Functional Analysis and Its Applications, Vol. 32, No. 2, 1998

A Refinement of Wigner's Semicircle Law in a Neighborhood of the Spectrum Edge for Random Symmetric Matrices*

Ya. G. Sinai and A. B. Soshnikov UDC 519.2

§1. Introduction and Statement of Results

We consider a Wigner ensemble of n-dimensional real random symmetric matrices $A_n = \|a_{ij}\|$, where $a_{ij} = a_{ji} = \xi_{ij}/\sqrt{n}$, $1 \le i \le j \le n$, and the ξ_{ij} are independent real random variables. We assume that the following conditions hold.

(i) The random variables ξ_{ij} have symmetric distribution laws, and

$$\mathbf{E}(\xi_{ij})^2 = 1/4 \quad \text{for } i < j, \quad \text{and} \quad \mathbf{E}(\xi_{ii})^2 \le \text{const}.$$

Here and in what follows, $\mathbf{E}$ denotes the expectation, and const stands for numbers independent of n.

(ii) All the moments of ξ_{ij} are finite and admit the estimate

$$\mathbf{E}(\xi_{ij})^{2m} \le (\text{const} \cdot m)^m; \tag{1.1}$$

this means that the distributions of the random variables ξ_{ij} decay at least as fast as Gaussian distributions.

Let λ_k, $k = 1, \dots, n$, be the eigenvalues of A_n. The Wigner limit theorem [1, 2] claims that the empirical distribution function for the numbers λ_k,

$$N_n(\lambda) = \frac{1}{n}\#\{k : \lambda_k < \lambda\},$$

converges in probability as $n \to \infty$ to a distribution corresponding to the Wigner semicircle law

$$\lim_{n\to\infty} N_n(\lambda) = \int_{-\infty}^{\lambda} \rho(u)\, du, \tag{1.2}$$

where

$$\rho(u) = \begin{cases} 0 & \text{for } u > 1 \text{ or } u < -1, \\ \frac{2}{\pi}\sqrt{1-u^2} & \text{for } -1 \le u \le 1. \end{cases}$$

Later, this result was strengthened by Marchenko, Pastur, L. Arnold, Wachter, Girko, and others [3–13].

Consider the r_n-neighborhood O_n of the right spectrum edge $\lambda = 1$, where $r_n n^{2/3} \to \infty$ as $n \to \infty$. For example, one can take $r_n \sim \text{const}/n^{\gamma}$, $\gamma < 2/3$. By formally applying the semicircle law, we find that the number of eigenvalues in O_n behaves as $\text{const} \cdot r_n^{3/2} n$. Let us renormalize the eigenvalues by setting

$$\lambda_k = 1 - \theta_k r_n$$

and place the mass

$$\mu_n(\theta_k) = \frac{1}{n r_n^{3/2}}$$

at each point θ_k. We thus obtain a measure μ_n on the real line such that $\mu_n(\mathbb{R}^1) = r_n^{-3/2}$. The main result of the present paper can be stated as follows.

* This work was supported by RFBR grant No. 96-01-0037 and NSF grant No. DMS-9706794 (Ya. G. Sinai) and by NSF grant No. DMS-9304580 (A. B. Soshnikov).

Princeton University, Dept. of Mathematics, Princeton, New Jersey, USA; Institute for Advanced Study, School of Mathematics, Princeton, New Jersey, USA. Translated from Funktsional′nyi Analiz i Ego Prilozheniya, Vol. 32, No. 2, pp. 56–79, April–June, 1998. Original article submitted September 5, 1997.

 0016–2663/98/3202–0114$12.50

Theorem 1 (Main Theorem). *As $n \to \infty$, the measures μ_n weakly converge in probability on each finite interval to a measure μ concentrated on the half-line $\mathbb{R}^+$ and absolutely continuous with respect to the Lebesgue measure. The density $d\mu(x)/dx$ has the form $(2\sqrt{2}/\pi)\sqrt{x}$ for $x \geq 0$ and is zero for $x < 0$.*

If $r_n > n^{\varepsilon - 2/3}$ for some $\varepsilon > 0$, then the measures μ_n weakly converge to μ with probability 1.

A similar theorem is valid for the eigenvalues in a neighborhood of the left spectrum edge $\lambda = -1$.

The main theorem has a series of consequences. Let $\lambda_{\max}(A_n)$ be the maximum eigenvalue of A_n. We write this value in the form

$$\lambda_{\max}(A_n) = 1 + \theta_{\max}(A_n)\, n^{-2/3}.$$

Corollary 1. *The random variables $\theta_{\max}(A_n)$ are uniformly bounded in probability. In other words, for each $\varepsilon > 0$ there exists a number M such that*

$$\mathbf{P}\{\theta_{\max}(A_n) \leq M\} \leq \varepsilon\,.$$

Corollary 2. *Let $\nu^+(A_n, x)$ be the number of eigenvalues lying to the right of $1 + n^{-2/3}x$, where x is an arbitrary real number. Then*

$$\mathbf{E}\,\nu^+(A_n\,, x) \leq \mathrm{const}\,(x)\,.$$

It also follows from the main theorem that the number of eigenvalues in O_n grows as $nr_n^{3/2}$. Hence, the average distance between eigenvalues in O_n decays as $n^{-1}r_n^{-1/2}$.

The main theorem can be derived from the following theorem.

Theorem 2. *Let $p_n \to \infty$ as $n \to \infty$ so that $p_n = o(n^{2/3})$. In this case*

$$\mathbf{E}\,(\mathrm{Tr}\, A_n^{p_n}) = \begin{cases} 2^{3/2} n (\pi p_n^3)^{-1/2}(1 + o(1)) & \text{if } p_n \text{ is even,} \\ 0, & \text{if } p_n \text{ is odd,} \end{cases}$$

and the distribution $\mathrm{Tr}\, A_n^{p_n} - \mathbf{E}\,(\mathrm{Tr}\, A_n^{p_n})$ weakly converges to the normal law $N(0, 1/\pi)$.

Theorem 2 is a refinement of a similar theorem proved by the authors in [14] for $p_n = o(n^{1/2})$.

Remark 1. We can also state a many-dimensional version of Theorem 2. Let

$$q_n^{(i)} = c^{(i)} p_n \cdot (1 + o(1)), \qquad i = 1, \ldots, l\,,$$

where the $c^{(i)}$ are positive constants and the differences $q_n^{(i_1)} - q_n^{(i_2)}$ are even. Then

$$\lim_{n \to \infty} \mathrm{Cov}\,(\mathrm{Tr}\, A_n^{q_n^{(i_1)}}, \mathrm{Tr}\, A_n^{q_n^{(i_2)}}) = \frac{2}{\pi} \frac{\sqrt{c_{i_1} c_{i_2}}}{c_{i_1} + c_{i_2}}, \qquad 1 \leq i_1, i_2 \leq l\,,$$

and the joint distribution of the random variables $\mathrm{Tr}\, A_n^{q_n^{(i)}} - \mathbf{E}\,(\mathrm{Tr}\, A_n^{q_n^{(i)}})$ converges to a multivariate normal distribution. On the other hand, if we choose $q_n^{(j_1)}$ and $q_n^{(j_2)}$ so that the difference $q_n^{(j_1)} - q_n^{(j_2)}$ is odd, then the centered traces are asymptotically independent as $n \to \infty$.

Theorem 2 will be proved in §§4 and 5. In §2 we derive the main theorem and the corollaries from Theorem 2, the latter being taken for granted.

The results of the present paper are also valid for a Wigner ensemble of complex self-adjoint matrices. This, as well as other remarks, is discussed in §3.

§2. The Derivation of the Main Theorem and its Corollaries from Theorem 2

To prove the main theorem, it suffices to establish the convergence of the Laplace transforms of the measures μ_n, that is, to show that

$$\int_{-\infty}^{\infty} e^{-c\theta} d\mu_n(\theta) = \frac{1}{nr_n^{3/2}} \sum_{k=1}^{n} e^{-c\theta_k} \xrightarrow[n \to \infty]{\mathbf{P}} \int_0^{+\infty} e^{-c\theta} \frac{2\sqrt{2}}{\pi} \sqrt{\theta}\, d\theta = \sqrt{\frac{2}{\pi c^3}}, \tag{2.1}$$

where $c > 0$.

We set $s_n = [(c/2)r_n^{-1}]$ and $p_n = 2s_n$ and introduce the following renormalization for the positive and negative eigenvalues:

$$\begin{aligned} \lambda_k &= 1 - \theta_k r_n\,, \qquad & \lambda_k &\ge 0\,,\\ \lambda_j &= -1 + \tau_j r_n\,, \qquad & \lambda_j &< 0\,. \end{aligned} \tag{2.2}$$

According to Theorem 2, the sequence $n^{-1}r_n^{-3/2}\,\mathrm{Tr}\,A^{2s_n}$ converges in probability to

$$\lim_{n\to\infty} \frac{1}{nr_n^{3/2}}\,\mathbf{E}(\mathrm{Tr}\,A^{2s_n}) = \lim_{n\to\infty} \frac{1}{nr_n^{3/2}}\,\frac{n}{\sqrt{\pi s_n^3}} = \frac{2\sqrt{2}}{\sqrt{\pi c^3}}\,,$$

and $n^{-1}r_n^{-3/2}\,\mathrm{Tr}\,A^{2s_n+1}$ converges in probability to zero.

