Universality of Random Matrices and Dyson Brownian Motion

Horng-Tzer Yau

Harvard University

Joint work with L. Erdős , and B. Schlein, J. Yin

Lecture 1: History of Universality and Main Results

Basic question: consider a large matrix whose elements are random variables with a given probability law. What can be said about the statistical properties of the eigenvalues of the matrix?

Gaussian unitary ensemble: $H = (h_{jk})_{1 \le j,k \le N}$ hermitian with

$$h_{jk} = \frac{1}{\sqrt{N}}(x_{jk} + iy_{jk}) \qquad \text{and} \qquad h_{jj} = \frac{2}{\sqrt{N}} x_{jj}$$

and where x_{jk}, $y_{jk}, j > k$, and x_{jj} are independent centered Gaussian random variables with variance $1/2$.

Classical ensembles: Gaussian unitary ensemble (GUE), Gaussian orthogonal ensemble (GOE), Gaussian symplectic ensemble(GSE), sample covariance ensembles

Probability density of eigenvalues (w.r.t. Lebesgue measure)

$$p_N(x_1, ..., x_N)$$

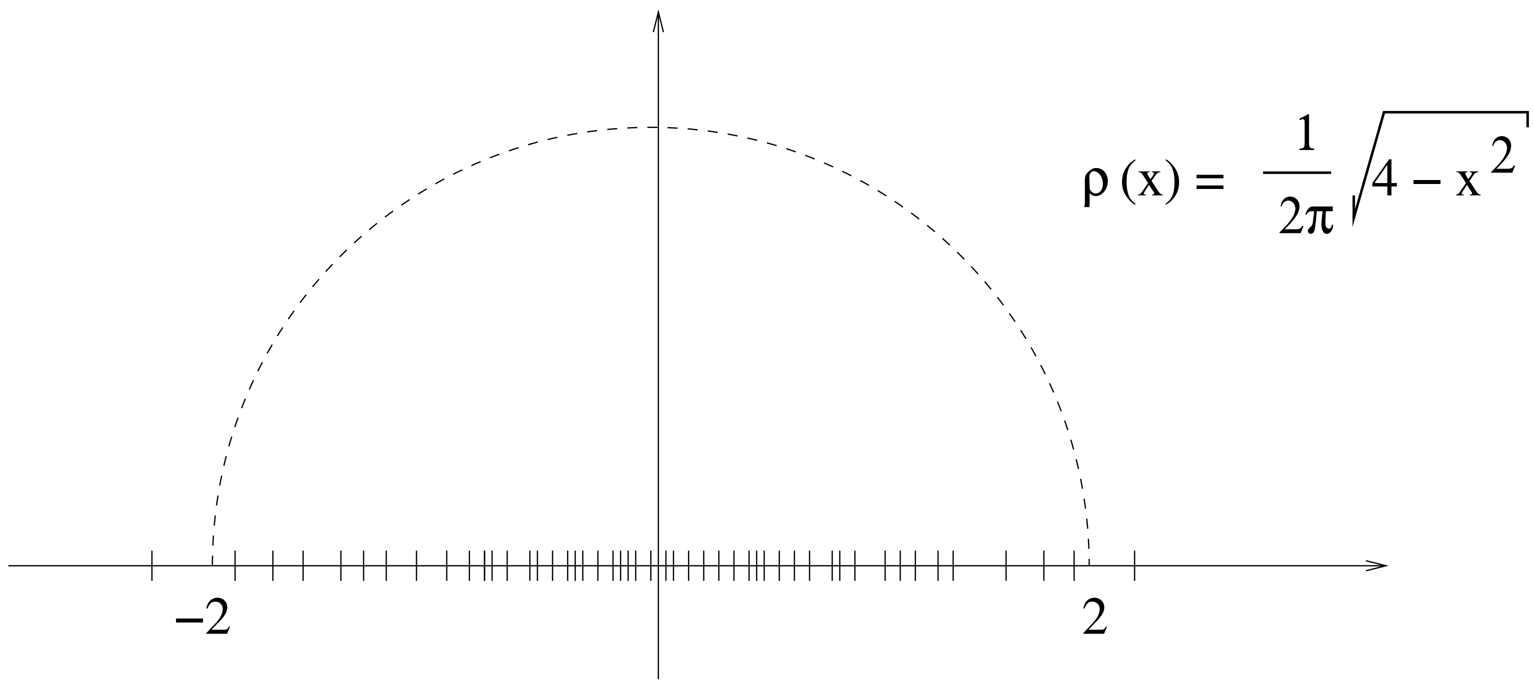
Correlation function for two eigenvalues:

$$p_N^{(2)}(x_1, x_2) = \int_{\mathbb{R}^{N-2}} p_N(x_1, x_2, ..., x_N) dx_3 ... dx_N$$

Density of states:

$$\rho_N(x) = \int_{\mathbb{R}^{N-1}} p_N(x, x_2, ..., x_N) dx_2 ... dx_N$$

.



Eigenvalues: $\lambda_1 \leq \lambda_2 \leq \ldots\ldots \lambda_N$, $\lambda_{i+1} - \lambda_i \sim 1/N$.

Global statistics: Density of state $\rho(x)$ follows the Wigner semi-circle law for GUE, GOE and GSE

Sample covariance ensembles, matrix of the form A^+A: Marchenko-Pastur law

Gaudin, Mehta, Wigner, Dyson: level correlation for local statistics

$$\lim_{N\to\infty} p_N^{(2)}\left(E+\frac{a_1}{N}, E+\frac{a_2}{N}\right)$$

$$= \det\left\{S(a_i-a_j)\right\}_{i,j=1}^{2}, \quad S(a)=\frac{\sin \pi a}{\pi a} \text{ (for GUE)}, \ |E|<2$$

Gap distribution

$$\mathbb{E}\,\frac{\#\text{ of e.v. pairs with gap less than } s \text{ near } E}{\#\text{ of eigenvalues near } E}$$

$$\to \int_0^s \frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\det(1-\mathcal{K}_\alpha)d\alpha$$

$$\mathcal{K}_\alpha(x,y)=\frac{\sin \pi(x-y)}{\pi(x-y)} \text{ as operator on } L^2((0,\alpha))$$

Edge Universality

λ_N: the largest eigenvalue of a random matrix.

