The Annals of Probability
1999, Vol. 27, No. 2, 913–944

ON THE NORM AND EIGENVALUE DISTRIBUTION OF LARGE RANDOM MATRICES

BY ANNE BOUTET DE MONVEL AND ALEXEI KHORUNZHY[1]

Université Paris 7 and Verkin Institute

We study the eigenvalue distribution of $N \times N$ symmetric random matrices $H_N(x, y) = N^{-1/2} h(x, y)$, $x, y = 1, \dots, N$, where $h(x, y)$, $x \le y$ are Gaussian weakly dependent random variables. We prove that the normalized eigenvalue counting function of H_N converges with probability 1 to a nonrandom function $\mu(\lambda)$ as $N \to \infty$. We derive an equation for the Stieltjes transform of the measure $d\mu(\lambda)$ and show that the latter has a compact support Λ_μ. We find the upper bound for $\limsup_{N\to\infty} \|H_N\|$ and study asymptotically the case when there are no eigenvalues of H_N outside of Λ_μ when $N \to \infty$.

1. Introduction. The first studies of $N \times N$ random matrices date back to the works in multivariate statistical analysis of the thirties and forties [see, e.g., the monograph by Anderson (1984)]. In the early fifties, Wigner used random matrices (RM) in nuclear physics, where the asymptotic behavior for large-N of the eigenvalue statistics plays an important role [see the collection of papers edited by Porter (1965)]. At present, random matrices are of great interest because of their applications in various fields of theoretical physics and also because of their rich mathematical content [see, e.g., the monographs and reviews by Cohen, Kesten and Newman (1986), Crisanti, Paladin and Vulpiani (1993), Bougerol and Lacroix (1985), Di Francesco, Ginsparg and Zinn-Justin (1995), Mehta (1991) and Voiculescu, Dykema and Nica (1992)].

In the RM theory, the following two classes of ensembles of Hermitian (or real symmetric) matrices have been most studied:

1. Ensembles of random matrices with jointly independent entries.
2. Ensembles of $N \times N$ matrices whose probability distribution is invariant with respect to the unitary (or orthogonal) transformations of $\mathbb{C}^N$ (or $\mathbb{R}^N$).

These two classes of ensembles can be regarded as different generalizations of the Gaussian unitary (or orthogonal) ensemble that plays a fundamental role in the RM theory [see, e.g., the monograph by Mehta (1991) and references therein]. This ensemble consists of Hermitian (or real symmetric, respectively) $N \times N$ matrices H_N whose entries $H_N(x, y)$, $1 \le x \le y \le N$ are independent Gaussian random variables with zero mathematical expectation

---

Received September 1997; revised August 1998.
[1]Supported in part by CNRS (France).
AMS 1991 *subject classifications*. Primary 60F15; secondary 15A18, 15A52.
Key words and phrases. Random matrices, eigenvalue distribution, spectral norm, limiting theorems.

and variance v^2 for $x \neq y$ and $2v^2$ for $x = y$ (for brevity we will consider the orthogonal case called GOE).

Ensembles (1) and (2) are opposite generalizations of the GOE in the sense that the matrices of (1) have independent arbitrarily distributed entries, while those of (2) possess a rather strong statistical dependence that does not decay even when the entries are far enough from one another in the matrix. Thus, it is natural to expect that the eigenvalue distributions of matrices (1) and (2) have different properties [see, e.g., Wigner (1955) and Boutet de Monvel, Pastur and Shcherbina (1995), respectively].

The present paper deals with ensembles that can be regarded as intermediate between (1) and (2). More precisely, we consider the case when the entries $H_N(x, y)$ of a symmetric $N \times N$ matrix H_N are weakly dependent Gaussian random variables; that is, we assume that the correlations between them vanish as the "distance" increases. This distance between $H_N(x, y)$ and $H_N(x', y')$, $x \le y$, $x' \le y'$ can be defined as the sum $|x - x'| + |y - y'|$.

We are interested in the asymptotic behavior as $N \to \infty$ of the spectral norm and eigenvalue distribution of H_N. These two characteristics are basic in spectral RM theory and play an important role in many applications [see, e.g., Bovier, Gayrard and Picco (1995), Crisanti, Paladin and Vulpiani (1993) and Isopi and Newman (1992)].

Given an $N \times N$ real symmetric matrix H_N, the eigenvalue distribution is described by the normalized eigenvalue counting function (NCF)

$$\sigma(\lambda; H_N) = \#\left\{\lambda_j^{(N)} \le \lambda\right\} N^{-1}, \qquad \lambda_1^{(N)} \le \cdots \le \lambda_N^{(N)}, \tag{1.1}$$

where $\lambda_j^{(N)} \equiv \lambda_j(H_N)$ are the eigenvalues of H_N. The spectral norm of H_N is defined as

$$\|H_N\| = \max_j \left|\lambda_j(H_N)\right|.$$

An important result of the RM theory is the semicircle law derived by Wigner (1955). For the case of GOE it can be formulated as follows. Consider the ensemble of real symmetric random matrices H_N with entries

$$H_N(x, y) = \frac{1}{\sqrt{N}} h(x, y), \qquad x, y = 1, \dots, N, \tag{1.2}$$

where $h(x, y)$, $x \le y$, $x, y \in \mathbb{N}$ are jointly independent random variables defined on the same probability space. Assume that the family $\{h(x, y)\}$ has a Gaussian distribution and satisfies the conditions

$$\mathbf{E} h(x, y) = 0 \tag{1.3a}$$

and

$$\mathbf{E} h(x, y) h(s, t) = v^2\left[\delta(x - s)\delta(y - t) + \delta(x - t)\delta(y - s)\right], \tag{1.3b}$$

where the symbol **E** denotes the mathematical expectation and δ is the Kronecker symbol:

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

Then $\sigma(\lambda; H_N)$ weakly converges with probability 1, when $N \to \infty$, to a nonrandom nondecreasing function $\sigma_v(\lambda)$,

$$\lim_{N\to\infty} \sigma(\lambda; H_N) = \sigma_v(\lambda), \tag{1.4a}$$

$$\sigma_v'(\lambda) = \begin{cases} (2\pi v^2)^{-1}\sqrt{4v^2 - \lambda^2}, & \text{if } |\lambda| \le 2v, \\ 0, & \text{if } |\lambda| > 2v. \end{cases} \tag{1.4b}$$

Here and below, by weak convergence of NCF's $\sigma(\lambda; H_N)$ we mean the weak convergence of the measures $d\sigma(\lambda; H_N)$ that are associated in a natural way with nonnegative nondecreasing functions.

