Bulk universality for Wigner matrices

Benjamin Schlein, University of Cambridge

ETH Zürich, June 10th 2009

Joint work with L. Erdős, J. Ramírez, and H.-T. Yau

Goal: describe statistical properties of eigenvalues of $N \times N$ matrices with random entries, in the limit $N \to \infty$.

Here we will restrict attention to **Wigner matrices** (entries are independent and identically distributed random variables).

Wigner matrices have been introduced by Wigner to describe excitation spectrum of heavy nuclei.

First step towards understanding of more complicated ensembles of random matrices (i.e. band matrices) and **random Schrödinger operators** (Anderson model) in the metallic phase.

Hermitian Wigner Matrices: $N \times N$ matrices $H = (h_{kj})_{1 \le k,j \le N}$ such that $H^* = H$ and

$$h_{kj} = \frac{1}{\sqrt{N}} \left(x_{kj} + i y_{kj} \right) \qquad \text{for all } 1 \le k < j \le N$$

$$h_{kk} = \frac{2}{\sqrt{N}} x_{kk} \qquad \text{for all } 1 \le k \le N$$

where x_{kj}, y_{kj} and x_{kk} $(1 \le k \le N)$ are iid with

$$\mathbb{E}\, x_{jk} = 0 \qquad \text{and} \qquad \mathbb{E}\, x_{jk}^2 = \frac{1}{2} \qquad \left(\Rightarrow \quad \mathbb{E}\, |h_{jk}|^2 = \frac{1}{N} \right)$$

Remark: scaling guarantees that eigenvalues λ_α remain finite as $N \to \infty$.

$$\mathbb{E} \sum_{\alpha=1}^{N} |\lambda_\alpha|^2 = \mathbb{E}\, \mathrm{Tr}\, H^2 = \mathbb{E} \sum_{j,k=1}^{N} |h_{jk}|^2 = N^2\, \mathbb{E}\, |h_{jk}|^2$$

$$\Rightarrow \quad \mathbb{E}|h_{jk}|^2 = O(N^{-1})$$

Semicircle Law (Wigner, 1955): for any $\delta > 0$,

$$\lim_{\eta \to 0} \lim_{N \to \infty} \mathbb{P} \left(\left| \frac{\mathcal{N}[E - \frac{\eta}{2}; E + \frac{\eta}{2}]}{N\eta} - \rho_{\text{sc}}(E) \right| \geq \delta \right) = 0$$

where

$$\mathcal{N}[I] = \text{number of eigenvalues in interval } I$$
$$\rho_{\text{sc}}(E) = \frac{1}{2\pi}\sqrt{1 - E^2/4}.$$

Remark 1: semicircle independent of density of entries.

Remark 2: Wigner result concerns the macroscopic density, that is the density in intervals containing order N eigenvalues.

Gaussian Unitary Ensemble (GUE): simplest example of hermitian Wigner ensemble.

Big advantage: joint eigenvalue distribution is explicit

$$p(\lambda_1, \dots, \lambda_N) = \text{const} \cdot \prod_{i<j}^{N} (\lambda_i - \lambda_j)^2 \, e^{-\frac{N}{2} \sum_{j=1}^{N} \lambda_j^2} .$$

Dyson's sine-kernel distribution for GUE: let

$$p^{(k)}(\lambda_1, \dots, \lambda_k) = \int \mathrm{d}\lambda_{k+1} \dots \mathrm{d}\lambda_N \, p(\lambda_1, \dots, \lambda_N)$$

be the k-point correlation function. Then

$$\frac{1}{\varrho_{sc}^k(E)} p^{(k)}\Big(E + \frac{x_1}{N\varrho_{sc}(E)}, .., E + \frac{x_k}{N\varrho_{sc}(E)}\Big) \to \det\left(\frac{\sin(\pi(x_i - x_j))}{\pi(x_i - x_j)}\right)_{i,j \le k}$$

Universality Conjecture: Dyson's sine-kernel describes local eigenvalue distribution of every hermitian Wigner ensemble.

