A PROBABILISTIC COMPUTATION OF A MEHTA INTEGRAL

LIVIU I. NICOLAESCU

ABSTRACT. We use the Kac-Rice formula to compute the Mehta integral describing the normalization constant arising in the statistics of the Gaussian Orthogonal Ensemble.

CONTENTS

1. INTRODUCTION

We denote by $\mathbf{Sym}(\mathbb{R}^m)$ the space of real symmetric $m \times m$ matrices. This is a Euclidean space with respect to the inner product $(A,B) := \operatorname{tr}(AB)$. This inner product is invariant with respect to the action of the orthogonal group $O(m)$ on $\mathbf{Sym}(\mathbb{R}^m)$.

We define

$$\ell_{ij}, \omega_{ij} : \mathbf{Sym}(\mathbb{R}^m) \to \mathbb{R}, \;\; \ell_{ij}(A) = a_{ij}, \;\; \omega_{ij}(A) := \begin{cases} a_{ij}, & i = j, \\ \sqrt{2}a_{ij}, & i < j. \end{cases} \tag{1.1}$$

The collection $(\omega_{ij})_{i \leq j}$ defines linear coordinates on $\mathbf{Sym}(\mathbb{R}^m)$ that are orthonormal with respect to the above inner product on $\mathbf{Sym}(\mathbb{R}^m)$. The volume density induced by this metric is

$$\operatorname{vol}\big[\, dA \,\big] := \prod_{i \leq j} d\omega_{ij} = 2^{\frac{1}{2}\binom{m}{2}} \prod_{i \leq j} d\ell_{ij}.$$

For any real numbers u, v such that

$$v > 0, mu + 2v > 0, \tag{1.2}$$

we denote by $\mathcal{S}_m^{u,v}$ the space $\mathbf{Sym}(\mathbb{R}^m)$ equipped with the centered Gaussian measure $\mathbf{\Gamma}_{u,v}\big[\, dA \,\big]$ uniquely determined by the covariance equalities

$$\mathbb{E}\big[\, \ell_{ij}(A)\ell_{k\ell}(A) \,\big] = u\delta_{ij}\delta_{k\ell} + v(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}), \;\; \forall 1 \leq i, j, \; k, \ell \leq m. \tag{1.3}$$

In particular we have

$$\mathbb{E}\big[\, \ell_{ii}^2 \,\big] = u + 2v, \;\; \mathbb{E}\big[\, \ell_{ii}\ell_{jj} \,\big] = u, \quad \mathbb{E}\big[\, \ell_{ij}^2 \,\big] = v, \;\; \forall 1 \leq i \neq j \leq m, \tag{1.4}$$

while all other covariances are trivial. The ensemble $\mathcal{S}^{0,v}$ is a rescaled version of the Gaussian Orthogonal Ensemble (GOE) and we will refer to it as GOE_m^v. The inequalities (1.2) guarantee that the covariance form defined by (1.3) is positive definite so that $\mathbf{\Gamma}_{u,v}$ is nondegenerate.

Date: Last revised August 12, 2024.

Key words and phrases. random matrices, Kac-Rice formula.

For $u>0$ the ensemble $\mathcal{S}_m^{u,v}$ can be given an alternate description. More precisely a random $A\in\mathcal{S}_m^{u,v}$ can be described as a sum

$$A=B+\;X\mathbb{1}_m,\;\;B\in\mathrm{GOE}_m^v,\;\;X\in\boldsymbol{N}(0,u),\;\;B\text{ and }X\text{ independent}.$$

We write this

$$\mathcal{S}_m^{u,v}=\mathrm{GOE}_m^v\,\hat{+}N(0,u)\mathbb{1}_m, \tag{1.5}$$

where $\hat{+}$ indicates a sum of *independent* variables.

In the special case GOE_m^v we have $u=0$ and

$$\boldsymbol{\Gamma}_{0,v}\big[\,dA\,\big]=\frac{1}{(4\pi v)^{\frac{m(m+1)}{4}}}e^{-\frac{1}{4v}\operatorname{tr}A^2}\operatorname{vol}\big[\,dA\,\big]. \tag{1.6}$$

Note that $\mathrm{GOE}_m^{1/2}$ corresponds to the Gaussian measure on $\mathbf{Sym}(\mathbb{R}^m)$ canonically associated to the inner product $(A,B)=\operatorname{tr}(AB)$.

We have a *Weyl integration formula* [2] which states that if $f:\mathbf{Sym}(\mathbb{R}^m)\to\mathbb{R}$ is a measurable function which is invariant under conjugation, then the value $f(A)$ at $A\in\mathbf{Sym}(\mathbb{R}^m)$ depends only on the eigenvalues $\lambda_1(A)\leq\cdots\leq\lambda_n(A)$ of A and we have

$$\mathbb{E}_{\mathrm{GOE}_m^v}\big[\,f(X)\,\big]=\frac{1}{\boldsymbol{Z}_m(v)}\int_{\mathbb{R}^m}f(\lambda_1,\dots,\lambda_m)\underbrace{\left(\prod_{1\leq i<j\leq m}|\lambda_i-\lambda_j|\right)\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}}_{=:Q_{m,v}(\lambda)}\,|d\lambda_1\cdots d\lambda_m|, \tag{1.7}$$

where the normalization constant $\boldsymbol{Z}_m(v)$ is defined by

$$\boldsymbol{Z}_m(v)=\int_{\mathbb{R}^m}\prod_{1\leq i<j\leq m}|\lambda_i-\lambda_j|\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}|d\lambda_1\cdots d\lambda_m|$$

$$=(2v)^{\frac{m(m+1)}{4}}\times\underbrace{\int_{\mathbb{R}^m}\prod_{1\leq i<j\leq m}|\lambda_i-\lambda_j|\prod_{i=1}^m e^{-\frac{\lambda_i^2}{2}}|d\lambda_1\cdots d\lambda_m|}_{=:\boldsymbol{Z}_m}.$$

