

Lectures on Random Matrices (Spring 2025)

Lecture 10: Dyson Brownian Motion

Leonid Petrov

Monday, March 24, 2025*

Abstract

This lecture begins the second half of the course after the Spring Break, and here we start with the Dyson Brownian Motion. In the later lectures we explore random growth models and their universal scaling limits which are closely connected to random matrix theory.

1 Motivations

1.1 Why introduce time?

Our previous lectures dealt with static matrix ensembles (e.g., GUE, GOE, and so on). However, there are both *physical* and *mathematical* reasons to study a dynamical model for random matrices. For instance:

1. In physics, one often interprets random matrices as Hamiltonians of quantum systems. It is natural to let these Hamiltonians vary in time and to describe how spectra evolve.
2. Such time-dependent models are vital for studying *universality results* in random matrix theory. Rigorous proofs of local eigenvalue correlations often involve coupling or evolving an ensemble toward (or away from) a known reference ensemble.
3. Dynamical extensions yield intriguing connections to 2D statistical mechanics, representation theory, and Markov chain interpretations such as *nonintersecting path ensembles*.

1.2 Simple example: 1×1 case

When $N = 1$, an $N \times N$ Hermitian matrix is just a single real entry. Thus GUE/GOE/GSE distributions each reduce to a real Gaussian variable with mean 0 and variance 1. If we allow *time*, the natural time evolution is standard *Brownian motion* $B(t)$ on \mathbb{R} .

Recall that a standard one-dimensional Brownian motion $B(t)$ is a continuous stochastic process with the following key properties:

1. **Continuity:** $t \mapsto B(t)$ is almost surely continuous.

*[Course webpage](#) • [Live simulations](#) • [TeX Source](#) • Updated at 05:45, Monday 24th March, 2025

2. **Independent increments:** For any $0 \leq s < t$, the increment $B(t) - B(s)$ is independent of the past $\{B(u) : 0 \leq u \leq s\}$.
3. **Gaussian increments:** $B(t) - B(s)$ is normally distributed with mean 0 and variance $t - s$; that is,

$$B(t) - B(s) \sim \mathcal{N}(0, t - s).$$

Thus, if the process starts at $B(0) = a$, then for any fixed $t > 0$,

$$B(t) \sim \mathcal{N}(a, t).$$

Our goal is to generalize this to the case of *matrix-valued* Brownian motion and, ultimately, to see how the *eigenvalues* of such a matrix evolve.

2 Matrix Brownian motion and its eigenvalues

2.1 Definition

Let $X(t)$ be an $N \times N$ matrix whose entries are i.i.d. real/complex Brownian motions (depending on $\beta = 1, 2$). For instance:

- If $\beta = 1$: $X(t)$ has entries that are i.i.d. real Brownian motions.
- If $\beta = 2$: $X(t)$ has entries that are i.i.d. complex Brownian motions (independent real and imaginary parts).

Since $X(t)$ may not be Hermitian, define

$$\mathcal{M}(t) = \frac{1}{\sqrt{2}}(X(t) + X^\dagger(t)).$$

Here $X^\dagger(t)$ is the conjugate transpose. Then $\mathcal{M}(t)$ is an *Hermitian* matrix (or real symmetric for $\beta = 1$).

Lemma 2.1. *If $\mathcal{M}(0) = A$ is a fixed deterministic matrix, then $\mathcal{M}(t)$ at time t is distributed as*

$$A + \sqrt{t} G_\beta,$$

where G_β is a random Hermitian matrix from the Gaussian ensemble with $\beta = 1$ or 2.

Sketch of proof. Straightforward observation. □

For the one-dimensional case, notice that $a + \sqrt{t}Z$, where $Z \sim \mathcal{N}(0, 1)$, is a Gaussian random variable with mean a and variance t , and every such Gaussian variable can be represented in this form.

