

$$= \frac{1}{2} \sum_{i,j} \| \vec{v}_i \|_j^2 + \frac{1}{2N} \sum_{i \neq j} \frac{(v_i(i) - v_j(j))^2}{(\lambda_i - \lambda_j)^2} \geq \frac{1}{2} \| \vec{v} \|_j^2.$$

Thus, $\nabla^2(\beta N f)$ $\geq \frac{1}{2} \beta N$. Here, we use a different normalization for the Dirichlet form: $D_\mu(f) = \frac{1}{\beta N} \int \| \nabla f \|^2 d\mu$. Then, the Bakry-Emery estimate gives $S_\mu(f) \leq 4 D_\mu(\sqrt{f})$.

This shows that the relaxation time is of order 1. (This is the global dynamics.)

We will show that the local dynamics converges to equilibrium ^{within} ~~at~~ time of order $N^{-1+\delta}$.

Thm 6.6. Assume the following conditions:

(i) There exists a constant $\beta > 0$ such that

$$Q := \sup_{0 \leq t \leq N} \frac{1}{N} \int \sum_{j=1}^N (\gamma_j - \bar{\gamma}_j)^2 f_t(d\bar{x}) \mu_G(d\bar{x}) \leq C N^{-2+2\beta}.$$

(ii) After time $1/N$, the solution to the equation $\partial_t f_t = L f_t$, $t \geq 0$, satisfies $S_\mu(f_{1/N}) \leq C N^m$ for some fixed m .

Fix any $\beta \geq 1$ and $n \in \mathbb{N}$. Let $F: \mathbb{R}^n \rightarrow \mathbb{R}$ be a smooth function with compact support.

Denote:

$$f_{i,n}(\bar{x}) = F(N(x_{i+1} - \bar{x}), N(x_{i+2} - \bar{x}), \dots, N(x_{i+n} - \bar{x})).$$

Then, for any $\beta \in (0, \frac{1}{2})$ and any sufficiently small constant $\epsilon > 0$, there exist constants $C, c > 0$ depending only on ϵ and F , such that for any $J \in \{1, 2, \dots, N-n\}$ we have

$$\sup_{t \geq N^{-1+2\beta+\epsilon}} \left| \int \frac{1}{|J|} \sum_{i \in J} f_{i,n}(\bar{x}) (f_t d\mu - d\mu) \right| \leq C N^\epsilon \sqrt{\frac{N^2 Q}{|J| t}} + C e^{-c N^\epsilon}.$$

In particular, if $t \geq N^{-1+2\beta+2\epsilon+\delta}$, we have

$$\left| \int \frac{1}{|J|} \sum_{i \in J} f_{i,n}(\bar{x}) (f_t d\mu - d\mu) \right| \leq \frac{C}{\sqrt{|J| N^{\delta-1}}} + C e^{-c N^\epsilon}.$$

Hence, the gap distribution, averaged over J indices, coincides for $f_t d\mu$ and $d\mu$ if $|J| N^{\delta-1} \rightarrow \infty$.

Averaged gap distribution on scale J corresponds to averaged energy distribution on scale $J \sim N^\delta$.

Lemma. Under the setting of Thm 6.6, suppose $\left| \int \frac{1}{|J|} \sum_{i \in J} f_{i,n}(\bar{x}) (f_t d\mu - d\mu) \right| \leq \frac{C}{\sqrt{|J| N^{\delta-1}}}$ for some constant $\delta > 0$. (62)

Suppose the eigenvalue rigidity holds: $\forall K > 0$, $\sup_{K \leq j \leq (1-K)N} |\lambda_j - \bar{\lambda}_j| < N^{-1+\delta}$.

Then, for any constant $\varepsilon > 0$ and $N^{-1} \ll b_N \ll 1$, we have that

$$\left| \int_{E-b}^{E+b} \frac{dE'}{2b} \int_{\mathbb{R}^n} d\vec{\alpha} O(\vec{\alpha}) (P_{f\mu, N}^{(n)} - P_{\mu, N}^{(n)}) (E' + \frac{\vec{\alpha}}{N p_{sc}(E)}) \right| \\ \leq N^{2\varepsilon} \left[\frac{N^{-1+\delta}}{b} + \sqrt{\frac{N^{-\delta}}{b}} \right].$$

Hence, the averaged energy distribution over $b > N^{-1+2\varepsilon+2\delta} \sqrt{N^{-\delta}} \approx N^{-1+2\varepsilon+2\delta}$ coincides for $f\mu$ and μ .

