Lecture 25: May 27, 2011

1. Recap from Last Time

We are working in the NCPS $(\mathcal{M}_n, \varphi_n)$ where

$$\mathcal{M}_n = \bigcap_{p\ge 1} L^p(\Omega, \mathscr{F}, \mathbb{P}\,;\, M_n(\mathbb{C})) \qquad \varphi_n(X) = \frac{1}{n}\mathbb{E}\,\mathrm{Tr}\,X.$$

Here $(\Omega, \mathscr{F}, \mathbb{P})$ is some complete probability space; e.g. $[0,1]$ with Lebesgue measure.

In $(\mathcal{M}_n, \varphi_n)$, we are looking at a particular random variable $X = X_n$ called a GUE_n: *Gaussian Unitary Ensemble*. It is defined thus: let $\{G_j^0\}_{1\le j\le n}$, $\{G_{jk}^1\}_{1\le j<j\le n}$, $\{G_{jk}^2\}_{1\le j<k\le n}$ be i.i.d. $N(0,1)$ random variables, and let $\alpha = \alpha_n > 0$. Then set

$$\begin{aligned}
X_{jj} &= \alpha G_j^0, && 1 \le j \le n \\
X_{jk} &= \frac{\alpha}{\sqrt{2}}(G_{jk}^1 + iG_{jk}^2), && 1 \le j < k \le n \\
X_{kj} &= \frac{\alpha}{\sqrt{2}}(G_{jk}^1 - iG_{jk}^2), && 1 \le j < k \le n
\end{aligned}$$

That is, $X_{kj} = \overline{X_{jk}}$, so $X = X^*$ is a Hermitian matrix, all of whose upper-triangular entries are independent complex Gaussians. Last time, we calculated that $\varphi_n(X_n) = 0$ while $\varphi_n(X_n^2) = n\alpha^2$, prompting us to set $\alpha = \frac{1}{\sqrt{n}}$ so as to standardize X_n. For higher moments, we calculated the general covariances of different entries, finding that

$$\mathbb{E}(X_{ij}X_{k\ell}) = \alpha^2\delta_{i\ell}\delta_{jk}. \tag{cov}$$

This turns out to be enough to calculate all higher moments, because of the following.

2. Wick's Formula

Theorem 1 (Wick's Formula)**.** *Let $G_1, \ldots, G_k$ be independent $N(0,1)$ random variables, and let $X_1, \ldots, X_k$ be linear functions of $G_1, \ldots, G_k$. Then*

$$\mathbb{E}(X_1 \cdots X_k) = \sum_{\pi \in \mathscr{P}_2(k)} \prod_{\{i,j\}\in\pi} \mathbb{E}(X_iX_j). \tag{Wick}$$

Remarks.

(1) The precise statement of the assumption is that, with $\mathbf{X} = (X_1, \ldots, X_k)$ and $\mathbf{G} = (G_1, \ldots, G_k)$, there exists a linear map $T\colon \mathbb{C}^k \to \mathbb{C}^k$ such that $\mathbf{X} = T(\mathbf{G})$. For our purposes, it is important that the map be allowed to be complex-valued. Note that the entries X_{ij} of our matrix are indeed linear functions of i.i.d. Gaussians, by definition.

(2) In terms of *classical* cumulant functionals c_n, the moment-cumulant formula says that

$$\mathbb{E}(X_1 \cdots X_k) = \sum_{\pi\in\mathscr{P}(k)} c_\pi(X_1, \ldots, X_k).$$

Now, if $\pi \in \mathscr{P}_2(k)$, then by definition

$$c_\pi(X_1, \ldots, X_k) = \prod_{\{i,j\}\in\pi} c_2(X_i, X_j).$$

Recall that $c_2(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$; in this case, the X_i are centered, and so $c_2(X_i, X_j) = \mathbb{E}(X_i X_j)$. Thus, Wick's formula actually says

$$\mathbb{E}(X_1 \cdots X_k) = \sum_{\pi \in \mathscr{P}_2(n)} c_\pi(X_1, \ldots, X_k).$$