It follows from (1.4) that $(-2v, 2v)$ is the support of the measure $d\sigma(\lambda)$. This means that the number $n(N)$ of eigenvalues that fall into this interval is proportional to N and $\lim_{N\to\infty} n(N)N^{-1} = 1$. In fact, a stronger statement is valid: with probability 1 all eigenvalues of H_N fall into this interval, because, according to the results of Bronk (1964) and Geman (1980),

$$\lim_{N\to\infty} \|H_N\| = 2v \quad \text{with probability } 1. \tag{1.5}$$

Analogous facts are known also for ensembles of random matrices with independent entries more general than GOE. Namely, the semicircle law (1.4) is valid for the ensemble of matrices H_N (1.2), where $h(x, y)$, $x \le y$ are arbitrary i.i.d. random variables satisfying (1.3) [see Pastur (1973) and Girko (1975) for the sufficiency and necessity of these conditions, respectively, and Girko (1988) for more details]. This ensemble is known as the Wigner ensemble of random matrices.

Relation (1.5) was also shown to be true for the Wigner ensemble. In this case, the condition $\mathbf{E}|h(x, y)|^4 < \infty$ is a sufficient and necessary one [see, e.g., Bai and Yin (1988) for a more general formulation]. Under more restrictive conditions on the probability distribution of $h(x, y)$, Boutet de Monvel and Shcherbina (1995) derived that the exponential bound

$$\begin{aligned} &\mathrm{Prob}\{\|H_N\| > 2v(1+\varepsilon)\} \\ &\quad = \exp\{-N^{\tau}\log(1+\varepsilon)(1+o(1))\}, \qquad N \to \infty \end{aligned} \tag{1.6}$$

is valid for any given fixed $\varepsilon > 0$ with some positive τ. A similar estimate follows from deep results recently obtained for the Wigner ensemble by Sinai and Soshnikov (1998). Relation (1.6) can be regarded as a generalization of estimates derived by Bronk (1964) for GUE and GOE.

In the present paper, our main goal is to understand how the statistical dependence between entries of random matrix H_N can change the limiting behavior of the spectral norm and eigenvalue distribution. Under rather natural conditions, we find the upper bound for $\limsup_{N\to\infty}\|H_N\|$, derive the estimate (1.6), and prove that the nonrandom limit $\mu(\lambda)$ of $\sigma(\lambda; H_N)$ exists. We show that in general the limit of $\|H_N\|$ and the upper (lower) bound of the support of $d\mu(\lambda)$ do not coincide. We prove that this coincidence takes place for a certain class of random matrices that includes the important particular case when the random field $\{H_N(x, y)\}$ can be regarded as a stationary one.

The paper is organized as follows. In Section 2 we formulate our main statements. In Section 3 we prove the theorems of Section 2 with the help of one key technical result (Theorem 3.1). Section 4 is devoted to the proof of Theorem 3.1. In Section 5 we formulate and prove auxiliary statements.

2. Main results and discussion. Let V be a nonrandom bounded symmetric nonnegative operator in $l^2(\mathbb{N})$ with real entries $V(x, y)$, $x, y \in \mathbb{N}$. Let $h(x, y)$, $x \le y$, $x, y \in \mathbb{N}$ be random variables defined on the same probability space Ω. We assume that the joint distribution of $\{h(x, y)\}$ is Gaussian with the following properties [cf. (1.3)]:

$$\mathbf{E}h(x, y) = 0 \tag{2.1a}$$

and

$$\mathbf{E}h(x, y)h(s, t) = V(x, s)V(y, t) + V(x, t)V(y, s). \tag{2.1b}$$

One can easily show that the matrix $C(x, y; s, t) = V(x, s)V(y, t) + V(x, t)V(y, s)$ is nonnegatively defined and therefore satisfies the covariance criterion [Loève (1978)]. We prove this in Lemma 5.9 of Section 5.

We introduce random symmetric $N \times N$ matrices H_N and V_N by the relations

$$H_N(x, y) = N^{-1/2}\begin{cases} h(x, y), & \text{if } x \le y, \\ h(y, x), & \text{if } x > y, \end{cases} \qquad x, y = 1, \dots, N \tag{2.2}$$

and $V_N(x, y) = V(x, y)$, $x, y = 1, \dots, N$ and define the spectral norm of H_N as

$$\|H_N\| = \max_{j=1,2,\dots,N} \left| \lambda_j^{(N)} \right|,$$

where $\lambda_j^{(N)}$ are the eigenvalues of H_N.

THEOREM 2.1. *Denote*

$$\tilde{V} \equiv \|V\|_{l^2} < \infty, \qquad v_1^{(+)} = \limsup_{N\to\infty} \frac{1}{N} \operatorname{Tr} V_N \le \tilde{V} \tag{2.3}$$

and assume that

$$v_1^{(-)} = \liminf_{N\to\infty} \frac{1}{N} \operatorname{Tr} V_N > 0, \tag{2.4}$$

where Tr *denotes the trace of a matrix. Then the inequality*

$$\limsup_{N\to\infty} \|H_N\| \le 2\sqrt{v_1^{(+)}\tilde{V}} \tag{2.5}$$

holds with probability 1.

REMARK. Using the technique developed to prove Theorem 2.1, it is possible to consider the ensembles (2.1) and (2.2) such that [cf. (2.4)]

$$N^{-1}\operatorname{Tr} V_N = O(N^{-\tau}), \qquad N\to\infty \text{ with some } \tau>0. \tag{2.6}$$

In this case, the upper bound for the norm $\|H_N\|$ remains the same as in (2.5), but the estimate for probability $\operatorname{Prob}\{\|H_N\| > 2\sqrt{v_1^{(+)}\tilde{V}}(1+\varepsilon)\}$ that we derive is changed [see (2.23) and proof of Theorem 3.1 in Section 4].

The quantities $v_1^{(-)}$ and $v_1^{(+)}$ coincide in an important particular case where the limit of the normalized eigenvalue counting function (1.1) of V_N exists,

$$\nu(\lambda) = \lim_{N\to\infty} \sigma(\lambda; V_N). \tag{2.7}$$

Assuming (2.7), we can study the limiting eigenvalue distribution of H_N in more detail.

