2

Wigner's Matrices; More Moments Estimates

In this chapter, we elaborate upon the previous computation of moments in two directions. First we give a better estimate of the error to the previous limit and prove a central limit theorem. Second, we consider the case where moments are taken at powers that blow up with the dimension of the matrices; we basically show that if this power is small compared to the square root of the dimension, the first-order contribution is still given, in the moment expansion, by graphs that are trees.

2.1 Central Limit Theorem

In the previous section, we proved Wigner's theorem by evaluating $\int x^p dL_{\mathbf{A}^N}(x)$ for $p \in \mathbb{N}$. We shall push this computation one step further here and prove a central limit theorem. Namely, setting

$$\int x^k d\bar{L}_{\mathbf{A}^N}(x) := \mathbb{E}\left[\int x^k dL_{\mathbf{A}^N}(x)\right],$$

we shall prove that

$$M_k^N := N\left(\int x^k dL_{\mathbf{A}^N}(x) - \int x^k d\bar{L}_{\mathbf{A}^N}(x)\right) = \sum_{i=1}^N \left(\lambda_i^k - \mathbb{E}[\lambda_i^k]\right)$$

converges in law to a centered Gaussian variable. Since in Part III we shall give a complete and detailed proof of the central limit theorem in the case of Gaussian entries with a weak interaction, we will be rather sketchy here. We refer to [7] for a complete and clear treatment and [6] for a simplified exposition of the full proof of the theorem we state below. To simplify, we assume here that $\mathbf{A}^N$ is a Wigner matrix with

$$A_{ij}^N = \frac{B_{ij}}{\sqrt{N}},$$

A. Guionnet, *Large Random Matrices: Lectures on Macroscopic Asymptotics*, Lecture Notes in Mathematics 1957, DOI: 10.1007/978-3-540-69897-5_2,

where $(B_{ij}, 1 \leq i \leq j \leq N)$ are independent real equidistributed random variables. Their marginal distribution μ has all moments finite (in particular (1.7) is satisfied) and satisfies

$$\int x d\mu(x) = 0 \text{ and } \int x^2 d\mu(x) = 1.$$

We shall show why the following statement holds.

Theorem 2.1. *Let*

$$\sigma_k^2 = k^2 [C_{\frac{k-1}{2}}]^2 + \frac{k^2}{2} [C_{\frac{k}{2}}]^2 \left[\int x^4 d\mu(x) - 1 \right] + \sum_{r=3}^{\infty} \frac{2k^2}{r} \left(\sum_{\substack{k_i \geq 0 \\ 2\sum_{i=1}^r k_i = k-r}} \prod_{i=1}^r C_{k_i} \right)^2 ,$$

In this formula, C_x equals zero if x is not an integer and otherwise is equal to the Catalan number.

Then, M_k^N converges in moments to the centered Gaussian variable with variance σ_k^2, i.e., for all $l \in \mathbb{N}$,

$$\lim_{N \to \infty} \mathbb{E}\left[(M_k^N)^l \right] = \frac{1}{\sqrt{2\pi}\sigma_k} \int x^l e^{-\frac{x^2}{2\sigma_k^2}} dx.$$

Remark. Unlike the standard central limit theorem for independent variables, the variance here depends on $\mu(x^4)$.

Outline of the proof.

- *We first prove that the statement is true when $l = 2$.* (It is clearly true for $k = 1$ since A_k^N is centered.) We thus want to show

$$\sigma_k^2 = \lim_{N \to \infty} \mathbb{E}\left[(M_k^N)^2 \right]. \tag{2.1}$$

Below (1.9), we proved that $\mathbb{E}\left[(A_k^N)^2 \right]$ is bounded, uniformly in N. Furthermore, we can write

$$\mathbb{E}\left[(M_k^N)^2 \right] = \frac{1}{N^k} \sum_{\mathbf{i}, \mathbf{i}'} [P(\mathbf{i}, \mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]$$

where the sum over $\mathbf{i}, \mathbf{i}'$ will hold on graphs $\tilde{G}(\mathbf{i}, \mathbf{i}') = (\tilde{V}(\mathbf{i}, \mathbf{i}'), \tilde{E}(\mathbf{i}, \mathbf{i}'))$ so that

$$|\tilde{V}(\mathbf{i}, \mathbf{i}')| \leq k, \quad |\tilde{E}(\mathbf{i}, \mathbf{i}')| \leq k.$$

