

Lectures 16– 17 : Gaussian Ensembles and Local Limits

Notes prepared by: Shirzad Malekpour and Sumeet Katariya

1 GUE Ensembles and the limit

Recall the joint eigenvalue density of gaussian ensembles, where $\beta = 2$, denotes GUE and is given by

$$P(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_n} e^{-\frac{1}{2} \sum_{i=1}^n \lambda_i^2} \prod_{i < j} |\lambda_i - \lambda_j|^2 \quad (1)$$

$$\Lambda_n = \{\lambda_1, \lambda_2, \dots, \lambda_n\} \quad (2)$$

We would like to prove a scaling limit for this probability distribution. We would like to show that, the scaled around zero, zoomed in the way that the distance between the neighboring eigenvalues are of order 1, will produce a nice point process. Specifically we would like to show that $\frac{\sqrt{n}}{2} \Lambda_n$ converges to a nice point process. Today we will prove that intensities converge and the limit is the determinant of a k by k matrix with the kernel functions ψ . Recall the following:

$$K^{(n)}(X, Y) = \sum_{k=0}^{n-1} \psi_k(X) \psi_k(Y) \quad (3)$$

Where the ψ_k is the k -th oscillator wave function. We have this finite kernel which was given by these oscillator wave functions which are given in terms of hermit polynomials and if we show that this normalized version of kernel $\frac{1}{\sqrt{n}} K^{(n)}(\frac{X}{\sqrt{n}}, \frac{Y}{\sqrt{n}})$ converges, this will give you the limit of joint intensities. Moreover we will prove that:

$$S^{(n)}(X, Y) = \frac{1}{\sqrt{n}} K^{(n)}(\frac{X}{\sqrt{n}}, \frac{Y}{\sqrt{n}}) \rightarrow \frac{\sin(X - Y)}{\pi(X - Y)} \quad (4)$$

Last time we sort of proved the statement using the following lemma:

Lemma 1. *Suppose that you have an extra variable v , and v will take values like $n, n + 1, n + 2, n - 1, n - 2 \dots$ for which $v - n$ will be a constant. Then we introduce the following rescaled wave function:*

$$\varphi_v(t) = n^{\frac{1}{4}} \psi_v(\frac{t}{\sqrt{n}}) \quad (5)$$

Then,

$$\lim |\varphi_v(t) - \frac{1}{\sqrt{n}} \cos(t - \frac{\pi v}{2})| = 0 \quad (6)$$

Uniformly in t on compacts.

Note that once this lemma is proved it will be more or less a simple computation to check that limit (4) holds. To prove this lemma, we need to know the Laplace Method:

1.0.1 Laplace Method

Suppose we want to identify $\int f(x)^s g(x) dx$ as $s \rightarrow \infty$. Where $f \geq 0$ and f has a global maximizer at a and near a it is locally quadratic. Also g should be a nice function, locally lipschitz near a . Then $\sqrt{s} f(a)^{-s} \int f(x)^s g(x) dx \rightarrow \sqrt{\frac{-2\pi f(a)}{f''(a)}} g(a)$ as $s \rightarrow \infty$.

1.1 Proof of the Lemma

Proof. Consider the following identity that we just rewrite using the fourirer integral. It is basically using the formula for the characteristic function for standard random variable.

$$e^{-\frac{X^2}{2}} = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{\xi^2}{2} - i\xi X} d\xi \quad (7)$$

Remember that we got the hermit polynomial by differentiating this exponential n times. So we differentiate both sides n times and since the integral is nice enough we can differentiate inside the integral. then we will have:

$$(-1)^n (e^{-\frac{X^2}{2}})^{(n)} = e^{-\frac{X^2}{2}} h_n(X) = \frac{1}{\sqrt{2\pi}} \int (i\xi)^n e^{-\frac{\xi^2}{2} - i\xi X} d\xi \quad (8)$$

Where the $h_n(X)$ is the n -hermit polynomial. Now write the oscillator function corresponding to index v .

$$\psi_v(X) = \frac{e^{-\frac{X^2}{2}} i^v}{(2\pi)^{\frac{3}{4}} \sqrt{v!}} \int \xi^v e^{-\frac{\xi^2}{2} - i\xi X} d\xi \quad (9)$$