The integral $\boldsymbol{Z}_m$ is usually referred to as *Mehta's integral*. Its value was first determined in 1960 by M. L. Mehta, [8]. Later Mehta observed that this integral was known earlier to N. G. de Brujin [4]. It was subsequently observed that Mehta's integral is a limit of the *Selberg integrals*, [2, Eq. (2.5.11)], [6, Sec. 4.7.1]. More precisely, we have

$$\boldsymbol{Z}_m=(2\pi)^{\frac{m}{2}}\prod_{j=0}^{m-1}\frac{\Gamma(\frac{j+3}{2})}{\Gamma(3/2)}=2^{\frac{3m}{2}}\prod_{j=0}^{m-1}\Gamma\Big(\frac{j+3}{2}\Big). \tag{1.8}$$

The goal of this note is to provide a probabilistic proof of (1.8).

We will determine $\boldsymbol{Z}_m$ inductively by computing explicitly the ratios $\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}$, $\forall m\geq 1$ and observing by immediate direct computation that

$$\boldsymbol{Z}_1=\int_{\mathbb{R}}e^{-t^2/2}dt=(2\pi)^{1/2}.$$

Here is the strategy. Any symmetric $(m+1)\times(m+1)$ matrix A determines a function on the unit sphere $S^m\subset\mathbb{R}^{m+1}$

$$\Phi_A:S^m\to\mathbb{R},\;\;\Phi_A(\boldsymbol{x})=\frac{1}{2}\big(\,A\boldsymbol{x},\boldsymbol{x}\,\big),$$

where $(-,-)$ is the canonical inner product on $\mathbb{R}^{m+1}$. When $A \in \mathrm{GOE}^v_{m+1}$, then with probability 1 the matrix A is simple and the Gaussian random function Φ_A is Morse.

To the random matrix $A \in \mathrm{GOE}^v_{m+1}$ we can associate two random measures on $\mathbb{R}$. The first is the *spectral measure*

$$\sigma_A = \sum_{\lambda \in \mathrm{Spec}(A)} \delta_\lambda,$$

where δ_x denotes the Dirac measure on $\mathbb{R}$ concentrated at x. The second one is the *discriminant measure*

$$\boldsymbol{D}_A = \sum_{\nabla\Phi_A(\boldsymbol{x})=0} \delta_{2\Phi_A(x)}.$$

The critical values of $2\Phi_A$ are precisely the eigenvalues of A and the critical points are the unit eigenvectors of A. The function is Morse iff A is simple, i.e., its eigenvalues are distinct. In this case to each critical value of A there corresponds exactly two critical points. With probability 1 we have

$$\boldsymbol{D}_A = 2\boldsymbol{\sigma}_A.$$

Then for any Borel subset $C \subset \mathbb{R}$ we have

$$\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big] = 2\mathbb{E}\big[\,\boldsymbol{\sigma}_A\big[\,C\,\big]\,\big]. \tag{1.9}$$

In particular

$$\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,\mathbb{R}\,\big]\,\big] = 2\mathbb{E}\big[\,\boldsymbol{\sigma}_A\big[\,\mathbb{R}\,\big]\,\big] = 2(m+1). \tag{1.10}$$

Using the Kac-Rice formula we will be able to express $\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,\mathbb{R}\,\big]\,\big]$ as an *explicit* multiple of the ratio $\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}$.

Here is the structure of the paper. Section 2 contains several probabilistic digressions. The first one concerns the expectation of the absolute value of characteristic polynomial of a random matrix $A \in \mathrm{GOE}$. The second one describes a version of the Kac-Rice formula needed in the proof. The last digression of this section is a well known classical result commonly referred to as the Gaussian regression formula. We give a coordinate free description of this result not readily available in traditional probabilistic sources, but very convenient to use in geometric applications. Then last section provides the details of the strategy outlined above.

