5 The resolvent several other ways

Recall the resolvent

$$G(z) = \frac{1}{N}\overline{\mathrm{Tr}(zI - H)^{-1}} \tag{1}$$

and its relation to the spectral density:

$$\rho(x) = \frac{1}{\pi}\lim_{\epsilon\to 0}\mathrm{Im}\,G(x - i\epsilon) \tag{2}$$

The resolvent is a powerful tool to find the spectral density, partially because it can be calculated so many different ways. We saw last week a field theory approach to its calculation. Today we will see some other ways to compute it.

The cavity method

The cavity method is an extremely general method for solving disordered systems problems at large size. Qualitatively, it is simple: you take a large system and examine one component of it, describing the behavior of the whole system as the behavior of the isolated piece, the rest of the system, and the interaction between the two. Reasoning that at large N the behavior of the rest of the system should not be changed by removing one piece, the average behavior of the isolated part can be self-consistently described.

Let's see how this woks for the resolvent. We consider a large GOE matrix H and consider the first row and column in isolation:

$$H = \begin{bmatrix} H_{11} & \boldsymbol{H}_1^T \\ \boldsymbol{H}_1 & H_r \end{bmatrix} \tag{3}$$

where H_r is the $N-1\times N-1$ rest of the matrix. Block matrices can be blockwise inverted like

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \tilde{M}_{11} & \tilde{M}_{12} \\ \tilde{M}_{21} & \tilde{M}_{22} \end{bmatrix} \tag{4}$$

where the $\tilde{M}_{ij}$ are functions of the M_{ij}. We only need the relation for the upper left block for our calculation, which is

$$\tilde{M}_{11}^{-1} = M_{11} - M_{12}M_{22}^{-1}M_{21} \tag{5}$$

which is the Schur complement formula. Writing this relationship for the upper left 1×1 block of the resolvent, we have

$$\frac{1}{[(zI-H)^{-1}]_{11}} = (z - H_{11}) - \boldsymbol{H}_1^T(zI - H_r)^{-1}\boldsymbol{H}_1 \tag{6}$$

Now, we take averages. For the first term we have

$$\overline{\frac{1}{[(zI-H)^{-1}]_{11}}} = \frac{1}{\overline{[(zI-H)^{-1}]_{11}}} = \frac{1}{G(z)} \tag{7}$$

where here we relied on kind of approximation that happens to be valid here: the resolvent is self-averaging and concentrates rapidly on its average, so that making the average of the inverse is the same as the inverse of the average. We also have used the fact that the matrix ensemble is invariant under permutation of the rows and columns, so that

$$\overline{(zI-H)^{-1}_{11}} = \overline{(zI-H)^{-1}_{ii}} \tag{8}$$

for any other index i. Then the next term gives

$$\overline{z - H_{11}} = z - \overline{H_{11}} = z \tag{9}$$

because the matrix element H_{11} is a centered Gaussian. Finally, the last term is

$$\begin{aligned}\overline{\boldsymbol{H}_1^T (zI-H_r)^{-1} \boldsymbol{H}_1} &= \overline{\sum_{ij} H_{1i} (zI-H_r)^{-1}_{ij} H_{1j}} = \sum_{ij} \overline{(zI-H_r)^{-1}_{ij}}\,\overline{H_{1i}H_{1j}} \\ &= \sum_{ij} \overline{(zI-H_r)^{-1}_{ij}} \frac{1}{2N}(\delta_{ij} + \delta_{1i}\delta_{1j}) \\ &= \frac{1}{2N} \sum_i \overline{(zI-H_r)^{-1}_{ii}} + \frac{1}{2N} \overline{(zI-H_r)^{-1}_{11}} = \frac{1}{2} G(z)\end{aligned} \tag{10}$$

where we have used the fact that the first row and column are independent of the rest of the matrix, and the fact that the resolvent of the $N-1 \times N-1$ submatrix is the same as that for the whole matrix when N is large. Putting these pieces together, we have

$$\frac{1}{G(z)} = z - \frac{1}{2} G(z) \tag{11}$$

which gives $G(z) = z - \sqrt{z^2 - 2}$ and $\rho(x) = \frac{1}{\pi}\sqrt{2 - x^2}$, as we saw in the last lecture.