The above two results establish the averaged gap/energy universality for $t \geq N^{-1+2\delta+2\varepsilon+\delta}$. Under the optimal rigidity, δ can be chosen as small as possible. Furthermore, ε and δ are arbitrary. This establishes the Dyson's conjecture: the local dynamics relaxes to equilibrium for $t \geq N^{-1+c}$ \forall constant $c > 0$.

Main ideas for the proof: Recall that

$$\langle \vec{v}, \nabla^2 f \vec{v} \rangle = \frac{1}{2} \sum_i v(i)^2 + \frac{1}{N} \sum_{i,j} \frac{(v(i) - v(j))^2}{(\lambda_i - \lambda_j)^2} \geq \frac{1}{2} \| \vec{v} \|^2.$$

The key intuition is that the relaxation time is in fact much shorter than order 1 for local variables that depend only on the eigenvalue differences. In particular, with the intuition that $\lambda_i - \lambda_j$ are close, the relaxation time in the direction $v(i) - v(j)$ should be much smaller than 1. However, this effect is hard to use directly in the real proof.

An idea is to add an "auxiliary strongly convex" potential to $f\mu$ to "speed up" the relaxation to local equilibrium. On the other hand, we will show that the effect of this speeding up on the local statistics can be controlled provided with the estimate

$$(+) \sup_{0 \leq t \leq N} \frac{1}{N} \int \sum_{j=1}^N (\lambda_j - \bar{\lambda}_j)^2 f_t(\vec{\lambda}) \mu_G(d\vec{\lambda}) \lesssim N^{-2+2\delta}.$$

We introduce the following auxiliary potential $\tilde{f}(\vec{\lambda}) := \frac{1}{2\tau} \sum_{j=1}^N (\lambda_j - \bar{\lambda}_j)^2$, $0 < \tau < 1$. (We will choose $\tau = t N^{-\varepsilon}$ in the proof.)

Then, the new Hamiltonian is $\tilde{H} = H + \tilde{f}\mu$, and the corresponding measure is $d\omega = w(\vec{\lambda}) d\vec{\lambda}$, $w(\vec{\lambda}) := e^{-\beta N \tilde{f}(\vec{\lambda})} / \tilde{Z}$, called the "local relaxation measure". The "local relaxation flow" with the generator characterized by the natural Dirichlet form w.r.t. w : $\tilde{L} = L - \sum_j b_j \partial_j$, $b_j |x_j| = \frac{x_j - \bar{x}_j}{\tau}$.

For small τ , $\tilde{f}\mu$ will substantially increase the lower bound on the Hessian, hence speeding up the dynamics so that the relaxation time is at most τ . In addition, we will compare the local statistics of the original system with those of the modified one.

It turns out that the difference is governed by $\|\nabla f_t\|^2$, which can be controlled with the estimate (†).

Proof of Theorem 6.6

Proposition I. Fix any $\beta \geq 1$. Consider the equation $\partial_t g_t = \tilde{L} g_t$, $t \geq 0$, with the reversible measure $w = e^{-\beta N \tilde{L}} / Z_N$. Suppose the initial condition g_0 satisfies $\int g_0 dw = 1$. Then, we have that

$$(1) \quad \partial_t D_w(\sqrt{g_t}) \leq -\frac{2}{t} D_w(\sqrt{g_t}) - \frac{1}{\beta N^2} \int \sum_{i,j=1}^N \frac{(\partial_i \sqrt{g_t} - \partial_j \sqrt{g_t})^2}{(x_i - x_j)^2} dw,$$

$$(2) \quad \frac{1}{\beta N^2} \int_0^\infty ds \int \sum_{i,j=1}^N \frac{(\partial_i \sqrt{g_s} - \partial_j \sqrt{g_s})^2}{(x_i - x_j)^2} dw \leq D_w(\sqrt{g_0}),$$

and the LSI $S_w(g_0) \leq C \tau D_w(\sqrt{g_0})$. Hence, the relaxation time to equilibrium is of order τ :

$$(3) \quad S_w(g_t) \leq e^{-Ct/\tau} S_w(g_0), \quad D_w(\sqrt{g_t}) \leq e^{-Ct/\tau} D_w(\sqrt{g_0}).$$

Proof: Performing the same calculation as in the proof of the Bakry - Emery estimate, we obtain that for $h = \sqrt{g_t}$,

$$\partial_t D_w(h_t) = \partial_t \frac{1}{\beta N} \int (\nabla h_t)^2 e^{-\beta N \tilde{L}} d\tilde{x} \stackrel{-\frac{2}{\beta N}}{\leq} \int (\nabla h_t) \cdot (\nabla^2 \tilde{L}) \cdot (\nabla h_t) e^{-\beta N \tilde{L}} d\tilde{x}.$$