2. PROBABILISTIC DIGRESSIONS

For any positive integer n we define the *normalized* 1-point correlation function $\rho_{n,v}(x)$ of GOE^v_n to be

$$\rho_{n,v}(x) = \frac{1}{\boldsymbol{Z}_n(v)}\int_{\mathbb{R}^{n-1}} Q_{n,v}(x,\lambda_2,\dots,\lambda_n)d\lambda_1\cdots d\lambda_n.$$

For any Borel measurable function $f:\mathbb{R}\to\mathbb{R}$ we have [5, §4.4]

$$\frac{1}{n}\mathbb{E}_{\mathrm{GOE}^v_n}\big[\,\mathrm{tr}\,f(X)\,\big] = \int_{\mathbb{R}} f(\lambda)\rho_{n,v}(\lambda)d\lambda. \tag{2.1}$$

The equality (2.1) characterizes $\rho_{n,v}$. We want to draw attention to a confusing situation in the existing literature on the subject. Some authors, such as M. L. Mehta [9], define the 1-point correlation function $R_n(x)$ by the equality

$$\mathbb{E}_{\mathrm{GOE}^{1/2}_n}\big[\,\mathrm{tr}\,f(X)\,\big] = \int_{\mathbb{R}} f(\lambda)R_n(\lambda)d\lambda.$$

The expected value of the absolute value of the determinant of of a random $A \in \mathrm{GOE}^v_m$ can be expressed neatly in terms of the correlation function $\rho_{m+1,v}$. More precisely, we have the following result first observed by Y.V. Fyodorov [7].

Lemma 2.1. *Suppose $v>0$. Then for any $c\in\mathbb{R}$ we have*

$$\mathbb{E}_{\mathrm{GOE}^v_m}\big[\,|\det(A-c\mathbb{1}_m)|\,\big]=\frac{e^{\frac{c^2}{4v}}\boldsymbol{Z}_{m+1}(v)}{\boldsymbol{Z}_m(v)}\rho_{m+1,v}(c)=\frac{e^{\frac{c^2}{4v}}(2v)^{\frac{m+1}{2}}\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,v}(c). \tag{2.2}$$

Proof. Using Weyl's integration formula we deduce

$$\mathbb{E}_{\mathrm{GOE}^v_m}\big[\,|\det(A-c\mathbb{1}_m)|\,\big]=\frac{1}{\boldsymbol{Z}_m(v)}\int_{\mathbb{R}^m}\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}|c-\lambda_i|\prod_{i\leq j}|\lambda_i-\lambda_j|d\lambda_1\cdots d\lambda_m$$

$$=\frac{e^{\frac{c^2}{4v}}}{\boldsymbol{Z}_m(v)}\int_{\mathbb{R}^m}e^{-\frac{c^2}{4v}}\prod_{i=1}^m e^{-\frac{\lambda_i^2}{4v}}|c-\lambda_i|\prod_{i\leq j}|\lambda_i-\lambda_j|d\lambda_1\cdots d\lambda_m$$

$$=\frac{e^{\frac{c^2}{4v}}\boldsymbol{Z}_{m+1}(v)}{\boldsymbol{Z}_m(v)}\frac{1}{\boldsymbol{Z}_{m+1}(v)}\int_{\mathbb{R}^m}Q_{m+1,v}(c,\lambda_1,\dots,\lambda_m)d\lambda_1\cdots d\lambda_m$$

$$=\frac{e^{\frac{c^2}{4v}}\boldsymbol{Z}_{m+1}(v)}{\boldsymbol{Z}_m(v)}\rho_{m+1,v}(c)=\frac{e^{\frac{c^2}{4v}}(2v)^{\frac{m+1}{2}}\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,v}(c).$$

□

We will need a special version of the Kac-Rice formula. Let (M,g) be a compact Riemann manifold. Denote by $\mathrm{vol}_g[-]$ the volume element on M determined by g and ∇^g the Levi-Civita connection of g. If $F\in C^2(M)$, then we define the Hessian of F at $\boldsymbol{p}\in M$ to be the linear operator

$$\mathrm{Hess}_F(\boldsymbol{p}):T_{\boldsymbol{p}}M\to T_{\boldsymbol{p}}M,\;\;\mathrm{Hess}_F(\boldsymbol{p})X=\nabla^g_X\nabla F,$$

where $\nabla^g F$ is the metric gradient of F.

Suppose that $F:M\to\mathbb{R}$ is a Morse function. For any subset $S\subset M$ we denote by $Z(S,dF)$ the number of critical points of F inside S and B is an open subset We denote by $\mathcal{D}(F)$ the *discriminant set* of F, i.e., the set of critical values of F. The *discriminant measure* of F is the pushforward

$$\boldsymbol{D}_F=\sum_{t\in\mathbb{R}}Z(F^{-1}(t),dF)\delta_t.$$

The discriminant measure is concentrated on $\mathcal{D}(F)$. For $\varphi\in C^0_{\mathrm{cpt}}(\mathbb{R})$ we set

$$\boldsymbol{D}_F\big[\,\varphi\,\big]:=\int_{\mathbb{R}}\varphi d\boldsymbol{D}_F.$$

When F is random, $\boldsymbol{D}_F\big[\,\varphi\,\big]$ is a random variable. We have the following result [1, Thm. 12.4.1].

