frac{\eta}{\sqrt{\kappa}}\right) \sim \sqrt{\kappa}.$$

In sum, we get $\mathrm{I} \sim \sqrt{\kappa+\eta}$.

② The Helffer – Sjöstrand formula

The H-S formula provides a way to control the ~~con~~ distance between two measures through bounding the difference between their Stieltjes transforms.

Lemma 4.3: Let $f \in C_c^1(\mathbb{R})$, and $\chi \in C_c^\infty(\mathbb{R})$ be a smooth cutoff function with support in $[-1,1]$ and with $\chi(y) \equiv 1$ for $|y| \leq 1/2$. Then, we have that

$$f(\lambda) = \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{iy f''(x)\chi(y) + i(f(x) + iyf'(x))\chi'(y)}{\lambda - (x+iy)} dx\,dy.$$

Rmk: The H-S formula can be regarded as an extension of the Cauchy integral formula:

$$f(\lambda) = \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{z-\lambda} dz, \quad C_\lambda \text{ encircles } \lambda.$$

Pf: Define $\tilde{f}(x+iy) := (f(x) + iyf'(x))\chi(y)$.

For $\partial_{\bar z} = (\partial_x + i\partial_y)/2$, we ~~to~~ claim that

$$(*) \qquad f(\lambda) = \frac{1}{\pi}\int_{\mathbb{R}^2} \frac{\partial_{\bar z}\tilde f(x+iy)}{\lambda-(x+iy)}\,dx\,dy.$$

Expanding $\partial_{\bar z}\tilde f(x+iy)$, we conclude the proof. The proof of $(*)$ is a simple application of the Green's theorem. With $\partial_{\bar z}(\lambda-z)^{-1}=0$, we can write that

$$\frac{1}{\pi}\int \frac{\partial_{\bar z}(\tilde f(x+iy))}{\lambda-(x+iy)}\,dx\,dy = \frac{1}{\pi}\int \partial_{\bar z}\, g(x+iy)\,dx\,dy, \quad \text{where } g(x+iy) = \frac{\tilde f(x+iy)}{\lambda-(x+iy)}.$$

$$= \frac{1}{2\pi}\int [\partial_x g(x+iy) + i\partial_y g(x+iy)]\,dx\,dy.$$

From the compact support condition of $g$ and Green's theorem, we get


$$\frac{1}{2\pi}\int_{|x+iy-\lambda|\ge\varepsilon} [\partial_x g(x+iy) + i\partial_y g(x+iy)]\,dx\,dy$$

$$= \frac{1}{2\pi}\oint_{C_\varepsilon} \cancel{f(x,y)}[i g(x,y)\,dx - g(x,y)\,dy] \cancel{dx\,dy} = \frac{1}{2\pi i}\oint_{C_\varepsilon}\frac{\tilde f(x+iy)}{x+iy-\lambda}\,dz$$

$$\to \tilde f(\lambda) = f(\lambda) \quad \text{as } \varepsilon\to 0.$$

Further, the integral $\int_{|x+iy-\lambda|\le\varepsilon}[\partial_x g(x+iy) + i\partial_y g(x+iy)]\,dx\,dy \to 0$ as $\varepsilon\to 0$. This gives $(*)$. $\square$

Rmk: The formula $(*)$ also holds for the choice $\tilde f(x+iy) = f(x)\chi(y)$. But, we need to add $iy f'(x)\chi(y)$ to make $\partial_{\bar z}\tilde f(x+iy) = O(|y|)$ for small $|y|$, which will be used in the following proof.

