On Gaussian Random Matrix Ensembles

Robert Ferydouni, University of California at Santa Cruz

Abstract

In this paper, we first provide the motivation from nuclear physics for studying random matrices, and then the necessary background to do so. We then define the three classic Gaussian random matrix ensmbles: Gaussian Orthogonal Ensemble (GOE), Gaussian Symplectic Ensemble (GSE) and the Gaussian Unitary Ensemble (GUE). The definitions will be motivated by the mathematical formulations of the physical properties we want the ensembles to capture in a model. We then give the probability density function a matrix from these ensembles, and then that of their eigenvalues.

1 Motivation

Nuclear reactors generate nuclear power by harnessing the energy released when a neutron collides and then reacts with a radioactive element's atom. Below we have two graphs plotting the data for such a reaction [2]. The first graph is the spectrum for a neutron reacting with a thorium atom, and the second is the spectrum for a neutron reacting with a uranium atom. The location of the peaks are called nuclear energy levels, i.e., the energy released throughout the reaction that we want to capture. A model describing how much energy is released during a neutron-nucleus reaction is really a model describing the way the nuclear energy levels are spread out.

Quantum observables are those measurable entities such as the position or momentum of a particle in a system, or the energy of a system. In our study, we will be concerned with obtaining the values of energy, particuarly the nuclear energy levels, where the system in question is the neutron nucleus reaction. Quantum mechanical theory says the system can be represented by a particular matrix H (with real or complex entries), called the Hamiltonian, and that the energy

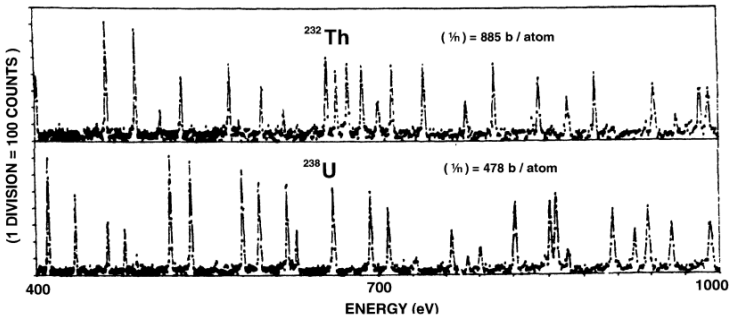

Figure 1: Spectrum of neutron reaction with thorium 232 and uranium 238 nuclei.

states of the system are the eigenvalues of H [3]. We want the eigenvalues to be real numbers, so that they have a clear physical interpretation when we say the energy can take such a value in our system. Recall by the spectral theorem in linear algebra, H has real eigenvalues (the number of which equaling the number of diagonal entries) if H is Hermitian (sometimes called self-adjoint). That is, if $H = H^*$, where $H^* = \overline{H^T} = \overline{H}^T$, H^T denotes the transpose of H, and $\overline{H}$ denotes the matrix obtained after taking complex conjugates of each entry in H. The matrices H^* is called the conjugate transpose of H.

Finding H however is no easy task as H may act on an infinite dimensional space to give infinitely many eigenvalues, and so is almost impossible to solve for. Therefore, loosely speaking, we can take a particular $N \times N$ matrix $H = (H_{kj})_{k,j=1}^N$ and look at the limiting behavior of eigenvalues as $N \to \infty$ [3]. Thus from here on, we assume our Hamiltonian H is an $N \times N$ matrix. The eigenvalues will depend on the entries of the matrix, which in turn depend on the assumptions we make about the physical system we are modeling. In the next section, we discuss some of the matrices we will encounter in our further analysis, and well as the probabilistic notions we will encounter in said analysis.

2 Preliminaries

2.1 Matrices

It is well known from elementary linear algebra that every operator acting on a vector space has a unique matrix corresponding to it (for a particular basis), and vice versa. In our consideration, we will thus exchange the use of the terms ”matrix” and ”operator” without distinction. Consider

the case H has only real entries. To ensure H has N real eigenvalues, the condition $H^* = H$ is equivalent to the condition $H = H^T$, since $H = \overline{H}$ already. That is, H need be a real symmetric matrix. If H has complex entries we need the original condition $H = H^*$ to ensure the existence of N real eigenvalues.

A matrix K is unitary if its conjugate transpose equals its inverse, i.e., $KK^* = 1 = K^*K$. If the entries of K are real, we have $K^* = K^T$ so that $KK^T = 1 = K^TK$, in which we say K is orthogonal. We will often see in our analysis that given a Hamiltonian H and a unitary matrix K, we consider transformations of the form $H \to K^{-1}HK = K^*HK$ when K is a general unitary matrix, or $H \to K^{-1}HK = K^THK$ when K is orthogonal (sometimes called conjugation). Note these transformations are a change of basis for the matrix H, and for appropriate K unitary or orthogonal this transformation diagonalizes H. Transformations of this kind are useful in physics because a change of basis allows one to do calculations in a more familiar coordinate system, and as we will see allows us to work directly with the eigenvalues of H when we parametrize its entries. Therefore when we eventually construct a notion of "the probability of a system H," we want the probability to be the same for H before and after making such transformations. If H is a real symmetric matrix, we should want to make such transformation with K orthogonal to ensure $K^{-1}HK = K^THK$ is still real symmetric, and not introduce unwanted complex numbers as in the case K would be a general unitary matrix. Hence if H complex Hermitian, we would diagonalize it by conjugating with a general unitary operator, and if H was real symmetric we would diagonlize it by conjugating with a real orthogonal matrix.

We will also see further below that there are cases where we cannot have our matrices be complex or real valued. It will instead be useful to consider families of matrices that are quaternion valued. First, we recall the definition of quaternions. Similarly to how $\mathbb{C}$ is an extension of $\mathbb{R}$ by taking numbers of the form $a + bi \in \mathbb{C}$ rather than just $a \in \mathbb{R}$, the quaternions can be thought of as an extension of $\mathbb{C}$ by taking numbers of the form $a + bi + cj + dk$ for $a, b, c, d \in \mathbb{R}$. Addition of quaternions is done component wise, while multiplication starts with defining $i^2 = j^2 = k^2 = -1$ (note how the condition $i^2 = -1$ agrees with that in the complex numbers), $ij = k, jk = i, ki = j$ and $ji = -k, kj = -i, ik = -j$, then extending with the distributive law. We let the conjugate of $q = a + bi + cj + dk$ be $q^R = a - bi - cj - dk$. One can then show $q^Rq = |q|^2 = a^2 + b^2 + c^2 + d^2$

so that $q^{-1} = \frac{q^R}{|q|^2}$. Now consdier

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, e_1 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, e_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, e_3 = e_1 e_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Then $1, e_1, e_2, e_3$ satisfy the same multiplication conditions as the $1, i, j, k$ in the quaternions. Thus we can identify the quaternions with linear combinations of the matrices $1, e_1, e_2, e_3$. Hence an $N \times N$ matrix with quaternion entries can be identified with a $2N \times 2N$ matrix where instead of elements of the form $a + bi + cj + dk$, we have 2×2 blocks that are linear combinations of $1, e_1, e_2, e_3$. In particular, we can identify q and q^R with

$$\begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} \text{ and } \begin{pmatrix} a - bi & -c - di \\ c - di & a + bi \end{pmatrix}$$

respectively. Note the matrix representing q^R is the conjugate transpose of that of q. Ultimately, we have introduced all this language for quaternions to obtain a condition when a quaternion valued matrix H yields N real eigenvalues. If $H = (q_{jk})_{j,k=1}^N$ is assumed to be an $N \times N$ matrix with quaternion entries q_{jk}, it can thus be identified with a $2N \times 2N$ matrix with complex entries. Let $H^R = (H^R)_{j,k=1}^N = (q_{kj}^R)_{j,k=1}^N$ be the dual of H, and call H self-dual if $H = H^R$. Then we see that H being self dual means the $2N \times 2N$ matrix we identify it with is Hermitian, as the matrix forms of q_{jk} and q_{kj}^R are the conjugate transpose of one another, so that H gives us $2N$ real eigenvalues. In fact these eigenvalues will come in pairs in 2×2 diagonal blocks on the diagonal of the $2N \times 2N$ form of H, so that we actually have N real eigenvalues. However, we will not prove this statement here.

