# Introduction to Random Matrix Theory

**Masoud Khalkhali**

**University of Western Ontario, Canada**

**Baby Steps Beyond the Horizon**

Banach Center, Sept 2022, Poland

# What is random matrix theory (RMT)?

- A random matrix is a matrix valued random variable.
- A matrix ensemble is a set of matrices + a probablity measure.
- RMT = Randomized linear algebra.

# What is random matrix theory (RMT)?

- A random matrix is a matrix valued random variable.
- A matrix ensemble is a set of matrices + a probablity measure.
- RMT = Randomized linear algebra.
- Linear algebra becomes probabilistic. Sample question: How the eigenvalues or eigenvectors of a random matrix are distributed?

# What is random matrix theory (RMT)?

- A random matrix is a matrix valued random variable.
- A matrix ensemble is a set of matrices + a probablity measure.
- RMT = Randomized linear algebra.
- Linear algebra becomes probabilistic. Sample question: How the eigenvalues or eigenvectors of a random matrix are distributed?
- Usually interested in: Large $N$ limits of eigenvalue distributions and their correlations when $N$= size of our matrices goes to infinity, and universality classes. Compare with law of large numbers and the central limit theorem in standard probability theory.

# Origins in nuclear physics (Eugen Wigner)

- Eigenvalue problem of quantum mechanics:

$$H\psi = E\psi$$

Hamiltonian $H$ is a Hermitian operator, $E$ = energy. Too difficult to solve!

# Origins in nuclear physics (Eugen Wigner)

- Eigenvalue problem of quantum mechanics:

$$H\psi = E\psi$$

  Hamiltonian $H$ is a Hermitian operator, $E =$ energy. Too difficult to solve!

- Wigner's idea (1950's): no good information on $H$ for the nuclei of heavy atoms, so assume $H = (h_{ij})$ is a random Hermitian $N \times N$ matrix, with $h_{ij}$ centered iid random variables for $i \leq j$ (Wigner's ensemble).

# Origins in nuclear physics (Eugen Wigner)

- Eigenvalue problem of quantum mechanics:

$$H\psi = E\psi$$

  Hamiltonian $H$ is a Hermitian operator, $E =$ energy. Too difficult to solve!

- Wigner's idea (1950's): no good information on $H$ for the nuclei of heavy atoms, so assume $H = (h_{ij})$ is a random Hermitian $N \times N$ matrix, with $h_{ij}$ centered iid random variables for $i \leq j$ (Wigner's ensemble).

- Simplest Wigner ensemble: Gaussian unitary ensemble (GUE) with probabiity density

$$P(H) = \frac{1}{Z_N} e^{-N \mathrm{Tr} \frac{H^2}{2}}$$

# Origins in agriculture and statistics (Wishart)

- Wishart ensemble (1928): Consider the map $M_{p\times n}(\mathbb{R}) \to M_{p\times p}(\mathbb{R})$

  $$X \mapsto C = \frac{1}{n} X X^t$$

  where columns of $X$ are independent Gaussian vectors $\sim \mathcal{N}(0, V)$. The set of all $C$ (covariance matrices) under the pushforward measure is the Wishart ensemble.

# Origins in agriculture and statistics (Wishart)

- Wishart ensemble (1928): Consider the map $M_{p\times n}(\mathbb{R}) \to M_{p\times p}(\mathbb{R})$

$$X \mapsto C = \frac{1}{n} X X^t$$

  where columns of $X$ are independent Gaussian vectors $\sim \mathcal{N}(0, V)$. The set of all $C$ (covariance matrices) under the pushforward measure is the Wishart ensemble.

- Let $n, p \to \infty$ with $\frac{p}{n} \to \lambda$. Marchenko-Pastur law gives the limiting eigenvalue distribution of $C$ (used in data analysis, machine learning, finance.)

# Warmup: Wigner's surmise, eigenvalue repulsion

- Let $A$ be a random real symmetric $2 \times 2$ matrix sampled from the probability distribution (GOE)

$$P(A) = \frac{1}{Z} e^{-\frac{1}{2}\mathrm{Tr}(A^2)}$$

- Find the eigenvalue spacing distribution of $A$, $s = \lambda_2 - \lambda_1$. This is given by

$$p(s) = \frac{1}{Z} \int e^{-\frac{1}{2}\mathrm{Tr}(A^2)} \delta(s - (\lambda_2 - \lambda_1)) dA$$