The probability distribution functions of λ_N for the classical Gaussian ensembles are identified by Tracy-Widom to be

$$\lim_{N\to\infty} \mathbb{P}(N^{2/3}(\lambda_N - 2) \le t) = F_\beta(t)$$

where F_β can be computed in terms of Painleve equations and $\beta = 1, 2, 4$ corresponds to the standard classical ensembles.

Edge universality is less universal than the bulk. TW distribution appears in random interfaces, random corners and random processes.

History, Applications and Conjectures

• **Nuclear Physics:** the excitation spectra of heavy nuclei are expected to have the same local statistical properties as the eigenvalues of GOE (Wigner, 1955).

• **Quantum Chaos Conjecture** (Bohigas-Giannoni-Schmit 1984, Berry-Tabor, 1977)

• **Riemann ζ-function:** Gap distribution of zeros of ζ function is given by GUE (Montgomery, 1973).

• **Anderson Model (1958):** V_ω random potential.

$$\text{random Schrödinger operator:} \quad H_\Lambda = -\Delta + \lambda V_\omega$$

the eigenvalue distributions of H in the **delocalization regime** (λ small) are given by GOE.

Frohlich-Spencer, Disertori-Spencer-Zirnbauer, Supersymmetry etc

KNOWN RESULTS FOR UNITARY ENSEMBLES

Unitary ensemble: Hermitian matrices with density

$$\mathcal{P}(H)\mathrm{d}H \sim e^{-\mathrm{Tr}\, NV(H)}\mathrm{d}H$$

Invariant under $H \to UHU^{-1}$ for any unitary U (GUE)
Joint density function of the eigenvalues is explicitly known

$$p(x_1, \ldots, x_N) = \text{const.} \prod_{i<j} (x_i - x_j)^{\beta} e^{-\sum_j NV(x_j)}$$

classical ensembles $\beta = 1, 2, 4$ GOE, GUE, GSE.

large N asymptotic of orthogonal polynomials

Mehta-Gaudin (1960-), Dyson (1962-76), via Hermite polynomials and general cases by Deift-Its-Zhou (1997), Pastur-Schcherbina (1997), Bleher-Its (1999), Deift-et al (1999-2009), Lubinsky (2008) . . .

Universality of Generalized Wigner Ensembles

$$H = (h_{kj})_{1\le k,j\le N}, \quad h_{ji} = h_{ij} \quad \text{independent}$$

$$\mathbb{E}h_{ij} = 0,\ \mathbb{E}|h_{ij}|^2 = \sigma_{ij}^2, \qquad \sum_i \sigma_{ij}^2 = 1,\ \sigma_{ij} \sim N^{-1/2}.$$

$$\delta \le C_{inf} \equiv \inf_{N,i,j}\{N\sigma_{ij}^2\} \le \sup_{N,i,j}\{N\sigma_{ij}^2\} \equiv C_{sup} \le \delta^{-1} \qquad \text{(A)}$$

$$\text{sub-exponential decay:} \quad \mathbb{E}e^{|\sqrt{N}h_{ij}|^\varepsilon} < \infty \qquad \text{(B)}$$

If h_{ij} are i.i.d. then it is called Wigner ensembles.

Wigner ensembles are completely different from the invariant ensembles due to the lack of explicit formula for the eigenvalue distribution. Only Gaussian are Wigner and invariant.

(A) rules out the band and sparse matrices, but some estimates of the proof are valid for these two ensembles.

Theorem [Erdoes-Schlein-Y-Yin, 2009-2010] The bulk universality holds for generalized Wigner ensembles satisfying (A) and (B), i.e., for $-2 < E < 2$

$$\lim_{N\to\infty} \left(p^{(k)}_{F,N} - p^{(k)}_{\mu,N}\right)\left(E + \frac{b_1}{N}, \ldots, E + \frac{b_k}{N}\right) \rightharpoonup 0$$

F	μ
generalized symmetric matrices	GOE
generalized hermitian	GUE
generalized self-dual quaternion	GSE
real covariance	real Gaussian Wishart
complex covariance	complex Gaussian Wishart

Edge universality: Two generalized Wigner ensembles have the same edge distributions if the first two moments of the two ensembles are the same.

Earlier results for edge universality: Soshnikov, Sodin (moment methods and generalization). Tao-Vu: Wigner ensembles with vanishing third moments.

Three Steps to the Universality of Random Matrices

Step 1. Local Semicircle Law, LSC (valid for more general class of ensembles)

Step 2. Universality of Gaussian divisible ensembles

$$H = H_0 + \sqrt{t}V, \qquad t > 0, \quad H_0 \text{ is Wigner } V \text{ is GUE}$$

i.e., the matrix entries have a substantial Gaussian component.

2a. Asymptotic analysis of explicit formulas (Johansson, Brezin-Hikami) for correlation functions for eigenvalues: valid only for Hermitian matrices

2b. Local ergodicity of Dyson Brownian motion

Gaussian convolution matrix

$$H = e^{-t/2}H_0 + (1 - e^{-t})^{1/2}V \sim H_0 + \sqrt{t}V$$

Step 3. Approximation by Gaussian divisible ensembles
(A density argument via perturbation expansion)
3a. Reverse heat flow (Need smoothness)
3b. Green function comparison theorem

Approach of this lecture: LSC + DBM + Green fn comparison on the bulk for generalized Wigner ensembles, all symmetry classes.

LSC + Green fn comparison on the edges (Surprising fact: Green function methods are good both on the bulk and on the edges!)

Tao-Vu's result: four moment theorem.

(1): LSC+ Johansson+ 4-moments $\Longrightarrow$ universality for Wigner Hermitian ensembles for distribution whose support contains at least three points (Bernoulli excluded).

(2) LSC+ 4-moments $\Longrightarrow$ universality for Wigner ensembles with the first four moments matched (three moments on the edges). No Johansson result for symmetric or quaternion case.

Outline of lectures

1. Statement of uniform local semicircle law and applications to rigidity of eigenvalues, eigenfunction delocalization estimates. Idea of proof of uniform local semicircle law.

2. Local ergodicity of DBM

3. Green function comparison theorem and universality.

4. Dyson's conjecture on local relaxation time of DBM.

Pseudo-equilibrium measures, local relaxation flow, logarithmic Sobolev inequality, entropy estimates.