For \tilde{L} , we can calculate its Hessian as

$$\nabla h \cdot \nabla^2 \tilde{L} \cdot \nabla h \geq \frac{1}{t} \sum_j (\partial_j h)^2 + \frac{1}{2N} \sum_{i,j} \frac{1}{(x_i - x_j)^2} (\partial_i h - \partial_j h)^2.$$

Plugging it into the above inequality, we conclude (1). Integrating over t , we conclude (2). The LSI and (3) can be derived in the same way as in the proof of Bakry - Emery. \square

The extra term on the RHS of (1) plays a key role in the proof of the following prop.

Proposition II. Let g be a probability density w.r.t. the local relaxation measure w , i.e., $\int g dw = 1$. Fix any $n \geq 1$ and a smooth function F with compact support. Recall that $F_{i,n}(\vec{x}) := F(N(x_{i+1} - x_i), \dots, N(x_{i+n} - x_i))$.

Then, for any $J \subseteq \{1, 2, \dots, N-n\}$ and $t \geq 0$, we have

$$\left| \int \frac{1}{|J|} \sum_{i \in J} F_{i,n}(\vec{x}) (g dw - dw) \right| \leq C \left(t \frac{D_w(\sqrt{g})}{|J|} \right)^{1/2} + C \sqrt{S_w(g)} e^{-Ct/\tau}.$$

Proof: For simplicity of notation, we take $n=1$, so $F_{i,1}(\vec{x}) = F(N(x_{i+1} - x_i))$. Let g_t be a solution to $\partial_t g_t = \tilde{L} g_t$, $t \geq 0$, with initial condition $g_0 = g$. Then,

$$\int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) (g - g_t) dw = \int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) (g - g_t) dw +$$

$$\int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) (g_t - 1) dw.$$

The second term can be bounded by

$$\int |g_t - 1| d\omega \stackrel{\uparrow}{\leq} C \sqrt{S_\omega(g_t)} \stackrel{\downarrow}{\leq} C e^{-ct/c} \sqrt{S_\omega(g_0)}.$$

Pinsker inequality

To estimate the first term, using $\partial_t g_t = \tilde{L} g_t$, we get that

$$\begin{aligned} & \int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) g_t d\omega - \int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) g_0 d\omega \\ &= \int_0^t ds \int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) \frac{g_t - g_0}{ds} d\omega \approx g_s d\omega \\ &= \int_0^t ds \int \frac{1}{|J|} \sum_{i \in J} G(N(x_{i+1} - x_i)) \tilde{L} g_s d\omega \\ &= \int_0^t ds \int \frac{1}{|J|} \sum_{i \in J} G'(N(x_{i+1} - x_i)) (\partial_{i+1} g_s - \partial_i g_s) d\omega. \quad (*) \end{aligned}$$

Using ~~$\partial_i g = 2\sqrt{g} \partial_i \sqrt{g}$~~ and the Cauchy-Schwarz inequality, we obtain that

$$\begin{aligned} |(*)| &\leq 2 \left[\int_0^t ds \int \frac{N^2}{|J|^2} \sum_{i \in J} |G'(N(x_{i+1} - x_i))|^2 (x_{i+1} - x_i)^2 g_s d\omega \right]^{1/2} \times \\ &\quad \underbrace{\left[\int_0^t ds \int \frac{1}{N^2} \sum_i \frac{1}{(x_i - x_{i+1})^2} (\partial_{i+1} \sqrt{g_s} - \partial_i \sqrt{g_s})^2 d\omega \right]^{1/2}} \\ &\leq C D_\omega(\sqrt{g_0}) \end{aligned}$$

Moreover, $|G'(N(x_{i+1} - x_i))|^2 |x_{i+1} - x_i|^2 \leq CN^{-2}$ because G is smooth and compactly supported. So

$$|(*)| \leq C \sqrt{D_\omega(g_0)} \cdot \left(\frac{t}{|J|} \right)^{1/2}. \quad \square$$

Proposition III. Consider the local relaxation measure ω for a given $\tau > 0$. Set $\gamma = \omega/\mu$ and $g_t = f_t/\gamma$ with f_t solving $\partial_t f_t = L f_t$. Suppose there is a const. $C > 0$ such that $S(f_t \mu | \omega) \leq N^C$.