# Warmup: Wigner's surmise, eigenvalue repulsion

- Let $A$ be a random real symmetric $2 \times 2$ matrix sampled from the probability distribution (GOE)

$$P(A) = \frac{1}{Z} e^{-\frac{1}{2}\mathrm{Tr}(A^2)}$$

- Find the eigenvalue spacing distribution of $A$, $s = \lambda_2 - \lambda_1$. This is given by

$$p(s) = \frac{1}{Z} \int e^{-\frac{1}{2}\mathrm{Tr}(A^2)} \delta(s - (\lambda_2 - \lambda_1)) dA$$

- A simple calculation shows (Wigner's surmise 1950's):

$$p(s) = \frac{s}{2} e^{-s^2/4}, \qquad s \geq 0$$

and $p(s) = 0$ for $s < 0$. This shows that eigenvalues are not independent (eigenvalue repulsion).

# Unitary invariant ensembles, eigenvalue repulsion

- Using Weyl integration formula, the integral

$$Z = \int_{\mathcal{H}_N} e^{-N\mathrm{Tr}(V(H))} dH,$$

can be reduced to integration over eigenvalues and gives the jpdf of eigenvalues

$$d\rho(\lambda_1, \cdots, \lambda_N) = \prod_{1 \le i < j \le N} |\lambda_j - \lambda_i|^2 \prod_{i=1}^{N} \left( e^{-NV(\lambda_i)} d\lambda_i \right)$$

- Vandermonde determinant indicates eigenvalue repulsion.

# Universality classes: Unitary Invariant and Wigner ensembles, Dyson 3-fold way (GUE, GOE, GSE)

- Two (almost exclusive) classes of ensembles on $\mathcal{H}_N$:
- (1) Unitary invariant ensembles, with $\mu = \frac{1}{Z} dH$,

$$Z = \int_{\mathcal{H}_N} e^{F(H)} dH,$$

  $F(H)$ is a unitary invariant function, $dH =$ Lebesgue measure on $\mathcal{H}_N$, and the unitary group $U_N$ acts by conjugation on $H_N$.
- (2) Wigner ensembles.

# Universality classes: Unitary Invariant and Wigner ensembles, Dyson 3-fold way (GUE, GOE, GSE)

- Two (almost exclusive) classes of ensembles on $\mathcal{H}_N$:
- (1) Unitary invariant ensembles, with $\mu = \frac{1}{Z} dH$,

$$Z = \int_{\mathcal{H}_N} e^{F(H)} dH,$$

  $F(H)$ is a unitary invariant function, $dH =$ Lebesgue measure on $\mathcal{H}_N$, and the unitary group $U_N$ acts by conjugation on $H_N$.
- (2) Wigner ensembles.

$$(1) \cap (2) = \text{Gaussian unitary ensemble (GUE)},$$

with $\mu = \frac{1}{Z} e^{-N\mathrm{Tr}(H^2)} dH$, and $Z = \int_{\mathcal{H}_N} e^{-N\mathrm{Tr}(H^2)} dH$.

# Eigenvalue distributions

- Eigenvalue density function: a probability distribution valued random variable:

$$\mu_N(H) = \frac{1}{N}\sum_{i=1}^{N} \delta(x - \lambda_i(H))$$

# Eigenvalue distributions

- Eigenvalue density function: a probability distribution valued random variable:

$$\mu_N(H) = \frac{1}{N}\sum_{i=1}^{N}\delta(x - \lambda_i(H))$$

- Mean density function:

$$\rho_N(x) = \langle \mu_N \rangle = \int_{\mathcal{H}_N} \mu_N(H)P(H)dH$$

# Eigenvalue distributions

- Eigenvalue density function: a probability distribution valued random variable:

$$\mu_N(H) = \frac{1}{N}\sum_{i=1}^{N}\delta(x-\lambda_i(H))$$

- Mean density function:

$$\rho_N(x) = \langle \mu_N \rangle = \int_{\mathcal{H}_N} \mu_N(H)P(H)dH$$

- Large $N$ limit

$$\rho = \lim \rho_N, \qquad N \to \infty$$

- In practice one computes (scaling) limits of tracial moments

$$\lim \langle \frac{1}{N}\mathrm{tr}(A^k)\rangle \quad N \to \infty$$

# Wigner semicircle law for Wigner matrices

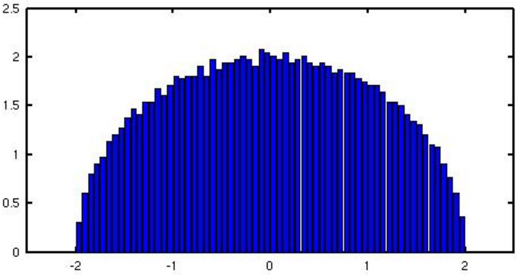


Figure: Histogram, sampled from a 3000x 3000 GUE matrix

- Wigner semicircle law (1950's), universality for Wigner ensembles:

$$\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2}, \qquad |x| \le 2$$

# Proof of semicircle law

- Moments of GUE are Gaussian integrals and easy to compute via Wick's theorem.