In other words: Wick's formula asserts that, for linear functions of independent Gaussian random variables, the only non-zero mixed cumulants are c_π with $\pi \in \mathscr{P}_2$. We already saw this for the case of a *single* Gaussian random variable:

$$c_k(X_i, \ldots, X_i) = \delta_{k2} \mathrm{Var} X_i.$$

Proof. Suppose we have proved Wick's formula for a mixture of the i.i.d. normal variables $\mathbf{G}$:

$$\mathbb{E}(G_{i_1} \cdots G_{i_k}) = \sum_{\pi \in \mathscr{P}_2(k)} \prod_{\{a,b\} \in \pi} \mathbb{E}(G_{i_a} G_{i_b}). \tag{Wick'}$$

(We do not assume that the i_a are all distinct.) Then since $\mathbf{X} = T(\mathbf{G})$ for linear T, $\mathbb{E}(X_1 \cdots X_k)$ is a linear combination of terms on the left-hand-side of (Wick'). Similarly, the right-hand-side of (Wick') is multi-linear in the entries, and so $\sum_{\pi \in \mathscr{P}_2(k)} \prod_{\{i,j\} \in \pi} \mathbb{E}(X_i X_j)$ is *the same* linear combination of terms on the right-hand-side of (Wick'). Hence, to prove formula (Wick), it suffices to prove formula (Wick').

Now, the G_i are independent, so the moment $\mathbb{E}(G_{i_1} \cdots G_{i_k})$ is easy to calculate. Let $\tau \in \mathscr{P}(k)$ be the partition determined by the indices $(i_1, \ldots, i_k)$: $a \sim_\tau b$ iff $i_a = i_b$. If $\tau = (B_1, \ldots, B_r)$, with index j_a common to B_a, then

$$\mathbb{E}(G_{i_1} \cdots G_{i_k}) = \mathbb{E}(G_{j_1}^{|B_1|} \cdots G_{j_r}^{|B_r|}) = \mathbb{E}(G_{j_1}^{|B_1|}) \cdots \mathbb{E}(G_{j_r}^{|B_r|})$$

since $j_1, \ldots, j_r$ are all distinct and $G_1, \ldots, G_k$ are independent. Now, the moments of a $N(0,1)$ variable have already been calculated in a previous lecture:

$$\mathbb{E}(X_{j_a}^\ell) = \begin{cases} (\ell-1)(\ell-3)\cdots(3)(1), & \ell \text{ even} \\ 0, & \ell \text{ odd} \end{cases} = \#\mathscr{P}_2(\ell).$$

Thus, the left-hand-side of (Wick') formula is

$$\mathbb{E}(G_{i_1} \cdots G_{i_k}) = \#\mathscr{P}_2(|B_1|) \cdots \#\mathscr{P}_2(|B_r|).$$

As for the right-hand-side, note that, because the G_i are centered and independent, for any $\pi \in \mathbb{P}_2(k)$

$$\prod_{\{a,b\} \in \pi} \mathbb{E}(G_{i_a} G_{i_b}) = 0 \text{ unless } i_a = i_b \ \forall \ \{a,b\} \in \pi.$$

I.e. the only terms contributing to the sum have $a \sim_\pi b \Rightarrow i_a = i_b \Rightarrow a \sim_\tau b$, so $\pi \leq \tau$ in the lattice $\mathscr{P}(k)$. In the case that $\pi \leq \tau$, all the terms in the product are equal to $\mathbb{E}(G_i^2) = 1$, and so we have

$$\sum_{\pi \in \mathscr{P}_2(k)} \prod_{\{a,b\} \in \pi\}} \mathbb{E}(G_{i_a} G_{i_b}) = \sum_{\substack{\pi \in \mathscr{P}_2(k) \\ \pi \leq \tau}} 1 = \#\{\pi \in \mathscr{P}_2(k) \, : \, \pi \leq \tau\}.$$

Now, if $\pi \leq \tau$, then any block of π is contained in a block of τ. Thus, we can decompose π as a product, $\pi \in \mathscr{P}_2(B_1) \times \cdots \times \mathscr{P}_2(B_r)$. Moreover, there is no "non-crossing" condition here, so any such product is a pairing that is $\leq \tau$. In other words

$$\#\{\pi \in \mathscr{P}_2(k) \,:\, \pi \leq \tau\} = \#\mathscr{P}_2(B_1) \times \cdots \times \mathscr{P}_2(B_r).$$