Semicircle law on microscopic scales: what can be said about density in small intervals?

Theorem 1 [Erdős-S.-Yau, 2008]: Suppose $\mathbb{E}\, e^{\nu x_{ij}^2} < \infty$ for some $\nu > 0$, and fix $|E| < 2$.

Then, for $\delta > 0$ sufficiently small,

$$\mathbb{P}\left(\left| \frac{\mathcal{N}\left[E - \frac{K}{2N}; E + \frac{K}{2N}\right]}{K} - \rho_{sc}(E) \right| \geq \delta \right) \leq C e^{-c\delta\sqrt{K}}$$

for all $K > 0$, uniformly in $N > N_0(\delta)$.

Semicircle law holds up to microscopic scales; this only involves average over finitely many eigenvalues.

On intermediate scales, if $\eta(N) \to 0$ such that $N\eta(N) \to \infty$, we have

$$\lim_{N\to\infty} \mathbb{P}\left(\left| \frac{\mathcal{N}\left[E - \frac{\eta(N)}{2}; E + \frac{\eta(N)}{2}\right]}{N\eta(N)} - \rho_{sc}(E) \right| \geq \delta \right) = 0$$

Delocalization of eigenvectors: let $\mathbf{v} = (v_1, \ldots, v_N)$ be an ℓ_2-normalized vector in $\mathbb{C}^N$. Distinguish two extreme cases:

Complete localization: one large component, for example

$$\mathbf{v} = (1, 0, \ldots, 0) \qquad \Rightarrow \qquad \|\mathbf{v}\|_p = 1, \text{ for all } 2 < p \le \infty$$

Complete delocalization: all components have same size,

$$\mathbf{v} = (N^{-1/2}, \ldots, N^{-1/2}) \quad \Rightarrow \quad \|\mathbf{v}\|_p = N^{-1/2+1/p} \ll 1$$

.

Theorem [Erdős-S.-Yau, 2008]: Suppose $\mathbb{E}\, e^{\nu x_{ij}^2} < \infty$ for some $\nu > 0$, fix $|E| < 2$, $K > 0$, $2 < p < \infty$. Then

$$\mathbb{P}\Big(\exists \mathbf{v} : H\mathbf{v} = \mu \mathbf{v}, |\mu - E| < \frac{K}{N}, \|\mathbf{v}\|_2 = 1, \|\mathbf{v}\|_p \ge M N^{-\frac{1}{2}+\frac{1}{p}}\Big) \le C e^{-c\sqrt{M}}$$

for all M, N large enough.

Level Repulsion [Erdős-S.-Yau, 2008]: Suppose $\mathbb{E}\, e^{\nu x_{ij}^2} < \infty$ for some $\nu > 0$, fix $|E| < 2$.

Fix $k \geq 1$, and assume that the probability density $h(x) = e^{-g(x)}$ of the matrix entries satisfies the bound

$$\left|\widehat{h}(p)\right| \leq \frac{1}{(1+Cp^2)^{\sigma/2}}, \quad \left|\widehat{hg''}(p)\right| \leq \frac{1}{(1+Cp^2)^{\sigma/2}} \quad \text{for } \sigma \geq 5+k^2.$$

Then there exists a constant $C_k > 0$ such that

$$\mathbb{P}\left(\mathcal{N}\left[E - \frac{\varepsilon}{2N}; E + \frac{\varepsilon}{2N}\right] \geq k\right) \leq C_k\, \varepsilon^{k^2}$$

for all N large enough, and all $\varepsilon > 0$.