Since $[P(\mathbf{i}, \mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]$ is uniformly bounded, the only contributing graphs to the leading order will be those such that $|\tilde{V}(\mathbf{i}, \mathbf{i}')| = k$. Then, since we always have $|\tilde{V}(\mathbf{i}, \mathbf{i}')| \leq |\tilde{E}(\mathbf{i}, \mathbf{i}')| + 1$, we have two cases:

• $|\tilde{E}(\mathbf{i}, \mathbf{i}')| = k - 1$ in that case the skeleton $\tilde{G}(\mathbf{i}, \mathbf{i}')$ will again be a tree but with one edge less than the total number possible; this means that one edge appears with multiplicity four and belongs to $\tilde{E}(\mathbf{i}) \cap \tilde{E}(\mathbf{i}')$, the other edges appearing with multiplicity 2. Hence, the graphs of $\tilde{E}(\mathbf{i})$ and $\tilde{E}(\mathbf{i}')$ are both trees (so that k must be even); there are $C_{\frac{k}{2}}^2$ such trees, and they are glued by a common edge, to choose among $\frac{k}{2}$ edges in each of the tree. Finally, there are two possible choices to glue the two trees according to the orientation. Thus, there are

$$2\left(\frac{k}{2}\right)^2 C_{\frac{k}{2}}^2 = \left(\frac{k^2}{2}\right) C_{\frac{k}{2}}^2$$

such graphs and then

$$P(\mathbf{i}, \mathbf{i}') - P(\mathbf{i})P(\mathbf{i}') = \int x^4 d\mu(x) - 1.$$

We hence obtain the contribution $(\frac{k^2}{2})C_{\frac{k}{2}}^2 (\int x^4 d\mu(x) - 1)$ to the variance.
• $|\tilde{E}(\mathbf{i}, \mathbf{i}')| = k$. In this case, the graph is no longer a tree and because $|\tilde{E}(\mathbf{i}, \mathbf{i}')| - |\tilde{V}(\mathbf{i}, \mathbf{i}')| = 1$, it contains exactly one cycle. This can be seen either by closer inspection of the arguments given after (1.1) or by using the formula that relates the genus of a graph and its number of vertices, faces and edges:

$$\sharp\text{vertices} + \sharp\text{faces} - \sharp\text{edges} = 2 - 2g \leq 2.$$

The faces are defined by following the boundary of the graph; each of these boundaries are exactly one cycle of the graph except one (since a graph has always one boundary) and therefore

$$\sharp\text{faces} = 1 + \sharp\text{cycles}.$$

So we get, for a connected graph with skeleton $(\tilde{V}, \tilde{E})$,

$$|\tilde{V}| \leq |\tilde{E}| + 1 - \sharp\text{cycles}. \tag{2.2}$$

In our case, $\sharp$vertices $=$ $\sharp$edges $= k$ and $\sharp$cycles ≥ 1 (since the graph is not a tree), so that the number of cycles must be exactly one. Counting the number of such graphs completes the proof of the convergence of $\mathbb{E}\left[(M_k^N)^2\right]$ to σ_k^2 (see [7] for more details).

- *Convergence to the Gaussian law.*
We next show that M_k^N is asymptotically Gaussian. This amounts to proving that $\lim_{N\to\infty} \mathbb{E}[(M_k^N)^{2l+1}] = 0$ whereas,

$$\lim_{N\to\infty} \mathbb{E}[(M_k^N)^{2l}] = \sharp\{\text{number of pair partitions of } 2l \text{ elements}\} \times \sigma_k^{2l}.$$

Again, we shall expand the expectation in terms of graphs and write for $l \in \mathbb{N}$,

$$\mathbb{E}[(M_k^N)^l] = \frac{1}{N^{\frac{kl}{2}}} \sum_{\mathbf{i}_1,\dots,\mathbf{i}_l} P(\mathbf{i}^1,\dots,\mathbf{i}^l)$$

with $P(\mathbf{i}^1,\dots,\mathbf{i}^l)$ given by

$$\mathbb{E}\Bigg[\Big(B_{i_1^1 i_2^1}\cdots B_{i_k^1 i_1^1} - \mathbb{E}[B_{i_1^1 i_2^1}\cdots B_{i_k^1 i_1^1}]\Big) \cdots \Big(B_{i_1^l i_2^l}\cdots B_{i_k^l i_1^l} - \mathbb{E}[B_{i_1^l i_2^l}\cdots B_{i_k^l i_1^l}]\Big)\Bigg].$$