Step 1: Strong Local semicircle law

Theorem [Erdös-Schlein-Y-Yin, 2010] Suppose assumptions (A) and (B) hold. Then for any $z = E + i\eta$ with $|E| \leq 5$, $N^{-1} \ll \eta \leq 10$ we have

$$\mathbb{P}\Big(\max_i |G_{ii}(z) - m_{sc}(z)| \gg \frac{1}{\sqrt{N\eta}} \Big) \ll N^{-\infty}$$

$$\mathbb{P}\Big(\max_{ij} |G_{ij}(z)| \gg \frac{1}{\sqrt{N\eta}} \Big) \ll N^{-\infty}$$

$$\mathbb{P}\Big(|m(z) - m_{sc}(z)| \gg \frac{1}{N\eta} \Big) \ll N^{-\infty}$$

where $G_{ij} = \frac{1}{H-z}(i,j)$ and $m(z) = \frac{1}{N}\mathrm{Tr}G = \frac{1}{N}\sum_i G_{ii}$.

$$\text{Stieltjes transform:} \quad m(z) = \int \frac{\varrho(x)\mathrm{d}x}{x - z}$$

These estimates are optimal and are first results on matrix elements of Green functions.

Corollary [Rigidity of Eigenvalues] Let γ_j be the classical location of the j-th eigenvalue, i.e.

$$\int_{-\infty}^{\gamma_j} \varrho_{sc}(x)\mathrm{d}x = \frac{j}{N}$$

Then

$$\mathbb{P}\Big(\exists\, j:\ |\lambda_j - \gamma_j| \geq (\log N)^L \Big[\min\big(\,j, N-j+1\,\big)\Big]^{-1/3} N^{-2/3}\Big) \ll N^{-\infty}$$

All estimates are optimal (up to log corrections)

Tao-Vu: For Wigner matrices with vanishing third moment:

$$\Big(\mathbb{E}\Big[|\lambda_j - \gamma_j|^2\Big]\Big)^{1/2} \leq \Big[\min\big(\,j, N-j+1\,\big)\Big]^{-1/3} N^{-1/6-\varepsilon_0}$$

with some small positive ε_0

From Stieltjes transform to counting function

Helffer-Sjöstrand formula: Let $f \in C^1(\mathbb{R})$. Let $\chi(y)$ be a cutoff function in $[-1, 1]$. Define the quasianalytic extension of f as

$$\widetilde{f}(x+iy) = (f(x) + iyf'(x))\chi(y),$$

then

$$f(\lambda) = \frac{1}{2\pi}\int_{\mathbb{R}^2} \frac{\partial_{\bar{z}}\widetilde{f}(x+iy)}{\lambda - x - iy}\mathrm{d}x\mathrm{d}y$$

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

Detour: Similar results without Assumption (A)

Theorem [Erdös-Yau-Yin, 2010] Let $\mathbb{E}h_{ij} = 0$, $\mathbb{E}|h_{ij}|^2 = \sigma_{ij}^2$ and Assumption (B) holds. Define $M := (\max_{ij} \sigma_{ij}^2)^{-1}$. Then for any $z = E + i\eta$ with $M\eta \gg 1$ and $\Big||E| - 2\Big| \geq \kappa_0 > 0$:

$$\mathbb{P}\Big(\max_i |G_{ii}(z) - m_{sc}(z)| \gg \frac{1}{\sqrt{M\eta}}\Big) \ll N^{-\infty}$$

$$\mathbb{P}\Big(\max_{ij} |G_{ij}(z)| \gg \frac{1}{\sqrt{M\eta}}\Big) \ll N^{-\infty}$$

$$\forall \varepsilon \qquad \mathbb{P}\Big(|m(z) - m_{sc}(z)| \gg \frac{N^\varepsilon}{M\eta}\Big) \ll N^{-\infty}$$

Eigenfn. delocalization estimate Denote the eigenfunction by $\mathbf{v}_\alpha = (v_\alpha(1), \ldots, v_\alpha(N))$. Then $\sup_j |v_\alpha(j)|^2 \leq C \log N/M$ with very high probability.

Disertori–Pinson-Spencer: Density of states for Gaussian band matrices. Related work on supersymmetric models

Local semicircle law

$H^{(i)}$ denotes the $(N-1) \times (N-1)$ minor of H after removing the i-th column/row and similarly for $H^{(ij)}$ etc. Let $\mathbf{a}^i$ be the i-th column. The Green functions are:

$$G = \frac{1}{H - z}, \qquad G^{(i)} = \frac{1}{H^{(i)} - z}$$

Then by inverting 2×2 matrix we have

$$G_{ii} = \left(\frac{1}{H - z}\right)_{ii} = \frac{1}{h_{ii} - z - \mathbf{a}^i \cdot G^{(i)} \mathbf{a}^i}$$

$$H = \begin{pmatrix} h & \mathbf{a}^* \\ \mathbf{a} & B \end{pmatrix}, \qquad h \in \mathbb{C}, \;\; \mathbf{a} \in \mathbb{C}^{N-1}, \;\; B \in \mathbb{C}^{(N-1)\times(N-1)}$$

$$(H - z)^{-1}(1,1) = \frac{1}{h - z - \mathbf{a} \cdot (B - z)^{-1} \mathbf{a}}$$

Crudest local semicircle law (for $\sigma_{ij}^2 = \frac{1}{N}$)

$$G_{ii} = \frac{1}{h_{ii} - z - \mathbb{E}_i\, \mathbf{a}^i \cdot G^{(i)} \mathbf{a}^i - Z_i}$$

$$Z_i := \mathbf{a}^i \cdot G^{(i)} \mathbf{a}^i - \mathbb{E}_i\, \mathbf{a}^i \cdot G^{(i)} \mathbf{a}^i$$

simple computation: $\quad \mathbb{E}_i\, \mathbf{a}^i \cdot G^{(i)} \mathbf{a}^i = \dfrac{N-1}{N} m^{(i)}$

By the interlacing property of the eigenvalues of H and $H^{(i)}$,

$$|m(z) - m^{(i)}(z)| \le \frac{1}{N\eta}, \qquad \eta = \operatorname{Im} z$$

$$\Longrightarrow G_{ii} = \frac{1}{-z - m(z) + \Omega_i}, \qquad \Omega_i = h_{ii} - Z_i + O\Big(\frac{1}{N\eta}\Big)$$

$$m(z) = \frac{1}{-z - m(z)} + \frac{1}{N} \sum_i \frac{1}{(z + m(z))^2} \Omega_i + \cdots$$