THEOREM 2.2. *Let V satisfies conditions* (2.3) *and* (2.7), *then*:

(i) *There exists a nonrandom function $\mu(\lambda)$ such that*

$$\lim_{N\to\infty} \sigma(\lambda; H_N) = \mu(\lambda) \tag{2.8}$$

with probability 1.

(ii) *The Stieltjes transform $f(z)$ of $d\mu(\lambda)$,*

$$f(z) = \int_{-\infty}^{\infty} (\lambda - z)^{-1}\, d\mu(\lambda), \qquad z\in\mathbb{C}\setminus\mathbb{R},$$

can be found from the relation

$$f(z) = \int_0^{\infty} \frac{d\nu(\lambda)}{-z-\lambda g(z)}, \tag{2.9a}$$

where $g(z)$ is a solution of the equation

$$g(z) = \int_0^{\infty} \frac{\lambda\, d\nu(\lambda)}{-z-\lambda g(z)}, \qquad z\in\mathbb{C}\setminus\mathbb{R}. \tag{2.9b}$$

This equation is uniquely solvable in the class $\mathbf{F}$ *of functions $\phi(z)$ analytic in $z\in\mathbb{C}\setminus\mathbb{R}$ and satisfying the conditions*

$$\lim_{\eta\to\infty} \eta\phi(i\eta) < \infty, \qquad \operatorname{Im}\phi(z)\operatorname{Im} z > 0, \qquad z\in\mathbb{C}\setminus\mathbb{R}.$$

(iii) *The support Λ_μ of the measure $d\mu(\lambda)$ satisfies the relation*

$$\Lambda_\mu \subset \left(-2\sqrt{v_1 v_m}, 2\sqrt{v_1 v_m}\right), \tag{2.10}$$

where

$$v_1 = \int_0^{\infty} \lambda\, d\nu(\lambda), \qquad v_m = \sup_{\lambda\in\Lambda_\nu} \lambda, \tag{2.11}$$

and Λ_ν is the support of the measure $d\nu(\lambda)$; also, if $\lambda\in\Lambda_\mu$, then $-\lambda\in\Lambda_\mu$.

REMARKS. (i) Equation (2.9) was derived by Khorunzhy and Pastur (1994) for random matrices (2.2) with Gaussian $h(x, y)$ satisfying (2.1), where $V(x, y) = u(x - y)$ and

$$u(-x) = u(x), \qquad x \in \mathbb{Z}, \sum_{x \in \mathbb{Z}} |u(x)| < \infty. \tag{2.12}$$

The latter inequality implies that the family of random variables $\{h(x, y)\}$ is weakly correlated, that is, that the dependence between $h(x, y)$ and $h(s, t)$ vanishes when $|x - s|$ or $|y - t|$ increases infinitely. Conditions (2.12) provide the existence of the limit (2.7) given by the relation

$$\nu(\lambda) = \text{meas}\{p \in [0, 1]: \tilde{u}(p) \leq \lambda\},$$

where $\tilde{u}(p) = \sum_{x \in \mathbb{Z}} u(x) \exp\{2\pi i x p\}$, $p \in [0, 1]$ [see, e.g., Grenander and Szegö (1958)].

(ii) Several analogues of (2.9) are known for some classes of random operators with statistically dependent coefficients [see, e.g., Wegner (1979), Khorunzhy and Pastur (1993) and Khorunzhy (1996)].

(iii) Condition (2.7) can be regarded as a certain form of the condition of weak statistical dependence for the family $\{h(x, y)\}$. Indeed, it follows from (2.3) and (2.7) that $N^{-1}\sum_{x, y=1}^{N}[V(x, y)]^2$ is bounded as $N \to \infty$. Therefore the N^2 terms $[V(x, y)]^2$ cannot be all of the same order of magnitude. This implies a decay of $V(x, y)$ when $|x - y| \to \infty$.

(iv) In case V has a diagonal form

$$V(x, y) = w(x)\delta(x - y), \qquad x, y \in \mathbb{N}, w(x) > 0, \tag{2.13}$$

the Gaussian random variables $h(x, y)$ are uncorrelated and, hence, are jointly independent. In this case, condition (2.7) is equivalent to the condition that the following limit exists:

$$\nu(\lambda) = \lim_{N \to \infty} \#\{x: w(x) \leq \lambda, x = 1, \dots, N\} N^{-1}. \tag{2.14}$$

This is true in the case where $w(x)$ is determined as a realization of an infinite sequence of i.i.d. random variables with distribution function $\nu(\lambda)$. Then the law of large numbers implies (2.14).

(v) Set

$$\nu(\lambda) = \begin{cases} 0, & \text{if } \lambda \leq v^2, \\ 1, & \text{if } \lambda > v^2. \end{cases} \tag{2.15}$$

Then (2.9) reduces to the equation

$$f(z) = \frac{1}{-z - v^2 f(z)}. \tag{2.16}$$

This equation was first derived by Marchenko and Pastur (1967) for the Wigner ensemble of random matrices.

(vi) Any function $\phi(z) \in \mathbf{F}$ admits a representation

$$\phi(z) = \int_{-\infty}^{\infty} (\lambda - z)^{-1} \, d\sigma(\lambda),$$

where $\sigma(\lambda)$ is a nonnegative nondecreasing function such that $\int_{-\infty}^{\infty} d\sigma(\lambda) < \infty$. This function can be found by the inversion formula

$$\sigma(b) - \sigma(a) = \lim_{\eta \downarrow 0} \frac{1}{\pi} \operatorname{Im} \int_a^b \phi(\lambda + i\eta)\, d\lambda, \tag{2.17}$$

where a and b are such that $\sigma(\lambda)$ is continuous at these points [see, e.g., Donoghue (1974)]. If the derivative σ' exists on the whole axis, then (2.17) can be rewritten as

$$\sigma'(a) = \frac{1}{\pi} \operatorname{Im} \phi(a + i0). \tag{2.18}$$

Using (2.18), one can easily derive from (2.16) an exact expression for $\sigma_v'(\lambda)$ (1.4b).

