From the mesoscopic to microscopic scale in random matrix theory

(fixed energy universality for random spectra)

With L. Erdős, H.-T. Yau, J. Yin
A spacially confined quantum mechanical system can only take certain discrete values of energy. Uranium-238:

Quantum mechanics postulates that these values are eigenvalues of a certain Hermitian matrix (or operator) $H$, the Hamiltonian of the system.

The matrix elements $H_{ij}$ represent quantum transition rates between states labelled by $i$ and $j$.

**Wigner’s universality idea (1956).** Perhaps I am too courageous when I try to guess the distribution of the distances between successive levels. 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.
**Wigner’s model**: the Gaussian Orthogonal Ensemble,

(a) Invariance by $H \mapsto U^* H U$, $U \in O(N)$.

(b) Independence of the $H_{i,j}$'s, $i \leq j$.

The entries are Gaussian and the spectral density is

$$
\frac{1}{Z_N} \prod_{i<j} |\lambda_i - \lambda_j|^\beta e^{-\beta \frac{N}{4} \sum_i \lambda_i^2}
$$

with $\beta = 1$ (2, 4 for invariance under unitary or symplectic conjugacy).

- Semicircle law as $N \to \infty$.
- Limiting bulk local statistics of GOE/GUE/GSE calculated by Gaudin, Mehta, Dyson.
Dyson’s description of the first experiments. All of our struggles were in vain. 82 levels were too few to give a statistically significant test of the model. As a contribution of the understanding of nuclear physics, random matrix theory was a dismal failure. By 1970 we had decided that it was a beautiful piece of work having nothing to do with physics.

As $N \to \infty$ and the nuclei statistics are performed over a large sample, the gap probability agree (resonance levels of 30 sequences of 27 different nuclei).
Fundamental belief in universality: macroscopic statistics depend on the model, but microscopic statistics only depend on the symmetries.

- GOE: Hamiltonians of systems with time reversal invariance
- GUE: no time reversal symmetry (e.g. application of a magnetic field)
- GSE: time reversal but no rotational symmetry

This is not proved for any realistic Hamiltonian.

The local universality is now known for random matrices. In the definition of the Gaussian ensembles, either keep:

- the independence of the entries (Wigner ensembles);
- or the conjugacy invariance (Invariant ensembles).

This talk is only about Wigner matrices, $N \times N$ matrices such that

$$
\mathbb{E}(X_{ij}) = 0, \quad \mathbb{E}(X_{ij}^2) = \frac{1}{N}, \quad \text{high moments are finite but arbitrary}.
$$

The developed techniques also apply to varying variances, covariance matrices, mean-field models of sparse random graphs.
Some statistics concerning point processes.

- **Correlation functions.** For a point process \( \chi = \sum_{i=1}^{N} \delta \lambda_i \):

  \[
  \rho_{k}^{(N)}(x_1, \ldots, x_k) = \lim_{\varepsilon \to 0} \varepsilon^{-k} \mathbb{P} \left( \chi(x_i, x_i + \varepsilon) = 1, 1 \leq i \leq k \right).
  \]

Gaudin, Dyson, Mehta (GUE for example) : for any \( E \in (-2, 2) \),

\[
\rho_{k}^{(N)} \left( E + \frac{u_1}{N \varrho(E)}, \ldots, E + \frac{u_k}{N \varrho(E)} \right) \xrightarrow{N \to \infty} \det_{k \times k} \frac{\sin(\pi(u_i - u_j))}{\pi(u_i - u_j)}.
\]

- **Counting numbers.**

  \[
  \lim_{N \to \infty} \mathbb{P} \left( |\lambda_i - E| \geq \frac{\alpha}{N \pi \varrho(E)} \right) = E(0, \alpha).
  \]

Jimbo-Miwa-Mori-Sato (GUE for example) : \( E(0, \alpha) \) exists, it is independent of \( E \) and satisfies a Painlevé equation.

Pointwise convergence of correlation functions cannot hold.

Recently universality was proved under various forms.