2.2 Eigenvalues as Markov process

We now focus on $\lambda_i(t)$, the (ordered) eigenvalues of $\mathcal{M}(t)$. Denote

$$\lambda(t) = (\lambda_1(t) \geq \dots \geq \lambda_N(t)).$$

Theorem 2.2. *As t varies, the process $\lambda(t)$ is a continuous-time Markov process in \mathbb{R}^N .*

Sketch of proof. Assume $\beta = 2$, the case $\beta = 1$ is similar. We need to show that $\lambda(t)$ depends on its future and past only through its instantaneous value. Using the independent increment property of $X(t)$, consider times $0 < u < t$. We have

$$\mathcal{M}(t) = \mathcal{M}(u) + (\mathcal{M}(t) - \mathcal{M}(u)).$$

Since $\mathcal{M}(u)$ diagonalizes to $\text{diag}(\lambda_1(u), \dots, \lambda_N(u))$ by some unitary U_u , we can write

$$U_u^\dagger \mathcal{M}(t) U_u = \text{diag}(\lambda_1(u), \dots, \lambda_N(u)) + U_u^\dagger (\mathcal{M}(t) - \mathcal{M}(u)) U_u.$$

The second term again has i.i.d. random entries (due to unitary invariance of GUE), independent of $\mathcal{M}(s)$ for $s \leq u$. Therefore, conditioned on $\mathcal{M}(s)$, $s \leq u$, the dependence only comes through $\lambda(u)$, and the eigenvalues $\lambda_i(s)$ for $s \geq u$ follow the same dynamics. This proves the Markov property. \square

3 Dyson Brownian Motion

We now describe the stochastic differential equation (SDE) for $\lambda(t)$ explicitly, following the classical result due to Dyson [Dys62]. Let us first briefly discuss what is an SDE.

3.1 Stochastic differential equations - an informal introduction

In order to describe the eigenvalues of a time-dependent Hermitian matrix, we rely on *stochastic differential equations* (SDEs). These are differential equations where one or more of the terms involve *random noise*. For simplicity, we start with the one-dimensional setup and later extend it to systems of equations such as those arising in Dyson Brownian Motion.

In an ordinary differential equation (ODE), a function $x(t)$ evolves according to a deterministic rule of the form

$$\frac{dx(t)}{dt} = b(x(t)),$$

where $b(\cdot)$ is a deterministic function called the *drift*. If one imposes an initial condition $x(0) = x_0$, then classical theorems guarantee that, under mild regularity assumptions, a unique solution exists for all $t \geq 0$.

An SDE generalizes this setup by adding a *stochastic (or noise) term* to the right-hand side. Concretely, suppose $W(t)$ is a standard one-dimensional Brownian motion. Then the simplest SDE has the form

$$dx(t) = \sigma dW(t),$$

where σ is a nonnegative constant. This equation may be formally interpreted as

$$\frac{dx(t)}{dt} = \sigma \frac{dW(t)}{dt},$$

but it should be emphasized that $\frac{dW}{dt}$ does not exist in the usual sense of classical calculus (Brownian motion is nowhere differentiable almost surely). Instead, one interprets the equation via the *Itô integral*

$$x(t) = x(0) + \int_0^t \sigma dW(s).$$

This integral is defined carefully through a limit of sums involving the increments $W(t_{k+1}) - W(t_k)$, yielding an *almost sure* continuous stochastic process $t \mapsto x(t)$.

More generally, one allows both *drift* and *diffusion* terms:

$$dx(t) = b(x(t)) dt + \sigma(x(t)) dW(t). \quad (3.1)$$

Here,

- $b(\cdot)$ is the *drift coefficient*, capturing deterministic motion;
- $\sigma(\cdot)$ is the *diffusion coefficient*, encoding how strongly the process is randomized by Brownian motion.

Under suitable Lipschitz and growth conditions on b and σ , one can show *existence and pathwise uniqueness* of strong solutions to (3.1). Concretely, this means there is almost surely a unique process $x(t)$ satisfying (3.1) for each realization of the Brownian motion $W(t)$. One constructs such a solution, for example, by an iterative limit of approximations. The simplest discrete-time approximation, analogous to Euler's method for ordinary differential equations. Over a small time step Δt , one approximates

$$x_{n+1} = x_n + b(x_n) \Delta t + \sigma(x_n) (W(t_{n+1}) - W(t_n)).$$

This scheme converges to the true solution pathwise under standard Lipschitz conditions on b and σ .

A major utility of SDEs is in performing *Itô calculus*. Suppose $x(t)$ solves the SDE (3.1) and let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a sufficiently smooth function. One might try to apply the usual chain rule to $f(x(t))$, but must account for the extra "noise" term. The correct extension is the *Itô formula*:

$$df(x(t)) = \frac{\partial f}{\partial x}(x(t)) dx(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x(t)) (dW(t))^2,$$

where $(dW(t))^2$ is interpreted as dt in a formal sense. Substituting (3.1) yields:

$$df(x(t)) = b(x(t)) \frac{\partial f}{\partial x}(x(t)) dt + \sigma(x(t)) \frac{\partial f}{\partial x}(x(t)) dW(t) + \frac{1}{2} \sigma^2(x(t)) \frac{\partial^2 f}{\partial x^2}(x(t)) dt.$$

This identity is an indispensable tool for analyzing stochastic processes, both in theoretical and applied contexts.