Remark: for GUE, we have

$$p(\lambda_1, \ldots, \lambda_N) \simeq \prod_{i<j} (\lambda_i - \lambda_j)^2 \quad \Rightarrow \quad \mathbb{P}(\mathcal{N}_\varepsilon \geq k) \simeq \varepsilon^{k^2}$$

Universality of sine-kernel: assume the density of the matrix elements $h(x) = e^{-g(x)} \le Ce^{-c|x|}$, with $g \in C^6(\mathbb{R})$ and

$$\sum_{j=1}^{6} |g^{(j)}(x)| \le C(1+x^2)^k$$

for some $k \in \mathbb{N}$.

Theorem [Erdős-Ramirez-S.-Yau, 2009]: For any $|u| < 2$ and for any bounded $O \in L_c^\infty(\mathbb{R}^2)$ with compact support, we have

$$\lim_{N\to\infty} \int_{\mathbb{R}^2} O(\alpha,\beta) \frac{1}{[\varrho_{sc}(u)]^2}\, p^{(2)}\Big(u + \frac{\alpha}{N\varrho_{sc}(u)}, u + \frac{\beta}{N\varrho_{sc}(u)}\Big) \mathrm{d}\alpha \mathrm{d}\beta$$

$$= \int_{\mathbb{R}^2} O(\alpha,\beta)\Big[1 - \Big(\frac{\sin \pi(\alpha-\beta)}{\pi(\alpha-\beta)}\Big)^2\Big] \mathrm{d}\alpha \mathrm{d}\beta.$$

Remarks:

• Assuming more regularity of density, our result extends to higher order correlation functions

$$\frac{1}{[\rho_{sc}(u)]^k}\, p^{(k)}\Big(u + \frac{x_1}{\rho_{sc}(u)N}, ..., u + \frac{x_k}{\rho_{sc}(u)N}\Big) \\ \to \det\left(\frac{\sin \pi(x_i - x_j)}{\pi(x_i - x_j)}\right)_{1\le i,j\le k}$$

• Shortly after we posted our paper, T. Tao and V. Vu submitted a paper with the same result. Assuming vanishing third moment and exponential decay of the matrix entries, they essentially do not require any regularity of density.

Their method makes use of our results on the microscopic convergence to the semicircle and on the delocalization of the eigenvectors.

Universality for Johansson Matrices: in 2001, K. Johansson proved universality for matrices of the form

$$H = H_0 + t^{\frac{1}{2}} V$$

where V is a GUE-matrix, and H_0 is an arbitrary Wigner matrix.

The matrix H can be obtained by letting every entry of H_0 evolve under a Brownian motion up to time t (more prec. t/N).

The distribution of the eigenvalues of the matrix evolves then according to Dyson's Brownian motion

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

where $\{B_\alpha : 1 \leq \alpha \leq N\}$ is a collection of independent Brownian motion.

The joint probability distribution of the eigenvalues $\mathbf{x} = (x_1, \ldots, x_N)$ of H is

$$p(\mathbf{x}) = \int \mathrm{d}\mathbf{y}\; q_t(\mathbf{x};\mathbf{y})\; p_0(\mathbf{y})$$

where p_0 is the distribution of the eigenvalues $\mathbf{y} = (y_1, \ldots, y_N)$ of H_0 and

$$q_t(\mathbf{x};\mathbf{y}) = \frac{N^{N/2}}{(2\pi t)^{N/2}} \frac{\Delta_N(\mathbf{x})}{\Delta_N(\mathbf{y})} \;\det\left(e^{-N(x_j-y_k)^2/2t}\right)_{j,k=1}^{N},$$

with the Vandermonde determinant

$$\Delta(\mathbf{x}) = \prod_{i<j}^{N} (x_i - x_j) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_N \\ \ldots & \ldots & \ldots & \ldots \\ x_1^N & x_2^N & \ldots & x_N^N \end{pmatrix}$$

This can be proven using the Harish-Chandra/Itzykson-Zuber formula

$$\int_{U(N)} e^{-\frac{N}{2t}\mathrm{Tr}\,(U^* R(\mathbf{x}) U - H_0(\mathbf{y}))^2} \mathrm{d}U = \frac{1}{\Delta(\mathbf{x})\Delta(\mathbf{y})} \;\det\left(e^{-\frac{N}{2t}(x_j-y_i)^2}\right)_{1\le i,j\le N}$$