One can show that a complex 2×2 matrix, q, is a quaternion if and only if

$$q^R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} q^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{1}$$

If H is self dual, each entry q_{jk} in H equals the conjugate of q_{kj}, i.e., q_{kj}^R. Since each q_{kj}^R can be written in the above way, we have $H = JH^T J^{-1}$, where J is the block diagonal matrix with diagonal

blocks $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Thus we can equivalently say H is self dual if the $2N \times 2N$ matrix B we identify it with satisfies $B = JB^T J^{-1}$.

2.2 Probability Triples and Random Variables

If the matrices we used in our analysis simply contained numbers as their entries (real, complex or quaternion), our system would be constant and thus deterministic. This would not account for the erratic behavior we see in the case of a neutron reacting with a nucleus. Thus our entries need to be functions to capture the variety in different possible values of the energy. However, they cannot be any kind of function. Philosophically, our models should capture the statistical behavior of a quantum system. To do so, we consider our matrix entries as functions that are also what are called random variables. These particular functions have an interpretation, in that it makes sense to ask: what is the probability the image of a random variable, acting on certain sets, lies in a particular range of values? To define what these functions are and how we will answer such probabilistic questions, we first introduce the necessary building blocks of probability theory.

In practical use in statistics, a random variable gives values to collected data on height, age, weight, neighborhood, etc. Thus the domain of a random variable needs to be abstract to be able to consider a function from qualitative notions to the real or complex numbers. The domain will be a probability triple, denoted by $(\Omega, \mathcal{F}, P)$. Ω is the sample space, i.e., the set containing all the events that can occur in a particular setting. We want to assign probabilities to subsets of Ω. However in general, this will not always be defined. We need to restrict our attention to the "measurable events" given by elements of $\mathcal{F}$, known as a σ-algebra. I.e., $\mathcal{F}$ is a set of subsets of Ω where it is well defined to assign probabilities to such subsets. $P : \mathcal{F} \to [0, 1]$ (though P can also take complex values and the codomain would be in $\mathbb{C}$) is the function that assign probabilities to elements of $\mathcal{F}$. Note that the values P can take lie in $[0, 1]$. Intuitively, the chance of an event occurring is between 0% and 100%. $\mathcal{F}$ is known as a σ-algebra and satisfies the following conditions: (1) $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$ (2) $A_1, A_2, ... \in \mathcal{F}$ implies $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$. Intuitively, the first condition means that if we can make sense of the probability of an event A occurring, we can make sense of the probability of A not occurring, while the second condition is so we do not have problems with measuring the

probability of infinitely many events. These conditions can be shown to imply $\mathcal{F}$ is closed under finite unions and intersections, countable intersections, and that it contains Ω and $\varnothing$. By definition $P(\varnothing) = 0$ (i.e., there is a 0% chance of "nothing" happening) and $P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} P(A_n)$ if the sequence of sets $\{A_n\}_{n\in\mathbb{N}}$ is pairwise disjoint.

Now let $(\Omega, \mathcal{F}, P)$ be a given probability space, R another sample space and $\mathcal{R}$ a σ-algebra of subsets of R. We define a random variable to be a function $X : \Omega \to R$ such that $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{R}$. That is, the preimage of measurable events in R are measurable in Ω. Consider the case $R = \mathbb{R}$ (respectively $R = \mathbb{C}$). To construct a notion of a σ-algebra on $\mathbb{R}$ or $\mathbb{C}$, we define the σ-algebra generated by a set A to be the intersection of all σ-algebras containing A, and we denote it $\sigma(A)$. There is at least one set in this intersection, precisely the power set of $\mathbb{R}$ or $\mathbb{C}$, so that the intersection always makes sense. In this case we typically take $\mathcal{R}$ to be the σ-algebra generated by all the open intervals in $\mathbb{R}$ (respectively open disks in $\mathbb{C}$). We call this the Borel σ-algebra and we denote it by $\mathcal{B}(\mathbb{R})$ (respectively $\mathcal{B}(\mathbb{C})$). For the rest of this section, we will develop the necessary probability theory just in the case we have real valued random variables for simplicity. The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ also contains all closed intervals, all half open and all half closed intervals. For example, if $X : \Omega \to \mathbb{R}$ is a random variable (defined on some abstract space Ω) and $(a, b] \in \mathbb{R}$, we can ask: what is $P(\{\omega \in \Omega : X(\omega) \in (a, b]\}) = P(X^{-1}((a, b]))$, i.e., what is the probability $a < X(\omega) \leqslant b$ for $\omega \in \Omega$?

As mentioned earlier, the random matrices we will analyze have entries that are real or complex valued random variables. That is, functions of the form $X : \Omega \to \mathbb{R}$ or $X : \Omega \to \mathbb{C}$ for some abstract underlying probability space $(\Omega, \mathcal{F}, P)$.

2.3 Distributions and Densities of Continuous Random Variables

Let $(\Omega, \mathcal{F}, P)$ be a given probability space. Note we do not yet have a probability measure for the space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Given a random variable $X : \Omega \to \mathbb{R}$ we can define a function $\mu : \mathcal{B}(\mathbb{R}) \to [0, 1]$ by $\mu(A) = P(X^{-1}(A))$, which is well defined since $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{B}(\mathbb{R})$ by assumption of X being a random variable. μ is called the law/distribution of X, and is actually a probability measure, making $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ a probability space. The probability measures we will use from this point will be distributions for random variables (and we will see soon of multiple random vari-

ables) so that we use the terms probability measure and probability distribution interchangably. It is possible for random variables to have values that are finite or countable. These are known as discrete random variables, and they will not be of concern here. Instead, we will focus on continuous random variables. Given a continuous random variable X with a distribution μ, using measure theory we can write

$$\mu(A) = \int_A \rho(x)dx$$

for an integrable function ρ. ρ is defined to be the density function of X. Intuitively, ρ gives the probability of X taking on particular values, and integrating over A gives the probability X has values in A. This notion agrees with the intuition behind the definition $P(X^{-1}(A)) = \mu(A)$. The probability density function satisfies

$$\int_{\mathbb{R}} \rho(X)dx = 1.$$

Intuitively, the probability of ρ taking any possible value is 100%.

Ultimately, we need to be able to compute probabilities for a matrix where each entry is a random variable. Thus we need the notion of a multivariate probability density. Let $X_1, ..., X_n$ be random variables where each X_j is defined on the probability triple $(\Omega_j, \mathcal{F}_j, P_j)$. We can consider the product space $(\Omega, \mathcal{F}, P)$, where $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathcal{F} = \sigma(\mathcal{F}_1 \times \times \mathcal{F}_n)$, and $P = P_1 \cdots P_n$ (known as the product measure). We can then define the joint probability distribution of $X_1, ..., X_N$ as $\mu(A_1, ..., x_n) = P(X_1^{-1}(A_1) \times \cdots \times X_n^{-1}(A_n))$. Similarly to the single variable case, we can say that if $X_1, ..., X_n$ are continuous, we can write

$$\mu(A_1, ..., A_n) = \int_{A_1} \cdots \int_{A_n} \rho(x_1, ..., x_n)dx_n \cdots dx_1$$

for an integrable function ρ. Here ρ is the joint probability density of $X_1, ..., X_n$. Intuitively, $\rho(X_1, ..., X_n)$ is the probability each X_j takes some specified values, so integrating ρ over $A_1 \times \cdots \times A_n$ gives the probability each X_j takes values in A_j. Also similarly to the single variable case, we have

$$\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} P(x_1, ..., x_n)dx_n \cdots dx_1 = 1.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $A_1, ..., A_N \in \mathcal{F}$ be measurable events. We say $A_1, ..., A_n$ are independent if $P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$. Intuitively, the probability of all the events

occurring depends only on the individual probability, and not where they may intersect. I.e., we do not need to avoid double counting by subtracting a term involving their intersections, as each event does not "depend" on the other. Similarly, a collection $X_1, ..., X_n$ of random variables, with each X_j having distribution μ_j and density ρ_j, are called independent if their joint distribution μ and joint density function ρ satisfies $\mu(A_1, ..., A_n) = \int_{A_1} \cdots \int_{A_n} \rho(x_1, ..., x_n) dx_N \cdots dx_1 = \int_{A_1} \rho_1(x_1) dx_1 \cdots \int_{A_N} \rho_N(n) dx_n = \mu_1(A_1) \cdots \mu_N(A_N)$. I.e., the joint distribution is the product measure of the individual distributions. This is if and only if $\rho = \rho_1 \cdots \rho_n$. The notion of independent random variables will be of central importance to our analysis below, as we will assume the random variables in the entries of our matrices will be independent. If $X_1, ...X_n$ have the same distribution (i.e., $\mu_1 = \cdots = \mu_n$) we say $X_1, ..., X_n$ are identically distributed. We call abbreviate the property of being independent and identically distributed by i.i.d.

