

## LECTURE 1

### 1. THE GAUSSIAN ORTHOGONAL ENSEMBLE

As the eigenvalues of a large matrix depend on the entries of the matrix in a rather complicated manner (they are the zeros of a polynomial of high degree whose coefficients are products of the matrix elements), it would seem that it's very difficult to express the distribution of the eigenvalues in terms of the distribution of the entries in general. Indeed, if you give me a large matrix with random entries whose joint distribution is arbitrary, it's likely to be impossible to say something explicit about the distribution of the eigenvalues.

Due to this, it's important to analyze certain prototype models - ones where the distribution of the entries is especially simple and allows some kind of exact solutions (it's perhaps a small miracle that such models even exist). Then one might be able to deduce qualitative properties of more general cases if it can be argued that the general ones can be approximated by the simpler ones.

Historically, this is perhaps how at least some parts of random matrix theory has developed: results are first proven for some prototype models, and later it has been noticed that some of these results are universal - can be extended to far more general ones. We will focus very heavily on the prototype models in these lectures.

**1.1. Definition of the GOE.** Perhaps the best known random matrix model is the one known as the Gaussian Orthogonal Ensemble (GOE) which is a model for random symmetric matrices (i.e. matrices  $A$  satisfying  $A^T = A$ ) whose entries are independent (up to the symmetricity constraint) normal random variables. The main reason for studying this particular model is that the distribution of its eigenvalues is very explicit and can be analyzed quite efficiently.

Let  $(H_{ii})_{i=1}^{\infty}$  be i.i.d. (that is independent and identically distributed) Gaussian random variables with zero expectation and variance 2 (i.e.  $H_{ii} \sim N(0, 2)$ ) and let  $(H_{ij})_{1 \leq i < j < \infty}$  be i.i.d. standard Gaussian random variables, i.e.  $H_{ij} \sim N(0, 1)$ , and assume that they are independent of  $(H_{ii})$ .

Let  $N$  be a positive integer. Then the probability distribution of the random (symmetric) matrix

$$(1) \quad \mathcal{H}_N = \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{12} & H_{22} & \cdots & H_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{1N} & H_{2N} & \cdots & H_{NN} \end{pmatrix}$$

is called the Gaussian Orthogonal Ensemble (GOE or  $\text{GOE}(N)$  for short).

The reason for having a different variance on the diagonal is that we can express the probability distribution of  $\mathcal{H}_N$  (that is the joint distribution of its entries) in a particularly useful way (the usefulness will become apparent when we begin the study of the eigenvalues).

**Lemma 1.** *The joint distribution of the elements  $(H_{i,j})_{1 \leq i \leq j \leq N}$  can be written as*

$$(2) \quad C_N e^{-\frac{1}{4} \text{Tr} H^2} dH,$$

where  $C_N$  is a normalizing constant whose precise value is not important to us,  $\text{Tr}$  denotes the trace of a matrix - i.e. the sum of its diagonal entries, or equivalently the sum of its eigenvalues. Here  $dH$  means

$$(3) \quad dH = \prod_{i \leq j} dH_{ij}.$$

To be more specific, what this means is for example if  $f$  is any say bounded continuous function defined on the space of  $N \times N$  symmetric matrices, then if we write  $\mathbb{E}f(\mathcal{H}_N)$  for the expectation of  $f(\mathcal{H}_N)$ , then

$$(4) \quad \mathbb{E}f(\mathcal{H}_N) = C_N \int f(H) e^{-\frac{1}{4} \text{Tr} H^2} dH.$$

In some situations this is more convenient than simply expressing the joint distribution of the entries as a product of Gaussian distributions.

*Proof of Lemma 1.* Recall that our entries are independent so their joint distribution is simply the product of their distributions. So their joint distribution is

$$(5) \quad \begin{aligned} & \prod_{i=1}^N \frac{1}{\sqrt{4\pi}} e^{-\frac{1}{4} H_{ii}^2} \prod_{1 \leq i < j \leq N} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} H_{ij}^2} \\ &= \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N}{2}} (2\pi)^{-\frac{N(N-1)}{4}} e^{-\frac{1}{4} [\sum_{i=1}^N H_{ii}^2 + 2 \sum_{1 \leq i < j \leq N} H_{ij}^2]} \\ &= \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N(N+1)}{4}} e^{-\frac{1}{4} [\sum_{i=1}^N H_{ii}^2 + \sum_{i \neq j} H_{ij}^2]} \\ &= \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N(N+1)}{4}} e^{-\frac{1}{4} [\sum_{i,j=1}^N H_{ij}^2]} \\ &= \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N(N+1)}{4}} e^{-\frac{1}{4} \sum_{i=1}^N [\sum_{j=1}^N H_{ij} H_{ji}]} \\ &= \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N(N+1)}{4}} e^{-\frac{1}{4} \sum_{i=1}^N (H^2)_{ii}} \\ &= \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N(N+1)}{4}} e^{-\frac{1}{4} \text{Tr}(H^2)}. \end{aligned}$$

