3. Wigner Matrices: Combinatorial Proof of Wigner's Semicircle Law

Wigner's semicircle law does not only hold for Gaussian random matrices, but more general for Wigner matrices; there we keep the independence and identical distribution of the entries, but allow arbitrary distribution instead of Gaussian. As there is no Wick formula any more, there is no clear advantage of the complex over the real case any more, hence we consider now the real one.

3.1. Definition: Let μ be a probability distribution on $\mathbb{R}$. A corresponding Wigner random matrix is of the form

$$A_N = \frac{1}{\sqrt{N}} (a_{ij})_{i,j=1}^{N}, \text{ where}$$

- $A_N = A_N^*$, i.e. $a_{ij} = a_{ji}$ $\forall i,j$
- $\{a_{ij} \mid i \geq j\}$ are independent
- each a_{ij} has distribution μ

3.2. Remarks: 1) In our combinatorial setting we will assume that all moments of μ exist; that the first moment is zero; and the second moment will usually be normalized to 1. In an analytic setting one can deal with more general situations; usually only the existence of the second moment is needed; and one can also allow a non-vanishing mean.

2) Often one also allows different distributions for the diagonal and the off-diagonal elements.

3) Even more general, one can give up the assumption of idendical distribution of the entries and replace this by uniform bounds on their moments.

4) We will now try to imitate our combinatorial proof from the Gaussian case also in this more general situation. Without a precise Wick formula for the higher moments of the entries, we will not aim at a precise genus expansion; it suffices to see that the leading contributions are still given by Catalan numbers

Consider a Wigner matrix

$$A_N = \frac{1}{\sqrt{N}} (a_{ij})_{i,j=1}^{N}$$

where μ has all moments and

$$\int x \, d\mu(x) = 0 \quad , \quad \int x^2 d\mu(x) = 1$$

Then we have

$$E[tr(A_N^m)] =$$

$$= \frac{1}{N^{m/2+1}} \sum_{i_1,\dots,i_m=1}^{N} E[a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_m i_1}]$$

$$= \frac{1}{N^{m/2+1}} \sum_{\sigma \in \mathcal{P}(m)} \sum_{\substack{i:[m]\to[N] \\ \ker i = \sigma}} E[\sigma]$$

where we group the appearing indices $i=(i_1,\dots,i_m)$ according to their "kernel", which is a partition σ of $\{1,\dots,m\}$

3.3. <u>Definition</u>: 1) A <u>partition</u> σ of $[n]$ is a decomposition of $[n]$ into disjoint, non-empty subsets (of arbitrary size), i.e., $\sigma = \{V_1,\dots,V_k\}$ where

- $V_i \subset [n] \quad \forall i$
- $V_i \neq \emptyset \quad \forall i$
- $V_i \cap V_j = \emptyset \quad \forall i \neq j$
- $\bigcup_{i=1}^{k} V_i = [n]$

The V_i are called <u>blocks</u> of σ.

The set of all partitions of $[n]$ is denoted by

$P(n) := \{\sigma \mid \sigma \text{ partition of } [n]\}$

2) For a multi-index $i = (i_1, \dots, i_n)$ we denote by $\ker i$ its <u>kernel</u>; this is the partition $\sigma \in P(n)$, such that we have

$$i_k = i_l \iff k \text{ and } l \text{ are in the same block of } \sigma$$

If we identify i with a function

$$i : [n] \to [N] \qquad \text{via } i(k) = i_k$$

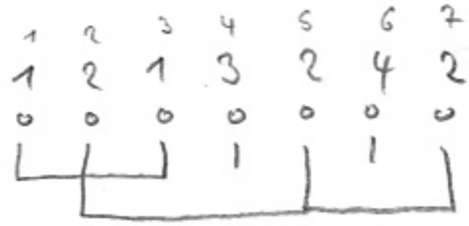
then we can also write

$$\ker i = \{ i^{-1}(1), i^{-1}(2), \dots, i^{-1}(N) \}$$

where we discard all empty sets

<u>3.4. Example:</u> For $i = (1,2,1,3,2,4,3)$

we have

$\ker i =$ 

$= \{(1,3), (2,5,7), (4), (6)\}$

$\in \mathcal{P}(7)$

<u>3.5. Remark:</u> The relevance of this kernel in our setting is that we have for $i = (i_1, \dots, i_m)$ and $j = (j_1, \dots, j_m)$:

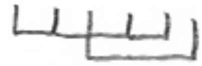
$\ker i = \ker j$ implies that

$$E[a_{i_1 i_2} \cdots a_{i_m i_1}] = E[a_{j_1 j_2} \cdots a_{j_m j_1}]$$

For example, for

$i = (1,1,2,1,1,2)$

$j = (2,2,7,2,2,7)$

we have $\ker i =$  $= \ker j$

and

$$E[a_{11} a_{12} a_{21} a_{11} a_{12} a_{21}] = E[a_{11}^2]\, E[a_{12}^4]$$

$$\shortparallel$$

$$E[a_{22} a_{27} a_{72} a_{22} a_{27} a_{72}] = E[a_{22}^2]\, E[a_{27}^4]$$

We denote this common value by

$$E(\sigma) := E[a_{i_1 i_2} \cdots a_{i_m i_1}] \quad \text{if } \ker i = \sigma$$

Thus we get

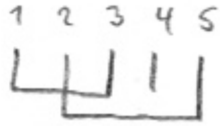
$$E[\mathrm{tr}(A_N^m)] = \frac{1}{N^{1+m/2}} \sum_{\sigma \in P(m)} E[\sigma] \cdot \#\{i : [m] \to \dots ;\ \ker i = \sigma\}$$

To understand the contribution corresponding to a $\sigma \in P(m)$ we associate to σ a graph $\mathcal{G}_\sigma$.

<u>3.6. Definition</u>: For $\sigma \overset{\{V_1,\dots,V_k\}}{=} \in P(m)$ we define a corresponding graph $\mathcal{G}_\sigma$ as follows. The vertices of $\mathcal{G}_\sigma$ are given by the blocks V_p of σ, and there is an edge between V_p and V_q if there is an $r \in [m]$ s.th. $r \in V_p$ and $\underset{(\mathrm{mod}\ m)}{r+1} \in V_q$

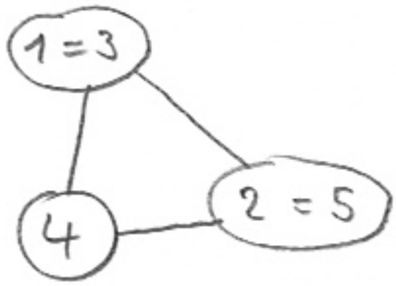
Another way of saying this is that we start with a graph with vertices $1, 2, \dots, m$ and edges $(1,2), (2,3), (3,4), \dots (m-1, m), (m, 1)$

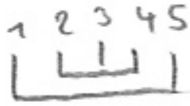
and then identify vertices according to the blocks of σ; we keep loops, but erase multiple edges.

<u>3.7. Examples:</u> For $\sigma =$  $= \{(1,3),(2,5),(4)\}$

we have

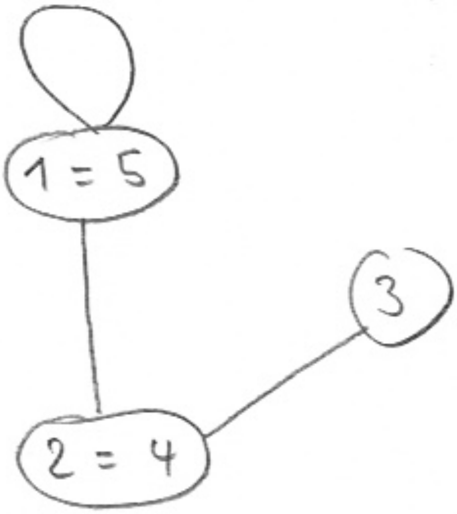
$\mathcal{G}_\sigma =$



whereas $\sigma =$  $= \{(1,5),(2,4),(3)\}$

gives

$\mathcal{G}_\sigma =$



and $\sigma =$  $= \{(1,3),(2),(4)\}$

gives

$\mathcal{G}_\sigma =$

The term $E[a_{i_1 i_2} \cdots a_{i_m i_1}]$ corresponds now to a walk in the graph $\mathcal{G}_\sigma$, with $\sigma = \ker$ along the edges, with steps $i_1 \to i_2 \to i_3 \to \cdots$
$\cdots \to i_{m-1} \to i_m \to i_1$

Hence we are using each edge in $\mathcal{G}_\sigma$ at least once.