We denote by $G(\mathbf{i}^1,\dots,\mathbf{i}^l) = (V(\mathbf{i}^1,\dots,\mathbf{i}^l), E(\mathbf{i}^1,\dots,\mathbf{i}^l))$ the corresponding graph; $V(\mathbf{i}^1,\dots,\mathbf{i}^l) = \{i_n^j, 1 \le j \le l, 1 \le n \le k\}$ and $E(\mathbf{i}^1,\dots,\mathbf{i}^l) = \{(i_n^j, i_{n+1}^j), 1 \le j \le l, 1 \le n \le k\}$ with the convention $i_{l+1}^j = i_1^j$. As before, $P(\mathbf{i}^1,\dots,\mathbf{i}^l)$ equals zero unless each edge appears with multiplicity 2 at least. Also, because of the centering, it vanishes if there exists a $j \in \{1,\dots,l\}$ so that $E(\mathbf{i}^1,\dots,\mathbf{i}^l) \cap E(\mathbf{i}^j)$ does not intersect $E(\mathbf{i}^1,\dots,\mathbf{i}^{j-1},\mathbf{i}^{j+1},\dots,\mathbf{i}^l)$. Let us decompose $G(\mathbf{i}^1,\dots,\mathbf{i}^l)$ into its connected components $(G_1,\dots,G_c)$. We claim that

$$|V(\mathbf{i}^1,\dots,\mathbf{i}^l)| \le c - l + \left[\frac{l(k+1)}{2}\right]. \tag{2.3}$$

This type of bound is rather intuitive; if a connected component G_i contains $G(\mathbf{i}^{j_1}),\dots,G(\mathbf{i}^{j_p})$, each gluing of the $G(\mathbf{i}^{j_l})$ should create either a cycle or an edge with multiplicity 4, the total number of vertices decreasing at least by one in each gluing. Hence, $|V(\mathbf{i}^1,\dots,\mathbf{i}^l)|$ should grow linearly with the number of connected components. The proof is given in Appendix 20.3 for completeness (see [6] or [7]). With (2.3), we conclude that the only indices that will contribute are such that

$$c - l + \left[\frac{l(k+1)}{2}\right] \ge \frac{kl}{2}$$

with $c \le [\frac{l}{2}]$. This implies that

$$\frac{kl}{2} \le \left[\frac{l}{2}\right] - l + \left[\frac{l(k+1)}{2}\right] \le \frac{l}{2} - l + \frac{l(k+1)}{2} = \frac{kl}{2}$$

resulting in all inequalities being equalities. Thus, to get a first-order contribution we must have l even and $c = \frac{l}{2}$. In that case, we write $(s_j, r_j)_{1\le j\le l}$ the pairing so that $(G(\mathbf{i}_{s_j}), G(\mathbf{i}_{r_j}))_{1\le j\le l}$ are connected for all $1 \le j \le l$ (with the convention $s_j < r_j$). By independence of the entries, we have

$$P(\mathbf{i}_1,\dots,\mathbf{i}_{2l}) = \prod_{j=1}^{l} P(\mathbf{i}_{s_j},\mathbf{i}_{r_j})$$

and so we have proved that

$$N^{-kl} \sum_{\mathbf{i}_1,\ldots,\mathbf{i}_{2l}} P(\mathbf{i}_1,\ldots,\mathbf{i}_{2l}) = \sum_{\substack{s_1<\cdots<s_l \\ r_j>s_j}} \left(N^{-k} \sum_{\mathbf{i}_1,\mathbf{i}_2} P(\mathbf{i}_1,\mathbf{i}_2) \right)^l + o(1)$$
$$= \sigma_k^{2l} \sum_{\substack{s_1<\cdots<s_l \\ r_j>s_j}} 1 + o(1)$$

which proves the claim since

$$\frac{1}{\sqrt{2\pi}} \int x^{2l} e^{-\frac{x^2}{2}} dx = \sum_{\substack{s_1<\cdots<s_l \\ r_j>s_j}} 1 = (2l-1)(2l-3)(2l-5)\cdots 1.$$

This completes the proof of the moments convergence.