Fix any small $\epsilon > 0$. For any $t \in [\tau N^\epsilon, N]$, the entropy and Dirichlet form satisfy

$$S(g_t \omega | \omega) \leq CN^2 Q \tau^{-1}, \quad D_\omega(\sqrt{g_t}) \leq CN^2 Q \tau^{-2},$$

where recall that $Q = \sup_{0 \leq t \leq N} \frac{1}{N} \int \sum_{j=1}^N (\lambda_j - \gamma_j)^2 f_t(\vec{\lambda}) \mu(d\vec{\lambda}) \quad (\leq CN^{-2+2\beta})$.

Proof: We calculate $\partial_t S(f_t \mu | \omega) = \partial_t S(f_t \mu | \gamma \mu)$

$$= \int (L f_t)(\log g_t) d\mu + \int f_t \frac{L f_t}{f_t} d\mu = \int (L f_t)(\log g_t) d\mu$$

$$= \int f_t L(\log g_t) d\mu = \int g_t + 4 L(\log g_t) d\mu = \int \gamma_t + 4 [g_t L(\log g_t) - g_t \frac{L g_t}{g_t}] d\mu$$

$$+ \int 4 L g_t d\mu.$$

We can check directly that $\ell(\log g_t) - \frac{\log g_t}{g_t} = -\frac{1}{\beta N} \frac{|\nabla g_t|^2}{g_t^2} = -\frac{4}{\beta N} |\nabla \sqrt{g_t}|^2 \cdot \frac{1}{g_t}$.

$$\begin{aligned} \text{Hence, we have } \partial_t S(f_t, \mu | w) &= -\frac{4}{\beta N} \int 4 |\nabla \sqrt{g_t}|^2 d\mu + \int 4 \log g_t d\mu, \quad (b_j = \frac{x_j - \bar{x}_j}{\tau}) \\ &= -\frac{4}{\beta N} \int |\nabla \sqrt{g_t}|^2 dw + \int \cancel{4 \log g_t}^0 d\mu + \sum_j \int b_j \partial_j g_t dw. \end{aligned}$$

Using $\partial_j g_t = 2 \sqrt{g_t} \partial_t \sqrt{g_t}$ and C-S, we get

$$\begin{aligned} \partial_t S(f_t, \mu | w) &\leq -4 D_w(\sqrt{g_t}) + 2 \sum_j \int b_j \sqrt{g_t} \partial_j \sqrt{g_t} dw \\ &\leq -2 D_w(\sqrt{g_t}) + C N \sum_j \int b_j^2 g_t dw \\ &= -2 D_w(\sqrt{g_t}) + C \frac{N^2}{\tau^2} \frac{1}{N} \int (x_j - \bar{x}_j)^2 g_t dw \\ &\leq -2 D_w(\sqrt{g_t}) + C N^2 Q \tau^{-2}. \quad (*) \end{aligned}$$

Applying the LSI w.r.t. dw , we get that

$$\partial_t S(f_t, \mu | w) \leq -C \tau^{-1} S(f_t, \mu | w) + C N^2 Q \tau^{-2}.$$

Applying the Gronwall's inequality, integrating this inequality from τ to t and using the assumption $S(f_\tau, \mu | w) \leq N^C$ and $t \geq \tau N^\varepsilon$, we obtain that

$$S(g_t, \mu | w) = S(f_t, \mu | w) \leq C N^2 Q \tau^{-1}. \quad (**) \quad \text{[Wavy line]} \quad \text{[Wavy line]}$$

To prove the second estimate, we notice

$$\begin{aligned} D_\mu(\sqrt{f_s}) &= \frac{1}{\beta N} \int \frac{|\nabla(g_s \psi)|^2}{g_s \psi} \frac{dw}{\psi} \leq \frac{C}{\beta N} \int \left[\frac{|\nabla g_s|^2}{g_s} + |\nabla \log \psi|^2 g_s \right] dw \\ &\leq C D_w(\sqrt{g_s}) + \frac{C}{\beta N} \int \frac{|\nabla(x_j - \bar{x}_j)|^2}{\tau^2} dt \leq C \frac{(\beta N)^2}{\beta N} \int \frac{|\nabla(x_j - \bar{x}_j)|^2}{\tau^2} dt \\ &\leq C D_w(\sqrt{g_s}) + C N^2 Q \tau^{-2}. \quad (*) \end{aligned}$$