Let's illustrate the ideas of distributions, densities, and independence in the context of random matrices. A random matrix $H = (H_{jk})_{j,k=1}^N$ is a matrix whose entries are random variables. We will take the entries H_{jk} to be in particular continuous and independent random variables. To say H has a probability distribution means we consider the joint probability distribution of its entries. Since the entries are continuous it makes sense to discuss the joint density function of the entries, which we sometimes refer to as the density of H. A distribution for H tells us the probability the entries of H lie in a some set of values. These values can be represented by a set of matrices G, where the j, k entry of $M \in G$ is one of the possible values H_{jk} can take. Since the entries of H are continuous, we have the distribution of H is given by the form $P(H \in G) = \int_G \rho(H) dH$. Note this integral is N^2 single variable integrals, since ρ is the joint density function for the N^2 entries in H. We will come back to this expression later.

2.4 Gaussian Random Variables

We lastly discuss Gaussian random variables, which will be the random variables we have as the entries in H, and discuss why we choose our entries to be Gaussian. We define a Gaussian random variable with mean μ and variance σ^2 to be the real random variable with density $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, and denote the distribution by $N(\mu, \sigma^2)$. Two particularly important cases in our consideration are when $\mu = 0$ and $\sigma^2 = 1$ so that the density becomes $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-x^2/2\right)$, and

when $\mu = 0$ and $\sigma^2 = 2$ so that the density becomes $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-2x^2\right)$. The corresponding distribution in the first case is called the standard normal distribution. The notion of mean and variance can be defined for much more general random variables than just the Gaussian. The mean μ of a random variable can be thought of as the average of its values, and the variance can be thought of as how spread out the values of the random variable are. The Gaussian is of particular importance because of the following theorem, known as the central limit theorem. Before giving the theorem, we say and sequence of random variables $\{X_n\}$ converges to and random variable X in distribution if for all $f \in C_0^\infty(\mathbb{R})$ (the set of compactly supported C^∞ functions on $\mathbb{R}$) if $\int_{\mathbb{R}} f d\mu_{X_n} \to \int_{\mathbb{R}} f d\mu_X$ as $n \to \infty$.

Theorem 2.4.1. *Let $X_1, X_2, ...$ be a sequence of independent and identically distributed random variables with mean $\mu < \infty$ and variance $\sigma^2 \in (0, \infty)$. Let $S_n = X_1 + \cdots + X_n$ and $Z_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$. Then Z_n converges in distribution to a Gaussian random variable with standard normal distribution $N(0, 1)$.*

Loosely speaking, the above result says that with an appropriate translation with the average and appropriate scaling with the square root of the variance, we can approximate the distribution of a general random variable (given its mean and variance are defined) with the standard normal distribution. This means the Gaussian random variable is, in a sense, a universal random variable. Hence it is reasonable to take our entries in the random matrices we consider the be Gaussian random variables.

3 Gaussian Ensembles

In this section we give the definitions for the three Gaussian ensembles, each of which will require a choice of what matrices we choose to include in the ensemble, which will be motivated by the physical system we wish to model. In each family of matrices, we will define it so that the matrices have a probability measure $P(H)$ satisfying particular properties. We define the entries of the matrices in each ensemble to have diagonal entries $N(0, 1)$ distributed, and off diagonal entries $N(0, 1/2)$ distributed (with $N(0, 1/2)$ distributed random variables in each off diagonal component if we have complex or quaternion entries). For a more detailed explanation of the following constructions and definitions, please refer to [4].

3.1 Gaussian Orthogonal Ensemble

When given a Hamiltonian system H, we can think of the energy in this system as a process over time, and then what it would mean to reverse this process with respect to time. This reversal is known as time reversal. In terms of matrices, the notion of time reversing a system H is given by the transformation $H \to KH^TK^{-1}$, for a unitary matrix K. Intuitively, taking the transpose of H reverses our system and conjugating with K is a change of basis to make our coordinates more tangible. To be time reversal invariant, the energy is the same in both the original system and the time reversed system, i.e., $H = KH^TK^{-1}$. The matrix K is also either symmetric or anti-symmetric, i.e., either $K = K^T$ or $K = -K^T$ respectively. Using the fact K is unitary, this holds if and only if $KK^* = I$ or $KK^* = -I$, where I is the $N \times N$ identity matrix. We refer to these two cases as the even spin and odd spin case respectively. The even spin case with time reversal invariance will be modeled by the GOE and the odd spin case with time reversal invariance will be modeled by the GSE. The terms "even spin" and "odd spin" are based on the notion of angular momentum, a notion we will not elaborate on further as it nor its above matrix formulation, is relevant to our discussion.

For the even spin case, $KK^* = I$ allows us to say without loss of generality that $K = I$, by considering transformations of the form $K \to UKU^T$ and $H \to UHU^T$ (essentially an appropriate change of basis). Then by the time reversal invariance of our Hamiltonian, $H = KH^TK^{-1} = H^T$, which implies H will be a symmetric matrix. To ensure H has real eigenvalues, recall it must have real entries. Also recall we want to be able to assign the same probability to a real symmetric matrix H before and after conjugating with real orthogonal matrices. A model for a time reversal invariant Hamiltonian in the even spin case is thus given by the GOE defined below.

Definition 3.1. The Gaussian Orthogonal Ensemble is the set of real symmetric matrices with a probability distribution P such that $P(H) = P(O^{-1}HO)$ for O orthogonal (invariant under orthogonal conjugation) and $P(H) = \Pi_{j \leqslant k} f(H_{jk})$ for some function f.

Note the second requirement is a statement that the entries of H are independent, but we instead only consider the below diagonal entries to avoid double counting, since H is symmetric so that $H_{jk} = H_{kj}$.

3.2 Gaussian Symplectic Ensemble

For the odd spin case, where $KK^* = -I$, we can take K without loss of generality to be a block diagonal matrix with diagonal blocks equal to $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then recall from our preliminaries that the condition $H = KH^TK^{-1}$ implies H is a self-dual quaternion matrix. With this observation, we therefore model an odd-spin, time reversal invariant system H using self-dual, quaternion valued.

Definition 3.2. The Gaussian Symplectic Ensemble is the set of self-dual Hermitian matrices with a probability distribution P such that $P(H) = P(W^{-}1HW)$ for W self-dual (invariant under self-dual conjugation) and $P(H) = \Pi_{j\leqslant k} f^{(0)}(H_{jk}^{(0)})\Pi_{\lambda=1}^{3}\Pi_{j<k} f^{(\lambda)}(H_{jk}^{(\lambda)})$ for some function f.

Note that since the entries in H are quaternion valued, i.e., of the form $q = a + bi + cj + dk$, we must also take products over each component. Hence the product includes terms of the form $H_{jk}^{(\lambda)}$ for $\lambda = 0, 1, 2, 3$, as opposed to the orthogonal case where each H_{jk} was real valued.

3.3 Gaussian Unitary Ensemble

The notion of time reversal invariance imposed extra conditions on our system, i.e., on our Hamiltonian H. The GUE is much simpler in the sense that it applies to systems without the condition of time reversal invariance, so that it can by modeled by arbitrary Hermitian matrices.

Definition 3.3. The Gaussian Unitary Ensemble is the set of Hermitian matrices with a probability distribution P such that $P(H) = P(U^{-1}HU)$ for U unitary (invariant under unitary conjugation) and $\rho(H) = \Pi_{k\leqslant j} f^{(0)}(H_{jk}^{(0)})\Pi_{j<k} f^{(1)}(H_{jk}^{(1)})$ for some function f.