With regard for (2.2), we obtain

$$\mathrm{Tr}\,A^{2s_n} = \sum_k (1 - r_n\theta_k)^{2[(c/2)r_n^{-1}]} + \sum_j (1 - \tau_j r_n)^{2[(c/2)r_n^{-1}]} \tag{2.3}$$

and

$$\mathrm{Tr}\,A^{2s_n+1} = \sum_k (1 - r_n\theta_k)^{2[(c/2)r_n^{-1}]+1} - \sum_j (1 - \tau_j r_n)^{2[(c/2)r_n^{-1}]+1}. \tag{2.4}$$

Note that if $|\theta_k| \le r_n^{-1/3}$ and $|\tau_j| \le r_n^{-1/3}$, then the corresponding terms in both (2.3) and (2.4) are $e^{-c\theta_j}(1+o(r_n^{1/3}))$ and $e^{-c\tau_j}(1+o(r_n^{1/3}))$. On the other hand, it follows from the estimate of the expectation $\mathbf{E}(\mathrm{Tr}\,A^{4s_n})$ given in Theorem 2 that the subsums in (2.3) and (2.4) over θ_k and τ_j such that

$$|\theta_k| > r_n^{-1/3}, \qquad |\tau_j| > r_n^{-1/3},$$

converge to zero in probability. The same holds for the subsums, over the above θ_k and τ_j, of the linear statistics $\sum_k e^{-c\theta_k}$ and $\sum_j e^{-c\tau_j}$. These considerations and formulas (2.3) and (2.4) yield

$$\left(\mathrm{Tr}\,A^{2s_n} + \mathrm{Tr}\,A^{2s_n+1} - 2\sum_k e^{-c\theta_k}\right)\frac{1}{nr_n^{3/2}} \xrightarrow[n\to\infty]{\mathbf{P}} 0\,,$$

and consequently

$$\frac{1}{nr_n^{3/2}}\sum_k e^{-c\theta_k} \xrightarrow[n\to\infty]{\mathbf{P}} \sqrt{\frac{2}{\pi c^3}}\,.$$

If $r_n = n^{-\gamma}$, where $\gamma < 2/3$, then it follows from Theorem 2 that the normalized traces $n^{-1}r_n^{-3/2}\,\mathrm{Tr}\,A^{2s_n}$ and $n^{-1}r_n^{-3/2}\,\mathrm{Tr}\,A^{2s_n+1}$ converge to nonrandom limits with probability 1, and hence the convergence of the measures μ_n in the main theorem also occurs with probability 1. It also follows from Theorem 2 that the fluctuations of the random variables $\sum_k e^{-c\theta_k}$ are of the order of a constant and converge in distribution to the normal law $N(0, 1/2\pi)$ as $n \to \infty$ provided that $r_n \ll n^{-2/5}$.

Below we prove Corollaries 1 and 2.

Let us represent the maximum eigenvalue in the form $\lambda_{\max} = 1 + \theta_{\max} n^{-2/3}$. Suppose that there exist sequences $n_i \to \infty$ and $L_i \to \infty$ and an $\varepsilon > 0$ such that

$$\mathbf{P}\{\lambda_{\max}(A_{n_i}) > 1 + n_i^{-2/3}L_i\} > \varepsilon > 0\,. \tag{2.5}$$

Let us consider $\mathrm{Tr}\,A_{n_i}^{p_{n_i}}$, where $p_{n_i} = 2[n^{2/3}/L_i^{1/2}]$ (for convenience, we can always assume that $L_i \ll n^{2/3}$) and apply Theorem 2. Then

$$\mathbf{E}(\mathrm{Tr}\,A_{n_i}^{p_{n_i}}) \le \frac{2}{\sqrt{\pi}}\,L_i^{3/2} \tag{2.6}$$

for sufficiently large n_i.

On the other hand, it would follow from inequality (2.5) that

$$\mathbf{E}(\mathrm{Tr}\,A_{n_i}^{p_{n_i}}) > \mathbf{E}\,\lambda_{\max}(A_{n_i})^{p_{n_i}} > \varepsilon\,(1 + n_i^{-2/3}L_i)^{2[n_i^{2/3}/\sqrt{L_i}]} \ge \varepsilon\, e^{\sqrt{L_i}}. \tag{2.7}$$

For sufficiently large L_i, the last inequality contradicts (2.6). The proof of Corollary 1 is complete.

Proof of Corollary 2. Corollary 2 can also be proved by contradiction. Suppose that there exist sequences $n_i \to \infty$ and $L_i \to \infty$ such that

$$\mathbf{E}\nu^+(A_{n_i}, x) > L_i, \tag{2.8}$$

where $\nu^+(A_n, x)$ is the number of eigenvalues lying to the right of $1 + n^{-2/3}x$, where $x > 0$ is arbitrary. To obtain a contradiction to (2.8), we take $p_{n_i} = 2[n_i^{2/3}/\sqrt{L_i}]$ and apply the estimate for the expectation (provided by Theorem 2) to $\operatorname{Tr} A_{n_i}^{p_{n_i}}$. We obtain $\mathbf{E}(\operatorname{Tr} A_{n_i}^{p_{n_i}}) \le (2/\sqrt{\pi})\, L_i^{3/4}$ for sufficiently large n_i. On the other hand,

$$\mathbf{E}\nu^+(A_{n_i}, x) \le \frac{\mathbf{E}(\operatorname{Tr} A_{n_i}^{p_{n_i}})}{(1 - x n_i^{-2/3})^{2[n_i^{2/3}/L_i^{1/2}]}} \le 2\mathbf{E}(\operatorname{Tr} A_{n_i}^{p_{n_i}}) \le \frac{4}{\sqrt{\pi}}\, L_i^{3/4}$$

for sufficiently large n_i, which contradicts (2.8) provided that L_i is also chosen to be sufficiently large. The proof of Corollary 2 is complete.

§3. Additional Remarks

1. Similar results hold for a Wigner ensemble of complex self-adjoint matrices $A_n = \{a_{kj}\}_{1 \le k, j \le n}$, where

$$a_{kj} = \overline{a_{jk}} = \frac{\operatorname{Re}\xi_{kj} + i \operatorname{Im}\xi_{kj}}{n^{1/2}}, \quad 1 \le k < j \le n, \qquad a_{kk} = \frac{\xi_{kk}}{n^{1/2}}, \quad k = 1, \dots, n,$$

$\mathbf{E}|\xi_{kj}|^2 = 1/4$ for $k < j$, and $\mathbf{E}\xi_{kk}^2 \le \text{const}$, as well as for an ensemble of positive definite matrices $B_n B_n^*$ near the right spectrum edge under the assumption that all the matrix elements $b_{ij} = n^{-1/2}\eta_{ij}$ are independent random variables.

2. For the case in which the matrix elements are Gaussian random variables, explicit formulas are known for many characteristics of the local distribution of eigenvalues [20, 16]. In particular, the distribution of the point random field $\{x_1, \dots, x_n\}$ given by the formula

$$\lambda_i = 1 + n^{-2/3} x_i, \qquad i = 1, \dots, n, \tag{3.1}$$

is known to have a limit as $n \to \infty$. The limit point random field (which is related to the local distribution of the eigenvalues near the spectrum edge) is determined by its k-point correlation functions, $k = 1, 2, \dots$. For an ensemble of Hermitian matrices such that

$$\operatorname{Re}\xi_{kj} \sim N(0, \tfrac{1}{8}), \quad \operatorname{Im}\xi_{kj} \sim N(0, \tfrac{1}{8}), \qquad 1 \le k < j \le n, \qquad \xi_{kk} \sim N(0, \tfrac{1}{4}),$$

which is known in the literature on random matrices [20] as the *Gaussian unitary ensemble*, the k-point correlation function is the determinant

$$\rho_k(y_1, \dots, y_k) = \det(K(y_i, y_j))_{1 \le i, j \le k} \tag{3.2}$$

of the k-dimensional matrix whose entries are the values of the kernel

$$K(x, y) = \frac{A(x)A'(y) - A'(x)A(y)}{x - y},$$

where $A(x)$ is the Airy function [16, 17].

Interestingly enough, for the ensemble of symmetric matrices with Gaussian entries

$$\xi_{ij} \sim N(0, \tfrac{1}{4}), \quad 1 \le i < j \le n, \qquad \xi_{ii} \sim N(0, \tfrac{1}{2}), \quad 1 \le i \le n$$

(the Gaussian orthogonal ensemble), the local distribution of eigenvalues differs from (3.2) [15, 20]. In both cases, the point random field is condensed as $x \to -\infty$, that is, the distance between neighboring eigenvalues tends to zero, and the spectral density $\rho_1(x)$ is asymptotically equivalent to the Wigner density as $x \to -\infty$. This suggests that the (local) Wigner law holds in a neighborhood of the spectral edge for any scale larger than $n^{-2/3}$. The main theorem of the present paper asserts that the same result holds for an arbitrary Wigner ensemble. Specifically, let us consider the point random field $\{x_1, \dots, x_n\}$ defined in (3.1), choose a sequence $R_n \to +\infty$ such that $R_n \ll n^{2/3}$, and count the points x_i lying to the right of $-R_n$: $N_n(-R_n) = \#\{x_i > -R_n\}$. By definition, $N_n(-R_n) = \#\{\lambda_i > 1 - n^{-2/3}R_n\}$. In this case, by virtue of the main theorem, the variable $N_n(-R_n)$ in the leading term is nonrandom and is equal to

$$n \int_{1-n^{-2/3}R_n}^{1} \frac{2}{\pi}\sqrt{1-t^2}\,dt \cdot (1+o(1)).$$

The order of fluctuation of the random variable $N_n(-R_n)$ is so far unknown. It is natural to assume, by analogy with the case of the local distribution of eigenvalues inside the spectrum for the Gaussian ensembles [21], that the fluctuation is of the order of $\sqrt{\log(R_n)}$. This conjecture is partially justified by the following corollary to Theorems 1, 2: for the smoothed counting functions $\sum_i e^{cx_i/R_n}$, $c > 0$, with $R_n \ll n^{2/3-2/5}$, the fluctuations are of the order of a constant, and the centered linear statistics

$$\sum_i e^{cx_i/R_n} - \mathbf{E}\Big(\sum_i e^{cx_i/R_n}\Big)$$

converge in distribution to the normal law $N(0, \frac{1}{2\pi})$.

3. The first results concerning the traces of high powers of Wigner matrices are due to Füredi and Komlóz [18] and A. Boutet de Monvel and Shcherbina [19]. In particular, Füredi and Komlóz treated the case of uniformly bounded random variables ξ_{ij} (whose distribution law is not required to be symmetric) and proved a formula for the leading term of the expectation $\mathbf{E}(\operatorname{Tr} A^{P_n})$ for $p_n \ll n^{1/6}$. Note that the technique of the present paper can be generalized to the case of nonsymmetrically distributed random variables ξ_{ij}. In this case, we face the necessity of studying paths passing through some edge an odd (> 1) number of times (see §§4, 5).