Taking integral over $[t-\tau, t]$ ~~over w~~ , we get

$$\int_{t-\tau}^t D_\mu(\sqrt{f_s}) ds \geq \int_{t-\tau}^t \left[\frac{1}{C} D_\mu(\sqrt{f_t}) - C N^2 Q \tau^{-2} \right] ds$$

$$(D_\mu(\sqrt{f_t}) \text{ is decreasing in } t) \Leftrightarrow \frac{1}{C} D_\mu(\sqrt{f_t}) - C N^2 Q \tau^{-1}.$$

On the other hand, with (*) and (**), we get that

$$\int_{t-\tau}^t D_\mu(\sqrt{g_s}) ds \leq C N^2 Q \tau^{-1} + [S(f_t, \mu | w) - S(f_{t-\tau}, \mu | w)] \leq C N^2 Q \tau^{-1}.$$

From the above two inequalities, we obtain that

$$D_\mu(\sqrt{f_t}) \leq C N^2 Q \tau^{-2}.$$

With a similar argument as in (*), we get that

$$D_W(\sqrt{g_t}) \leq C D_W(\sqrt{f_t}) + CN^2 Q t^{-2} \leq CN^2 Q t^{-2}. \quad \square$$

Theorem 6.6 is a consequence of the above three propositions. We choose $\tau = tN^{-\epsilon}$ (recall that $t \geq N^{-1+2\beta+2\epsilon+\delta}$), $\psi = w/\mu$ and $g_t = f_t/\psi$. To apply Prop. III, we still need to verify the assumption $S(f_{t\tau} \mu | w) \leq N^C$. By definition,

$$\begin{aligned} S(f_{t\tau} \mu | w) &= \int f_{t\tau} \log f_{t\tau} d\mu - \int f_{t\tau} \log \psi d\mu \\ &= S_\mu(f_{t\tau}) - \int f_{t\tau} \log \psi d\mu, \end{aligned}$$

where $-\log \psi = \frac{\beta N}{2\tau} \sum_{j=1}^N (x_j - \bar{x}_j)^2 + \log \frac{\tilde{Z}}{Z}$. We know $\tilde{Z} \leq Z$ since $\tilde{f}_{t\tau} \geq f_t$. Hence,

$$-\int f_{t\tau} \log \psi d\mu \leq CN\tau^{-1} \int \sum_{j=1}^N (x_j - \bar{x}_j)^2 f_{t\tau}(\bar{x}_j) d\mu(\bar{x}_j) \leq CN^2 \tau^{-1} Q \leq N^C.$$

It remains to bound $S_\mu(f_{t\tau})$ for a "good initial condition" as a Wigner matrix.

Lemma: Let $\beta=1, 2$. Suppose the initial data f_0 of the DBM is given by the eigenvalue distribution of a Wigner matrix. Then, $\forall \tau \geq 0$, we have

$$S_\mu(f_{t\tau}) \leq CN^2 [1 - \log(1 - e^{-\tau})].$$

Proof: We only consider the real case with $\beta=1$. Recall that $f_{t\tau} \mu$ is the eigenvalue distribution of $H_t = e^{-\tau/2} H_0 + (1 - e^{-\tau})^{1/2}$ GOE. Notice that $f_{t\tau} \mu$ is the marginal of μ_{H_t} by integrating out the eigenvector distribution, and μ is the marginal of GOE. Then, using the fact that the entropy is decreasing w.r.t. taking a marginal, we get

$$S_\mu(f_{t\tau}) \leq S(\mu_{H_t} \otimes \mu_{\text{GOE}}).$$

Notice that both μ_{H_t} and μ_{GOE} are product measures of the laws of matrix elements. Then, using the additivity of entropy, $S(\mu_{H_t} \otimes \mu_{\text{GOE}})$ is equal to the sum of the relative entropies of the matrix elements. Let's consider the off-diagonal elements:

for $\gamma = 1 - e^{-\tau}$, let p_γ be the probability density of $(1-\gamma)^{1/2} g = e^{-\tau/2} g$, where g is the random variable for an off-diagonal matrix element. Then, the prob. density of an off-diagonal element of H_t is given as: $g_t = p_\gamma * g_{\text{GOE}}/N$, where g_{GOE} denotes the Gaussian distribution

$$g_{\text{GOE}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2}).$$

Then, we calculate that

$$S(g_t | g_{\text{GOE}}/N) = S\left(\int dy p_\gamma(y) g_{\text{GOE}}(\cdot - y) | g_{\text{GOE}}/N\right)$$

$$(\text{Jensen}) \leq \int dy p_\gamma(y) S(g_{\text{GOE}}(\cdot - y) | g_{\text{GOE}}/N).$$