To handle matrix-valued processes, one must consider multi-dimensional (or matrix-dimensional) analogs of (3.1). For instance, if $X(t) \in \mathbb{R}^n$ is an n -dimensional stochastic process, the SDE becomes

$$dX(t) = b(X(t)) dt + \sigma(X(t)) dW(t),$$

where $b(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\sigma(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$. Here $W(t)$ is an n -dimensional Brownian motion, and the product $\sigma(X(t)) dW(t)$ is understood as a matrix-vector multiplication in each small time increment. Existence, uniqueness, and Itô's formula all generalize naturally under suitable regularity assumptions.

Summary Although SDEs can be introduced rigorously via measure-theoretic tools, the above *informal* derivation and discussion provide a workable framework for many typical computations. The key points are:

- Brownian motion's roughness prevents classical differential calculus, so new techniques (Itô integrals) are needed.
- The Itô formula extends the classical chain rule by adding a second-order correction term.
- Existence and uniqueness theorems ensure that SDEs define well-posed dynamical systems in a stochastic setting.
- Extending to matrix-valued (or multi-dimensional) settings is conceptually straightforward but requires careful linear algebraic bookkeeping and additional regularity arguments.

Equipped with these ideas, we can rigorously address how the eigenvalues of a random matrix evolve over continuous time, culminating in the Dyson Brownian Motion description of Hermitian ensembles.

3.2 Heuristic derivation of the SDE for the Dyson Brownian Motion

Let $\mathcal{M}(t)$ be an $N \times N$ Hermitian matrix evolving as $\mathcal{M}(0) = A$ plus i.i.d. Gaussian increments in time. Denote its ordered eigenvalues at time t by

$$\lambda_1(t) \geq \dots \geq \lambda_N(t).$$

We aim to find an SDE for $\lambda_i(t)$.

For a small increment Δt , we have

$$\mathcal{M}(t + \Delta t) = \mathcal{M}(t) + \Delta \mathcal{M},$$

where the entries of $\Delta \mathcal{M}$ are (approximately) independent $\mathcal{N}(0, \Delta t)$ random variables (real or complex). Suppose we diagonalize $\mathcal{M}(t) = U \text{diag}(\lambda_1(t), \dots, \lambda_N(t)) U^\dagger$.

Sketch of the computation. Search for the i -th eigenvalue of the form

$$\lambda = \lambda_i(T) + \Delta \lambda \quad [\text{expect } \Delta \lambda \approx O(\sqrt{\Delta t})].$$

We want to solve

$$\det \begin{pmatrix} \lambda_1(T) - \lambda_i(T) + B_1(\Delta t) - \Delta \lambda & \cdots & -\Delta \lambda B_{i1}(\Delta t) \\ \vdots & \ddots & \vdots \\ -\Delta \lambda B_{1i}(\Delta t) & \cdots & \lambda_N(T) - \lambda_i(T) + B_N(\Delta t) - \Delta \lambda \end{pmatrix} = 0.$$

In this matrix only $N - 1$ diagonal elements — excluding the (i, i) entry — are of finite size; the remaining off-diagonal elements are small. We have

$$\det = \prod_{m=1}^N [\lambda_m(T) - \lambda_i(T) + B_m(\Delta t) - \Delta \lambda] - \sum_{j \neq i} \left(\prod_{\substack{m \neq j \\ m \neq i}} [\lambda_m(T) - \lambda_i(T) + B_m(\Delta t) - \Delta \lambda] \right) \frac{1}{2} B_{ji}^2(\Delta t) + o(\Delta t).$$

Here, the first product (diagonal part) involves all N diagonal-like terms, and the sum over $j \neq i$ ($N - 1$ diagonal elements) accounts for corrections from the off-diagonal blocks. Higher-order terms are $o(\Delta t)$.

