Universality of local spectral statistics of random matrices

László Erdős
Ludwig-Maximilians-Universität, Munich, Germany

CRM, Montreal, Mar 19, 2012

Joint with P. Bourgade, B. Schlein, H.T. Yau, and J. Yin
“Perhaps I am now too courageous when I try to guess the distribution of the distances between successive levels (of energies of heavy nuclei). Theoretically, the situation is quite simple if one attacks the problem in a simpleminded fashion. The question is simply what are the distances of the characteristic values of a symmetric matrix with random coefficients.”

Eugene Wigner, 1956

Nobel prize 1963
INTRODUCTION

**Basic question [Wigner]:** 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? Do some universal patterns emerge and what determines them?

$$H = \begin{pmatrix} h_{11} & h_{12} & \ldots & h_{1N} \\ h_{21} & h_{22} & \ldots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1} & h_{N2} & \ldots & h_{NN} \end{pmatrix} \Rightarrow (\lambda_1, \lambda_2, \ldots, \lambda_N) \text{ Eigenvalues?}$$

$$N = \text{size of the matrix, will go to infinity.}$$

**Analogy:** Central limit theorem: $$\frac{1}{\sqrt{N}}(X_1 + X_2 + \ldots + X_N) \sim N(0, \sigma^2)$$
**Gaussian Unitary Ensemble (GUE):**

\[ H = (h_{jk})_{1 \leq j,k \leq N} \text{ hermitian } N \times N \text{ matrix with} \]

\[ h_{jk} = \overline{h_{kj}} = \frac{1}{\sqrt{N}} (x_{jk} + i y_{jk}) \quad \text{and} \quad h_{kk} = \frac{\sqrt{2}}{\sqrt{N}} x_{kk} \]

where \( x_{jk}, y_{jk} \) (for \( j < k \)) and \( x_{kk} \) are independent standard Gaussian

Eigenvalues: \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N \).

They are of order one:

\[ \mathbb{E} \frac{1}{N} \sum_i \lambda_i^2 = \mathbb{E} \frac{1}{N} \text{Tr}H^2 = \frac{1}{N} \sum_{ij} \mathbb{E}|h_{ij}|^2 = 2 \]

at least in average sense.

Hermitian can be replaced with symmetric or quaternion self-dual (GOE, GSE)
Wigner semicircle law

\[ \rho(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \]

Typical evaleve spacing (gap): \( \delta_i = \lambda_{i+1} - \lambda_i \sim \frac{1}{N} \) (in the bulk)

**Observations:**  
1) Semicircle density.  
2) Level repulsion.

Holds for other symmetry classes GUE, GOE, GSE.

For Wishart matrices, i.e. matrices of the form \( H = AA^* \), where the entries of \( A \) are i.i.d.: Marchenko-Pastur law
• **E. Wigner (1955):** The excitation spectra of heavy nuclei have the same **spacing distribution** as the eigenvalues of GOE. Experimental data for excitation spectra of heavy nuclei: \(^{238}U\)

![Graph of excitation spectra](image)

*Typical Poisson statistics:*

\[ \begin{array}{ccccccccccc}
| & | & | & | & | & | & | & | & | \\
\end{array} \]

*Typical random matrix eigenvalues*

\[ \begin{array}{ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
Level spacing (gap) histogram for different point processes.

NDE – Nuclear Data Ensemble, resonance levels of 30 sequences of 27 different nuclei.
SINE KERNEL FOR CORRELATION FUNCTIONS

Probability density of the eigenvalues: \( p(x_1, x_2, \ldots, x_N) \)

The \( k \)-point correlation function is given by

\[
p_N^{(k)}(x_1, x_2, \ldots, x_k) := \int_{\mathbb{R}^{N-k}} p(x_1, \ldots, x_k, x_{k+1}, \ldots, x_N) \, dx_{k+1} \ldots dx_N
\]