Then by normalizing this wave function we will have:

$$\varphi_v(t) = \frac{e^{-\frac{t^2}{4n}} i^v n^{\frac{1}{4}}}{(2\pi)^{\frac{3}{4}} \sqrt{v!}} \int \xi^v e^{-\frac{\xi^2}{2} - \frac{i\xi t}{\sqrt{n}}} d\xi \quad (10)$$

Now we want to get rid of the \sqrt{n} inside the integral. So we use the change of variables. Let $Z = \frac{\xi}{\sqrt{n}}$. Then,

$$\varphi_v(t) = \frac{\sqrt{n} (2\pi)^{\frac{1}{4}} n^{\frac{1}{4} + \frac{v}{2}} e^{-\frac{t^2}{4n}}}{2\pi \sqrt{v!}} \int (Z e^{-\frac{Z^2}{2}})^n i^v e^{-iZt} Z^{v-n} dZ \quad (11)$$

Remember that $v - n$ inside the integral is a constant. Now Let's try to simplify the multiplier. Using the stirling formula we have:

$$\frac{n^{\frac{1}{4} + \frac{v}{2}}}{\sqrt{v!}} \sim \frac{n^{\frac{1}{4} + \frac{n}{2}}}{\sqrt{n!}} \sim e^{\frac{n}{2}} (2\pi)^{-\frac{1}{4}} \quad (12)$$

Also note that t is in a compact interval and so it is bounded. Therefore the term $e^{-\frac{t^2}{4n}}$ goes to 1. Replacing these in (11) we then get:

$$\varphi_v(t) \sim \frac{\sqrt{n}}{2\pi} e^{\frac{n}{2}} \int (Z e^{-\frac{Z^2}{2}})^n i^v e^{-iZt} Z^{v-n} dZ \quad (13)$$

Note that we are finding a real valued function. So it makes sense to only look at the real part of the integral:

$$\varphi_v(t) \sim \frac{\sqrt{n}}{2\pi} e^{\frac{n}{2}} \int (Z e^{-\frac{Z^2}{2}})^n \operatorname{Re}(i^v e^{-iZt}) Z^{v-n} dZ \quad (14)$$

But we can write $i^v e^{-iZt}$ as $e^{i(\frac{\pi v}{2} - Zt)}$. Hence (14) will reduce to:

$$\frac{\sqrt{n}}{2\pi} e^{\frac{n}{2}} \int_{-\infty}^{\infty} (Z e^{-\frac{Z^2}{2}})^n \cos(Zt - \frac{v\pi}{2}) Z^{v-n} dZ \quad (15)$$

Now we are ready to use the Laplace formula. To do that we need to consider different cases. First consider the case when n is even. Then the function $(Z e^{-\frac{Z^2}{2}})^n$ is even. Moreover, $\cos(Zt - \frac{v\pi}{2}) Z^{v-n}$ is also even. so we can rewrite the integral as:

$$\frac{\sqrt{n}}{\pi} e^{\frac{n}{2}} \int_0^{\infty} (Z e^{-\frac{Z^2}{2}})^n \cos(Zt - \frac{v\pi}{2}) Z^{v-n} dZ \quad (16)$$

Now by applying the Laplace formula to $f(X) = X e^{-\frac{X^2}{2}}$ and $g(X) = \cos(Xt - \frac{v\pi}{2}) X^{v-n}$ we get the result. Note that the global maximizer for $f(X)$ is at 1 and $f(1) = e^{-\frac{1}{2}}$ and also $\frac{-f'(1)}{f''(1)} = \frac{1}{2}$. Then by letting $s = n, a = 1$ we have $f^{-s}(a) = e^{\frac{n}{2}}$ and then the desired result is proved:

$$\varphi_v(t) \rightarrow \frac{1}{\sqrt{\pi}} \cos(t - \frac{v\pi}{2}) \quad (17)$$

(17) is not correct in general. you have to pick the v that provides the appropriate parity. Also this convergence is in the sense that the difference is going to zero:

$$|\varphi_v(t) - \frac{1}{\sqrt{\pi}} \cos(t - \frac{v\pi}{2})| \rightarrow 0 \quad (18)$$

and from this we conclude that $S^{(n)}(X, Y) \rightarrow \frac{\sin(X-Y)}{\pi(X-Y)}$. We are not done with the proof yet. We need to consider all other cases, i.e n =odd. This will have similar computation. Next, to be more precise you need to consider all possible cases and check the parity of $\cos()$, but the main idea will remain the same. Furthermore you need this for a family of functions , g , because the claim is saying this holds uniformly in t .

