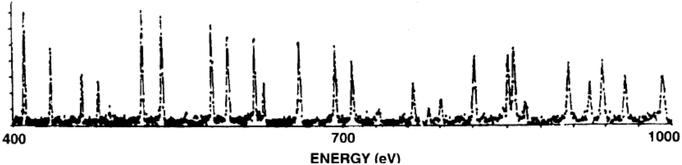
Homogenization of the Dyson Brownian Motion

P. Bourgade, joint work with L. Erdős, J. Yin, H.-T. Yau

Cincinnati symposium on probability theory and applications, September 2014

A spacially confined quantum mechanical system can only take on 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.

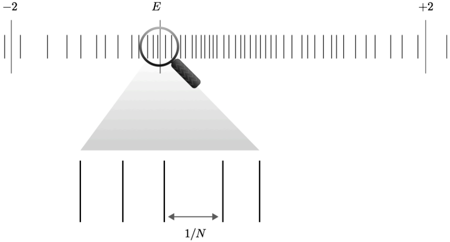
The matrix elements H_{ij} represent quantum transition rates between states labelled by i and j.

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 $H \mapsto U^* H U$, $U \in \mathrm{O}(N)$.
(b) Independence of the $H_{i,j}$'s, $i \leq j$.
The entries are Gaussian and the spectral density is

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

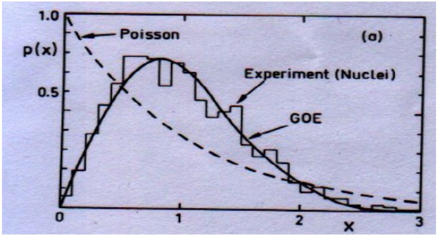
with $\beta = 1$ (2, 4 for invariance under unitary or symplectic conjugacy).



- 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.

When $N \to \infty$ and the nuclei statistics performed over a large sample, the gap probability agree (resonance levels of 30 sequences of 27 different nuclei).



Fundamental belief in universality : the macroscopic statistics (like the equilibrium measure) depend on the models, but the microscopic statistics are independent of the details of the systems except 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

Correlation functions. For a point process $\chi = \sum \delta_{\lambda_i}$:

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

For deterministic systems, $\mathbb{P}$ is an averaging over the energy level in the semiclassical limit.

Gaudin, Dyson, Mehta : for any $E \in (-2, 2)$ then ($\beta = 2$ for example)

$$\rho_k^{(N)}\left(E + \frac{u_1}{N\varrho(x)}, \dots, E + \frac{u_k}{N\varrho(x)}\right) \underset{N\to\infty}{\longrightarrow} \det_{k\times k} \frac{\sin(\pi(u_i - u_j))}{\pi(u_i - u_j)}.$$

Wigner matrix : symmetric, Hermitian (or symplectic), entries have variance $1/N$, some large moment is finite.

The Wigner-Dyson-Mehta conjecture. Correlation functions of symmetric Wigner matrices (resp. Hermitian, symplectic) converge to the limiting GOE (resp. GUE, GSE).

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 ε and $s = N^{-1+\varepsilon}$,

$$\lim_{N\to\infty} \frac{1}{\varrho(E)^k} \int_E^{E+s} \frac{\mathrm{d}x}{s} \int \mathrm{d}\mathbf{v} F(\mathbf{v}) \rho_k^{(N)} \left(x + \frac{\mathbf{v}}{N\varrho(E)} \right) \mathrm{d}\mathbf{v} = \int \mathrm{d}\mathbf{v} F(\mathbf{v}) \rho_k^{(\mathrm{GOE})} (\mathbf{v})$$

Johansson (2001)	Hermitian class, fixed E, Gaussian divisible entries
Erdős Schlein Péché Ramirez Yau (2009)	Hermitian class, fixed E Entries with density
Tao Vu (2009)	Hermitian class, fixed E Entries with 3rd moment=0
Erdős Schlein Yau (2010)	Any class, averaged E

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

Key input for all recent results : **rigidity of eigenvalues** (Erdős Schlein Yau) : $|\lambda_k - \gamma_k| \leq N^{-1+\varepsilon}$ in the bulk. Optimal rigidity ?

Related developments : gaps universality by Erdős Yau (2012).