We obtain (2.5) and (2.8) and (2.9) under the rather weak and natural conditions (2.3) and (2.4) and (2.7). In brief, we require that the matrix V possess those spectral characteristics that we expect to exist for H_N in the limit $N \to \infty$. We see that the location of the support of the limiting eigenvalue distribution function is determined by the product $v_1 v_m$, while the upper bound of the spectral norm $\|H_N\|$ is determined by the variables v_1 and $\|V\|$. It is natural to assume that the lower bound of $\|H_N\|$ also depends on the norm $\|V\|$ and therefore in general there exist eigenvalues of H_N lying outside of the support Λ_μ.

A trivial example of a matrix H_N with this property is provided by (2.1) and (2.2) when $V(x, y)$ has the form of (2.13), with

$$w(x) = \begin{cases} v', & \text{if } x = 1, \\ v, & \text{if } x > 1. \end{cases} \tag{2.19}$$

In this case it is easy to see that $\nu(\lambda)$ is given by (2.15) and, hence, the density of the limit eigenvalue distribution is given by (1.4) and has the support $(-2v, 2v)$. On the other hand, we have for the vector $e_1(x) = \delta(x-1)$, with probability 1,

$$\|H_N e_1\|^2 = \frac{1}{N} \sum_{x=1}^{N} [h(x,1)]^2 \to vv' \qquad \text{as } N \to \infty.$$

This relation implies that $\limsup_{N \to \infty} \|H_N\|^2 \geq vv'$. Thus, for $v' > 4v$ one can find with probability 1 eigenvalues of H_N outside $(-2v, 2v)$ in the limit $N \to \infty$.

THEOREM 2.3. *Let V satisfies the conditions of Theorem* 2.2 *and*

$$\frac{1}{N} \operatorname{Tr} V_N^r \leq \int_0^\infty \lambda^r \, d\nu(\lambda), \qquad r \in \mathbb{N} \tag{2.20}$$

for all $N \in \mathbb{N}$. Then with probability 1,

$$\lim_{N \to \infty} \|H_N\| = \chi_\mu, \tag{2.21}$$

where $\chi_\mu = \sup_{\lambda \in \Lambda_\mu} |\lambda|$.

It is easy to show that (2.20) is satisfied when V is a difference matrix

$$V(x, y) = u(x - y), \tag{2.22}$$

with nonnegative entries $u(x) > 0$ such that (2.12) holds (see Lemma 5.5 in Section 5). It should be noted that all our results remain valid in the case where condition (2.1b) is replaced by the condition

$$\mathbf{E}h(x, y)h(s, t) = V(x, s)V(y, t).$$

In this case the particular form (2.22) of V describes the random field $\{h(x, y),\ x \le y\}$ that can be regarded as a version of a stationary random field. In this connection, let us note that the matrices H_N resemble the metrically transitive operators introduced and studied by Pastur and Figotin (1992). These operators have the property that their spectra coincide with the support of the limiting eigenvalue distribution function. Thus, our observations lead to the conjecture that the same property for random matrices is also related with stationarity of the probability distribution of their entries. It could be interesting to develop a more precise formulation of this conjecture.

To conclude, we remark that as a by-product of the proof of Theorems 2.1 and 2.3 [see estimates (3.14) and (3.24)] we obtain the estimate

$$\mathrm{Prob}\{\|H_N\| > \chi_\mu(1 + \varepsilon_N)\} \le \exp\left[-\left(\frac{N}{3\theta_N}\right)^{1/6} \varepsilon_N(1 + o(1))\right], \tag{2.23}$$

as $N \to \infty$, where $\theta_N = \|V_N\|/(N^{-1}\ \mathrm{Tr}\, V_N)$ and ε_N is an arbitrarily chosen sequence such that $N^{1/6}(3\theta_N)^{-1/6}\varepsilon_N \to \infty$ as $N \to \infty$.

In the case of GOE, estimate (2.23) takes the form

$$\begin{aligned} &\mathrm{Prob}\{\|H_N\| > 2v(1 + \varepsilon_N)\} \\ &\quad \le \exp\left[-(N)^\tau \varepsilon_N/3(1 + o(1))\right], \qquad N \to \infty, \end{aligned} \tag{2.24}$$

with $\tau = 1/6$. In this case it is known that (2.24) holds also for $\tau > 1/6$. However, it is not hard to observe that the upper bound here is $2/3$. In particular, this follows from results obtained by Tracy and Widom (1994) that the maximal eigenvalue of A_N is located in the vicinity of the point $2v(1 + N^{-2/3})$. In this connection, it would be interesting to find an optimal improvement of (2.23).

3. Infinite system of moment relations. To prove the theorems of Section 2, we study the asymptotic behavior of the moments

$$M_p(N) = \mathbf{E}\overline{H_N^p}, \qquad \overline{H_N^p} = N^{-1}\ \mathrm{Tr}\, H_N^p = \int_{-\infty}^{\infty} \lambda^p \, d\sigma(\lambda; H_N),\ p \in \mathbb{N}.$$

We derive an infinite system of relations that involves the moments $M_p(N)$ and certain terms vanishing for $N \to \infty$. We develop a technique that allows us to estimate these terms for $p = O(N^\tau)$ with some $\tau > 0$, $N \to \infty$. Using these estimates in the case of finite p, we obtain the convergence of $M_p(N)$ as $N \to \infty$ and the convergence of $d\sigma(\lambda; H_N)$ with probability 1. The estimates derived for the case of infinitely increasing p allow us to study the deviations of $\|H_N\|$.

Let us introduce the variables

$$M_p^{(q)}(N) = \mathbf{E}\overline{H_N^p V_N^q}, \qquad p, q \in \mathbb{Z}_+, \mathbb{Z}_+ = \mathbb{N} \cup \{0\}, M_0^{(0)}(N) = 1.$$

In the sequel we omit the subscripts N when this does not lead to confusion. We compute the average

$$\mathbf{E}H^p(x, y) = \sum_{s=1}^{N} \mathbf{E}H^{p-1}(x, s)H(s, y)$$

with the help of the following statement.

PROPOSITION 3.1. *Let random variables $\gamma_1, \dots, \gamma_n$ have a joint Gaussian distribution with zero mathematical expectation. Then*

$$\mathbf{E}\gamma_j F(\gamma_1, \dots, \gamma_n) = \sum_{i=1}^{n} \mathbf{E}\gamma_j\gamma_i \mathbf{E}\frac{\partial F}{\partial \gamma_i}, \tag{3.1}$$

where $F(x_1, \dots, x_n)$ is a nonrandom smooth function such that all integrals in (3.1) *exist.*

One can easily prove this statement by integration by parts.