$\square$

Exercise 2.2. *Show that Theorem 2.1 implies that M_k^N converges weakly to the centered Gaussian variable with variance σ_k^2. Hint: control tails to approximate bounded continuous functions by polynomials.*

Bibliographical Notes. Johansson [120] proved a rather general central limit theorem for the spectral measure of Gaussian random matrices (and more generally for particles interacting via a Coulomb gas potential). It was generalized to β-ensembles and Laguerre ensembles in [82] by using tri-diagonal representation of the classical ensembles [81]. The strategy of moments developed here follows an article of Anderson and Zeitouni [7] (see a generalization in [177]). Central limit theorems were also obtained in the case of Ginibre ensembles (with spectral measure converging to the so-called circular law) in [169].

We shall see in Part III that this kind of theorem generalizes to the multi-matrix setting that we shall introduce in the next chapter.

2.2 Estimates of the Largest Eigenvalue of Wigner Matrices

In this section, we derive estimates on the largest eigenvalue of a Wigner matrix with real entries $A_{ij}^N = N^{-\frac{1}{2}} B_{ij}$ with $(B_{ij}, 1 \le i \le j \le N)$ independent equidistributed centered random variables with marginal distribution P. The idea is to improve the moments estimates of the previous chapter.

We shall assume that P is a symmetric law (see the recent article [166] for a relaxation of this hypothesis):

$$P(-x \in .) = P(x \in .).$$

We take the normalization $E[x^2] = 1$. Further, we assume that P has sub-Gaussian tail, i.e., that there exists a finite constant c such that for all $k \in \mathbb{N}$,

$$E[x^{2k}] \leq (ck)^k.$$

We follow the article of S. Sinaï and A. Soshnikov [179] to prove the following result:

Theorem 2.3 (S. Sinaï–A. Soshnikov [179]). *For all $\epsilon > 0$, all $N \in \mathbb{N}$, there exists a finite function $o(s, N)$ such that $\lim_{N\to\infty} \sup_{N^\epsilon \leq s \leq N^{\frac{1}{2}-\epsilon}} o(s, N) = 0$ and*

$$\mathbb{E}[\mathrm{Tr}((A^N)^{2s})] = \frac{N2^{2s}}{\sqrt{\pi s^3}}(1 + o(s, N)). \tag{2.4}$$

As a consequence, for all $\epsilon > 0$, if we let $\lambda_{max}(A^N)$ denote the spectral radius of A^N,

$$\lim_{N\to\infty} P(|\lambda_{\max}(A^N) - 2| \geq \epsilon) = 0.$$

A previous result of the same nature (but under weaker hypothesis (the symmetry hypothesis of the distribution of the entries being removed) under which the moments estimate (2.4) holds for a smaller range of s) was proved by Komlós and Füredi [93]. A later result of Soshnikov [180] improves the range of s under which (2.4) holds to s of order less than $n^{\frac{2}{3}}$, a result that captures the fluctuations of $\lambda_{\max}(A^N)$. We emphasize here that the proof below heavily depends on the assumption that the distribution of the entries is symmetric.

Proof. Let us first derive the convergence in probability from the moment estimates. First, note that

$$P(\lambda_{\max}(A^N) \leq 2 - \epsilon) \leq P\left(\int f(x) dL_{A^N} = 0\right)$$

for all functions f supported on $]2-\epsilon, \infty[$. Taking f bounded continuous, null on $]-\infty, 2-\epsilon]$ and strictly positive in $[2-\frac{\epsilon}{2}, 2]$, we see that $P(\int f(x)dL_{A^N} = 0)$ goes to zero by Theorem 1.15. For the upper bound on $\lambda_{\max}(A^N)$, we shall use Chebychev's inequality and the moment estimates (2.4) as follows:

$$\begin{aligned} P(\lambda_{\max}(A^N) \geq 2 + \epsilon) &\leq \frac{1}{(2+\epsilon)^{2s}} \mathbb{E}[\lambda_{\max}(A^N)^{2s}] \leq \frac{1}{(2+\epsilon)^{2s}} \mathbb{E}[\mathrm{Tr}((A^N)^{2s})] \\ &\leq \frac{N2^{2s}}{(2+\epsilon)^{2s}\sqrt{\pi s^3}}(1 + o(s, N)) \end{aligned}$$

where the right-hand side goes to zero with N when $s = N^\epsilon$ for some $\epsilon > 0$.

To prove the moment estimates we shall again expand the moments and count contributing paths, in particular estimate more precisely contributions from paths that are not trees. Yet, the central point of the proof is to show that these paths give a negligible contribution. We follow the presentation of [179].

