Derivation of the Wigner surmise

Abstract

We derive the standard Wigner surmise for the Gaussian Orthogonal Ensemble (GOE, Dyson index $\beta = 1$). The steps are explicit so every constant can be checked. At the end we also state the result for $\beta = 2$ (GUE).

1 Why to study the level spacing distribution? BGS conjecture

The study of level spacing distributions in quantum chaotic systems is pivotal for understanding the underlying quantum dynamics and their classical counterparts. Level spacing distribution provides crucial insights into the energy spectrum statistics of a quantum system, which can reveal whether the system exhibits chaotic behavior. According to the Bohigas-Giannoni-Schmit (BGS) conjecture, quantum systems whose classical analogs are chaotic display spectral statistics that align with predictions from random matrix theory (RMT). This conjecture serves as a bridge between quantum chaos and RMT, suggesting that the energy levels of quantum chaotic systems are correlated in a manner similar to the eigenvalues of random matrices.

Comparing the level spacing distribution of a quantum chaotic system with RMT predictions not only tests the validity of the BGS conjecture but also enhances our comprehension of quantum chaos as a whole. Such comparisons can help distinguish between truly chaotic systems and those that might merely exhibit random or pseudo-random behavior. Furthermore, exploring deviations from RMT predictions can signal the presence of novel physical phenomena or symmetries not accounted for in classical descriptions, thereby enriching our understanding of quantum systems in diverse physical contexts.

2 The appropriate probability measure for Time-Reversal Symmetric Systems. Gaussian Orthogonal Ensemble (GOE)

In quantum physics, a time-reversal symmetric quantum chaotic system is often modeled using the Gaussian Orthogonal Ensemble (GOE) of random matrices. This is because the GOE captures the essential statistical properties of the spectrum of such systems. Time-reversal symmetry implies that the Hamiltonian, which governs the system's dynamics, is invariant under time reversal. In mathematical terms, this means the Hamiltonian is represented by a real symmetric matrix. The GOE consists of real symmetric matrices with elements that are random variables following a Gaussian distribution. This ensemble is invariant under orthogonal transformations, which means any such transformation applied to a GOE matrix results in another matrix belonging to the GOE. This invariance reflects the lack of any preferential direction or orientation in the system, which is a hallmark of chaotic behavior. Thus, the GOE provides a natural and effective framework for modeling the statistical properties of time-reversal symmetric quantum chaotic systems.

2.1 Why orthogonal?

To demonstrate that time-reversal symmetry implies the Hamiltonian is represented by a real symmetric matrix, we begin by considering the properties of time-reversal symmetry in quantum mechanics. In general, the time-reversal operator $\mathcal{T}$ is an anti-unitary operator, which can be expressed as $\mathcal{T} = UK$, where U is a unitary operator and K denotes complex conjugation.

For a quantum system to be time-reversal symmetric, the Hamiltonian H must satisfy the condition $\mathcal{T}H\mathcal{T}^{-1} = H$. If we apply $\mathcal{T} = UK$ to the Hamiltonian, we have:

$$\mathcal{T}H\mathcal{T}^{-1} = UKH(UK)^{-1} = UKH(KU^{-1}).$$

Since K denotes complex conjugation, for any operator A, we have $KAK = A^*$. Thus, the above expression becomes:

$$UKH(KU^{-1}) = U(H^*)U^{-1}.$$

The condition for time-reversal symmetry then reads:

$$U(H^*)U^{-1} = H.$$

This implies that H^* and H are related via a similarity transformation involving a unitary operator U. To satisfy this condition for all possible U, H must be such that the transformation results in the same operator, i.e., $H^* = H$. Therefore, the Hamiltonian must be real (since $H^* = H$ implies H is real) and symmetric ($H = H^\dagger$ and for real matrices, $H = H^T$).

Consequently, a time-reversal symmetric Hamiltonian in quantum mechanics is represented by a real symmetric matrix. This is consistent with the properties of the Gaussian Orthogonal Ensemble (GOE), which consists of real symmetric matrices following a Gaussian distribution. The GOE captures the statistical properties of time-reversal symmetric quantum chaotic systems by reflecting the invariance under orthogonal transformations and the absence of a preferential direction or orientation in the system.

