CENTRAL LIMIT THEOREM FOR TRACES OF LARGE RANDOM SYMMETRIC MATRICES WITH INDEPENDENT MATRIX ELEMENTS

Ya. Sinai[1] and A. Soshnikov[2]

Dedicated to the memory of R. Mane

Abstract

We study Wigner ensembles of symmetric random matrices $A = (a_{ij})$ $i, j = 1, \dots, n$ with matrix elements $a_{ij}, \quad i \le j$ being independent symmetrically distributed random variables

$$a_{ji} = \frac{\xi_{ij}}{n^{\frac{1}{2}}} .$$

We assume that $\mathrm{Var}\, \xi_{ij} = \frac{1}{4}$, for $i < j$, $\mathrm{Var}\, \xi_{ii} \leqslant$const and that all higher moments of ξ_{ij} also exist and grow not faster than the Gaussian ones. Under formulated conditions we prove the central limit theorem for the traces of powers of A growing with n more slowly than $\sqrt{n}$. The limit of $\mathrm{Var}(\mathrm{Trace}\ A^p)$, $1 \ll p \ll \sqrt{n}$, does not depend on the fourth and higher moments of ξ_{ij} and the rate of growth of p, and equals to $\frac{1}{\pi}$.

As a corollary we improve the estimates on the rate of convergence of the maximal eigenvalue to 1 and prove central limit theorem for a general class of linear statistics of the spectra.

1. Introduction and formulation of the results.

We revisit the classical ensemble of random matrices introduced by E. Wigner in the fifties ([1], [2]): the components $a_{ij} = a_{ji} = \frac{\xi_{ij}}{\sqrt{n}}$ of the real symmetric $n \times n$ matrices A are such that:

[1]Mathematics Department, Princeton University, Princeton, NJ 08544–1000 USA, and Landau Institute of Theoretical Physics, Moscow, Russia, *sinai@math.princeton.edu*

[2]Institute for Advanced Study, Olden Lane, Princeton, NJ 08540, USA

current address: Caifornia Institute of Technology, Sloan 253-37, Pasadena, CA 91125, USA

(i) $\{\xi_{ij}\}_{1\le i\le j\le n}$ are independent random variables;

(ii) the laws of distribution for ξ_{ij} are symmetric;

(iii) each moment $E\xi_{ij}^p$ exists and $E|\xi_{ij}^p| \le C_p$, C_p is a constant depending only on p; (ii) implies that all odd moments of ξ_{ij} vanish;

(iv) the second moments of $\xi_{ij}, i < j$, are equal $\frac{1}{4}$; for $i = j$ they are uniformly bounded.

Studying the empirical distribution function $F_n(\lambda) = \frac{1}{n}\#\{\lambda_i \le \lambda, \quad i = 1,\dots,n\}$, of the eigenvalues of A, Wigner (see [1], [2]) proved the convergence of moments of $F_n(\lambda)$

$$\langle\lambda^p\rangle_n = \int\limits_{-\infty}^{\infty} \lambda^p dF_n(\lambda) = \frac{1}{n}\sum_{i=1}^{n} \lambda_i^p = \frac{1}{n} \cdot \text{Trace}\, A^p$$

to the moments of nonrandom distribution function

$$F(\lambda) = \begin{cases} 1\,, & \lambda \ge 1 \\ \dfrac{2}{\pi} \displaystyle\int\limits_{-1}^{\lambda} \sqrt{1-x^2}\, dx\,, & -1 \le \lambda \le 1 \\ 0\,, & \lambda \le -1 \end{cases}$$

in probability, i.e.

$$\frac{1}{n}\ \text{Trace}\ A^p \xrightarrow[n\to\infty]{Pr} \quad \mu_p = \begin{cases} \frac{(2s)!}{s!\ (s+1)} \cdot \frac{1}{4^s}\,, & \text{if } p = 2s, \\ 0\,, & \text{if } p = 2s+1\,. \end{cases} \tag{1.1}$$

Later, under more general conditions, the convergence in (1.1) was proven to be with probability 1 (see [3] – [6]). This statement is sometimes called the semicircle Wigner law.

The proof by Wigner resembles the method of moments in the theory of sums of independent random variables. In the late sixties and early seventies Marchenko and Pastur proposed a more powerful technique based on the analysis of matrix elements of resolvents $(A - z \cdot Id)^{-1}$, which allowed them to generalize Wigner's results to the case of Lindeberg–Feller type random variables: for any $c > 0$

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{1\le i\le j\le n} E(\xi_{ij}^2 \cdot \chi(|\xi_{ij}| > c\sqrt{n})) = 0$$

([6] – [9], see also [10] – [12]). Similar results for random band matrices have been obtained in [13].

Due to strong correlations between eigenvalues, the fluctuations $\frac{1}{n}\sum_{i=1}^{n}\lambda_i^p - \mu_p$ are of order $\frac{1}{n}$ (not $\frac{1}{\sqrt{n}}$!), and Trace $A^p - E(\text{Trace } A^p)$ converges in distribution to the normal law $\mathcal{N}(\ 0, \sigma_p)$, ($p$ is fixed), where the variance σ_p depends on the second and fourth moments of ξ_{ij} ([6]). The purpose of this paper is to extend these results to the case of powers of A growing with n.

Main Theorem. *Consider Wigner ensemble of symmetric random matrices* (i) – (iv) *with the additional assumption*

$$E\,\xi_{ij}^{2\kappa} \le (const\,\kappa)^{\kappa} \quad , \qquad const > 0 \tag{1.2}$$

uniformly in i, j *and* κ, *meaning that the moments of* ξ_{ij} *grow not faster than the Gaussian ones. Then if* $1 << p << \sqrt{n}$

$$E\,(\text{Trace } A^p) = \begin{cases} \frac{1}{\sqrt{\pi}} \cdot \frac{n}{s^{3/2}} \cdot (1 + o(1)), & p = 2s \\ \\ 0 & , \quad p = 2s+1 \end{cases} \tag{1.3}$$

and Trace $A^p - E(\text{Trace } A^p)$ *converges in distribution to the normal law with mathematical expectation zero and variance* $\frac{1}{\pi}$. *Moreover, if* $[[e^t p]]$ *is defined as the nearest integer* p' *to* $e^t p$ *such that* $p' - p$ *is even, then the random process*

$$\eta_p(t) = \text{Trace } A^{[[e^t p]]} - E \ \text{Trace } A^{[[e^t p]]}$$

converges in the finite–dimensional distributions to the stationary random process $\eta(t)$ *with zero mean and covariance function*

$$E\,\eta(t_1)\cdot\eta(t_2) = \frac{1}{\pi\cosh(\frac{t_1 - t_2}{2})} \,. \tag{1.4}$$

Remark 1. It also follows from our results that if $p' - p$ is odd, p', p grow to infinity with n more slowly than $\sqrt{n}$, and $0 < \text{const}_1 < \frac{p'}{p} < \text{const}_2$, then the distributions of Trace $A^p - E(\text{Trace } A^p)$, Trace $A^{p'} - E(\text{Trace } A^{p'})$ are asymptotically independent. The reason for this can be best seen when p, p' are consecutive integers $2s, 2s+1$. Let us also, in addition to $p \ll n^{1/2}$, assume for simplicity $n^{2/5} \ll p$.