1. *Moments expansion.* As usual, we write

$$\mathbb{E}[\mathrm{Tr}\left((\mathbf{A}^N)^{2s}\right)] = \frac{1}{N^s} \sum_{i_0,\dots,i_{2s-1}=1}^{N} \mathbb{E}[B_{i_0 i_1} \cdots B_{i_{2s-1},i_0}]. \tag{2.5}$$

   We let E denote the set of edges of the graph, i.e., the undirected collection of couples $\{(i_p, i_{p+1}), p = 0, \dots, 2s-1\}$. Because we assumed the law of the B_{ij}'s symmetric, only indices such that each edge in E appears an even number of times will contribute. We call a *closed path* the sequence $P : i_0 \to i_1 \to \cdots \to i_{2s-1} \to i_0$. An *even path* is a closed path where each edge appears with even multiplicity; they are the only contributing paths.
2. *Descriptions of paths.* We will say that the ℓth step $i_{\ell-1} \to i_\ell$ of a path P is *marked* if during the first ℓ steps of P, the edge $\{i_{\ell-1}, i_\ell\}$ appears an odd number of times (note here that the ℓth step is counted, and so a step is marked iff the edge $\{i_{\ell-1}, i_\ell\}$ appears an even number of times in the previous steps, in particular if it does not appear). The step is *unmarked* otherwise. For even paths, the number of marked and unmarked edges is equal to s. The complete set of vertices $\mathcal{V}$ is the collection $\{1, \dots, N\}$ of all possible values of the points $(i_k, 0 \le k \le 2s-1)$. We say that a vertex $i \in \mathcal{V}$ belongs to the subset $\mathcal{N}_k = \mathcal{N}_k(P)$ if the number of times we arrive at i via marked edges equals k. Note that no vertex of the path except i_0 can belong to $\mathcal{N}_0$. Moreover, $\mathcal{N}_p = 0$ for $p > s$ (since there are at most s edges). Note that if we let $n_k = \sharp\mathcal{N}_k$, since $(\mathcal{N}_0, \dots, \mathcal{N}_s)$ is a partition of $\mathcal{V}$, $\sum_{k=0}^{s} n_k = N$. Moreover, $(\mathcal{N}_0, \dots, \mathcal{N}_s)$ also induces a partition of the edges and hence

$$\sum_{k=0}^{s} k n_k = s.$$

   We say that P is of type $(n_0, n_1, \dots, n_s)$ if $n_k = \sharp\mathcal{N}_k = \sharp\mathcal{N}_k(P)$ for all $k \in \{0, \dots, s\}$. We finally say that a path is a *simple even path* if $i_0 \in \mathcal{N}_0$ and P is of type $(N-s, s, 0, \dots, 0)$. Observe that in a simple even path, each edge appears only twice (since there are at most s different edges in P and here exactly s since there are s different vertices in $\mathcal{N}_1$). Also, we see that the graph corresponding to P has exactly s vertices in $\mathcal{N}_1$ plus $i_0 \in \mathcal{N}_0$ and so exactly $s+1$ vertices. Hence, the skeleton $(V, \tilde{E})$ of the graph drawn by P satisfies the relation $|V| = |\tilde{E}| + 1$ and hence is a tree. The strategy of the proof will be to show that simple even paths dominate the expectation when $s = o(\sqrt{N})$.
3. *Contribution of simple even paths.* Considering (2.5), we see that for simple even paths, $\mathbb{E}[B_{i_0 i_1} \cdots B_{i_{2s-1} i_0}] = 1$. Moreover, given a simple even path, we have N possible choices for i_0, $N-1$ for the first new vertex encountered when following P, $N-2$ for the second new vertex encountered,

etc. Since we have $C_s = (2s)!/s!(s+1)!$ simple even paths (see Property 1.10), we get the contribution

$$C_1^N = \frac{1}{N^s} N(N-1)\cdots(N-s)\frac{(2s)!}{s!(s+1)!} = \frac{2^{2s}N}{\sqrt{\pi s^3}}(1+o_1(s,N))$$

where we have used Stirling's formula and found

$$o_1(s,N) = -\frac{1}{N}\sum_{k=1}^{s} k + \frac{1}{s} \approx \frac{s^2}{2N} + \frac{1}{s}.$$