Now that we proved the bulk scaling limit near 0, let's consider this problem for a general c , where $|c| < 2$. Now we have to center about the $c\sqrt{n}$ and we have to blow up the picture by $\sqrt{n}\sigma(c)$ where $\sigma(c)$ was $\frac{1}{2\pi}\sqrt{4-c^2}\mathbf{1}(|c| < 2)$. The proof for this case differs from previous in the following: We need to prove that $\frac{1}{\sqrt{n}}K^{(n)}(c\sqrt{n} + \frac{X}{\sqrt{n}}, c\sqrt{n} + \frac{Y}{\sqrt{n}}) \rightarrow \frac{\sin(\pi\sigma(c)(X-Y))}{\pi(X-Y)}$. For this case we will get the same result. The term $c\sqrt{n}$ may also be replaced by any sequence of the form c_n which grows like $c\sqrt{n}$.

For the case of analyzing the edges, you need another multiplier, 6 here. To understand the intensities we can look at the kernels and for this case it will be $2\sqrt{n} + \frac{X}{n^{\frac{1}{6}}}$. From this point analysis are almost the same as before but a little more complicated integrals. \square

2 Couple of statements

These limits will give the joint intensities, but we may look for actual probabilities.

Theorem 2. Let $A \subset \mathbb{R}$ Then,

$$\mathcal{P}_n(\bigcap_{i=1}^n \{\lambda_i \in A\}) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^c} \int_{A^c} \dots \int_{A^c} \det K^{(n)}(X_i, X_j) dX_1 \dots dX_n \quad (19)$$

The matrix inside the integral is a k by k matrix built by finite kernels. Remember that this matrix was the product of two matrices:

$$K^{(n)}(X_i, X_j) \Big|_{i,j=1}^k = \psi_k \times_n \psi_n \times_k \quad (20)$$

In all these expressions n is fixed so the rank $\leq n$ which means that even though we have an infinite sum here but high terms will be zero and it will reduce to finite sum. Remember that we had:

$$\mathcal{P}_n(\lambda_1, \dots, \lambda_n) = \frac{1}{n!} \det K^n(X_i, X_j) \quad (21)$$

Therefore (19) will be:

$$\int_A \int_A \dots \int_A \det K^{(n)}(X_i, X_j) \prod dX_i \quad (22)$$

Now using Jacobi identity and note that ψ functions are orthonormal we get (22) equal to:

$$\det \int \psi_i(X) \psi_j(X) dX \rightarrow \det(\delta_{ij} - \int_{A^c} \psi_i(X) \psi_j(X) dX) \quad (23)$$

Now we have difference of two matrices and using simple linear algebra we get:

$$1 + \sum_{k=1}^n \sum_{0 \leq v_1 \leq v_2 \dots v_k \leq n-1} (-1)^k \det \int_{A^c} \psi_{v_i}(X) \psi_{v_j}(X) dX \quad (24)$$

Then again using Jacobi identity and Cauchy-Binet formula the results can be proved.

The huge expression in (19) is called Fredholm determinant. To find this you need a kernel function. Assume $K : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ and also assume K is bounded, i.e. $\|K\| = \sup_{X,Y} |K(X,Y)| < \infty$. Also Our reference measure will be Lebesgue measure. Then:

$$\Delta_n(K) = \int \dots \int \det K(X_i, X_j) dv(X_1) \dots dv(X_n) \quad (25)$$

and the Fredholm determinant corresponding to this kernel function K is defined by:

$$\Delta(K) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \Delta_n(K) \quad (26)$$

Assuming that $\Delta_0(K) = 1$. Convergence can be shown but it is not presented here. the following lemma proves why this is finite:

Lemma 3.

$$|\det F(X_i, X_j) - \det G(X_i, X_j)| \leq n^{1+\frac{n}{2}} \|F - G\| \max(\|F\|, \|G\|)^{n-1} \quad (27)$$

For example if we plug $G = 0$ we will have: $|\det F(X_i, X_j)| \leq n^{1+\frac{n}{2}} \|F\|^n$. Now if K is bounded and we use the mentioned bounds because of the $n!$ in the denominator we can conclude that the sum is convergent. Note that the limit of these probabilities can be expressed in terms of Fredholm determinant involving sine kernels.

Proof of the Lemma

Proof. Define

$$H_i^{(k)}(x, y) = \begin{cases} G(x, y) & i < k \\ F(x, y) - G(x, y) & i = k \\ F(x, y) & i > k \end{cases}$$

noting that, by the linearity of the determinant with respect to rows,

$$\det F(x_i, y_j) - \det G(x_i, y_j) = \sum_{k=1}^n \det H_i^{(k)}(x_i, y_j). \quad (28)$$

Now, the Hadamard Inequality states that for any column vectors v_1, \dots, v_n of length n with complex entries, it holds that

$$\det[v_1 \cdots v_n] \leq \prod_{i=1}^n \sqrt{\bar{v}_i^T v_i} \leq n^{n/2} \prod_{i=1}^n |v_i|_\infty$$

Considering the vectors $v_i = v_i^{(k)}$ with $v_i(j) = H_i^{(k)}(x_i, y_j)$, and applying Hadamard's inequality, one gets

$$\left| \det H_i^{(k)}(x_i, y_j) \right| \leq n^{n/2} \|F - G\| \cdot \max(\|F\|, \|G\|)^{n-1}.$$

Substituting in (28) yields (27). □

Now using this lemma we can prove the following for GUE.

Theorem 4. *Gap probabilities in GUE:*

We can ask for the following probability:

$$\lim \mathcal{P}_n \left((\sqrt{n}\lambda_1 \dots \sqrt{n}\lambda_n) \notin \left(-\frac{t}{2}, \frac{t}{2}\right) \right) \quad (29)$$

This probability is the of limiting point process not having any points in $(-\frac{t}{2}, \frac{t}{2})$. If they are not in $(-\frac{t}{2}, \frac{t}{2})$ they should be in its complement. Then (29) will be:

$$= \lim \mathcal{P}_n \left(\bigcap \{ \sqrt{n}\lambda_i \in [-\frac{t}{2}, \frac{t}{2}]^c \} \right) \quad (30)$$

$$= \lim \mathcal{P}_n \left(\bigcap \{ \lambda_i \in [\frac{-t}{2\sqrt{n}}, \frac{t}{2\sqrt{n}}]^c \} \right) \quad (31)$$

Let $A = [-\frac{t}{2}, \frac{t}{2}]$. Then we have:

$$= \lim \left[1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{\frac{A}{\sqrt{n}}} \int_{\frac{A}{\sqrt{n}}} \dots \int_{\frac{A}{\sqrt{n}}} \det K^{(n)}(X_i, X_j) dX_1 \dots dX_k \right] \quad (32)$$

$$= \lim \left[1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \int_A \dots \int_A \det S^{(n)}(X_i, X_j) dX_1 \dots dX_k \right] \quad (33)$$

Now we have Fredholm determinant for finite. So we can look at the limit of the Kernel:

$$= \lim \left[1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \int_A \dots \int_A \det \frac{1}{\pi} \frac{\sin(X_i - X_j)}{(X_i - X_j)} dX_1 \dots dX_k \right] \quad (34)$$

Where $\frac{1}{\pi} \frac{\sin(X_i - X_j)}{(X_i - X_j)}$ is the $K_{\text{sine}}(X_i, X_j)$. It is called Fredholm determinant and it is reduced to:

$$\det(I - K_{\text{sine}} \mathbf{1}_A) = P(A \text{ is empty for the limiting point process}) \quad (35)$$

2.1 Analysis

One can try to study this determinant. There are various sophisticated methods to analyze:

1. Jimbo-Miwa-Mori-Sato

$$\lim P\left(\left(-\frac{t}{2}, \frac{t}{2}\right) \text{ empty}\right) = 1 - F(t) \quad (36)$$

Where $1 - F(t) = \exp(\int_0^t \sigma(X) dX)$ and σ is the solution of the following differential equation:

$$(t\sigma'') + 4(t\sigma' - \sigma)(t\sigma' - \sigma + (\sigma')^2) = 0 \quad (37)$$

Assuming σ be a function of t . And the boundary condition is given by asymptotic near zero:

$$\sigma(t) = -\frac{t}{\pi} - \frac{t^2}{\pi^2} - \frac{t^3}{\pi^3} + O(t^4) \quad (38)$$

as $t \rightarrow 0$.

Edge scaling hasn't been discussed yet. The following is a similar result for edge case:

2. Tracy-Widom

Remember that we have $n^{\frac{1}{6}}(2\sqrt{n} - \Lambda_n)$ for this case. We zoom in near the edge of the same semi-circle. Then there will be a smallest (considering the minus sign) point. So we will look at the probability that the smallest number is bigger than t .

$$P(\text{edge is empty for } (-\infty, 0)) \rightarrow F_2(t). \quad (39)$$

Where $F_2(t) = \exp(-\int_t^\infty (X - t)^2 q^2(X) dX)$ and q is the solution to the following differential equation:

$$q'' = tq + 2q^3 \quad (40)$$

and the boundary condition is $q(t) \sim Ai(t)$ as $t \rightarrow \infty$.

In the case of $\beta = 2$ there are explicit formulas for probabilities although they are a little bit complicated. For $\beta = 1, 4$ you will have vandermond and vandermond⁴. You can not do exactly the same calculation but still Bulk scaling limits with more or less explicit but more complicated expressions can be done.

3 Edge Sclaing for the general β -ensembles

For a fixed $\beta > 0$ we have $\Lambda_n = \{\lambda_1, \dots, \lambda_n\}$ and then the joint intensities will be:

$$\mathcal{P}_n(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,\beta}} \Delta(\lambda)^\beta e^{-\frac{\beta}{4} \sum_{i=1}^n \lambda_i^2} \quad (41)$$

We would like to understand the point process limit of this ensemble. It can be proved that the empirical spectral will converge to the semi-circle. We apply the same scaling near the edge $n^{\frac{1}{6}}(2\sqrt{n} - \Lambda_n)$. We would like to show that $n^{\frac{1}{6}}(2\sqrt{n} - \Lambda_n)$ converges weakly to a random point process $(\lambda_0, \lambda_1, \dots)$ where $\lambda_0 \leq \lambda_1 \leq \dots$.

This is an infinite sequence and one way to prove it is to do it for finite marginals. For each k we look at first $k+1$ eigenvalues and using rescale property in limit you will get certain random vector:

$$n^{\frac{1}{6}}(2\sqrt{n} - \lambda_n^{(n)}), n^{\frac{1}{6}}(2\sqrt{n} - \lambda_{n-1}^{(n)}), \dots, n^{\frac{1}{6}}(2\sqrt{n} - \lambda_{n-k}^{(n)}) \rightarrow (\lambda_0, \lambda_1, \dots, \lambda_k) \quad (42)$$

For general β -ensembles, instead of looking at joint density we can look at another representation. Recall the tridiagonal matrix:

$$M_n = \begin{bmatrix} a_n & b_{n-1} & & & & \\ b_{n-1} & a_{n-1} & b_{n-2} & & & \\ & b_{n-2} & \ddots & \ddots & & \\ & & \ddots & \ddots & b_1 & \\ & & & & b_1 & a_1 \end{bmatrix}_{n \times n} \quad (43)$$

Where $a_k \sim \frac{\sqrt{2}}{\beta} N(0, 1)$ and $b_k \sim \frac{1}{\sqrt{\beta}} \mathcal{X}_{\beta k}$. The joint eigenvalue density of M_n is given by $\mathcal{P}_{n,\beta}(\lambda_1, \dots, \lambda_n)$.