The main contribution to the Trace A^p comes from the eigenvalues at the $O(\frac{1}{p})$ distance from the endpoints of the Wigner semicircle distribution. If we consider rescaling

$$\lambda_j = 1 - \frac{x_j}{p} \quad , \qquad j = 1, 2 \ldots$$

for the positive eigenvalues, and

$$\lambda_i = -1 + \frac{y_i}{p} \quad , \qquad i = n, n-1, \ldots$$

for the negative ones, then

$$\text{Trace } A^p = \sum_j e^{-x_j} + \sum_i e^{-y_i} + o(1)\ .$$

We can analogously write

$$\text{Trace } A^{p'} = \sum_j e^{-x_j} - \sum_i e^{y_i} + o(1)\ .$$

Now the asymptotic independence of the distributions of the eigenvalues in the parts of the spectrum far apart from each other, and the identical distribution of the sums $\sum_j e^{-x_j}$, $\sum_i e^{-y_i}$ due to the central symmetry of the model imply

$$\text{Cov}(\text{Trace } A^p\ ,\ \text{Trace } A^{p'}) \underset{n\to\infty}{\to} 0\ .$$

Remark 2. The fact that the covariance function (1.4) of the limiting Gaussian process $\eta(t)$ in the Main Theorem does not depend on the fourth and higher moments of $\{\xi_{ij}\}$, supports the conjecture of the local universality of the distribution of eigenvalues in different ensembles of random matrices (see also [11], [12]).

We derive from the Main Theorem the central limit theorem for a more general class of linear statistics (see also [6] and [12], where central limit theorem was proven for the traces of the resolvent $(A - z \cdot Id)^{-1}$ under the condition $|Im\ z| > 1$).

Corollary 1. *Let $f(z)$ be an analytic function on the closed unit disk $|z| \leq 1$. Then $\sum_{i=1}^{n} f(\lambda_i) - E(\sum_{i=1}^{n} f(\lambda_i))$ converges in distribution to the Gaussian random variable $N(0, \sigma_f)$.*

Remark 3. In general, the limiting Gaussian distribution may be degenerate, i.e. $\sigma_f = 0$.

Another corollary concerns the rate of convergence of the maximal eigenvalue to 1. Under assumption of uniform boundedness of random variables ξ_{ij} (not necessarily symmetrically distributed), Z. Füredi and T. Komlóz proved in [14]) that with probability 1

$$\lambda_{\max}(A) = 1 + O\left(n^{-1/6} \log\ n\right)$$

Z. D. Bai and Y. Q. Yin showed in [15] the a.e. convergence of $\lambda_{\max}(A)$ to 1 assuming only the existence of the fourth moments of ξ_{ij}. The main ingredients of proofs of both results were the estimates of the mathematical expectations of the traces of high powers of A. In particular, Z. Füredi and T. Komlóz proved (1.3) for $p \ll n^{1/6}$.

Corollary 2. *Under the conditions of the Main Theorem*

$$\lambda_{\max}(A) = 1 + o(n^{-1/2} \log^{1+\epsilon} n)$$

for any $\epsilon > 0$ with probability 1.

Remark 4. C. Tracy and H. Widom proved recently (see [16]) that for the Gaussian Orthogonal Ensemble

$$\lambda_{\max}(A) = 1 + O(n^{-2/3}) \tag{1.5}$$

and calculated the limiting distribution function

$$G(x) = \lim_{n\to\infty} P\{\lambda_{\max} < 1 + \frac{x}{n^{2/3}}\}$$

which can be expressed in terms of Painleve II functions. One can expect the same kind of asymptotics (1.5) in the general case.

Remark 5. The technique used in this paper can be modified to extend our results to the case of not necessarily symmetrically distributed random variables $\xi_{ij},\ i \leqslant j$, with a less strict condition on the growth of higher moments

$$|E\xi_{ij}^k| \leqslant \ (\text{const}\, k)^k\ . \qquad 1.2'$$

Remark 6. The Main Theorem also holds for the Wigner ensemble of hermitian matrices

$$A = (a_{j\kappa}) j, \kappa = 1,\ \dots n\ ,$$
$$a_{jk} = \overline{a_{\kappa j}} = \frac{\xi_{j\kappa} + i \cdot \eta_{j\kappa}}{\sqrt{n}}\ , \quad \text{where } \xi_{j\kappa}, \eta_{j\kappa},\ 1 \le j \le \kappa \le n$$

are independent random variables, (property (iv) reads as Var $\xi_{j\kappa}$ + Var $\eta_{j\kappa} = \frac{1}{4}$); and for the ensemble of *covariance* matrices $A \cdot A^t$, where the entries of A are independent random variables satisfying the conditions of the Main Theorem.

The plan of the remaining part of our paper is the following. Sections 2, 3, and 4 are devoted to the proof of the Main Theorem. We evaluate E Trace A^p in §2, the variance Var Trace A^p in §3, and the moments of higher orders in §4. The combinatorial technique developed in §2 will be used throughout the sections 3 and 4 as well. We discuss corollaries of the Main Result and concluding remarks in section 5.

The authors are grateful to E. I. Dinabrug for many critical comments on the text. A. Soshnikov also would like to thank K. Johansson and B. Khoruzhenko for useful discussions. Ya. Sinai and A. Soshnikov thank National Science Foundation (grants DMS–9304580 and DMS–9706794) for the financial support.

2. Mathematical expectation of Trace A^p.

The main result of this section is the following theorem.

Theorem 1. *$E(\text{Trace } A^{2s}) = \frac{n}{\sqrt{\pi s^3}}(1+o(1))$ as $n \to \infty$ assuming $p \to \infty$ more slowly than $\sqrt{n}$.*

Since

$$E(\text{Trace } A^p) = \frac{1}{n^{p/2}} \sum_{i_0,i_1,\,\dots\,,i_{p-1=1}}^{n} E\,\xi_{i_0 i_1}\xi_{i_1 i_2}\,\cdot\dots\cdot\xi_{i_{p-1} i_0} \tag{2.1}$$

we will study in detail different types of closed paths $\mathcal{P}:\ i_0 \to i_1 \to \dots \to i_{p-1} \to i_0$ of length p (loops are allowed) on the set of vertices $\{1, 2, \dots, n\}$, and the contributions of the corresponding terms to (2.1).

We shall call edges of the path $\mathcal{P}$ such pairs (i,j) that $i, j \in \{1, 2, \dots, n\}$ and $\mathcal{P}$ has either the step $i \to j$ or the step $j \to i$. Mathematical expectation $E\xi_{i_0 i_1}, \cdots, \xi_{i_{p-1} i_0}$ is non–zero only if each edge in $\mathcal{P}$ appears even number of times. Such paths will be called even. Even paths exist only if p is even, therefore the mathematical expectation of the traces of odd powers of A equal zero. From now on we assume p to be even, $p = 2s$. The case we are interested in is $s = s(n),\ \ s(n) \to \infty,\ \ \frac{s(n)}{\sqrt{n}} \to 0$ as $n \to \infty$. We are going to show that similar to the case of fixed p the main contribution to $E(\text{Trace } A^p)$ comes from simple even paths (see the definition below). For this we shall introduce some classification of all closed paths by constructing for each $\mathcal{P}$ a partition of the set of all vertices $\{1, 2, \cdots, n\}$ onto subsets $\mathcal{N}_0, \mathcal{N}_1 \cdots \mathcal{N}_s$.