The cavity method is often the simplest method for deriving a given spectral density, but it relies on assumptions that are sometimes not true. It is often not clear when those assumptions fail.

The replica method

The replica method is extremely versatile, and in practice is one of the easiest ways to derive spectral densities. It relies on an integral identity for a symmetric matrix A, or

$$A^{-1}_{ij} = \sqrt{\frac{\det A}{(2\pi)^N}} \int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}} s_i s_j \tag{12}$$

This identity can be derived by differentiating a Gaussian integral by A: on one hand,

$$\frac{\partial}{\partial A_{ij}} \int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}} = -\frac{1}{2} \int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}} s_i s_j \tag{13}$$

whereas on the other hand,

$$\frac{\partial}{\partial A_{ij}} \int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}} = \frac{\partial}{\partial A_{ij}} \sqrt{\frac{(2\pi)^N}{\det A}} = -\frac{1}{2}\sqrt{\frac{(2\pi)^N}{(\det A)^3}} \frac{\partial}{\partial A_{ij}} \det A \tag{14}$$
$$= -\frac{1}{2}\sqrt{\frac{(2\pi)^N}{(\det A)^3}} \det A A^{-1}_{ij} = -\frac{1}{2}\sqrt{\frac{(2\pi)^N}{(\det A)}} A^{-1}_{ij}$$

Equating (13) and (14) and rearranging gives (12). The first factor is annoying, but we can cancel it with another integral:

$$\int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}} = \sqrt{\frac{(2\pi)^N}{\det A}} \tag{15}$$

so that

$$A^{-1}_{ij} = \frac{\int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}} s_i s_j}{\int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T A \boldsymbol{s}}} \tag{16}$$

We can therefore write

$$G(z) = \frac{1}{N}\overline{\mathrm{Tr}(zI-H)^{-1}} = \frac{1}{N}\overline{\frac{\int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T(zI-H)\boldsymbol{s}} \|\boldsymbol{s}\|^2}{\int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T(zI-H)\boldsymbol{s}}}} \tag{17}$$

where the denominator is necessary to cancel the leading factor of πs and $\det A$. Now, taking the average is a problem: we can average the numerator because it it is the linear exponential of a Gaussian random variable, but the denominator poses a challenge. We treat it with replicas, and the trivial identity

$$\lim_{n\to 0} x^{n-1} = \frac{1}{x} \tag{18}$$

If we call the integration variable in the numerator $\boldsymbol{s}_1$, we can write

$$G(z) = \frac{1}{N}\lim_{n\to 0}\overline{\int d\boldsymbol{s}_1\, e^{-\frac{1}{2}\boldsymbol{s}_1^T(zI-H)\boldsymbol{s}_1}\|\boldsymbol{s}_1\|^2 \left(\int d\boldsymbol{s}\, e^{-\frac{1}{2}\boldsymbol{s}^T(zI-H)\boldsymbol{s}}\right)^{n-1}} \tag{19}$$
$$= \frac{1}{N}\lim_{n\to 0}\overline{\int \left(\prod_{a=1}^{n} d\boldsymbol{s}_a\right) \|\boldsymbol{s}_1\|^2 e^{-\frac{1}{2}\sum_{a=1}^{n}\boldsymbol{s}_a^T(zI-H)\boldsymbol{s}_a}}$$

where the $n-1$ replicas from the denominator have integration variables $\boldsymbol{s}_2$ through $\boldsymbol{s}_n$.

