# 3 From the JPDF to the spectral density

Last time we derived the JPDF of eigenvalues for the GOE ensemble, which was

$$\rho(x_1, \ldots, x_N) = \frac{1}{\mathcal{Z}_N} e^{-\frac{1}{2}\sum_{i=1}^N x_i^2 + \frac{1}{2}\sum_{i\neq j}^N \log|x_i - x_j|} \tag{1}$$

Today we are going to massage this JPDF to work out the spectral density, which is the probability distribution of any single eigenvalue and can be defined by

$$\rho(x) = \int dx_2 \cdots dx_N \, \rho(x, x_2, \ldots, x_N) \tag{2}$$

An alternative way to define the spectral density is the average of the density function

$$n(x) = \frac{1}{N}\sum_{i=1}^N \delta(x - x_i) \tag{3}$$

This gives

$$\begin{aligned} \overline{n(x)} &= \int dx_1 \cdots dx_N \, \rho(x_1, \ldots, x_N) n(x) \\ &= \frac{1}{N}\sum_{i=1}^N \int dx_1 \cdots dx_N \, \rho(x_1, \ldots, x_N)\delta(x - x_i) \\ &= \rho(x) \end{aligned} \tag{4}$$

Before we get started in calculating the spectral density, a note. We started (for notational convenience) using GOE matrices filled with Gaussian random variables with unit variance. However, this will result in a spectral density whose width increases with $N$. In order to arrive at an answer that is well-defined in the large-$N$ limit, we should instead scale the variance of the random variables filling our matrices with $N^{-1}$. Since the Jacobian is unaffected by changing the variance, we simply have

$$\rho(x_1, \ldots, x_N) = \frac{1}{\mathcal{Z}_N} e^{-\frac{1}{2}N\sum_{i=1}^N x_i^2 + \frac{1}{2}\sum_{i\neq j}^N \log|x_i - x_j|} \tag{5}$$

where now the partition function $\mathcal{Z}_N$ has a different value, and is given by

$$\mathcal{Z}_N = \int dx_1 \ldots dx_N \, e^{-\frac{1}{2}N\sum_{i=1}^N x_i^2 + \frac{1}{2}\sum_{i\neq j}^N \log|x_i - x_j|} = \int dx_1 \ldots dx_N \, e^{-N^2\mathcal{V}(\boldsymbol{x})} \tag{6}$$

where the potential $\mathcal{V}$ is defined by

$$\mathcal{V}(\boldsymbol{x}) = \frac{1}{2}\frac{1}{N}\sum_{i=1}^N x_i^2 - \frac{1}{2}\frac{1}{N^2}\sum_{i\neq j}^N \log|x_i - x_j| \tag{7}$$

We're now clearly in the realm of a statistical mechanics problem. The potential is defined such that if the eigenvalues take values of order one, then also $\mathcal{V}$ will be of

order one, because of the factors of $N$ inside. The 'temperature' of this stat mech problem is $T = N^{-2}$, so that the large-$N$ limit is also a simultaneous zero-temperature limit. We can find the configuration of the eigenvalues by minimizing the free energy

$$F = -\log \mathcal{Z}_N \tag{8}$$

How do we compute $\mathcal{Z}_N$ in the large-$N$ limit? We will employ a common strategy, which is to replace the individual 'particles' or eigenvalues with a density, or precisely the $n(x)$ defined earlier. We can make this replacement by multiplying the integrand of $\mathcal{Z}_N$ by

$$1 = \int \mathcal{D}n\, \delta\left(n(x) - \frac{1}{N}\sum_{i=1}^{N} \delta(x - x_i)\right) \tag{9}$$

where the functional integration is only over positive, normalized functions $n$. We can then write $\mathcal{V}(\boldsymbol{x})$ as a functional of $n$ alone, since

$$\sum_{i=1}^{N} f(x_i) = N \int dx\, n(x) f(x) \tag{10}$$

$$\sum_{i,j=1}^{N} g(x_i, x_j) = N^2 \int dx\, dx\, n(x) n(x') g(x, x') \tag{11}$$

This gives

$$\mathcal{V}(\boldsymbol{x}) \mapsto \mathcal{V}[n] = \frac{1}{2}\int dx\, n(x) x^2 - \frac{1}{2}\int dx\, dx'\, n(x) n(x') \log|x - x'| \tag{12}$$

