Lectures on Random Matrices (Spring 2025)
Lecture 2: Wigner semicircle law

Leonid Petrov

Wednesday, January 15, 2025*

Notes for the lecturer

PREP:

1. Start: Catalan number formula
2. Moments of SC need to be computed
3. SC is uniquely determined by its moments; need Carleman criterion to show that the moments determine the distribution.
4. from expected moments to the full convergence, some analysis needed

1 Recap

We are working on the Wigner semicircle law.

1. Wigner matrices W: real symmetric random matrices with iid entries X_{ij}, $i > j$ (mean 0, variance σ^2); and iid diagonal entries X_{ii} (mean 0, some other variance and distribution).
2. Empirical spectral distribution (ESD)
$$\nu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i/\sqrt{n}},$$
which is a random probability measure on $\mathbb{R}$.
3. Semicircle distribution μ_{sc}:
$$\mu_{\mathrm{sc}}(dx) = \frac{1}{2\pi}\sqrt{4-x^2}\,dx, \qquad x \in [-2,2].$$
4. Computation of expected traces of powers of W. We showed that
$$\int_{\mathbb{R}} x^k\,\nu_n(dx) \to \#\left\{\text{rooted planar trees with } k/2 \text{ edges}\right\}.$$

*Course webpage • TeX Source • Updated at 15:58, Tuesday 14th January, 2025

2 Two computations

First, we finish the combinatorial part, and match the limiting expected traces of powers of W to moments of the semicircle law.

2.1 Moments of the semicircle law

We also need to match the Catalan numbers to the moments of the semicircle law. Let $k = 2m$, and we need to compute the integral

$$\int_{-2}^{2} x^{2m} \frac{1}{2\pi} \sqrt{4 - x^2}\, dx.$$

By symmetry, we write:

$$\int_{-2}^{2} x^{2m} \rho(x)\, dx = \frac{2}{\pi} \int_{0}^{2} x^{2m} \sqrt{4 - x^2}\, dx.$$

Using the substitution $x = 2\sin\theta$, we have $dx = 2\cos\theta\, d\theta$. The integral becomes:

$$\frac{2}{\pi} \int_0^{\pi/2} (2\sin\theta)^{2m} (2\cos\theta)(2\cos\theta\, d\theta) = \frac{2^{2m+2}}{\pi} \int_0^{\pi/2} \sin^{2m}\theta \cos^2\theta\, d\theta.$$

Using $\cos^2\theta = 1 - \sin^2\theta$, we split the integral:

$$\frac{2^{2m+2}}{\pi} \left(\int_0^{\pi/2} \sin^{2m}\theta\, d\theta - \int_0^{\pi/2} \sin^{2m+2}\theta\, d\theta \right).$$

Using the standard formula (cf. Problem B.1)

$$\int_0^{\pi/2} \sin^{2n}\theta\, d\theta = \frac{\pi}{2} \frac{(2n)!}{2^{2n}(n!)^2}, \tag{2.1}$$

we compute each term:

$$\frac{2^{2m+2}}{\pi} \left(\frac{\pi}{2} \frac{(2m)!}{2^{2m}(m!)^2} - \frac{\pi}{2} \frac{(2m+2)!}{2^{2m+2}((m+1)!)^2} \right).$$

After simplification, this becomes C_m, the m-th Catalan number.

2.2 Counting trees and Catalan numbers

Throughout this section, for a random matrix trace moment of order k, we use $m = k/2$ as our main parameter. Note that m can be arbitrary (not necessarily even).

Definition 2.1 (Dyck Path)**.** A *Dyck path* of semilength m is a sequence of $2m$ steps in the plane, each step being either $(1, 1)$ (up step) or $(1, -1)$ (down step), starting at $(0, 0)$ and ending at $(2m, 0)$, such that the path never goes below the x-axis. We denote an up step by U and a down step by D.