The main error term is in terms of $\sum_i Z_i$!

$$G_{ii} = \frac{1}{-z - m(z) + \Omega_i}, \qquad \Omega_i = h_{ii} - Z_i + O\Big(\frac{1}{N\eta}\Big)$$

Expand and sum up, with $\Omega = \max_i \Omega_i \ll 1$

$$m(z) = -\frac{1}{z + m(z)} + O(\Omega)$$

Large deviation (CLT) estimate: $\quad |Z_i| \lesssim \frac{\mathrm{Im}\, m^{(i)}}{\sqrt{N\eta}} \sim \frac{\mathrm{Im}\, m}{\sqrt{N\eta}}$

$$m(z) = -\frac{1}{z + m(z)} + O\Big(\frac{\mathrm{Im}\, m}{\sqrt{N\eta}}\Big)$$

By the stability analysis of this equation, and $m_{sc} = -\frac{1}{z+m_{sc}}$,

$$|m - m_{sc}| \le C\frac{1}{\sqrt{N\eta}}, \qquad \Big||E| - 2\Big| > 0$$

Formulas relating Green functions:

$$G_{ij} = -G_{jj} G_{ii}^{(j)} K_{ij}, \quad K_{ij} = h_{ij} - z\delta_{ij} - \mathbf{a}^i \cdot G^{(ij)} \mathbf{a}^j$$

$$G_{ii} = G_{ii}^{(j)} + \frac{G_{ij}G_{ji}}{G_{jj}}, \quad G_{ij} = G_{ij}^{(k)} + \frac{G_{ik}G_{kj}}{G_{kk}} \quad i, j, k, \ \text{different}$$

Rule of thumb: $G_{ii} \sim O(1)$, G_{ij} is small.

Lemma For any even number p and away from the edges, we have

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} Z_i\right|^p \leq \frac{C_p}{(N\eta)^p}$$

for sufficiently large N, i.e., with high probability:

$$\frac{1}{N}\sum_{i=1}^{N} Z_i \leq (N\eta)^{-1}$$

Step 2: Local Ergodicity of Dyson Brownian Motion

Gaussian convolution matrix $H = e^{-t/2}H_0 + (1-e^{-t})^{1/2}V$ where V is a GUE matrix can be obtained by evolving matrix elements

by an OU process: $\partial_t u_t(x) = L u_t(x), \qquad L = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{x}{2}\frac{\mathrm{d}}{\mathrm{d}x}$

Denote the probability density of the eigenvalues by $f_t(\mathbf{x})\mu(\mathrm{d}\mathbf{x})$.

GUE : $\mu = Ce^{-N\mathcal{H}(\mathbf{x})}\mathrm{d}\mathbf{x}, \quad \mathcal{H}(\mathbf{x}) = \sum_{i=1}^{N} x_i^2/2 - N^{-1}\sum_{i\neq j}\log|x_j - x_i|$

$\partial_t f_t = \mathscr{L} f_t, \; \int f\mathscr{L}f d\mu = \frac{1}{2N}\int |\nabla f|^2 d\mu,$ gradient flow of GUE

$$\mathscr{L} = \frac{1}{2N}\sum_{i=1}^{N}\partial_i^2 + \sum_{i=1}^{N}\left(-\frac{1}{2}x_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{x_i - x_j}\right)\partial_i,$$

BM of matrix elements $\Longrightarrow$ DBM on eigenvalues.
Johansson (essentially): local equilibrium is reached $t = O(1)$.

$$(SDE) \qquad \mathrm{d}x_i = \frac{1}{\sqrt{N}}\mathrm{d}B_i + \Big(-\frac{1}{2}x_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{x_i - x_j}\Big)\mathrm{d}t$$

Derivation of DBM (for BM case, $h_{ij} = N^{-1/2}B_{ij}$ for simplicity):

$$\frac{\partial \lambda_\alpha}{\partial h_{ik}} = \bar{u}_\alpha(i)u_\alpha(k), \qquad \frac{\partial u_\alpha(i)}{\partial h_{\ell k}} = \sum_{\beta\neq\alpha}\frac{1}{\lambda_\alpha - \lambda_\beta}\bar{u}_\beta(\ell)u_\alpha(j)u_\beta(i)$$

$$\frac{\partial^2 \lambda_\alpha}{\partial h_{ik}\partial h_{\ell j}} = \sum_{\beta\neq\alpha}\frac{1}{\lambda_\alpha - \lambda_\beta}u_\beta(\ell)\bar{u}_\alpha(j)\bar{u}_\beta(i)u_\alpha(k) + c.c.$$

$$\mathrm{d}\lambda_\alpha = \frac{1}{\sqrt{N}}\sum_{ik}\bar{u}_\alpha(i)u_\alpha(k)\mathrm{d}B_{ik} + \frac{1}{N}\sum_{ik}\sum_{\beta\neq\alpha}\frac{1}{\lambda_\alpha - \lambda_\beta}u_\beta(i)\bar{u}_\alpha(k)\bar{u}_\beta(i)u_\alpha(k)\mathrm{d}t$$

$$= \frac{1}{\sqrt{N}}\mathrm{d}B_\alpha + \frac{1}{N}\sum_{\beta\neq\alpha}\frac{1}{\lambda_\alpha - \lambda_\beta}\mathrm{d}t$$

where one can check that $\sum_{ik}\bar{u}_\alpha(i)u_\alpha(k)\mathrm{d}B_{ik}$ are independent white noises for different α's.

Assumption III. There exists an $\varepsilon > 0$ such that

$$\sup_{t \geq N^{-2\mathfrak{a}}} \int \frac{1}{N} \sum_{j=1}^{N} (x_j - \gamma_j)^2 f_t(\mathrm{d}\mathbf{x}) \mu(\mathrm{d}\mathbf{x}) \leq C N^{-1-2\mathfrak{a}}$$

with a constant C uniformly in N.

The typical spacing between neighboring points is of order $1/N$ away from the spectral edges. Assumption III guarantees that typically the random points x_j remain in the $N^{-1/2-\mathfrak{a}}$ vicinity of their classical location. (The notation of eigenvalues are changed to x_j).