Theorem 2.2. *Suppose that $F:M\to\mathbb{R}$ is a C^2 Gaussian random function function satisfying the ampleness condition*

$$\textit{for any } \boldsymbol{p}\in M \textit{ the Gaussian vector } F(\boldsymbol{p})\oplus dF(\boldsymbol{p})\in T^*_{\boldsymbol{p}}M \textit{ is nondegenerate.} \tag{A}$$

We denote by $\mathbb{P}_{F(p)}$ the probability distribution of the random variable $F(\boldsymbol{p})$ and by $p_{dF(p)}$ the probability density *of the Gaussian vector $dF(\boldsymbol{p})$.*

Then F is a.s. *Morse and, for any function $\varphi\in C^0_{\mathrm{cpt}}(\mathbb{R})$ we have*

$$\begin{gathered}\mathbb{E}\big[\,\boldsymbol{D}_F[\varphi]\,\big]\\ =\int_M\left(\int_{\mathbb{R}}\mathbb{E}\big[\,\big|\det\mathrm{Hess}_F(\boldsymbol{p})\big|\,\|\,dF(\boldsymbol{p})=0,\;F(\boldsymbol{p})=t\,\big]\varphi(t)\mathbb{P}_{F(\boldsymbol{p})}\big[\,dt\,\big]\right)p_{dF(\boldsymbol{p})}(0)\,\mathrm{vol}_g\,\big[\,d\boldsymbol{p}\,\big]\\ =\int_M\mathbb{E}\big[\,\big|\det\mathrm{Hess}_F(\boldsymbol{p})\big|\varphi\big(\,F(\boldsymbol{p})\,\big)\,\|\,dF(\boldsymbol{p})=0\,\big].\end{gathered}\tag{2.3}$$

Above, $\mathbb{E}\big[\,-\,\|\,-\,\big]$ denotes appropriate conditional expectations.

When applying the Kac-Rice formula we need to evaluate certain conditional expectations. In the Gaussian case this is readily achieved using the classical Gaussian regression formula. In the remainder of this section we describe this Gaussian regression in a form convenient in geometric applications.

Suppose that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are finite dimensional vector spaces. Consider two random vectors

$$X:(\Omega,\mathcal{S},\mathbb{P})\to\boldsymbol{X},\;\;Y:(\Omega,\mathcal{S},\mathbb{P})\to\boldsymbol{Y},$$

where $(\Omega,\mathcal{S},\mathbb{P}))$ is a probability space. The mean or expectation of X is the vector

$$m(X)=\mathbb{E}\big[\,X\,\big]=\int_\Omega X(\omega)\mathbb{P}\big[\,d\omega\,\big]\in\boldsymbol{X},$$

whenever the integral is well defined. The random vector X is called centered if $m(X)=0$.

The *covariance form* of Y and X is the bilinear form

$$\mathrm{Cov}\,\big[\,Y,X\,\big]:\boldsymbol{Y}^*\times\boldsymbol{X}^*\to\mathbb{R}$$

given by

$$\mathrm{Cov}\,\big[\,Y,X\,\big](\eta,\xi)=\mathrm{Cov}\,\big[\,\langle\eta,Y\rangle,\langle\xi,X\rangle\,\big],\;\;\forall\eta\in\boldsymbol{Y}^*,\;\xi\in\boldsymbol{X}^*.$$

If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are equipped with inner products $(-,-)_{\boldsymbol{X}}$ and respectively $(-,-)_{\boldsymbol{Y}}$, then we can identify $\mathrm{Cov}\,\big[\,Y,X\,\big]$ with a linear operator $C_{Y,X}:\boldsymbol{X}\to\boldsymbol{Y}$. Concretely, if $(\boldsymbol{e}_i)_{i\in I}$ and $(\boldsymbol{f}_j)_{j\in J}$ are *orthonormal* bases of $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$, and we set $X_i:=(\boldsymbol{e}_i,X)_{\boldsymbol{X}}$, $Y_j:=(\boldsymbol{f}_j,Y)_{\boldsymbol{Y}}$, then in these bases the operator $C_{Y,X}$ is described by matrix $(c_{ji})_{(j,i)\in J\times I}$, where $c_{ji}:=\mathrm{Cov}\,\big[\,Y_j,X_i\,\big]$. Hence

$$C_{Y,X}\boldsymbol{e}_i=\sum_j c_{ji}\boldsymbol{f}_j.$$

We will refer to $C_{Y,X}$ as the *correlator* of Y with X. Some times, for typographical reasons, we will use the alternate notation $\mathrm{Corr}\,\big[\,Y,X\,\big]:=C_{Y,X}$.

The variance operator of X is $\mathrm{Var}\,\big[\,X\,\big]:=C_{X,X}$. We say that X is nondegenerate if its variance operator is invertible. Observe that $C_{X,Y}:\boldsymbol{Y}\to\boldsymbol{X}$ is the adjoint of $C_{Y,X}$, $C_{X,Y}=C^*_{Y,X}$.

If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are equipped with inner products, then $\boldsymbol{X}\oplus\boldsymbol{Y}$ is equipped with the direct sum of these inner products and in this case $\mathrm{Var}\,\big[\,X\oplus Y\,\big]:\boldsymbol{X}\oplus\boldsymbol{Y}\to\boldsymbol{X}\oplus\boldsymbol{Y}$ admits the block decomposition

$$\mathrm{Var}\,\big[\,X\oplus Y\,\big]=\left[\begin{array}{cc}\mathrm{Var}\,\big[\,X\,\big]&C_{X,Y}\\ C_{Y,X}&\mathrm{Var}\,\big[\,Y\,\big]\end{array}\right].$$

The random vectors X,Y are said to be *jointly Gaussian* if the random vector $X\oplus Y$ is Gaussian.