For $p_n \ll n^{1/2}$, the main theorem and Theorem 2 remain valid for nonsymmetrically distributed ξ_{ij}. Moreover, for $p_n \ll n^{1/2}$, condition (1.1) on the growth of the moments of the random variables ξ_{ij} can be weakened by requiring that the distribution functions of ξ_{ij} should decay at infinity at most exponentially, that is, by replacing (1.1) by

$$\mathbf{E}(\xi_{ij})^{2k} \le (\mathrm{const} \cdot k)^{2k}. \tag{1.1'}$$

§4. The Expectation $\mathbf{E}(\operatorname{Tr} A_n^{p_n})$

Theorem 2 will be proved by the moment method. To this end, we must prove that the moments of the random variable $\operatorname{Tr} A_n^{p_n} - \mathbf{E}(\operatorname{Tr} A_n^{p_n})$ converge to the corresponding moments of the normal distribution $N(0, \frac{1}{\pi})$. In this section, we estimate the leading term of the expectation $\mathbf{E}(\operatorname{Tr} A_n^{p_n})$. Clearly,

$$\mathbf{E}(\operatorname{Tr} A_n^{2s_n}) = \frac{1}{n^{s_n}} \sum_{\mathcal{P}} \mathbf{E}\xi_{i_0 i_1} \cdot \xi_{i_1 i_2} \cdot \ldots \cdot \xi_{i_{2s_n-1}, i_0}. \tag{4.1}$$

The sum in (4.1) is taken over all closed paths $\mathcal{P} = \{i_0, i_1, \ldots, i_{2s_n-1}, i_0\}$ with a distinguished origin in the set $\{1, \ldots, n\}$. It is convenient to regard the set of vertices $\{1, \ldots, n\}$ as a nonoriented graph in which any two vertices are joined by an unordered edge. Since the distribution of the random variables ξ_{ij} is symmetric, it follows that the only paths contributing to (4.1) are those for which the number of occurrences of each edge is even. In what follows, we consider only such paths, which are said to be *even*.

Definition 1. An instant k is said to be *marked* if the (unordered) edge $\{i_{k-1}, i_k\}$ occurs an odd number of times up to the instant k (inclusive). The other instants are said to be *unmarked*.

Each even path contains equally many (namely, $s_n = p_n/2$) marked and unmarked instants. To each path $\mathcal{P}$ we assign a trajectory $X = \{x(0), x(1), \ldots, x(2s_n)\}$ of a random walk on the positive half-line, where $x(0) = 0$ and the difference $x(k) - x(k-1)$ is equal to 1 or -1 depending on whether the kth instant is or is not marked. Obviously, $x(t) \ge 0$ for all $0 \le t \le p_n$, and we have $x(p_n) = 0$ for even paths.

Definition 2. An even path $\mathcal{P}$ is called a *path without self-intersections* if, for any two distinct marked instants k' and k'', one has $i_{k'} \neq i_{k''}$.

Any path without self-intersections has the following structure. First, there is a series of marked instants at which the path passes through distinct vertices (the number of vertices is equal to the length of the series). Then, there is a series of unmarked instants at which the path passes through some of these vertices in the reverse order. Then, there follows a new series of marked instants and the corresponding series of vertices, etc. It is important that, for paths without self-intersections, the trajectory is uniquely determined at the unmarked instants provided the trajectory at the marked edges is known. Each edge is passed twice in such paths.

The subsum of (4.1) over the paths without self-intersections is equal to

$$Z(0) = \frac{1}{n^{s_n}} \sum_{\substack{\mathcal{P} \text{ without} \\ \text{self-intersections}}} \mathbf{E}\xi_{i_0 i_1} \cdots \xi_{i_{2s_n-1} i_0} = \frac{1}{n^{s_n}} \frac{1}{4^{s_n}} \frac{(2s_n)!\, n(n-1)\cdots(n-s_n)}{s_n!(s_n+1)!}. \tag{4.2}$$

Here $(2s_n)!/(s_n!(s_n+1)!)$ is the number of trajectories, of length $2s_n$ and with $x(0) = x(2s_n) = 0$, of the simplest random walk on the nonnegative half-line, and the product $n(n-1)\cdots(n-s_n)$ specifies the number of ways in which the initial vertex and the vertices at the marked instances can be chosen. It was shown in [14] that for $s_n = o(n^{1/2})$, the main contribution to $\mathbf{E}(\operatorname{Tr} A_n^{2s_n})$ is due to the paths without self-intersections, that is,

$$\mathbf{E}(\operatorname{Tr} A_n^{2s_n}) = Z(0) \cdot (1 + o(1)).$$

This is no longer valid in the general case, and we must study the statistics of self-intersections.

Definition 3. A marked instant m is called an *instant of self-intersection* if there exists a marked instant $m' < m$ such that $i_{m'} = i_m$.

Definition 4. A vertex i is called a *vertex of simple* (*respectively, triple, quadruple*, etc.) *intersection* if there are exactly two (respectively, three, four, etc.) marked instances i_m such that $i_m = i$.

According to Definition 4, all the vertices split into $s_n + 1$ disjoint subsets:

$$\{1, \ldots, n\} = \bigsqcup_{k=0}^{s_n} N_k,$$

where N_k is the subset of vertices of k-fold intersection. In other words, recasting Definition 4, we say that a vertex i belongs to the class N_k, $k \geq 0$, if there are exactly k marked instants $m_1, \ldots, m_k$ such that $i_{m_j} = i$, $j = 1, \ldots, k$. All but one of the vertices from N_0 do not belong to $\mathcal{P}$. The initial point i_0 of the path may be the only exception provided that it is not visited at marked instants at the intermediate steps. Let $n_k = \#(N_k)$. We can readily see that

$$\sum_{k=0}^{s_n} n_k = n, \qquad \sum_{k=0}^{s_n} k n_k = s_n. \tag{4.3}$$

We say that $\mathcal{P}$ is a *path of type* $(n_0, n_1, \ldots, n_{s_n})$. For example, every even path without self-intersections is a path of type $(n - s_n, s_n, 0, \ldots 0)$. It was shown in [14] that the subsum over the paths of type $(n_0, n_1, \ldots, n_{s_n})$ is bounded above by

$$\frac{1}{n^{s_n}} \frac{n!}{n_0!\, n_1! \cdots n_{s_n}!}\, n \frac{(2s_n)!}{s_n!(s_n+1)!} \frac{s_n!}{\prod_{k=1}^{s_n}(k!)^{n_k}}\, U(n_0, n_1, \ldots, n_{s_n}), \tag{4.4}$$

where

$$U(n_0, n_1, \ldots, n_{s_n}) = \max_{\substack{\mathcal{P} \text{ is of type} \\ (n_0, n_1, \ldots, n_{s_n})}} \mathbf{E}\left(\prod_{l=0}^{s_n} \xi_{i_l i_{l+1}}\right) W_n,$$

and W_n is the number of ways in which the trajectory can be chosen at unmarked instants provided that the vertices at the marked instants have already been chosen. For example, $W_n = 1$ for the paths without self-intersections.

In [14], we used the following estimate for $U(n_0, n_1, \dots, n_{s_n})$:

$$U(n_0, n_1, \dots, n_{s_n}) \le \left(\frac{1}{4}\right)^{n_1} \prod_{k=2}^{s_n} (\text{const}_1 \cdot k)^{2kn_k}.$$

This estimate is far from being optimal if $n_k > 0$ for large k. Below (in Lemma 1) we show that

$$U(n_0, n_1, \dots, n_{s_n}) \le \left(\frac{1}{4}\right)^{n_1} \prod_{k=2}^{s_n} (\text{const}_2 \cdot k)^{kn_k}. \tag{4.5}$$

By performing the summation over all paths of type $(n_0, n_1, \dots, n_{s_n})$ such that

$$\sum_{k=2}^{s_n} kn_k \ge 10\,\frac{s_n^2}{n}$$

and by using the estimates (4.4) and (4.5), one can readily find that the subsum over such paths is $o(1)$, and hence it is small compared with the value (established in Theorem 2) of the total sum

$$\mathbf{E}(\operatorname{Tr} A_n^{2s_n}) = \frac{n}{\sqrt{\pi s_n^3}}\,(1 + o(1)).$$

In what follows, we study the sum over the paths for which we always have

$$\sum_{k=2}^{s_n} kn_k < 10\,\frac{s_n^2}{n}\,. \tag{4.6}$$

Let $M = \sum_{k=2}^{s_n} (k-1)n_k$. If a path $\mathcal{P}$ contains only simple self-intersections, then $M = n_2$. We shall show that in our case ($s_n = o(n^{2/3})$) it is the paths with simple self-intersections that make the main contribution to $\mathbf{E}(\operatorname{Tr} A_n^{2s_n})$. Let us denote the sum over paths with M simple self-intersections by $Z(M)$. The following assertion holds.

Proposition 1.

$$Z(M) = \frac{n}{\sqrt{\pi s_n^3}}\,e^{-s_n^2/2n}\,\frac{1}{M!}\left(\frac{s_n^2}{2n}\right)^M (1 + o(1)) \tag{4.7}$$

uniformly with respect to $0 \le M \le 10 s_n^2/n$.

Remark 2. By formula (4.7), the number M of self-intersections for typical paths is of the order of s_n^2/n and is equal to $(s_n^2/2n)(1 + O(\sqrt{n}/s_n))$ for $s_n \gg \sqrt{n}$.

Proof of Proposition 1. Let us calculate $Z(M)$ as follows. First, we choose $s_n = p_n/2$ marked instants t_i and assume that $0 \le t_1 < t_2 < \dots < t_{s_n} < 2s_n$. Recall that to any choice of marked instants there corresponds a trajectory $x(t)$, $0 \le t \le 2s_n$, of the simplest random walk on the positive half-line, namely,

$$\begin{gathered}
x(0) = x(2s_n) = 0, \\
x(t_j) - x(t_j - 1) = 1, \qquad j = 1, \dots, s_n, \\
x(t) - x(t-1) = -1 \qquad \text{for } t \ne t_1, \dots, t_{s_n}.
\end{gathered}$$

Then, among the marked instants we choose M instants of self-intersection $t_{j_1}, t_{j_2}, \dots, t_{j_M}$ (that is, we choose indices $j_1, \dots, j_M$ such that $1 \le j_1 < j_2 < \dots < j_M \le s_n$). After this, we choose the origin of the path and the vertices occurring at the marked instants. (The origin can be chosen in n ways, and then we successively choose, in $n-1, n-2, \dots, n - s_n + M$ ways, the vertices occurring at the marked instants that are not instants of self-intersection.) At the instants of self-intersection, the vertices are chosen as follows.