In the first step we simply took the constants out of the product and changed a product of exponentials into an exponential of sums. In the second step we used the fact that  $H_{ij} = H_{ji}$  so that  $\sum_{i < j} H_{ij}^2 = \sum_{i > j} H_{ij}^2 =$

$\frac{1}{2} \sum_{i \neq j} H_{ij}^2$ . We then combined the two sums and interpreted  $\sum_{i,j} H_{ij}^2$  as  $\text{Tr}H^2$ . We also note that the constant is

$$(6) \quad C_N = \frac{1}{\sqrt{2}} (2\pi)^{-\frac{N(N+1)}{4}}.$$

In calculating it, we used the fact that the number of pairs  $(i, j)$  with  $1 \leq i < j \leq N$  is  $N(N-1)/2$ .  $\square$

One reason the GOE is so special is that it has a high degree of symmetry. Namely it is invariant under conjugation by orthogonal matrices. More precisely, if  $\mathcal{H}_N$  is a GOE matrix and  $O_N$  is a  $N \times N$  orthogonal matrix, then the distribution of  $O_N^T \mathcal{H}_N O_N$  is the same as the distribution of  $\mathcal{H}_N$ . This follows from the Gaussian nature of the entries, as well as the way we chose the variances. Also it is why the name of the GOE has the O (for invariance under orthogonal transformations) in it. We'll prove this now.

**Lemma 2.** *Let  $\mathcal{H}_N$  be a  $GOE(N)$  matrix, and let  $O_N$  be a non-random  $N \times N$  orthogonal matrix. Then the distribution of  $\mathcal{H}_N$  is the same as the distribution of  $O_N \mathcal{H}_N O_N^T$ .*

*Proof.* The claim essentially says that for any nice enough function  $f$  defined on the space of  $N \times N$  symmetric matrices (say continuous and bounded)  $\mathbb{E}f(O_N \mathcal{H}_N O_N^T) = \mathbb{E}f(\mathcal{H}_N)$ . Let us thus write the first expectation in integral form and try to transform it into the integral form of the latter one.

Using the facts that  $\text{Tr}(AB) = \text{Tr}(BA)$  and  $O_N^T O_N = I$ , where  $I$  is the  $N \times N$  identity matrix (this essentially the definition of an orthogonal matrix), we see that  $\text{Tr}[O_N H O_N^T]^2 = \text{Tr}O_N H^2 O_N^T = \text{Tr}H^2$ . Thus we only need to check that  $d[O_N H O_N^T] = dH$ , or more precisely, the absolute value of the Jacobian determinant from the transformation  $H \mapsto O_N H O_N^T$  is one.

Let us write down explicitly the entries of  $O_N H O_N^T$ :

$$(7) \quad (O_N H O_N^T)_{ij} = \sum_{k,l=1}^N O_{N,ik} H_{kl} (O_N^T)_{lj} = \sum_{k,l=1}^N O_{N,ik} O_{N,jl} H_{kl}.$$

The Jacobian matrix of this transformation will be a matrix whose columns (labelled by pairs  $(i, j)$  with  $i \leq j$ ) are derivatives of  $(O_N H O_N^T)_{ij}$  with respect to the variables  $H_{kl}$  (the pairs  $(k, l)$  with  $k \leq l$  labelling the rows of the matrix). These derivatives are simple to calculate from (7):

$$(8) \quad \frac{\partial}{\partial H_{kl}} (O_N H O_N^T)_{ij} = O_{N,ik} O_{N,jl}.$$

If we write  $J(O_N)$  for the Jacobian matrix, we see that its entries are the following

$$(9) \quad J(O_N)_{(i,j),(k,l)} = O_{N,ik} O_{N,jl}.$$

Note that  $\det J(O_N)$  really just depends on  $O_N$ , and not on  $H$  in any way. Our goal is to see that  $|\det J(O_N)| = 1$  for all orthogonal matrices  $O_N$ .

The first ingredient we'll need for proving this is that if we perform two changes of variables, then by the product rule of differentiation, the Jacobian of the combination of the two change of variables will be the product of the two individual ones. In our case, this means that if  $U_N$  is another orthogonal matrix, then  $J(O_N U_N) = J(O_N)J(U_N)$  so  $\det(J(O_N U_N)) = \det J(O_N) \det J(U_N)$ .