We can ignore the divergent 'self-energy' due to $x = x'$ in the logarithm because it is a mistake of less importance in $N$. Then we have

$$\mathcal{Z} = \int \mathcal{D}n\, e^{-N^2 \mathcal{V}[n]} \int dx_1 \ldots dx_N\, \delta\left(n(x) - \frac{1}{N}\sum_{i=1}^{N} \delta(x - x_i)\right) \tag{13}$$

This form is one that appears many times in statistical physics and will appear more times in this course. The first factor is the weight due to the energy of the problem, and depends only on the order parameter $n$. The second factor consists only of integrating the definition of $n$ over all of the original variables $\boldsymbol{x}$. Literally, it is the volume of configurations consistent with a given order parameter value $\boldsymbol{x}$. This gives the entropy.

In this problem, we note that the solution we seek at large $N$ is also a 'zero-temperature' limit. Therefore, the entropy won't matter in our final calculation, since at zero temperature the lowest-energy configuration dominates. But, it is still instructive to compute the entropy, since the way we do it will arise more times. The volume is

$$I[n] = \int dx_1 \ldots dx_N\, \delta\left(n(x) - \frac{1}{N}\sum_{i=1}^{N} \delta(x - x_i)\right) \tag{14}$$

First, we represent the function $\delta$ using a functional Fourier representation, or

$$I[n] = \int \mathcal{D}\hat{n} \int dx_1 \ \ldots \ dx_N \ \exp\left[N \int dx\, \hat{n}(x) \left(n(x) - \frac{1}{N}\sum_{i=1}^{N} \delta(x - x_i)\right)\right] \quad (15)$$

Next, we represent the exponential of the sum as the product of exponentials, which factorizes the integrals over the eigenvalues:

$$\begin{aligned} I[n] &= \int \mathcal{D}\hat{n}\, e^{N\int dx\, \hat{n}(x) n(x)} \prod_{i=1}^{N} \int dx_i\, e^{-\int dx\, \hat{n}(x)\, \delta(x - x_i)} \quad (16)\\ &= \int \mathcal{D}\hat{n}\, e^{N\int dx\, \hat{n}(x) n(x)} \prod_{i=1}^{N} \int dx_i\, e^{-\hat{n}(x_i)} \\ &= \int \mathcal{D}\hat{n}\, e^{N\int dx\, \hat{n}(x) n(x)} \left(\int dx\, e^{-\hat{n}(x)}\right)^N \\ &= \int \mathcal{D}\hat{n}\, e^{N\int dx\, \hat{n}(x) n(x) + N \log \int dx\, e^{-\hat{n}(x)}} \\ &= \int \mathcal{D}\hat{n}\, e^{N\mathcal{S}[\hat{n}]} \end{aligned}$$

where we have defined an effective action $\mathcal{S}$ by

$$\mathcal{S}[\hat{n}] = \int dx\, n(x)\hat{n}(x) + \log \int dx\, e^{-\hat{n}(x)} \quad (17)$$

Because the volume is an integral of the form of an exponential varying with a diverging quantity $N$, we can evaluate it using the saddle-point method. A saddle point of the volume is given by

$$0 = \frac{\delta \mathcal{S}[\hat{n}]}{\delta \hat{n}(x)} = n(x) - \frac{e^{-\hat{n}(x)}}{\int dx'\, e^{-\hat{n}(x')}} \quad (18)$$

Trying to solve, we find

$$e^{-\hat{n}(x)} = n(x) \int dx'\, e^{-\hat{n}(x')} \quad (19)$$

or

$$\hat{n}(x) = -\log\left[n(x) \int dx'\, e^{-\hat{n}(x')}\right] = -\log n(x) - \log \int dx'\, e^{-\hat{n}(x')} \quad (20)$$

In fact this is solved by $\hat{n}(x) = -\log n(x)$ ignoring the last term. Why? Because

$$\log \int dx'\, e^{-\hat{n}(x')} = \log \int dx'\, e^{\log n(x')} = \log \int dx'\, n(x') = \log 1 = 0 \quad (21)$$

due to the normalization of $n$. Therefore we have, to leading exponential order in $N$, the entropy

$$\log I[n] = N\mathcal{S}[\hat{n}] = -N \int dx\, n(x) \log n(x) + o(N) \quad (22)$$

This is a very general result for the entropy of a probability distribution – this is the Shannon entropy of the distribution. Now we see why it won't matter for the spectral density: the entropic factor $I$ grows like $e^N$, while the energetic one grows like $e^{N^2}$.