Taking into account (3.1), we derive the relation

$$\begin{aligned} &\mathbf{E}H^{p-1}(a, b)H(x, y) \\ &\quad = \frac{1}{N}\sum_{i=0}^{p-2} \mathbf{E}\{[H^{p-2-i}V](a, x)[VH^i](y, b) \\ &\qquad\qquad + [H^{p-2-i}V](a, y)[VH^i](x, b)\} \end{aligned} \tag{3.2}$$

(see Lemma 5.1 for the proof). Applying (3.2) to $M_p^{(q)}(N)$, we obtain

$$\mathbf{E}\overline{H^p V^q} = \sum_{i=0}^{p-2} \mathbf{E}\big\{\overline{H^{p-2-i}V}\,\overline{H^i V^{q+1}} + N^{-1}\overline{H^{p-2-i}VH^iV^{q+1}}\big\}. \tag{3.3}$$

Let us introduce the notation

$$\langle \xi \rangle = \xi - \mathbf{E}\xi,$$

where ξ is an arbitrary random variable with finite expectation. Taking into account the fact that $\mathbf{E}\overline{H^{2k+1}V^q} = 0$ [see also (4.2) and Lemma 5.2 for the proof], we derive from (3.3) that

$$\begin{aligned} M_{2k}^{(q)}(N) &= \sum_{j=0}^{k-1} M_{2k-2-2j}^{(1)}(N) M_{2j}^{(q+1)}(N) \\ &\quad + \Phi_{2k-2}^{(q)}(N) + \Psi_{2k-2}^{(q)}(N), \end{aligned} \tag{3.4}$$

where

$$\Phi^{(q)}_{2k-2}(N) = N^{-1} \sum_{i=0}^{2k-2} \mathbf{E}\overline{H^{2k-2-i} V H^i V^{q+1}}$$

and

$$\Psi^{(q)}_{2k-2}(N) = \sum_{i=0}^{2k-2} \mathbf{E}\langle \overline{H^{2k-2-i} V}\rangle \langle \overline{H^i V^{q+1}}\rangle.$$

The key observation is that $\Phi^{(q)}_{2k-2}(N)$ and $\Psi^{(q)}_{2k-2}(N)$ are $o(M^{(q)}_{2k-2}(N))$ as $N \to \infty$. To show this, we introduce variables

$$L^{(b)}_{p,N}(x,y) = \sum_{\alpha_i \ge 0,\, \Sigma_{i=1}^{b} \alpha_i = p} [H^{\alpha_1} V H^{\alpha_2} \cdots V H^{\alpha_b}](x,y)$$

and

$$D^{(m,q)}_{p,N}(b_1,\dots,b_m) = \sum_{\beta_i > 0,\, \Sigma_{i=1}^{m} \beta_i = p} \left| \mathbf{E}\langle \overline{L^{(b_1)}_{\beta_1} V}\rangle \cdots \langle \overline{L^{(b_{m-1})}_{\beta_{m-1}} V}\rangle \langle \overline{L^{(b_m)}_{\beta_m} V^{q+1}}\rangle \right|$$

and formulate our main technical result.

THEOREM 3.1. *Let* $\theta = \sup_N \theta_N$, *with* $\theta_N = \tilde{V}_N / v_1^{(N)}$ *and*

$$v_1^{(N)} = \overline{V_N} = N^{-1} \operatorname{Tr} V_N, \qquad \tilde{V}_N = \max_{j=1,\dots,N} |\lambda_j(V_N)|.$$

If $N \ge N_0 \ge 3 \cdot 2^{18}\theta$, *then*

$$0 < \mathbf{E}\overline{L^{(b)}_{2k,N} V^q_N} \le (2k+2)^{2(b-1)} (b-1)^{b-1} \tilde{V}^{b-1} M^{(q)}_{2k}(N) \tag{3.5}$$

and

$$D^{(m,q)}_{2k,N}(b_1,\dots,b_m) \le \frac{(4k+4)^{2B_m+m} (B_m)^{B_m}}{N^m} \tilde{V}^{B_m-1} M^{(q+1)}_{2k}(N) \tag{3.6}$$

for all $q \ge 0$ *and* $b, k, B_m = b_1 + \cdots + b_m$ *such that* $2\max\{b, B_m, 4k+4\} \le (N/3\theta)^{1/6}$.

REMARKS. (1) To explain the form of the estimates (3.5) and (3.6), let us note that $L^{(b)}_{2k,N}$ is the sum of $T(2k,b) \equiv \binom{2k+b-1}{b-1}$ terms, where $T(2k,b)$ is the number of all possible distributions of $2k$ identical balls into b boxes. Each of the terms involved can be estimated from above by the same expression. So, to simplify the subsequent computations, we replace $T(2k,b)$ by its upper bound. This increases the value of the exponent of $4k+4$ in the right-hand sides of the estimates (3.5) and (3.6). In turn, this leads to decrease of the exponent τ in the estimate for the possible growth of $k \le N^{\tau}$. This is why we consider $\tau = 1/6$ as far from the optimal exponent.

(2) The factor N^m that appears in the right-hand side of (3.6) reflects a special property of the moments $M^{(q)}_{2k}(N)$ and, hence, of the measure $d\sigma(\lambda; H_N)$. In the theoretical physics literature this property is known as the strong self-averaging property [see, e.g., Lifshitz, Gredeskul and Pastur

(1988)]. In the RM theory, it was first observed by Berezin (1973). Here the strong self-averaging means that the variance of the variables

$$\int_{-\infty}^{\infty} \rho(\lambda)\, d\sigma(\lambda; H_N), \qquad \rho \in C_0^{\infty}(\mathbb{R}),$$

is $o(N^{-1})$. This is true for a wide class of random matrices. Actually, (3.6) is an improvement of this statement. The form of (3.6) is based on the assertion that the random variable

$$N\left[\int_{-\infty}^{\infty} \rho(\lambda)\, d\sigma(\lambda; H_N) - \mathbf{E}\int_{-\infty}^{\infty} \rho(\lambda)\, d\sigma(\lambda; H_N)\right]$$

converges to a gaussian random variable as $N \to \infty$ [see, e.g., Girko (1988), Khorunzhy, Khoruzhenko and Pastur (1996)].