Divide by $\prod_{m \neq i} [\lambda_m(T) - \lambda_i(T) + B_m(\Delta t) - \Delta \lambda]$ to obtain

$$o(\Delta t) = -\Delta \lambda + B_i(\Delta t) - \sum_{j \neq i} \frac{\frac{1}{2} B_{ji}^2(\Delta t)}{\lambda_j(T) - \lambda_i(T) + B_j(\Delta t) - \Delta \lambda}.$$

Hence, to leading order in small Δt , we can ignore $\Delta \lambda$ in the denominator, replace $B_{ji}^2(\Delta t)$ by Δt as its expectation, ignore the random correction (as in Itô calculus), and obtain the desired SDE. We do not go into further details here, but the details are abundant in the literature, including the original work of Dyson [Dys62]. \square

Definition 3.1 (Dyson Brownian Motion). Fix $\beta > 0$ and initial data $(\lambda_1(0) \geq \dots \geq \lambda_N(0))$. The *Dyson Brownian Motion* is the unique strong solution to the system of SDEs

$$d\lambda_i(t) = \frac{\beta}{2} \sum_{j \neq i} \frac{dt}{\lambda_i(t) - \lambda_j(t)} + dW_i(t), \quad i = 1, \dots, N, \quad (3.2)$$

with the $W_i(t)$ being independent real standard Brownian motions. For $\beta = 1, 2, 4$, this coincides with the eigenvalue process of matrix Brownian motion (GOE, GUE, GSE).

Remark 3.2. Equation (3.2) succinctly captures the key idea that the eigenvalues repel each other. Note the singular drift term $\frac{1}{\lambda_i - \lambda_j}$ which pushes λ_i away from collisions with λ_j . This repulsion is so strong (for all $\beta > 0$) that eigenvalues will not cross (and thus remain ordered) with probability one.

4 Mapping the $G\beta E$ densities with the Dyson Brownian Motion

If the Dyson Brownian motion starts from zero¹ $\lambda_1(0) = \dots = \lambda_N(0) = 0$, we expect that at time t , the density of eigenvalues is $G\beta E$,

$$\propto \prod_{i < j} |\lambda_i - \lambda_j|^\beta \exp \left\{ -\frac{1}{2t} \sum_i \lambda_i^2 \right\}.$$

This is evident for $\beta = 1, 2, 4$, when we have a matrix model, but not so much for other β . For other β , we would like to

- Make sense of the SDE and its solutions. We skip this part in the course.
- Make a computation checking that the above density is preserved under the SDE (3.2).

¹And then the particles immediately repel each other and stay ordered for the whole time.

For example, in the $N = 1$ case, $d\lambda = dW(t)$ is a Markov process and one wants to show that

$$p(t, \lambda) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{\lambda^2}{2t}\right)$$

is preserved in the evolution. To verify this, one computes the generator of the semigroup, which for Brownian motion is

$$\frac{1}{2} \frac{\partial^2}{\partial \lambda^2}.$$

One then checks that

$$\frac{\partial}{\partial t} p(t, \lambda) = \frac{1}{2} \frac{\partial^2}{\partial \lambda^2} p(t, \lambda).$$

This is a direct computation.

For larger N , one needs to write down the corresponding generator and check that the same type of equation is satisfied. See Problem [J.3](#).

5 Determinantal structure for $\beta = 2$

5.1 Transition density

Theorem 5.1 ($\beta = 2$ Dyson Brownian Motion Transition Probabilities). *For $\beta = 2$, let $\lambda(t) = (\lambda_1(t) \geq \dots \geq \lambda_N(t))$ follow Dyson Brownian Motion starting at $\lambda(0) = \mathbf{a} = (a_1 \geq \dots \geq a_N)$. Then for each fixed time $t > 0$,*

$$P(\lambda(t) = \mathbf{x} \mid \lambda(0) = \mathbf{a}) = N! \left(\frac{1}{\sqrt{2\pi t}}\right)^N \prod_{1 \leq i < j \leq N} \frac{x_i - x_j}{a_i - a_j} \det \left[\exp\left(-\frac{(x_i - a_j)^2}{2t}\right) \right]_{i,j=1}^N,$$

where $x_1 \geq \dots \geq x_N$.

Proof. Consider an $N \times N$ Hermitian matrix process $X(t)$ whose entries perform independent complex Brownian motions (so that $X(t)$ is distributed as $A + \sqrt{t}$ GUE at each fixed time, with $A = \text{diag}(a_1, \dots, a_N)$). Its eigenvalues $\lambda_1(t) \geq \dots \geq \lambda_N(t)$ evolve exactly according to the $\beta = 2$ Dyson Brownian Motion.