So if we analyze the rescaled eigenvalues of this matrix, it is like analyzing the rescaled eigenvalues of β -ensembles. Therefore we apply the same rescaling. We want to show that the eigenvalues of $n^{\frac{1}{6}}(2\sqrt{n}I_n - M_n)$ converges to something nice. And this is the same statement we had before.

Remember that $\mathcal{X}_{\beta k}^2 \sim \Gamma(\frac{1}{2}, \frac{\beta k}{2})$. In particular if n is an integer, then $\mathcal{X}_n^2 = X_1^2 + \dots + X_n^2$ where X_i are iid $N(0, 1)$. Now using law of large numbers and central limit Theorem and assuming that we have standard normal random variabls we will have the following equations:

$$\mathcal{X}_n = \sqrt{\sum_{i=1}^n X_i^2} \quad (44)$$

$$= \sqrt{n} \sqrt{\frac{\sum_{i=1}^n X_i^2}{n}} \quad (45)$$

$$= \sqrt{n} \sqrt{\frac{\sum_{i=1}^n (X_i^2 - 1)}{n} + 1} \quad (46)$$

$$\simeq \sqrt{n} \left(1 + \frac{\sum_{i=1}^n (X_i^2 - 1)}{2n} \right) \quad (47)$$

$$= \sqrt{n} + \frac{\sum_{i=1}^n (X_i^2 - 1)}{2\sqrt{n}} \quad (48)$$

The second term in (48) is asymptotically $N(0, \frac{1}{2})$ and finally it means that $\mathcal{X}_n \simeq \sqrt{n} + \frac{1}{\sqrt{2}}\zeta$ where ζ is a standard normal random variable. Using moment generating function it can be proved that

this result also holds for non-integer indices.

When we look at largest eigenvalues of the matrix and look at expected values, we will have zeors on diagonals and \sqrt{k} on the off-diagonals. Therefore, we will have large values on left top corner of the matrix and vanishing values on bottom right. More specifically we have:

$$b_k \sim \sqrt{k} + \frac{1}{\sqrt{2\beta}}\zeta \quad (49)$$

$$b_{n-k} = \sqrt{n-k} + \frac{1}{\sqrt{2\beta}}\zeta_k \simeq \sqrt{n} - \frac{k}{2\sqrt{n}} + \frac{1}{\sqrt{2\beta}}\zeta_k \quad (50)$$

if $k \ll n$.

Now we have to look at the tridiagonal matrix $n^{\frac{1}{6}}(2\sqrt{n}I_n - M_n)$:

Diagonal entries: $n^{\frac{1}{6}}(2\sqrt{n} - \sqrt{\frac{2}{\beta}}\rho_i)$

Off-diagonal entries: $-n^{\frac{1}{6}}\sqrt{n} + \frac{k}{2\sqrt{n}}n^{\frac{1}{6}} + n^{\frac{1}{6}}\frac{1}{\sqrt{2\beta}}\zeta_k$

If we try to match these terms, notice that main terms are of the same order,Then we can rewrite:

$$n^{\frac{1}{6}}(2\sqrt{n}I_n - M_n) = n^{\frac{2}{3}} \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix} + \frac{n^{-\frac{1}{3}}}{2} \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 2 & & \\ & 2 & \ddots & \ddots & \\ & & \ddots & \ddots & n-1 \\ & & & n-1 & 0 \end{bmatrix} \quad (51)$$

$$+ \frac{n^{\frac{1}{6}}}{\sqrt{2\beta}} \begin{bmatrix} 2\rho_0 & \zeta_1 & & & \\ \zeta_1 & 2\rho_1 & \zeta_2 & & \\ & \zeta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \zeta_{n-1} \\ & & & \zeta_{n-1} & 2\rho_{n-1} \end{bmatrix} \quad (52)$$

3.1 Edelman-Sutton Conjecture

If you look at this matrix and let $n \rightarrow \infty$ then it will approximate an operator:

$$\partial_{XX} + X + \frac{2}{\sqrt{\beta}}B'(X) \quad (53)$$

Where $B(X)$ is the standard brownian motion¹. The first term is differentiation twice. The second term is just multiplication by function X and the last term is multiplication by derivative of a standard brownian motion. Note that although brownian motion is not differentiable but, by looking at it as a distribution and using integration by parts you can still make sense of it. The whole argument was made rigorous by Ramirez-Rider-Virag (07).