$$\langle h_{ij} h_{kl} \rangle = \frac{1}{N} \delta_{il} \delta_{jk}$$

- Wick's theorem:

$$\langle \frac{1}{N} \mathrm{Tr}(H^{2k}) \rangle = \sum \prod \langle h_{ij} h_{kl} \rangle$$

  where the summation is over the set of indices $i_1, i_2, \ldots, i_{2k} = 1, \ldots, 2k$ and the product is over the pairings of these indices. The odd moments are zero by symmetry.

- Feynman's theorem: summation can be reduced to summing over gluings of a 2k gon

# Genus expansion

- Using polygon gluings, we obtain

$$\langle \frac{1}{N} \mathrm{Tr}(H^{2k}) \rangle = \sum_{\sigma} N^{v(\sigma)-k-1} = \sum_{g=0}^{\infty} \varepsilon_g(k) N^{-2g}$$

  where the first sum is over the set of 1-face maps and $g$ is the genus of corresponding oriented closed surface.

- It can be shown that the leading term

$$\varepsilon_0(k) = \text{ number of planar gluings of a 2k gon}$$

  is equal to the number of non-crossing gluings which is equal to the k-th Catalan number

$$\varepsilon_0(k) = C_k = \frac{1}{k+1} \binom{2k}{k}$$

# Large $N$ limit

- From genus expansion it follows that

$$\lim_{N\to\infty}\langle\frac{1}{N}\mathrm{Tr}(H^{2k})\rangle = \varepsilon_0(k) = \frac{1}{k+1}\binom{2k}{k}$$

- The moments of the semincircle distribution

$$\langle x^{2k}\rangle_\rho = \frac{1}{2\pi}\int_{-2}^{2} x^{2k}\sqrt{4-x^2}dx = \frac{1}{k+1}\binom{2k}{k}$$

- It follows that

$$\lim_{N\to\infty}\langle\frac{1}{N}\mathrm{Tr}(H^{2k})\rangle = \langle x^{2k}\rangle_\rho$$

And this is the simplest version of the Wigner semicircle law.

# Genus expansion in general: summing over discrete surfaces ('t Hooft, Brezin-Itzykson-Parisi-Zuber)

- Fix a polynomial $V(x) = \sum \frac{t_k}{k} x^k$. Consider the formal matrix integral
$$Z_N = \int_{\mathcal{H}_N} e^{-N\mathrm{Tr}(V(H))} dH,$$

## Genus expansion in general: summing over discrete surfaces ('t Hooft, Brezin-Itzykson-Parisi-Zuber)

- Fix a polynomial $V(x) = \sum \frac{t_k}{k} x^k$. Consider the formal matrix integral
$$Z_N = \int_{\mathcal{H}_N} e^{-N\mathrm{Tr}(V(H))} dH,$$
- Topological expansion of $F_N = \log Z_N$
$$F_N = \sum_{g \geq 0} (N)^{2-2g} F_g\,, \qquad F_g = \sum_{[\mathcal{M}] \in \mathbb{M}_{\emptyset}^{g}} \mathrm{weight}(\mathcal{M})$$

where $\mathbb{M}_{\emptyset}^{g}$ = set of isomorphism classes of the Feynman weighted connected closed maps of genus $g$.

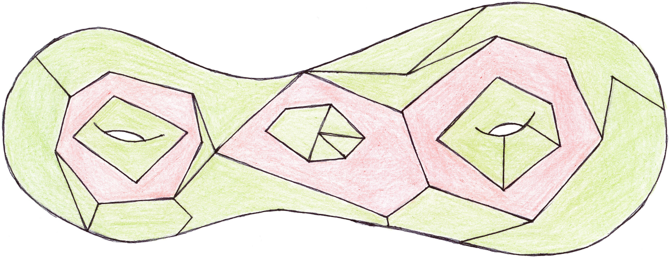

Figure: A polygonalization of a genus 2 surface (S. Azarfar and M K. Random finite noncommutative geometries and topological recursion, arXiv:1906.09362)

Genus expansion leads to a quick proof of the Wigner law, links with geometry of moduli spaces of curves, topological recursion (Eynard-Orantin), 2d gravity, recursion formula for volumes of moduli spaces of Riemann surfaces (Mirzakhani recursion).