The k-point correlation function of p is therefore given by

$$p^{(k)}(x_1,\ldots,x_k) = \int q_t^{(k)}(x_1,\ldots,x_k;\mathbf{y})\, p_0(\mathbf{y})\mathrm{d}\mathbf{y}$$

where

$$\begin{aligned} q_t^{(k)}(x_1,\ldots,x_k;\mathbf{y}) &= \int q_t(\mathbf{x};\mathbf{y})\,\mathrm{d}x_{k+1}\ldots\mathrm{d}x_N \\ &= \frac{(N-k)!}{N!}\,\det\Big(K_{t,N}(x_i,x_j;\mathbf{y})\Big)_{1\le i,j\le k} \end{aligned}$$

with

$$\begin{aligned} K_{N,t}(u,v;\mathbf{y}) &= N\frac{e^{N(v^2-u^2)/2t}}{(2\pi i)^2(v-u)t} \\ &\times \int_\gamma \mathrm{d}z \int_\Gamma \mathrm{d}w(1-e^{N(v-u)z/t})\prod_{j=1}^{N}\frac{w-y_j}{z-y_j} \\ &\times \frac{1}{z}\Big(w+z-v-\frac{t}{N}\sum_j \frac{y_j}{(w-y_j)(z-y_j)}\Big)e^{N(w^2-2vw-z^2+2uz)/2t} \end{aligned}$$

where γ is the union of two horizontal lines and Γ is a vertical line in the $\mathbb{C}$-plane.

Convergence of k-point correlation follows from

$$\frac{1}{N\varrho(u)}K_{t,N}(u+\frac{\alpha}{N\varrho(u)}, u+\frac{\beta}{N\varrho(u)}; \mathbf{y}) \to \frac{\sin\pi(\beta-\alpha)}{\pi(\beta-\alpha)} \qquad \text{for a.e. } \mathbf{y}$$

To prove convergence of $K_{t,N}$ Johansson uses identity

$$\frac{1}{N\varrho(u)}K_{t,N}\Big(u, u+\frac{\tau}{N\varrho}; \mathbf{y}\Big) = N\int_\gamma \frac{\mathrm{d}z}{2\pi i}\int_\Gamma \frac{\mathrm{d}w}{2\pi i} h_N(z,w)g_N(z,w)e^{N(f_N(w)-f_N(z))}$$

with

$$f_N(z) = \frac{1}{2t}(z^2-2uz) + \frac{1}{N}\sum_j \log(z-y_j)$$

$$g_N(z,w) = \frac{1}{tz}[w+z-u] - \frac{1}{Nz}\sum_j \frac{y_j}{(w-y_j)(z-y_j)}$$

$$h_N(z,w) = \frac{1}{\tau}\Big(e^{-\tau w/t\varrho} - e^{-\tau(w-z)/t\varrho}\Big)$$

and performs a detailed asymptotic saddle analysis.

Universality with small Gaussian part: what if $t = t(N) \to 0$? Consider

$$t = N^{-1+\varepsilon}$$

Same integral representation but asymptotic analysis is more delicate and requires microscopic convergence to the semicircle.

Recall that

$$\frac{1}{N\varrho(u)} K_{t,N}\Big(u, u + \frac{\tau}{N\varrho(u)}; \mathbf{y}\Big) = N \int_\gamma \frac{\mathrm{d}z}{2\pi i} \int_\Gamma \frac{\mathrm{d}w}{2\pi i} h_N(z,w) g_N(z,w) e^{N(f_N(w) - f_N(z))}.$$

with

$$f_N(z) = \frac{1}{2t}(z^2 - 2uz) + \frac{1}{N}\sum_j \log(z - y_j)$$

Saddles are determined by the equation

$$f_N'(z) = \frac{1}{t}(z-u) + \frac{1}{N}\sum_j \frac{1}{z-y_j} = 0$$

There are two complex conjugated solutions $z = q_N^{\pm}$.