Special case: \( k = 1 \) (density)

\[
\varrho_N(x) := p_N^{(1)}(x) = \int_{\mathbb{R}^{N-1}} p(x, x_2, \ldots, x_N) \, dx_2 \ldots dx_N
\]

It allows to compute expectation of observables with one eigenvalue:

\[
\mathbb{E} \frac{1}{N} \sum_{i=1}^{N} O(\lambda_i) = \int O(x) \varrho_N(x) \, dx \to \frac{1}{2\pi} \int O(x) \sqrt{4 - x^2} \, dx
\]

Higher \( k \) computes observables with \( k \) values.
Local level correlation statistics for GUE  [Gaudin, Dyson, Mehta]

\[
\lim_{N \to \infty} \frac{1}{[\rho(E)]^2} p_N^{(2)} \left( E + \frac{x_1}{N \rho(E)}, E + \frac{x_2}{N \rho(E)} \right) = \det \left\{ S(x_i - x_j) \right\}_{i,j=1}^2

\]

for any |E| < 2 (bulk spectrum), where \( S(x) := \frac{\sin \pi x}{\pi x} \)

\[
= 1 - \left( \frac{\sin \pi (x_1 - x_2)}{\pi (x_1 - x_2)} \right)^2
\]

(\( \Rightarrow \) Level repulsion)

9
\( k \)-point correlation functions are given by \( k \times k \) determinants:

\[
\lim_{N \to \infty} \frac{1}{[\rho(E)]^k} p_N^{(k)} \left( E + \frac{x_1}{N \rho(E)}, E + \frac{x_2}{N \rho(E)}, \ldots, E + \frac{x_k}{N \rho(E)} \right) = \det \left\{ S(x_i - x_j) \right\}_{i,j=1}^k
\]

The limit is independent of \( E \) as long as \( E \) is in the bulk spectrum, i.e. \(|E| < 2\).

Gap distribution (original question of Wigner) is obtained from correlation functions by the exclusion-inclusion formula.

**Main question:** going beyond Gaussian towards universality!

There are two almost disjoint directions of generalization: Gaussian is the common intersection.
GENERALIZATION NO.1: INvariant ensembles

**Unitary ensemble**: Hermitian matrices with density

\[ P(H)dH \sim e^{-\text{Tr} V(H)}dH \]

Invariant under \( H \rightarrow UHU^{-1} \) for any unitary \( U \)

Joint density function of the eigenvalues is explicitly known

\[ p(\lambda_1, \ldots, \lambda_N) = \text{const.} \prod_{i<j} (\lambda_i - \lambda_j)^\beta e^{-\sum_j V(\lambda_j)} \]

classical ensembles \( \beta = 1, 2, 4 \) (orthogonal, unitary, symplectic symmetry classes; GOU, GUE, GSE for Gaussian case, \( V(x) = x^2/2 \))

Correlation functions can be explicit computed via orthogonal polynomials due to the Vandermonde determinant structure.

large \( N \) asymptotic of orthogonal polynomials \( \Rightarrow \) local eigenvalue statistics indep of \( V \). But density of e.v. depends on \( V \).
GENERALIZATION NO.2:  
(GENERALIZED) WIGNER ENSEMBLES

\[ H = (h_{ij})_{1 \leq i, j \leq N}, \quad \bar{h}_{ji} = h_{ij} \quad \text{independent} \]

\[ \mathbb{E} h_{ij} = 0, \quad \mathbb{E} |h_{ij}|^2 = \sigma_{ij}^2, \quad \sum_i \sigma_{ij}^2 = 1, \]

\[ \frac{c}{N} \leq \sigma_{ij}^2 \leq \frac{C}{N} \]

Moment condition: \[ \mathbb{E} |\sqrt{N} h_{ij}|^{4+\varepsilon} < C \]

If \( h_{ij} \) are i.i.d. then it is called Wigner ensemble.