2.2 Why gaussian?

A random matrix H from the Gaussian Orthogonal Ensemble (GOE) is defined such that:

$$H_{ij} \sim \begin{cases} \mathcal{N}(0,1), & \text{if } i = j, \\ \mathcal{N}(0,1/2), & \text{if } i \neq j, \end{cases} \tag{1}$$

where $\mathcal{N}(\mu, \sigma^2)$ represents a Gaussian distribution with mean μ and variance σ^2. The term "Gaussian" in GOE comes from this distribution characteristic. Let us show in this section how to arrive at Eq. (1) from first principles. We will only assume that we are drawing matrices from an ensemble that is invariant under orthogonal transformations. In other words, if we select a matrix from this ensemble and subsequently apply an orthogonal transformation, the result is another matrix that also belongs to the same ensemble.

For a set of matrices to be a truly statistical ensemble we need to endow it with a probability measure. We are looking for such a probability measure $P(H)$ that is invariant under orthogonal transformations O:

$$P(OHO^T) = P(H). \tag{2}$$

Let us consider 2×2 matrices. An arbitrary real and symmetric matrix H writes

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix}. \tag{3}$$

In this case, the probability measure is the product of the probabilities for each matrix element:

$$P(H) = p_{11}(H_{11})p_{12}(H_{12})p_{22}(H_{22}). \tag{4}$$

Then, an orthogonal transformation writes

$$O = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Since orthogonal transformations form a group, which is denoted $O(d)$, the composition of two orthogonal transformations is another orthogonal transformation. Therefore, it suffices to consider only infinitesimal orthogonal transformations, as every other transformation can be composed as infinitely many infinitesimal orthogonal transformations. Thus,

$$\lim_{\theta \to 0} O = \begin{pmatrix} 1 & -\theta \\ \theta & 1 \end{pmatrix}. \tag{5}$$

If we denote $H' = \lim_{\theta\to 0} OHO^T$ one obtains:

$$\begin{aligned} H'_{11} =& H_{11} - 2\theta H_{12} && (6)\\ H'_{12} =& H_{12} + \theta H_{12}(H_{11} - H_{22}) && (7)\\ H'_{22} =& H_{22} + 2\theta H_{22}. && (8) \end{aligned}$$

Now let us investigate how $P(H)$ changes under an infinitesimal transformation. We make a Taylor expansion of $P(H)$ and keep terms up to the linear order in θ:

$$\begin{aligned} P(H) &\approx P(H) + \frac{\partial P(H)}{\partial H_{11}}\Delta H_{11} + \frac{\partial P(H)}{\partial H_{12}}\Delta H_{12} + \frac{\partial P(H)}{\partial H_{22}}\Delta H_{22} && (9)\\ &= P(H) + \frac{\mathrm{d}p_{11}(H_{11})}{\mathrm{d}H_{11}} p_{12}(H_{12})p_{22}(H_{22})\Delta H_{11} + p_{11}(H_{11})\frac{\mathrm{d}p_{12}(H_{12})}{\mathrm{d}H_{12}} p_{22}(H_{22})\Delta H_{12} && (10)\\ &\quad + p_{11}(H_{11})p_{12}(H_{12})\frac{\mathrm{d}p_{22}(H_{22})}{\mathrm{d}H_{22}}\Delta H_{22} && (11)\\ &= P(H) + P(H)\theta\Bigg(-2H_{12}\frac{\mathrm{d}\ln[p_{11}(H_{11})]}{\mathrm{d}H_{11}} + (H_{11} - H_{22})\frac{\mathrm{d}\ln[p_{12}(H_{12})]}{\mathrm{d}H_{12}} + 2H_{12}\frac{\mathrm{d}\ln[p_{22}(H_{22})]}{\mathrm{d}H_{22}}\Bigg). && (12) \end{aligned}$$