In the case where $i_0 \notin \mathcal{N}_0$ but $n_1 = s, n_2 = 0 \cdots, n_s = 0$, we must have $i_0 \in \mathcal{N}_1$. This means that we have one cycle and one different vertex less in the graph of an even path. Note that if we split the vertex i_0 into two vertices as in Figure 2.1, the new vertex being attached to the marked edge, then the old i_0 belongs to $\mathcal{N}_0$ and the new vertex to $\mathcal{N}_1$ and we are back to the case where $i_0 \in \mathcal{N}_0$.
There are s possibilities for the position of the marked edge incoming in i_0, but we are losing $N-s$ possibilities to choose a different vertex. Hence, the contribution to this term is bounded by

$$C_2^N \le \frac{s}{N-s} E[x^4] C_1^N$$

where the last term comes from the possibility that one edge attached to i_0 now has multiplicity 4.

4. *Contribution of paths that are not simple.* If a path is not as in the previous paragraph, there must be an $n_k \ge 1$ for $k \ge 2$. Let us count the number of these paths.
Given $n_0, n_1, \dots, n_s$, we have $\frac{N!}{n_0!n_1!\cdots n_s!}$ ways to choose the values of the vertices. Then, among the n_0 vertices in $\mathcal{N}_0$, we have at most n_0 ways to choose the vertex corresponding to i_0 (if $i_0 \in \mathcal{N}_0$).
Being given the values of the vertices, a path is uniquely described if we know the order of appearance of the vertices at the marked steps, the times when the marked steps occur and the choice of end points of the unmarked steps. The moments of time when marked steps occur can be coded by