**Universality conjecture (Dyson, Wigner, Mehta etc):** If \( h_{ij} \) are independent, then the local eigenvalues statistics are the same as for the Gaussian ensembles.
Several previous results for invariant ensembles

Dyson (1962-76), Gaudin-Mehta (1960- ) classical Gaussian ensembles via Hermite polynomials


All these results are limited to invariant ensembles and to the classical values of $\beta = 1, 2, 4$ (OP Method). For non-classical values, there is no underlying matrix ensemble, but the Gibbs measure

$$p(\lambda_1, \ldots, \lambda_N) = \text{const.} \prod_{i<j} (\lambda_i - \lambda_j)^\beta e^{-\beta N \sum_j V(\lambda_j)}$$

can still be studied ("log-gas"). $\implies$ PROBLEM 1.

No previous results for Wigner (apart from Johansson’s for hermitian matrices with Gaussian convolution)

Universality of Wigner matrices? $\implies$ PROBLEM 2.
**PROBLEM 1: NON-CLASSICAL β-ENSEMBLES**

\[ p(\lambda_1, \ldots, \lambda_N) = \text{const.} \prod_{i<j} (\lambda_i - \lambda_j)^\beta e^{-\beta N \sum_j V(\lambda_j)} \]

Limit density \( \varrho \) is the unique minimizer of

\[ I(\nu) = \int_\mathbb{R} V(t)\nu(t)dt - \int_\mathbb{R} \int_\mathbb{R} \log |t - s|\nu(s)\nu(t)dtds. \]

**Theorem** [Bourgade-E-Yau, 2011] Let \( \beta > 0 \) and \( V \) be real analytic. Let \( p^{(k)}_{V,N} \) and \( p^{(k)}_{G,N} \) be the \( k \)-point correlation functions for \( V \) and for the Gaussian case, \( V(x) = x^2/2 \).

Fix \( E \in \text{int}(\text{supp}\varrho) \), \( E' \in \text{int}(\text{supp}\varrho_{sc}) \) and \( \varepsilon \equiv N^{-1/2} \), then

\[
\int_{E-\varepsilon}^{E+\varepsilon} dx \frac{1}{2\varepsilon \varrho(E)^k} p^{(k)}_{V,N}(x + \frac{\alpha_1}{N\varrho(E)}, \ldots, x + \frac{\alpha_k}{N\varrho(E)}) - \int_{E'-\varepsilon}^{E'+\varepsilon} dx \frac{1}{2\varepsilon \varrho_{sc}(E')^k} p^{(k)}_{G,N}(x + \frac{\alpha_1}{N\varrho_{sc}(E')}, \ldots, x + \frac{\alpha_k}{N\varrho_{sc}(E')}) \rightarrow 0.
\]

weakly in \( \alpha_1, \ldots, \alpha_k \) as \( N \rightarrow \infty \).
**PROBLEM 2: NON-ININVARIANT WIGNER ENSEMBLES**

**Theorem** [E-Schlein-Yau-Yin, 2009-2010] The bulk universality holds for generalized Wigner ensembles i.e., for \(|E| < 2, \varepsilon = N^{-1+\delta}, \delta > 0\)

\[
\lim_{N \to \infty} \int_{E-\varepsilon}^{E+\varepsilon} \frac{dx}{2\varepsilon} \left( p_{F,N}^{(k)} - p_{\mu,N}^{(k)} \right) \left( x + \frac{b_1}{N}, \ldots, x + \frac{b_k}{N} \right) = 0 \text{ weakly}
\]

<table>
<thead>
<tr>
<th>(F)</th>
<th>(\mu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>generalized symmetric matrices</td>
<td>GOE</td>
</tr>
<tr>
<td>generalized hermitian</td>
<td>GUE</td>
</tr>
<tr>
<td>generalized self-dual quaternion</td>
<td>GSE</td>
</tr>
<tr>
<td>real covariance</td>
<td>real Gaussian Wishart</td>
</tr>
<tr>
<td>complex covariance</td>
<td>complex Gaussian Wishart</td>
</tr>
</tbody>
</table>

Variances can vary in this theorem.