At the first marked instant t_{j_1} of self-intersection, we can choose one of the vertices that occur in the path before the instant t_{j_1}, and there are exactly $j_1 - 1$ such vertices because our trajectory $x(t)$ of the random walk made $j_1 - 1$ steps to the right for $0 \le t < t_{j_1}$. At the next instant of self-intersection, t_{j_2}, there are exactly $j_2 - 2$ possibilities of choosing a vertex (as long as we deal with simple self-intersections, the vertex chosen at the instant t_{j_1} cannot occur at any later instant of self-intersection). Likewise, at the last marked instant of self-intersection, t_{j_M}, there are $j_M - M$ possibilities of choosing the next vertex. For a path with self-intersections, the choice of vertices at unmarked instants (the choice of the "backward trajectory") may be ambiguous. For the "first return" from a vertex of simple self-intersection, one of the following three edges can be chosen:

(a) the edge used to arrive at the vertex for the first time;

(b) the edge used to leave the vertex;

(c) the edge used to arrive at the vertex for the second time.

We shall show below that only the third possibility is realized for typical paths since, by the instant of self-intersection, the edges (a) and (b) have already been passed twice. Let $Z_1(M)$ be the sum over the paths with M simple self-intersections such that the edge used to arrive at a vertex of self-intersection for the last time is always chosen for returning. If each edge of $\mathcal{P}$ is passed twice, then

$$\mathbf{E}\left(\prod_{l=0}^{2s_n-1} \xi_{i_l i_{l+1}}\right) = \left(\frac{1}{4}\right)^{s_n}. \tag{4.8}$$

In the general case, some edges can be used four times. Let m be the number of such edges. In this case,

$$\mathbf{E}\left(\prod_{l=0}^{2s_n-1} \xi_{l_l i_{l+1}}\right) \le \left(\frac{1}{4}\right)^{s_n-2m} (\mathrm{const}_2)^m. \tag{4.9}$$

Let us show that the main contribution to $Z_1(M)$ is due to paths in which each edge is passed twice. In the Appendix, we study the following characteristic of $\mathcal{P}$: the maximum number of vertices that can be visited at marked instants from a given vertex. Let us denote this number by $\nu_n(\mathcal{P})$. By definition, each vertex i of the path $\mathcal{P}$ is the left end of at most $\nu_n(\mathcal{P})$ marked edges. In the Appendix, we prove that $\nu_n(\mathcal{P})$ cannot grow too fast for typical paths; for instance, it grows no faster than s_n^γ for any γ, $0 < \gamma \le 1$. In other words, the sum over paths with $\nu_n(\mathcal{P}) > s_n^\gamma$ is $o(1)$ (see Lemma 2). In what follows, we always assume that

$$\nu_n(\mathcal{P}) < s_n^{1/4}. \tag{4.10}$$

Let $j_{u_1}, \dots, j_{u_m}$ be the indices of the instants of self-intersection corresponding to edges that are passed four times. In this case,

$$Z_1(M) = \sum_{X=\{x(t)\}} \sum_{1\le j_1<\cdots<j_M\le s_n} n(n-1)\cdots(n-s_n+M) \times (j_1-1)(j_2-2)\cdots(j_M-M)\,\frac{1}{n^{s_n}}\,\mathbf{E}\left(\prod_{l=0}^{2s_n-1} \xi_{i_l i_{l+1}}\right). \tag{4.11}$$

The probability of choosing a vertex of self-intersection corresponding to an edge used four times is of the order of $\nu_n(\mathcal{P})/s_n$, and the average number of such vertices is $O(M\nu_n(\mathcal{P})/s_n) = o(1)$ because $\nu_n(\mathcal{P}) < s_n^{1/4}$ and $M < 10s_n^2/n \ll s_n^{1/2}$. More precisely, the above argument can be carried out as follows:

$$Z_1(M) = \sum_{X} \sum_{1\le j_1<\cdots<j_M\le s_n} n(n-1)\cdots(n-s_n+M) \times (j_1-1)(j_2-2)\cdots(j_M-M)\,\frac{1}{n^{s_n}}\left(\frac{1}{4}\right)^{s_n} + Z_1'(M), \tag{4.12}$$

where

$$0 \le Z_1'(M) \le \sum_X \sum_{1\le j_1<\cdots<j_M\le s_n} \sum_{m=1}^{M} n(n-1)\cdots(n-s_n+M)$$
$$\times \sum_{1\le u_1<\cdots<u_m\le M} (j_1-1)\cdots\widehat{(j_{u_1}-u_1)}\cdots\widehat{(j_{u_m}-u_m)}\cdots(j_M-M)$$
$$\times (s_n^{1/4})^m \frac{1}{n^{s_n}}\left(\frac{1}{4}\right)^{s_n-2m} (\mathrm{const}_2)^m \tag{4.13}$$

(we write $\widehat{(j_{u_1}-u_1)}\cdots\widehat{(j_{u_m}-u_m)}$ to indicate that these terms are absent in the product). The subsum $Z_1'(M)$ corresponds to the paths that contain edges passed four times. Let $B(M)$ be the subsum

$$\sum_{1\le j_1<\cdots<j_M\le s_n} (j_1-1)\cdots(j_M-M)$$

in (4.12). The estimate

$$B(M) = \sum_{0\le j_1\le\cdots\le j_M\le s_n-M} j_1\cdots j_M = \frac{1}{M!}\left(\sum_0^{s_n-M} j\right)^M \left(1+O\left(\frac{s_n}{n^{2/3}}\right)\right)$$
$$= \frac{1}{M!}\left(\frac{(s_n-M)(s_n-M+1)}{2}\right)^M \left(1+O\left(\frac{s_n}{n^{2/3}}\right)\right) = \frac{(s_n^2/2)^M}{M!}(1+o(1)) \tag{4.14}$$

holds uniformly with respect to $0 \le M \le 10s_n^2/n$. It follows from (4.14) that

$$Z_1(M) = \frac{(2s_n)!}{s_n!(s_n+1)!}\, n \prod_{k=1}^{s_n-M}\left(1-\frac{k}{n}\right)\frac{B(M)}{n^M}\left(\frac{1}{4}\right)^{s_n}(1+o(1)) + Z_1'(M)$$
$$= \frac{1}{\sqrt{\pi}}\frac{n}{s_n^{3/2}}\exp\left(-\frac{s_n^2}{2n}\right)\frac{(s_n^2/2n)^M}{M!}(1+o(1)) + Z_1'(M)\,. \tag{4.15}$$

Similar estimates for $Z_1'(M)$ prove that the relation

$$Z_1'(M) \le O\left(\frac{M s_n^{1/4}}{s_n}\right)\frac{1}{\sqrt{\pi}}\frac{n}{s_n^{3/2}}\exp\left(-\frac{s_n^2}{2n}\right)\frac{(s_n^2/2n)^M}{M!} \tag{4.16}$$

holds uniformly with respect to $0 \le M \le 10s_n^2/n$. Thus,

$$Z_1(M) = \frac{1}{\sqrt{\pi}}\frac{n}{s_n^{3/2}}\exp\left(-\frac{s_n^2}{2n}\right)\frac{(s_n^2/2n)^M}{M!}(1+o(1)),$$

and, by summing over all M, we obtain

$$Z_1 = \sum_{M=0}^{s_n} Z_1(M) = \sum_{M=0}^{10s_n^2/n} Z_1(M) + o(1) = \frac{1}{\sqrt{\pi}}\frac{n}{s_n^{3/2}}(1+o(1))\,. \tag{4.17}$$

The vertices of self-intersection for which there are several possibilities of returning will be called *nonclosed*. Let $r = r(\mathcal{P})$ be the number of such vertices. Moreover, let $t_{j_{k_l}}$, $l = 1, \ldots, r$, be the instants of self-intersection corresponding to the nonclosed vertices. Obviously, a vertex is nonclosed only if the edge used to arrive at this vertex has not been used for returning (i.e., at an unmarked instant) by the current instant of self-intersection. Hence, the number of possibilities of choosing a nonclosed vertex of self-intersection at an instant t does not exceed $x(t)$. We must also take account of the fact that in the general case the geometry of self-intersections can be more complicated (along with simple self-intersections, the path may contain self-intersections that are triple, quadruple, etc.).

Let $\mathcal{P}$ contain M instants of self-intersection, that is,

$$M = \sum_{k=2}^{s_n} (k-1)\, n_k \,.$$

We denote the indices of the instants of simple self-intersection by $j_1, \ldots, j_{n_2}$, $1 \le j_1 < \cdots < j_{n_2} \le s_n$, and the number of nonclosed vertices by r. Let $j_{l_1}, \ldots, j_{l_r}$ be the indices of the instants of simple self-intersection corresponding to the nonclosed vertices. The pairs of indices of the marked instants corresponding to the vertices of class N_3 will be denoted by

$$(j_{1,1}^{(2)}, j_{1,2}^{(2)}),\ (j_{2,1}^{(2)}, j_{2,2}^{(2)}),\ \ldots,\ (j_{n_3,1}^{(2)}, j_{n_3,2}^{(2)})$$

and ordered so that

$$1 \le j_{1,1}^{(2)} < j_{2,1}^{(2)} < \cdots < j_{n_3,1}^{(2)} \le s_n\,, \qquad j_{k,1}^{(2)} < j_{k,2}^{(2)}\,, \quad k = 1, \ldots, n_3\,. \tag{4.18}$$

We assume that the kth vertex of triple self-intersection is visited for the second time (at a marked instant) at $t_{j_{k,1}^{(2)}}$ and, for the third time, at an instant $t_{j_{k,2}^{(2)}}$.