Observe that for the $P(H)$ to remain invariant for all θ (for all rotations) and for all matrices H the term in parenthesis should vanish,

$$\begin{aligned} -2H_{12}\frac{\mathrm{d}\ln[p_{11}(H_{11})]}{\mathrm{d}H_{11}} + (H_{11} - H_{22})\frac{\mathrm{d}\ln[p_{12}(H_{12})]}{\mathrm{d}H_{12}} + 2H_{12}\frac{\mathrm{d}\ln[p_{22}(H_{22})]}{\mathrm{d}H_{22}} &= 0 && (13)\\ \frac{2}{H_{11} - H_{22}}\left(-\frac{\mathrm{d}\ln[p_{11}(H_{11})]}{\mathrm{d}H_{11}} + \frac{\mathrm{d}\ln[p_{22}(H_{22})]}{\mathrm{d}H_{22}}\right) &= -\frac{1}{H_{12}}\frac{\mathrm{d}\ln[p_{12}(H_{12})]}{\mathrm{d}H_{12}}, && (14) \end{aligned}$$

since the left-hand side is independent of H_{12}, we can equate it to a constant:

$$\begin{aligned} \frac{2}{H_{11} - H_{22}}\left(-\frac{\mathrm{d}\ln[p_{11}(H_{11})]}{\mathrm{d}H_{11}} + \frac{\mathrm{d}\ln[p_{22}(H_{22})]}{\mathrm{d}H_{22}}\right) &= A && (15)\\ \frac{\mathrm{d}\ln p_{11}}{\mathrm{d}H_{11}} + \frac{A}{2}H_{11} &= \frac{\mathrm{d}\ln p_{22}}{\mathrm{d}H_{22}} + \frac{A}{2}H_{22} && (16) \end{aligned}$$

both sides now can be equated to another constant B, and arrive at the set of differential equations:

$$\begin{aligned} -\frac{1}{H_{12}}\frac{\mathrm{d}\ln[p_{12}(H_{12})]}{\mathrm{d}H_{12}} &= A && (17)\\ \frac{\mathrm{d}\ln p_{11}}{\mathrm{d}H_{11}} + \frac{A}{2}H_{11} &= B && (18)\\ \frac{\mathrm{d}\ln p_{22}}{\mathrm{d}H_{22}} + \frac{A}{2}H_{22} &= B. && (19) \end{aligned}$$

Solving this equations

$$\begin{aligned} p_{11}(H_{11}) &= C_{11}e^{-AH_{11}^2/4 + BH_{11}} && (20)\\ p_{12}(H_{12}) &= C_{12}e^{AH_{12}^2/2} && (21)\\ p_{22}(H_{22}) &= C_{22}e^{-AH_{22}^2/4 + BH_{22}}. && (22) \end{aligned}$$

Finally,

$$P(H) = Ce^{-A(H_{11}^2 + H_{22}^2 + 2H_{12}^2)/4 + B(H_{11} + H_{22})} \tag{23}$$

Without loss of generality we can impose $B = 0$. In consequence, we have arrived at our goal: the probability distribution for the diagonal matrix elements become a Gaussian distribution with mean zero and variance $2/A$, and the probability distribution for the off-diagonal matrix elemenets become also a Gaussian distribution with mean zero as well, but variance $1/A$. Without loss of generality we can take $A = 2$ and recover the definition made in Eq. 1.

3 Let's derive the Wigner surmise for the GOE

3.1 Joint PDF of eigenvalues for a 2×2 matrix sampled from the GOE

The eigenvalues of H in Eq. (3) are given by:

$$\lambda_{1,2} = \frac{1}{2}\left(H_{11} + H_{22} \pm \sqrt{(H_{11} - H_{22})^2 + 4H_{12}^2}\right)$$