Definition 1. The ℓ^{th} step $i_{\ell-1} \to i_\ell,\ \ \ell = 1, \dots,\ p$ of the path $\mathcal{P}$ is called marked if during the first ℓ steps of $\mathcal{P}$ the edge $\{i_{\ell-1}, i_\ell\}$ appeared odd number of times.

The first step obviously is marked, and for even paths the number of marked steps is equal to the number of unmarked ones.

Definition 2. The vertex i, $\quad 1 \le i \le n$, belongs to the subset $\mathcal{N}_\kappa = \mathcal{N}_\kappa(\mathcal{P}),\ \ 0 \le k \le s$, if the number of times we arrived at i by marked steps equals to k. In other words, $i \in \mathcal{N}_\kappa$ if the vertex i was k times the right end of marked steps. All but one vertices from $\mathcal{N}_0$ do not belong to $\mathcal{P}$. The only possible exception can be the starting point i_0 of the path, if it was not visited by the path at the intermediate steps.

Put $n_k = \#(\mathcal{N}_\kappa)$. It follows easily from the definitions that

$$\sum_{k=0}^{s} n_k = n, \quad \sum_{k=0}^{s} k \cdot n_k = s\ .$$

We shall call $\mathcal{P}$ to be of the type $(n_0, n_1, \dots, n_s)$.

Definition 3. Any path of the type $(n-s,\ s,\ 0,\ \dots,\ 0)$ and such that $i_0 \in \mathcal{N}_o$ will be called simple even path.

For simple even paths each edge appears twice, once in one direction, and once in the other direction. However, not all paths with the last property are simple even paths.

We shall estimate the number of closed even paths of each type and their contribution to $E(\text{Trace } A^P)$. First of all, let us remark that the set of n vertices can be decomposed onto $(s+1)$ subsets $\mathcal{N}_0,\ \mathcal{N}_1, \ldots,\ \mathcal{N}_s$ with $n_0, n_1, \ldots,\ n_s$ elements by $\dfrac{n!}{n_0! n_1! \ \ldots \ n_s!}$ ways. If this partition is given and the path $\mathcal{P}$ is such that $i_0 \in \mathcal{N}_0$, then i_0 can be chosen by n_0 ways, and $\mathcal{N}_1 \cup \mathcal{N}_2 \cup \ldots, \cup \mathcal{N}_s$ is the set of remaining vertices of $\mathcal{P}$.

Different paths with the same partition $\{\mathcal{N}_0, \mathcal{N}_1, \ldots \mathcal{N}_s\}$, the same initial point and the same order of appearance of the vertices at the marked steps, differ by the choice of the moments of time when marked steps occur, and by the choice of end points of the unmarked steps. Since at any moment of time, the number of marked steps is greater or equal to the number of unmarked steps, we can code the choice of marked steps by a random walk of length $p = 2s$ on the set of non–negative integers which starts and ends at the origin. Namely, at the k^{th} step our random walk goes to the right if the k^{th} step is marked, and to the left otherwise. The number of such walks is $\frac{(2s)!}{s! \cdot (s+1)!}$ (see [17]). For simple even paths we have no freedom of choosing unmarked steps (we just go back), and if the partition $\mathcal{N}_0, \mathcal{N}_1, \cdots,\ N_s$, the initial point i_0 and the order of the appearance of the vertices are given, the constructed correspondence between paths and random walks is 1 to 1.

Let us first calculate the contribution to E Trace A^p from simple even paths. Suppose that a random walk, a partition $\{\mathcal{N}_0, \ldots,\ \mathcal{N}_s\}$ and initial point are chosen. Then the related number of paths is $\frac{s!}{\prod_{k=1}^{s} (\kappa!)^{n_k}}$. Indeed, the the number of marked steps is s, the number of the vertices of the type $\mathcal{N}_k$ is $k \cdot n_k$, then $\frac{s!}{\prod_{k=1}^{s} (k n_k)!}$ is the number of ways to decompose the set of s points onto s subsets of cardinality $k \cdot n_k$ each. As soon as each of these subsets is chosen the number of ways to write down each of n_k points k times is $\frac{(k n_k)!}{(k!)^{n_k}}$. Therefore, the total number of paths is $\frac{s!}{\prod_{k=1}^{s} (k!)^{n_k}}$.

For simple even paths $E\xi_{i_0 i_1} \cdot \ldots \cdot \xi_{i_p i_0} = \left(\frac{1}{4}\right)^s$ and the contribution of these paths to $E(\text{Trace } A^p)$ equals to

$$\frac{1}{n^s} \cdot \frac{n!}{(n-s)! \quad s!} \cdot (n-s) \cdot \frac{(2s)!}{s!(s+1)!} \cdot \frac{s!}{1!} \cdot \left(\frac{1}{4}\right)^s = \frac{n}{\sqrt{\pi s^3}}(1+o(1))\,. \tag{2.2}$$

If $n_1 = s, \quad n_2 = n_3 = \ldots = n_s = 0$ and $i_0 \notin \mathcal{N}_0$, we have a "double loop" when the second half of the path $\mathcal{P}$ repeats the first one. The expression for the contribution of this set of paths is different because of the difference in the number of positions of the initial point, which now belongs to $\mathcal{N}_1$, and by this reason is $\frac{n-s}{s}$ times smaller than (2.2) and can be neglected.

We shall show that the contribution to (2.1) of all terms with $n_1 < s$ is also smaller compared with (2.2).

If $n_1 < s$, then some of $n_k, k \geq 2$ are non–zero and the choice of the end points at the unmarked steps from the vertices of type $\mathcal{N}_k,\ k \geq 2$ may be non–unique. Namely, if the left end, which we denote by j, of the unmarked step belongs to $\mathcal{N}_k$. we have at most $2k$ possibilities for the right end of the step: it can be shown that we can count all the possibilities by assigning to each marked step arriving at j zero, one or two of them.

For the paths of $(n_0, n_1, \dots, n_s)$ type in view of condition (1.2)

$$\left|E\xi_{i_0 i_1} \dots \xi_{i_{p-1} i_0}\right| \le \prod_{k=1}^{s} (\text{const} \cdot k)^{k \cdot n_k} . \tag{2.3}$$

The subsum in (2.1) corresponding to $(n_0, n_1, \dots, n_s)$ can be estimated from above by

$$\frac{1}{n^s} \cdot \frac{n!}{n_o!\, n_1! \, \dots n_s!} \cdot n \cdot \frac{(2s)!}{s!\,(s+1)!} \cdot \frac{s!}{\prod_{k=2}^{s} (k!)^{n_k}}$$

$$\cdot \prod_{k=2}^{s} (2k)^{k \cdot n_k} \cdot \prod_{k=2}^{s} (\text{const } k)^{k\, n_k} \le n \cdot \frac{(2s)!}{s!\,(s+1)!} \cdot \frac{1}{4^s} \tag{2.4}$$

$$\cdot \left[\frac{n(n-1)\dots(n-n_0+1)}{n^s} \cdot \frac{1}{n_1! \dots n_s!} \cdot \frac{s!}{\prod_{k=2}^{s} (k\, e^{-1})^{k\, n_k}} \cdot \prod_{k=2}^{s} (2 \text{ const } k^2)^{k \cdot n_k} 4^s \right] .$$