Likewise, by

$$(j_{1,1}^{(3)}, j_{1,2}^{(3)}, j_{1,3}^{(3)}), \ldots, (j_{n_4,1}^{(3)}, j_{n_4,2}^{(3)}, j_{n_4,3}^{(3)})$$

we denote the triples of indices of the self-intersection instants corresponding to the vertices of class N_4, and we assume that

$$1 \le j_{1,1}^{(3)} < j_{2,1}^{(3)} < \cdots < j_{n_4,1}^{(3)} \le s_n\,, \quad j_{k,1}^{(3)} < j_{k,2}^{(3)} < j_{k,3}^{(3)}\,, \quad k = 1, \ldots, n_4\,, \tag{4.19}$$

etc. Let Z_2 be the sum over the paths admitting only simple self-intersections and having at least one nonclosed vertex, and let Z_3 be the sum over paths admitting nonsimple self-intersections. In what follows, we show that Z_2 and Z_3 are small compared with Z_1, namely,

$$Z_2 = o(n/s_n^{3/2})\,, \qquad Z_3 = o(n/s_n^{3/2})\,. \tag{4.20}$$

First, let us consider Z_2. By the above reasoning, we have the upper bound

$$\begin{aligned} Z_2 \le \sum_X \sum_{M=1}^{10 s_n^2/n} \sum_{1 \le j_1 < \cdots < j_M \le s_n} \sum_{r=1}^{M} \sum_{1 \le l_1 < \cdots < l_r \le M} & n(n-1)\cdots(n-s_n+M) \\ \times (j_1 - 1)\cdots\widehat{(j_{l_1} - l_1)}\cdots\widehat{(j_{l_r} - l_r)}\cdots(j_M - M) & \\ \times x(t_{j_{l_1}})\cdots x(t_{j_{l_r}}) \cdot \frac{1}{n^{s_n}} \cdot \mathbf{E}\left(\prod_{k=0}^{2s_n-1} \xi_{i_k i_{k+1}}\right) & \cdot W_n + o(1)\,. \end{aligned} \tag{4.21}$$

The last factor, W_n, in (4.21) is the number of ways to continue a trajectory at the unmarked instants provided that the vertices for the marked instants have already been chosen.

If all self-intersections in $\mathcal{P}$ are simple and there are exactly r nonclosed vertices, then $W_n \le 3^r$. The double sum

$$\sum_{1 \le j_1 < \cdots < j_M \le s_n} \sum_{1 \le l_1 < \cdots < l_r \le M} (j_1 - 1)\cdots\widehat{(j_{l_1} - l_1)}\cdots\widehat{(j_{l_r} - l_r)}\cdots(j_M - M) \times x(t_{j_{l_1}})\cdots x(t_{j_{l_r}})$$

is bounded above by

$$\frac{1}{M!}\left(\sum_{j=1}^{s_n}(j-1)\right)^{M-r}\left(\sum_{l=1}^{s_n} 1\right)^{r} \frac{1}{r!}\left(\sum_{k=1}^{M} 1 \cdot \max_{0 \le t \le 2s_n} x(t)\right)^{r} \le \frac{1}{M!}\left(\frac{s_n^2}{2}\right)^{M} \frac{1}{r!}\left(\frac{2M}{\sqrt{s_n}} \max_{0 \le t \le 2s_n} \frac{x(t)}{\sqrt{s_n}}\right)^{r}.$$

Thus, the subsum Z_2 over paths in which each edge is passed twice does not exceed

$$n\,\frac{(2s_n)!}{s_n!\,(s_n+1)!}\left(\frac{1}{4}\right)^{s_n} e^{-s_n^2/2n}(1+o(1))\left(\sum_{M=0}^{\infty}\frac{(s_n^2/2n)^M}{M!}\right)\frac{\sum_X\sum_{r=1}^{\infty}\frac{1}{r!}\left(\frac{60s_n^{3/2}}{n}\max\frac{x(t)}{\sqrt{s_n}}\right)^r}{\sum_X 1}+o(1)$$

$$\leq\frac{1}{\sqrt{\pi}}\,\frac{n}{s_n^{3/2}}(1+o(1))\left(\mathbf{E}_X\left(\exp\left(\frac{60s_n^{3/2}}{n}\max_{0\leq t\leq 2s_n}\frac{x(t)}{\sqrt{s_n}}\right)-1\right)\right). \tag{4.22}$$

The last factor (the mathematical expectation) in (4.22) tends to zero as $n\to\infty$. Indeed, the distribution of a normalized trajectory $x_n(t)=(x([2s_nt])+(2s_nt-[2s_nt])(x([2s_nt]+1)-x([2s_nt])))(2s_n)^{-1/2}$, $0\leq t\leq 1$, of the random walk converges to the conditional distribution of the Bessel process on the positive half-line under the condition $b(1)=0$. Consequently,

$$\mathbf{E}_X\left(\exp\left(\varepsilon\max_{0\leq t\leq 2s_n}\frac{x(t)}{\sqrt{2s_n}}\right)\right)\xrightarrow[n\to\infty]{}\mathbf{E}\left(\exp\left(\varepsilon\max_{0\leq t\leq 1}b(t)\right)\,\Big|\,b(1)=0\right)$$

for an arbitrarily small $\varepsilon>0$, and since the coefficient $60s_n^{3/2}/n$ in formula (2.22) tends to zero as $n\to\infty$, it follows that

$$0\leq\mathbf{E}_X\left(\exp\left(\frac{60s_n^{3/2}}{n}\max_{0\leq t\leq 2s_n}\frac{x(t)}{\sqrt{s_n}}\right)-1\right)\leq\mathbf{E}_X\left(\exp\left(\varepsilon\max_{0\leq t\leq 2s_n}\frac{x(t)}{\sqrt{s_n}}\right)-1\right)$$

$$\leq\left(\mathbf{E}\left(\exp\left(\varepsilon\max_{0\leq t\leq 1}b(t)\,\Big|\,b(1)=0\right)\right)-1\right)(1+\varepsilon) \tag{4.23}$$

for sufficiently large n. By choosing ε to be sufficiently small, we can ensure that the right-hand side of (4.23) is as close to zero as desired. Hence,

$$Z_2=o(Z_1)+Z_2', \tag{4.24}$$

where Z_2' is the subsum of Z_2 over the paths in which at least one edge is used four times. The sum Z_2' can be analyzed in the same way as Z_1', and we finally obtain $Z_2'=o(Z_1)$. This, together with (4.24), yields $Z_2=o(Z_1)$.

Now let us consider the sum Z_3 over the paths that admit nonsimple self-intersections. Using the above notation, we can write out the following estimate for Z_3:

$$Z_3\leq\sum_{X=\{x(t)\}}\sum_{M=1}^{10s_n^2/n}\sum_{n_2+2n_3+\cdots+Mn_{M+1}=M}\sum_{r=0}^{n_2}\sum_{1\leq j_1<\cdots<j_{n_2}\leq s_n}\sum_{1\leq l_1<\cdots<l_r\leq n_2}$$

$$\sum_{\substack{(j_{1,1}^{(2)},j_{1,2}^{(2)}),\ldots,(j_{n_3,1}^{(2)},j_{n_3,2}^{(2)}),\\ 1\leq j_{1,1}^{(2)}<j_{2,1}^{(2)}<\cdots<j_{n_3,1}^{(2)}\leq s_n,\\ j_{q_2,1}^{(2)}<j_{q_2,2}^{(2)},\ q_2=1,\ldots,n_3}}\ \sum_{\substack{(j_{1,1}^{(3)},j_{1,2}^{(3)},j_{1,3}^{(3)}),\ldots,(j_{n_4,1}^{(3)},j_{n_4,2}^{(3)},j_{n_4,3}^{(3)}),\\ 1\leq j_{1,1}^{(3)}<j_{2,1}^{(3)}<\cdots<j_{n_4,1}^{(3)}\leq s_n,\\ j_{q_3,1}^{(3)}<j_{q_3,2}^{(3)}<j_{q_3,3}^{(3)},\ q_3=1,\ldots,n_4}}\cdots$$

$$n(n-1)\cdots(n-s_n+M)(j_1-1)\cdots(\widehat{j_{l_1}-l_1})\cdots(\widehat{j_{l_r}-l_r})\cdots(j_{n_2}-n_2)$$

$$\times\prod_{k=1}^{r}x(t_{j_{l_k}})\cdot\prod_{q_2=1}^{n_3}(j_{q_2,1}^{(2)}-1)\cdot\prod_{q_3=1}^{n_4}(j_{q_3,1}^{(3)}-1)\cdots\frac{1}{n^{s_n}}\cdot\mathbf{E}\left(\prod_{u=0}^{2s_n-1}\xi_{i_ui_{u+1}}\right)W_n\,. \tag{4.25}$$

Here W_n is again the number of ways in which a trajectory can be continued at the unmarked instants if it is given at the marked instants. In [14], we have essentially used the following estimate:

$$W_n\leq\prod_{k=3}^{M}(2k)^{kn_k}3^r. \tag{4.26}$$

This estimate would suffice here if the random variables $\{\xi_{ij}\}$ were uniformly bounded. In the general case of the overexponential growth of the moments of ξ_{ij}, the factor $\mathbf{E}(\prod_{u=0}^{2s_n-1}\xi_{i_ui_{u+1}})$ may be large since so may be the contribution of the edges used many times, and we need a somewhat finer estimate for the last two factors in (4.25), which is proved below.

Lemma 1.

$$\mathbf{E}\biggl(\prod_{u=0}^{2s_n-1}\xi_{i_u i_{u+1}}\biggr)W_n \le \Bigl(\frac{1}{4}\Bigr)^{s_n}\prod_{k=3}^{M}(\mathrm{const}_3\cdot k)^{kn_k}3^r. \tag{4.27}$$

Proof of the lemma. In [14], we used the estimate (4.26) for W_n and the estimate

$$\mathbf{E}\biggl(\prod_{u=0}^{2s_n-1}\xi_{i_u i_{u+1}}\biggr) \le \Bigl(\frac{1}{4}\Bigr)^{s_n}\prod_{\text{the edges }\{i,j\}\text{ of }\mathcal{P}}(4\,\mathrm{const}\cdot l(\{i,j\}))^{l(\{i,j\})}, \tag{4.28}$$

where $\{i,j\}$ stands for an edge of $\mathcal{P}$ and $2l(\{i,j\})$ for the number of occurrences of the edge $\{i,j\}$ in $\mathcal{P}$. The estimate (4.26) follows from the fact that, for a vertex of k-fold intersection, there are at most $2k$ possibilities for each "return." The estimate (4.28) is a direct consequence of condition (1.1) imposed on the growth of the moments of the random variable $\{\xi_{ij}\}$. The meaning of Lemma 1 is that these two factors cannot be large simultaneously. If an edge is passed more than twice, say, $2l$ times ($2l > 2$), which results in the appearance of the factor $(4\,\mathrm{const}\cdot l)^l$ on the right-hand side in (4.28), then we can reduce the estimate for W_n on the right-hand side in (4.26) by a factor of $l!$ because the "returns" occur l times along the same edge. Thus,

$$\begin{aligned}\mathbf{E}\biggl(\prod_{u=0}^{2s_n-1}\xi_{i_u i_{u+1}}\biggr)W_n &\le \Bigl(\frac{1}{4}\Bigr)^{s_n}{\prod_{\{i,j\}}}'(4\,\mathrm{const}\cdot l(\{i,j\}))^{l(\{i,j\})}\frac{1}{l(\{i,j\})!}\prod_{k=2}^{M}(2k)^{kn_k}3^r\\ &\le \Bigl(\frac{1}{4}\Bigr)^{s_n}{\prod_{\{i,j\}}}'(\mathrm{const}_1)^{l(\{i,j\})}\prod_{k=3}^{M}(2k)^{kn_k}3^r\\ &\le \Bigl(\frac{1}{4}\Bigr)^{s_n}(\mathrm{const}_1)^{\sum_{k=3}^{M}kn_k}\prod_{k=3}^{M}(2k)^{kn_k}3^r.\end{aligned} \tag{4.29}$$

This is just the estimate from Lemma 1. In (4.29), the product $\prod'_{\{i,j\}}$ is taken over the edges $\{i,j\}$ such that $l(\{i,j\})\ge 4$.