The next thing to note is that

$$\begin{aligned}
 (10) \quad [J(O_N)^T]_{(i,j),(k,l)} &= J(O_N)_{(k,l),(i,j)} \\
 &= O_{N,ki} O_{N,lj} \\
 &= (O_N^T)_{ik} (O_N^T)_{jl} \\
 &= J(O_N^T)_{(i,j),(k,l)},
 \end{aligned}$$

so that  $J(O_N)^T = J(O_N^T)$ . As transposing a matrix does not affect its determinant, we see that  $\det J(O_N^T) = \det J(O_N)$ .

Finally we note that of course if we perform no change of variables, then the Jacobian matrix is simply the identity matrix:  $J(I) = I$  and  $\det J(I) = 1$ . Putting these remarks together, we see the following:

$$\begin{aligned}
 (11) \quad [\det J(O_N)]^2 &= \det J(O_N) \det J(O_N^T) \\
 &= \det J(O_N O_N^T) \\
 &= \det J(I) \\
 &= 1.
 \end{aligned}$$

Thus  $|\det J(O_N)| = 1$ . □

A fact we'll not show here is that the only symmetric random matrix whose entries are independent (up to the symmetricity constraint) and whose distribution is invariant under orthogonal transformations is the GOE - so the orthogonal invariance is a very special property of the Gaussian nature of the matrices.

## 2. EIGENVALUE DISTRIBUTION OF THE GOE

The main goal of this section is to prove the following theorem which gives an explicit formula for the distribution of the eigenvalues of a GOE matrix.

**Theorem 3.** *Let  $\mathcal{H}_N$  be a  $GOE(N)$  matrix. Let  $(\lambda_1^{(N)}, \dots, \lambda_N^{(N)})$  be the eigenvalues of  $\mathcal{H}_N$ . The probability distribution of  $(\lambda_1^{(N)}, \dots, \lambda_N^{(N)})$  (on  $\mathbb{R}^N$ ) is*

$$(12) \quad \frac{1}{Z_N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j| \prod_{j=1}^N e^{-\frac{1}{4}\lambda_j^2} d\lambda_j,$$

where  $Z_N$  is a normalization constant we won't calculate (though it can be calculated explicitly by making use of so-called Selberg integrals).

The proof is rather involved and in fact there are a couple things we'll skip and refer to the book [1] on. The proof itself is not that important to what will come later in the course but the result is fundamental. Working out the case  $N = 2$  explicitly might be an instructive exercise.

*Proof.* To set the stage, note that we want to prove the for a bounded continuous function  $f : \mathbb{R}^N \rightarrow \mathbb{R}$ ,

$$(13) \quad \mathbb{E}f\left(\lambda_1^{(N)}, \dots, \lambda_N^{(N)}\right) = \frac{1}{Z_N} \int_{\mathbb{R}^N} f(\lambda_1, \dots, \lambda_N) \prod_{i < j} |\lambda_i - \lambda_j| \prod_{j=1}^N e^{-\frac{1}{4}\lambda_j^2} d\lambda_j.$$

To do this, we point out that  $e^{-\frac{1}{4}\text{Tr}\mathcal{H}_N^2} = e^{-\frac{1}{4}\sum_j(\lambda_j^{(N)})^2}$  so what we'll need to show is that  $dH$  can actually be written in a form where such a product over the eigenvalues appears. To do this, we'll want to write a generic symmetric matrix  $H$  in terms of its eigenvalues and some other parameters. We then make a corresponding change of variables in the integral, and the product will come out from calculating the Jacobian determinant. There are some difficulties because this change of variables will not be valid on the whole space of symmetric matrices, but we'll have to restrict to a subset of it, but it turns out that this subset is almost the whole space in the sense that the probability of it is one. We begin by studying this representation of  $H$  in terms of its eigenvalues and other parameters. We then calculate the Jacobian, and refer to [1] about the issues with the change of variables not being valid on the entire space.

The change of variables we'll want to make use of is essentially the one coming from diagonalizing  $\mathcal{H}_N$ . The problem with this is that if we want to change the integration variables from  $H$  to  $(O, \Lambda)$ , where  $O$  is orthogonal,  $\Lambda$  is diagonal, and  $H = O^T \Lambda O$ , then this mapping is not a bijection (one to one and onto). When changing integration variables, this is important to avoid overcounting and to make sure that it also is valid everywhere in the region of integration. Also this mapping may not be smooth on the whole space of symmetric matrices, so the Jacobian might not be well defined.

The idea is to then to restrict to a subset of the space of symmetric matrices where this type of change of variables is unique and smooth. On this set, it'll turn out that the Jacobian factorizes nicely into a  $\Lambda$ -dependent part and an  $O$ -dependent part which will give us our claim once we show that the event that  $\mathcal{H}_N^{(1)}$  is not in this good subset has probability zero.