A direct calculation yields that

$$S(g_{\text{GOE}}(\cdot - y) | g_{\text{GOE}}/N) = \log \frac{N}{6} + \frac{y^2}{2\sigma^2} + \frac{6^2}{2N^2} - \frac{1}{2}.$$

Applying it to our case, we get

$$S(g_{\gamma/N}(\cdot - y) | g_{1/N}) = \frac{1}{2}(Ny^2 - \log r + r - 1).$$

Thus, $S(g_t | g_{1/N}) \leq \int dy p_\gamma(y) \frac{1}{2}(Ny^2 - \log r + r - 1)$. Using $\int y^2 p_\gamma(y) dy = \frac{1}{N}$, we conclude the proof. \square

With the above lemma, we can apply Prop. II & III to get that (with $g = f_t \otimes g$, $g dw = f_t d\mu$, $t = \tau N^{-\varepsilon}$)

$$\left| \int \frac{1}{|J|} \sum_{i \in J} f_{i,n}(\bar{x}) (f_t d\mu - g dw) \right| \leq C \left(t \frac{D_w(\sqrt{g})}{|J|} \right)^{1/2} + C \sqrt{S_w(g)} e^{-cN^\varepsilon}$$

$$[D_w(\sqrt{g}) = D_w(\sqrt{g_t})] \rightarrow \leq C \left(t \frac{N^2 Q}{|J| \tau^2} \right)^{1/2} + N^\varepsilon e^{-cN^\varepsilon} \leq C N^\varepsilon \left(\frac{N^2 Q}{|J| t} \right)^{1/2} + C e^{-cN^\varepsilon}.$$

Hence, $f_t d\mu$ and dw are close for any initial data. Applying this estimate to the Gaussian initial data, we can also compare $d\mu$ and dw . This concludes the proof.

Section 7 Green function comparison

We now finish Step 3 of the three step strategy. That is, given a Wigner H , we can find a Wigner H_0 and $t \geq N^{-1+\epsilon}$ such that the local eigenvalue statistics of H_t match those of $H_t = e^{-t/2} H_0 + (-e^{-t})^{1/2} H_G$, $H_G \stackrel{d}{=} \text{GOE/GUE}$. We have shown the bulk universality of H_t . Then, Step 3 implies the bulk universality of H .

To compare local statistics of two Wigner ensembles, we use the following "correlation function comparison theorem". For the detailed setting, suppose we have two Wigner ensembles, labelled by v and w , respectively: H^v and H^w . Denote their eigenvalues by λ_i^v and λ_i^w , and the corresponding Stieltjes transforms by m^v and m^w , i.e.,

$$m^v(z) = \frac{1}{N} \sum_{j=1}^N \frac{1}{\lambda_j^v - z}.$$

Let $P_{v,N}^{(n)}$ and $P_{w,N}^{(n)}$ be the n -point correlation functions of $\{\lambda_i^v\}$ and $\{\lambda_i^w\}$.

Theorem 7.1 Let $\kappa' > 0$ be a small const. Suppose for some small constants $\delta, \theta > 0$ such that the following two conditions hold.

(i) $\forall \epsilon > 0$ and fixed $k \in \mathbb{N}$,

$$|\mathbb{E} [\operatorname{Im} m^v(E + iN^{-1+\epsilon})]^k + \mathbb{E} [\operatorname{Im} m^w(E + iN^{-1+\epsilon})]^k| \leq C$$

for $|E| \leq 2 - \kappa$.

(ii) For any sequence $\beta_j = E_j \pm i\eta_j$, $j = 1, \dots, n$, with $|E_j| \leq 2 - \kappa$ and $\eta_j = N^{-1-\theta}$ for some $\theta \leq \delta$, we have

$$|\mathbb{E} [\operatorname{Im} m^v(\beta_1) \cdots \operatorname{Im} m^v(\beta_n)] - \mathbb{E} [\operatorname{Im} m^w(\beta_1) \cdots \operatorname{Im} m^w(\beta_n)]| \leq N^{-\delta}. \quad (*)$$

Then, for any $n \in \mathbb{N}$, there are constants $C_n = C_n(\delta, \theta)$ s.t. for any $|E| \leq 2 - 2\kappa$ and C^1 function $O: \mathbb{R}^n \rightarrow \mathbb{R}$ with compact support,

$$\left| \int_{\mathbb{R}^n} O(\vec{a}) (P_{v,N}^{(n)} - P_{w,N}^{(n)}) (E + \frac{\vec{a}}{N}) d\vec{a} \right| \leq C_n N^{-C_n}.$$

Roughly speaking, this theorem says that if the finite dimensional marginals of the Stieltjes transforms are close, then the local correlation functions are close. It is crucial that in $(*)$, β_i has imaginary part $N^{-1-\theta}$ below the scale of bulk eigenvalue gaps N^{-1} . Controlling the Stieltjes transforms on this very short scale is necessary to identify and compare local correlation functions at the scale N^{-1} .