Again, note that since the entries in H are complex valued, i.e., of the form $q = a + bi$, we must also take products over each component. Hence the product includes terms of the form $H_{jk}^{(\lambda)}$ for $\lambda = 0, 1$, as opposed to the orthogonal case where each H_{jk} was real valued.

4 The Joint Probability Density Function for Gaussian Ensembles

In the previous section, we have defined the Gaussian ensembles with a probability density function that we have not given an explicit expression for. In this section, we find an explicit expression

for the joint density function of the entries for H, and show (at least in the GOE case) that it satisfies the independence and invariance property we want it to.

4.1 Consequence of Independence for the Joint Density Function

First consider the case $H = (H_{jk})_{j,k=1}^{N}$ is from the GOE for the sake of being concrete. Now, finding the joint density of the matrix entries H_{jk} amounts to finding the joint density for $1 \leqslant j \leqslant k \leqslant N$ by the symmetry of H. Since the matrix entries are independent, the joint density function of the entries in H will be the product of the density functions of the individual entries for such j, k. Recall the N diagonal entries are Gaussian random variables with mean 0 and variance 1, and the lower diagonal entries are Gaussian random variables with mean 0 and variance $\frac{1}{2}$. From the preliminaries, these random variables have densities $\exp\left(-\frac{1}{2}x^2\right)$ and e^{-x^2} respectively (up to a normalization constant). Thus the joint density of H is of the form

$$\exp\left(-\frac{1}{2}H_{11}^2\right)\exp\left(-\frac{1}{2}H_{22}^2\right)\cdots\exp\left(-\frac{1}{2}H_{NN}^2\right)\exp\left(-H_{21}^2\right)\exp\left(-H_{31}^2\right)\exp\left(-H_{32}^2\right)\cdots\exp\left(-H_{N(N-1)}^2\right)$$

$$= \exp\left(-\frac{1}{2}H_{11}^2 - \frac{1}{2}H_{22}^2\cdots - \frac{1}{2}H_{NN}^2 - A_{21}^2 - H_{31}^2 - H_{32}^2\cdots - H_{N(N-1)}^2\right)$$

$$= \exp\left(-\frac{1}{2}\sum_{j=1}^{N}H_{jj}^2 - \sum_{1\leqslant j<k\leqslant N}H_{jk}^2\right)$$

But note

$$\mathrm{Tr}(H^2) = -\frac{1}{2}\sum_{j,k=1}^{N}(H^2)_{jk} = -\frac{1}{2}\sum_{j=1}^{N}(H^2)_{jj} - \frac{1}{2}\sum_{1\leqslant j<k\leqslant N}(H^2)_{jk} - \frac{1}{2}\sum_{1\leqslant k<j\leqslant N}(H^2)_{jk}$$

$$= -\frac{1}{2}\sum_{j=1}^{N}(H^2)_{jk} - \sum_{1\leqslant j<k\leqslant N}(H^2)_{jk}$$

where the last equality is by the symmetry of H. Thus we see that the joint density function of H is given by $\exp\left(-\frac{1}{2}\mathrm{Tr}(H^2)\right)$.

Now consider a matrix $H = (H_{jk})_{j,k=1}^{N}$ from the GUE. The computations are almost identical to the above, except that the lower diagonal entries are complex numbers whose real and imaginary parts are Gaussian random variables with mean 0 and variance $\frac{1}{2}$, instead of just being a Gaussian

random variables with mean 0 and variance $\frac{1}{2}$. Thus the density of H is of the form

$$\exp\left(-\frac{1}{2}H_{11}^2\right)\exp\left(-\frac{1}{2}H_{22}^2\right)\cdots\exp\left(-\frac{1}{2}H_{NN}^2\right)\exp\left(-(\operatorname{Re}H_{21})^2\right)\exp\left(-(\operatorname{Im}H_{21})^2\right)\exp\left(-(\operatorname{Re}H_{31})^2\right)$$

$$\cdot\exp\left(-(\operatorname{Im}H_{31})^2\right)\exp\left(-(\operatorname{Re}H_{32})^2\right)\exp\left(-(\operatorname{Im}H_{32})^2\right)\cdots\exp\left(-(\operatorname{Re}H_{N(N-1)})^2\right)\exp\left(-(\operatorname{Im}H_{N(N-1)})^2\right)$$

$$= \exp\Big(-\frac{1}{2}H_{11}^2-\frac{1}{2}H_{22}^2\cdots-\frac{1}{2}H_{NN}^2-(\operatorname{Re}H_{21})^2-(\operatorname{Im}H_{21})^2$$

$$-(\operatorname{Re}H_{31})^2-(\operatorname{Im}H_{31})^2-(\operatorname{Re}H_{32})^2-(\operatorname{Im}H_{32})^2\cdots-(\operatorname{Re}H_{N(N-1)})^2-(\operatorname{Im}H_{N(N-1)})^2\Big)$$

$$= \exp\left(-\frac{1}{2}\sum_{j=1}^{N}H_{jj}^2-\sum_{1\leqslant j<k\leqslant N}\left((\operatorname{Re}H_{jk})^2+(\operatorname{Im}H_{jk}^2)\right)\right)$$

But note

$$\operatorname{Tr}(H^2)=-\frac{1}{2}\sum_{j,k=1}^{N}(H^2)_{jk}=-\frac{1}{2}\sum_{j=1}^{N}(H^2)_{jj}-\frac{1}{2}\sum_{1\leqslant j<k\leqslant N}(H^2)_{jk}-\frac{1}{2}\sum_{1\leqslant k<j\leqslant N}(H^2)_{jk}$$

$$=-\frac{1}{2}\sum_{j=1}^{N}(H^2)_{jk}-\sum_{1\leqslant j<k\leqslant N}(H^2)_{jk}\overline{(H^2)_{jk}}=-\frac{1}{2}\sum_{j=1}^{N}H_{jj}^2-\sum_{1\leqslant j<k\leqslant N}\left((\operatorname{Re}H_{jk})^2+(\operatorname{Im}H_{jk}^2)\right)$$

where the last equality is since H equals its conjugate transpose. Thus we see that the joint density function of H is given by $\exp\left(-\frac{1}{2}\operatorname{Tr}(H^2)\right)$. We obtain $\exp\left(-\frac{1}{2}\operatorname{Tr}(H^2)\right)$ as the density for the case H is from the GSE, where the computation is similar but would need the consideration of two extra components.

4.2 Invariance Under Conjugation Transformations

The results in this section are based on [1]. Recall we can write the probability measure of H as $P(H\in G)=\int_G\rho(H)dH$. We will not discuss precisely what dH means, but it can be thought as the form $\Pi_{j\leqslant k}dH_{jk}$, so that it represents a volume form for the values of the probability of H. Consider the case H is from the GOE. To show the GOE is invariant under orthogonal transformations, we require that for all orthogonal matrices O we have $P(H\in G)=P(O^THO\in G)$. I.e.,

$$\int_G\exp\left(-\frac{1}{2}\operatorname{Tr}(H^2)\right)dH=\int_G\exp\left(-\frac{1}{2}\operatorname{Tr}((O^THO)^2)\right)d(O^THO)$$

Since $\mathrm{Tr}((O^T HO)^2) = \mathrm{Tr}(O^T HOO^T AO) = \mathrm{Tr}(O^T HO) = \mathrm{Tr}(O^T H^2 O) = \mathrm{Tr}(H^2 O^T O) = \mathrm{Tr}(H^2 O^{-1} O) = \mathrm{Tr}(H^2)$, we have the above becomes

$$\int_G \exp\left(-\frac{1}{2}\mathrm{Tr}(H^2)\right) dH = \int_G \exp\left(-\frac{1}{2}\mathrm{Tr}(H^2)\right) d(O^T HO)$$

Thus showing invariance of the probability distribution for the GOE under orthogonal transformations amounts to showing $dH = d(O^T HO)$. This essentially means that if the entries of H have a particular probability distribution, these are preserved in the new matrix after the transformation $H \to O^T HO$. To show this, we have a couple lemmas in order.