Fixed (averaged) energy universality. For any $k \geq 1$, smooth $F : \mathbb{R}^k \to \mathbb{R}$, for arbitrarily small $\varepsilon$ and $s = N^{-1+\varepsilon}$,

$$
\lim_{N \to \infty} \frac{1}{\varrho(E)^k} \int_E^{E+s} \frac{dx}{s} \int \, d\mathbf{v} F(\mathbf{v}) \rho_k^{(N)} \left( x + \frac{\mathbf{v}}{N\varrho(E)} \right) d\mathbf{v} = \int d\mathbf{v} F(\mathbf{v}) \rho_k^{(\text{GOE})} (\mathbf{v}).
$$
Johansson (2001)  Hermitian class, fixed $E$

Erdős Schlein Péché Ramirez Yau (2009)  Hermitian class, fixed $E$

Tao Vu (2009)  Hermitian class, fixed $E$

Erdős Schlein Yau (2010)  Any class, averaged $E$

For symmetric matrices, Tao and Vu’s four moments theorem states that universality holds (including for symmetric matrices) if the Wigner matrix first four moments are 0,1,0,3.

This does not include Jimbo, Miwa, Mori, Sato relations for gaps in the spectrum of Bernoulli matrices, for example.

Assume $\lambda_1 \leq \cdots \leq \lambda_N$ and define $\int_{-2}^{\gamma_i} d\varrho = \frac{i}{N}$.

Key input for all recent results: **rigidity of eigenvalues** (Erdős Yau Yin): $|\lambda_k - \gamma_k| \leq N^{-1+\epsilon}$ in the bulk. Is this the optimal rigidity? Are fluctuations in the bulk Gaussian?
The true size of fluctuations is suggested by the following theorem.

**Theorem (Gustavsson, O’Rourke)**

Consider GOE/GUE/GSE. Let $k_0$ a bulk index and $k_{i+1} \sim k_i + N \theta_i$, $0 < \theta_i < 1$. Then the normalized eigenvalues fluctuations

$$X_i = \frac{\lambda_{k_i} - \gamma_{k_i}}{\sqrt{\log N}} \sqrt{\frac{\beta (4 - \gamma_{k_i}^2)}{N}}$$

converge to a Gaussian vector with covariance

$$\Lambda_{ij} = 1 - \max\{\theta_k, i \leq k < j\}.$$

In particular, $\lambda_i - \gamma_i$ has Gaussian fluctuations of size $\sqrt{\frac{\log N}{N}}$.

Extension to Wigner matrices hold under the four moment matching assumption (O’Rourke, Dallaporta, Vu).

Proof: determinantal point processes a la Costin-Lebowitz (GUE) + decimation relations (GOE, GSE).
This is a **Log-correlated field**, as explained below. The main difficulty proving fixed energy universality is the very slow decay of correlations:

\[
\langle N\lambda_i, N\lambda_j \rangle \sim \log \left( \frac{N}{1 + |i - j|} \right).
\]

Each eigenvalue is localized on a very small window, almost regular spacing: smooth density and translation invariance are not necessarily intuitive for Wigner matrices.

Log-correlated random fields appear in the study of random surfaces, Liouville quantum gravity, models of turbulence...

**Theorem (with Erdös, Yau, Yin)**

Fixed energy universality holds for Wigner matrices from all symmetry classes. Individual eigenvalues fluctuate as a Log-correlated Gaussian field.
The **Dyson Brownian Motion** (DBM, \(dH_t = \frac{dB_t}{\sqrt{N}} - \frac{1}{2} H_t dt\)) is an essential interpolation tool in our proof, as in the Erdős Schlein Yau approach to universality, which can be summarized as follows:

\[
\begin{align*}
H_0 &
\Downarrow \\
\tilde{H}_0 \quad \xrightarrow{\text{(DBM)}} \quad \tilde{H}_t \\
\end{align*}
\]

\(\xrightarrow{\text{(DBM)}}\) : for \(t = N^{-1+\varepsilon}\), the eigenvalues of \(\tilde{H}_t\) satisfy averaged universality.