③ The convergence speed of the ESD

Define the empirical distribution

$$F_N(E) := \int_{-\infty}^{E} \rho_{H_N}(x)\,dx = \frac{1}{N}\#\{k: \lambda_k \le E\}.$$

Define the distribution function of the semicircle density

$$F_{sc}(E) := \int_{-\infty}^{E}\rho_{sc}(x)\,dx.$$

Theorem 4.4 ~~Under~~ Given the averaged local law $|m_N(z) - m_{sc}(z)| \prec \frac{1}{N\eta}$ uniformly in $|E|\le\varepsilon^{-1}$ and $|\eta|\ge N^{-1+\varepsilon}$ for any small constant $\varepsilon>0$, we have that

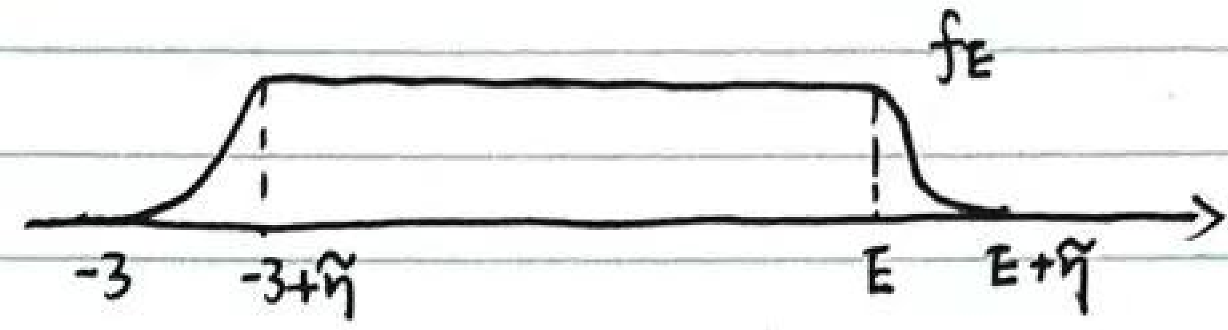
$$\sup_{E\in\mathbb{R}} |F_N(E) - F_{sc}(E)| \prec \frac{1}{N}.$$

Pf: We already know that $H_N$ has no eigenvalue outside $[-2-N^{-\frac{2}{3}+\varepsilon}, 2+N^{-\frac{2}{3}+\varepsilon}]$. So outside this interval, $F_N(E) - F_{sc}(E) = 0$ w.h.p.

We only need to control $F_N(E) - F_{sc}(E)$ for $E \in [-2-N^{-\frac{2}{3}+\varepsilon}, 2+N^{-\frac{2}{3}+\varepsilon}]$.
We pick $\tilde{\eta} = N^{-1+\varepsilon}$ for a small constant $\varepsilon > 0$. We know that

$$|m_N(z) - m_{sc}(z)| = \left| \int \frac{dF_N(x)}{x-z} - \int \frac{dF_{sc}(x)}{x-z} \right| \prec \frac{1}{N\eta} \quad \text{for } E \in (-3,3) \text{ and } \eta \geq \tilde{\eta}.$$

Given $E \in [-2-N^{-\frac{2}{3}+\varepsilon}, 2+N^{-\frac{2}{3}+\varepsilon}]$, we define a smooth cutoff function $f_E(x)$ such that $\mathrm{supp}(f_E) \subseteq [-3, E+\tilde{\eta}]$, $f_E(x) \equiv 1$ for $x \in [-3+\tilde{\eta}, E]$, ~~and~~ $\|f_E'\|_\infty \leq C\tilde{\eta}^{-1}$, $\|f_E''\|_\infty \leq C\tilde{\eta}^{-2}$, and $f_E \uparrow$ on $[-3, -3+\tilde{\eta}]$, $f_E \downarrow$ on $[E, E+\tilde{\eta}]$.




Note that $\int f_E(x)\, \rho_{H_N}(x)\, dx$ and $\int f_E(x)\, \rho_{sc}(x)\, dx$ approximates $F_N(E)$ and $F_{sc}(E)$.
We now bound

$$\int f_E(x) [\rho_{H_N}(x) - \rho_{sc}(x)] dx \overset{H-S}{=} \frac{1}{\pi} \int_{\mathbb{R}^2} \int_{\mathbb{R}} \frac{iy f_E''(x)\chi(y) + i(f_E(x) + iy f_E'(x))\chi'(y)}{\lambda - (x+iy)} [\rho_{H_N}(\lambda) - \rho_{sc}(\lambda)]\, d\lambda\, dx\, dy$$