Consider now what kind of non-uniqueness and smoothness problems one might have in the mapping  $H \mapsto (\Lambda, O)$  where  $H = O^T \Lambda O$ . First of all, recall that questions about eigenvectors and eigenvalues are simpler if all of the eigenvalues are distinct. It's perhaps natural to expect that an event where two eigenvalues are equal would have probability zero as this set is of lower dimension than the full set. For a simpler analogue of such a question, consider a random vector  $(X, Y)$  where  $X$  and  $Y$  are independent normal random variables. Then the event  $X = Y$  will have probability zero. While in our case there are more complicated correlations, it might still seem

reasonable to expect such a situation. We'll be more precise about this later on. Let us assume for now that all of the eigenvalues are distinct. We recall from linear algebra that we can write our matrix  $H$  as

$$(14) \quad H = \sum_{j=1}^N \lambda_j v_j v_j^T = O^T \Lambda O$$

where  $Hv_j = \lambda_j v_j$ ,  $v_j$  are normalized to be of unit length,  $O$  is the matrix whose rows are  $v_j^T$  and  $\Lambda$  is the diagonal matrix whose entries are  $\lambda_j$  (in the same order as the rows of  $O$ ). Now any matrix which is obtained by permuting simultaneously the  $\lambda_j$ s and  $v_j$ s results in the same matrix  $H$ . Thus we have  $N!$  such representations. Let us thus fix one representation by demanding that if  $\Lambda_{ij} = \lambda_i \delta_{ij}$ , then  $\lambda_1 > \lambda_2 > \dots > \lambda_N$ .

We also point out that if we replace  $v_j$  by  $-v_j$ , the matrix  $H = \sum_j \lambda_j v_j v_j^T$  does not change. Moreover, the event that say one of the diagonal elements of  $O$  is zero would seem again to have probability zero as the set is lower dimensional than the full set. Thus let us also require that all of the diagonal entries of  $O$  are positive.

This set will not quite be sufficient for us, but let us add further constraints as we run into problems. So we define the set  $A_N$  to be the set of symmetric matrices  $H$  which can be written in the form  $H = O^T \Lambda O$ , where  $O$  is orthogonal with strictly positive diagonal entries, and  $\Lambda$  is a diagonal matrix satisfying  $\Lambda_{11} > \Lambda_{22} > \dots > \Lambda_{NN}$ .

Let us now try to parametrize the matrix  $O$  appearing here - it has  $N \times N$  entries, but these satisfy lots of constraints as the rows of  $O$  are orthonormal vectors, so let us try to find a way to parametrize these constraints (express the entries of  $O$  in terms of parameters without any constraints). Let us look at the first row of  $O$ . We write it as

$$(15) \quad v_1^T = (O_{11} \ O_{12} \ \dots \ O_{1N}) = O_{11}(1 \ p_{12} \ \dots \ p_{1N}).$$

As this is a unit vector, we have  $O_{11}^2 \left[ 1 + \sum_{j=2}^N p_{1j}^2 \right] = 1$  so we can solve for  $O_{11}$  in terms of the  $p_{1j}$ . As we demanded that  $O_{11} > 0$ , we have

$$(16) \quad O_{11} = \frac{1}{\sqrt{1 + \sum_{j=2}^N p_{1j}^2}}.$$

We'll want to view the variables  $p_{1j}$  as the "unconstrained" real parameters - note that each choice of them produces a unique unit vector  $v_1^T$  and  $v_1^T$  depends smoothly on the parameters. Also as  $p_{1j}$  run through  $\mathbb{R}$ , the first column of  $O$  runs through the space of unit vectors whose first entry is positive. Let us then look at the second row of  $O$ . We write

$$(17) \quad v_2^T = (O_{21} \ O_{22} \ \dots \ O_{2N}) = O_{22}(q_{21} \ 1 \ p_{23} \ \dots \ p_{2N}).$$

Here we have different notation for the entry before the 1 in the vector as we'll not interpret it as an independent parameter but something depending on the  $p_{ij}$ . From the condition that  $v_2^T$  is a unit vector and  $O_{22} > 0$ , we find

$$(18) \quad O_{22} = \frac{1}{\sqrt{1 + q_{21}^2 + \sum_{j=3}^N p_{2j}^2}}.$$

From the condition that  $v_1^T$  and  $v_2^T$  are orthogonal, we can solve

$$(19) \quad q_{21} = - \left[ p_{12} + \sum_{j=3}^N p_{1j} p_{2j} \right].$$

Thus each  $\{p_{ij} : i \in \{1, 2\}, j > i\}$  produces a unique pair of orthonormal vectors  $v_1^T$  and  $v_2^T$ , and these vectors depend smoothly on the parameters  $p_{ij}$ .