In view of the inequality $s! \le n_1! \cdot s^{s-n_1}$, and $\sum_{k=1}^{s} k\, n_k = s$, $\sum_{k=1}^{s} n_k = n - n_0$, the r.h.s. of (2.4) is not greater than

$$n \cdot \frac{(2s)!}{s!\,(s+1)!} \cdot \frac{1}{4^s} \cdot \left(n^{-\sum_{k=2}^{s} (k-1) \cdot n_k} \cdot \frac{1}{n_1! \quad n_s!} \cdot s! \prod_{k=1}^{s} (8e \text{ const} \cdot k)^{k \cdot n_k} \right)$$

$$\le n \frac{(2s)!}{s!\,(s+1)!} \cdot \frac{1}{4^s} \left(\prod_{k=2}^{s} \frac{1}{n_k!} \left[\frac{(8e \cdot \text{const} \cdot k \cdot s)^k}{n^{k-1}} \right]^{n_k} \right) .$$

The sum of the last expression over all non–negative integers $n_2, n_3, \dots, n_k$ such that

$$0 < \sum_{k=2}^{s} k \cdot n_k \le s$$

is not greater than

$$n \cdot \frac{(2s)!}{s!\,(s+1)!} \cdot \left(\frac{1}{4}\right)^s \cdot \left(\exp\left(\sum_{k=2}^{s} \frac{(8e \cdot \text{const} \cdot k \cdot s}{n^{k-1})^k} \right) - 1 \right) . \tag{2.5}$$

Since for $s \ll n^{1/2}$

$$\sum_{k=2}^{s} \frac{(8e \text{ const } k \cdot s)^k}{n^{k-1}} = O\left(\frac{s^2}{n}\right) = o(1) ,$$

(2.5) is small compared with (2.2). Theorem 1 is proven. As one can see the core of the proof of Theorem 1 is the following Proposition which will be useful in the next sections as well.

Proposition 1. *The main contribution to the number of all even paths of length* p *on the set of* n *vertices* $\{1, 2, \ldots, n\}$ *where* $p = o(n^{1/2})$ *as* $n \to \infty$ *is given by simple even paths, i. e.*

$$\frac{\#_{n,p}\ (simple\ even\ paths)}{\#_{n,p}\ (even\ paths)} \underset{n\to\infty}{\longrightarrow} 1\ .$$

3. Variance of Trace A^p.

The main result of this section is the following theorem.

Theorem 2. *Let* $p = o(\sqrt{n})$, *Then* $\mathrm{Var}(\mathrm{Trace}\ A^p) \leq$ *const for all* n *and* $\mathrm{Var}(\mathrm{Trace}\ A^p) \to \frac{1}{\pi}$ *as* $n \to \infty$, $p \to \infty$, $\frac{p}{\sqrt{n}} \to 0$.

The formula for the variance of Trace A^p has the form

$$\mathrm{Var}\,(\mathrm{Trace}\ A^p) = E(\mathrm{Trace}\ A^p)^2 - (E\ \mathrm{Trace}\ A^p)^2 = \sum_{i_o, i_1, \ldots, i_{p-1}=1}^{n} \ \sum_{j_0, j_1, \ldots, j_{p-1}=1}^{n} \frac{1}{n^p}$$

$$\cdot \left(E \prod_{\ell=1}^{p} \xi_{i_{\ell-1} i_\ell} \cdot \prod_{m=1}^{p} \xi_{j_{m-1} j_m} - E \prod_{\ell=1}^{p} \xi_{i_{\ell-1} i_\ell} \cdot E \prod_{m=1}^{p} \xi_{j_{m-1} j_m} \right) \qquad (3.1)$$

where we use the convention $i_p = i_0, j_p = j_0$. The only non-zero terms in (3.1) come from pairs of paths

$$P\{i_o \to i_1 \to \ldots \to i_{p-1} \to i_0\}\ , \mathcal{P}' = \{j_0 \to j_1 \to \ldots \to j_{p-1} \to j_0\}$$

such that the following conditions are satisfied:

(a) $\mathcal{P}$ and $\mathcal{P}'$ have at least one edge in common.
(b) Each edge appears in the union of $\mathcal{P}$ and $\mathcal{P}'$ an even number of times.

Definition 4. Any pair of paths satisfying (a) and (b) is called correlated. A correlated pair is called simply correlated if each edge appears in the union of $\mathcal{P}$ and $\mathcal{P}'$ only twice.

Proposition 2. *Let* $1 << p << \sqrt{n}$. *Then in the main order the number of correlated pairs equals to the number of simply correlated pairs and is*

$$\frac{1}{\pi} \cdot n^p \cdot 2^{2p} \cdot (1 + o(1))\ . \qquad (3.2)$$

Remark. A slight modification of the proof, which takes into account the weights

$$E\left(\prod_{\ell=1}^{p} \xi_{i_{\ell-1} i_\ell} \cdot \prod_{m=1}^{p} \xi_{j_{m-1} j_m}\right) - E\left(\prod_{\ell=1}^{p} \xi_{i_{\ell-1} i_\ell}\right) E\left(\prod_{j=1}^{m} \xi_{j_{m-1} j_m}\right)$$

ascribed to the correlated paths gives

$$\mathrm{Var}\,(\mathrm{Trace}\ A^p) \underset{n\to\infty}{\longrightarrow} \frac{1}{\pi}. \tag{3.3}$$

Proof of Proposition 2. To calculate the number of correlated pairs, we construct for each such a pair an even path of the length $2p-2$. The corresponding map of correlated pairs to such paths will not be 1 to 1, and in general an even path of length $2p-2$ has many preimages. We will study the number of preimages in more detail, and then count even paths of length $2p-2$ (with corresponding multiplicities) in a way, similar to the one used in section 2 for calculating E Trace A^p.

Let us now construct the map. Consider the ordered pair of correlated paths $\mathcal{P}$, $\mathcal{P}'$ (see Figure 1). Each edge appears in the union of $\mathcal{P}$ and $\mathcal{P}'$ an even number of times. In Fig. 1 we have shown this for simplicity only for the joint edge, which we define below:

FIGURE 1

Definition 5. The first edge along $\mathcal{P}$ which also belongs to $\mathcal{P}'$ is called joint edge of the (ordered) correlated pair $\mathcal{P}, \mathcal{P}'$.

The new even path of length $2p-2$ is constructed in the following way (see Fig. 2).