We also have a similar result at the spectral edge (universality of Tracy-Widom distribution)
Adjacency matrix $A = (a_{ij})$, real symmetric with

$$a_{ij} = \frac{\gamma}{q} \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p, \end{cases}$$

where $q := \sqrt{pN}$ and $\gamma = (1 - p)^{-1/2}$ so that $\text{Var} \ a_{ij} = N^{-1}$.

Note that $\mathbb{E} a_{ij} \neq 0 \implies$ there is a large eigenvalue.

Each column typically has $pN = q^2$ nonvanishing entries.

Bulk is given by the (analogue of) the sine-kernel.

Edge is given by the Airy kernel and Tracy-Widom.

Single outlier is Gaussian.
**RECENT RESULTS ON BULK UNIVERSALITY**

1. **Hermitian** ensemble with $C^6$ distribution. [EPRSY 2009]. (Brezin-Hikami, contour integral and reverse heat flow approach)

2. **Hermitian** Wigner ensemble with probability law supported on at least three points [Tao-Vu] (Extension to Bernoulli in [ERSTVY]). **Symmetric** ensemble with the first 4 moments of matrix elements matching the GOE [Tao-Vu] (4-moment approach)

3. **Symmetric** ensemble with three point condition [E-Schlein-Yau]. (Dyson Brownian Motion (DBM) flow approach)

4. **Generalized symmetric or hermitian Wigner** ensembles (the variances were allowed to vary) [E-Yau-Yin].

5. **Erdős-Rényi sparse matrices** with $pN \gg N^{2/3}$ [E-Knowles-Yau-Yin]

Similar development for real and complex sample covariance ensembles [E-Schlein-Yau-Yin], [Tao-Vu], [Peche], and also for edge univ.
KEY STEPS IN OUR PROOF FOR THE WIGNER CASE

**Step 1.** Good local semicircle law including a control near the edge.

**Method:** System of self-consistent equations for the Green function, control the error by large deviation methods.

**Step 2.** Universality for Wigner matrices with a small ($\sim N^{-\epsilon}$) Gaussian component.

**Method:** Modify DBM to speed up its local relaxation, then show that the modification is irrelevant for statistics involving differences of eigenvalues.

**Step 3.** Universality for arbitrary Wigner matrices.

**Method:** Remove the small Gaussian component in Step 2 by resolvent perturbation theory and moment matching.
**Step 1: LOCAL SEMICIRCLE LAW**

Green function:  
\[ G_{ij} = \frac{1}{H - z}(i, j), \quad m(z) = \frac{1}{N} \text{Tr} G = \frac{1}{N} \sum_i G_{ii} \]

Let \( m_{sc} \) be the Stieltjes transform of the semicircle measure, i.e.,

\[ m_{sc}(z) = \int \frac{\rho_{sc}(x) \, dx}{x - z}, \quad \rho_{sc}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \]

**Theorem** [Erdős-Y-Yin, 2010] For any \( z = E + i\eta \) with \( \eta \gtrsim N^{-1} \) the following holds with exponentially high probability:

\[ |m(z) - m_{sc}(z)| \lesssim \frac{1}{N\eta} \]

where \( \lesssim \) means up to \((\log N)\#\) factors. Estimates are optimal.
Step 2: DYSON BROWNIAN MOTION

Gaussian convolution matrix interpolates between Wigner and GUE.

Evolve the matrix elements with an OU process:

\[
    dH_t = \frac{1}{\sqrt{N}} dB_t - \frac{1}{2} H_t dt \quad H_t \sim e^{-t/2} H_0 + (1 - e^{-t})^{1/2} V.
\]

\[
    d\lambda_i = \frac{1}{\sqrt{N}} dB_i + \left( -\frac{1}{2} \lambda_i + \frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} \right) dt
\]

Idea: Equilibrium is the invariant ensemble (GUE, etc.) with known local statistics.