\(\Downarrow\) : Density argument. For any \(t \ll 1\), there exists \(\tilde{H}_0\) such that \(H_0\) and \(\tilde{H}_t\) have the same statistics on the microscopic scale.

What makes the Hermitian universality easier? The \(\xrightarrow{\text{(DBM)}}\) step is replaced by HCIZ formula: correlation functions of \(\tilde{H}_t\) are explicit only for \(\beta = 2\).
A few curiosities about the proof of fixed energy universality.

(i) A game coupling three Dyson Brownian Motions.

(ii) Homogenization allows to obtain microscopic statistics from mesoscopic ones.

(iii) Need of a second order type of Hilbert transform. Emergence of new explicit kernels for any Bernstein-Szegő measure. These include Wigner, Marchenko-Pastur, Kesten-McKay.

(iv) The relaxing time of DBM depends on the Fourier support of the test function: the step $\text{DBM} \rightarrow$ becomes the following.

$$\tilde{F}(\lambda, \Delta) = \sum_{i_1, \ldots, i_k = 1}^{N} F \left( \{ N(\lambda_{i_j} - E) + \Delta, 1 \leq j \leq k \} \right)$$

Fact

If $\text{supp} \hat{F} \subset B(0, 1/\sqrt{\tau})$, then for $t = N^{-\tau}$,

$$\mathbb{E} \tilde{F}(\lambda_t, 0) = \mathbb{E} \tilde{F}(\lambda^{\text{GOE}}, 0).$$
First step: coupling two Dyson Brownian Motions. Let \( x(0) \) be the eigenvalues of \( \tilde{H}_0 \) and \( y(0), z(0) \) those of two independent GOE.

\[
\frac{dx_i}{dy_i}/dz_i = \sqrt{\frac{2}{N}} dB_i(t) + \frac{1}{N} \left( \sum_{j \neq i} \frac{1}{x_i/y_i/z_i - x_j/y_j/z_j} - \frac{1}{2} \frac{x_i/y_i/z_i}{x_j/y_j/z_j} \right) dt
\]

Let \( \delta_\ell(t) = e^{t/2}(x_\ell(t) - y_\ell(t)) \). Then we get the parabolic equation

\[
\partial_t \delta_\ell(t) = \sum_{k \neq \ell} B_{k\ell}(t) (\delta_k(t) - \delta_\ell(t)), \quad B_{k\ell}(t) = \frac{1}{N(x_k(t) - x_\ell(t))(y_k(t) - y_\ell(t))}.
\]

By the de Giorgi-Nash-Moser method, Caffarelli-Chan-Vasseur and Erdős-Yau, this PDE is Hölder-continuous for \( t > N^{-1+\varepsilon} \), i.e.

\[
\delta_\ell(t) = \delta_{\ell+1}(t) + O(N^{-1-\varepsilon}), \text{ i.e. gap universality:}
\]

\[
x_{\ell+1}(t) - x_\ell(t) = y_{\ell+1}(t) - y_\ell(t).
\]

This is not enough for fixed energy universality.
Second step: homogenization. The continuum-space analogue of our parabolic equation is

\[ \partial_t f_t(x) = (K f_t)(x) := \int_{-2}^{2} \frac{f_t(y) - f_t(x)}{(x - y)^2} \varrho(y) dy. \]

\( K \) is some type of second order Hilbert transform.

Fact

Let \( f_0 \) be a smooth continuous-space extension of \( \delta(0) : f_0(\gamma_\ell) = \delta_\ell(0) \). Then for any small \( \tau > 0 \) \((t = N^{-\tau})\) there exists \( \varepsilon > 0 \) such that

\[ \delta_\ell(t) = (e^{tK} f_0)_\ell + O(N^{-1-\varepsilon}). \]

Proof. Rigidity of the eigenvalues and the Duhamel formula.
Third step : the continuous-space kernel.

1. For the translation invariant equation

\[ \partial_t g_t(x) = \int_{\mathbb{R}} \frac{g_t(y) - g_t(x)}{(x - y)^2} dy, \]

the fundamental solution is the Poisson kernel \( p_t(x, y) = \frac{c_t}{t + (x-y)^2} \).