Proposition 2.3 (Gaussian regression formula)**.** *Suppose that X,Y are Gaussian vectors valued in the Euclidean spaces $\boldsymbol{X}$ and respectively $\boldsymbol{Y}$. Assume additionally that*

(i) *the random vectors X,Y are jointly Gaussian and,*

(ii) *X is nondegenerate.*

Define the regression operator

$$R_{Y,X} : \boldsymbol{X} \to \boldsymbol{Y}, \quad R_{Y,X} := C_{Y,X} \operatorname{Var}[X]^{-1} \tag{2.4}$$

Then the following hold.

(a) The conditional expectation $\mathbb{E}\big[\, Y \,\|\, X \,\big]$ is the Gaussian vector described by the linear regression formula

$$\mathbb{E}\big[\, Y \,\|\, X \,\big] = m(Y) - R_{Y,X} m(X) + R_{Y,X} X. \tag{2.5}$$

(b) For any $x \in X$

$$\mathbb{E}\big[\, Y \big|\, X = x \,\big] = m(Y) - R_{Y,X} m(X) + R_{Y,X} x.$$

(c) The random vector vector $Z = Y - \mathbb{E}\big[\, Y \,\|\, X \,\big]$ is Gaussian and independent of X. It has mean 0 and variance operator

$$\Delta_{Y,X} = \operatorname{Var}\big[\, Y \,\big] - D_{Y,X} : \boldsymbol{Y} \to \boldsymbol{Y}, \quad D_{Y,X} = C_{Y,X} \operatorname{Var}[X]^{-1} C_{X,Y}. \tag{2.6}$$

Moreover, for any bounded measurable function $f : \boldsymbol{Y} \to \mathbb{R}$ and any $x \in \boldsymbol{X}$ we have

$$\mathbb{E}\big[\, f(Y) \big|\, X = x \,\big] = \mathbb{E}\big[\, f\big(\, Z + m(Y) - R_{Y,X} m(X) + R_{Y,X} x \,\big)\,\big]. \tag{2.7}$$

In particular, if X and Y are centered we have

$$\mathbb{E}\big[\, f(Y) \big|\, X = x \,\big] = \mathbb{E}\big[\, f\big(\, Z + R_{Y,X} x \,\big)\,\big]. \tag{2.8}$$

For a proof we refer to [3, Prop. 2.1].

3. The computation of the Mehta integral

As explained in the introduction, when A runs in the Gaussian ensemble GOE^{v}_{m+1} we obtain a Gaussian function

$$\Phi = \Phi_A : S^m \to \mathbb{R}.$$

This function is invariant under the natural $O(m+1)$-action on S^m.

Lemma 3.1. *The Gaussian function Φ_A is* a.s. *Morse.*

Proof. It suffices to show that the Gaussian section $\nabla \Phi_A$ of TS^m i satisfies the ampleness condition (**A**). Let $\boldsymbol{x} \in S^m$. If $\operatorname{Proj}_{\boldsymbol{x}} : \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ the orthogonal projection onto $T_{\boldsymbol{x}} S^m$, then

$$\nabla \Phi_A(\boldsymbol{x}) = \operatorname{Proj}_{\boldsymbol{x}} A\boldsymbol{x} = A\boldsymbol{x} - (A\boldsymbol{x}, \boldsymbol{x})\boldsymbol{x}.$$

The map

$$\mathbf{Sym}_{m+1}(\mathbb{R}) \ni A \mapsto A\boldsymbol{x} \in \mathbb{R}^{m+1}$$

is onto and thus the map

$$\mathbf{Sym}_{m+1}(\mathbb{R}) \ni A \mapsto \operatorname{Proj}_{\boldsymbol{x}} A\boldsymbol{x} \in T_{\boldsymbol{x}} S^m$$

is also onto, thus proving that the gradient $\nabla \Phi_A(\boldsymbol{x})$ is nondegenerate since the Gaussian ensemble $\mathrm{GOE}_{m+1,v}$ is nondegenerate. □

The spectral measure of A is

$$\boldsymbol{\sigma}_A := \sum_{\lambda \in \operatorname{Spec}(A)} \operatorname{mult}(\lambda) \delta_\lambda.$$

The discriminant measure of Φ_A is

$$\boldsymbol{D}_A = \sum_{\nabla \Phi_A(\boldsymbol{x}) = 0} \delta_{2\Phi_A(x)}.$$

With probability 1 we have $\boldsymbol{D}_A = 2\boldsymbol{\sigma}_A$. Then for any Borel subset $C \subset \mathbb{R}$ we have

$$\frac{1}{(m+1)}\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big] = \frac{2}{m+1}\mathbb{E}\big[\,\boldsymbol{\sigma}_A\big[\,C\,\big]\,\big] = 2\int_C \rho_{m+1,v}(\lambda)d\lambda. \tag{3.1}$$

Using the Kac-Rice formula (2.3) we will give an alternate description of the left-hand-side of the above equality. We will need to describe explictly the integrand in this formula.

For $\boldsymbol{x} \in S^m$ we denote by $\operatorname{Hess}_A(\boldsymbol{x})$ the Hessian of Φ_A at $\boldsymbol{x}$ viewed as a symmetric operator $T_{\boldsymbol{x}}S^m \to T_{\boldsymbol{x}}S^m$.