The density of X at time t , viewed as a random matrix, is proportional to

$$\exp\left(-\frac{1}{2t} \text{Tr}(X - A)^2\right).$$

If we replace A by $U A U^\dagger$ for any fixed unitary U , the law of X remains the same (this follows from the unitary invariance of the GUE). Thus the distribution of the eigenvalues of X is unchanged by such conjugation.

One writes

$$\int_{U(N)} \exp\left(-\frac{1}{2t} \text{Tr}(X - U A U^\dagger)^2\right) dU = (\text{const.}) \times [\text{HCIZ integral in the variables } (X, A)],$$

which by the Harish–Chandra–Itzykson–Zuber formula leads to a product of determinants and a factor that is precisely

$$\exp\left(-\frac{1}{2t}\sum_{i=1}^N x_i^2 - \frac{1}{2t}\sum_{i=1}^N a_i^2\right) \frac{\det\left[\exp\left(\frac{x_i a_j}{t}\right)\right]}{\prod_{i<j}(x_i - x_j)(a_i - a_j)},$$

where x_1, \dots, x_N are the eigenvalues of X .

To convert this matrix distribution into a distribution on eigenvalues alone, we multiply by the usual Vandermonde Jacobian $\prod_{i<j}(x_i - x_j)^2$ (which comes from integrating out the unitary degrees of freedom). This produces exactly

$$N! \left(\frac{1}{\sqrt{2\pi t}}\right)^N \prod_{i<j} \frac{x_i - x_j}{a_i - a_j} \det\left[\exp\left(-\frac{(x_i - a_j)^2}{2t}\right)\right].$$

Hence we obtain the stated transition probability for the Dyson Brownian Motion at $\beta = 2$. \square

Remark 5.2. The factor $N! \left(\frac{1}{\sqrt{2\pi t}}\right)^N$ arises naturally from normalizing the Gaussian increments and accounts for the ordering $\lambda_1 \geq \dots \geq \lambda_N$. The determinant and product factors encode the eigenvalue “repulsion” characteristic of $\beta = 2$ random matrices.

5.2 Determinantal correlations

Theorem 5.3 (Determinantal structure for $\beta = 2$ DBM). *Let $\{x_1(t), \dots, x_n(t)\}$ be the eigenvalues at time $t > 0$ of the $\beta = 2$ Dyson Brownian Motion started at initial locations (a_1, \dots, a_n) at time 0. Equivalently, these $x_i(t)$ are the eigenvalues of*

$$A + \sqrt{t} G,$$

where $A = \text{diag}(a_1, \dots, a_n)$ and G is a random Hermitian matrix from the GUE. Then the (random) point configuration $\{x_i(t)\}$ forms a determinantal point process with correlation kernel

$$K_t(x, y) = \frac{1}{(2\pi i)^2 t} \oint \oint \exp\left(\frac{w^2 - 2 y w}{2t}\right) \Big/ \exp\left(\frac{z^2 - 2 x z}{2t}\right) \prod_{i=1}^n \frac{w - a_i}{z - a_i} \frac{dw dz}{w - z}.$$

Here z goes around all the points a_1, \dots, a_n , and the w contour passes from $-i\infty$ to $i\infty$, to the right of the z contour.

- If $a_1 = \dots = a_n = 0$ and $t = 1$, this kernel reduces to the familiar correlation kernel of the GUE (see [Lecture 6](#)).
- One can use this formula to study the Baik–Ben Arous–Péché (BBP) [\[BBP05\]](#) phase transition for $\beta = 2$, which deals with finite rank perturbations of the GUE random matrix ensemble. Indeed, rank r perturbation corresponds to taking $a_1, \dots, a_r \neq 0$, and $a_{r+1} = \dots = a_n = 0$.

5.3 On the proof of determinantal structure

The idea of the proof of Theorem 5.3 is to represent the measure (the transition density) as a product of determinants. In general, if a measure is given as a product of determinants, there is a well-studied method (biorthogonal ensembles and, more generally, the Eynard–Mehta theorem) to compute the determinantal correlation kernel. We refer to [BR05], [Bor11] for a detailed exposition in the discrete case (which is arguably more transparent). The first step for the Dyson Brownian Motion is as follows.