FIGURE 2

The new path of length $2p-2$ is constructed in the following way (see Fig. 2). We begin walking along the first path until we reach the left point of the joint edge, then switch to the second path and make other $p-1$ steps if the directions of $\mathcal{P}$ and $\mathcal{P}'$ along the joint edge are opposite. If the directions coincide, we walk along $\mathcal{P}'$ in the inverse direction. After $p-1$ steps along $\mathcal{P}'$, we will arrive at the right point of the joint edge, then switch to the first path and go until the final point i_0. The new path

$$\ell_0 = i_0 \to \ell_1 = i_1 \to \ldots \to \ell_r = i_r = j_m \to \ldots \\ \to \ell_{r+p-1} = i_{r+1} = j \to \ldots \to \ell_{2p-3} = i_{p-1} \to \ell_0 = i_0$$

is even and is exactly what we need. Now assume that we have an even path $\ell_0 \to \ell_1 \to \ldots \to \ell_{2p-3} \to \ell_0$. We shall estimate in how many different ways it can be obtained from correlated pairs of paths of length p. To construct the pair of correlated paths, we have to choose some vertex ℓ_r in the first half of the path, $0 \le r \le p-1$, and connect it with the vertex ℓ_{r+p-1} (see Fig. 2). We also have to choose the starting point of the second path of the correlated pair and the direction of the path. This can be done in not more than $2p$ ways. The edge $\{l_r, l_{r+p-1}\}$ can be the joint edge of $\mathcal{P}$ and $\mathcal{P}'$ if it is the first along $\mathcal{P}$ common edge of $\mathcal{P}$ and $\mathcal{P}'$. This condition is easier to formulate in terms of a simple walk on the positive semi–axis which corresponds to the sequence of the marked steps of $\{\ell_0 \to \ell_1 \to \ldots \to \ell_{2p-3} \to \ell_0\}$. We recall that at the moment of time t random walk jumps to the right, if during the first t steps we met the edge $\{\ell_{t-1}, \ell_t\}$ an odd number of times; otherwise it jumps to the left; we start from the origin at the zeroth moment of time and return there at the final $(2p-2)$–th step.

The necessary condition for $\{\ell_r, \ell_{r+p-1}\}$ to be the joint edge of $\mathcal{P}$ and $\mathcal{P}'$ is that during the time interval $[r, r+p-1]$ (half of the whole walk), the trajectory does not descent lower than $x(r)$. It is also a sufficient condition in a typical situation, i.e. when $\{\ell_0 \to \ell_r \to \ldots \to \ell_{2p-3} \to \ell_0\}$ is a simple even path.

Lemma 1. *The sum over r, $0 \le r \le p-1$ of the number of walks of length $2p-2$ on* *the positive semi–axis such that*

$$\begin{aligned} x(t) \geq 0, t = 0, 1, \ \ldots \ , \ 2p-2; \\ x(t+1) - x(t) = \pm 1; \\ x(0) = x(2p-2) = 0 \end{aligned}$$

and

$$x(t) \geq x(r) \qquad \textit{if} \qquad r \leq t \leq r+p-1$$

is $2^{2p-2} \cdot \frac{2}{\pi} \cdot \frac{1}{p} \cdot (1 + o\,(1))$.

Remark. The probabilistic meaning of Lemma 1 is that the mathematical expectation of the number of moments of time t, for which

$$x(\tau) \geq x(t), \qquad t \leq \tau \leq t+p-1 \,,$$

is equal in the main order to $2 \cdot \sqrt{\frac{p}{\pi}}$.

Proof of Lemma 1. Fix some r, $0 \leq r \leq p-1$, and assume that a trajectory does not descend lower than $x(r)$ during the interval of time $r \leq t \leq r+p-1$. Let us first consider the non–degenerate case when $x(r) > 0$. Since $x(2p-2) = 0$, there exists a moment of time $r+p-1+\ell$, $0 \leq \ell \leq p-1-r$, such that

$$x(t) \geq x(r) \quad \text{for} \quad r \leq t \leq r+p-1+\ell$$

and

$$x(r+p-1+\ell+1) < x(r) = x(r+p-1+\ell) \,.$$

If we now fix ℓ and freeze the trajectory outside the interval of time $[r, \ r+p-1+\ell]$, then the number of trajectories satisfying the last two inequalities equals

$$2^{p-1+\ell} \cdot \frac{(p-1+\ell)\,!}{\left(\frac{p-1+\ell}{2}\right)!\left(\frac{p-1+\ell}{2}+1\right)!} = 2^{p-1+\ell} \frac{1}{\sqrt{\pi\left(\frac{p-1+\ell}{2}\right)^3}} \cdot (1+o(1)) \,. \tag{3.4}$$

From the other side, the number of trajectories on $[0, r] \cup [r+p-1+\ell, \quad 2p-2]$ such that $x(t) \geq 0$ for all moments of time, and $x(0) = x(2p-2) = 0$, $x(r) = x(r+p-1+\ell)$, $x(r+p+\ell) - x(r+p-1+\ell) = -1$ is

$$\begin{aligned} &\frac{1}{2} \cdot 2^{p-1-\ell} \cdot \frac{(p-1-\ell)!}{\left(\frac{p-1-\ell}{2}\right)!\left(\frac{p-1-\ell}{2}+1\right)!} \cdot (1+o(1)) \\ &\qquad = \frac{1}{2} \cdot 2^{p-1-\ell} \cdot \frac{1}{\sqrt{\pi \cdot \left(\frac{p-1-\ell}{2}\right)^3}} \cdot (1+o(1)) \,. \end{aligned} \tag{3.5}$$

Remark. (3.5) holds for typical ℓ, i.e. when $p-1-\ell \gg 1$; otherwise we can multiply r.h.s. of (3.5) by 2 and use it as an upper bound. It is easy to see that the contribution from ℓ such that $p-1-\ell \ll$ const is negligible. Multiplying (3.4) and (3.5) and making the summation over r, $0 \le r \le p-1$ and ℓ, $0 \le \ell \le p-1-r$ (depending on whether $p-1$ is even or odd, ℓ takes only even or odd values) we arrive at

$$\begin{aligned}
&\frac{1}{2} \cdot 2^{2p-2} \frac{1}{\pi} \sum_{r=1}^{p-1} \sum_{\ell=0}^{p-1-r} \frac{8}{(p-1+\ell)^{3/2} \cdot (p-1-\ell)^{3/2}} \cdot (1+o\,(1)) \\
&= 2^{2p} \cdot \frac{1}{\pi} \sum_{\ell=0}^{p-2} \frac{1}{(p-1+\ell)^{3/2}} \cdot \frac{1}{(p-1-\ell)^{1/2}} \cdot (1+o(1)) \\
&= 2^{2p} \cdot \frac{1}{\pi} \cdot \frac{1}{4(p-1)} \cdot \int_0^{1/2} \frac{1}{(1-y)^{3/2}} \cdot \frac{1}{y^{1/2}}\, dy \cdot (1+o(1)) \\
&= 2^{2p} \cdot \frac{1}{\pi} \cdot \frac{1}{2(p-1)} \cdot (1+o(1)) \,.
\end{aligned} \tag{3.6}$$

The case $x(r)=0$ can be considered in a similar way. The corresponding sum will be $O(2^{2p} \cdot p^{-3/2})$ and is negligible compared to (3.6).

Taking into account the factor $2p$ standing for the choice of direction and initial point of $\mathcal{P}'$, we get the statement of Lemma 1.