We keep doing this: write

$$(20) \quad v_k^T = (O_{k1} \cdots O_{kN}) = O_{kk}(q_{k1} \cdots q_{k,k-1} \ 1 \ p_{k,k+1} \cdots p_{k,N}).$$

The idea is then to solve for  $q_{k1}, \dots, q_{k,k-1}$  in terms of the parameters  $\{p_{ij} : i \leq k-1, j > i\}$  from the conditions that  $v_k^T$  is orthogonal to  $v_j^T$  for  $j \leq k-1$ .  $O_{kk}$  is expressed in terms of everything else from the condition that  $v_k^T$  is a unit vector. We encounter a problem here as nothing ensures that the linear system of equations that  $q_{k1}, \dots, q_{k,k-1}$  satisfies has a solution. But it will turn out that the event where the system of equations does not have a solution is again "lower dimensional" and has zero probability. Let us look more carefully at this linear system of equations. The condition that  $v_k^T$  is orthogonal to  $v_j^T$  for  $j < k$  can be written as (we choose to write the entries of  $v_j^T$  in terms of  $O$  as this will allow us to formulate our final condition in terms of  $O$  which is more convenient)

$$(21) \quad \sum_{l=1}^{k-1} O_{jl} q_{kl} + O_{j,k} + \sum_{l=k+1}^N O_{jl} p_{kl} = 0.$$

In matrix form, we write this as

$$(22) \quad \begin{pmatrix} O_{11} & O_{12} & \cdots & O_{1,k-1} \\ O_{21} & O_{22} & \cdots & O_{2,k-1} \\ \vdots & \vdots & \cdots & \vdots \\ O_{k-1,1} & O_{k-1,2} & \cdots & O_{k-1,k-1} \end{pmatrix} \begin{pmatrix} q_{k1} \\ q_{k2} \\ \vdots \\ q_{k,k-1} \end{pmatrix} \\ = - \begin{pmatrix} O_{1k} + \sum_{l=k+1}^N O_{1l} p_{kl} \\ O_{2k} + \sum_{l=k+1}^N O_{2l} p_{kl} \\ \vdots \\ O_{k-1,k} + \sum_{l=k+1}^N O_{k-1,l} p_{kl} \end{pmatrix}$$

Thus the condition that  $q_{k1}, \dots, q_{k,k-1}$  can be uniquely expressed in terms of  $(O_{ij})_{i < k}$  for all  $k$  is that the determinant of the matrix

$$(23) \quad \begin{pmatrix} O_{11} & O_{12} & \cdots & O_{1,k-1} \\ O_{21} & O_{22} & \cdots & O_{2,k-1} \\ \vdots & \vdots & \cdots & \vdots \\ O_{k-1,1} & O_{k-1,2} & \cdots & O_{k-1,k-1} \end{pmatrix}$$

is non-zero for all  $k$ . This again seems like a constraint which is satisfied with probability one as the determinant being zero is a set of lower dimension. Let us now define  $B_N$  to be the set of symmetric matrices  $H$  which can be written as  $H = O^T \Lambda O$ , where  $\Lambda$  is diagonal and satisfies  $\Lambda_{ii} > \Lambda_{i+1,i+1}$  for all  $i$ , and  $O$  is orthogonal and satisfies:  $O_{ii} > 0$  for all  $i$ , and  $\det(O_{ij})_{i,j=1}^k \neq 0$  for all  $k = 1, \dots, N$ . On this set, we see (inductively) that  $O$  can be written in terms of the parameters  $(p_{ij})_{i < j}$ , and let us write  $\tilde{B}_N$  for the subset of  $\mathbb{R}^{N(N-1)/2}$  which is the projection of  $B_N$  under the bijection relating  $O$  and  $(p_{ij})_{i < j}$ .

Recall that our goal is to calculate the Jacobian determinant which consists of partial derivatives so we'll need to know that  $O$  depends smoothly on the  $p_{ij}$ . This can again be argued recursively. Note  $O_{11}$  depends smoothly on the  $p_{1j}$ , so the first row of  $O$  depends smoothly on the  $p_{ij}$ . Now consider the  $k$ th row and assume that all the rows above it depend smoothly on the  $p_{ij}$ . From the linear system the  $q_{kj}$  satisfy, we see that they depend only on things on rows above the  $k$ th row, and on the  $p_{kj}$ . Moreover, from Cramer's rule, we see that the only problems with differentiability would occur if the determinant of the matrix  $(O_{ij})_{i,j \leq k-1}$  vanished but we restricted to the set  $B_N$  where this does not happen. Thus the  $q_{kj}$  are smooth and we see by induction that this is true for all rows.