Definition 2.2 (Rooted Plane Tree)**.** A *rooted plane tree* is a tree with a designated root vertex where the children of each vertex have a fixed left-to-right ordering. The size of such a tree is measured by its number of edges, which we denote by m.

Definition 2.3 (Catalan Numbers)**.** The sequence of *Catalan numbers* $\{C_m\}_{m\geq 0}$ is defined recursively by:

$$C_0 = 1, \quad C_{m+1} = \sum_{j=0}^{m} C_j C_{m-j} \quad \text{for } m \geq 0. \tag{2.2}$$

Alternatively, they have the closed form:

$$C_m = \frac{1}{m+1}\binom{2m}{m} = \binom{2m}{m} - \binom{2m}{m+1}. \tag{2.3}$$

These numbers appear naturally in the moments of random matrices, where $m = k/2$ for trace moments of order k.

Lemma 2.4. *Formulas* (2.2) *and* (2.3) *are equivalent.*

Proof. One can check that the closed form satisfies the recurrence relation by direct substitution. The other direction involves generating functions. Namely, (2.2) can be rewritten for the generating function

$$C(z) = \sum_{m=0}^{\infty} C_m z^m$$

as

$$C(z) = 1 + zC(z)^2.$$

Solving for $C(z)$, we get

$$C(z) = \frac{1 \pm \sqrt{1-4z}}{2z}. \tag{2.4}$$

We need to pick the solution which is nonsingular at $z = 0$, and it corresponds to the minus sign. Taylor expansion of the right-hand side of (2.4) at $z = 0$ gives the closed form. ☐

Remark 2.5. Catalan numbers enumerate many (too many!) combinatorial objects. For a comprehensive treatment, see [Sta15].

Proposition 2.6 (Dyck Path–Rooted Tree Correspondence)**.** *For any m, there exists a bijection between the set of Dyck paths of semilength m and the set of rooted plane trees with m edges.*

Proof. Given a Dyck path of semilength m, we build the corresponding rooted plane tree as follows (see Figure 1 for an illustration):

1. Start with a single root vertex
2. Read the Dyck path from left to right:
   - For each up step (U), add a new child to the current vertex
   - For each down step (D), move back to the parent of the current vertex

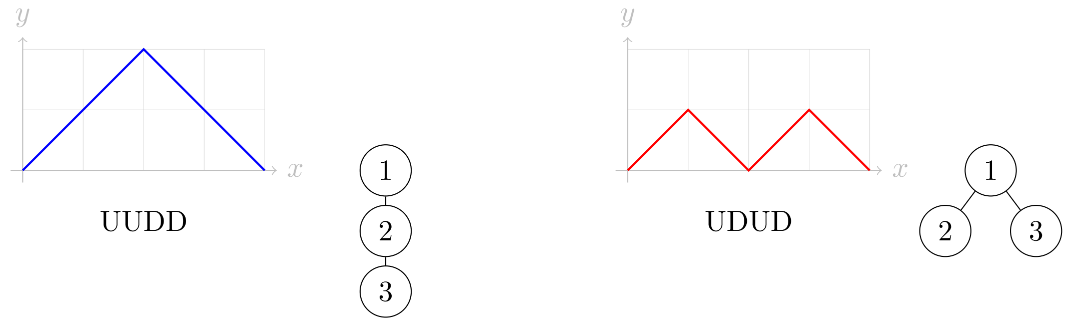

Figure 1: The two possible Dyck paths of semilength $m = 2$ and their corresponding rooted plane trees.

3. The order of children is determined by the order of up steps

This is clearly a bijection, and we are done. $\square$

It remains to show that the Dyck paths or rooted plane trees are counted by the Catalan numbers, by verifying the recursion (2.2) for them. By Proposition 2.6, it suffices to consider only Dyck paths.

Proposition 2.7. *The number of Dyck paths of semilength m satisfies the Catalan recurrence* (2.2).