Theorem [Erdös-Schlein-Y-Yin, 2009] Suppose that Assumption III hold for the solution of $\partial f_t = \mathscr{L} f_t$. Let $|E| < 2$ and $b > 0$. Then $\forall \delta > 0$, $\forall k \geq 2$ we have

$$\lim_{b \to 0} \lim_{N \to \infty} \sup_{t \geq N^{-2\mathfrak{a}+\delta}} \int_{E-b}^{E+b} \frac{\mathrm{d}E'}{2b} \times \left(p_{t,N}^{(k)} - p_{\mu,N}^{(k)} \right) \left(E' + \frac{\alpha_1}{N\varrho(E)}, \ldots, E' + \frac{\alpha_k}{N\varrho(E)} \right) \rightharpoonup 0$$

where $\mathfrak{a} > 0$ is the exponent from Assumption III.

Explicit error estimates were obtained.

Strong local semicircle law $\Longrightarrow \mathfrak{a} = 1/2 - \varepsilon \Longrightarrow$ local relaxation time to DBM is $N^{-1+\varepsilon}$ (Dyson's conjecture).

$t \ll N^{-1}$ then the eigenvalue does not move on the scale of N^{-1}.

Once Step 2 is completed, the matrices fall into the universality class are in some sense dense. Thus we need a density argument.

Step 3 Approximation by Gaussian divisible ensembles:

Reverse heat flow: Given a smooth distribution $u(x)\mathrm{d}x$, choose

$$g_t = \left(1 - t\Delta + \frac{(-t\Delta)^2}{2!} + \dots\right)u \approx e^{-t\Delta}u$$

then

$$e^{t\Delta}g_t$$

will have a Gaussian component and will be very close to u since $t \sim N^{-\varepsilon}$ is small.

3c [Erdös-Y-Yin '10] (Green Function Comparison theorem) Suppose that the first four moments of v_{ij} and w_{ij} for two random matrices almost match. Let $q_j = E_j + i\eta, j = 1, 2$ with $|E_j| < 2$ and $\eta \geq N^{-1-\varepsilon}$. Then we have

$$\left(\mathbb{E}_{\mathbf{v}} - \mathbb{E}_{\mathbf{w}}\right)\left[\left(\frac{1}{q_1 - H}\right)_{ij}\left(\frac{1}{q_2 - H}\right)_{k\ell}\right] \to 0$$

$$\Longrightarrow \left(p^{(2)}_{\mathbf{v},N} - p^{(2)}_{\mathbf{w},N}\right)\left(E + \frac{\alpha_1}{N}, E + \frac{\alpha_2}{N}\right) \rightharpoonup 0.$$

Bulk universality: For any matrix ensemble H, find another matrix ensembles H_0 such that the first four moments of H and $H_0 + \sqrt{t}V$ are the same.

Proof

$$\mathbb{E}\,\frac{1}{N}\mathrm{Tr}\frac{1}{H^{(\mathbf{v})}-z}-\mathbb{E}\,\frac{1}{N}\mathrm{Tr}\frac{1}{H^{(\mathbf{w})}-z}$$
$$=\sum_{\gamma=1}^{N^2}\left[\mathbb{E}\,\frac{1}{N}\mathrm{Tr}\frac{1}{H_\gamma-z}-\mathbb{E}\,\frac{1}{N}\mathrm{Tr}\frac{1}{H_{\gamma-1}-z}\right].$$

$H_{\gamma-1}$ with H_γ differ only in the (i,j) and (j,i) matrix elements:

$$H_{\gamma-1}=Q+\frac{1}{\sqrt{N}}V, \qquad V:=v_{ij}|\mathbf{e}_i\rangle\langle\mathbf{e}_j|+v_{ji}|\mathbf{e}_j\rangle\langle\mathbf{e}_i|$$

$$H_\gamma=Q+\frac{1}{\sqrt{N}}W, \qquad W:=w_{ij}|\mathbf{e}_i\rangle\langle\mathbf{e}_j|+w_{ji}|\mathbf{e}_j\rangle\langle\mathbf{e}_i|,$$

$$Q_{ij}=Q_{ji}=0$$

resolvent identity:

$$\frac{1}{z-Q-A} = \frac{1}{z-Q} + \frac{1}{z-Q}A\frac{1}{z-Q} + \frac{1}{z-Q}A\frac{1}{z-Q}A\frac{1}{z-Q} + \ldots$$

Power counting: $A \sim h_{ij} \sim N^{-1/2}$. So four moments matching implies a factor of $N^{-2-\delta}$ provided that all green functions in the resolvent expansions are finite. This is guaranteed by the LSC. Without LSC up to $\eta \sim N^{-1}$, this expansion could be badly divergent.

Edge universality: Power counting changes due to the density on the edge is $N^{-2/3}$.

$3b$ Four-Moment Theorem [Tao-Vu 2009]: If the first four moments of the single entry distribution of two Wigner matrices match, then the local e.v. statistics coincide.

Inputs:

1. Local semicircle law (ESY, no smoothness asumption)

2. eigenfunction estimates (ESY, no smoothness assumption)

3. Level repulsion (ESY proof needs smoothness, TV's estimate is much weaker but sufficient for their purpose and they removed the smoothness assumption.)

Main Difficulty: Overlaps of singularities in the perturbation of eigenvalues and eigenfunctions.

logarithmic Sobolev inequality and convexity

Recall entropy of a probability density f w.r.t a measure μ:

$$S(f\mu|\mu) = S(f) = \int f \log f \mathrm{d}\mu, \quad \mu = e^{-\mathcal{H}}/Z$$

Suppose $\partial_t f_t = \mathscr{L} f_t$ and μ is invariant, i.e., $\int \mathscr{L} f_t \mathrm{d}\mu = 0$.

$$\begin{aligned}\partial_t S(f_t) &= \int (\mathscr{L} f_t) \log f_t \mathrm{d}\mu + \int f_t \frac{\mathscr{L} f_t}{f_t} \mathrm{d}\mu \\ &= -\frac{1}{2} \int \frac{(\nabla f_t)^2}{f_t} \mathrm{d}\mu = -4D(h_t), \quad h_t := \sqrt{f_t}\end{aligned}$$

$$\partial_t D(\sqrt{f_t}) = -\frac{1}{2} \int (\nabla h)(\nabla^2 \mathcal{H}) \nabla h - \frac{1}{2} \int \sum_{ij} \Big(\partial_{ij} h - \frac{(\partial_i h)(\partial_j h)}{h} \Big)^2,$$

Bakry-Emery: $\nabla^2 \mathcal{H} \geq \tau^{-1} \Longrightarrow \partial_t D_\mu(\sqrt{f_t}) \leq -\tau^{-1} D_\mu(\sqrt{f_t})$

$\Longrightarrow$ LSI: $S(f\mu|\mu) \leq \tau D_\mu(\sqrt{f}), \quad S(f_t\mu|\mu) \leq e^{-Ct/\tau} S(f_0\mu|\mu).$

For DBM: $\nabla^2 \mathcal{H} \geq 1$(from x_j^2) relaxation time is O(1)

Global equilibrium is reached in time scale of O(1) .