Let us write  $O(p)$  for this function from  $\tilde{B}_N$  to the space of  $N \times N$  orthogonal matrices. Also let  $\Delta_N = \{\lambda \in \mathbb{R}^N : \lambda_1 > \lambda_2 > \dots > \lambda_N\}$ . We then define the function  $H : \tilde{B}_N \times \Delta_N$  to the space of symmetric matrices by

$$(24) \quad H(p, \lambda) = O(p)^T \Lambda(\lambda) O(p),$$

where  $\Lambda(\lambda)_{ij} = \lambda_i \delta_{ij}$ . Let us now begin calculating the Jacobian of the change of variables  $H = H(p, \lambda)$ . We thus want to calculate  $\partial_{p_{ij}} H(p, \lambda)$  for all  $i < j$  and  $\partial_{\lambda_i} H(p, \lambda)$ . Let us start with the latter one which is simpler. We have



$$\begin{aligned}
(25) \quad [\partial_{\lambda_i} H(p, \lambda)]_{k,l} &= [O(p)^T \partial_{\lambda_i} \Lambda(\lambda) O(p)]_{k,l} \\
&= \sum_{r,s=1}^N O_{rk}(p) \partial_{\lambda_i} \Lambda_{rs}(\lambda) O_{sl}(p) \\
&= \sum_{r,s=1}^N O_{rk}(p) \delta_{ri} \delta_{si} O_{sl}(p) \\
&= O_{ik}(p) O_{il}(p).
\end{aligned}$$

We'll find it a bit more convenient to express the derivatives conjugated by  $O$  - using the orthogonality of  $O$  (i.e.  $O^T O = O O^T = I$ )

$$(26) \quad [O(p) \partial_{\lambda_i} H(p, \lambda) O(p)^T]_{kl} = \delta_{ki} \delta_{li}.$$

Consider then the  $p$ -derivatives. First of all, we have

$$(27) \quad \partial_{p_{ij}} H(p, \lambda) = [\partial_{p_{ij}} O(p)]^T \Lambda(\lambda) O(p) + O(p)^T \Lambda(\lambda) [\partial_{p_{ij}} O(p)]$$

so if we multiply from the left by  $O(p)$  and right by  $O(p)^T$ , we find

$$(28) \quad O(p) \partial_{p_{ij}} H(p, \lambda) O(p)^T = [O(p) \partial_{p_{ij}} O(p)^T] \Lambda(\lambda) + \Lambda(\lambda) [O(p) \partial_{p_{ij}} O(p)^T].$$

Using orthogonality of  $O$  (i.e. as  $O(p) O(p)^T = I$ , its derivative vanishes) we find by the product rule of differentiation that

$$(29) \quad \partial_{p_{ij}} O(p) O(p)^T = -O(p) \partial_{p_{ij}} O(p)^T$$

which implies that

$$(30) \quad O(p) \partial_{p_{ij}} H(p, \lambda) O(p)^T = [O(p) \partial_{p_{ij}} O(p)^T] \Lambda(\lambda) - \Lambda(\lambda) [O(p) \partial_{p_{ij}} O(p)^T].$$

In component form this is:

$$\begin{aligned}
(31) \quad (O(p) \partial_{p_{ij}} H(p, \lambda) O(p)^T)_{kl} &= \sum_{r=1}^N [O(p) \partial_{p_{ij}} O(p)^T]_{kr} \Lambda_{rl}(\lambda) \\
&\quad - \sum_{r=1}^N \Lambda_{kr}(\lambda) [O(p) \partial_{p_{ij}} O(p)^T]_{rl} \\
&= (\lambda_l - \lambda_k) [O(p) \partial_{p_{ij}} O(p)^T]_{kl}.
\end{aligned}$$

Let us now organize the array  $(p_{ij})_{i < j}$  into a sequence  $(\pi_l)_{l=1}^{N(N-1)/2}$  by enumerating the pairs  $(i, j)$  in some arbitrary way that we'll fix from now on. We then define for  $k = 1, \dots, N(N+1)/2$

$$(32) \quad z_k = \begin{cases} \lambda_k, & k \leq N \\ \pi_{k-N}, & k > N \end{cases}.$$

The Jacobian matrix of our transformation is then

$$(33) \quad J(\lambda, p) = \begin{pmatrix} \frac{\partial H_{11}}{\partial z_1} & \cdots & \frac{\partial H_{NN}}{\partial z_1} & \frac{\partial H_{12}}{\partial z_1} & \cdots & \frac{\partial H_{N-1,N}}{\partial z_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial H_{11}}{\partial z_{N(N+1)/2}} & \cdots & \frac{\partial H_{NN}}{\partial z_{N(N+1)/2}} & \frac{\partial H_{12}}{\partial z_{N(N+1)/2}} & \cdots & \frac{\partial H_{N-1,N}}{\partial z_{N(N+1)/2}} \end{pmatrix}$$