Proof. We need to show that the number of Dyck paths of semilength $m+1$ is given by the sum in the right-hand side of (2.2). Consider a Dyck path of semilength $m+1$, and let the *first* time it returns to zero be at semilength $j+1$, where $j = 0, \ldots, m$. Then the first and the $(2j+1)$-st steps are, respectively, U and D. From 0 to $2j+2$, the path does not return to the x-axis, so we can remove the first and the $(2j+1)$-st steps, and get a proper Dyck path of semilength j. The remainder of the Dyck path is a Dyck path of semilength $m-j$. This yields the desired recurrence. $\square$

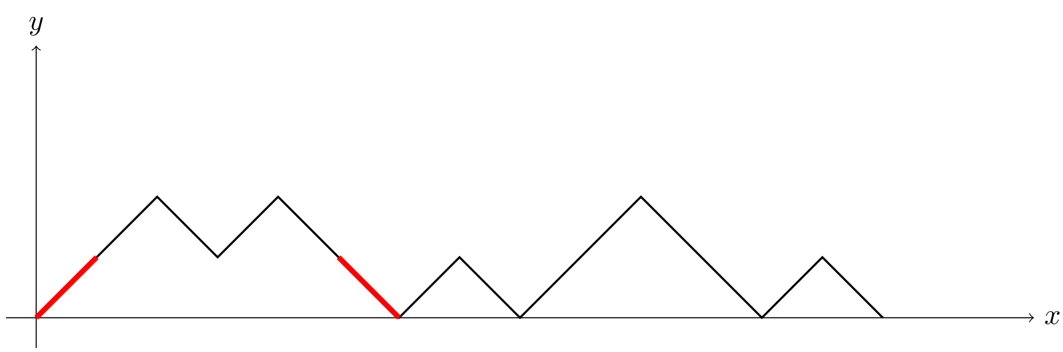

Figure 2: Illustration of a Dyck path decomposition for the proof of Proposition 2.7.

3 Analysis steps in the proof

We are done with combinatorics, and it remains to justify that the computations lead to the desired semicircle law from Lecture 1.

Let us remember that so far, we showed that

$$\lim_{n\to\infty} \frac{1}{n^{k/2+1}} \mathbb{E}\left[\operatorname{Tr} W^k\right] = \begin{cases} \sigma^{2m} C_m & \text{if } k = 2m \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Here, W is real Wigner (unnormalized) with mean 0, where its off-diagonal entries are iid with variance σ^2.

3.1 The semicircle distribution is determined by its moments

We use (without proof) the known Carleman's criterion for the uniqueness of a distribution by its moments.

Proposition 3.1 (Carleman's criterion [ST43, Theorem 1.10], [Akh65])**.** *Let X be a real-valued random variable with moments $m_k = \mathbb{E}[X^k]$ of all orders. If*

$$\sum_{k=1}^{\infty} (m_{2k})^{-1/(2k)} = \infty,$$

then the distribution of X is uniquely determined by its moments $(m_k)_{k\geq 1}$.

Remark 3.2. By the Stone-Wierstrass theorem, the semicircle distribution is the unique distribution with compact support with these moments. However, we need to guarantee that there are no distributions on $\mathbb{R}$ with the same moments.

Now, the moments satisfy the asymptotics

$$m_{2k} = C_k \sigma^{2k} \sim \frac{4^k}{k^{3/2}\sqrt{\pi}} \sigma^{2k},$$

so

$$\sum_{k=1}^{\infty} (m_{2k})^{-1/(2k)} \sim \sum_{k=1}^{\infty} \left(\frac{k^{3/2}\sqrt{\pi}}{4^k}\right)^{1/2k} \sigma^{-1}.$$

The k-th summands converges to $1/(2\sigma)$, so the series diverges.

Remark 3.3. See also Problem A.4 from Lecture 1 on an example of a distribution not determined by its moments.

3.2 Convergence to the semicircle law

Recall [Bil95, Theorem 30.2] that convergence of random variables in moments plus the fact that the limiting distribution is uniquely determined by its moments implies convergence in distribution. However, we need weak in probability convergence, which deals with random variables

$$\int_{\mathbb{R}} f(x)\, \nu_n(dx), \qquad f \in C_b(\mathbb{R}),$$

and we did not compute the moments of these random variables.