$$= \frac{1}{\pi} \int_{\mathbb{R}^2} [iy f_E''(x)\chi(y) + i(f_E(x) + iy f_E'(x))\chi'(y)] (m_N(z) - m_{sc}(z))\, dx\, dy$$

$$= -\frac{1}{\pi} \int_{\mathbb{R}^2} [y f_E''(x)\chi(y) + f_E(x)\chi'(y)]\, \mathrm{Im}(m_N(z) - m_{sc}(z))\, dx\, dy$$

$$- \frac{1}{\pi} \int_{\mathbb{R}^2} y f_E'(x) \chi'(y)\, \mathrm{Re}[m_N(z) - m_{sc}(z)]\, dx\, dy = \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

Note $\chi'(y) \neq 0$ for $1/2 \leq |y| \leq 1$. Furthermore, $f_E'(x)$ is of order $O(\frac{1}{\tilde{\eta}})$ on two intervals of size $\tilde{\eta}$. For $1/2 \leq |y| \leq 1$, we have $|\mathrm{Re}[m_N(z) - m_{sc}(z)]| \prec \frac{1}{N}$. Hence, we have

$$|\mathrm{III}| \prec \frac{1}{N} \times \frac{1}{\tilde{\eta}} \times \tilde{\eta} = \frac{1}{N}.$$

Similarly, we have $|\mathrm{II}| \prec \frac{1}{N}$. For I, by symmetry, we only need to consider the upper complex plane with $y \geq 0$. We consider two regimes: $0 < y \leq \tilde{\eta}$ and $y > \tilde{\eta}$. On $0 < y \leq \tilde{\eta}$, we have $|\mathrm{Im}\, m_{sc}(z)| \leq 1$ and $|\mathrm{Im}\, m_N(z)| \leq C\tilde{\eta}/y$ w.h.p. using the inequality $y\, \mathrm{Im}\, m_N(E+iy) \leq \tilde{\eta}\, \mathrm{Im}\, m_N(E+i\tilde{\eta})$ for $y \leq \tilde{\eta}$.
Furthermore, $f_E''(x)$ is of order $O(\tilde{\eta}^{-2})$ on two intervals of size $\tilde{\eta}$. Hence,

$$\int_{x \in [-3, E+\tilde{\eta}]} \int_{0 < y \leq \tilde{\eta}} |y f_E''(x)\chi(y)|\, |\mathrm{Im}(m_N(z) - m_{sc}(z))|\, dx\, dy$$

$$\leq C\tilde{\eta} \int_{x \in [-3, -3+\tilde{\eta}] \cup [E, E+\tilde{\eta}]} \int_{0 < y \leq \tilde{\eta}} |f_E''(x)|\, dx\, dy \leq C\tilde{\eta} \cdot \frac{1}{\tilde{\eta}^2} \cdot \tilde{\eta}^2 = O(\tilde{\eta}).$$

For the regime $\tilde{\eta} < y \le 1$, we use integration by parts in $x$ to get

$$\int dx \int_{\tilde{\eta} < y \le 1} dy \quad y f_E''(x) \chi(y) \ \mathrm{Im}\,(m_N(z) - m_{sc}(z))\, dx\, dy$$

$$= -\int dx \int_{\tilde{\eta} < y \le 1} dy \quad y f_E'(x) \chi(y) \ \partial_x \mathrm{Im}\,(m_N(z) - m_{sc}(z))\, dx\, dy$$