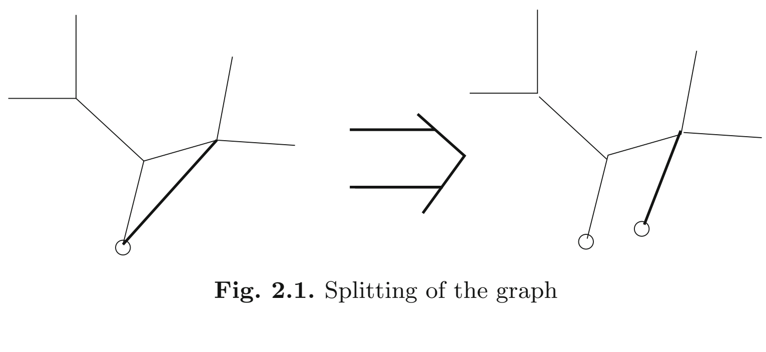

Fig. 2.1. Splitting of the graph

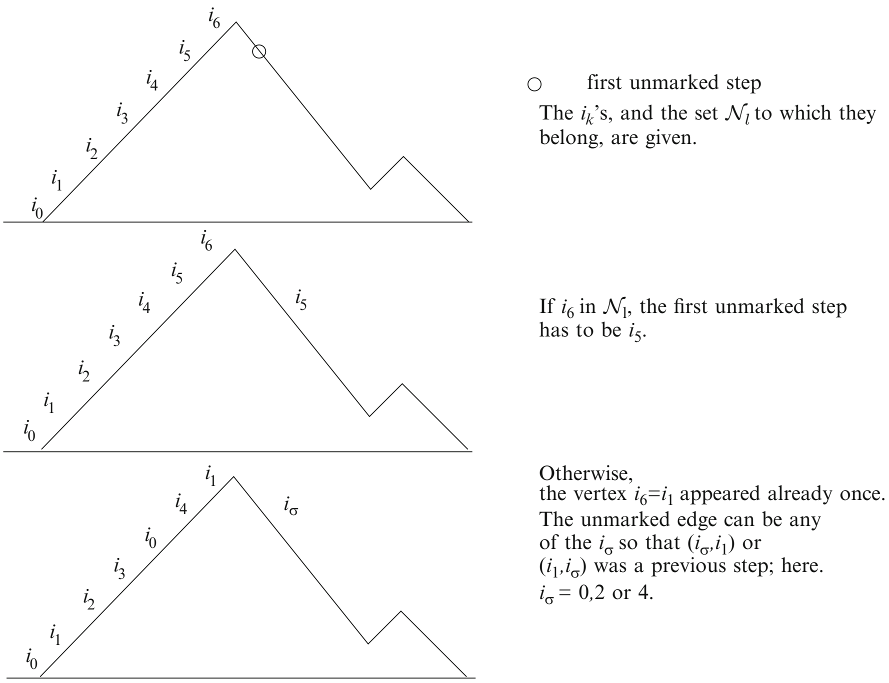

Fig. 2.2. Counting unmarked steps

a Dick path by adding $+1$ when the step is marked and -1 otherwise. Hence, there are $C_s = (2s)!/s!(s+1)!$ choices for the times of marked steps. Once we are given this path, we have s marked steps. The marked steps are partitioned into s sets corresponding to the $\mathcal{N}_k$, $1 \le k \le s$, with cardinality $n_k k$ each. Hence, we have $\frac{s!}{\prod_{k=1}^{s}(n_k k)!}$ possibilities to assign the sets into which the end points of the marked steps are. Finally, we have $(n_k k)!/(k!)^{n_k}$ ways to partition the set $\mathcal{N}_k$ into k copies of the same point of $\mathcal{N}_k$. So far, we have prescribed uniquely the marked steps and the set to which they belong.
To prescribe the unmarked steps, we still have an indeterminate. In fact, let us follow the Dick path of the marked steps till the first decreasing part corresponding to unmarked steps. Let i_ℓ be the vertex assigned to the last step. Then, if i_ℓ appeared only once in the past path (in the edge $(i_{\ell-1}, i_\ell)$), we have no choice and the next vertex in the path has to be $i_{\ell-1}$. This is the case in particular if $i_\ell \in \mathcal{N}_1$. If now $i_\ell \in \mathcal{N}_k$ for $k \ge 2$, the undirected step (i_p, i_ℓ) for some i_p may have occurred already at most $2k$ times (since it could occur either as a step (i_p, i_ℓ) or a step (i_ℓ, i_p), the later happening also less than k times since it requires that a marked step arrived at i_ℓ before). We have thus at most $2k$ choices now for the next vertex; one of the i_p among the at most $2k$ vertices such that the

step (i_p, i_ℓ) or (i_ℓ, i_p) were present in the past path. Once this choice has been made, we can proceed by induction since this choice comes with the prescription of the set $\mathcal{N}_l$ in which the vertex i_p belongs. Hence, since we have kn_k vertices in each set, we see that we have at most $\prod_{k=2}^s (2k)^{kn_k}$ choices for the end points of the unmarked steps.
Coming back to (2.5) we see that if the path is of type $(n_0, \ldots, n_s)$, entries appear at most n_k times with multiplicity $2k$ for $1 \le k \le s$. Thus Hölder's inequality gives

$$\mathbb{E}[B_{i_0 i_1} \cdots B_{i_{2s-1} i_0}] \le \prod_{k=1}^{s} \mathbb{E}[x^{2k}]^{n_k} \le \prod_{k=2}^{s} (ck)^{kn_k}$$

where we used that $\mathbb{E}[x^2] = 1$. This shows that the contribution of these paths can be bounded as follows.

$$\begin{aligned} E_{n_0,\ldots,n_s} &= \sum_{i_0,\cdots i_{2s-1}:P \text{ of type}(n_0,\ldots,n_s)} \mathbb{E}[B_{i_0 i_1} \cdots B_{i_{2s-1} i_0}] \\ &\le \frac{1}{N^s} n_0 \frac{N!}{n_0! n_1! \cdots n_s!} \frac{(2s)!}{s!(s+1)!} \frac{s!}{\prod_{k=1}^s (n_k k)!} \\ &\qquad \prod_{k=1}^{s} \frac{(n_k k)!}{(k!)^{n_k}} \prod_{k=2}^{s} (2k)^{kn_k} \prod_{k=2}^{s} (ck)^{kn_k} \\ &\le n_0 \frac{N(N-1)\cdots(n_0+1)}{N^s} \frac{(2s)!}{s!(s+1)!} \frac{1}{n_1! \cdots n_s!} \\ &\qquad \frac{s!}{\prod_{k=1}^s (ke^{-1})^{n_k k}} \prod_{k=1}^{s} (2ck^2)^{kn_k} \\ &\le N N^{N-n_0-s} \frac{(2s)!}{s!(s+1)!} \frac{s!}{n_1! \cdots n_s!} \prod_{k=2}^{s} (2cek)^{kn_k} \end{aligned}$$

where we have used that $(k!)^{n_k} \ge (ke^{-1})^{kn_k}$. Since $s = \sum_{k=1}^s kn_k$ and $N = \sum_k n_k$, we have $N - n_0 - s = \sum_{k=2}^s (1-k) n_k$. Using $s! \le (s)^s$, we obtain the bound

$$E_{n_0,\ldots,n_s} \le N \frac{(2s)!}{s!(s+1)!} \prod_{k=2}^{s} \frac{1}{n_k!} (N^{1-k} (2ceks)^k)^{n_k}.$$