So we have

$$\mathcal{Z} = \int \mathcal{D}n\, e^{-N^2\mathcal{V}[n] - N\int dx\, n(x)\log n(x) + o(N)} = \int \mathcal{D}n\, e^{-N^2\mathcal{V}[n] + o(N^2)} \tag{23}$$

Given what we just did, you can guess where this is going: we can determine the form for the spectral density by evaluating the integral in $n$ with the saddle-point method in the large $N$ limit. Then, the value of the function $n$ at the saddle will be the spectral density! To do this, we have to make one change. When evaluating the integral by a variational means, we have to respect the constraints of the integrand. In this case, $n$ is positive and normalized. The positivity won't end up being an issue, but the normalization is. In order to make the normalization explicit, we introduce it as an explicit $\delta$ function,

$$\delta\left(N^2\left[\int dx\, n(x) - 1\right]\right) = \int d\kappa\, e^{N^2\kappa\left[\int dx\, n(x) - 1\right]} \tag{24}$$

Then we have

$$\mathcal{Z} = \int \mathcal{D}n\, d\kappa\, e^{-N^2\left(\mathcal{V}[n] - \kappa\left[\int dx\, n(x) - 1\right]\right) + o(N^2)} = \int \mathcal{D}n\, d\kappa\, e^{-N^2\mathcal{S}[n,\kappa] + o(N^2)} \tag{25}$$

where we have defined the new effective action

$$\mathcal{S}[n,\kappa] = \mathcal{V}[n] - \kappa\left[\int dx\, n(x) - 1\right] \tag{26}$$

We can now evaluate the partition function using the saddle-point method. The extremal conditions are

$$0 = \frac{\delta\mathcal{S}}{\delta n(x)} = \frac{x^2}{2} - \int dx'\, n(x')\log|x - x'| - \kappa \tag{27}$$

$$0 = \frac{\partial\mathcal{S}}{\partial\kappa} = 1 - \int dx\, n(x) \tag{28}$$

We need to solve the first of these integral equations for $n$ as a function of $\kappa$, then choose $\kappa$ so that the result is normalized. In this sense $\kappa$ acts as a Lagrange multiplier. The free energy is then

$$F = N^2\mathcal{S}[n,\kappa] \tag{29}$$

and the extremal problem is equivalent to minimizing the free energy.

Differentiating the equation gives

$$0 = x - \int dx'\, \frac{n(x')}{x - x'} \tag{30}$$

We expect this equation to be solved for $n$ with finite support, and in fact a single *compact* support $[a, b]$. There is a theorem of Tricomi that says equations of the form

$$\int_a^b dx' \, \frac{f(x')}{x - x'} = g(x) \tag{31}$$

can be solved by

$$f(x) = \frac{C - \int_a^b \frac{dt}{\pi} \frac{\sqrt{(t-a)(b-t)}}{x-t} g(t)}{\pi\sqrt{(x-a)(b-x)}} \tag{32}$$

for some constant $C$. Using $g(t) = t$ and fixing $C$ by the normalization gives

$$n(x) = \frac{1}{\pi\sqrt{(x-a)(b-x)}} \left[1 - x^2 + \frac{1}{2}(a+b)x + \frac{1}{8}(b-a)^2\right] \tag{33}$$

Since we don't know the bounds $a$ and $b$, how do we find them? Well, we know that the free energy should be minimized. Inserting this solution $n$ into the free energy, we have a new expression that is a function of $a$ and $b$, and we can minimize this function. After many nasty integrals and also a nasty extremal problem, one finds the non-nasty answer $a = -\sqrt{2}$ and $b = \sqrt{2}$. Inserting these into $n$ gives

$$n(x) = \frac{1}{\pi\sqrt{(x-\sqrt{2})(\sqrt{2}-x)}} \left[1 - x^2 + \frac{1}{8}(2\sqrt{2})^2\right] \tag{34}$$

$$= \frac{1}{\pi} \frac{2 - x^2}{\sqrt{2 - x^2}} = \frac{1}{\pi}\sqrt{2 - x^2} \tag{35}$$

which is the Wigner semicircle.