The gaps are much more stable statistics than the fixed energy ones :

$$\langle \lambda_i, \lambda_j \rangle \sim \log \frac{N}{1+|i-j|}, \text{ almost crystal. } \langle \lambda_{i+1}-\lambda_i \lambda_{j+1}-\lambda_j \rangle \sim \frac{1}{1+|i-j|^2}.$$

Theorem. 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, $\mathrm{d}H_t = \frac{\mathrm{d}B_t}{\sqrt{N}} - \frac{1}{2}H_t\mathrm{d}t$) is an essential interpolation tool, as in the Erdős Schlein Yau approach to universality, summarized as :

$$\begin{array}{ccc} & & H_0 \\ & & \updownarrow \\ \widetilde{H}_0 & \overset{(\mathrm{DBM})}{\longrightarrow} & \widetilde{H}_t \end{array}$$

$\overset{(\mathrm{DBM})}{\longrightarrow}$: for $t = N^{-1+\varepsilon}$, the eigenvaues of $\widetilde{H}_t$ satisfy averaged universality.

$\updownarrow$: Density argument. For any $t \ll 1$, there exists $\widetilde{H}_0$ s.t. the resolvents of H_0 and $\widetilde{H}_t$ have the same statistics on the microscopic scale.

What makes the Hermitian universality easier ? The $\overset{(\mathrm{DBM})}{\longrightarrow}$ step is replaced by HCIZ formula : correlation functions of $\widetilde{H}_t$ are explicit only for $\beta = 2$.

A few facts about the proof of fixed energy universality.

(i) A game coupling 3 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 $\overset{(\mathrm{DBM})}{\longrightarrow}$ becomes the following.

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

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

$$\mathbb{E}\, \widetilde{F}(\boldsymbol{\lambda}_t, 0) = \mathbb{E}\, \widetilde{F}(\boldsymbol{\lambda}^{(\mathrm{GOE})}, 0).$$

First step : coupling 3 DBM. Let $\mathbf{x}(0)$ be the eigenvalues of $\widetilde{H}_0$ and $\mathbf{y}(0), \mathbf{z}(0)$ those of two indepndent GOE.

$$\mathrm{d}x_i/\mathrm{d}y_i/\mathrm{d}z_i = \sqrt{\frac{2}{N}}\mathrm{d}B_i(t) + \frac{1}{N}\left(\sum_{j\neq i}\frac{1}{x_i/y_i/z_i - x_j/y_j/z_j} - \frac{1}{2}x_i/y_i/z_i\right)\mathrm{d}t$$

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

$$\partial_t\delta_\ell(t) = \sum_{k\neq\ell}\mathcal{B}_{k\ell}(t)\left(\delta_k(t) - \delta_\ell(t)\right),$$

where $\mathcal{B}_{k\ell}(t) = \frac{1}{N(x_k(t)-x_\ell(t))(y_k(t)-y_\ell(t))} > 0$. By the de Giorgi-Nash-Moser method (+Caffarelli-Chan-Vasseur+Erdős-Yau), this PDE is Hölder-continuous for $t > N^{-1+\varepsilon}$, i.e. $\delta_\ell(t) = \delta_{\ell+1}(t) + \mathrm{O}(N^{-1+\varepsilon})$, i.e. gap universality.

This is not enough for fixed energy universality.

Second step : homogenization. The continuum-space analogue of our parabolic equation is

$$\partial_t f_t(x) = (\mathcal{K} f_t)(x) := \int_{-2}^{2} \frac{f_t(y) - f_t(x)}{(x-y)^2} \varrho(y) \mathrm{d}y.$$

$\mathcal{K}$ is some type of second order Hilbert transform.

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

$$\delta_\ell(t) = \left(e^{t\mathcal{K}} f_0\right)_\ell + \mathrm{O}(N^{-1+\varepsilon}).$$

Proof. Rigidity of the eigenvalues, optimal Wegner estimates (for level-repulsion), and the Hölder regularity of the discrete-space parabolic equation.

Third step : the continuous-space kernel.

1. For the translation invariant equation

$$\partial_t g_t(x) = \int_{\mathbb{R}} \frac{g_t(y) - g_t(x)}{(x-y)^2} \mathrm{d}y,$$

   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 curvture cannot be neglected. Fortunately, $\mathcal{K}$ can be fully diagonalized and ($x = 2\cos\theta$, $y = 2\cos\phi$)

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

   Called the Mehler kernel by Biane in free probability context, never considered as a second-order Hilbert transform fundamental solution.