Note that different edges in $\mathcal{G}_\sigma$ correspond to independent random variables; hence: if we use an edge only once in our walk then $E[\sigma] = 0$ because the expectation factorises into a product with one factor being a first moment of a_{ij}, which is assumed to be zero.

Thus every edge must be used at least twice, but this implies

$$\#\text{edges in } \mathcal{G}_\sigma \leq \frac{\#\text{steps in the walk}}{2} = \frac{m}{2}$$

Since the number of i with the same kernel is easy to count,

$$\#\{i : [m] \to [N] : \ker i = \sigma\} = N(N-1)(N-2) \cdots (N - \#\sigma + 1)$$

where

$\#\sigma$ = number of blocks in σ, we get finally

$$(*) \quad E[\mathrm{tr}(A_N^m)] = \frac{1}{N^{1+m/2}} \sum_{\substack{\sigma \in \mathcal{P}(m) \\ \#E(\mathcal{G}_\sigma) \le \frac{m}{2}}} E[\sigma] \cdot \underbrace{N(N-1)\cdots(N-\#\sigma+}_{\sim N^{\#\sigma} \text{ for } N \to \infty}$$

We have to understand what the constraint on the number of edges in $\mathcal{G}_\sigma$ gives us for the number of vertices in $\mathcal{G}_\sigma$ (which is the same as $\#\sigma$)

For this recall (or prove as an exercise) the following result about graphs.

<u>3.8. Proposition:</u> Let $\mathcal{G} = (V, E)$ be a connected finite graph with vertices V and edges E. (We allow loops and multiedges.) Then we have that

$$\#V \le \#E + 1$$

and we have equality if and only if $\mathcal{G}$ is a tree (i.e. a connected graph without cycles)

This yields then

3.9 Theorem (Wigner's semicircle law for Wigner matrices; averaged version):

Let A_N be a Wigner matrix corresponding to μ which has all moments, with mean zero and second moment 1. Then we have for all $m \in \mathbb{N}$

$$\lim_{N\to\infty} E[tr(A_N^m)] = \frac{1}{2\pi}\int_{-2}^{+2} x^m \sqrt{4-x^2}\, dx$$

Proof: From (*) we get

$$\lim_{N\to\infty} E[tr(A_N^m)] = \sum_{\sigma\in P(m)} E[\sigma]\cdot \lim_{N\to\infty} N^{\#\sigma - \frac{m}{2}-1}$$

In order to have $E[\sigma]\neq 0$ we can restrict to σ with $\# E(\mathcal{G}_\sigma) \leq \frac{m}{2}$, which implies by 3.8 that

$$\# V(\mathcal{G}_\sigma) \leq \#(\mathcal{G}_\sigma)+1 \leq \frac{m}{2}+1$$

Hence all terms converge and the only contribution in the limit $N\to\infty$ comes from those σ, where we have equality, i.e.

$$\# V(\mathcal{G}_\sigma) = \#(\mathcal{G}_\sigma)+1 = \frac{m}{2}+1$$

Hence $\mathcal{G}_\sigma$ must be a tree and in our walk we use each edge exactly twice (necessarily in opposite direction). For such a σ we have $E[\sigma] = 1$; thus

$$\lim_{N\to\infty} E[\mathrm{tr}(A_N^m)] = \#\{\sigma \in P(m) \mid \mathcal{G}_\sigma \text{ is a tree}\}$$

We will check in an exercise that the latter number is also counted by Catalan numbers. □

<u>3.10 Remark</u>: Note that our $\mathcal{G}_\sigma$ are not just abstract trees, but they are coming with the walks, which encode

- a starting point, i.e. the $\mathcal{G}_\sigma$ are rooted trees
- a cyclic order of the outgoing edges at a vertex, which gives us a planar drawing of our graph

Hence what we have to count are rooted planar trees.

Examples: i) $m = 2$:

ii) $m = 4$:

iii) $m = 6$:

those are different as plana trees

Note also that a rooted plana tree determines uniquely the corresponding walk :