By the convergence to the semicircle on scales of order $N^{-1+\varepsilon}$, we have, with high probability,

$$q_N^{\pm} = q^{\pm} + O(tN^{-\varepsilon/2})$$

where

$$q^{\pm} = u(1-2t) \pm 2ti\sqrt{1-u^2} + O(tN^{-\varepsilon/2})$$

are the two solutions of

$$\frac{1}{t}(q^{\pm}-u) + \int \frac{\varrho_{sc}(y)}{q^{\pm}-y}\mathrm{d}y = 0$$

The integration paths can be shifted to pass through the saddles.

Only important contribution arises from z, w both close to $q_{N,\pm}$.

Contribution from saddles can be computed through local change of variable which makes the exponent quadratic (Laplace method).

As $N \to \infty$, saddle contribution leads to sine-kernel.

This gives convergence to sine-kernel for Wigner ensembles

$$H = H_0 + t^{\frac{1}{2}} V \qquad \text{for } t = N^{-1+\varepsilon}$$

Remark: If matrix elements evolve through Ornstein-Uhlenbeck process, this method proves sine-kernel for Wigner ensembles

$$H = e^{-t/2} H_0 + (1 - e^{-t})^{1/2} V \qquad \text{for } t = N^{-1+\varepsilon}$$

Removal of Gaussian part: let $h(x)$ be the density of the matrix elements of H_0.

The matrix elements of $H = H_0 + t^{\frac{1}{2}} V$ have density

$$h_t(x) = (e^{tL} h)(x), \qquad \text{with} \qquad L = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

Then

$$\int \frac{|h_t(x) - h(x)|^2}{h(x)} dx \le Ct^2$$

Letting $F = h^{\otimes N^2}$ and $F_t = (e^{tL}h)^{\otimes N^2}$ we find

$$\int \frac{|F_t - F|^2}{F} dx_1 \ldots dx_{N^2} \le CN^2 t^2$$

It is only small for $t \ll N^{-1}$.

Hence $t = N^{-1+\varepsilon}$ is still not enough

Idea of time reversal: we would like to write

$$h = e^{tL} f \qquad \text{with} \qquad f = e^{-tL} h$$

as a Gaussian convolution.

But the heat equation cannot be reversed.

$$\Rightarrow \quad \text{approximate inversion of heat semigroup}$$

Define $v_t = (1 - tL)h$. Then

$$e^{tL} v_t = e^{tL}(1 - tL)h \simeq h + t^2 L^2 h \qquad \Big(\text{while} \quad e^{tL} h \simeq h + tLh\Big)$$

Therefore

$$\int \frac{|h_t - h|^2}{h} dx \le C t^4$$

Hence, if $F = h^{\otimes N^2}$ and $F_t = h_t^{\otimes N^2}$, we find

$$\int \frac{|F_t - F|^2}{F} dx_1 \ldots dx_{N^2} \le C N^2 t^4 \ll 1 \qquad \text{for } t = N^{-1+\varepsilon}$$

Rate of convergence can be improved with better approximation of inverse evolution.

To prove that the 2-point correlation functions of F_t and F are asymptotically equal, we need $v_t = (1 - tL + t^2L^2/2)h$.

$$\Rightarrow \qquad e^{tL}v_t - h \simeq t^3L^3h$$

This explains the condition $h \in C^6(\mathbb{R})$.

To prove convergence for higher marginals, we need higher order approximations (and hence more regularity).

For this idea to work, we need v_t to be a probability density.

v_t is automatically normalized, with correct mean and variance.

We have to make sure that v_t is non-negative! This leads to assumption $h(x) = e^{-g(x)}$ where g does not grow too fast.