Lemma 4.2.1. *Let* $X = (X_1, ..., X_m)^T$ *be a random vector whose components are i.i.d with distribution* $N(0,1)$*, and* $Y = (X_1, ..., X_m)^T$ *where* $Y = TX$ *for a fixed* $m \times m$ *real matrix* T*. Then* Y *is also a random vector whose components are i.i.d with distribution* $N(0,1)$ *if and only if* $T^T T = I$ *(i.e.,* T *is orthogonal).*

We say a $N \times N$ matrix O is an **elementary orthogonal matrix** if

$$O = P^{-1} \begin{pmatrix} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} & 0 \\ 0 & I_{N-2} \end{pmatrix} P$$

where P is a permutation matrix, I_{N-2} is the $N-2 \times N-2$ identity matrix, $\alpha^2 + \beta^2 = 1$ and the zeros in the matrix indicate all remaining entries besides the diagonal blocks are 0. We will reduce the problem of showing invariance by conjugating a matrix from the GOE with an elementary orthogonal matrix instead of conjugating with an arbitrary orthogonal matrix.. This is justified by the following lemma.

Lemma 4.2.2. *Each orthogonal matrix is the product of elementary orthogonal matrices and a diagonal matrix with diagonal entries* ± 1*.*

We also make use of the following lemma.

Lemma 4.2.3. *Let* X *be the matrix*

$$X = \begin{pmatrix} X_1 & \frac{1}{\sqrt{2}} X_{12} \\ \frac{1}{\sqrt{2}} X_{12} & X_2 \end{pmatrix}.$$

where X_1, X_2, X_{12} *are i.i.d random variables with distribution* $N(0,1)$. *Let* O *be the matrix*

$$O = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

where $\alpha^2 + \beta^2 = 1$, *and* $Y = O^T X O$. *Then* Y *can be written in the form*

$$Y = \begin{pmatrix} Y_1 & \frac{1}{\sqrt{2}} Y_{12} \\ \frac{1}{\sqrt{2}} Y_{12} & Y_2 \end{pmatrix}$$

and Y_1, Y_2, Y_{12} *are i.i.d random variables with distribution* $N(0,1)$.

Proof. Observe $Y^T = (O^T X O)^T = O^T X^T (O^T)^T = O^T X^T O = O^T X O = Y$ since X is symmetric, so that Y is also symmetric. Adding a factor of $\frac{1}{\sqrt{2}}$ gives us the above matrix form of Y. Now observe

$$XO = \begin{pmatrix} X_1 & \frac{1}{\sqrt{2}} X_{12} \\ \frac{1}{\sqrt{2}} X_{12} & X_2 \end{pmatrix} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} \alpha X_1 + \beta \frac{1}{\sqrt{2}} X_{12} & -\beta X_1 + \alpha \frac{1}{\sqrt{2}} X_{12} \\ \alpha \frac{1}{\sqrt{2}} X_{12} + \beta X_2 & -\beta \frac{1}{\sqrt{2}} X_{12} + \alpha X_2 \end{pmatrix}$$

So that

$$\begin{pmatrix} Y_1 & \frac{1}{\sqrt{2}} Y_{12} \\ \frac{1}{\sqrt{2}} Y_{12} & Y_2 \end{pmatrix} = O^T X O = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} \alpha X_1 + \beta \frac{1}{\sqrt{2}} X_{12} & -\beta X_1 + \alpha \frac{1}{\sqrt{2}} X_{12} \\ \alpha \frac{1}{\sqrt{2}} X_{12} + \beta X_2 & -\beta \frac{1}{\sqrt{2}} X_{12} + \alpha X_2 \end{pmatrix}$$

$$= \begin{pmatrix} \alpha^2 X_1 + \alpha\beta \frac{1}{\sqrt{2}} X_{12} + \beta\alpha \frac{1}{\sqrt{2}} X_{12} + \beta^2 X_2 & -\alpha\beta X_1 + \alpha^2 \frac{1}{\sqrt{2}} X_{12} - \beta^2 \frac{1}{\sqrt{2}} X_{12} + \beta\alpha X_2 \\ -\beta\alpha X_1 - \beta^2 \frac{1}{\sqrt{2}} X_{12} + \alpha^2 \frac{1}{\sqrt{2}} X_{12} + \alpha\beta X_2 & \beta^2 X_1 - \beta\alpha \frac{1}{\sqrt{2}} X_{12} - \alpha\beta \frac{1}{\sqrt{2}} X_{12} + \alpha^2 X_2 \end{pmatrix}$$

$$= \begin{pmatrix} \alpha^2 X_1 + \sqrt{2}\beta\alpha X_{12} + \beta^2 X_2 & -\alpha\beta X_1 + \frac{1}{\sqrt{2}}(\alpha^2 - \beta^2) X_{12} + \beta\alpha X_2 \\ -\beta\alpha X_1 + \frac{1}{\sqrt{2}}(\alpha^2 - \beta^2) X_{12} + \alpha\beta X_2 & \beta^2 X_1 - \sqrt{2}\beta\alpha \frac{1}{\sqrt{2}} X_{12} + \alpha^2 X_2 \end{pmatrix}$$

Thus

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_{12} \end{pmatrix} = \begin{pmatrix} \alpha^2 & \beta^2 & \sqrt{2}\alpha\beta \\ \beta^2 & \alpha^2 & -\sqrt{2}\alpha\beta \\ -\sqrt{2}\alpha\beta & \sqrt{2}\alpha\beta & \alpha^2 - \beta^2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_{12} \end{pmatrix}$$

Letting the above transformation matrix be denoted by T, it follows that $T^T T = I$ so that the result

follows from the first lemma of this section. □

We are now ready to prove that the probability distribution on the GOE is invariant under orthogonal transformations.

Proof. We can instead show invariance under transformations by elementary orthogonal matrices and diagonal matrices by the second of the previous lemmas, and then apply the first and third of our previous lemmas. Let H be from the GOE. Then the diagonal entries have $N(0,1)$ distribution and the off diagonal entries have $N(0,1/2)$ distribution. We can write H as

$$\begin{pmatrix} \begin{pmatrix} H_1 & \frac{1}{\sqrt{2}}X_{12} \\ \frac{1}{\sqrt{2}}X_{12} & H_2 \end{pmatrix} & \begin{pmatrix} \frac{1}{\sqrt{2}}X_{13} & \frac{1}{\sqrt{2}}X_{14} & \cdots \\ \frac{1}{\sqrt{2}}X_{23} & \frac{1}{\sqrt{2}}X_{24} & \cdots \end{pmatrix} \\ \begin{pmatrix} \frac{1}{\sqrt{2}}X_{13} & \frac{1}{\sqrt{2}}X_{23} \\ \frac{1}{\sqrt{2}}X_{14} & \frac{1}{\sqrt{2}}X_{24} \\ \vdots & \vdots \end{pmatrix} & \begin{pmatrix} H_3 & & \\ & \ddots & \\ & & H_N \end{pmatrix} \end{pmatrix}$$

where each X_{jk} (without the factor of $\frac{1}{\sqrt{2}}$) will be $N(0,1)$ distributed. Without loss of generality, one can take $P = I$ in the definition of elementary orthogonal matrices, so that we can consider

$$Y := \begin{pmatrix} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} & 0 \\ 0 & I_{N-2} \end{pmatrix} X \begin{pmatrix} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} & 0 \\ 0 & I_{N-2} \end{pmatrix}$$

We let

$$Y = \begin{pmatrix} \begin{pmatrix} Y_1 & \frac{1}{\sqrt{2}}Y_{12} \\ \frac{1}{\sqrt{2}}Y_{12} & Y_2 \end{pmatrix} & \begin{pmatrix} \frac{1}{\sqrt{2}}Y_{13} & \frac{1}{\sqrt{2}}Y_{14} & \cdots \\ \frac{1}{\sqrt{2}}Y_{23} & \frac{1}{\sqrt{2}}Y_{24} & \cdots \end{pmatrix} \\ \begin{pmatrix} \frac{1}{\sqrt{2}}Y_{13} & \frac{1}{\sqrt{2}}Y_{23} \\ \frac{1}{\sqrt{2}}Y_{14} & \frac{1}{\sqrt{2}}Y_{24} \\ \vdots & \vdots \end{pmatrix} & \begin{pmatrix} Y_3 & & \\ & \ddots & \\ & & Y_N \end{pmatrix} \end{pmatrix}$$