Cauchy - Riemann condition

$$= \int dx \int_{\tilde{\eta} < y \le 1} dy \quad y f_E'(x) \chi(y) \ \partial_y \mathrm{Im}\,(m_N(z) - m_{sc}(z))\, dx\, dy$$

$$= -\int dx \quad \tilde{\eta} f_E'(x) \ \mathrm{Im}\,(m_N(x+i\tilde{\eta}) - m_{sc}(x+i\tilde{\eta}))\, dx$$

$$- \int dx \int_{\tilde{\eta} < y \le 1} dy \quad [\chi(y) + y\chi'(y)]\, f_E'(x) \ \mathrm{Im}(m_N(z) - m_{sc}(z))\, dx\, dy$$

$$= J_1 + J_2 + J_3 .$$

Using the averaged local law, $|m_N(x+i\tilde{\eta}) - m_{sc}(x+i\tilde{\eta})| \prec \frac{1}{N\tilde{\eta}}$ and that $f_E'(x)$ is of order $O(\tilde{\eta}^{-1})$ on two intervals of size $\tilde{\eta}$, we get that

$$|J_1| \prec C \frac{1}{N\tilde{\eta}} \cdot \tilde{\eta} \cdot \frac{1}{\tilde{\eta}} \cdot \tilde{\eta} = O\left(\frac{1}{N}\right).$$

$J_3$ can be bounded in the same way as III: $|J_3| \prec \frac{1}{N} \times \frac{1}{\tilde{\eta}} \times \tilde{\eta} = \frac{1}{N}$.

Finally, $J_1$ can be bounded as

$$|J_1| \prec \frac{1}{\tilde{\eta}} \times \tilde{\eta} \times \int_{\tilde{\eta} < y \le 1} \chi(y) \frac{1}{Ny}\, dy \le \frac{1}{N} \log \frac{1}{\tilde{\eta}} \le \frac{\log N}{N} \prec \frac{1}{N}.$$

To summarize, we have shown that

$$\int f_E(x)\, [\rho_{H_N}(x) - \rho_{sc}(x)]\, dx \prec \frac{1}{N}.$$

Our goal is to bound $\int 1_{(-\infty, E]}(x)\, [\rho_{H_N}(x) - \rho_{sc}(x)]\, dx$. We bound

$\int |f_E(x) - 1_{(-\infty,E]}(x)|\, \rho_{H_N}(x)\, dx$ and $\underbrace{\int |f_E(x) - 1_{(-\infty,E]}(x)|\, \rho_{sc}(x)}\, dx$ separately.

$\le \int_E^{E+\tilde{\eta}} \rho_{H_N}(x)\, dx$ w.h.p. because there is no eigenvalue of $H_N$ inside $(-\infty, 3]$.

$\le C\tilde{\eta}$ because $\rho_{sc}(x) = 0$ on $(-\infty, 3]$, $\rho_{sc}(x) \le C$, $|f_E(x) - 1_{(-\infty,E]}(x)|$ is non-zero on two intervals of size $\tilde{\eta}$.

Moreover, $\int_E^{E+\tilde{\eta}} \rho_{H_N}(x)\, dx \le \int [f_{E+\tilde{\eta}}(x) - f_{E-\tilde{\eta}}(x)]\, \rho_{H_N}(x)\, dx$

$$\prec \frac{1}{N} + \int [f_{E+\tilde{\eta}}(x) - f_{E-\tilde{\eta}}(x)]\, \rho_{sc}(x)\, dx \le C\tilde{\eta}.$$

In sum, we obtain that $|F_N(E) - F_{sc}(E)| = \left|\int 1_{(-\infty,E]}(x)\,(\rho_{H_N}(x) - \rho_{sc}(x))\,dx\right| \prec \frac{1}{N}$.
The above arguments are uniformly in E because the local law is uniform in E. □

④ Rigidity of eigenvalues (Proof of Theorem 4.1)
Recall that

$$\frac{j}{N} = \int_{-\infty}^{\lambda_j} \rho_N(x)\,dx, \qquad \frac{j}{N} = \int_{-\infty}^{\gamma_j} \rho_{sc}(x)\,dx. \quad (*)$$

We already know that $\int_{-\infty}^{E}[\rho_{H_N}(x) - \rho_{sc}(x)]\,dx$ is small. We want to turn this estimate into bounds of $|\lambda_j - \gamma_j|$.
By symmetry, we consider $j \le N/2$. We aim to show that

$$|\lambda_j - \gamma_j| \prec j^{-1/3} N^{-2/3} \quad \text{uniformly in } 1 \le j \le \frac{N}{2}.$$

From (*), we get $0 = \int_{-\infty}^{\lambda_j} \rho_N(x)\,dx - \int_{-\infty}^{\gamma_j} \rho_{sc}(x)\,dx = \int_{-\infty}^{\lambda_j} (\rho_N(x) - \rho_{sc}(x))\,dx - \int_{\lambda_j}^{\gamma_j} \rho_{sc}(x)\,dx$.