Here we prove that the moments of the random variable $(\overline{H^{p_N}} - \mathbf{E}\overline{H_N^{p_N}})^k$ are of the order $N^{-(p+k)}$ when $N \to \infty$ and $p + k \to \infty$ simultaneously. This means that $d\sigma(\lambda; H_N)$ converges to a nonrandom limit much faster than the strong self-averaging property predicts. Let us also note that a similar observation is made by Sinai and Soshnikov (1998), who have proved that in Wigner ensemble, the random variable $\overline{H_N^{p_N}} - \mathbf{E}\overline{H_N^{p_N}}$ converges to a Gaussian random variable when $N, p \to \infty$, $p = N^{\tau}$, $\tau < 1/2$.

(3) Theorem 3.1 remains valid when θ is replaced by θ_N. This means that we can replace condition (2.4) by condition (2.6), with $\tau > 0$ such that $N/\theta_N \to \infty$.

We prove Theorem 3.1 in Section 4. It follows from (3.5) and (3.6) that

$$\left|\Phi_{2k-2}^{(q)}(N)\right| \le \mathbf{E}\frac{1}{N}\overline{L_{2k-2,N}^{(2)} V^{q+1}} \le \frac{\theta(4k)^2}{N} M_0^{(1)}(N) M_{2k-2}^{(q+1)}(N) \tag{3.7}$$

and

$$\left|\Psi_{2k-2}^{(q)}(N)\right| \le D_{2k-2,N}^{(2,q)}(1,1) \le \frac{4\theta(4k)^6}{N^2} M_0^{(1)}(N) M_{2k-2}^{(q+1)}(N) \tag{3.8}$$

for all $N > N_0$ such that $(N/3\theta)^{1/6} \ge 4k$. Taking these estimates into account, we return to the proof of Theorem 2.1.

PROOF OF THEOREM 2.1. Let us first note that given any positive number $\varepsilon < 1$, there exists $N_1(\varepsilon)$ such that

$$\frac{\theta(4k)^2}{N} + \frac{4\theta(4k)^6}{N^2} < \varepsilon \qquad \forall\, N > \max\{N_0, N_1\}. \tag{3.9}$$

Then (3.7) and (3.8) imply that

$$(1-\varepsilon)\sum_{j=0}^{k-1} M_{2k-2-2j}^{(1)}(N) M_{2j}^{(q+1)}(N) \le M_{2k}^{(q)}(N) \tag{3.10a}$$

and

$$M_{2k}^{(q)}(N) \le (1+\varepsilon) \sum_{j=0}^{k-1} M_{2k-2-2j}^{(1)}(N) M_{2j}^{(q+1)}(N), \tag{3.10b}$$

where $q \in \mathbb{Z}_+$ and $k \in \mathbb{N}$.

Modifying slightly the arguments used by Boutet de Monvel and Shcherbina (1995), we introduce a sequence of positive numbers $m_k^{(q)}(N,\varepsilon)$ determined for a fixed positive $\varepsilon < 1$ by the relations

$$m_k^{(q)}(N,\varepsilon) = (1+\varepsilon) \sum_{j=0}^{k-1} m_{k-1-j}^{(1)}(N,\varepsilon) m_j^{(q+1)}(N,\varepsilon), \qquad k \ge 1, \tag{3.11a}$$

$$m_0^{(q)}(N,\varepsilon) = (1+\varepsilon)^{q/2} \overline{V_N^q}, \qquad m_0^{(0)}(N,\varepsilon) = 1. \tag{3.11b}$$

It is easy to see that the inequalities (3.10) imply the estimates

$$m_k^{(q)}(N,-\varepsilon) \le M_{2k}^{(q)}(N) \le m_k^{(q)}(N,\varepsilon). \tag{3.12}$$

Equation (3.11a) resembles the system of relations derived by Wigner (1955) for the moments of the semicircle distribution (see Section 5, Lemma 5.4). From this observation, we obtain

$$m_k^{(q)}(N,\varepsilon) \le \left[(1+\varepsilon)2l_N\right]^{2k} (1+\varepsilon)^{q/2} \tilde{V}_N^q, \qquad l_N = \sqrt{v_1^{(N)} \tilde{V}_N} \tag{3.13}$$

(see Lemma 5.6 for the proof). Thus we get

$$M_{2k}^{(0)}(N) \le \left[(1+\varepsilon)2l_N\right]^{2k}$$

for $k < 8^{-1}(N/3\theta)^{1/6}$ when N is large enough. Let us show that this implies (2.5).

In view of the definition of the moments $M_{2k}^{(0)}(N)$, we can write for any $a > 0$,

$$\begin{aligned} M_{2k}^{(0)}(N) &\ge \mathbf{E} \int_{\mathbb{R}\setminus(-a,a)} \lambda^{2k}\, d\sigma(\lambda; H_N) \\ &\ge \frac{a^{2k}}{N} \mathbf{E}\#\left\{ \left|\lambda_j^{(N)}\right| \ge a \right\} \\ &\ge \frac{a^{2k}}{N} \operatorname{Prob}\{\|H_N\| \ge a\}. \end{aligned}$$

Then

$$\begin{aligned} &\operatorname{Prob}\{\|H_N\| \ge (1+2\varepsilon)2l_N\} \\ &\quad \le N \inf_k \frac{M_{2k}^{(0)}(N)}{\left[(1+2\varepsilon)2l_N\right]^{2k}} \\ &\quad = N \exp\left[-\left(\frac{N}{3\theta}\right)^{1/6} \log\left(1 + \frac{\varepsilon}{1+\varepsilon}\right)\right]. \end{aligned} \tag{3.14}$$

An elementary calculation shows that

$$\sum_{N=1}^{\infty} \text{Prob}\{\|H_N\| \geq (1+2\varepsilon)2l_N\} < \infty.$$

Since ε is arbitrary positive, then the Borel–Cantelli lemma implies (2.5). □

PROOF OF THEOREM 2.2. Let us prove Theorem 2.2. In view of condition (2.7), the following limits exist:

$$\lim_{N\to\infty} m_0^{(q)}(N,\varepsilon) = \hat{m}_0^{(q)}(\varepsilon) \equiv (1+\varepsilon)^{q/2}\int \lambda^q \, d\nu(\lambda). \tag{3.15a}$$

This fact together with Lemma 5.4 implies that for any fixed $k, q \in \mathbb{Z}_+$,

$$\lim_{N\to\infty} m_k^{(q)}(N,\varepsilon) = \hat{m}_k^{(q)}(\varepsilon), \tag{3.15b}$$

where the moments $\hat{m}_k^{(q)}(\varepsilon)$ satisfy the analogue of system (3.11) where (3.11b) is replaced by the right-hand side of (3.15a).

