

# 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: $1 \times 1$ 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 03:59, Monday 24<sup>th</sup> 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  $G_\beta$  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  $\sigma$  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  $\sigma$ , 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  $\Delta t$ , 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  $\sigma$ .

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  $\Delta t$ , 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

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

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  $\Delta t$ , we can ignore  $\Delta \lambda$  in the denominator, replace  $B_{ji}^2(\Delta t)$  by  $\Delta t$  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  $\lambda_i$  away from collisions with  $\lambda_j$ . This repulsion is so strong (for all  $\beta > 0$ ) that eigenvalues will not cross (and thus remain ordered) with probability one.

## 4 Transition densities

### 4.1 Starting from zero

If the Dyson Brownian motion starts from zero<sup>1</sup>  $\lambda_1(0) = \dots = \lambda_N(0) = 0$ , 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  $\beta$ . For other  $\beta$ , we would like to show that

---

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

## 5 Determinantal structure for $\beta = 2$

## 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

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

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

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

### 6.3 Symmetry

- The expression must be symmetric in the eigenvalues  $\{a_1, \dots, a_N\}$  (since any permutation of those entries can be achieved by a suitable unitary conjugation).
- Similarly, it is symmetric in the eigenvalues  $\{b_1, \dots, b_N\}$ .
- One can argue that if  $a_i = a_j$  for some  $i \neq j$  but  $b_i \neq b_j$ , the integral must vanish (or degenerate) appropriately, which implies the presence of Vandermonde factors in the denominator.

By general principles of polynomial (or analytic) symmetry, one shows

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

where  $\Phi_N$  is a constant (independent of the specific eigenvalues).

### 4. Determining the Constant $\Phi_N$

One standard method to fix  $\Phi_N$ :

1. Set some  $\{a_i\}$  or  $\{b_j\}$  to special values (e.g. let one subset go to 0 or let all but one eigenvalue be 0).
2. Expand both sides in power series, compare coefficients, or use known integrals from representation theory or from the “Weingarten formula” approach.

It is well known that the result is

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



## 5. Connection to Dyson Brownian Motion

In the context of Dyson Brownian Motion (for  $\beta = 2$ ), one writes the law of  $\mathcal{M}(t) = A + \sqrt{t}G_{\text{UE}}$  and integrates out the angular parts via

$$\int_{U(N)} \exp(\text{Tr}(A U \Lambda U^\dagger)) dU,$$

where  $\Lambda$  is diagonal with eigenvalues  $\{\lambda_1, \dots, \lambda_N\}$ . The HCIZ formula then produces the classic determinant–Vandermonde factor, yielding the transition density for the eigenvalues of  $\mathcal{M}(t)$ . This leads to the “ $\exp(a_i \lambda_j)$ ” structure in the joint density or correlation kernels discussed in Lecture 10.

## 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

1. **Generator for the Dyson SDE.** For  $\beta > 0$ , write the generator  $\mathcal{L}$  of the Markov process (3.2) in differential form. That is, show that for any  $C^2$  test function  $f(\lambda_1, \dots, \lambda_N)$ ,

$$\frac{d}{dt} \mathbb{E}[f(\lambda(t))] = \mathbb{E}[\mathcal{L}f(\lambda(t))],$$

where

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

Verify formally that the density from ?? is stationary for  $\mathcal{L}$  when  $\lambda(0) = 0$ .

2. **Limit shape for Non-Intersecting Bridges.** Let  $b_1(t), \dots, b_N(t)$  be  $N$  real Brownian bridges from  $a_1 < \dots < a_N$  at  $t = 0$  to  $z_1 < \dots < z_N$  at  $t = 1$ , conditioned on no intersections. Simulate numerically for  $N = 10$  and  $a_i = i$ ,  $z_i = 2i$ . Describe visually how the paths arrange themselves and how they may approximate a continuous shape as  $N$  grows.

3. **Determinantal representation of transition densities (general  $\beta = 2$ ).** Fill in details for the matrix integral approach that leads to the transition probability formula in ??, including the key use of the unitary integral (the HCIZ formula). You may take the HCIZ formula as known or provide a reference. Make explicit the factor by which  $P(\lambda(t) = x | \lambda(0) = a)$  is normalized.
4. **DBM short time expansions.** Let  $\lambda_i(0) = 0$  for all  $i$ , and consider (3.2) for small  $t$ . Expand the drift and diffusion terms to second order in  $t$ , and compare with the approach in ??. Why do the  $\frac{1}{\lambda_i - \lambda_j}$  terms not blow up when  $\lambda_i \approx \lambda_j$  for small  $t$ ?

## References

[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](mailto:lenia.petrov@gmail.com)