Global equilibrium is reached in time \(O(1)\) (convexity, Bakry-Emery).

For local statistics, only local equilibrium needs to be achieved which is much faster. Our main result proves Dyson’s conjecture on Dyson’s Brownian motion:
“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.”

Freeman Dyson, 1962

on the approach to equilibrium of Dyson Brownian Motion

Global equilibrium is reached in time scale of $O(1)$. Local equilibrium was believed to be reached in $O(N^{-1})$. 
OUTLOOK: UNIVERSALITY CONJECTURES

• **Quantum Chaos Conjecture** (vague)
  classical dynamics with potential $V$  
  e.v. gap of $-\Delta + V$

  chaotic  
  GOE statistics

  integrable  
  Poisson statistics


• **Anderson Model (1958):** $V_\omega$ random potential on $\mathbb{R}^d$ or $\mathbb{Z}^d$

  random Schrödinger operator: $H = -\Delta + \lambda V_\omega$

Depending on $\lambda$ and $d$, there are two distinct regimes.
I: **Insulator regime:** Pure point spectrum in infinite volume limit. Eigenvectors are exponentially localized. No diffusion/transport. Poisson local statistics.

II: **Conductor regime:** absolutely continuous spectrum. Eigenvectors are delocalized (do not decay). Diffusive transport. GOE local statistics.

I. is relatively well understood, II. is not.
**Conjectured Dichotomy:** There are essentially two different behaviors for local eigenvalue statistics of disordered quantum systems:

**A:** Poisson statistics, for systems with little or no correlations.

**B:** Random matrix statistics: for systems with high correlations.

**Fundamental belief of universality:** The macroscopic statistics (like density of states) depend on the models, but the microscopic statistics are independent of the details of the systems except the symmetries.

Our results on Wigner matrices verify this conjecture for random matrices, but it is still mean field.

**Major goal:** move towards random Schrödinger.
SUMMARY

1. We proved bulk universality for general $\beta > 0$ ensemble with real analytic potential.

2. We proved bulk and edge universality for generalized Wigner matrices (varying variance and even singular distribution – Erdős-Rényi sparse matrices)
1. Universality for sparser matrices (eventually \(pN \sim O(1)\)).

2. **Random band matrices:** \(H\) is symmetric with independent but not identically distributed entries with mean zero and variance

\[
\mathbb{E} |h_{k\ell}|^2 = W^{-1} e^{-|k-\ell|/W}
\]

**Conjecture** (even Gaussian case is open)

Narrow band, \(W \ll \sqrt{N} \implies\) localization, Poisson statistics

Broad band, \(W \gg \sqrt{N} \implies\) delocalization, GOE statistics

3. **Spectral statistics for Anderson model.**
Zeros of the Riemann-zeta function (Detour)

\[ \zeta(s) = \sum_n \frac{1}{n^s}, \quad \zeta\left(\frac{1}{2} + i\gamma_n\right) = 0, \quad \hat{\gamma}_n = \frac{1}{2\pi} \gamma_n \log \gamma_n, \quad \delta_n = \hat{\gamma}_n + 1 - \hat{\gamma}_n \]

**Figure 3**
Probability density of the normalized spacings \( \delta_n \). Solid line: GUE prediction. Scatter plot: empirical data based on zeros \( \gamma_n \), \( 1 \leq n \leq 10^5 \).

**Figure 4**
Probability density of the normalized spacings \( \delta_n \). Solid line: GUE prediction. Scatter plot: empirical data based on zeros \( \gamma_n \), \( 10^{12} + 1 \leq n \leq 10^{15} - 10^5 \).
Figure 1. Nearest neighbor spacings among 70 million zeroes beyond the $10^{20}$-th zero of zeta, verses $\mu_1(\text{GUE})$.

[Odlyzko, ATT, 1989]