Hence, $\left|\int_{\lambda_j}^{\gamma_j} \rho_{sc}(x)\,dx\right| \le |F_N(\lambda_j) - F_{sc}(\lambda_j)| \prec \frac{1}{N}$.

Note that $\rho_{sc}(x) \sim \sqrt{(x+2)(2-x)} \sim \sqrt{\kappa_x}$ for $x \in [-2,1]$. Then, we can derive that $F_{sc}(x) \sim \kappa_x^{3/2}$ and $\rho_{sc}(x) \sim \sqrt{\kappa_x} \sim F_{sc}(x)^{1/3}$, $x \in [-2,1]$. This implies that,

$$\gamma_j + 2 \sim \left(\frac{j}{N}\right)^{2/3}, \quad \rho_{sc}(\gamma_j) \sim \left(\frac{j}{N}\right)^{1/3}, \quad 1 \le j \le \frac{N}{2}.$$

[ If we know that $\rho_{sc}(\gamma_j)$ and $\rho_{sc}(\lambda_j)$ are comparable, then we have

$$\frac{1}{N} \succ \left|\int_{\lambda_j}^{\gamma_j} \rho_{sc}(x)\,dx\right| \gtrsim \rho_{sc}(\gamma_j)\,|\gamma_j - \lambda_j| \Rightarrow |\gamma_j - \lambda_j| \prec \frac{1}{N}\frac{1}{\rho_{sc}(\gamma_j)} = j^{-\frac{1}{3}} N^{-\frac{2}{3}}.$$ ]

To give a rigorous proof, we first consider $j \ge N^{\frac{\varepsilon}{2}}$. Since $F_{sc}(\gamma_j) = \frac{j}{N} \ge N^{-1+\frac{\varepsilon}{2}}$ and $F_{sc}(\gamma_j) = F_N(\lambda_j)$, we have

$$F_N(\lambda_j) = F_{sc}(\lambda_j) + O_\prec\left(\frac{1}{N}\right), \qquad |F_{sc}(\lambda_j) - F_{sc}(\gamma_j)| \prec \frac{1}{N} \le N^{-\frac{\varepsilon}{2}} F_{sc}(\gamma_j).$$

Hence, $|F_{sc}(\lambda_j) - F_{sc}(\gamma_j)| \le N^{-\frac{\varepsilon}{4}} F_{sc}(\gamma_j)$ w.h.p. This shows that $F_{sc}(\lambda_j) \sim F_{sc}(\gamma_j)$, implying that $\rho_{sc}(\lambda_j) \sim \rho_{sc}(\gamma_j)$. This completes the proof.

For indices $j \le N^{\varepsilon/2}$, we use the trivial bound: $\lambda_1 \le \lambda_j \le \lambda_{N^{\varepsilon/2}}$.
But we already know that $\lambda_1 \ge -2 - N^{-2/3+\varepsilon/2}$ w.h.p. and $\lambda_{N^{\varepsilon/2}} \le -2 + N^{\varepsilon/6}\left(\frac{N^{\varepsilon/2}}{N}\right)^{2/3} = -2 + N^{-\frac{2}{3}+\varepsilon/2}$ w.h.p.

In addition, we have $|\gamma_j + 2| \le C\left(\frac{N^{\varepsilon/2}}{N}\right)^{2/3} = CN^{-\frac{2}{3}+\frac{\varepsilon}{3}}$.

Hence, $|\lambda_j - \gamma_j| \le CN^{-\frac{2}{3}+\frac{\varepsilon}{2}} \le N^{-\frac{2}{3}+\varepsilon}$ w.h.p.