Denote by $(x^0, x^1, \dots, x^m)$ the canonical Euclidean coordinates on $\mathbb{R}^{m+1}$. Since Φ_A is $O(m+1)$ invariant, the distribution of $\operatorname{Hess}_A(\boldsymbol{x})$ is independent of $\boldsymbol{x}$ so it suffices to determine it at any point of our choosing. Suppose that $\boldsymbol{x}$ is the north pole

$$\boldsymbol{x} = \boldsymbol{n} = (1, 0, \dots, 0) \in \mathbb{R}^{m+1}$$

Then $T_{\boldsymbol{n}}S^m = \{x^0 = 0\}$ and $\boldsymbol{x}_* := \big(x^1, \dots, x^m\big)$ are orthonormal coordinates on $T_{\boldsymbol{n}}S^m$. The coordinates $\boldsymbol{x}_*$ also define local coordinates on S^m. More precisely, the upper hemisphere

$$S^m_+ := \big\{\,\boldsymbol{x} \in S^m;\;\; x^0 > 0\,\big\}$$

admits the parametrization

$$\boldsymbol{x}^* \mapsto \boldsymbol{x}\big(\,\boldsymbol{x}_*\,\big) = \big(\,x^0(\boldsymbol{x}_*), \boldsymbol{x}^*\,\big) \in S^m, \;\; x^0(\boldsymbol{x}_*) = \sqrt{1 - \|\boldsymbol{x}_*\|^2}.$$

The round metric on S^m satisfies

$$g_{ij} = \delta_{ij} + O\big(\,\|\boldsymbol{x}_*\|^2\,\big) \;\text{ near } \boldsymbol{n}. \tag{3.2}$$

On the upper hemisphere we will view Φ_A as a function of $\boldsymbol{x}_*$.

If $A = (a_{ij})_{0 \le i,j \le m}$, then in the coordinates $\boldsymbol{x}_*$ we have

$$\Phi_A(\boldsymbol{x}) = \frac{1}{2}a_{00}\big(\,1 - \|\boldsymbol{x}_*\|^2\,\big) + \frac{1}{2}\sum_{j=1}^m a_{jj}(x^j)^2 + \sum_{0 \le j < k \le m} a_{jk}x^jx^k,$$

$$= \frac{1}{2}a_{00}^2 + \frac{1}{2}\sum_{j=1}^m \big(\,a_{jj} - a_{00}\,\big)(x^j)^2 + \sum_{0 \le j < k \le m} a_{jk}x^jx^k,$$

$$\nabla\Phi_A(\boldsymbol{n}) = d\Phi_A(\boldsymbol{x}_*)|_{\boldsymbol{x}_*=0} = \sum_{j=1}^m a_{0j}dx^j.$$

Since $A \in \mathrm{GOE}^v_{m+1}$, covariance kernel of Φ_A is

$$\mathcal{K}_A(\boldsymbol{n}, \boldsymbol{x}) = \mathbb{E}\big[\,\Phi_A(\boldsymbol{n})\Phi_A(\boldsymbol{x})\,\big] = \frac{1}{4}\big(\,1 - \|\boldsymbol{x}_*\|^2\,\big)\mathbb{E}\big[\,a_{00}^2\,\big] = \frac{v}{2}\big(\,1 - \|\boldsymbol{x}_*\|^2\,\big).$$

Denote by A_* the $m \times m$ matrix $A_* = (a_{ij})_{1 \le i \le m}$. Note that $A_* \in \mathrm{GOE}^v_m$. Using (3.2)

$$\operatorname{Hess}_A(\boldsymbol{n}) = A_* - a_{00}\mathbb{1}_m.$$

Since a_{00} is independent of A_* we deduce from (1.5) that $\operatorname{Hess}_A(0) \in \mathcal{S}^{2v,v}_m$, where $\mathcal{S}^{u,v}_m$ is the $O(m)$-invariant Gaussian ensemble defined by (1.3). If we set

$$L_{ij} = \ell_{ij}\big(\,\operatorname{Hess}_A(\boldsymbol{n})\,\big), \;\; \Omega_{iij} = \omega_{ij}\big(\,\operatorname{Hess}_A(\boldsymbol{n})\,\big),$$

where ℓ_{ij} and ω_{ij} are defined by (1.1), then

$$\mathbb{E}\big[\,L_{ij}L_{k\ell}(A)\,\big] = 2v\delta_{ij}\delta_{k\ell} + v(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}), \;\; \forall 1 \le i, j, .k, \ell \le m.$$

Note that $\nabla\Phi_A(\boldsymbol{n}) = (a_{01},\dots,a_{0n})$ is independent of $\Phi_A(\boldsymbol{n})$ and of $\operatorname{Hess}_A(\boldsymbol{n})$ and since $A \in \mathrm{GOE}^v_{m+1}$ we deduce from (1.3) that

$$\operatorname{Var}\big[\,\Phi_A\,\big] = v\mathbb{1}_m. \tag{3.3}$$

Moreover the correlator $\operatorname{Corr}\big[\,\operatorname{Hess}_A(\boldsymbol{n}),\Phi_A(\boldsymbol{n}\,\big] : \mathbb{R} \to \mathbf{Sym}_m(\mathbb{R})$ is given by

$$\mathbb{R} \ni x \mapsto -x\mathbb{1}_m \in \mathbf{Sym}_m(\mathbb{R}).$$

Set

$$W = \begin{bmatrix} \Phi_A(\boldsymbol{n}) \\ \nabla\Phi_A(\boldsymbol{n}) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}a_{00} \\ a_{01} \\ \vdots \\ a_{0m} \end{bmatrix}.$$