This is clearly equal to the value of the left-hand-side already calculated, thus proving (Wick'), and thus (Wick). □

Exercise. Wick's formula is often stated with the condition that $\mathbf{X} = (X_1, \ldots, X_n)$ is a *jointly-Gaussian* random vector. That is: there is a positive-definite $n \times n$ matrix C such that

$$\mu_{\mathbf{X}}(d\mathbf{t}) = (2\pi)^{-n/2}(\det C)^{-1/2} e^{-\frac{1}{2}\langle \mathbf{t}, C^{-1}\mathbf{t}\rangle}\, d\mathbf{t}.$$

The matrix C is the *covariance matrix*, since indeed $C_{ij} = \mathbb{E}(X_i X_j)$. Since C is assumed positive-definite, $C^{-1/2}$ exists as a positive-definite matrix. Show that the entries of the vector $\mathbf{G} = C^{-\frac{1}{2}}(\mathbf{X})$ are i.i.d. $N(0,1)$, so this is a special case of the conditions given above.

3. The Genus Expansion

Let us now proceed to calculate the (matrix) moments of our GUE_n $X = X_n$ in the NCPS $(\mathcal{M}_n, \varphi_n)$. Let $k \geq 2$. First, matrix multiplication gives

$$[X^k]_{ii} = \sum_{1 \leq i_1, \ldots, i_k \leq n} X_{ii_2} X_{i_2 i_3} \cdots X_{i_{k-1} i_k} X_{i_k i}.$$

So, taking the trace and expectation,

$$\varphi_n(X_n^k) = \frac{1}{n} \sum_{1 \leq i_1, \ldots, i_k \leq n} \mathbb{E}(X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_k i_1}).$$

The X_{ij} are linear functions of a family of independent $N(0,1)$ variables; hence, we may apply Wick's formula (Wick) to them. That is,

$$\mathbb{E}(X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_k i_1}) = \sum_{\pi \in \mathscr{P}_2(k)} \prod_{\{a,b\} \in \pi} \mathbb{E}(X_{i_a i_{a+1}} X_{i_b i_{b+1}})$$

where we agree that $i_{k+1} = i_1$. But we already calculated all such covariances, in Equation (cov); that is

$$\mathbb{E}(X_{i_a i_{a+1}} X_{i_b i_{b+1}}) = \alpha^2 \delta_{i_a i_{b+1}} \delta_{i_{a+1} i_b}.$$

There are $k/2$ such factors in each product in the sum, so factoring out the α^2 terms and switching the (independent) order of summation, we get

$$\varphi_n(X^k) = \frac{\alpha^k}{n} \sum_{\pi \in \mathscr{P}_2(k)} \sum_{1 \leq i_1, \ldots, i_k \leq n} \prod_{\{a,b\} \in \pi} \delta_{i_a i_{b+1}} \delta_{i_{a+1} i_b}.$$

To understand the internal product, it is convenient to think of $\pi \in \mathscr{P}_2(k)$ as a *permutation* of $[k]$: the blocks of π are the cycles of this partition. So $\{a,b\} \in \pi$ means $\pi(a) = b$ and $\pi(b) = a$. Then we can write

$$\prod_{\{a,b\} \in \pi} \delta_{i_a i_{b+1}} \delta_{i_b i_{a+1}} = \prod_{a=1}^{k} \delta_{i_a i_{\pi(a)+1}}.$$

(Note: not every permutation arises this way; such π satisfy $\pi^2 = 1$ and have no fixed-points.) Utilizing the permutation language, let's go one step further and use the rotation $\gamma = (1, 2, \ldots, k)$ (i.e. $\gamma(\ell) = (\ell + 1) \bmod k$) to rewrite out moment as

$$\varphi_n(X^k) = \frac{\alpha^k}{n} \sum_{\pi \in \mathscr{P}_2(k)} \sum_{1 \le i_1, \ldots, i_k \le n} \prod_{a=1}^{k} \delta_{i_a i_{\gamma\pi(a)}}.$$