3. Explicit kernels can be obtained for all Bernstein-Szego measures,

$$\varrho(x) = \frac{c_{\alpha,\beta}(1-x^2)^{1/2}}{(\alpha^2 + (1-\beta^2)) + 2\alpha(1+\beta)x + 4\beta x^2}.$$

Fourth step : microscopic from mesoscopic. Homogenization yields

$$\delta_\ell(t) = \int k_t(x,y) f_0(y)\varrho(y)\mathrm{d}y + \mathrm{O}(N^{-1+\varepsilon})$$

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

$$N x_\ell(t) = N y_\ell(t) - \Psi_t(\mathbf{y}_0) + \Psi_t(\mathbf{x}_0),$$

where $\Psi_t(\mathbf{x}_0) = \sum h(N^\tau(x_i(0)-E))$ for some smooth h. We wanted to prove

$$\mathbb{E}\,\widetilde{F}(\mathbf{x}_t, 0) = \mathbb{E}\,\widetilde{F}(\mathbf{z}_t, 0) + \mathrm{o}(1).$$

We reduced it to

$$\mathbb{E}\,\widetilde{F}(\mathbf{y}_t, -\Psi_t(\mathbf{y}_0)+\Psi_t(\mathbf{x}_0)) = \mathbb{E}\,\widetilde{F}(\mathbf{y}_t, \Psi_t(\mathbf{y}_0)+\Psi_t(\mathbf{z}_0)) + \mathrm{o}(1).$$

where $\Psi_t(\mathbf{y}_0)$, $\Psi_t(\mathbf{x}_0)$ and $\Psi_t(\mathbf{z}_0)$ are mesoscopic observables and independent.

Fifth step and conclusion : CLT for GOE beyond the natural scale. Do $\Psi_t(\mathbf{x}_0)$ and $\Psi_t(\mathbf{y}_0)$ have the same distribution ? No, their variance depend on their fourth moment.

A stronger result holds : $\mathbb{E}\,\widetilde{F}(\mathbf{y}_t, -\Psi_t(\mathbf{y}_0)+c)$ does not depend on the constant c.

We know that $\mathbb{E}\,\widetilde{F}(\mathbf{y}_t, -\Psi_t(\mathbf{y}_0)+\Psi_t(\mathbf{z}_0)+c) = \mathbb{E}\,\widetilde{F}(\mathbf{y}_t, -\Psi_t(\mathbf{y}_0)+\Psi_t(\mathbf{z}_0))$ for all c (why ?).

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

The proof uses algebraic ideas of Johansson and rigidity of β-ensembles.

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.

What is the optimal rigidity of eigenvalues ?

Theorem (Gustavsson, O'Rourke). Let $\boldsymbol{\lambda}$ be the ordered eigenvalues of a Gaussian ensemble, k_0 a bulk index and $k_{i+1} \sim k_i + N^{\theta_i}$, $0 < \theta_i < 1$. Then the nornalized eigenvalues fluctuations

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

converge to a Gaussian vector with vovariance

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

In particlar, $\lambda_i - \gamma_i$ has fluctuations $\frac{\sqrt{\log N}}{N}$.

Proof : determinantal point processes a la Coston-Lebowitz (GUE) + decimation relations (GOS, GSE).

Theorem. Same log-correlated Gaussian limit for any Wigner matrix.

Proof. By homogenization we have

$$\frac{N(x_\ell(t)-\gamma_\ell)}{\sqrt{\log N}} = \frac{N(y_\ell(t)-\gamma_\ell)}{\sqrt{\log N}} + \frac{\Psi_t(\mathbf{y}(0))}{\sqrt{\log N}} - \frac{\Psi_t(\mathbf{x}(0))}{\sqrt{\log N}}.$$

The fluctuations of $\Psi_t(\mathbf{y}(0))$ are of order $\sqrt{\tau \log N}$. The fluctuations of $\Psi_t(\mathbf{x}(0))$ are of the same order $\sqrt{\tau \log N}$. Take arbitrarily small τ and the result follows.

Other eigenvalues possible applications of homogenization of DBM :

1. Largest gap amongst bulk eigenvalues of Wigner matrices is universal.
2. Extreme deviation from typical location is universal.

Unexpected applications for eigenvectors.