Note that

$$\operatorname{Var}\big[\,W\,\big] = \operatorname{Diag}\big(\,\frac{v}{2},\underbrace{2v,\dots,2v}_{m}\,\big).$$

Denote by $\overline{\operatorname{Hess}}_A(\boldsymbol{n})$ the random symmetric matrix with variance given by the regression formula

$$\operatorname{Var}\big[\overline{\operatorname{Hess}}_A(\boldsymbol{n})\big] = \operatorname{Var}\big[\,\operatorname{Hess}_A(\boldsymbol{n})\,\big] - \operatorname{Corr}\big[\,\operatorname{Hess}_A(\boldsymbol{n}),W\,\big]\operatorname{Var}\big[\,W\,\big]^{-1}\operatorname{Corr}\big[\,W,\operatorname{Hess}_A(\boldsymbol{n})\,\big].$$

Set

$$\bar{L}_{ij} = \ell_{ij}\big(\overline{\operatorname{Hess}}_A(\boldsymbol{n})\,\big),\;\; \bar{\Omega}_{ij} := \omega_{ij}\big(\overline{\operatorname{Hess}}_A(\boldsymbol{n})\,\big),$$

and

$$C_{ij|k} := \operatorname{Cov}\big[\,\Omega_{ij},W_k\,\big],\;\; 1 \le i \le j \le m,\;\; 0 \le k \le m.$$

Note that

$$C_{ij|k} = 0,\;\; \forall i,j,\;\; \forall k > 0,\;\; C_{ij|k} = 0,\;\; \forall i < j,\;\; \forall k \ge 0,$$

and

$$C_{ii|0} = \frac{1}{2}\mathbb{E}\big[\,(a_{ii}-a_{00})a_{00}\,\big] = -\frac{1}{2}\mathbb{E}\big[\,a_{00}^2\,\big] = -v.$$

If we write

$$\operatorname{Var}\big[\,W\,\big]^{-1} = \big(\,t_{ab}\,\big)_{0\le a,b\le m},$$

then

$$\mathbb{E}\big[\bar{\Omega}_{ij}\bar{\Omega}_{k\ell}\,\big] = \mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big] - \sum_{a,b=0}^m C_{ij|a}t_{ab}C_{k\ell|b} = \mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big] - \frac{2}{v}C_{ij|0}C_{k\ell|0}$$

For $i \ne j$

$$\mathbb{E}\big[\bar{\Omega}_{ii}\bar{\Omega}_{jj}\,\big] = \mathbb{E}\big[\,\Omega_{ii}(\boldsymbol{x})\Omega_{jj}(\boldsymbol{x})\,\big] - 2v = 0$$

$$\mathbb{E}\big[\bar{\Omega}_{ii}^2\,\big] = 2v,$$

$$\mathbb{E}\big[\bar{\Omega}_{ij}\bar{\Omega}_{k\ell}\,\big] = \mathbb{E}\big[\,\Omega_{ij}(\boldsymbol{x})\Omega_{k\ell}(\boldsymbol{x})\,\big],\;\; \forall 1 \le k \le \ell.$$

We deduce that $\overline{\operatorname{Hess}}_A \in \mathrm{GOE}^v_m$ so $\overline{\operatorname{Hess}}_A$ has the same distribution as A.

The regression operator

$$R_{\operatorname{Hess}_A,W} = \operatorname{Corr}\big[\,\operatorname{Hess}_A,W\,\big]\operatorname{Var}\big[\,W^{-1}\,\big] : \mathbb{R}^{m+1} \to \mathbf{Sym}_m(\mathbb{R})$$

is

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix} \mapsto -2w_0\mathbb{1}_m.$$

Using the regression formula we deduce

$$\mathbb{E}\big[\,|\operatorname{Hess}_A(\boldsymbol{n})|\,\|\,W(\boldsymbol{n})=(t/2,0)\,\big]=\mathbb{E}_{\mathrm{GOE}_{m,v}}\big[\,|\det(A-vt)|\,\big].$$