The advantage of this transformation is that it reduces dependence on H to linear in an exponential. For Gaussian H we know how to average this, since for Gaussians we have

$$\overline{e^{\sum_{ij} A_{ij} H_{ij}}} = e^{\frac{1}{2}\sum_{ijkl} A_{ij}\overline{H_{ij}H_{kl}}A_{kl}} \tag{20}$$

For us this gives

$$\overline{e^{-\frac{1}{2}\sum_{a=1}^{n}\sum_{ij=1}^{N} s_{ai}s_{aj}H_{ij}}} = e^{\frac{1}{8}\sum_{a=1}^{n}\sum_{ij=1}^{N} s_{ai}s_{aj}\sum_{b=1}^{n}\sum_{kl=1}^{N} s_{bk}s_{bl}\overline{H_{ij}H_{kl}}}$$
$$= e^{\frac{1}{8}\sum_{a=1}^{n}\sum_{ij=1}^{N} s_{ai}s_{aj}\sum_{b=1}^{n}\sum_{kl=1}^{N} s_{bk}s_{bl}\frac{1}{2N}(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})}$$
$$= e^{\frac{N}{8}\sum_{a,b=1}^{n}(\frac{\boldsymbol{s}_a\cdot\boldsymbol{s}_b}{N})^2} \tag{21}$$

and therefore

$$G(z) = \lim_{n\to 0}\int\left(\prod_{a=1}^{n} d\boldsymbol{s}_a\right)\frac{\boldsymbol{s}_1\cdot\boldsymbol{s}_1}{N}e^{-\frac{N}{2}z\sum_{a=1}^{n}\frac{\boldsymbol{s}_a\cdot\boldsymbol{s}_a}{N}+\frac{N}{8}\sum_{a,b=1}^{n}(\frac{\boldsymbol{s}_a\cdot\boldsymbol{s}_b}{N})^2} \tag{22}$$

Averaging over H has coupled the previously independent replicas together. Now, we notice that the resolvent depends only on the scalar products $\boldsymbol{s}_a\cdot\boldsymbol{s}_b$. We therefore look to replace them with an order parameter

$$Q_{ab} = \frac{\boldsymbol{s}_a\cdot\boldsymbol{s}_b}{N} \tag{23}$$

which is called the overlap matrix. We do this by inserting a δ-function

$$\int\prod_{ab} dQ_{ab}\,\delta\big(\tfrac{1}{2}(NQ_{ab}-\boldsymbol{s}_a\cdot\boldsymbol{s}_b)\big) = \int dQ\,d\tilde{Q}\,e^{\sum_{ab}\frac{1}{2}\tilde{Q}_{ab}(NQ_{ab}-\boldsymbol{s}_a\cdot\boldsymbol{s}_b)} \tag{24}$$

which gives us

$$G(z) = \lim_{n\to 0}\int dQ\,d\tilde{Q}\,Q_{11}e^{-\frac{N}{2}z\sum_{a=1}^{n}Q_{aa}+\frac{N}{8}\sum_{a,b=1}^{n}Q_{ab}^2+\frac{N}{2}\sum_{ab}\tilde{Q}_{ab}Q_{ab}}\int\left(\prod_{a=1}^{n}d\boldsymbol{s}_a\right)e^{-\frac{1}{2}\sum_{ab}\tilde{Q}_{ab}\boldsymbol{s}_a\cdot\boldsymbol{s}_b}$$
$$= \lim_{n\to 0}\int dQ\,d\tilde{Q}\,Q_{11}e^{-\frac{N}{2}z\sum_{a=1}^{n}Q_{aa}+\frac{N}{8}\sum_{a,b=1}^{n}Q_{ab}^2+\frac{N}{2}\sum_{ab}\tilde{Q}_{ab}Q_{ab}}\left(\frac{(2\pi)^n}{\det\tilde{Q}}\right)^{\frac{N}{2}}$$
$$= \lim_{n\to 0}(2\pi)^{\frac{nN}{2}}\int dQ\,d\tilde{Q}\,Q_{11}e^{-\frac{N}{2}z\sum_{a=1}^{n}Q_{aa}+\frac{N}{8}\sum_{a,b=1}^{n}Q_{ab}^2+\frac{N}{2}\sum_{ab}\tilde{Q}_{ab}Q_{ab}-\frac{N}{2}\log\det\tilde{Q}}$$