The Dyson vector flow

Coupled eigenvalues/eigenvectors dynamics when the entrie of H are Brownian motions :

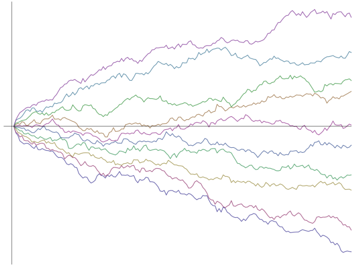
$$\mathrm{d}\lambda_k = \frac{\mathrm{d}B_{kk}}{\sqrt{N}} + \left(\frac{1}{N}\sum_{\ell\neq k}\frac{1}{\lambda_k-\lambda_\ell}\right)\mathrm{d}t$$

$$\mathrm{d}u_k = \frac{1}{\sqrt{N}}\sum_{\ell\neq k}\frac{\mathrm{d}B_{k\ell}}{\lambda_k-\lambda_\ell}u_\ell - \frac{1}{2N}\sum_{\ell\neq k}\frac{\mathrm{d}t}{(\lambda_k-\lambda_\ell)^2}u_k$$

Let $c_{k\ell}=\frac{1}{N}\frac{1}{(\lambda_k-\lambda_\ell)^2}$. If all $c_{k\ell}$'s were equal, $U=(u_1,\dots,u_N)$ would be the Brownian motion on the unitary group.

Such eigenvector flows were discovered by Norris, Rogers, Williams (Brownian motion on GL_N), Bru (real Wishart), Anderson, Guionnet, Zeitouni (symmetric and Hermitian).

A random walk in a dynamic random environment

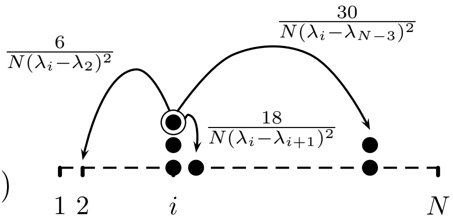


Definition of the (real) eigenvector moment flow.

The eigenvalues trajectory is a parameter ($c_{i,j}(t) = \frac{1}{N}\frac{1}{(\lambda_i(t)-\lambda_j(t))^2}$).

Configuration $\boldsymbol{\eta}$ of n points on $[\![1,N]\!]$. Number of particles at x : η_x.
Configuration obtained by moving a particle from i to j : $\boldsymbol{\eta}^{ij}$.
Dynamics given by $\partial_t f = \mathscr{B}(t) f$ where

$$\mathscr{B}(t)f(\boldsymbol{\eta}) = \sum_{i\neq j} c_{ij}(t) 2\eta_i(1+2\eta_j)\left(f(\boldsymbol{\eta}^{i,j}) - f(\boldsymbol{\eta})\right)$$



Properties of the eigenvector moment flow

Let $z_k = \sqrt{N}\langle \mathbf{q}, u_k\rangle$, random and time dependent. For a configuration $\boldsymbol{\eta}$ with j_k points at i_k, let

$$f_{t,\boldsymbol{\lambda}}(\boldsymbol{\eta}) = \mathbb{E}\left(\prod_k z_{i_k}^{2j_k} \mid \boldsymbol{\lambda}\right) / \, \mathbb{E}\left(\prod_k \mathscr{N}_{i_k}^{2j_k}\right).$$

Fact 1 : $\partial_t f_{t,\boldsymbol{\lambda}}(\boldsymbol{\eta}) = \mathscr{B}(t) f_{t,\boldsymbol{\lambda}}(\boldsymbol{\eta})$.

QUE+Normality of the eigenvectors is equivalent to fast relaxation to equilibrium of the eigenvector moment flow.
This PDE analysis is made possible thanks to an explicit reversible measure for $\mathscr{B}$

Fact 2 :

- GOE : $\pi(\boldsymbol{\eta}) = \prod_{x=1}^{N} \phi(\eta_x)$ where $\phi(k) = \prod_{i=1}^{k}\left(1 - \frac{1}{2k}\right)$
- GUE : π is uniform

Perturbative analysis in non-perturbative regime.

Let $(M_N)_{N\geq 0}$ be deterministic with eigenvalues satisfying the local semicircle law, eigenvectors $(e_k)_k$. What do the eigenvectors $(u_k(t))_k$ of $M_N + \sqrt{t}\,\mathrm{GOE}$ look like ?

If $1/N \ll t \ll 1$, neither perturbative regime nor free-probability regime.

Theorem. The coordinates $(\langle u_k(t), e_j\rangle)_j$ are independent Gaussian with variance

$$\mathbb{E}\left(\langle u_k(t), e_j\rangle^2\right) \sim \frac{1}{(Nt)^2 + (\gamma_k - \gamma_j)^2}$$

Proof : the eigenvector moment flow describes the evolution of the variances : $f_t(k) = \mathbb{E}(|\langle u_k(t), e_j\rangle|^2 \mid \boldsymbol{\lambda}(\cdot))$ satisfies

$$\partial_t f_t(k) = \frac{1}{N} \sum_{j\neq k} \frac{f_t(j) - f_t(k)}{(\lambda_j(t) - \lambda_k(t))^2}.$$

Then use homogenization for DBM.