Lemma 5.4 implies that the difference between $m_k^{(q)}(-\varepsilon)$ and $m_k^{(q)}(\varepsilon)$ vanishes when ε vanishes. This means that

$$\lim_{N\to\infty} M_{2k}^{(q)}(N) = \hat{m}_k^{(q)}(0) \equiv \hat{m}_k^{(q)}, \tag{3.16}$$

where the $\hat{m}_k^{(q)}$ satisfy the system of relations

$$\hat{m}_k^{(q)} = \sum_{j=0}^{k-1} \hat{m}_{k-1-j}^{(1)} \hat{m}_j^{(q+1)}, \qquad k, q \geq 1, \tag{3.17a}$$

$$\hat{m}_0^{(q)} = \int \lambda^q \, d\nu(\lambda), \qquad q \geq 1,\ \hat{m}_0^{(0)} = 1. \tag{3.17b}$$

Relation (3.16) with $q = 0$ can be rewritten in the form

$$\lim_{N\to\infty} \mathbf{E}\int \lambda^{2k} \, d\sigma(\lambda; H_N) = \hat{m}_k^{(0)}.$$

It follows from Lemma 5.2 of Section 5 that $\mathbf{E}\int \lambda^{2k+1}\, d\sigma(\lambda; H_N) = 0$. Let us show that there exists a unique limiting measure $d\mu(\lambda)$ with odd moments zero and even moments $\hat{m}_k^{(0)}$ bounded by $(2\tilde{V})^{2k}$ [see (3.13)].

Let us define the functions

$$f^{(q)}(z) = -\sum_{k=0}^{\infty} \hat{m}_k^{(q)} z^{-2k-1}, \qquad q \in \mathbb{Z}_+, \tag{3.18}$$

which are analytic in the region

$$U = \{z : |\text{Im } z| \geq 2\tilde{V}^2 + 1\}. \tag{3.19}$$

Then it is easy to show that (3.17) is equivalent to the system of equations

$$-zf^{(q)}(z) = \hat{m}_0^{(q)} + f^{(1)}(z) f^{(q+1)}(z), \qquad q \in \mathbb{Z}_+. \tag{3.20}$$

The system (3.20) has a unique solution that is also the unique solution of the system

$$f^{(q)}(z) = \int \frac{\lambda^q \, d\nu(\lambda)}{-z - \lambda f^{(1)}(z)}, \qquad q \in \mathbb{Z}_+ \tag{3.21}$$

(see Lemma 5.7 for the proof). Denoting $f^{(0)}(z) \equiv f(z)$ and $f^{(1)}(z) \equiv g(z)$, we obtain (2.9). In Lemma 5.8 we prove (iii) of Theorem 2.2.

At this point, we have derived the weak convergence (2.8) in average. To prove that (2.8) holds with probability 1, we need to show that $\lim_{N\to\infty} \overline{H_{2k}^{(N)}} = \hat{m}_k^{(0)}$ with probability 1. This can easily be proved in view of the convergence (3.16) of the mathematical expectations $M_{2k}^{(q)}(N) = \mathbf{E}\overline{H_{2k}^{(N)} V_N^q}$ and the estimate of the variance

$$(3.22)\quad \mathbf{E}\langle \overline{H_{2k}^{(N)} V_N^q}\rangle \langle \overline{H_{2k}^{(N)} V_N^q}\rangle = O(N^{-2}), \qquad N \to \infty \text{ for fixed } k, q \in \mathbb{Z}_+,$$

which follows from relation (3.6) of Theorem 3.1. Let us note that in Lemma 5.2 we prove that

$$\mathbf{E}\langle \overline{H_{2k+1}^{(N)} V_N^q}\rangle \langle \overline{H_{2k+1}^{(N)} V_N^q}\rangle = 0$$

[see also equality (4.3)].

The Borel–Cantelli lemma implies that each moment of the measure $d\sigma(\lambda; H_N)$ converges with probability 1,

$$\lim_{N\to\infty} \overline{H_p^{(N)}} = \begin{cases} \hat{m}_k^{(0)}, & \text{if } p = 2k, \\ 0, & \text{if } p = 2k+1. \end{cases}$$

The moments $\hat{m}_k^{(0)}$ uniquely determine the measure $d\mu(\lambda)$ with Stieltjes transform $f(z)$. This proves Theorem 2.2. □

Proof of Theorem 2.3. Condition (2.20) implies that the inequality

$$(3.23)\qquad m_k^{(q)}(N, \varepsilon) \le \hat{m}_k^{(q)}(\varepsilon)$$

holds for all $N \in \mathbb{N}$. It follows from Lemmas 5.7 and 5.8 that the moments $\hat{m}_k^{(0)}(\varepsilon)$ uniquely determine the measure with compact support $d\mu_\varepsilon(\lambda)$. Let us denote by $\chi_\mu^{(\varepsilon)}$ the upper edge of this support. Then (3.23), together with the second inequality of (3.12), implies that for large N we have

$$M_{2k}^{(0)}(N) \le \left[\chi_\mu^{(\varepsilon)}\right]^{2k} \quad \text{for all } k \le (N/3\theta)^{1/6}.$$