Note that the lower right block is unaffected because we multiply it with I_{N-2}, so that $Y_3, ..., Y_N$

are i.i.d with distribution $N(0,1)$. Observe

$$\begin{pmatrix} Y_{1m} \\ Y_{2m} \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} H_{1m} \\ H_{2m} \end{pmatrix}$$

for each $m = 1, ..., N$. Thus using this fact and the matrix from the proof of the previous lemma, we have

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_{12} \\ Y_{13} \\ Y_{23} \\ \vdots \\ Y_{1n} \\ Y_{2n} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \alpha^2 & \beta^2 & \sqrt{2}\alpha\beta \\ \beta^2 & \alpha^2 & -\sqrt{2}\alpha\beta \\ -\sqrt{2}\alpha\beta & \sqrt{2}\alpha\beta & \alpha^2 - \beta^2 \end{pmatrix} & & & \\ & \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} & & \\ & & \ddots & \\ & & & \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_{12} \\ X_{13} \\ X_{23} \\ \vdots \\ X_{1n} \\ X_{2n} \end{pmatrix}$$

with zeros in the remaining entries in the above matrix transformation. Since each block is an orthogonal matrix, the whole matrix is orthogonal. Thus, since $X_1, X_2, X_{12}, X_{13}, X_{23}, ..., X_{1n}, X_{2n}$ are i.i.d with distribution $N(0,1)$, we have that $Y_1, Y_2, Y_{12}, Y_{13}, Y_{23}, ..., Y_{1n}, Y_{2n}$ are i.i.d with distribution $N(0,1)$ by our first lemma. Since the entries from the lower right block in Y are also i.i.d with distribution $N(0,1)$, hence showing the result. □

5 The Joint Probability Density Function for the Eigenvalues of Gaussian Ensembles

Given an expression for $P(H)$, we are now ready to obtain our ultimate goal of the joint probability density function for the eigenvalues of a matrix from the GOE, GSE and GUE. This result is summarized in Theorem 5.2.1. The following arguments are based on [1]. For a more detailed explanation of the proof of Theorem 5.2.1, please refer to [4].

5.1 Parametrizing The Matrix Entries

Let $H = (H_{jk})_{j,k=1}^{N}$ be a matrix from the GOE so that it is real symmetric, and $\lambda_1, ..., \lambda_N$ its eigenvalues. Consider the number of independent parameters for the matrix entries. There are N diagonal entries, and $\frac{N(N-1)}{2}$ lower diagonal entries that also give us the above diagonal entries by the symmetry of H. Thus H can be parametrized by $N + \frac{N(N-1)}{2} = \frac{N(N+1)}{2}$ independent parameters. Since H is from the GOE it can be diagonalized by an orthogonal matrix O as follows

$$H = O^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} O,$$

where the matrix between O^T and O is diagonal and has zeros in the remaining entries. We want to obtain the $\frac{N(N+1)}{2}$ parameters for A in terms of the entries of the O and diagonal matrix above. There are N eigenvalues that give us N parameters from the diagonal entries. The remaining $l := \frac{N(N+1)}{2} - N = \frac{N(N-1)}{2}$ come from O after appropriate transformation. Thus we can parametrize the entries of H by $\lambda_1, ..., \lambda_N$ and $p_1, ..., p_l$ parameters.

Now let $H = (H_{jk})_{j,k=1}^{N}$ be a matrix from the GUE, so that it is Hermitian. We have $N(N)$ total entries, so that the number of extra parameters to account for in addition to the N eigenvalues will be $N(N) - N = N(N-1)$. Hence in this case $l = N(N-1)$. Note we cannot only look at the lower triangular entries as in the last example since an arbitrary Hermitian matrix is not symmetric. Similarly to above, we can obtain the N parameters through the eigenvalues and the remaining $N(N-1)$ parameters through a unitary matrix after diagonalizing H with a unitary matrix U as follows

$$H = U^* \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} U.$$

Now let $H = (H_{jk})_{j,k=1}^{N}$ be a matrix from the GSE, so that is it quaternion self-dual. We have $N(N)$ total entries, so that the number of extra entries to account for in addition to the N eigenvalues will be $N(N) - N = N(N-1)$. But since we only need to account for the lower diagonal entries by the fact H is self-dual we have $\frac{N(N-1)}{2}$ entries in our consideration thus far. However,

since each entry is quaternion and has four components, we have $l = 4 \cdot \frac{N(N-1)}{2} = 2N(N-1)$ independent parameters we need to parameterize our matrix in addition to our N eigenvalues.

5.2 Calculating the Joint Density for the Eigenvalues

We will now derive the joint density function for the eigenvalues of H when H is in the GOE, GUE or GSE. Let $\lambda_1, ..., \lambda_N$ be the real eigenvalues of H. Let $\Lambda \subseteq \mathbb{R}^N$, $\lambda = (\lambda_1, ..., \lambda_N) \in \mathbb{R}^N$, and $\theta = (\theta_1, ..., \theta_l)$ be the vector containing the remaining l parameters. By the continuity of the entries of H, the joint eigenvalue density function $\rho(\lambda_1, ..., \lambda_N)$ will satisfy

$$P((\lambda_1, ..., \lambda_N) \in \Lambda) = \int_\Lambda \rho(\lambda_1, ..., \lambda_N) d\lambda$$

But note, recalling the density for a matrix H from the GOE,

$$P((\lambda_1, ..., \lambda_N) \in \Lambda) = \int_{\{H\text{: the eigenvalues of } H \text{ are contained in } \Lambda\}} \exp\left(-\frac{1}{2}\operatorname{Tr}(H^2)\right) dH$$

Now consider the function $(H_{jk})_{j \leqslant k} \to (\lambda, \theta)$, i.e., our parametrization of the matrix entries. This change of variables yields the Jacobian matrices

$$J(\lambda, \theta) = \det\begin{pmatrix} \frac{\partial H_{jk}}{\partial \lambda_i} & \frac{\partial H_{jk}}{\partial \theta_l} \end{pmatrix}$$

$$J(\lambda, \theta) = \det\begin{pmatrix} \frac{\partial (\operatorname{Re} H)_{jk}}{\partial \lambda_i} & \frac{\partial (\operatorname{Im} H)_{jk}}{\partial \lambda_i} & \frac{\partial (\operatorname{Re} H)_{jk}}{\partial \theta_l} & \frac{\partial (\operatorname{Im} H)_{jk}}{\partial \theta_l} \end{pmatrix}$$

$$J(\lambda, \theta) = \det\begin{pmatrix} \frac{\partial H^{(1)}_{jk}}{\partial \lambda_i} & \frac{\partial H^{(2)}_{jk}}{\partial \lambda_i} & \frac{\partial H^{(3)}_{jk}}{\partial \lambda_i} & \frac{\partial H^{(4)}_{jk}}{\partial \lambda_i} & \frac{\partial H^{(1)}_{jk}}{\partial \theta_l} & \frac{\partial H^{(2)}_{jk}}{\partial \theta_l} & \frac{\partial H^{(3)}_{jk}}{\partial \theta_l} & \frac{\partial H^{(4)}_{jk}}{\partial \theta_l} \end{pmatrix}$$

in the GOE, GUE and GSE case respectively, where $(H^{(m)}_{jk}$ is the m^{th} component of the quaternion entry H_{jk} in the GSE case, $i = 1, ..., N$ and l is the number of remaining parameters in addition to the N eigenvalues. Thus, letting R denote the space of all possible values for θ, the above integral becomes

$$= \int_\Lambda \int_R \exp\left(-\frac{1}{2}(\lambda_1^2 + \cdots + \lambda_N^2)\right) |J(\lambda, \theta)| \, d\theta d\lambda = \int_\Lambda \exp\left(-\frac{1}{2}(\lambda_1^2 + \cdots + \lambda_N^2)\right) \int_R |J(\lambda, \theta)| \, d\theta d\lambda$$

Thus by equating our two expressions for $P((\lambda_1, ..., \lambda_N) \in \Lambda)$ we must have

$$\rho(\lambda_1, ..., \lambda_N) = \exp\left(-\frac{1}{2}(\lambda_1^2 + \cdots + \lambda_N^2)\right) \int_R |J(\lambda, \theta)| \, d\theta \tag{1}$$

We now give an example to illustrate the computation of $|J(\lambda, \theta)| \, d\theta$, thus obtaining an expression for $\rho(\lambda_1, ..., \lambda_N)$.