Let $\mathbf{i}\colon [k] \to [n]$ be the function $\mathbf{i}(a) = i_a$; so we are summing over all $\pi \in \mathscr{P}_2(k)$ and all functions $\mathbf{i}\colon [k] \to [n]$. Fixing a π and $\mathbf{i}$, the product

$$\prod_{a=1}^{k} \delta_{\mathbf{i}(a)\mathbf{i}(\gamma\pi(a))} \neq 0 \quad \text{iff} \quad \mathbf{i} \text{ is constant on the cycles of } \gamma\pi \text{, in which case it } = 1.$$

That is to say:

$$\varphi_n(X^k) = \frac{\alpha^k}{n} \sum_{\pi \in \mathscr{P}_2(k)} \#\{\mathbf{i}\colon [k] \to [n] \,:\, \mathbf{i} \text{ is constant on the cycles of } \gamma\pi\}.$$

This number (inside the sum) is trivial to count: we choose one value (out of n) for each of the cycles in $\gamma\pi$, with no restriction on repeats; thus

$$\varphi_n(X^k) = \frac{\alpha^k}{n} \sum_{\pi \in \mathscr{P}_2(k)} n^{\#\text{ cycles in } \gamma\pi}.$$

For a permutation $\sigma \in S_k$, let $\#\sigma$ denote the number of cycles in σ. Now, remembering that $\alpha = n^{-1/2}$, we combine to find the **genus expansion**:

$$\varphi_n(X_n^k) = \sum_{\pi \in \mathscr{P}_2(k)} n^{\#(\gamma\pi) - k/2 - 1}. \tag{GE1}$$

Why do we call this a "genus" expansion? To answer that question, let's first note that $\mathscr{P}_2(k) = \varnothing$ if k is odd, so take $k = 2m$ even. Then we have

$$\varphi_n(X_n^{2m}) = \sum_{\pi \in \mathscr{P}_2(2m)} n^{\#(\gamma\pi) - m - 1}.$$

Now, let us explore the exponent of n in the terms in this formula. There is a beautiful geometric way to visualize $\#(\gamma\pi)$. Draw a regular $2m$-gon, and label its vertices in cyclic order $v_1, v_2, \ldots, v_{2m}$. Its edges may be identifies as (cyclically)-adjacent pairs of vertices: $e_1 = v_1v_2$, $e_2 = v_2v_3$, until $e_{2m} = v_{2m}v_1$. A pairing $\pi \in \mathscr{P}_2(2m)$ can now be used to glue the edges of the $2m$-gon together to form a compact surface. Note: by convention, we always identify edges in "tail-to-head" orientation. For example, if $\pi(1) = 3$, we identify e_1 with e_3 by gluing v_1 to v_4 (and ergo v_2 to v_3). This convention forces the resultant compact surface to be orientable.

Lemma 2. *Let S_π be the compact surface obtained by gluing the edges of a $2m$-gon according to the pairing $\pi \in \mathscr{P}_2(2m)$ as described above; let G_π be the image of the $2m$-gon in S_π. Then the number of distinct vertices in G_π is equal to $\#(\gamma\pi)$.*

Proof. Since e_i is glued to $e_{\pi(i)}$, by the "tail-to-head" rule this means that v_i is identified with $v_{\pi(i)+1}$ (with addition modulo $2m$); that is, v_i is identified with $v_{\gamma\pi(i)}$ for each $i \in [2m]$. Now, edge $e_{\gamma\pi(i)}$ is glued to $e_{\pi\gamma\pi(i)}$, and by the same argument, the vertex $v_{\gamma\pi(i)}$ (now the tail of the edge in question) gets identified with $v_{\gamma\pi\gamma\pi(i)}$. Continuing this way, we see that v_i is identifies with precisely those v_j for which $j = (\gamma\pi)^\ell(i)$ for some $\ell \in \mathbb{N}$. Thus, the cycles of $\gamma\pi$ count the number of distinct vertices after the gluing. □