Local equilibrium is believed to be reached in $O(N^{-1})$.

F. Dyson on approach to equilibrium of DBM (1962):

The picture of the gas coming into equilibrium in two well-separated stages, with microscopic and macroscopic time scales, is suggested with the help of physical intuition. A rigorous proof that this picture is accurate would require a much deeper mathematical analysis.

Now we compute $\partial_t D(\sqrt{f_t})$. Let $h := \sqrt{f}$ for simplicity, then

$$\partial_t h_t = \frac{1}{2h_t}\partial h_t^2 = \frac{1}{2h_t}\mathscr{L} h_t^2 = \mathscr{L} h_t + \frac{1}{2h_t}(\nabla h_t)^2.$$

In the last step we used that $\mathscr{L} h^2 = (\nabla h)^2 + 2h\mathscr{L} h$ that can be seen directly from $\mathscr{L} = \frac{1}{2}\Delta - \frac{1}{2}(\nabla \mathcal{H})\nabla$.

We compute (dropping the t subscript for brevity and $\int = \int \mathrm{d}\mu$)

$$\partial_t D(\sqrt{f_t}) = \frac{1}{2}\partial_t \int (\nabla h)^2 \mathrm{d}\mu$$

$$= \int (\nabla h)(\nabla \mathscr{L} h) + \frac{1}{2}\int (\nabla h)\cdot \nabla \frac{(\nabla h)^2}{h}$$

$$= \int (\nabla h)[\nabla, \mathscr{L}]h + \int (\nabla h)\mathscr{L}(\nabla h) + \frac{1}{2}\int \sum_{ij} \partial_i h \Big[\frac{2(\partial_j h)\partial_i\partial_j h}{h} - \frac{(\partial_j h)^2 \partial_i h}{h^2}\Big]$$

$$= -\frac{1}{2}\int (\nabla h)(\nabla^2 \mathcal{H})\nabla h - \frac{1}{2}\int \sum_{ij} (\partial_i\partial_j h)^2$$

$$+ \frac{1}{2}\int \sum_{ij} \Big[\frac{2(\partial_j h)(\partial_i h)\partial_{ij} h}{h} - \frac{(\partial_j h)^2(\partial_i h)^2}{h^2}\Big]$$

$$= -\frac{1}{2}\int (\nabla h)(\nabla^2 \mathcal{H})\nabla h - \frac{1}{2}\int \sum_{ij} \Big(\partial_{ij} h - \frac{(\partial_i h)(\partial_j h)}{h}\Big)^2,$$

where we used the commutator

$$[\nabla, \mathscr{L}] = -\frac{1}{2}(\nabla^2 \mathcal{H})\nabla.$$

Approach to equilibrium: $\nabla^2 \mathcal{H} \geq \tau^{-1}$

$$\partial_t S(f\mu|\mu) = -D_\mu(\sqrt{f_t}), \qquad \partial_t D_\mu(\sqrt{f_t}) \leq -\tau^{-1} D_\mu(\sqrt{f_t})$$

$$\Longrightarrow \quad S(f_t\mu|\mu) \leq e^{-Ct/\tau} S(f_0\mu|\mu).$$

$\tau \sim 1$ for DBM and there is no room to improve.

Key Idea: Pseudo equilibrium measure $\omega = \psi\mu$ is almost an equilibrium. Then

$$\partial_t S(f_t\mu|\omega) = \partial_t S(g_t\omega|\omega) = -D_\omega(\sqrt{g_t}) + \Omega_t, \quad g_t = f_t/\psi$$

where Ω_t is the error term due to that ω is non-equilibrium.

Choose ω so that

(1) ω is more convex than μ so that the LSI w.r.t. ω improves.

(2) the error Ω_t is still not too large.

(3) ω and μ have almost identical local equilibrium.

This implies that $D_\omega(\sqrt{g_t})$ is small provided Ω_t is small. Thus $f_t \sim \omega$ and we can deduce the local statistics of f_t from ω.

Note: We never compared two dynamics! We estimate the Dirichlet form of f_t w.r.t. the pseudo equilibrium measure ω with f_t given by DBM.

Local relaxation flow

pseudo equilibrium measure, $\omega \sim e^{-N\widetilde{\mathcal{H}}}$, $\widetilde{\mathcal{H}} = \mathcal{H} + \mathbf{W}$,

$$\mathbf{W} = \sum_{j=1}^{N} W_j(x_j), \quad W_j(x_j) = \frac{1}{2\tau}(x_j - \gamma_j)^2,$$

Local relaxation flow:

$$\partial_t u_t = \widetilde{\mathscr{L}} u_t, \quad \text{(density of system at time } t) = u_t\,\omega$$

$$-\int f \widetilde{\mathscr{L}} f \mathrm{d}\omega = \frac{1}{2N} \int |\nabla f|^2 \, \mathrm{d}\omega := D_\omega(f) \quad \text{gradient flow w.r.t. } \omega$$

$$\widetilde{\mathscr{L}} = \mathscr{L} - \sum_j b_j \partial_j, \quad b_j = W_j'(x_j) = \tau^{-1}(x_j - \gamma_j)$$

$\nabla^2 \widetilde{\mathcal{H}} \geq \tau^{-1} \Longrightarrow$ relaxation time of $\widetilde{\mathscr{L}}$ is τ.

Let $\omega = \psi\mu$ and define $f_t = g_t\psi$ so that $f_t\mu = g_t\omega$. We will call ω pseudo equilibrium measure. ψ is time independent.