Note that if we write  $Q^{(r)} = O(p)\partial_{z_r}H(\lambda, p)O(p)^T$ , then

$$(34) \quad \frac{\partial H_{ij}}{\partial z_r} = (O(p)^T Q^{(r)} O(p))_{ij} = \sum_{l,m=1}^N O_{li} O_{mj} Q_{lm}^{(r)},$$

and by (26) and (31) we have

$$(35) \quad Q_{lm}^{(r)} = \begin{cases} \delta_{lr}\delta_{mr}, & r \leq N \\ (\lambda_m - \lambda_l)[O(p)\partial_{z_r}O(p)^T]_{lm}, & r \geq N \end{cases}.$$

We point out that for each  $r$ ,  $Q_{lm}^{(r)} = Q_{ml}^{(r)}$ . Now note that each column of  $J$  can be written as

$$(36) \quad \begin{pmatrix} \partial_{z_1} H_{ij} \\ \partial_{z_2} H_{ij} \\ \vdots \\ \partial_{z_{N(N+1)/2}} H_{ij} \end{pmatrix} = \sum_{l,m=1}^N O_{li} O_{mj} \begin{pmatrix} Q_{lm}^{(1)} \\ Q_{lm}^{(2)} \\ \vdots \\ Q_{lm}^{(N(N+1)/2)} \end{pmatrix}.$$

Now by the multilinearity of the determinant, we can express each column this way and then pull out the sums and  $O$ -terms out of the determinant to find

$$(37)$$

$\det J(\lambda, p)$

$$(38) \quad = \sum_{l_{(1,1)}, m_{(1,1)}} O_{l_{(1,1)},1} O_{m_{(1,1)},1} \cdots \sum_{l_{(N-1,N)}, m_{(N-1,N)}} O_{l_{(N-1,N)},N-1} O_{m_{(N-1,N)},N} \\ \times \begin{vmatrix} Q_{l_{(1,1)},m_{(1,1)}}^{(1)} & \cdots & Q_{l_{(N,N)},m_{(N,N)}}^{(1)} & Q_{l_{(1,2)},m_{(1,2)}}^{(1)} & \cdots & Q_{l_{(N-1,N)},m_{(N-1,N)}}^{(1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Q_{l_{(1,1)},m_{(1,1)}}^{(N(N+1)/2)} & \cdots & Q_{l_{(N,N)},m_{(N,N)}}^{(N(N+1)/2)} & Q_{l_{(1,2)},m_{(1,2)}}^{(N(N+1)/2)} & \cdots & Q_{l_{(N-1,N)},m_{(N-1,N)}}^{(N(N+1)/2)} \end{vmatrix}$$

If any of the  $(l, m)$  indices on two distinct columns are equal, then the determinant here vanishes. Moreover, as  $Q_{lm}^{(r)} = Q_{ml}^{(r)}$ , then if for two different columns (indexed by say  $(l, m)$  and  $(l', m')$ , we have  $(l, m) = (m', l')$  then also the determinant vanishes. If we consider unordered pairs  $\{l, m\}$ , then for  $1 \leq l, m \leq N$ , there are precisely  $N(N+1)/2$  of these. This is exactly the same amount of columns as the determinant we are interested in. So up to permuting the columns, there is only a single way we can assign values to the indices  $(l, m)$  without the determinant vanishing, more precisely, we have

$$(39) \quad \begin{aligned} & \begin{vmatrix} Q_{l(1,1),m(1,1)}^{(1)} & \cdots & Q_{l(N,N),m(N,N)}^{(1)} & Q_{l(1,2),m(1,2)}^{(1)} & \cdots & Q_{l(N-1,N),m(N-1,N)}^{(1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Q_{l(1,1),m(1,1)}^{(N(N+1)/2)} & \cdots & Q_{l(N,N),m(N,N)}^{(N(N+1)/2)} & Q_{l(1,2),m(1,2)}^{(N(N+1)/2)} & \cdots & Q_{l(N-1,N),m(N-1,N)}^{(N(N+1)/2)} \end{vmatrix} \\ &= \alpha_{\mathbf{1}, \mathbf{m}} \begin{vmatrix} Q_{1,1}^{(1)} & \cdots & Q_{N,N}^{(1)} & Q_{1,2}^{(1)} & \cdots & Q_{N-1,N}^{(1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Q_{1,1}^{(N(N+1)/2)} & \cdots & Q_{N,N}^{(N(N+1)/2)} & Q_{1,2}^{(N(N+1)/2)} & \cdots & Q_{N-1,N}^{(N(N+1)/2)} \end{vmatrix}, \end{aligned}$$