This connection is very fortuitous, because it allows us to neatly describe the exponent $\#(\gamma\pi) - m - 1$. Consider the surface $S = S_\pi$ described above. The *Euler characteristic* $\chi(S)$ of S is a well-defined even integer, which (miraculously) can be defined as follows: if G is any imbedded polygonal complex in S, then

$$\chi(S) = V(G) - E(G) + F(G)$$

where $V(G)$ is the number of vertices in G, $E(G)$ is the number of edges in G, and $F(G)$ is the number of faces of G. What's more, $\chi(S)$ is related to another topological invariant of S: its *genus*. Any orientable compact surface is homeomorphic to a g-holed torus for some $g \geq 0$ (the $g = 0$ case is the sphere); this $g = g(S)$ is the genus of the surface. It is a theorem (due to Cauchy – which is why we name it after Euler) (?) that

$$\chi(S) = 2 - 2g(S).$$

Now, consider our imbedded complex G_π in S_π. It is a quotient of a $2m$-gon, which has only 1 face, therefore $F(G_\pi) = 1$. Since we identify edges in pairs, $E(G_\pi) = (2m)/2 = m$. Thus, by the above lemma,

$$2 - 2g(S_\pi) = \chi(S_\pi) = \#(\gamma\pi) - m + 1.$$

Thus: $\#(\gamma\pi) - m - 1 = -2g(S_\pi)$. Returning to Equation (GE1), we therefore have

$$\varphi_n(X_n^{2m}) = \sum_{\pi \in \mathscr{P}_2(2m)} n^{-2g(S_\pi)} = \sum_{g \geq 0} \varepsilon_g(m) \frac{1}{n^{2g}}, \tag{GE2}$$

where

$$\varepsilon_g(m) = \#\{\text{genus-}g \text{ surfaces obtained by gluing pairs of edges in a } 2m\text{-gon}\}.$$

Since the genus of any surface is ≥ 0, this shows in particular that

$$\varphi_n(X_n^{2m}) = \varepsilon_0(m) + O\left(\frac{1}{n^2}\right).$$

So, we have shown that the moments of a GUE_n all converge! It remains only to calculate $\varepsilon_0(m)$; i.e. which $\pi \in \mathscr{P}_2(2m)$ yield a sphere when used to identify edges in a $2m$-gon?

Proposition 3. *The genus of S_π is 0 if and only if $\pi \in NC_2(2m)$.*

Proof. First, suppose that π has a crossing. As Figure 1 below demonstrates, this means that the surface S_π has an imbedded double-ring that cannot be imbedded in S^2. Thus, S_π must have genus ≥ 1.

On the other hand, suppose that π is non-crossing. Then, as we know, π has the following recursive property: there is an interval $\{i, i+1\}$ (addition modulo $2m$) in π, and $\pi - \{i, i+1\}$ is in $NC_2(\{1, \ldots, i-1, i+2, \ldots, 2m\})$. As Figure 2 below shows, gluing along this interval pairing $\{i, i+1\}$ first, we reduce to the case of $\pi - \{i, i+1\}$ gluing a $2(m-1)$-gon *within the plane*.

Since $\pi - \{i, i+1\}$ is also non-crossing, we proceed inductively until there are only two pairings left. It is easy two check that the two non-crossing pairings $\{\{1, 2\}, \{3, 4\}\}$ and

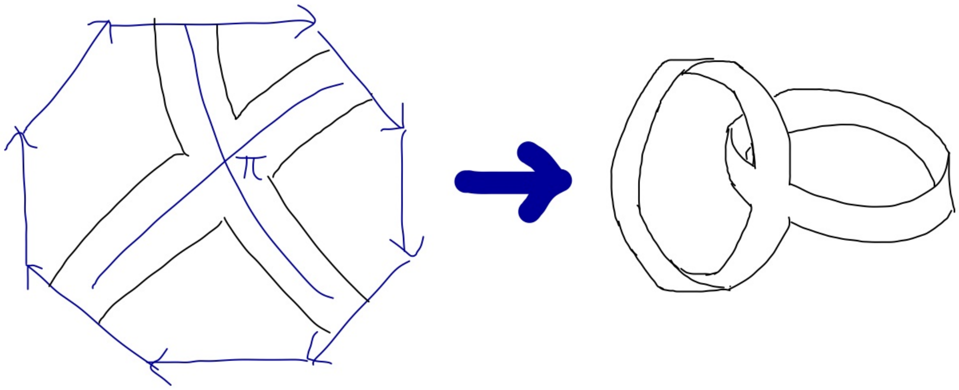

FIGURE 1. If π has a crossing, then the above surface (with boundary) is imbedded into S_π. Since this two-strip surface does not imbed in S^2, it follows the genus of S_π is ≥ 1.