The subsequent estimates for Z_3 are similar to those for Z_2 and Z_1. Namely, in (4.25) we consider the subsum

$$\begin{aligned}&\sum_{1\le j_1<\dots<j_{n_2}\le s_n}\ \sum_{1\le l_1<\dots<l_r\le n_2}\ \sum_{\substack{(j^{(2)}_{1,1},j^{(2)}_{1,2}),\dots,(j^{(2)}_{n_3,1},j^{(2)}_{n_3,2}),\\ 1\le j^{(2)}_{1,1}<\dots<j^{(2)}_{n_3,1}\le s_n,\\ j^{(2)}_{q_2,1}<j^{(2)}_{q_2,2},\ q_2=1,\dots,n_3}}\ \sum_{\substack{(j^{(3)}_{1,1},j^{(3)}_{1,2},j^{(3)}_{1,3}),\dots,\\ (j^{(3)}_{n_4,1},j^{(3)}_{n_4,2},j^{(3)}_{n_4,3}),\\ 1\le j^{(3)}_{1,1}<\dots<j^{(3)}_{n_4,1}\le s_n}}\cdots\\ &(j_1-1)\dots\widehat{(j_{l_1}-l_1)}\cdots\widehat{(j_{l_r}-l_r)}\cdots(j_{n_2}-n_2)\\ &\times\prod_{k=1}^{r}x(t_{j_{l_k}})\cdot\prod_{q_2=1}^{n_3}(j^{(2)}_{q_2,1}-1)\cdot\prod_{q_3=1}^{n_4}(j^{(3)}_{q_3,1}-1)\cdots\prod_{q_M=1}^{n_{M+1}}(j^{(M)}_{g_M,1}-1).\end{aligned} \tag{4.30}$$

The subsum (4.30) is majorized by

$$\begin{aligned}&\frac{1}{n_2!}\biggl(\sum_{j_1=1}^{s_n}(j_1-1)\biggr)^{n_2-r}\frac{1}{r!}\biggl(\sum_{j_2=1}^{s_n}1\biggr)^{r}\biggl(\sum_{l=1}^{M}\max_{0\le t\le 2s_n}x(t)\cdot 1\biggr)^{r}\\ &\quad\times\frac{1}{n_3!}\biggl(\frac{1}{2!}\sum_{j_3=1}^{s_n}(j_3-1)\cdot\sum_{j_4=1}^{s_n}1\biggr)^{n_3}\frac{1}{n_4!}\biggl(\frac{1}{3!}\sum_{j_5=1}^{s_n}(j_5-1)\cdot\sum_{j_6=1}^{s_n}1\cdot\sum_{j_7=1}^{s_n}1\biggr)^{n_4}\cdots\\ &\qquad\le\prod_{k=2}^{s_n}\frac{1}{n_k!}\biggl(\frac{s_n^k}{2}\,\frac{1}{(k-1)!}\biggr)^{n_k}\frac{1}{r!}\biggl(\frac{2M}{\sqrt{s_n}}\max\frac{x(t)}{\sqrt{s_n}}\biggr)^{r}.\end{aligned} \tag{4.31}$$

It follows from (4.31) and Lemma 1 that

$$
\begin{aligned}
Z_3 \le n\,\frac{(2s_n)!}{s_n!\,(s_n+1)!}\left(\frac{1}{4}\right)^{s_n}\sum_{M=1}^{10s_n^2/n}\ \sum_{n_2+2n_3+\cdots+Mn_{M+1}=M}\ \prod_{i=1}^{s_n-M}\left(1-\frac{i}{n}\right)\\
\times\frac{1}{n_2!}\left(\frac{s_n^2}{2n}\right)^{n_2}\cdot\prod_{k=3}^{s_n+1}\frac{1}{n_k!}\left(\frac{s_n^k}{2n^{k-1}}\,\frac{(\mathrm{const}_2\cdot k)^k}{(k-1)!}\right)^{n_k}\\
+\,n\,\frac{(2s_n)!}{s_n!(s_n+1)!}\left(\frac{1}{4}\right)^{s_n}\sum_{M=1}^{10s_n^2/n}\ \sum_{n_2+2n_3+\cdots+Mn_{M+1}=M}\ \prod_{i=1}^{s_n-M}\left(1-\frac{i}{n}\right)\\
\times\frac{1}{n_2!}\left(\frac{s_n^2}{2n}\right)^{n_2}\cdot\prod_{k=3}^{s_n}\frac{1}{n_k!}\left(\frac{s_n^k}{2n^{k-1}}\,\frac{\mathrm{const}_2\cdot k^k}{(k-1)!}\right)^{n_k}\mathbf{E}_X\left(\exp\left(\frac{60s_n^{3/2}}{n}\max_{0\le t\le 2s_n}\frac{x(t)}{\sqrt{s_n}}\right)-1\right)+o(1)\\
\le\frac{n}{\sqrt{\pi}s_n^{3/2}}\left\{\left(\exp\left(\sum_{k=3}^{s_n+1}\frac{s_n^k}{2n^{k-1}}\,\mathrm{const}_5^k\right)-1\right)\right.\\
\left.+\,\mathbf{E}_X\left(\exp\left(\frac{60s_n^{3/2}}{n}\max_{0\le t\le 2s_n}\frac{x(t)}{\sqrt{s_n}}\right)-1\right)\right\}=o\left(\frac{n}{s_n^{3/2}}\right).
\end{aligned}\tag{4.32}
$$

Thus, we have proved formulas (4.20). In conjunction with (4.17), this implies the following estimate for the leading term of the expectation:

$$
\mathbf{E}(\mathrm{Tr}\,A_n^{2s_n})=\frac{n}{\sqrt{\pi s_n^3}}(1+o(1))\qquad\text{for }\ s_n\ll n^{2/3}.\tag{4.33}
$$

As a by-product, we obtain the following estimate for the number of even paths of length $2s_n$.

Proposition 2. *Let $1\ll s_n\ll n^{2/3}$. In this case, the number of even paths of length $2s_n$ is equal to* $n^{s_n+1}4^{s_n}\pi^{-1/2}s_n^{-3/2}(1+o(1))$.

§5. Estimates for the Variance and for the Higher Moments of the Random Variable $\mathrm{Tr}\,A_n^{p_n}$

The treatment of the variance and the higher moments of $\mathrm{Tr}\,A^{p_n}$ essentially follows [14]. For the reader's convenience, we give here the main ideas of the proof.

Proposition 3. *Let $1\ll p_n\ll n^{2/3}$. In this case,* $\mathrm{Var}(\mathrm{Tr}\,A_n^{p_n})\le\mathrm{const}_6$ *for all n, and we have the relation* $\mathrm{Var}(\mathrm{Tr}\,A_n^{p_n})\to 1/\pi$ *as $n\to\infty$, $p_n\to\infty$, and $p_nn^{-2/3}\to 0$.*

The formula for the variance of the random variable $\mathrm{Tr}\,A^{p_n}$ reads

$$
\begin{aligned}
\mathrm{Var}(\mathrm{Tr}\,A_n^{p_n}) &= \mathbf{E}(\mathrm{Tr}\,A_n^{p_n})^2-(\mathbf{E}(\mathrm{Tr}\,A_n^{p_n}))^2\\
&=\sum_{i_0,i_1,\dots,i_{p_n-1}=1}^{n}\ \sum_{j_0,j_1,\dots,j_{p_n-1}=1}^{n}\frac{1}{n^{p_n}}\cdot\left(\mathbf{E}\left(\prod_{l=1}^{p_n}\xi_{i_{l-1}i_l}\cdot\prod_{m=1}^{p_n}\xi_{j_{m-1}j_m}\right)\right.\\
&\qquad\left.-\,\mathbf{E}\left(\prod_{l=1}^{p_n}\xi_{i_{l-1}i_l}\right)\cdot\mathbf{E}\left(\prod_{m=1}^{p_n}\xi_{j_{m-1}j_m}\right)\right),
\end{aligned}\tag{5.1}
$$

where it is assumed that $i_{p_n}=i_0$ and $j_{p_n}=j_0$.

Since the symmetrically distributed random variables ξ_{ij}, $i\le j$, are independent, it follows that most of the terms in (5.1) vanish. The term corresponding to a pair

$$
\mathcal{P}=\{i_0\to i_1\to\cdots\to i_{p_n-1}\to i_0\},\qquad \mathcal{P}'=\{j_0\to j_1\to\cdots\to j_{p_n-1}\to j_0\}
$$

of closed paths of length p_n is nonzero if and only if

(a) the paths $\mathcal{P}$ and $\mathcal{P}'$ have at least one common (nondirected) edge;

(b) the number of occurrences of each edge in the union $\mathcal{P}\cup\mathcal{P}'$ is even (here the union $\mathcal{P}\cup\mathcal{P}'$ is understood as the union of the sets of edges of $\mathcal{P}$ and $\mathcal{P}'$ with regard to the multiplicities of their passage).