To complete the argument, let us show that for each fixed integer $k \geq 1$, we have almost sure convergence of the moments (of a random distribution, so that the $Y_{n,k}$'s are random variables):

$$Y_{n,k} := \int_{\mathbb{R}} x^k \, \nu_n(dx) \xrightarrow[n\to\infty]{\text{a.s.}} m_k, \qquad n \to \infty,$$

where m_k are the moments of the semicircle distribution, and ν_n is the ESD corresponding to the scaling of the eigenvalues as $\lambda_i/\sqrt{n}$.

As typical in asymptotic probability, we not only need the expectation of $Y_{n,k}$, but also their variances, to control the almost sure convergence. Recall that we showed $\mathbb{E}(Y_{n,k}) \to m_k$. Let us assume the following:

Proposition 3.4 (Variance bound)**.** *For each fixed integer $k \geq 1$ and large enough n, we have*

$$\operatorname{Var}(Y_{n,k}) \leq \frac{m_k}{n^2}.$$

We will prove Proposition 3.4 in Section 4 below. Let us finish the proof of convergence to the semicircle law modulo Proposition 3.4.

3.2.1 A concentration bound and the Borel–Cantelli lemma

From Chebyshev's inequality,

$$\mathbb{P}\Big(\big|Y_{n,k} - \mathbb{E}[Y_{n,k}]\big| \geq n^{-\frac{1}{4}}\Big) \leq \operatorname{Var}[Y_{n,k}]\sqrt{n} = O(n^{-\frac{3}{2}}),$$

where in the last step we used Proposition 3.4.

Hence the probability that $|Y_{n,k} - \mathbb{E}[Y_{n,k}]| > n^{-\frac{1}{4}}$ is summable in n. By the Borel–Cantelli lemma, with probability 1 only finitely many of these events occur. Since $\mathbb{E}[Y_{n,k}] \to m_k$, we conclude

$$\big|Y_{n,k} - m_k\big| \;\leq\; \big|Y_{n,k} - \mathbb{E}[Y_{n,k}]\big| + \big|\mathbb{E}[Y_{n,k}] - m_k\big| \xrightarrow[n\to\infty]{} 0 \quad \text{almost surely.}$$

3.2.2 Tightness of $\{\nu_n\}$ and subsequential limits

Since $|Y_{n,k}| = \big|\int x^k \, \nu_n(dx)\big|$ stays almost surely bounded for each k, one readily checks (Problem B.4) that almost surely, for each fixed k,

$$\nu_n\big(\{x : |x| > M\}\big) \;\leq\; \frac{C}{M^k}. \tag{3.1}$$

By choosing k large, we see that ν_n puts arbitrarily little mass outside any large interval $[-m, m]$. Thus, the sequence of probability measures $\{\nu_n\}$ is *tight*. By Prokhorov's theorem [Bil95, Theorem 25.10], there exists a subsequence ν_{n_j} converging weakly to some probability measure ν^*. We will now characterize all subsequential limits ν^* of ν_n.

3.2.3 Characterizing the limit measure

We claim that $\nu^* = \mu_{\mathrm{sc}}$, the semicircle distribution (and in particular, this measure is not random). Indeed, fix k. Since x^k is a bounded function on a sufficiently large interval, and $\nu_{n_j} \to \nu^*$ weakly, we have

$$\int_{\mathbb{R}} x^k \, \nu_{n_j}(dx) \;\to\; \int_{\mathbb{R}} x^k \, \nu^*(dx).$$

On the other hand, we have already shown

$$\int_{\mathbb{R}} x^k \, \nu_{n_j}(dx) \;=\; Y_{n_j,k} \;\xrightarrow[j\to\infty]{\text{a.s.}}\; m_k \;=\; \int_{\mathbb{R}} x^k \, \mu_{\mathrm{sc}}(dx).$$

Thus

$$\int_{\mathbb{R}} x^k \, \nu^*(dx) \;=\; m_k \;=\; \int_{\mathbb{R}} x^k \, \mu_{\mathrm{sc}}(dx) \qquad \text{for all } k \ge 1.$$

By Proposition 3.1, the measure ν^* is uniquely determined by its moments. Hence ν^* must coincide with μ_{sc}.