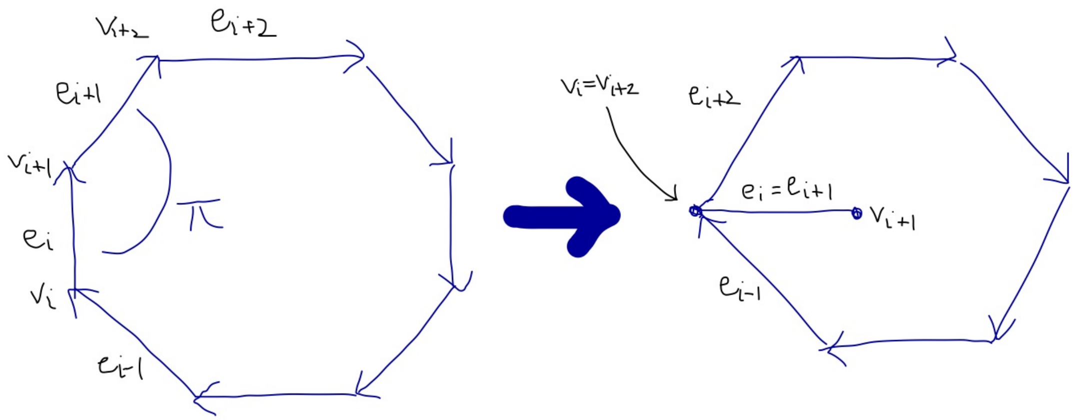

FIGURE 2. Given interval $\{i, i+1\} \in \pi$, the gluing can be done in the plane reducing the problem to size $2(m-1)$.

$\{\{1,4\},\{2,3\}\}$ of a square produce a sphere upon gluing. Thus, if π is non-crossing, $S_\pi = S^2$ which has genus 0. $\square$

Thus, since $\varepsilon_g(m)$ counts the number of $\pi \in \mathscr{P}_2(2m)$ such that $g(S_\pi) = g$, it follows that:

Corollary 4. *In the genus expansion,* $\varepsilon_0(m) = C_m = \frac{1}{m+1}\binom{2m}{m}$.

Exercise. As we did with $\mathscr{P}_2(k)$, we can identify any partition $\pi \in \mathscr{P}(k)$ as a permutation in S_k: the blocks of π are the cycles of the permutation. (To make this well-defined, we must choose an order for each block; canonically, we give each block increasing order.) Recall the Kreweras complementation map $K\colon NC(k) \to NC(k)$. Show that, for $\pi \in NC_2(k)$, $K^{-1}(\pi) = \gamma\pi$.

Exercise. Without reference to genus (or topology in general), show combinatorially that Equation (GE1) yields that $\varphi_n(X_n^{2m}) = C_m + o(1)$. Do this by showing that, for any $\pi \in$

$\mathscr{P}_2(2m)$, $\#\gamma\pi \leq m+1$, and $\#\gamma\pi = m+1$ if and only if $\pi \in NC_2(2m)$. [*Hint*: Use the properties of the Kreweras complementation map.]

4. Wigner's Law

Theorem 5 (Wigner's Semicircular Law)**.** *Let X_n be a GUE_n. Let s be a standard semicircular random variable in some non-commutative probability space. Then $X_n \xrightarrow{\mathscr{D}} s$ as $n \to \infty$.*

Proof. The definition of convergence in distribution is simply $\mu_{X_n} \to \mu_s$ pointwise: i.e. for each non-commutative polynomial P, $\mu_{X_n}(P) \to \mu_s(P)$. In this case, where we have but one self-adjoint variable, the law acts on $\mathbb{C}[X]$, polynomials in one-variable (which are actually commutative), so it suffices to prove that $\mu_{X_n}(X^k) \to \mu_s(X^k)$ for all $k \in \mathbb{N}$. Since $\mu_s(X^k) = 0$ if k is odd and $= C_{k/2}$ if k is even, the genus expansion

$$\mu_{X_n}(X^{2m}) = \varphi_n(X_n^{2m}) = C_m + O\left(\frac{1}{n^2}\right)$$

provides the correct limit for k even; all odd moments of X_n are 0 for all n, completing the proof. □

This is an extremely important example, because of what it says. Recall from Lecture 5, an Exercise asked you to show that, if X is a self-adjoint (deterministic) matrix, the law μ_X (identified as a measure on $\mathbb{R}$) is

$$\mu_X = \frac{1}{n} \sum_{\lambda \in \mathrm{Eig}(X)} \delta_\lambda.$$

In our present situation, we are dealing with a NCPS of random matrices; but the situation is not so different.