Lemma 5.4 (Density representation). *Let $P_t(x \rightarrow y)$ be the transition probability kernel of standard Brownian motion,*

$$P_t(x \rightarrow y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right).$$

Then the density of the eigenvalues (x_1, \dots, x_N) of DBM started at (a_1, \dots, a_N) at time 0 admits the representation

$$\lim_{s \rightarrow \infty} \left(\frac{1}{Z}\right) \det\left[P_t(a_i \rightarrow x_j)\right]_{i,j=1}^N \det\left[P_s(x_i \rightarrow k-1)\right]_{i,k=1}^N. \quad (5.1)$$

Remark 5.5. This representation (5.1) is related to an alternative description of the $\beta = 2$ Dyson Brownian Motion as an ensemble of noncolliding Brownian motions (that is, independent Brownian motions, conditioned to never collide).

Proof of Lemma 5.4. The first determinant (as $s \rightarrow \infty$) matches the determinant we have in Theorem 5.1. It remains to analyze the second determinant

$$\det\left[P_s(x_j \rightarrow k-1)\right]_{j,k=1}^N = \det\left[\frac{1}{\sqrt{2\pi s}} \exp\left(-\frac{((k-1)-x_j)^2}{2s}\right)\right]_{j,k=1}^N.$$

We may ignore the factor $\frac{1}{\sqrt{2\pi s}}$ in each entry since it does not depend on x_j . Inside the exponential,

$$-\frac{((k-1)-x_j)^2}{2s} = -\frac{x_j^2}{2s} + \frac{x_j(k-1)}{s} - \frac{(k-1)^2}{2s}.$$

Thus, up to the factor $\exp\left(-\frac{(k-1)^2}{2s}\right)$ (which depends only on k and hence is independent of each x_j), we can factor out $\exp\left(-\frac{x_j^2}{2s}\right)$ from row j . Consequently, the nontrivial part of the determinant becomes

$$\det\left[e^{\frac{x_j(k-1)}{s}}\right]_{j,k=1}^N.$$

Recognize this as a Vandermonde-type determinant in the variables $e^{x_j/s}$. Indeed,

$$\det\left[e^{\frac{x_j(k-1)}{s}}\right]_{j,k=1}^N = \prod_{1 \leq i < j \leq N} \left(e^{\frac{x_i}{s}} - e^{\frac{x_j}{s}}\right).$$

As $s \rightarrow \infty$, we expand $e^{\frac{x_i}{s}} = 1 + \frac{x_i}{s} + O\left(\frac{1}{s^2}\right)$, so each difference $(e^{\frac{x_i}{s}} - e^{\frac{x_j}{s}}) \sim \frac{x_i - x_j}{s}$. Hence,

$$\prod_{1 \leq i < j \leq N} \left(e^{\frac{x_i}{s}} - e^{\frac{x_j}{s}}\right) \sim \frac{1}{s^{\frac{N(N-1)}{2}}} \prod_{1 \leq i < j \leq N} (x_i - x_j).$$

Combining all these factors and matching with the first determinant (as $s \rightarrow \infty$) verifies the claimed product form, up to overall constants that do not depend on the variables x_j . This completes the proof. \square

Then, the product of determinants idea (biorthogonal ensembles) applies to the density (5.1) before the limit $s \rightarrow \infty$, and simplifies after taking the limit. We omit the details here, see Problem J.5.

6 Derivation of the HCIZ integral

In this section, we give a self-contained derivation of the Harish–Chandra–Itzykson–Zuber (HCIZ) integral from first principles, in a form commonly used in Random Matrix Theory and particularly in the derivation of Dyson Brownian Motion transition densities.

6.1 Statement of the HCIZ formula

Let A and B be two $N \times N$ Hermitian matrices with (real) eigenvalues

$$\text{Spec}(A) = (a_1, \dots, a_N), \quad \text{Spec}(B) = (b_1, \dots, b_N).$$

We want to compute the integral

$$\mathcal{I}(A, B) := \int_{U(N)} \exp(\text{Tr}(A U B U^\dagger)) dU,$$

where $U(N)$ is the group of $N \times N$ unitary matrices equipped with its normalized Haar measure dU . The Harish–Chandra–Itzykson–Zuber formula states that

$$\int_{U(N)} e^{\text{Tr}(A U B U^\dagger)} dU = \left(\prod_{k=1}^{N-1} k! \right) \frac{\det[e^{a_i b_j}]_{i,j=1}^N}{\prod_{1 \leq i < j \leq N} (a_j - a_i) \prod_{1 \leq i < j \leq N} (b_j - b_i)},$$

up to conventions for the normalization of the Haar measure. Many references fix the normalization constant as above.