Explicación de la diagonalización

$$\begin{aligned} H_{11} &= \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta, \\ H_{22} &= \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta, \\ H_{12} &= (\lambda_1 - \lambda_2)\cos\theta \sin\theta. \end{aligned}$$

$$\begin{aligned} P(H)\,\mathrm{d}H &= P(H_{11}, H_{12}, H_{22})\,\mathrm{d}H_{11}\,\mathrm{d}H_{12}\,\mathrm{d}H_{22} \\ &= P(E_1, E_2, \theta)\left|\frac{\partial(H_{11}, H_{22}, H_{12})}{\partial(E_1, E_2, \theta)}\right|\mathrm{d}E_1\,\mathrm{d}E_2\,\mathrm{d}\theta \\ &= e^{-(E_1^2+E_2^2)/2}|E_1 - E_2|\,\mathrm{d}E_1\,\mathrm{d}E_2\,\mathrm{d}\theta \end{aligned}$$

$$P(E_1, E_2) = 2\pi\, e^{-(E_1^2+E_2^2)/2}|E_1 - E_2|$$

$s = E_1 - E_2$ and $r = (E_1 + E_2)/2$,

$$P(E_1, E_2)\,\mathrm{d}E_1\,\mathrm{d}E_2 = 2\pi\, e^{-(r^2+s^2/4)}|s|\left|\frac{\partial(E_1, E_2)}{\partial(r, s)}\right|\mathrm{d}r\,\mathrm{d}s$$

$$P(s) = \int_{-\infty}^{\infty} 2\pi\, e^{-(r^2+s^2/4)}|s|\left|\frac{\partial(E_1, E_2)}{\partial(r, s)}\right|\mathrm{d}r \tag{24}$$

To find the joint PDF of the eigenvalues, we perform a change of variables:

$$(H_{11}, H_{22}, H_{12}) \to (\lambda_1, \lambda_2, \theta)$$

where θ is an angle parameterizing the eigenvectors. A standard choice is to define new variables:

$$\begin{aligned} H_{11} &= x + u \\ H_{22} &= x - u \\ H_{12} &= v \end{aligned}$$

The eigenvalues then simplify to:

$$\lambda_{1,2} = x \pm \sqrt{u^2 + v^2} = x \pm r$$

where we've defined the "radial" variable $r = \sqrt{u^2 + v^2}$.

The Jacobian J for this transformation $(H_{11}, H_{22}, H_{12}) \to (x, u, v)$ is a constant (its determinant is 1). The volume element becomes:

$$dH_{11}dH_{22}dH_{12} = dxdudv$$

Now, we change to "polar" coordinates in the (u, v) plane: $u = r\cos\theta$, $v = r\sin\theta$. The volume element becomes:

$$dxdudv = dx(rdrd\theta)$$

We consider the change of variables

$$H_{11} = x + u$$
$$H_{22} = x - u$$
$$H_{12} = v$$

$$P(x, u, v) = \frac{1}{\pi^{3/2}} e^{-(x^2+u^2+v^2)} \tag{25}$$

jacobiano aquí Now we consider $r^2 = u^2 + v^2$, and $u = r\cos\theta$ and $v = r\sin\theta$

$$P(x, r, \theta)\,\mathrm{d}x\,\mathrm{d}r\,\mathrm{d}\theta = \frac{1}{\pi^{3/2}} e^{-(x^2+r^2)} r\,\mathrm{d}x\,\mathrm{d}r\,\mathrm{d}\theta \tag{26}$$

We observe that $r = s'/2$, then

$$P(x, s, \theta)\,\mathrm{d}x\,\mathrm{d}s\,\mathrm{d}\theta = \frac{1}{4\,\pi^{3/2}} e^{-(x^2+s'^2/4)} s'\,\mathrm{d}x\,\mathrm{d}s'\,\mathrm{d}\theta\,. \tag{27}$$

Finally, to obtain the *raw* level spacing distribution

$$P_{\mathrm{raw}}(s') = \frac{1}{4\,\pi^{3/2}} \int_{-\infty}^{\infty} \int_0^{2\pi} s'\, e^{-(x^2+s'^2/4)}\,\mathrm{d}\theta\,\mathrm{d}x \tag{28}$$
$$= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} s'\, e^{-(x^2+s'^2/4)}\,\mathrm{d}x \tag{29}$$
$$= \frac{s'}{2} e^{-s'^2/4} \tag{30}$$