Example 5.2.1. *Let* H *be a* 2×2 *matrix given by* $H = O^T \Lambda O$ *where*

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ and } O = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for $\theta \in (-\pi, \pi)$ *and real random variables* λ_1, λ_2. *Then*

$$\Lambda O = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \lambda_1 \cos\theta & -\lambda_1 \sin\theta \\ \lambda_2 \sin\theta & \lambda_2 \cos\theta \end{pmatrix}$$

So that

$$H = O^T \Lambda O = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \lambda_1 \cos\theta & -\lambda_1 \sin\theta \\ \lambda_2 \sin\theta & \lambda_2 \cos\theta \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 \cos\theta \cos\theta + \lambda_2 \sin\theta \sin\theta & -\lambda_1 \cos\theta \sin\theta + \lambda_2 \cos\theta \sin\theta \\ -\lambda_1 \cos\theta \sin\theta + \lambda_2 \cos\theta \sin\theta & \lambda_1 \sin\theta \sin\theta + \lambda_2 \cos\theta \cos\theta \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta & (\lambda_2 - \lambda_1) \cos\theta \sin\theta \\ (\lambda_2 - \lambda_1) \cos\theta \sin\theta & \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta \end{pmatrix}$$

Note H *is real symmetric so that it is a member of the GOE. Let*

$$\begin{pmatrix} H_{11} & H_{21} \\ H_{12} & H_{22} \end{pmatrix} := H = \begin{pmatrix} \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta & (\lambda_2 - \lambda_1) \cos\theta \sin\theta \\ (\lambda_2 - \lambda_1) \cos\theta \sin\theta & \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta \end{pmatrix}$$

Then the Jacobian matrix is given by

$$\begin{pmatrix} \frac{\partial H_{11}}{\partial \lambda_1} & \frac{\partial H_{11}}{\partial \lambda_2} & \frac{\partial H_{11}}{\partial \theta} \\ \frac{\partial H_{22}}{\partial \lambda_1} & \frac{\partial H_{22}}{\partial \lambda_2} & \frac{\partial H_{22}}{\partial \theta} \\ \frac{\partial H_{12}}{\partial \lambda_1} & \frac{\partial H_{12}}{\partial \lambda_2} & \frac{\partial H_{12}}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos^2\theta & \sin^2\theta & 2\lambda_1\cos\theta(-\sin\theta) + 2\lambda_2\sin\theta\cos\theta \\ \sin^2\theta & \cos^2\theta & 2\lambda_1\sin^2\theta\cos\theta + 2\lambda_2\cos\theta(-\sin\theta) \\ -\cos\theta\sin\theta & \cos\theta\sin\theta & (\lambda_2-\lambda_1)(-\sin\theta\sin\theta + \cos\theta\cos\theta) \end{pmatrix}$$

$$= \begin{pmatrix} \cos^2\theta & \sin^2\theta & 2(\lambda_2-\lambda_1)\cos\theta\sin\theta \\ \sin^2\theta & \cos^2\theta & 2(\lambda_1-\lambda_2)\cos\theta\sin\theta \\ -\cos\theta\sin\theta & \cos\theta\sin\theta & (\lambda_2-\lambda_1)(\cos^2\theta - \sin^2\theta) \end{pmatrix}$$

Hence

$$J(\lambda,\theta) = \det\begin{pmatrix} \cos^2\theta & \sin^2\theta & 2(\lambda_2-\lambda_1)\cos\theta\sin\theta \\ \sin^2\theta & \cos^2\theta & 2(\lambda_1-\lambda_2)\cos\theta\sin\theta \\ -\cos\theta\sin\theta & \cos\theta\sin\theta & (\lambda_2-\lambda_1)(\cos^2\theta - \sin^2\theta) \end{pmatrix}$$

$$= \cos^2\theta\det\begin{pmatrix} \cos^2\theta & 2(\lambda_1-\lambda_2)\cos\theta\sin\theta \\ \cos\theta\sin\theta & (\lambda_2-\lambda_1)(\cos^2\theta-\sin^2\theta) \end{pmatrix} - \sin^2\theta\det\begin{pmatrix} \sin^2\theta & 2(\lambda_1-\lambda_2)\cos\theta\sin\theta \\ -\cos\theta\sin\theta & (\lambda_2-\lambda_1)(\cos^2\theta-\sin^2\theta) \end{pmatrix}$$

$$+2(\lambda_2-\lambda_1)\cos\theta\sin\theta\det\begin{pmatrix} \sin^2\theta & \cos^2\theta \\ -\cos\theta\sin\theta & \cos\theta\sin\theta \end{pmatrix}$$

$$= \cos^2\theta\left((\lambda_2-\lambda_1)(\cos^4\theta - \sin^2\theta\cos^2\theta) - 2(\lambda_1-\lambda_2)\cos^2\theta\sin^2\theta\right)$$

$$-\sin^2\theta\left((\lambda_2-\lambda_1)(\sin^2\theta\cos^2\theta - \sin^4\theta) + 2(\lambda_1-\lambda_2)\cos^2\theta\sin^2\theta\right)$$

$$+2(\lambda_2-\lambda_1)\cos\theta\sin\theta\left(\cos\theta\sin^3\theta + \cos^3\theta\sin\theta\right)$$

$$= \left((\lambda_2-\lambda_1)(\cos^6\theta - \sin^2\theta\cos^4\theta) - 2(\lambda_1-\lambda_2)\cos^4\theta\sin^2\theta\right)$$

$$-\left((\lambda_2-\lambda_1)(\sin^4\theta\cos^2\theta - \sin^6\theta) + 2(\lambda_1-\lambda_2)\cos^2\theta\sin^4\theta\right)$$

$$+\left(2(\lambda_2-\lambda_1)\cos^2\theta\sin^4\theta + 2(\lambda_2-\lambda_1)\cos^4\theta\sin^2\theta\right)$$

$$= (\lambda_2-\lambda_1)(\cos^6\theta - \sin^2\theta\cos^4\theta) + 2(\lambda_2-\lambda_1)\cos^4\theta\sin^2\theta$$

$$+2(\lambda_2-\lambda_1)\cos^2\theta\sin^4\theta - (\lambda_2-\lambda_1)(\sin^4\theta\cos^2\theta - \sin^6\theta)$$

$$+2(\lambda_2-\lambda_1)\cos^2\theta\sin^4\theta+2(\lambda_2-\lambda_1)\cos^4\theta\sin^2\theta$$

$$=(\lambda_2-\lambda_1)(\cos^6\theta-\sin^2\theta\cos^4\theta)+4(\lambda_2-\lambda_1)\cos^4\theta\sin^2\theta$$

$$+4(\lambda_2-\lambda_1)\cos^2\theta\sin^4\theta-(\lambda_2-\lambda_1)(\sin^4\theta\cos^2\theta-\sin^6\theta)$$

$$=(\lambda_2-\lambda_1)\Big[(\cos^6\theta-\sin^2\theta\cos^4\theta)+4\cos^4\theta\sin^2\theta+4\cos^2\theta\sin^4\theta-(\sin^4\theta\cos^2\theta-\sin^6\theta)\Big]$$

Let $f(\theta)$ *be the function in the brackets that only depends on* θ*. Thus observe*

$$\rho(\lambda_1,\lambda_2)=\exp\left(-\frac{1}{2}(\lambda_1^2+\lambda_2^2)\right)\int_R|J(\lambda,\theta)|\,d\theta=\exp\left(-\frac{1}{2}(\lambda_1^2+\lambda_2^2)\right)\int_{-\pi}^{\pi}|(\lambda_2-\lambda_1)f(\theta)|\,d\theta$$

$$=\exp\left(-\frac{1}{2}(\lambda_1^2+\lambda_2^2)\right)|\lambda_2-\lambda_1|\int_{-\pi}^{\pi}|f(\theta)|\,d\theta=C\exp\left(-\frac{1}{2}(\lambda_1^2+\lambda_2^2)\right)|\lambda_2-\lambda_1|$$

for some constant C*.*

The fact the Jacobian matrix ends up in the form $C|\lambda_2-\lambda_1|$ above is not an anomaly, and actually highlights a specific case of a more general phenomenon. This is shown in the following theorem.