New we can see clearly where this is going: the integral now only involves the order parameter matrices Q and $\tilde{Q}$ with dependence on an exponential raised to the power N. Therefore, we can used the saddle-point method. First we consider the integral in $\tilde{Q}$. The extremal condition is

$$\frac{\partial\mathcal{S}}{\partial\tilde{Q}_{ab}} = \frac{1}{2}Q_{ab}-\frac{1}{2}\frac{1}{\det\tilde{Q}}\frac{\partial}{\partial\tilde{Q}_{ab}}\det\tilde{Q} = \frac{1}{2}Q_{ab}-\frac{1}{2}\tilde{Q}_{ab}^{-1} \tag{25}$$

which implies $\tilde{Q} = Q^{-1}$. This gives

$$G(z) = \lim_{n\to 0}\int dQ\,Q_{11}e^{-\frac{N}{2}z\sum_{a=1}^{n}Q_{aa}+\frac{N}{8}\sum_{a,b=1}^{n}Q_{ab}^2+\frac{N}{2}\log\det Q+\frac{Nn}{2}(1+\log(2\pi))} \tag{26}$$

This is clearly set up for us to use a saddle-point approximation. For the matrix Q, we have

$$0 = \frac{\partial \mathcal{S}}{\partial Q_{ab}} = -\frac{1}{2} z \delta_{ab} + \frac{1}{4} Q_{ab} + \frac{1}{2} Q^{-1}_{ab} \tag{27}$$

Multiplying both sides by $2Q$ gives the matrix equation

$$0 = \frac{1}{2} Q^2 - zQ + I \tag{28}$$

which is really n^2 equations. How can we hope to make sense of this mess?

This requires making an ansatz for the structure of Q. Experience tells us that, because our initial problem was a quadratic Hamiltonian, Q will have a property called *replica symmetry*, meaning that its structure is $Q_{ab} = q_d \delta_{ab} + q_0$. This is a matrix with one constant $q_d + q_0$ on the diagonal and another constant q_0 on all off-diagonals. Inserting this into our matrix equation, we have

$$\begin{aligned} 0 &= \frac{1}{2} \sum_c Q_{ac} Q_{cb} - z Q_{ab} + \delta_{ab} \\ &= \frac{1}{2} \sum_c (q_d \delta_{ac} + q_0)(q_d \delta_{cb} + q_0) - z(q_d \delta_{ab} + q_0) + \delta_{ab} \\ &= \frac{1}{2} (q_d^2 \delta_{ab} + 2 q_d q_0 + n q_0^2) - z(q_d \delta_{ab} + q_0) + \delta_{ab} \\ &= \left[\frac{1}{2} q_d^2 - z q_d + 1 \right] \delta_{ab} + \left[q_d q_0 + \frac{1}{2} n q_0^2 - z q_0 \right] \end{aligned} \tag{29}$$

This gives us two equations, one for the diagonal and one for the off-diagonal. The off diagonal equation is solved by $q_0 = 0$, whereas the diagonal one is a quadratic equation for q_d with solutions

$$q_d = z \pm \sqrt{z^2 - 2} \tag{30}$$

Since $Q_{ab} = q_d \delta_{ab}$ with zero off-diagonal, we have

$$\sum_{a=1}^{n} Q_{aa} = n q_d \qquad \sum_{ab=1}^{n} Q_{ab}^2 = n q_d^2 \qquad \log\det Q = n \log q_d \qquad Q_{11} = q_d \tag{31}$$

and therefore we have

$$G(z) = \lim_{n \to 0} q_d e^{Nn\left[-\frac{1}{2} z q_d + \frac{1}{8} q_d^2 + \frac{1}{2} \log q_d + \frac{1}{2}(1 + \log(2\pi))\right]} \tag{32}$$

In the limit $n \to 0$, the whole exponential factor goes to one regardless of the value of q_d. We therefore have

$$G(z) = q_d = z \pm \sqrt{z^2 - 2} \tag{33}$$

and the semicircle distribution as before.

The replica method is extremely powerful for treating mean-field cases where subleading corrections in N are not desired or necessary. It can be difficult to successfully use in situations beyond the spectral density, e.g., for two-point functions, because it can require application of replica symmetry breaking.