We next sum over all $n_i \ge 0$ so that at least one $n_i \ge 1$ for $i \in \{2, \ldots, s\}$. This gives, with $\gamma_k := N^{1-k}(2ceks)^k$,

$$\sum_{n_0,\ldots,n_s:\max_{j\geq 2} n_j \geq 1} E_{n_0,\ldots,n_s} \leq N\frac{(2s)!}{s!(s+1)!}\sum_{k=2}^{s}(e^{\gamma_k}-1)\prod_{\ell\neq k} e^{\gamma_\ell}$$
$$\leq N\frac{(2s)!}{s!(s+1)!}e^{\sum_{\ell\geq 2}\gamma_\ell}(\sum_{\ell\geq 2}\gamma_\ell)$$

where we used that $e^x - 1 \leq xe^x$ for all $x \geq 0$. Note that in the range of s where $s^2 \leq N^{1-\epsilon}$, if we choose K big enough so that $K\epsilon \geq 1$,

$$\sum_\ell \gamma_\ell = \sum_{2\leq\ell\leq s} N^{1-\ell}(2ce\ell s)^\ell$$
$$\leq NK(2cesKN^{-1})^2 + N\sum_{K+1\leq\ell\leq s}(2ces^2N^{-1})^\ell$$
$$\leq \text{constant}(N^{-1}K^2s^2 + N(2ceN^{-\epsilon})^{K+1}) \leq \text{constant } N^{-\epsilon}$$

goes to zero as N goes to infinity. Thus, we conclude that

$$\sum_{n_0,\ldots,n_s} E_{n_0,\ldots,n_s} \leq CC_1^N N^{-\epsilon}.$$

Hence, in the regime s^2/N going to zero, the contribution of the indices $\{i_0,\ldots,i_{2s-1}\}$ associated with a path of type $(n_0,\ldots,n_s)$ with some $n_k \geq 1$ for some $k \geq 2$ is negligible compared to the contribution of simple even paths.

□

Exercise 2.4. *The extension of Theorem 2.3 to Hermitian Wigner matrices satisfying the same type of hypotheses is left to the reader as an exercise.*

Bibliographical Notes. Soshnikov [181] elaborated on his combinatorial estimation of moments to prove that the largest eigenvalue fluctuations follow the Tracy–Widom law, by estimating moments of order $N^{\frac{2}{3}}$ when the entries are symmetrically distributed and have sub-Gaussian tails. By approximation, Ruzmaikina [174] could weaken the later hypothesis to the case where the entries have only the eighteenth (thirty-sixth according to [9]) moment finite. The case where the entries are not symmetrically distributed is still mysterious, despite recent progress by Péché and Soshnikov [166] who prove the universality of moments of order much larger than $\sqrt{N}$ (but still much smaller than $N^{\frac{2}{3}}$). A rather different result was proved by Johansson [121]; he showed the universality of the fluctuations of the eigenvalues in the bulk for matrices whose entries are the convolution of a Gaussian law with a law with finite six moments. Similar results are expected to hold for the largest eigenvalues. It is well known [15] that the largest eigenvalue of a Wigner matrix converges to 2 if and only if the entries have fourth moments. It is expected that the fluctuations follow the Tracy–Widom law when the fourth moment is finite. What happens when the entries have less finite moments is described in [9, 184]. Also, the case where one adds a finite rank perturbation to the matrix was studied in [16]; if the perturbation is sufficiently small the fluctuations still follows the Tracy–Widom law, whereas if it is large, they will be Gaussian.

Other classical ensembles were studied; for instance Wishart matrices [14, 27, 183, 190].

In the next chapter, we shall consider polynomials in several random matrices; it was shown in [111] that the spectral radius of polynomials in several independent matrices following the GUE converge to the expected limit (that is the edge of the support of the limiting spectral measure of this polynomial). This was generalized to the case of matrices interacting via a convex potential in [106].

Springer

http://www.springer.com/978-3-540-69896-8

Large Random Matrices: Lectures on Macroscopic Asymptotics
École d'Été de Probabilités de Saint-Flour XXXVI – 2006
Guionnet, A.
2009, XII, 294 p., Softcover
ISBN: 978-3-540-69896-8