Proposition 6. *Let $X \in (\mathcal{M}_n, \varphi_n)$ be self-adjoint. Its eigenvalues $\lambda_1, \ldots, \lambda_n$ are random variables. The distribution μ_X, identified as a measure, is given by*

$$\mu_X = \sum_{j=1}^{n} \mathbb{E}_\omega \delta_{\lambda(\omega)}.$$

To be clear: if $f \in C_c(\mathbb{R})$ is any continuous compactly-supported function, the measure $\nu = \mathbb{E}_\omega \delta_{\lambda(\omega)}$ is given by

$$\int_{\mathbb{R}} f\, d\nu = \mathbb{E} \int_{\mathbb{R}} f\, d\delta_\lambda = \mathbb{E}(f(\lambda)).$$

Proof. **Exercise.** □

Given a random Hermitian matrix X with eigenvalues $\lambda_1, \ldots, \lambda_n$, the *random measure* $\frac{1}{n}(\delta_{\lambda_1} + \cdots + \delta_{\lambda_n})$ is called the **empirical eigenvalue distribution** of X. So, the law μ_X in $(\mathcal{M}_n, \varphi_n)$ is the **expected** empirical eigenvalue distribution. Wigner's law says that this converges to the semicircular distribution for a GUE.

In fact, while this is the precise statement of the theorem that Wigner proved in 1955, the theorem has been greatly improved since then. In the 1960s, it was shown that one need not average the empirical measure to get convergence.

Theorem 7. *Let X_n be a GUE_n, and let μ_n be its empirical eigenvalue distribution (which is a random measure). Then $\mu_n \to \frac{1}{2\pi}\sqrt{4-t^2}\mathbb{1}_{|t|\leq 2}\,dt$ almost surely as $n \to \infty$.*

That is: there is almost sure convergence of the random measure, not just convergence of its expectation. This is proved using a standard technique, as expressed in the following exercise.

Exercise. Let Y_n, Y be L^1 random variables such that $\mathbb{E}(Y_n) \to \mathbb{E}(Y)$. Suppose $\mathrm{Var} Y_n \to 0$. Show that $Y_n \to Y = \mathbb{E}(Y)$ $a.s.$

So, the second version of Wigner's law was proved by showing that $\mathrm{Var}_\omega\left(\frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j(\omega)}\right) \to 0$ as $n \to \infty$. (This proof is not as combinatorial as the ones we've seen so far.)

Further generalizations of the theorem involve removing the conditions on the matrix X_n. One can replace independence of the entries with somewhat weaker requirements, like long-range independence. More interesting is the fact that the normal distribution of the entries is not at all required. This version of the theorem (the most general to date) is due to Soshnikov, from the last 10 years. We state it here for the real-valued case.

Theorem 8. *Let X_n be a random symmetric matrix such that $[X_n]_{ij}$ are i.i.d. for $1 \leq i \leq j \leq n$, with $\mathbb{E}([X_n]_{ij}) = 0$ and $\mathbb{E}([X_n]_{ij}^2) = \frac{1}{n}$. Then the empirical eigenvalue distribution of X_n converges almost surely to Wigner's semicircle law as $n \to \infty$.*

This is a so-called *universality* law: the limit distribution does not depend on the distribution of the entries – only on their first and second moments, much like in the central limit theorem.

The question of why this is all so important can now be understood with a picture. Let X_n be a Hermtian matrix whose upper-diagonal entries are independent random variables all sampled from the same distribution (with properly normalized first and second moments). Wigner's law says that the empirical eigenvalue distribution of X_n is close to the semicircle law. What does this mean? The empirical eigenvalue distribution puts a point mass at each eigenvalue; if we "smooth it out" a little bit (by averaging it over a grid of small bins), what we get is a *histogram of eigenvalues*. The following plot is the actual histogram of eigenvalues of (one instance of a) GUE_{4000}, next to a plot of the semicircular density (scaled by a factor of 4000 to match up to the histogram):