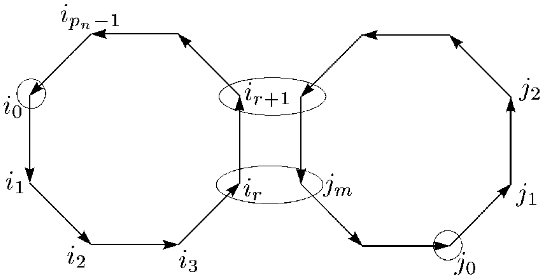

Fig. 1

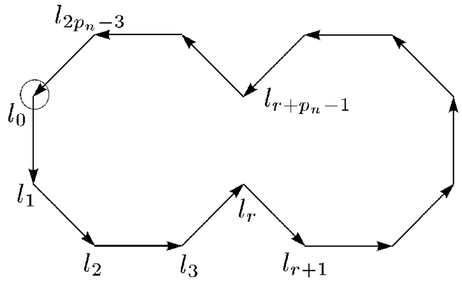

Fig. 2

Definition 5. An ordered pair of paths satisfying conditions (a) and (b) is said to be *correlated.* A correlated pair is said to be *simply correlated* if each edge occurs in the union $\mathcal{P} \cup \mathcal{P}'$ exactly twice.

The computation of $\mathrm{Var}(\mathrm{Tr}\, A^{p_n})$ is essentially reduced to that of the number of correlated pairs of paths counted according to the corresponding statistical weight. In particular, our argument will prove the following assertion.

Proposition 4. *Let $1 \ll p_n \ll n^{2/3}$. In this case, the number of correlated pairs of paths in the leading term is equal to the number of simply correlated pairs of paths and coincides with $\pi^{-1} n^{p_n} 2^{2p_n}(1 + o(1))$.*

The technique developed in §4 permits one to find the leading term of the number of even paths of length $2s_n$. Indeed, if the events $\xi_{ij} = \pm 1/2$ for the random variables ξ_{ij}, $i \le j$, are equiprobable, then

$$\mathbf{E}(\mathrm{Tr}\, A^{2s_n}) = 4^{s_n} n^{s_n} \cdot (\text{the number of even paths of length } 2s_n),$$

and Proposition 2 follows from (4.33).

In what follows, to each correlated pair of paths of length p_n, we assign an even path of length $2p_n - 2$. This correspondence is not one-to-one, and each path of length $2p_n - 2$ has several preimages in general. As a result, the calculation of the number of the correlated pairs of paths is reduced to that of the number of even paths counted according to the multiplicities equal to the numbers of the corresponding preimages. The construction of the above correspondence is illustrated in Figs. 1 and 2.

Definition 6. By the *first common edge* of an ordered correlated pair we mean the first edge in $\mathcal{P}$ that belongs also to $\mathcal{P}'$.

The new path of length $2p_n - 2$ will consist of edges of the paths $\mathcal{P}$ and $\mathcal{P}'$and will be denoted by $\mathcal{P} \vee \mathcal{P}'$. The path $\mathcal{P} \vee \mathcal{P}'$ is constructed as follows: first we go along $\mathcal{P}$ to the left end of the first common edge of $\mathcal{P}$ and $\mathcal{P}'$, then jump to $\mathcal{P}'$ and make $p_n - 1$ steps on $\mathcal{P}'$ (along $\mathcal{P}'$ if $\mathcal{P}$ and $\mathcal{P}'$ have opposite direction on the common edge and in the direction opposite to that of $\mathcal{P}'$ if $\mathcal{P}$ and $\mathcal{P}'$ have the same direction on the common edge). After $p_n - 1$ steps on $\mathcal{P}'$, we again arrive at the common edge of $\mathcal{P}$ and $\mathcal{P}'$. Then we jump to $\mathcal{P}$ and finish the path $\mathcal{P} \vee \mathcal{P}'$ by going along the first path. The path

$$\mathcal{P} \vee \mathcal{P}' = \{l_0 = i_0 \to \cdots \to l_r = i_r = j_m \to l_{r+p_n-1} = i_{r+1} = j \to \cdots \to l_{2p_n-3} \to l_0 = i_0\}$$

thus obtained, where $\{i_r, i_{r+1}\}$ is a common edge of the paths $\mathcal{P}$ and $\mathcal{P}'$, is even because the pair $(\mathcal{P}, \mathcal{P}')$ is correlated.

Now let us estimate the number of ways in which a given path of length $2p_n - 2$ can be obtained from various correlated pairs. First, we must choose some vertex l_r in the first half of the path, $0 \le r \le p_n - 1$, and join it with the vertex l_{r+p_n-1}. The choice is made so that the edge $\{l_r, l_{r+p_n-1}\}$ is the first common edge of $\mathcal{P}$ and $\mathcal{P}'$. Moreover, we must choose the beginning of the second path $\mathcal{P}'$, which can be done in p_n ways in a typical situation, and the direction of motion along $\mathcal{P}'$, which can be done in two ways. It is convenient to restate the condition that $\{l_r, l_{r+p_n-1}\}$ should be the first common edge of the paths $\mathcal{P}$ and $\mathcal{P}'$ in terms of the random walk

$$X = \{x(t) \ge 0,\ 0 \le t \le 2p_n - 2,\ x(0) = x(2p_n - 2) = 0\}$$

on the positive half-line. Recall that $x(t) - x(t-1) = 1$ if the number of occurrences of the edge $\{l_{t-1}, l_t\}$ in the first t steps is odd, and $x(t) - x(t-1) = -1$ otherwise. In this case, a necessary condition for the edge $\{l_r, l_{r+p_n-1}\}$ to be the first common edge of the pair $(\mathcal{P}, \mathcal{P}')$ is that for $r \le t \le r + p_n - 1$ (half the total time of the walk) the trajectory $x(t)$ should not fall below $x(r)$. For typical paths (i.e., those with simple self-intersections and without nonclosed vertices) this condition is also sufficient. Let $K_n(x(\cdot))$ be the number of instants r_j, $0 \le r_j \le p_n - 1$, $j = 1, \dots, K_n$, such that $x(t) \ge x(r_j)$ for $r_j \le t \le r_j + p_n + 1$. It was shown in [14, Lemma 1] that

$$\mathbf{E}_X K_n = 2\sqrt{\frac{p_n}{\pi}}\,(1 + o(1)). \tag{5.2}$$

We have shown in §4 that the uniform distribution on the space of closed even paths of length $2p_n - 2$ induces a distribution on the space of trajectories of the random walk such that the latter distribution tends to a uniform distribution as $n \to \infty$. Following the lines of §4, we can readily show that

$$(\text{the number of correlated pairs of length } p_n) = n^{p_n} \frac{(2p_n - 2)!}{(p_n - 1)!\,p_n!}\,\mathbf{E}_X(K_n(x(\cdot))) \cdot 2p_n \cdot (1 + o(1)). \tag{5.3}$$

Proposition 4 follows from (5.2) and (5.3). When calculating $\operatorname{Var}(\operatorname{Tr} A_n^{p_n})$, one must also take into account the statistical weight

$$\frac{1}{n^{p_n}}\left(\mathbf{E}\left(\prod_{l=1}^{p_n} \xi_{i_{l-1}i_l} \cdot \prod_{m=1}^{p_n} \xi_{j_{m-1}j_m}\right) - \mathbf{E}\left(\prod_{l=1}^{p_n} \xi_{i_{l-1}i_l}\right) \cdot \mathbf{E}\left(\prod_{m=1}^{p_n} \xi_{j_{m-1}j_m}\right)\right). \tag{5.4}$$

The factor (5.4) is equal to 2^{-2p_n} provided that the following two conditions are satisfied:

(a) the path $\mathcal{P} \vee \mathcal{P}'$ contains each edge exactly twice;

(b) the path $\mathcal{P} \vee \mathcal{P}'$ does not contain the unordered edge $\{l_r, l_{r+p_n-1}\}$.

The proof of the fact that the paths satisfying conditions (a) and (b) are typical and make the main contribution to $\operatorname{Var}(\operatorname{Tr} A^{p_n})$ essentially reproduces the argument in §4. The proof of Proposition 3 is complete.

The case of higher moments can also be considered in a manner similar to that in [14]. To prove Theorem 2, we must show that

$$\mathbf{E}(\operatorname{Tr} A_n^{p_n} - \mathbf{E}(\operatorname{Tr} A_n^{p_n}))^{2k} = \frac{(2k-1)!!}{\pi^k} + o(1), \tag{5.5}$$

$$\mathbf{E}(\operatorname{Tr} A_n^{p_n} - \mathbf{E}(\operatorname{Tr} A_n^{p_n}))^{2k+1} = o(1). \tag{5.6}$$

The following identity holds, which is similar to (5.1):

$$\mathbf{E}(\operatorname{Tr} A_n^{p_n} - \mathbf{E}(\operatorname{Tr} A_n^{p_n}))^L = \frac{1}{n^{p_n L/2}}\,\mathbf{E}\prod_{m=1}^{L}\left(\sum_{i_0^{(m)}, i_1^{(m)}, \dots, i_{p_n-1}^{(m)}=1}^{n}\left(\prod_{r=1}^{p_n} \xi_{i_{r-1}^{(m)} i_r^{(m)}} - \mathbf{E}\prod_{r=1}^{p_n} \xi_{i_{r-1}^{(m)} i_r^{(m)}}\right)\right). \tag{5.7}$$

Let us consider an arbitrary set of L closed paths

$$\mathcal{P}_m = \{i_0^{(m)} \to i_1^{(m)} \to \dots \to i_{p_n}^{(m)} = i_0^{(m)}\}, \qquad m = 1, \dots, L,$$

of length p_n.

Definition 7. We say that paths $\mathcal{P}_{m'}$ and $\mathcal{P}_{m''}$ *intersect each other by an edge* if $\mathcal{P}_{m'}$ and $\mathcal{P}_{m''}$ have a common (nondirected) edge.

Definition 8. A subset $\mathcal{P}_{m_{j_1}}, \mathcal{P}_{m_{j_2}}, \dots, \mathcal{P}_{m_{j_k}}$ of the set of paths is called a *cluster of intersecting paths* if the following conditions hold:

(a) for each pair $\mathcal{P}_{m_i}$, $\mathcal{P}_{m_j}$ there exists a chain of paths from this subset such that $\mathcal{P}_{m_i}$ is the first path in the chain, $\mathcal{P}_{m_j}$ is the last path in the chain, and any two neighboring paths intersect each other by an edge;

(b) property (a) is violated if we add an arbitrary new path to this subset.

It follows from the definition that the sets of edges corresponding to different clusters are disjoint. This, in conjunction with the independence of the random variables $\{\xi_{ij}\}_{l\le j}$, implies that the expectation in (5.7) can be represented as the product of the expectations of the factors corresponding to various clusters. Formulas (5.5) and (5.6) follow from the fact that the main contribution to (5.7) corresponds to the situation in which each of the clusters contains exactly two paths. (Obviously, if at least one cluster consists of a single path, then the expectation of the corresponding term is zero.)