To enforce $\langle s\rangle = 1$, we perform a change of scale. Let $s = s'/\langle s'\rangle = s'/\sqrt{\pi}$. This new variable s has mean 1. The distribution for s is found by requiring $P(s)ds = P_{\mathrm{raw}}(s)ds'$, so $P(s) = P_{\mathrm{raw}}(s')ds'/ds = \sqrt{\pi}P_{\mathrm{raw}}(s')$. Now, we rescale to unit mean spacing **aclarar** Compute the mean raw spacing:

$$\langle s'\rangle = \int_0^{\infty} s'\, P_{\mathrm{raw}}(s')\, ds' \tag{31}$$
$$= \frac{1}{2} \int_0^{\infty} s'^2 e^{-s'^2/(4)}\, ds' \tag{32}$$
$$= \sqrt{\pi}. \tag{33}$$

The change of variables gives $P(s) = \langle s'\rangle\, P_{\mathrm{raw}}(s' = \langle s'\rangle s)$, thus obtaining the *Wigner surmise for GOE ($\beta = 1$)* with unit mean spacing is

$$\boxed{P_{\mathrm{Wigner},\beta=1}(s) = \frac{\pi}{2}\, s\, \exp\Big(-\frac{\pi}{4}s^2\Big)\,, \qquad s \geq 0.} \tag{34}$$

The full level spacing distribution $P(s)$ for an $N \times N$ matrix in the large-N limit is complicated. Wigner's brilliant insight was that the essential local statistics, like the repulsion between two adjacent levels, could be captured by studying the simplest non-trivial case: a 2×2 random matrix. The surmise (an educated guess) is that the result for the 2×2 case will be very close to the result for large matrices. This turns out to be remarkably true.

This derivation showed us: 1. **The Origin of Repulsion**: The term $|\lambda_1 - \lambda_2|$ comes directly from the Jacobian of the transformation to eigenvalue space. It is a geometric consequence of the symmetry of the matrix ensemble. 2. **The Form**: The distribution is linear in s for small s ($P(s) \to 0$ as $s \to 0$), showing strong linear repulsion, and has a Gaussian tail for large s. 3. **Universality**: While derived for a 2×2 matrix, this $P(s)$ is an astonishingly good approximation for the spacing distribution of large GOE matrices and, most importantly, for the spectra of real physical systems whose classical counterparts are chaotic (e.g., stadium billiards, heavy nuclei). This is the profound connection between quantum chaos and random matrix theory.

This result for the GOE is often called the **Wigner-Dyson distribution**. The corresponding results for the Gaussian Unitary Ensemble (GUE) and Gaussian Symplectic Ensemble (GSE) are $P(s) \propto s^2 e^{-s^2}$ and $P(s) \propto s^4 e^{-s^2}$, respectively, reflecting their different symmetries and degree of level repulsion.

5. Statement for GUE ($\beta = 2$)

For GUE (complex Hermitian, Dyson index $\beta = 2$) the joint eigenvalue density has factor $|\lambda_2 - \lambda_1|^\beta$ with $\beta = 2$. The normalized unit-mean spacing distribution is

$$\boxed{p_{\text{Wigner},\beta=2}(s) = \frac{32}{\pi^2}\, s^2\, \exp\left(-\frac{4}{\pi} s^2\right).} \tag{35}$$

Here one sees the small-s repulsion $p(s) \propto s^\beta$ and a Gaussian tail.

6. Remarks

- The Wigner surmise is exact for 2×2 random matrices and gives an excellent approximation to the nearest-neighbor spacing distribution of large random matrices (GOE/GUE) after unfolding the spectrum (rescaling by the local mean level density).
- The small-s behavior $p(s) \sim s^\beta$ encodes level repulsion: eigenvalues avoid crossing. β depends on symmetry class ($\beta = 1$ real time-reversal, $\beta = 2$ complex unitary, $\beta = 4$ symplectic).
- To apply to a physical spectrum one must first *unfold* it (divide raw spacings by the local mean spacing $\rho(E)^{-1}$) so that the mean spacing is 1.