Now let's try to analyze the terms. If you have a nice twice differentiable function f and you look at a finite mesh $\{h, 2h, 3h, \dots, kh, \dots\}$. Then let $f_k = f(kh)$. By making h smaller and smaller

¹See end of notes for a review of Brownian motion

your approximation of function and its first and second derivative will be more precise. Now we can approximate the second derivative by:

$$f''(kh) \simeq \frac{f((k+1)h) + f((k-1)h) - 2f(kh)}{h^2} = \frac{f_{k+1} + f_{k-1} - 2f_k}{h^2} \quad (54)$$

So now if we look at f as a vector $(f_1, f_2, \dots)^T$, we can get $(f''_1, f''_2, \dots)^T$ by multiplying the function values by a tridiagonal matrix. Let $f''_k = f''(kh)$ then the tridiagonal matrix will be:

$$\frac{1}{h^2} \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix} \quad (55)$$

This is basically an approximation of second derivative on a mesh of size h . It means that if you look at:

$$n^{\frac{2}{3}} \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix} \quad (56)$$

in (52), it is just the $-\partial_{XX}$ on a mesh of size $n^{-\frac{1}{3}}$.

For the second term in (52) we can take $n^{-\frac{1}{3}}$ term inside the matrix:

$$\frac{n^{-\frac{1}{3}}}{2} \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 2 & & \\ & 2 & \ddots & \ddots & \\ & & \ddots & \ddots & n-1 \\ & & & n-1 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & n^{-\frac{1}{3}} & & & \\ n^{-\frac{1}{3}} & 0 & 2n^{-\frac{1}{3}} & & \\ & 2n^{-\frac{1}{3}} & \ddots & \ddots & \\ & & \ddots & \ddots & (n-1)n^{-\frac{1}{3}} \\ & & & (n-1)n^{-\frac{1}{3}} & 0 \end{bmatrix} \quad (57)$$

Now we will multiply this matrix by $(f_1, f_2, \dots)^T$. For small h , f_k and f_{k+1} are close together and we can move the off-diagonals to diagonals with negligible error.

$$\sim \begin{bmatrix} n^{-\frac{1}{3}} & 0 & & & \\ 0 & 2n^{-\frac{1}{3}} & 0 & & \\ & 0 & \ddots & \ddots & \\ & & \ddots & \ddots & \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \end{bmatrix} \quad (58)$$

This is basically the same as multiplication with X on the same mesh. Now consider the last term. Again by hand-waving computation we can make it diagonal. Assuming f is a nice function we

will have:

$$\frac{n^{\frac{1}{6}}}{\sqrt{2\beta}} \begin{bmatrix} 2\rho_0 & \zeta_1 & & & & & & & & & \\ \zeta_1 & 2\rho_1 & \zeta_2 & & & & & & & & \\ & \zeta_2 & \ddots & \ddots & & & & & & & \\ & & \ddots & \ddots & \zeta_{n-1} & & & & & & \\ & & & & \zeta_{n-1} & 2\rho_{n-1} & & & & & \end{bmatrix} \sim \frac{n^{\frac{1}{6}}}{\sqrt{2\beta}} \begin{bmatrix} N(0,1) & 0 & & & & & & & & & \\ 0 & N(0,1) & 0 & & & & & & & & \\ & 0 & \ddots & \ddots & & & & & & & \\ & & \ddots & \ddots & 0 & & & & & & \\ & & & & 0 & N(0,1) & & & & & \\ & & & & & 0 & N(0,1) & & & & \end{bmatrix} \quad (59)$$

Now the question is that why this is the same as multiplication with white noise? What we are doing is:

$$\frac{n^{\frac{1}{6}}}{\sqrt{2\beta}} \begin{bmatrix} N(0,1) & 0 & & & & & & & & & \\ 0 & N(0,1) & 0 & & & & & & & & \\ & 0 & \ddots & \ddots & & & & & & & \\ & & \ddots & \ddots & 0 & & & & & & \\ & & & & 0 & N(0,1) & & & & & \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \sum f(kh)n^{\frac{1}{6}}\zeta_1 \quad (60)$$

$$\rightarrow \int f dB \simeq \sum f(kh)[B(k+1)h - B(kh)] \quad (61)$$

If you look at the scaling parameter of brownian motion, then the step sizes of $n^{\frac{1}{3}}$ will correspond to normals with $n^{\frac{1}{6}}$ variance which is exactly the multiplier. Then the limit is integration by standard brownian motion.