The condition $(*)$ can be replaced by a condition that is easier to check: let $F: \mathbb{R}^n \rightarrow \mathbb{R}$ be a function such that for any multi-index $a = (a_1, \dots, a_n)$ with $1 \leq |a| = |a_1| + \dots + |a_n| \leq 5$ and sufficiently small constant $\epsilon > 0$,

$$\max \{ |\partial^2 F(x_1, \dots, x_n)| : \max_j |x_j| \leq N^\varepsilon \} \leq N^{C_0 \varepsilon},$$

$$\max \{ |\partial^3 F(x_1, \dots, x_n)| : \max_j |x_j| \leq N^2 \} \leq N^{C_0},$$

for some constant C_0 . For any such F , we have

$$(+) \quad |IE F(Im m^V(z_1) \dots Im m^V(z_n)) - IE F(Im m^W(z_1), \dots, Im m^W(z_n))| \leq N^{-\delta}.$$

Given (+), we can approximate $IE [Im m^W(z_1) \dots Im m^W(z_n)]$ by $IE F(Im m^W(z_1), \dots, Im m^W(z_n))$ with $F(x_1, \dots, x_n) = x_1 x_2 \dots x_n$ if $\max_j |x_j| \leq N^c$ and it is smoothly cutoff to 0 in the regime $\max_j |x_j| \geq N^c$ for some small constant $c > 0$.

To see this, recall that $\eta Im m(E+i\eta) \leq \eta' Im m(E+i\eta')$ for $0 < \eta \leq \eta'$. Then, (i) of Theorem 7.1 implies that for any $\eta \leq N^{-1+\varepsilon}$,

$$IE [\eta Im m(E+i\eta)]^k \leq IE [N^{-1+\varepsilon} Im m(E+iN^{-1+\varepsilon})]^k \\ \leq (CN^{-1+\varepsilon})^k$$

$$\Rightarrow IE [Im m(E+i\eta)]^k \leq (C \frac{N^\varepsilon}{N\eta})^k \quad \text{for any fixed } k \in \mathbb{N}.$$

Let $\eta = N^{-1-\varepsilon}$ and $z = E+i\eta$. Consider $n=1$ for simplicity. By choosing $c = \frac{26}{22}$,

$$|IE F(Im m^V(z)) - IE Im m^V(z)| \leq IE Im m^V(z) \cdot \mathbf{1}(Im m^V(z) \geq N^c) \\ \leq N^{-C(k-1)} IE [Im m^V(z)]^k \leq CN^{k(6+\varepsilon)-C(k-1)} \leq N^{-\delta}$$

as long as ε is small enough and k is large enough. Similar arguments apply to general n .

The conditions on F and $\partial^3 F$ will be used in the Green's function comparison argument. The proof of Theorem 7.1 can be found in Theorem 10.15.3 of Erdős-Yau.

I. The Lindeberg replacement trick

The Green's function comparison uses a classical "Lindeberg replacement trick", which was developed in Lindeberg's own proof of CLT. To explain the basic idea of this argument and the importance of moment matching conditions, we use it to prove the following version of the Berry-Esseen theorem.

Theorem 7.2 (Berry-Esseen theorem) Let X be a random variable with mean 0, variance 1, and finite 3rd moment. Let φ be smooth with uniformly bounded third-order derivatives up to third order. Let X_1, \dots, X_n be i.i.d. copies of X and

$$Z_n := \frac{X_1 + \dots + X_n}{\sqrt{n}}.$$

Then, we have that

$$\mathbb{E} \varphi(Z_n) = \mathbb{E} \varphi(G) + O\left(\frac{1}{\sqrt{n}} \mathbb{E} |x|^3 \cdot \|\varphi''\|_{\infty}\right),$$

where $G \sim N(0, 1)$ is standard Gaussian.