Since $2\Phi_A(\boldsymbol{n})=a_{00}$ is Gaussian with variance $2v$, we deduce. from the Kac-Rice formula (2.3) that for any Borel subset $C\subset\mathbb{R}$ we have

$$\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big]=\int_C\rho_A(t)\boldsymbol{\gamma}_{2v}\big[\,dt\,\big],$$

where

$$\begin{aligned}\rho_A(t)&=\int_{S^m}\mathbb{E}\big[\,|\operatorname{Hess}_A(\boldsymbol{x})|\,\|\,2\Phi_A(\boldsymbol{x})=t,\ \nabla\Phi_A(\boldsymbol{x})=0\,\big]p_{\nabla\Phi_A(\boldsymbol{x})}(0)d\boldsymbol{x}\\&\overset{(3.3)}{=}\big(\,2\pi v\,\big)^{-m/2}\int_{S^m}\mathbb{E}\big[\,|\operatorname{Hess}_A(\boldsymbol{x})|\,\|\,W(\boldsymbol{x})=(t/2,0)\,\big]d\boldsymbol{x}\\&=\big(\,2\pi v\,\big)^{-m/2}\mathbb{E}\big[\,|\operatorname{Hess}_A(\boldsymbol{n})|\,\|\,W(\boldsymbol{n})=(t/2,0)\,\big]\operatorname{vol}\big[\,S^m\,\big]\\&=\big(\,2\pi v\,\big)^{-m/2}\operatorname{vol}\big[\,S^m\,\big]\mathbb{E}_{\mathrm{GOE}^v_m}\big[\,|\det(A-vt)|\,\big].\end{aligned}\tag{3.4}$$

In Lemma 2.1 we showed

$$\mathbb{E}_{\mathrm{GOE}_{m,v}}\big[\,\det(A-vt)|\,\big]=(2v)^{\frac{m+1}{2}}e^{\frac{v^2t^2}{4v}}\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,v}(vt).$$

Assume now that $v=1$.

$$\mathbb{E}_{\mathrm{GOE}^1_m}\big[\,\det(A-t)|\,\big]=e^{\frac{t^2}{4}}2^{\frac{m+1}{2}}\pi^{-1/2}\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\rho_{m+1,1}(t)$$

Since $\boldsymbol{\gamma}_{2v}\big[\,dt\,\big]=e^{-\frac{t^2}{4}}\frac{dt}{\sqrt{4\pi}}$ we deduce

$$\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big]=\big(\,2\pi\,\big)^{-m/2}2^{\frac{m+1}{2}}\operatorname{vol}\big[\,S^m\,\big]\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}\int_C\rho_{m+1,1}(t)\frac{dt}{\sqrt{4\pi}}.$$

On the other hand, we deduce from (3.1) that

$$\frac{1}{(m+1)}\mathbb{E}\big[\,\boldsymbol{D}_A\big[\,C\,\big]\,\big]=2\int_C\rho_{m+1,1}(t)dt,$$

so that

$$\frac{\big(\,2\pi\,\big)^{-m/2}2^{\frac{m+1}{2}}\operatorname{vol}\big[\,S^m\,\big]}{(m+1)}\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}(4\pi)^{-1/2}=2.$$

Using the fact that

$$\frac{\operatorname{vol}\big[\,S^m\,\big]}{m+1}=\frac{\pi^{\frac{m+1}{2}}}{\Gamma(\frac{m+3}{2})}$$

we deduce

$$\frac{\boldsymbol{Z}_{m+1}}{\boldsymbol{Z}_m}=\frac{\Gamma(\frac{m+3}{2})}{\pi^{\frac{m+1}{2}}}\cdot\frac{2(2\pi)^{m/2}(4\pi)^{1/2}}{2^{\frac{m+1}{2}}}=2^{3/3}\Gamma\Big(\frac{m+3}{2}\Big).$$

Note that

$$\boldsymbol{Z}_1=\int_{\mathbb{R}}e^{-t^2/2}dt=(2\pi)^{1/2}.$$

We deduce immediately the equality (1.8)

$$\boldsymbol{Z}_m=\boldsymbol{Z}_1\prod_{j=1}^{m-1}\frac{\boldsymbol{Z}_{j+1}}{\boldsymbol{Z}_j}=2^{\frac{3m}{2}}\prod_{j=0}^{m-1}\Gamma\Big(\frac{j+3}{2}\Big).$$

REFERENCES

[1] R. Adler, R. J. E. Taylor: *Random Fields and Geometry*, Springer Monographs in Mathematics, Springer Verlag, 2007.
[2] G. W. Anderson, A. Guionnet, O. Zeitouni: *An Introduction to Random Matrices*, Cambridge University Press, 2010.
[3] J.-M. Azaïs, M. Wschebor: *Level Sets and Extrema of Random Processes*, John Wiley & Sons, 2009.
[4] N.G. de Brujin: *On some multiple integrals involving determinants*, J. Ind. Math. Soc., **19**(1955), 133-151.
[5] P. Deift, D. Gioev: *Random Matrix Theory: Invariant Ensembles and Universality*, Courant Lecture Notes, vol. 18, Amer. Math. Soc., 2009.
[6] P. J. Forrester: *Log-Gases and Random Matrices*, London Math. Soc. Monographs, Princeton University Press, 2010.
[7] Y. V. Fyodorov: *Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices*, Phys. Rev. Lett, **92**(2004), 240601; Erratum: **93**(2004), 149901.
[8] M. L. Mehta: *On the statistical proerties of level-spacings in nuclear spectra*, Nucl. Phys.. **18**(1960), 395-419.
[9] M. L. Mehta: *Random Matrices*, 3rd Edition, Elsevier, 2004.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NOTRE DAME, NOTRE DAME, IN 46556-4618.
Email address: nicolaescu.1@nd.edu
URL: http://www.nd.edu/~lnicolae/