4 Side Remarks on Brownian Motion and White Noise

4.1 Brownian Motion

A Brownian motion process is characterized by three facts:

1. $B(0) = 0$
2. $B(t)$ is almost surely continuous.
3. $B(t)$ has independent increments with distribution $B(t) - B(s) \sim \mathcal{N}(0, t - s)$ (for $0 \leq s \leq t$).

The condition that it has independent increments means that if $0 \leq s_1 \leq t_1 \leq s_2 \leq t_2$ then $B(t_1) - B(s_1)$ and $B(t_2) - B(s_2)$ are independent random variables.

The following facts can be proved:

1. Brownian motion exists.

2. The Brownian motion paths are almost surely nowhere differentiable.

White noise formally is the derivative of Brownian motion. However, since Brownian motion is not differentiable, to speak about white noise, we need the notion of distributions or generalized functions.

4.2 Distributions (Generalized functions)

Consider f to be a linear functional on the space of smooth functions i.e. f maps a smooth, compactly supported function $\varphi(x)$ to $\int f(x)\varphi(x)dx$. If we think of a function as a linear functional, we can include many more functions. For eg: the dirac delta function δ_0 is such that $\varphi \rightarrow \varphi(0)$. If we have a linear functional $T : \varphi \rightarrow T\varphi$, we can define the derivative of this functional in a manner analogous to ordinary functions. For ordinary differentiable functions, we know that $\int f'(x)\varphi(x)dx = -\int f(x)\varphi'(x)dx$. (We can neglect the boundary terms because we are concerned only with compactly supported functions). Hence, we define $T'\varphi = -T\varphi'$. Thus, for the dirac delta, we have that the derivative of δ_0 is a linear functional that maps $\varphi \rightarrow -\varphi'(0)$. This way, we can define the derivative of any continuous function $f(x)$.

Since Brownian motion is continuous, we can define its derivative in the sense of distributions. Thus, $B'(x)$ is a random distribution (generalized function). It maps $\varphi \rightarrow \int \varphi B'(x)dx = \int \varphi dB$, and this random object is distributed as $\mathcal{N}(0, \int \varphi^2)$.

4.3 Scaling Limits

One way to think about Brownian motion is as a scaling limit of discrete random variables. Let τ_1, τ_2, \dots be iid random variables, such that $P(\tau_i = \pm 1) = \frac{1}{2}$. CLT gives us $\frac{1}{\sqrt{n}} \sum \tau_i \rightarrow \mathcal{N}(0, 1)$. This is a random walk, and if you rescale it i.e.

$$S_\epsilon(t) = \sqrt{\epsilon} \sum_{\epsilon i \leq t} \tau_i$$

As $\epsilon \rightarrow 0, S_\epsilon(\cdot) \Rightarrow$ Brownian motion

White noise is a distribution such that if you feed a test function into it, you get a normal random variable out.

$$\sqrt{\epsilon} \sum \varphi(k\epsilon)T_k \Rightarrow \mathcal{N}(0, \int \varphi^2), \text{ where } \varphi \text{ is a compactly supported smooth function}$$

Another way to think about white noise is to take an orthonormal basis in L_2 , say $\varphi_1, \varphi_2, \dots$. Choose an iid sequence of standard normals ξ_1, ξ_2, \dots . We define $\int \varphi_i dB = \xi_i$. We define the action of white noise as follows-

$$f \in L_2 \quad f = \sum c_i \varphi_i \quad \sum c_i^2 < \infty$$

$$\int f dB = \sum c_i \xi_i$$

Brownian motion can be obtained from white noise by

$$B(x) = \int 1_{[0,x]} dB$$