Now we are ready to conclude the proof of Proposition 2 and Theorem 2. Calculations similar to that from Theorem 1 prove that the main contribution to the number of correlated pairs and to Var(Trace A^p) is due to simple correlated pairs. In this case the weight

$$E \prod_{\ell=1}^{p} \xi_{i_{\ell-1} i_\ell} \cdot \prod_{m=1}^{p} \xi_{j_{m-1} j_m} - E \prod_{\ell=1}^{p} \xi_{i_{\ell-1} i_\ell} \cdot E \prod_{m=1}^{p} \xi_{j_{m-1} j_m}$$

equals to 2^{-2p} if $\{\ell_r,\ \ell_{r+p-1}\}$ is not the edge of $\{\ell_0 \to \ell_1 \to \ldots \to \ell_{2p-3} \to \ell_0\}$. Otherwise the weight is either

$$2^{-2p-4} \cdot E\, \xi^4_{i_r i_{r+1}} \qquad \textit{or} \qquad 2^{-2p-4} \cdot (E\, \xi^4_{i_r i_{r+1}} - (E\, \xi_{i_r i_{r+1}})^2) \,.$$

It is easy to see that the ratio of the number of simple even paths of length $2p-2$ that have an edge $\{\ell_r,\ \ell_{r+p-1}\}$ for some $0 \le r \le p-1$ to the whole number of simple even paths of length $2p-2$ tends to zero when $p \to \infty$, $n \to \infty$ which means that in the limit $n \to \infty$ the variance Var(Trace A^p) does not depend on the fourth moments of $\{\xi_{ij}\}$. It follows also from our calculations that

$$\lim_{n \to \infty} \mathrm{Var}(\mathrm{Trace}\ A^p) = \frac{1}{\pi}, 1 \ll p \ll n^{1/2} \,. \tag{3.7}$$

In a similar way one can show that

$$\lim_{n\to\infty} \mathrm{Cov}(\text{Trace } A^{[t_1p]}, \text{ Trace } A^{[t_2p]}) = \begin{cases} \frac{2\sqrt{t_1t_2}}{\pi(t_1+t_2)} & \text{if we choose } [t_1p]-[t_2p] \text{ even} \\ 0 & \text{if we choose } [t_1p]-[t_2p] \text{ odd} \end{cases} . \tag{3.8}$$

4. Higher moments of Trace A^p.

To finish the proof of the Central Limit Theorem for Trace A^p, we have to show that

$$E(\text{Trace } A^p - E\,\text{Trace } A^p)^{2k} = (2k-1)\,!! \cdot (\pi^{-k} + o(1)) \tag{4.1}$$

$$E(\text{Trace}\, A^p - E \text{ Trace } A^p)^{2k+1} = o\,(1)\,. \tag{4.2}$$

The main idea is rather straightforward: the analogue of (3.1) is

$$E(\text{Trace } A^p - E \text{ Trace } A^p)^L =$$
$$\frac{1}{n^{\frac{pL}{2}}} \cdot E \prod_{m=1}^{L} \left(\sum_{i_0^{(m)}, i_1^{(m)}, \ldots, i_{p-1}^{(m)}=1}^{n} \left(\prod_{r=1}^{p} \xi_{i_{r-1}^{(m)} i_r^{(m)}} - E \prod_{r=1}^{p} \xi_{i_{r-1}^{(m)} i_r^{(m)}} \right) \right) \tag{4.3}$$

as before (we use in (4.3) the convention $i_p^{(m)} = i_0^{(m)}$, $m = 1\,, \ldots\,, L$).

Consider the set of closed paths of the length p

$$\mathcal{P}_m = \{i_0^{(m)} \to i_1^{(m)} \to i_2^{(m)} \to \ldots \to i_{p-1}^{(m)} \to i_p^{(m)}\}, \quad m = 1\,, \ldots\,, L\,.$$

We shall call $\mathcal{P}_{m'}$ and $\mathcal{P}_{m''}$ connected, if they share some common edge.

Definition 6. A subset of paths $\mathcal{P}_{m_{j_1}}, \quad \mathcal{P}_{m_{j_2}}, \ldots, \mathcal{P}_{m_{j_\kappa}}$ is called a cluster of correlated paths if

1) for any pair $\mathcal{P}_{m_i}$, $\mathcal{P}_{m_j}$ from the subset one can find a chain of paths $\mathcal{P}_{m_s}$, also belonging to the subset, which starts with $\mathcal{P}_{m_i}$, ends with $\mathcal{P}_{m_j}$, and is such that any two neighboring paths are connected;
2) the subset $\mathcal{P}_{m_{j_1}}, \mathcal{P}_{m_{j_2}}, \ldots, \mathcal{P}_{m_{i_\kappa}}$ cannot be enlarged with the preservation of 1).

It is clear that the sets of edges corresponding to different clusters are disjoint. By this reason and because of the independence of ξ_{ij}, the mathematical expectation in (4.3) decomposes into the product of mathematical expectations corresponding to different clusters. We shall show that the main contribution to (4.3) comes from products where all clusters consist exactly of two paths.

Formulas (4.1), (4.2) essentially follow from Lemma 2.

Lemma 2.

$$E\frac{1}{n^{p\ell/2}}\prod_{m=1}^{\ell}{}^{\star}\left(\sum_{i_0^{(m)},\,\dots,\,i_{p-1}^{(m)}=1}^{n}\left(\prod_{r=1}^{p}\xi_{i_{r-1}^{(m)}i_r^{(m)}}-E\prod_{r=1}^{p}\xi_{i_{r-1}^{(m)}i_r^{(m)}}\right)\right) = \begin{cases}\frac{1}{\pi}+o(1) & , \ \textit{if}\ \ell=2\\ o(1) & , \ \textit{if}\ \ell>2\end{cases} \tag{4.4}$$

where the product $\prod^{\star}$ *in (* 4.4*) is taken over the paths which form a cluster.*

Proof of Lemma 2. The case $\ell=2$ was actually considered in §3. Assume now $\ell>2$. Similar to §3 we are going to construct for each correlated cluster $\{\mathcal{P}_1,\dots,\mathcal{P}_\ell\}$ an even path of the length behaving roughly as $\ell\cdot p$, estimate the multiplicity with which the newly constructed path appears, and apply the technique used in §2 and §3. As soon as the construction procedure is explained and the multiplicity problem is investigated, the last part of the proof is very much the same as in the previous sections.

THE CONSTRUCTION PROCEDURE.

Regular Step. Take the path $\mathcal{P}_1=\mathcal{P}_{\alpha_1}$ (i.e. we set $\alpha_1=1$) and find the first edge along $\mathcal{P}_1$ that also belongs to some other path, say $\mathcal{P}_{\alpha_2}$. Since $\{\mathcal{P}_1,\dots,\mathcal{P}_\ell\}$, form a cluster such edge always exists. We will denote the new path by $\{i,j\}$. Using $\mathcal{P}_{\alpha_1}$ and $\mathcal{P}_{\alpha_2}$ construct the new path of the length $2p-2$ as it was explained in §3 (see Fig. 1, 2). We will denote the new path by $\mathcal{P}_{\alpha_1}\vee\mathcal{P}_{\alpha_2}$. Now if $\mathcal{P}_{\alpha_1}\vee\mathcal{P}_{\alpha_2}$ and $(\ell-2)$ remaining paths still form a cluster, we perform again the Regular Step at the next stage of the construction procedure, and so on.

However it can happen in general, that the edge we just threw away was the only one that connected $\mathcal{P}_{\alpha_1}$, $\mathcal{P}_{\alpha_2}$ with the rest of the cluster and $\mathcal{P}_{\alpha_1}\vee\mathcal{P}_{\alpha_2}$ does not have this edge anymore. In particular, it implies that $\{i,j\}$ appears both in each $\mathcal{P}_{\alpha_1}$ and $\mathcal{P}_{\alpha_2}$ only once. In this case we have to modify the construction procedure.