Lemma 2.

$$\mathbf{E}\frac{1}{n^{p_n l/2}}\prod_{m=1}^{l}{}^{*}\Bigg(\sum_{i_0^{(m)},\dots,i_{p_n-1}^{(m)}=1}^{n}\Bigg(\prod_{r=1}^{p_n}\xi_{i_{r-1}^{(m)}i_r^{(m)}}-\mathbf{E}\Bigg(\prod_{r=1}^{p_n}\xi_{i_{r-1}^{(m)}i_r^{(m)}}\Bigg)\Bigg)\Bigg)=\begin{cases}1/\pi+o(1) & \textit{for } l=2,\\ o(1) & \textit{for } l>2,\end{cases} \tag{5.8}$$

where the product $\prod^*$ *in* (5.8) *is taken over the paths forming a cluster.*

The proof of Lemma 2 is given in [14]. Note that the case $l=2$ corresponds to Proposition 3 of the present paper. The case $l>2$ can be treated in a similar way. To each cluster of l paths we assign an even path of length $l_{p_n}-q$, where $l\le q\le 2l$. The number of preimages under this correspondence is bounded above by K_n^{l-1}, where K_n is the number of instants t_i, $i=1,\dots,K_n$, such that $t_i\le(l-1)p_n-q$ and $x(t)\ge x(t_i)$ for $t_i\le t\le t_i+p_n-1$. The corresponding analog of (5.2) is given by the inequality

$$\mathbf{E}_X K_n^{l-1}\le \mathrm{const}_l\cdot p_n^{(l-1)/2}.$$

Appendix

In this Appendix, we prove that the sum over paths in which some edge is passed at least four times is small compared with the total sum (4.1).

Let $\mathcal{P}=\{i_0\to i_1\to\cdots\to i_{2s_n}=i_0\}$, and let $\nu_n(\mathcal{P})$ be the maximum number of vertices into which one can get at marked instants from a given vertex. By definition, each vertex i of the path $\mathcal{P}$ is the left end of at most $\nu_n(\mathcal{P})$ marked edges. In what follows, we show that, for typical paths, $\nu_n(\mathcal{P})$ grows slower than any positive power of s_n.

Lemma 3. *For each* γ, $0<\gamma<1$, *the subsum in* (4.1) *over the paths* $\mathcal{P}$ *such that* $\nu_n(\mathcal{P})>s_n^\gamma$, $0<\gamma<1$, *is* $o(1)$ *and hence is negligibly small compared with the total sum.*

To understand why $\nu_n(\mathcal{P})$ cannot be large for typical paths, let us consider a path with simple self-intersections and without nonclosed vertices. In this case, if a vertex is the left end of ν_n marked edges, then, for the corresponding random walk $x(t)$, $0\le t\le 2s_n$, on the positive half-line, there exists a time interval $[t_1,t_2]$ on which the trajectory $x(t)$ falls $\nu_n/2$ times into the level $x(t_1)$ but never falls below this level. Indeed, to make each new step from the vertex i_{t_1}, we must first return to this vertex along the trajectory, and the coefficient $1/2$ of ν_n responds to the fact that i_{t_1} can be a vertex of self-intersection. Obviously, the probability of such trajectories of the random walk decays exponentially as $\nu_n\to\infty$, that is, there exists a constant const_7 such that the fraction of such trajectories does not exceed

$$(2s_n)^2 e^{-\mathrm{const}_7\cdot\nu_n}. \tag{A1}$$

Let η_n be the number of edges passed four times by a path with simple self-intersections. In this case,

$$\mathbf{E}\prod_{u=0}^{2s_n-1}\xi_{i_u i_{u+1}}\le\left(\frac{1}{4}\right)^{s_n}(8\,\mathrm{const})^{2\eta_n}.$$

For a given ν_n, we can readily write out an upper bound both for the number of edges passed four times and for the quantity $(8\,\mathrm{const})^{2\eta_n}$. We still denote the total number of instants of self-intersection by M. As was shown in §4, it suffices to consider the case $M \le 10 s_n^2/n$. For each of the M instants of self-intersection, the number of possible choices of the vertex of self-intersection that is an end of a fourfold edge does not exceed ν_n; at the same time, the number of all possible choices of the vertex of self-intersection is of the order of s_n in the general case. Thus, the mean value of η_n is of the order of $(\nu_n/s_n)M = O((\nu_n/\sqrt{s_n})(s_n^{3/2}/n))$. A similar argument shows that the mean value of $(8\,\mathrm{const})^{2\eta_n}$ does not exceed $s_n^{1/2}(8\,\mathrm{const})^{(\nu_n/\sqrt{s_n})(20 s_n^{3/2}/n)}$. As a result, we see that the subsum of Z_1 over the paths with fourfold intersections and with $\nu_n > s_n^\gamma$ behaves as

$$Z_1 \cdot O\Big(\max_{s_n^\gamma < \nu_n \le s_n} \Big\{ (8\,\mathrm{const})^{(20 s_n^{3/2}/n)(\nu_n/\sqrt{s_n})} (2s_n)^2 e^{-\,\mathrm{const}_7 \cdot \nu_n} \Big\}\Big) = o(Z_1)\,.$$

For an arbitrary even path, let N_n be the number of unmarked instants for which the choice of the continuation ("return") of the trajectory is ambiguous. By definition, $N_n \le r + \sum_{k=3}^{s_n} k n_k$. It follows from the reasoning in §4 that the probability of choosing the vertices of the path so that $N_n > N$ is $O((\mathrm{const}_5 \cdot s_n/n^{2/3})^N)$ uniformly with respect to N. The above N_n instants divide the interval $[0, 2s_n]$ into $N_n + 1$ subintervals such that at least one of them contains a subsubinterval $[t_1, t_2]$ on which the trajectory falls at least $\nu_n/(N_n+1)$ times into the level $x(t_1)$ but never falls below it. The fraction of such trajectories $x(\cdot)$ does not exceed $(2s_n)^2 e^{-\,\mathrm{const}_7 \cdot \nu_n/(N_n+1)}$ (cf. (A1)), and by carrying out computations similar to those used in §4 for Z_1, Z_2, and Z_3, we find that the subsum over the paths with $\nu_n > s_n^\gamma$ is bounded above by

$$Z_1 \cdot \left(\sum_{\nu_n = s_n^\gamma}^{s_n} \sum_{N_n=0}^{s_n} \left(\frac{\mathrm{const}_5 \cdot s_n}{n^{2/3}} \right)^{N_n} e^{-\,\mathrm{const}_7 \cdot \nu_n/(N_n+1)} (8\,\mathrm{const})^{\frac{\nu_n}{\sqrt{s_n}} \frac{20 s_n^{3/2}}{n}} \right). \tag{A2}$$

Elementary computations show that the second factor in (A2) is $o(1)$. The proof of Lemma 3 is complete.

The sum over the paths with edges of multiplicity ≥ 4 and with $\nu_n < s_n^\gamma$ was studied in §4 (see the estimates for Z_1' and Z_2'), where it was shown that this sum is also $o(Z_1)$.

References

1. E. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," Ann. Math., **62**, 548–564 (1955).
2. E. Wigner, "On the distribution of the roots of certain symmetric matrices," Ann. Math. **67**, 325–328 (1958).
3. V. A. Marchenko and L. A. Pastur, "The distribution of the eigenvalues in some ensembles of random matrices," Mat. Sb., **72**, 507–536 (1967).
4. L. A. Pastur, "On the spectrum of random matrices," Teor. Mat. Fiz., **10**, 102–112 (1972).
5. L. A. Pastur, "The spectra of random self-adjoint operators," Usp. Mat. Nauk, **28**, 3–64 (1973).
6. L. Arnold, "On the asymptotic distribution of the eigenvalues of random matrices," J. Math. Anal. Appl., **20**, 262–268 (1967).
7. L. Arnold, "On Wigner's semicircle law for the eigenvalues of random matrices," Z. Wahrscheinlichkeitstheorie Verw. Gebiete, **19**, 191–198 (1971).
8. K. W. Wachter, "The strong limits of random matrix spectra for sample matrices of independent elements," Ann. Probab., **6**, No. 1, 1–18 (1978).
9. V. L. Girko, The Spectral Theory of Random Matrices [in Russian], Nauka, Moscow, 1988.
10. Z. D. Bai, "Convergence rate of expected spectral distributions of large random matrices. I. Wigner Matrices," Ann. Probab., **21**, 625–648 (1993).
11. 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," in: Phase Transitions and Critical Phenomena, Vol. 15 (C. Domb and J. L. Lebowitz, eds.), Academic Press, London, 1992.
12. A. M. Khorunzhy, B. A. Khoruzhenko, and L. A. Pastur, "Asymptotic properties of large random matrices with independent entries," J. Math. Phys., **37**, 5033–5059 (1996).

13. A. Khorunzhy, "On smoothed density of states for Wigner random matrices," Random Oper. Stochastic Equations, **5**, No. 2, 147–162 (1997).
14. Ya. Sinai and A. Soshnikov, "Central limit theorem for traces of large random symmetric matrices with independent matrix elements," Bol. Soc. Brasil. Mat., **29**, No. 1, 1–24 (1998).
15. C. Tracy and H. Widom, "On orthogonal and symplectic matrix ensembles," Commun. Math. Phys., **177**, 727–754 (1996).
16. C. Tracy and H. Widom, "Level-spacing distribution and the Airy kernel," Commun. Math. Phys., **159**, 151–174 (1994).
17. P. J. Forrester, "The spectrum edge of random matrix ensembles," Nucl. Phys. B, **402**, 709–728 (1994).
18. Z. Füredi and J. Komlóz, "The eigenvalues of random symmetric matrices," Combinatorica, **1**, No. 3, 233–241 (1981).
19. A. Boutet de Monvel and M. V. Shcherbina, "On the norm of random matrices," Mat. Zametki, **57**, No. 5, 688–698 (1995).
20. M. L. Mehta, Random Matrices, Academic Press, New York, 1991.
21. O. Costin and J. L. Lebowitz, "Gaussian fluctuations in random matrices," Phys. Rev. Lett., **75**, No. 1, 69–72 (1995).

Translated by V. E. Nazaikinskii