2. For us, \( t \) will be close to 1, so the edge curvature cannot be neglected. Fortunately, \( K \) can be fully diagonalized and \( (x = 2 \cos \theta, y = 2 \cos \phi) \)

\[ k_t(x, y) = \frac{c_t}{|e^{i(\theta+\phi)} - e^{-t/2}|^2|e^{i(\theta-\phi)} - e^{-t/2}|^2}. \]

Called the Mehler kernel by Biane in free probability context, not considered as a second-order Hilbert transform fundamental solution.

3. Explicit kernels can be obtained for all Bernstein-Szego measures, \( \varrho \)

\[ \varrho(x) = \frac{c_{\alpha, \beta}(1 - x^2)^{1/2}}{(\alpha^2 + (1 - \beta^2)) + 2\alpha(1 + \beta)x + 4\beta x^2}. \]
Fourth step: from mesoscopic to microscopic. Homogenization yields

\[ \delta_\ell(t) = \int k_t(x, y)f_0(y)\varrho(y)dy + O(N^{-1-\varepsilon}) \]

The LHS is microscopic-type of statistics, the RHS is mesoscopic. This yields, up to negligible error,

\[ Nx_\ell(t) = Ny_\ell(t) - \Psi_t(y_0) + \Psi_t(x_0), \]

where \( \Psi_t(x_0) = \sum h(N^\tau(x_i(0) - E)) \) for some smooth \( h \). We wanted to prove

\[ \mathbb{E} \tilde{F}(x_t, 0) = \mathbb{E} \tilde{F}(z_t, 0) + o(1). \]

We reduced it to

\[ \mathbb{E} \tilde{F}(y_t, -\Psi_t(y_0) + \Psi_t(x_0)) = \mathbb{E} \tilde{F}(y_t, \Psi_t(y_0) + \Psi_t(z_0)) + o(1). \]

The observables \( \Psi_t(y_0), \Psi_t(x_0) \) and \( \Psi_t(z_0) \) are mesoscopic and independent, while \( x_t \) and \( z_t \) are microscopic and dependent.
Fifth step and conclusion: CLT for GOE beyond the natural scale. Do $\Psi_t(x_0)$ and $\Psi_t(y_0)$ have the same distribution? No, their variance depend on their fourth moment.

A stronger result holds: $\mathbb{E} \tilde{F}(y_t, -\Psi_t(y_0) + c)$ does not depend on $c$.

We know that $\mathbb{E} \tilde{F}(y_t, -\Psi_t(y_0) + \Psi_t(z_0) + c) = \mathbb{E} \tilde{F}(y_t, -\Psi_t(y_0) + \Psi_t(z_0))$.

**Exercise**

Let $X$ be a random variable. If $\mathbb{E} g(X + c) = 0$ for all $c$, is it true that $g \equiv 0$?

Not always. But true if $X$ is Gaussian (by Fourier).

**Lemma**

$$\mathbb{E} (e^{i\lambda \Psi_t(z(0))}) = e^{-\frac{\lambda^2}{2} \tau \log N} + O(N^{-1/100}).$$

The proof uses algebraic ideas of Johansson and rigidity of $\beta$-ensembles (with Erdős and Yau).

By Parseval, proof when the support of $\hat{F}$ has size $1/\sqrt{\tau}$. This is why DBM needs to be run till time almost 1.
The homogenization idea has also applications to eigenvectors perturbations in non-perturbative regime.

Some problems about microscopic statistics of random matrices:

1. Are extreme gaps and extreme deviations universal?
2. Log-correlated field for $\beta$-ensembles?
3. Classification of underlying spectral measures such that $\mathcal{K}$ is integrable?

Main problem in the field now: beyond the mean-field case, universality for sparse+geometry-dependent models of random matrices.

Curiosities:

- Dynamical ideas to solve time independent problems: usually study the limits of the flow at $t \to \infty$.
- Homogenization allows to access microscopic statistics from mesoscopic ones.