Theorem 5.2.1. *The joint probability density function for the eigenvalues of a random matrix from a GOE, GSE and GUE is given by*

$$\rho(\lambda_1,...,\lambda_N)=C_{N,\beta}\exp\left(-\frac{1}{2}\beta\sum_{j=1}^{N}\lambda_j^2\right)\Pi_{j<k}\left|x_j-x_k\right|^{\beta}$$

where $\beta=1$*,* $\beta=4$*,* $\beta=2$ *respectively, and* $C_{N,\beta}$ *is chosen so that*

$$\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}\rho(\lambda_1,...,\lambda_N)\,dx_1\cdots dx_n=1$$

We will prove this theorem in the GOE case. Before doing so, we note that the value of β corresponds to the number of components in the entries of a matrix from the GOE, GSE and GUE respectively. From a physical point of view, the term $\Pi_{j<k}\left|x_j-x_k\right|^{\beta}$ indicates there is an inherent repelling effect in the statistical behavior of the eigenvalues.

Proof. Let H be in the GOE with eigenvalues $\lambda_1, ..., \lambda_N$. To show

$$\rho(\lambda_1, ..., \lambda_N) = C_N \exp\left(-\frac{1}{2}\sum_{j=1}^{N} \lambda_j^2\right) \Pi_{j<k} \left|x_j - x_k\right|$$

it suffices to show $\int_R |J(\lambda, \theta)| \, d\theta$ as in (1) of this section equals $C\Pi_{j<k} \left|x_j - x_k\right|$ for some constant C. Given H is in the GOE, it can be diagonalized with an orthogonal matrix as follows:

$$H = O^T \Lambda O = O^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} O$$

where $O^T O = I$, $i = 1, ..., N$ and $l = 1, ..., \frac{N(N-1)}{2}$. Let E_i be the diagonal matrix such that there is a 1 in the i^{th} diagonal entry and zeros elsewhere. We recall that differentiation of a matrix is done componentwise and respects the linearity of matrix multiplication. Then for each i we see

$$\frac{\partial H}{\partial \lambda_i} = \frac{\partial}{\partial \lambda_i}\left(O^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} O\right) = O^T \left(\frac{\partial}{\partial \lambda_i}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}\right) O = O^T E_i O \tag{2}$$

Recall from our parametrization of H, the extra parameters $\theta_1, .., \theta_N$ come from O and not Λ. Thus Λ does not depend on θ_l for each l. Hence by the product rule, we have for each l that

$$\frac{\partial H}{\partial \theta_l} = \frac{\partial O^T}{\partial \theta_l}\Lambda O + O^T \Lambda \frac{\partial O}{\partial \theta_l} = \frac{\partial O^T}{\partial \theta_l} O O^T \Lambda O + O^T \Lambda \frac{\partial O}{\partial \theta_l}$$

Using the fact $\frac{\partial O^T}{\partial \theta_l} O + O^T \frac{\partial O}{\partial \theta_l}$ is the zero matrix, which implies $\frac{\partial O^T}{\partial \theta_l} O = -O^T \frac{\partial O}{\partial \theta_l}$, we have that the above becomes

$$\frac{\partial H}{\partial \theta_l} = -O^T \frac{\partial O}{\partial \theta_l} O^T \Lambda O + O^T \Lambda \frac{\partial O}{\partial \theta_l} = O^T \Lambda \frac{\partial O}{\partial \theta_l} - O^T \frac{\partial O}{\partial \theta_l} O^T \Lambda O$$

$$\implies O\frac{\partial A}{\partial \theta_l} = O O^T \Lambda \frac{\partial O}{\partial \theta_l} - O O^T \frac{\partial O}{\partial \theta_l} O^T \Lambda O = \Lambda \frac{\partial O}{\partial \theta_l} - \frac{\partial O}{\partial \theta_l} O^T \Lambda O$$

$$\implies O\frac{\partial H}{\partial \theta_l} O^T = \Lambda \frac{\partial O}{\partial \theta_l} O^T - \frac{\partial O}{\partial \theta_l} O^T \Lambda O O^T = \Lambda \frac{\partial O}{\partial \theta_l} O^T - \frac{\partial O}{\partial \theta_l} O^T \Lambda \tag{3}$$

Letting $C_l = \frac{\partial O}{\partial \theta_l} O^T$, we have by (2) and (3) that

$$O\frac{\partial H}{\partial \lambda_i}O^T = E_i \text{ and } O\frac{\partial H}{\partial \theta_l}O^T = \Lambda C_l - C_l\Lambda$$

for each i and l. By the fact the determinant of an orthogonal matrix is ± 1 and that the determinant is multiplicative, we have

$$|J(\lambda,\theta)| = \left|\det\begin{pmatrix} \frac{\partial H_{jk}}{\partial \lambda_i} & \frac{\partial H_{jk}}{\partial \theta_l}\end{pmatrix}\right| = \left|\det\begin{pmatrix} (E_i)_{jk} & (\Lambda C_l - C_l\Lambda)_{jk}\end{pmatrix}\right|$$

$$= \left|\det\begin{pmatrix} \delta_{i,j}\delta_{i,k} & (\lambda_j - \lambda_k)(C_l)_{jk}\end{pmatrix}\right| = \Pi_{j<k}\left|\lambda_j - \lambda_k\right| f(\theta)$$

where f is a function depending only on θ. Thus $\int_R |J(\lambda,\theta)|\,d\theta = \int_R \Pi_{j<k}\left|\lambda_j - \lambda_k\right| f(\theta) d\theta = \Pi_{j<k}\left|\lambda_j - \lambda_k\right| \int_R f(\theta) d\theta = C\Pi_{j<k}\left|\lambda_j - \lambda_k\right|$ for some constant C. □

We make the following connection between the above theorem to $X_1, ..., X_N$ independent and identically distributed random variables each with $N(0,1)$ distribution. By the identically distributed assumption, each X_j has density $\rho_{X_j}(t_j) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t_j^2\right)$. Since $X_1, ..., X_n$ are independent and continuous, the joint density function for $X_1, ..., X_N$ is given by

$$\rho(t_1, ..., t_N) = \rho_{X_1}(t_1)\cdots\rho_{X_N}(t_N) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t_1^2\right)\cdots\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t_N^2\right)$$

$$= \frac{1}{(2\pi)^{N/2}}\exp\left(\left(-\frac{1}{2}t_1^2\right) + \cdots + \left(-\frac{1}{2}t_1^2\right)\right) = \frac{1}{(2\pi)^{N/2}}\exp\left(-\frac{1}{2}\sum_{j=1}^{N} t_j^2\right)$$

which is exactly the exponential term in the joint density function for the eigenvalues of a random matrix from the GOE.

6 Summary and Consequence

We began by asking the question of how nuclear energy levels are distributed, and noted that quantum theory allows us to consider the asymptotics of a Hermitian matrix to model these energy levels. We then gave the foundational notions of probability theory, and most importantly considered the joint probability density function, which allows us to calculate probabilities of ran-

dom variables taking on given values, as well as types of matrices we would encounter in the three ensembles.

We described the notion of time reversal invariance for an energy system H to define the GOE in the even spin case and the GSE in the odd spin case, and used standard Hermitian matrices to define the GUE. We derived an expression for the joint density in each case by the assumption of independence, and discussed the invariance under conjugation of the probability distribution. By parametrizing H, one can obtain the joint probability density functions for the eigenvalues of H in the GOE, GSE and GUE case. With this joint probability density function $P(\theta_1, ..., \theta_N)$ and the knowledge that the distribution of eigenvalues is the same as the distribution of energy levels the Hamiltonian H represents as $N \to \infty$, we can ask what the probability is of energy being certain values by instead looking at what $\rho(\lambda_1, ...\lambda_N)$ equals as $N \to \infty$.

References

[1] Torsten Ehrhardt. Math 233 lectures, 2022.

[2] H. I. Liou, H. S. Camarda, S. Wynchank, M. Slagowitz, G. Hacken, F. Rahn, and J. Rainwater. Neutron-resonance spectroscopy. *Phys. Rev. C*, 5:974–1001, Mar 1972.

[3] Yi-Kai Liu. Statistical behavior of the eigenvalues of random matrices. 2001.

[4] Madan Lal Mehta. *Random matrices*. Elsevier, 2004.