Theorem: [Relative Entropy w.r.t pseudo equilibrium measure] Let $g_t = f_t/\psi$. Assume that $S_\omega(g_0) \le CN^m$ with some fixed m. Then we have

$$\partial_t S_\omega(g_t) \le -D_\omega(\sqrt{g_t}) + CNQ\tau^{-2} \le -C\tau^{-1}S_\omega(g_t) + CNQ\tau^{-2}$$

Hence we obtain the estimates:

$$S_\omega(g_\tau) \le CN\tau^{-1}Q, \qquad D_\omega(\sqrt{g_\tau}) \le CN\tau^{-2}Q.$$

where

$$Q := \sup_{t\ge 0} \sum_j \int (x_j - \gamma_j)^2 f_t \mathrm{d}\mu.$$

General relative entropy calculation

$$\partial_t \int h f_t d\mu = \int (\mathscr{L} h) f_t d\mu = \int (\mathscr{L} h) g_t d\omega = \int h \mathscr{L}_\omega^* g_t d\omega$$

$$\frac{\partial g_t}{\partial t} = \mathscr{L}_\omega^* g_t,$$

$$\begin{aligned}
\frac{d}{dt} S_\omega(g_t) &= \int (\mathscr{L}_\omega^* g_t)(\log g_t)\, d\omega \\
&= \int (g_t \psi_t)\, \mathscr{L}\, (\log g_t)\, d\mu \\
&= -\frac{1}{2} \int \nabla (g_t \psi_t)\, \nabla (\log g_t)\, d\mu \\
&= -\frac{1}{2} \Big[\int \psi_t (\nabla g_t)^2 / g_t\, d\mu + \int (\nabla g_t)(\nabla \psi_t)\, d\mu \Big]
\end{aligned}$$

$$\partial_t S_\omega(g_t) = -\frac{2}{N}\sum_j \int (\partial_j \sqrt{g_t})^2 \, \mathrm{d}\omega + \sum_j \int b_j \partial_j g_t \, \mathrm{d}\omega, \qquad b_j = \nabla W_j = \frac{x_j - \gamma_j}{\tau}$$

From the Schwarz inequality

$$\partial_t S_\omega(g_t) \le -D_\omega(\sqrt{g_t}) + CN \sum_j \int b_j^2 g_t \, \mathrm{d}\omega \le -D_\omega(\sqrt{g_t}) + CNQ\tau^{-2}$$

Together with LSI, we have

$$\partial_t S_\omega(g_t) \le -C\tau^{-1} S_\omega(g_t) + CNQ\tau^{-2}$$

which, after integrating it from $t = 0$ to τ (actually slightly bigger than τ so that the initial entropy disappears) proves the bound on $S_\omega(g_\tau)$.

The estimate for D follows from integration from τ to 2τ.

Theorem [Relaxation of the local relaxation flow] Fact:

$$\left\langle \mathbf{v}, \nabla^2 \widetilde{\mathcal{H}}(\mathbf{x})\mathbf{v} \right\rangle \geq \frac{1}{\tau}\|\mathbf{v}\|^2 + \frac{1}{N}\sum_{i<j}\frac{(v_i - v_j)^2}{(x_i - x_j)^2}, \qquad \mathbf{v} \in \mathbb{R}^N.$$

Consider the forward equation

$$\partial_t q_t = \widetilde{\mathscr{L}} q_t, \qquad t \geq 0,$$

with initial condition $q_0 = q$ and with reversible measure ω. Then

$$\partial_t D_\omega(\sqrt{q_t}) \leq -\frac{1}{\tau} D_\omega(\sqrt{q_t}) - \frac{1}{2N^2}\int \sum_{i,j=1}^{N} \frac{(\partial_i\sqrt{q_t} - \partial_j\sqrt{q_t})^2}{(x_i - x_j)^2} \mathrm{d}\omega,$$

$$\frac{1}{2N^2}\int_0^\infty \mathrm{d}s \int \sum_{i,j=1}^{N} \frac{(\partial_i\sqrt{q_s} - \partial_j\sqrt{q_s})^2}{(x_i - x_j)^2} \mathrm{d}\omega \leq D_\omega(\sqrt{q})$$

and the logarithmic Sobolev inequality holds

$$S_\omega(q) \leq C\tau D_\omega(\sqrt{q})$$

Thus the time to equilibrium is of order τ:

$$S_\omega(q_t) \leq e^{-Ct/\tau} S_\omega(q_0).$$

Theorem [Gap distribution-Dirichlet form estimate] Let q be a probability density w.r.t. ω and G is bounded smooth function with compact support. Then we have

$$\Big| \int \frac{1}{N} \sum_i G(N(x_i - x_{i+1})) \Big[q - 1 \Big] \mathrm{d}\omega \Big| \le C \Big(\frac{D_\omega(\sqrt{q})\tau}{N} \Big)^{1/2} \qquad (*)$$

where τ is the time to equilibrium for $\widetilde{\mathscr{L}}$.
(The estimate (*) is time independent !)

As a comparison we show what entropy control alone can reach:

Theorem: We have, for any $|E_i| < 2$

$$\Big| \frac{1}{N} \sum_i \int G\Big(N(x_i - E_i)\Big) \Big[q - 1 \Big] \mathrm{d}\omega \Big| \le C \sqrt{\tau S_\omega(q)}$$

Recall

$$\partial f_t = \mathscr{L} f_t, \qquad \mathscr{L} = \frac{1}{2N}\sum_i \Big[\Delta_i - \nabla_i \mathcal{H} \cdot \nabla_i\Big] \tag{1}$$

$$D_\omega(\sqrt{g_\tau}) \le C N \tau^{-2} Q, \quad Q := \sup_{t\ge 0} \sum_j \int (x_j - \gamma_j)^2 f_t \mathrm{d}\mu \le N^{-2\mathfrak{a}} \tag{2}$$

$$\Big| \int \frac{1}{N}\sum_i G(N(x_i - x_{i+1}))\Big[q-1\Big]\mathrm{d}\omega\Big| \le C\Big(\frac{D_\omega(\sqrt{q})\tau}{N}\Big)^{1/2} \tag{3}$$