6.2 Reduction to the diagonal case

The integrand $\exp(\text{Tr}(A U B U^\dagger))$ depends on U only via conjugation. Exploiting the Haar measure's bi-invariance:

1. Diagonalize $A = V_A \text{diag}(a_1, \dots, a_N) V_A^\dagger$.
2. Diagonalize $B = V_B \text{diag}(b_1, \dots, b_N) V_B^\dagger$.
3. Notice

$$\text{Tr}(A U B U^\dagger) = \text{Tr}\left(\text{diag}(a) (V_A^\dagger U V_B) \text{diag}(b) (V_B^\dagger U^\dagger V_A)\right).$$

Setting $W = V_A^\dagger U V_B$ preserves the Haar measure. Thus

$$\int_{U(N)} e^{\text{Tr}(A U B U^\dagger)} dU = \int_{U(N)} e^{\text{Tr}(\text{diag}(a) W \text{diag}(b) W^\dagger)} dW.$$

Therefore, we may assume $A = \text{diag}(a)$ and $B = \text{diag}(b)$. In that case,

$$\text{Tr}(A U B U^\dagger) = \sum_{i,j=1}^N a_i b_j |U_{ij}|^2.$$

Hence

$$\int_{U(N)} \exp\left(\text{Tr}(A U B U^\dagger)\right) dU = \int_{U(N)} \exp\left(\sum_{i,j=1}^N a_i b_j |U_{ij}|^2\right) dU. \quad (6.1)$$

6.3 Symmetry

Let $f(A, B)$ denote the right-hand side of (6.1). We have established that $f(A, B)$ must be:

1. Symmetric in the eigenvalues $\{a_1, \dots, a_N\}$ of A
2. Symmetric in the eigenvalues $\{b_1, \dots, b_N\}$ of B
3. Analytic in all variables when the eigenvalues are distinct

When some eigenvalues coincide, the function must behave appropriately. Specifically:

Lemma 6.1. *If $a_i = a_j$ for some $i \neq j$, then $f(A, B)$ must be invariant under permuting the corresponding b_i and b_j .*

Proof. When eigenvalues coincide, the corresponding eigenvectors can be chosen arbitrarily within the degenerate subspace. This means that when $a_i = a_j$, we can apply a unitary transformation that effectively swaps the roles of b_i and b_j without changing the integral. \square

Remark on rigor. To make these symmetry arguments fully rigorous, one notes that $f(A, B)$ can be extended to an analytic function of the eigenvalues (even when they are treated as complex variables close to the real axis). Moreover, if some $a_i = a_j$, the existence of a unitary acting within the degenerate subspace justifies the required symmetry in (b_i, b_j) . One also checks that $f(A, B)$ remains finite in the limit $(a_j - a_i) \rightarrow 0$ or $(b_j - b_i) \rightarrow 0$, enforcing vanishing at a rate that compensates for the factor in the denominator.

This constraint, combined with analyticity, forces $f(A, B)$ to vanish as $(a_j - a_i) \rightarrow 0$ or $(b_j - b_i) \rightarrow 0$ at a rate that exactly cancels the denominator's singularity. The form of the answer must therefore be:

$$f(A, B) = \frac{g(A, B)}{\prod_{1 \leq i < j \leq N} (a_j - a_i) \prod_{1 \leq i < j \leq N} (b_j - b_i)},$$

where $g(A, B)$ is analytic and antisymmetric in the $\{a_i\}$ and in the $\{b_i\}$ variables.

6.4 Conclusion of the argument

By the fundamental theorem of antisymmetric polynomials, $g(A, B)$ must be expressible as a product of the Vandermonde determinants and a symmetric function. Moreover, by examining the behavior under the scaling $A \mapsto tA$ and $B \mapsto B/t$, one shows that the only function with the correct analytic properties and scaling behavior is

$$g(A, B) = C_N \cdot \det[e^{a_i b_j}]_{i,j=1}^N,$$

where C_N is a constant depending only on N . One can alternatively pin this down by checking that $f(A, B)$ satisfies a certain heat equation in A (or B), and thus matches the known solution $\det[e^{a_i b_j}]$ up to a constant.

