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 13:53, 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. $\square$

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

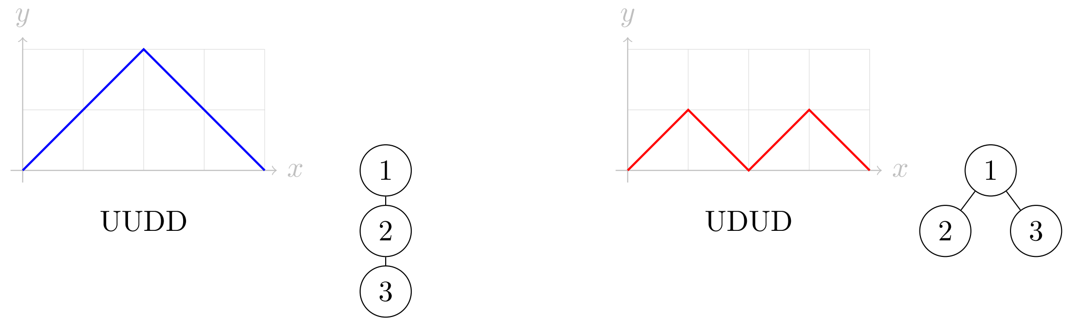

Figure 1: The two possible Dyck paths of semilength $m = 2$ and their corresponding rooted plane trees.

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. $\square$

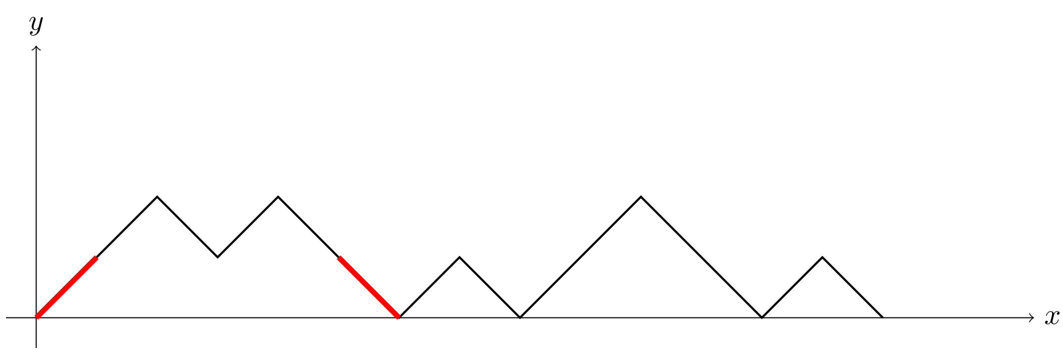

Figure 2: Illustration of a Dyck path decomposition for the proof of Proposition 2.7.

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 ([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

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 $p > q \geq 0$. In how many ways can these votes be counted such that A is always strictly ahead of B in partial tallies?

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

[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`