Proof: Note that G has the same distribution as $\frac{Y_1 + \dots + Y_n}{\sqrt{n}}$, where Y_i 's are independent $N(0, 1)$ random variables. Hence, we only need to bound

$$\mathbb{E} \varphi\left(\frac{X_1 + \dots + X_n}{\sqrt{n}}\right) - \mathbb{E}\left(\frac{Y_1 + \dots + Y_n}{\sqrt{n}}\right) = \sum_{i=0}^{n-1} [\mathbb{E} \varphi(Z_{n,i+1}) - \mathbb{E} \varphi(Z_{n,i})],$$

where $Z_{n,i} = \frac{x_1 + \dots + x_i + Y_{i+1} + \dots + Y_n}{\sqrt{n}}$, i.e., we replace ~~the first~~ y_1, \dots, y_i with x_1, \dots, x_i in $Z_{n,i}$.

Then, it suffices to show that

$$\mathbb{E} \varphi(Z_{n,i+1}) - \mathbb{E} \varphi(Z_{n,i}) = O\left(\mathbb{E} |x|^3 \frac{\|\varphi''\|_{\infty}}{n^{3/2}}\right).$$

Note that $Z_{n,i+1}$ is different from $Z_{n,i}$ only by one entry, i.e., a $\frac{Y_{i+1}}{\sqrt{n}}$ is replaced by $\frac{x_{i+1}}{\sqrt{n}}$.

Define

$$S_{n,i} = \frac{x_1 + \dots + x_i + Y_{i+2} + \dots + Y_n}{\sqrt{n}}, \quad Z_{n,i+1} = S_{n,i} + \frac{x_{i+1}}{\sqrt{n}}, \quad \begin{cases} S_{n,i} \text{ is independent} \\ \text{of } x_{i+1} \text{ and } Y_{i+1}. \end{cases}$$

Using Taylor expansion, $\varphi(Z_{n,i}) = \varphi(S_{n,i}) + \varphi'(S_{n,i}) \frac{Y_{i+1}}{\sqrt{n}} + \frac{1}{2} \varphi''(S_{n,i}) \frac{Y_{i+1}^2}{n} + O\left(\frac{1}{n^{3/2}} \|\varphi''\|_{\infty}\right)$,

$$\varphi(Z_{n,i+1}) = \varphi(S_{n,i}) + \varphi'(S_{n,i}) \frac{x_{i+1}}{\sqrt{n}} + \frac{1}{2} \varphi''(S_{n,i}) \frac{x_{i+1}^2}{n} + O\left(\frac{1}{n^{3/2}} \|\varphi''\|_{\infty}\right).$$

Since the moments of x_{i+1} and Y_{i+1} match to second order. We get

$$\mathbb{E} \varphi(Z_{n,i+1}) - \mathbb{E} \varphi(Z_{n,i}) = O\left(\|\varphi''\|_{\infty} \frac{\mathbb{E} |Y_{i+1}|^3 + \mathbb{E} |x_{i+1}|^3}{n^{3/2}}\right) = O\left(\mathbb{E} |x|^3 \frac{\|\varphi''\|_{\infty}}{n^{3/2}}\right). \quad \square$$

II. Green's function comparison theorem (four-moment theorem)

In Theorem 7.1, (i) follows from the local semicircle law. The condition (ii) follows from the next theorem, which, roughly speaking, shows that two Wigner ensembles have almost identical local statistics provided that their matrix elements have matching "first four moments".

Theorem 7.3 Suppose we have two Wigner matrices H^V and H^W , with matrix elements

$v_{ij}^V = \frac{1}{\sqrt{n}} v_{ij}$, $w_{ij}^W = \frac{1}{\sqrt{n}} w_{ij}$. Suppose v_{ij} and w_{ij} have finite moments up to arbitrary order, i.e., $\forall k \in \mathbb{N}$,

$$\mathbb{E} |v_{ij}|^k + \mathbb{E} |w_{ij}|^k \leq c_k \text{ for a constant } c_k > 0.$$

Suppose v_{ij} and w_{ij} have matching first three moments, i.e.,

$$\mathbb{E} \bar{v}_{ij}^s v_{ij}^{k-s} = \mathbb{E} \bar{w}_{ij}^s w_{ij}^{k-s}, \quad k=1, 2, 3, \quad s=0, 1, \dots, k,$$

and almost identical fourth moments:

$$|\mathbb{E} \bar{v}_{ij}^s v_{ij}^{4-s} - \mathbb{E} \bar{w}_{ij}^s w_{ij}^{4-s}| \leq N^{-\delta}, \quad s=0, 1, 2, 3, 4,$$

for some constant $\delta > 0$. Suppose the local semicircle laws hold for H^V and H^W . (71)