Modified step. Consider all paths that contain the edge $\{i,j\}$ and denote them $\mathcal{P}_{\alpha_1}$, $\mathcal{P}_{\alpha_2},\dots,\mathcal{P}_{\alpha_{\ell_1}}$. We remark that the number of such paths ℓ_1 must be greater than 2. If ℓ_1 is even, we shall construct the new path of length $\ell_1\cdot p-\ell_1$ in the following way: first we go along $\mathcal{P}_{\alpha_1}=\mathcal{P}_1$ until we reach vertex i of the edge $\{i,j\}$, then we switch to the path $\mathcal{P}_{\alpha_2}$ and make (in the appropriate direction) $(p-1)$ steps until we reach vertex j, then we switch to $\mathcal{P}_{\alpha_3}$ and make other $(p-1)$ steps until we reach i, then we switch to $\mathcal{P}_{\alpha_4}$, and so on. Finally, we return to $\mathcal{P}_{\alpha_1}$ at the vertex j and finish the walk along $\mathcal{P}_{\alpha_1}$. We shall denote the new path by $\mathcal{P}_{\alpha_1}\vee\mathcal{P}_{\alpha_2}\vee\dots\vee\mathcal{P}\alpha_{\ell_1}$. This path and the remaining $\ell-\ell_1$ paths again form a cluster, and we can apply Regular, or Modified Step at the next stage.

Consider now the case of odd ℓ_1. Since $\{i,j\}$ appears in the union of all paths an even number of times, at least one of the paths $\mathcal{P}_{\alpha_{\ell_1}}, \dots, \mathcal{P}_{\alpha_{\ell_1}}$ contains $\{i,j\}$ two times or more. We will denote such path by $\mathcal{P}_{\alpha_2}$ (if there are several, take one with the minimal subindex). It should also be noted that $\mathcal{P}_1$ contains $\{i,j\}$ only once. Now $\mathcal{P}_{\alpha_1} \vee \mathcal{P}_2$ still contains $\{i,j\}$, and $\mathcal{P}_{\alpha_1} \vee \mathcal{P}_2$ and $(\ell - 1)$ remaining paths form a cluster. This means that we are allowed to apply the Regular Step at the next stage; but if we wish to finish the part of the construction procedure corresponding to $\{i,j\}$ faster, we can apply the Modified Step to the paths containing $\{i,j\}$ once again. (Now we have an even number of such paths.)

As a final result of so defined procedure, we will get an even path of length $(\ell p - q)$, which we denote by $\mathcal{P}_{\alpha_1} \vee \mathcal{P}_{\alpha_2} \vee \dots \vee \mathcal{P}_{\alpha\ell}$ with $\ell \leq q \leq 2\ell$ and $q = 2\ell_1 + 2\ell_2 + \dots + 2\ell_r + 2g$, where g is the number of times we used Regular Step, r is the number of times we used Modified Step; first time "combining" $2\ell_1$ paths, second time combining $2\ell_2$ paths, and so on.

Now we come to the question in how many ways we can get an even path $\mathcal{P}$ of length $\ell p - q$ from the correlated cluster of ℓ paths each of the length p. We need only a rough estimate of this multiplicity from above. First of all we have to choose

- the order in which we are taking $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ to construct $\mathcal{P}_{\alpha_1} \vee \dots \vee \mathcal{P}_{\alpha_\ell}$ (since $\mathcal{P}_{\alpha_1}$ is always equal to $\mathcal{P}_1$ we can do it in $(\ell - 1)!$ ways);
- moments of time when we use Regular Steps and Modified Steps (the whole number of steps is not greater than ℓ so we can do it in no more than 2^ℓ ways);
- how many paths we combine together on each of the Modified Steps (ℓ^ℓ is a trivial estimate).

What is really important here is that the number of all such choices is bounded by some constant depending only on ℓ.

Now we have to estimate the number of ways in which we can reconstruct from $\mathcal{P}$ the paths $\mathcal{P}_{\alpha_\ell}, \mathcal{P}_{\alpha_{\ell-1}}, \dots, \mathcal{P}_{\alpha_2}$. We will write down the trivial upper bound for the number of choices of the edges to be thrown away at all stages of the construction procedure except the first one: we can choose the left point of each of these edges in at most $(\ell \cdot p)$ ways (we will recall that the right point automatically lies $(p-1)$ steps further along $\mathcal{P}$), and we can choose the direction and the starting point of the path that we split off from $\mathcal{P}$ in $(2 \cdot p)$ ways. The only nontrivial estimate appears for the number of choices of the edge thrown away at the first step of the construction procedure. And here we essentially repeat the arguments that we used in section 3. Namely, consider the simple walk of length $(\ell p - q)$ on the positive semi–axis associated with $\mathcal{P}_{\alpha_1} \vee \mathcal{P}_{\alpha_2} \vee \cdots \vee \mathcal{P}_{\alpha_\ell}$. The necessary condition on $\{i,j\} = \{i_r^{(1)}, i_{r+1}^{(1)}\}$ to be the first edge of $\mathcal{P}$ that also belongs to another path from the cluster can be written in terms of simple walk as

$$x(t) \geq x(r) \quad \text{for} \quad r \leq t \leq r + (\ell p - q) - p - 1 \,.$$

A trivial modification of Lemma 1 gives us a factor $O(p^{1/2})$ for the number of choices of $i = i_r^{(1)}$. We also have $2 \cdot p$ choices for the direction and the starting point of $\mathcal{P}_{\alpha_2}$. Altogether, the above arguments allow us to write the upper bound for the multiplicity of $\mathcal{P}$ as

$$\text{const}_\ell \cdot p^{1/2} \cdot p \cdot p^{2\ell_1+2\ell_2+\cdots+2\ell_r+2g-2} = \text{const}_\ell \cdot p^{q-1/2}\,.$$

Taking into account that the number of even paths of length $(\ell p - q)$ is $O\left(\frac{n^{(\ell p-q)/2+1}}{p^{3/2}}\right)$, we obtain the estimate of the number of correlated clusters containing ℓ paths as $o(n^{\frac{p\ell}{2}})$. Lemma 2 is proven. Formulas (4.1), (4.2) are straightforward corollaries of this lemma. Factor $(2k-1)!! = (2k-1)\cdot(2k-3)\ \ldots 1$ appears as a number of partitions of $\{1, 2, \ldots, 2k\}$ onto two–element subsets corresponding to correlated pairs of paths.

The Central Limit Theorem for Trace $(A^p) - E$ Trace (A^p) is proven. The statement about finite–dimensional distributions

$$\big(\text{Trace } A^{[[e^{t_1}p]]} - E\,\text{Trace } A^{[[e^{t_1}p]]}, \ldots, \ \text{Trace } A^{[[e^{t_m}p]]} - E\ \text{Trace } A^{[[e^{t_m}p]]}\big)$$

can be obtained in the same way.