where  $\alpha_{\mathbf{1}, \mathbf{m}}$  is zero if a pair of indices was the same, or otherwise  $\pm 1$  (with the sign depending on how we had to permute the rows). The determinant is now independent of the summation indices so we can take it out of the sum. Also it is the only thing that carries the  $\lambda$ -dependence so our task is to calculate this determinant. To do this, note that the structure of  $Q_{lm}^{(r)}$  implies that this determinant can be written in the form

$$(40) \quad \begin{vmatrix} I_{N \times N} & 0 \\ * & D_{N(N-1)/2} \end{vmatrix},$$

where  $I_{N \times N}$  is the  $N \times N$  identity matrix,  $0$  is a matrix of  $N$  rows and  $N(N-1)/2$  columns filled with zeroes,  $*$  is a matrix of  $N(N-1)/2$  columns and  $N$  rows that we won't care about  $D_{N(N-1)/2}$  is a matrix whose entries are of the form  $Q_{ij}^{(r)}$ , where  $i < j$  and  $r > N$ . Due to this structure, the determinant of the whole matrix actually equals  $\det(D_{N(N-1)/2})$ . Note that the  $(i, j)$  row of  $D_{N(N-1)/2}$  is proportional to  $\lambda_i - \lambda_j$ , so we can pull it out of the determinant for all  $i < j$  and as this leaves no  $\lambda$ -dependence in the determinant, we find

$$(41) \quad \det(D_{N(N-1)/2}) = \prod_{i < j} (\lambda_i - \lambda_j) F(p).$$

We conclude that we can actually write

$$(42) \quad \det J(\lambda, p) = \prod_{i < j} (\lambda_i - \lambda_j) \tilde{F}(p),$$

for some function  $\tilde{F}$ .

To complete our proof, we need to know that the event  $B_N$  (and equivalently  $\tilde{B}_N$ ) has probability one. For this, we refer to similar results in [1]: Lemma 2.5.5, Lemma 2.5.6, and Lemma 2.5.7. Using these results we see that

$$(43) \quad \mathbb{E}f\left(\lambda_1^{(N)}, \dots, \lambda_N^{(N)}\right) = \int_{B_N} f(\lambda_1, \dots, \lambda_N) C_N e^{-\frac{1}{4}\text{Tr}H^2} dH \\ + \int_{B_N^C} f(\lambda_1, \dots, \lambda_N) C_N e^{-\frac{1}{4}\text{Tr}H^2} dH,$$

and the latter integral is zero because it is bounded by a constant times the probability of the event  $B_N$  (the function  $f$  is bounded).

Recall now that originally, when we wanted to calculate  $\mathbb{E}f(\lambda_1^{(N)}, \dots, \lambda_N^{(N)})$ , we assumed nothing about the ordering of the eigenvalues, but our calculation assumes that they are ordered. That being said, our argument would have worked perfectly well for any ordering of the eigenvalues. The absolute value of the Jacobian determinant would have been again of the form  $\prod_{i<j} |\lambda_i - \lambda_j| G(p)$  for some non-negative function  $G$  which can depend on the ordering. Also the suitable set  $\tilde{B}_N$  would depend on the ordering. We conclude that

$$(44) \quad \mathbb{E}f\left(\lambda_1^{(N)}, \dots, \lambda_N^{(N)}\right) = \frac{1}{Z_N} \int_{\mathbb{R}^N} f(\lambda_1, \dots, \lambda_N) \prod_{i<j} |\lambda_i - \lambda_j| e^{-\frac{1}{4}\sum_{j=1}^N \lambda_j^2} \prod_{j=1}^N d\lambda_j,$$

where the constant  $1/Z_N$  comes from summing the the  $p$ -integrals for different orderings and the constant  $C_N$  in the distribution of  $\mathcal{H}_N$ . Also we could forget the condition about the eigenvalues being distinct because the integral over the set where at least two are equal vanishes because the corresponding set is not  $N$ -dimensional.  $\square$

**Remark 4.** *As we saw that when the eigenvalues were ordered, the distribution of the whole matrix factored into the product of the distribution of the eigenvalues and the distribution of the eigenvectors, we see that the eigenvalues and eigenvectors are independent. Using the invariance under orthogonal matrices, it wouldn't be hard to check that the law of the eigenvectors is essentially given by the so called Haar measure on the orthogonal group, but we don't assume this to be a familiar concept to all of the students following the course, so we don't go into further detail. For further details, see [1, Corollary 2.5.4].*

**Remark 5.** *Note that our proof would work just as well if the original distribution was of the form  $e^{-\text{Tr}V(H)} dH$ , where  $V$  is some nice enough function (here  $\text{Tr}V(H)$  means  $\sum_j f(\lambda_j)$ ).*

## REFERENCES

- [1] G.W. Anderson, A. Guionnet, and O. Zeitouni: *An introduction to random matrices. Cambridge Studies in Advanced Mathematics, 118.* Cambridge University Press, Cambridge, 2010.