Remark 3.5. In Sections 3.2.2 and 3.2.3 we tacitly assumed that we choose an elementary outcome ω, and view ν_n as measures depending on ω. Then, since the convergence of moments is almost sure, ω belongs to a set of full probability. The limiting measure ν^* must coincide with μ_{sc} for this ω, and thus, ν^* is almost surely nonrandom.

Any subsequence of $\{\nu_n\}$ has a further sub-subsequence convergent to ν. By a standard diagonal argument, this forces $\nu_n \to \nu$ in the weak topology (almost surely). This completes the proof that the ESD of our Wigner matrix (rescaled by $\sqrt{n}$) converges to the semicircle distribution weakly almost surely, modulo Proposition 3.4. (See also Problem B.4.1 for the weakly in probability convergence.)

4 Proof of Proposition 3.4: bounding the variance

There is one more "combinatorial" step in the proof of the semicircle law: we need to show that the variance of the moments of the ESD is bounded by m_k/n^2.

Recall that

$$Y_{n,k} = \int_{\mathbb{R}} x^k \, \nu_n(dx) = \frac{1}{n^{1+\frac{k}{2}}} \sum_{i_1,\ldots,i_k=1}^{n} X_I, \quad \text{where} \quad X_I = X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_k i_1}.$$

Here we use the notation I for the multi-index $(i_1, \ldots, i_k)$, and throughout the computation below, we use the notation $I \in [n]^k$, where $[n] = \{1, \ldots, n\}$. We have

$$\operatorname{Var}\bigl(Y_{n,k}\bigr) = \frac{1}{n^{2+k}} \operatorname{Var}\Bigl(\sum_{I\in[n]^k} X_I\Bigr) = \frac{1}{n^{2+k}} \sum_{I,J\in[n]^k} \operatorname{Cov}\bigl(X_I, X_J\bigr).$$

We claim that the sum of all covariances is bounded by a constant times n^k, which then implies $\operatorname{Var}\bigl(Y_{n,k}\bigr) \le \text{const} \cdot n^k/n^{2+k} = O\bigl(\frac{1}{n^2}\bigr)$.

Step 1. Identifying when $\mathrm{Cov}\big(X_I, X_J\big)$ can be nonzero. For each k-tuple $I = (i_1, i_2, \dots, i_k) \in [n]^k$, the product

$$X_I = X_{i_1 i_2}\, X_{i_2 i_3}\, \dots\, X_{i_k i_1}$$

is the product of the entries of our Wigner matrix corresponding to the directed "edges" $(i_1 \to i_2), (i_2 \to i_3), \dots, (i_k \to i_1)$. Similarly, X_J is determined by the edges of another closed directed walk J.

1. If I and J use disjoint collections of matrix entries, then X_I and X_J are independent, and hence $\mathrm{Cov}(X_I, X_J) = 0$.
2. If there is an edge (say, $X_{i_1 i_2}$) which appears *only once* in exactly one of I or J but not both, then that edge factor is independent and forces $\mathrm{Cov}(X_I, X_J) = 0$ since $\mathbb{E}[X_{i_1 i_2}] = 0$. Indeed, for example if $X_{i_1 i_2}$ appears only in X_I, then

$$\mathbb{E}\left[X_I\right] = \mathbb{E}\left[X_{i_1 i_2}\right] \cdot \mathbb{E}\left[\text{other factors}\right] = 0, \qquad \mathbb{E}\big[X_I X_J\big] = \mathbb{E}[X_{i_1 i_2}] \cdot \mathbb{E}\big[\text{other factors}\big] = 0.$$

Thus, the only way we could get a nonzero covariance is if *every* edge that appears in $I \cup J$ appears at least twice overall. Graphically, let us represent each k-tuple I by a directed closed walk in the complete graph on $[n]$. Then the union $I \cup J$ must be a subgraph in which every directed edge has total multiplicity ≥ 2.