Therefore, we have established that

$$\int_{U(N)} e^{\text{Tr}(A U B U^\dagger)} dU = \Phi_N \frac{\det[e^{a_i b_j}]_{i,j=1}^N}{\prod_{1 \leq i < j \leq N} (a_j - a_i) \prod_{1 \leq i < j \leq N} (b_j - b_i)},$$

where $\Phi_N = C_N$ is a normalization constant independent of the eigenvalues. Through a separate calculation (see Problem J.4), often involving either a small-time heat-kernel expansion or a rank-one reduction, one can determine that

$$\Phi_N = \prod_{k=1}^{N-1} k!. \quad (6.2)$$

J Problems (due 2025-04-29)

J.1 Collisions

Show that two independent standard 1D Brownian motions, started at $a_1 \neq a_2$, almost surely intersect.

J.2 Estimate on the modulus of continuity

Let $B(t)$ be a standard 1D Brownian motion with $B(0) = 0$, defined as a process with independent increments and $B(t) - B(s) \sim \mathcal{N}(0, t - s)$, without any continuity assumptions.

Show that

$$\mathbb{E} |B(t) - B(s)|^2 \leq |t - s|$$

implies that that one can take an almost surely continuous modification of the function $t \mapsto B(t)$.

J.3 Generator for Dyson Brownian Motion

Consider the Dyson Brownian Motion (Definition 3.1) for general $\beta > 0$. The invariant measure for this process when started from zero is expected to be the distribution with density proportional to:

$$p_\beta(\lambda_1, \dots, \lambda_N) \propto \prod_{i < j} |\lambda_i - \lambda_j|^\beta \exp \left\{ -\frac{1}{2} \sum_{i=1}^N \lambda_i^2 \right\}.$$

Prove that this density is invariant under the Dyson SDE (3.2) by showing

$$\mathcal{L} p_\beta = 0,$$

where \mathcal{L} is the infinitesimal generator of the process. Specifically, compute:

$$\mathcal{L} \rho = \frac{1}{2} \sum_{i=1}^N \frac{\partial^2}{\partial \lambda_i^2} \rho - \frac{\beta}{2} \sum_{i=1}^N \sum_{j \neq i} \frac{\partial}{\partial \lambda_i} \left(\frac{1}{\lambda_i - \lambda_j} \rho \right),$$

and verify that it indeed annihilates p_β .

J.4 Constant in the HCIZ formula

Show that in the Harish–Chandra–Itzykson–Zuber formula, the constant Φ_N is given by

$$\Phi_N = \prod_{k=1}^{N-1} k!,$$

by *directly* evaluating the left-hand side for the special case

$$A = \text{diag}(x, 0, \dots, 0), \quad B = \text{diag}(y, 0, \dots, 0).$$

In this rank-one case, note that

$$\text{Tr}(A U B U^\dagger) = x y |U_{11}|^2.$$

You can then reduce the integral to one over the distribution of the first column of U , which is a vector uniformly distributed on the complex unit sphere \mathbb{C}^N (under the normalized Haar measure). Use the known Jacobian for this parametrization to perform the integral and match it with the right-hand side evaluated at $(a_1, b_1) = (x, y)$ and $(a_2 = \dots = a_N = b_2 = \dots = b_N = 0)$.

J.5 Biorthogonal ensembles

Derive Theorem 5.3 from Lemma 5.4 using the orthogonalization process similar to Lecture 5, and then taking the limit as $s \rightarrow \infty$.

J.6 Scaling of the kernel

Let $a_i = 0$ in Theorem 5.3. Find α such that $t^\alpha K_t(x/\sqrt{t}, y/\sqrt{t})$ is independent of t . Can you explain this value of α ?

References

- [BBP05] J. Baik, G. Ben Arous, and S. Péché, *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices*, Ann. Probab. **33** (2005), no. 5, 1643–1697. arXiv:math/0403022 [math.PR]. [↑8](#)
- [Bor11] A. Borodin, *Determinantal point processes*, Oxford handbook of random matrix theory, 2011. arXiv:0911.1153 [math.PR]. [↑9](#)
- [BR05] A. Borodin and E.M. Rains, *Eynard–Mehta theorem, Schur process, and their Pfaffian analogs*, J. Stat. Phys **121** (2005), no. 3, 291–317. arXiv:math-ph/0409059. [↑9](#)
- [Dys62] F.J. Dyson, *A Brownian motion model for the eigenvalues of a random matrix*, Jour. Math. Phys. **3** (1962), no. 6, 1191–1198. [↑3](#), [6](#)

L. PETROV, UNIVERSITY OF VIRGINIA, DEPARTMENT OF MATHEMATICS, 141 CABELL DRIVE, KERCHOF HALL, P.O. BOX 400137, CHARLOTTESVILLE, VA 22904, USA
E-mail: lenia.petrov@gmail.com