Choose $q = g_\tau = f_\tau/\psi$. Then

$$\Big| \int \frac{1}{N}\sum_i G(N(x_i - x_{i+1}))\Big[g_\tau-1\Big]\mathrm{d}\omega\Big| \le C\Big(\frac{D_\omega(\sqrt{g_\tau})\tau}{N}\Big)^{1/2}$$

$$\le C\sqrt{\frac{Q}{\tau}} \le \frac{C}{\sqrt{N^{2\mathfrak{a}}\tau}} \to 0, \quad \text{if} \quad N^{2\mathfrak{a}}\tau \to \infty$$

$$\int \frac{1}{N}\sum_i G(N(x_i - x_{i+1})) f_\tau \mathrm{d}\mu = \int \frac{1}{N}\sum_i G(\dots)\mathrm{d}\omega, \qquad \text{indep. of initial data}$$

Proof.

$$\int \frac{1}{N}\sum_i G(N(x_i - x_{i+1}))\Big[q-1\Big]\mathrm{d}\omega = \int \frac{1}{N}\sum_i G(\ldots)\Big[(q-q_t)+(q_t-1)\Big]\mathrm{d}\omega$$

For the second term use entropy,

$$\int \frac{1}{N}\sum_i |G(\ldots)||q_t - 1|\mathrm{d}\omega \le C\sqrt{S_\omega(q_t)} \le C\sqrt{S_\omega(q)}e^{-ct/\tau}$$

For the first term, integrate its derivative

$$\int \frac{1}{N}\sum_i G(\ldots)(q - q_t)\mathrm{d}\omega = \int_0^t \mathrm{d}s \int \frac{1}{N}\sum_i G(\ldots)\partial_s q_s \mathrm{d}\omega$$

$$= \int_0^t \mathrm{d}s \int \frac{1}{N}\sum_j (\partial_j q_s)\ \partial_j \Big[\frac{1}{N}\sum_i G(N(x_i - x_{i+1}))\Big]$$

$$= \int_0^t \mathrm{d}s \int \frac{1}{N}\sum_i (\partial_i q_s - \partial_{i+1} q_s)\ G'(N(x_i - x_{i+1}))$$

$$= \int_0^t \mathrm{d}s \int \frac{1}{N}\sum_i (\partial_i \sqrt{q_s} - \partial_{i+1}\sqrt{q_s})\ \sqrt{q_s}\ G'(N(x_i - x_{i+1}))$$

After a Schwarz

$$\Big|\int \frac{1}{N}\sum_i G(N(x_i - x_{i+1}))\Big[q - q_t\Big]\mathrm{d}\omega\Big|$$

$$= \Big|\int_0^t \mathrm{d}s \int \frac{1}{N}\sum_i (\partial_i\sqrt{q_s} - \partial_{i+1}\sqrt{q_s})\ \sqrt{q_s}\ G'(N(x_i - x_{i+1}))\Big|$$

$$\leq \left[\int_0^t \mathrm{d}s \int \sum_i \frac{[\partial_i\sqrt{q_s} - \partial_{i+1}\sqrt{q_s}]^2}{N^2(x_i - x_{i+1})^2}\mathrm{d}\omega\right]^{1/2}$$

$$\times \left[\int_0^t \mathrm{d}s \int \sum_i G'(N(x_i - x_{i+1}))^2(x_i - x_{i+1})^2\, q_s \mathrm{d}\omega\right]^{1/2}$$

$$\leq \left[\int_0^t \mathrm{d}s \int \sum_{i,j} \frac{[\partial_i\sqrt{q_s} - \partial_j\sqrt{q_s}]^2}{N^2(x_i - x_j)^2}\mathrm{d}\omega\right]^{1/2} \sqrt{\frac{t}{N}}$$

The first term is estimated by $D_\omega(\sqrt{q})$ from the additional term in entropy dissipation term. $\square$

Three Steps to the Universality of Random Matrices

Step 1. Local Semicircle Law

Step 2. Strong local ergodicity of Dyson Brownian motion

Step 3. Approximation via the Green function comparison thm

• Universality comes from the local ergodicity of DBM — model independent and involves no explicit formula.

• Universality holds for all classical ensembles; the variances are allowed to be different (but bounded). The only condition on the distribution is subexponential decay (which can be weakened).

• Step 2 and 3 are now very simple and completely general. Only Step 1 may be model dependent.

Main Open Problems

Universality for band matrices, random Schrödinger, non-classical β-ensembles.

Random band matrices: H is symmetric with independent but not identically distributed entries with mean zero and variance

$$\mathbb{E}\, W|h_{k\ell}|^2 = e^{-|k-\ell|/W}$$

Narrow band, $W \ll \sqrt{N} \Longrightarrow$ localization, Poisson statistics

Broad band, $W \gg \sqrt{N} \Longrightarrow$ delocalization, GOE statistics

Even the Gaussian case is open.

Lemma [Large Deviation Estimates] Let a_i be $(1 \le i \le N)$ independent random complex variables with mean zero, variance σ^2 and are subexponential decay. Let A_i, $B_{ij} \in \mathbb{C}$. Then

$$\mathbb{P}\left\{\left|\sum_{i=1}^{N} \overline{a}_i B_{ii} a_i - \sum_{i=1}^{N} \sigma^2 B_{ii}\right| \gg \sigma^2 \Big(\sum_{i=1}^{N} |B_{ii}|^2\Big)^{1/2}\right\} \le C N^{-\infty},$$

$$\mathbb{P}\left\{\left|\sum_{i \ne j} \overline{a}_i B_{ij} a_j\right| \gg \sigma^2 \Big(\sum_{i \ne j} |B_{ij}|^2\Big)^{1/2}\right\} \le C N^{-\infty},$$

From the Lemma: $\quad EZ_i^2 \le \dfrac{1}{N^2} \displaystyle\sum_{k,j \ne i} |G_{kj}^{(i)}|^2, \quad G^{(i)} = \dfrac{1}{H^{(i)} - z}$

identity: $\quad |G|^2 = \dfrac{1}{\eta} \mathrm{I}m G, \quad |G|^2 = GG^*$

$$\frac{1}{N^2} \sum_{kl \ne i} |G_{kl}^{(i)}|^2 \le \frac{1}{N\eta} \mathrm{I}m\, m^{(i)},$$

Since $m^{(i)} \sim m \sim O(1)$, we have $\Omega \sim |Z_i| \lesssim \frac{C}{\sqrt{N\eta}}$