Step 2. Counting the contributions to the sum. Denote by $q = |V(I \cup J)|$ the number of distinct vertices (out of the possible n) involved in the union $I \cup J$. We can pick which q vertices are in the union in $\binom{n}{q} \approx n^q$ ways (times some constant), and then specify how the edges form *two* closed walks of length k.

We split into two cases:

- $\boldsymbol{q \leq k}$. In this situation, each pair (I, J) with $|V(I \cup J)| = q$ contributes a term to $\sum_{I,J} \mathrm{Cov}(X_I, X_J)$. However, because each edge is used at least twice in the union $I \cup J$, and each walk has length k, there are at most k edges in a single walk. One can check that such "doubly used edges" cannot produce too large a contribution. A more direct combinatorial check (cf. the argument in the lecture notes) shows these terms together are of order $O(n^k)$.
- $\boldsymbol{q \geq k+1}$. In this case, to use at most k edges in each of the two walks, one sees that the union $I \cup J$ *must* form a "double tree" structure: the same edge is traced out in both I and J, and each vertex has total in-degree and out-degree equal to 2 when counting both walks. Geometrically, removing extra edges shows that I is (itself) a double tree, and so is J. Then one verifies (see the lecture notes) that this forces at least one single edge to appear exactly once—which we already saw makes the covariance vanish!

  Hence no further nontrivial contributions appear from the case $q \geq k+1$.

Combining these two cases, we conclude that the total number of pairs (I, J) with nonzero covariance is of order at most n^k, and each covariance is at most a bounded constant in absolute value. Thus

$$\sum_{I,J \in [n]^k} \mathrm{Cov}\big(X_I, X_J\big) = O(n^k).$$

Multiplying by $1/n^{2+k}$ completes the proof:

$$\operatorname{Var}(Y_{n,k}) = \frac{1}{n^{2+k}} \sum_{I,J\in[n]^k} \operatorname{Cov}\big(X_I, X_J\big) \le \frac{C\,n^k}{n^{2+k}} = \frac{C}{n^2}. \quad \square$$

B Problems (due 2025-02-15)

B.1 Standard formula

Prove formula (2.1):

$$\int_0^{\pi/2} \sin^{2n}\theta\, d\theta = \frac{\pi}{2}\frac{(2n)!}{2^{2n}(n!)^2}.$$

B.2 Tree profiles

Show that the expected height of a uniformly random Dyck path of semilength m is of order $\sqrt{m}$.

B.3 Ballot problem

Suppose candidate A receives p votes and candidate B receives q votes, where $p > q \ge 0$. In how many ways can these votes be counted such that A is always strictly ahead of B in partial tallies?

B.4 Bounding probability in the proof

Show inequality (3.1).

B.4.1 Almost sure convergence and convergence in probability

Show that in Wigner's semicircle law, the weakly almost sure convergence of random measures ν_n to μ_{sc} implies weak convergence in probability.

References

[Akh65] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Translated by N. Kemmer, Hafner Publishing Co., New York, 1965. ↑5

[Bil95] P. Billingsley, *Probability and Measure*, 3rd ed., Wiley Ser. Probab. Stat., Wiley, 1995. ↑5, 6

[ST43] J. Shohat and J. D. Tamarkin, *The Problem of Moments*, American Mathematical Society (AMS), 1943. ↑5

[Sta15] R. Stanley, *Catalan numbers*, Cambridge University Press, 2015. ↑3

L. Petrov, University of Virginia, Department of Mathematics, 141 Cabell Drive, Kerchof Hall, P.O. Box 400137, Charlottesville, VA 22904, USA

E-mail: lenia.petrov@gmail.com