5. Corollaries of the Main Theorem.

The proof of corollaries formulated in the introduction is rather straightforward. First we prove Corollary 2. Let us choose $p = 2\,[\,\frac{1}{2}\,\frac{n^{1/2}}{\log^{\epsilon/2} n}],\quad \epsilon > 0$. Then

$$\begin{aligned} Pr\left\{\lambda_{\max}(A) \geq 1 + \frac{\log^{1+\epsilon} n}{n^{1/2}}\right\} &\leq Pr\left\{\text{Trace } (A^p) \geq \left(1 + \frac{\log^{1+\epsilon} n}{n^{1/2}}\right)^p\right\} \\ &\leq Pr\left\{\text{Trace } (A^p) \geq \frac{1}{2}\exp(\log^{1+\epsilon/2} n)\right\} \qquad (5.1) \\ &\leq \frac{E \text{ Trace } A^p}{\frac{1}{2}\exp(\log^{1+\epsilon/2} n)} = o(n\exp(-\log^{1+\epsilon/2} n)) \end{aligned}$$

which implies

$$\sum_{n=1}^{\infty} Pr\left\{\lambda_{\max}(A) \geq 1 + \frac{\log^{1+\epsilon} n}{n^{1/2}}\right\} < \infty\,.$$

The statement of corollary 2 now follows from Borel–Cantelli lemma. We finish the paper with the proof of Corollary 1.

Let us denote the linear statistics

$$\sum_{i=1}^{n} f(\lambda_i)$$

by $S_n(f)$. Then we can write

$$\begin{aligned} S_n(f) = S_n(f) \cdot \chi\{|\lambda_i| \leq 1 + \frac{\log^{1+\epsilon} n}{\sqrt{n}};\ i = 1, \ldots, n\} \\ + S_n(f) \cdot \chi\{|\lambda_i| \leq 1 + \frac{\log^{1+\epsilon} n}{\sqrt{n}}, i = 1, \ldots, n\}^c\,, \qquad (5.2) \end{aligned}$$

here by $\{\ \}^c$ we denote the complement of the set. It follows from (5.1) that probability of the complement of the event in (5.2) decays faster than any power of n and

$$E\,S_n(f) - E\left(S_n(f)\cdot\chi\left\{|\lambda_i| \le 1 + \frac{\log^{1+\epsilon} n}{\sqrt{n}};\ i = 1,\ldots,n\right\}\right) \underset{n\to\infty}{\longrightarrow} 0$$

Clearly it is enough to prove the Central Limit Theorem for

$$S_n(f)\cdot\chi\left\{|\lambda_i| \le 1 + \frac{\log^{1+\epsilon} n}{\sqrt{n}}\ ;\ i = 1,\ldots,n\right\}\ .$$

To use our results for the traces of powers of A we write the Taylor series for $f(x)$:

$$f(x) = \sum_{k=0}^{\infty} a_k \cdot x^k\ .$$

The analyticity of $f(x)$ in the closed unit disk implies the exponential decay of the series coefficients:

$$|a_k| \le c\cdot(1-\delta)^k, \quad c = c(\delta) > 0, \quad 0 < \delta < 1\ . \tag{5.3}$$

We choose a positive big enough integer M and write

$$\begin{aligned}
S_n(f)\chi&\left\{|\lambda_i| \leqslant 1 + \frac{\log^{1+\epsilon} n}{\sqrt{n}};\ i = 1,\ldots,n\right\} \\
&- E\left(S_n(f)\cdot\chi\left\{|\lambda_i| \le 1 + \frac{\log^{1+\epsilon} n}{\sqrt{n}};\ i = 1,\ldots,n\right\}\right) \\
&= \sum_{k=0}^{\infty} a_k\cdot(\text{Trace } A^k\chi\{\ldots\} - E(\text{ Trace } A^k\chi\{\ldots\})) \\
&= \sum_{k=0}^{M} a_k\cdot(\text{Trace }\ A^k\cdot\chi\{\ldots\} - E(\text{ Trace } A^k\cdot\chi\{\ldots\})) \\
&\quad + \sum_{k=M+1}^{n^{1/10}} a_k(\text{Trace } A^k\cdot\chi\{\ldots\} - E(\text{Trace }\ A^k\cdot\chi\{\ldots\})) \\
&\qquad + \sum_{k>n^{1/10}} a_k(\text{Trace } A^k\cdot\chi\{\ldots\} - E(\text{ Trace } A^k\cdot\chi\{\ldots\}))\ .
\end{aligned} \tag{5.4}$$

We proved in §3 that

$$\text{Var (Trace } A^p)\,, \qquad p \ll \sqrt{n}$$

are uniformly bounded. This allows us to estimate from above the variance of the second subsum in (5.4) by $\mathfrak{D}\cdot(1-\delta)^M$, where $\mathfrak{D}$ is some constant. Similar arguments imply that the variance of the sum of the first two terms in (5.4) has a finite limit as $n\to\infty$ which we denote σ_f. We also claim that the third term in (5.4) goes to zero, because of (5.2). Finally we remark that for any fixed M the Central Limit Theorem holds for the first subsum. Making M as large as we wish, we derive the Central Limit Theorem for the linear statistics $S_n(f)$.

References

[1] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. **62** (1955), 548–564.

[2] ———, *On the distribution of the roots of certain symmetric matrices*, Ann. of Math. **67** (1958), 325–327.

[3] L. Arnold, *On the asymptotic distribution of the eigenvalues of random matrices*, J. Math. Analysis and Appl. **20** (1967), 262–268.

[4] ———, *On Wigner's Semicircle Law for the eigenvalues of random matrices*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **19** (1971), 191–198.

[5] K. W. Wachter, *The strong limits of random matrix spectra for sample matrices of independent elements*, Ann. of Probability. **6, No. 1** (1978), 1–18.

[6] V. L. Girko, *Spectral Theory of Random Matrices*, Nauka, Moscow. (In Russian), 1988.

[7] V. A. Marchenko, Pastur, L. A., *Distribution of eigenvalues in certain sets of random matrices*, Math. Sbornik **72, No. 4** (1967), 507–536.

[8] L. A. Pastur, *On the spectrum of random matrices*, Theor. Mat. Fiz. **10** (1972), 102–112.

[9] ———, *The spectra of random self–adjoint operators*, Russ. Math. Surv. **28** (1973), 3–64.

[10] A. M. Khorunzhy, B. A. Khoruzhenko, L. A. Pastur and M. V. Shcherbina, *The large–n limit in statistical mechanics and the spectral theory of disordered systems*, Phase Transitions and Critical Phenomena (C. Domb and J. L. Lebowitz, eds.), vol. 15, Academic Press, London, 1992, pp. 73–239.

[11] A. M. Khorunzhy, B. A. Khoruzhenko, L. A. Pastur, *On the $1/N$ corrections to the green functions of random matrices with independent entries*, J. Phys. A: Math. Gen. **28** (1995), L31–L35.

[12] ———, *Asymptotic properties of large random matrices with independent entries*, J. Math. Phys. **37 No. 10** (1996), 5033–5059.

[13] S. A. Molchanov, L. A. Pastur, A. M. Khorunzhy, *Limiting eigenvalue distribution for band random matrices*, Teor. Mat. Fiz. **90** (1992), 108–118.

[14] Z. Füredi, J. Komlóz, *The eigenvalues of random symmetric matrices*, Combinatorica **1, No. 3** (1981), 233–241.

[15] Z. D. Bai, Y. Q. Yin, *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*, Ann. of Probability **16, No. 4** (1988), 1729–1741.

[16] C. Tracy, H. Widom, *On orthogonal and symplectic matrix ensembles*, Commun. Math. Phys. **177** (1996), 727–754.

[17] W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley, New York, 1966.