Repeating the arguments of the proof of Theorem 2.1, we obtain for arbitrary $\varepsilon > 0$ the inequality

$$(3.24)\quad \operatorname{Prob}\left\{\|H_N\| > \chi_\mu^{(\varepsilon)}(1 + \varepsilon_N)\right\} \le \exp\left[-\left(\frac{N}{3\theta_N}\right)^{1/6} \varepsilon_N(1 + o(1))\right],$$

as $N \to \infty$, hence

$$(3.25)\qquad \limsup_{N\to\infty} \|H_N\| \le \chi_\mu^{(\varepsilon)}.$$

Lemma 5.4 implies that $\lim_{\varepsilon\to 0} \chi_\mu^{(\varepsilon)} = \chi_\mu$. Relation (3.24) implies that

$$\limsup_{N\to\infty} \|H_N\| \le \chi_\mu \quad \text{with probability 1.}$$

On the other hand, Theorem 2.2 implies that, with probability 1, any interval $(\chi_\mu - \Delta, \chi_\mu)$ with fixed $\Delta > 0$ contains eigenvalues of H_N in the limit $N \to \infty$, and then

$$\limsup_{N\to\infty} \|H_N\| \geq \chi_\mu - \Delta \quad \text{with probability 1.}$$

This proves Theorem 2.3. Let us note that estimate (2.23) follows from (3.24) because the right-hand side of this estimate does not depend on $\varepsilon > 0$. □

4. Proof of recurrent estimates. In this section we prove Theorem 3.1. Let us consider the mathematical expectation

$$\mathbf{E}\overline{L_p^{(b)} V^q} = N^{-1} \sum_{x,s} \sum_{\alpha_i \geq 0, \sum_{i=1}^b \alpha_i = p} \mathbf{E}(H^{\alpha_1} V \cdots H^{\alpha_{b-1}} V H^{\alpha_b})(x,t) V^q(t,x).$$

We apply (5.1) to the last average (see Lemma 5.1 in Section 5) and obtain

$$\begin{aligned} \mathbf{E}\overline{L_p^{(b)} V^q} = &\sum_{s=0}^{b-1} \sum_{t=0}^{b-s-1} \sum_{i=0}^{p-2} \mathbf{E}\overline{L_{p-2-i}^{(b-s-t)} V}\, \mathbf{E}\overline{L_i^{(t+1)} V^{q+s+1}} \\ &+ N^{-1} \sum_{s=0}^{b-1} (b-s) \mathbf{E}\overline{L_{p-2}^{(b-s+1)} V^{q+s+1}} \\ &+ \sum_{s=0}^{b-1} \sum_{t=0}^{b-s-1} \sum_{i=0}^{p-2} \mathbf{E}\langle \overline{L_{p-2-i}^{(b-s-t)}} \rangle \mathbf{E} \langle \overline{L_i^{(t+1)} V^{q+s+1}} \rangle. \end{aligned} \tag{4.1}$$

Taking into account the inequality (5.3) from Section 5 and the relations

$$\mathbf{E}\overline{L_{2k+1,N}^{(b)} V_N^q} = 0, \qquad \mathbf{E}\overline{L_{2k,N}^{(b)} V_N^q} > 0 \tag{4.2}$$

and

$$D_{2j+1,N}^{(2,q)}(b_1, b_2) = 0, \tag{4.3}$$

which are proved in Lemma 5.2, we derive from (4.1) our first main inequality,

$$\begin{aligned} \mathbf{E}\overline{L_{2k}^{(b)} V^q} \leq &\sum_{s=0}^{b-1} \tilde{V}^s \sum_{t=0}^{b-s-1} \sum_{j=0}^{k-1} \mathbf{E}\overline{L_{2k-2-2j}^{(b-s-t)} V}\, \mathbf{E}\overline{L_{2j}^{(t+1)} V^{q+1}} \\ &+ N^{-1} \sum_{s=0}^{b-1} (b-s) \tilde{V}^s \mathbf{E}\overline{L_{2k-2}^{(b-s+1)} V^{q+1}} \\ &+ \sum_{s=0}^{b-1} \sum_{t=0}^{b-s-1} D_{2k-2}^{(2,q+s)}(b-s-t, t+1). \end{aligned} \tag{4.4}$$

The second relation concerns the variables $D_p^{(m)}$. To derive it, we use the identity

$$\mathbf{E}\langle \xi_1 \rangle \langle \xi_2 \rangle \equiv \mathbf{E}\langle \xi_1 \rangle \xi_2 \tag{4.5}$$

and write the equality

$$\mathbf{E}\langle \overline{L_{\beta_1}^{(b_1)} V} \rangle \cdots \langle \overline{L_{\beta_{m-1}}^{(b_{m-1})} V} \rangle \langle L_{\beta_m}^{(b_m)}(x,y) \rangle = \mathbf{E}\langle \Lambda_{m-1} \rangle L_{\beta_m}^{(b_m)}(x,y), \tag{4.6}$$

where we denote $\Lambda_{m-1} = \langle \overline{L_{\beta_1}^{(b_1)} V} \rangle \cdots \langle \overline{L_{\beta_{m-1}}^{(b_{m-1})} V} \rangle$. Using relations (3.2) and (5.1), we obtain

$$\begin{aligned}
&\mathbf{E}\langle \Lambda_{m-1} \rangle L_{\beta_m}^{(b_m)}(x,y) \\
&\quad = \sum_{s=0}^{b_m-1} \left(\sum_{t=0}^{b_m-s-1} \sum_{i=0}^{\beta_m-2} \mathbf{E}\Big\{ \langle \Lambda_{m-1} \rangle \overline{L_{\beta_m-2-i}^{(b_m-s-t)} V} \big[V^{s+1} L_i^{(t+1)} \big](x,y) \Big\} \right. \\
&\qquad\qquad \left. + \frac{b_m - s}{N} \mathbf{E}\langle \Lambda_{m-1} \rangle \big[L_{\beta_m-2}^{(t+1)} V^{s+1} \big](x,y) \right) \\
&\qquad + \sum_{s=0}^{b_m-1} \sum_{r=1}^{m-1} \mathbf{E}\Big\{ Q_1 \cdots Q_{r-1} \frac{2b_r}{N} \big[L_{\beta_r+\beta_m-2}^{(b_r+b_m-s+1)} V^{s+1} \big](x,y) \\
&\qquad\qquad \times Q_{r+1} \cdots Q_{m-1} \Big\},
\end{aligned} \tag{4.7}$$

where Q_r stands for $\overline{\langle L_{\beta_r}^{(b_r)} V \rangle}$. Let us note that the first mathematical expectation in the right-hand side of (4.7) is of the form $\mathbf{E}\langle T \rangle YZ$. We rewrite it in a form appropriate for our use, with the help of the identity

$$\mathbf{E}\langle T \rangle YZ = \mathbf{E}T\langle Y \rangle \mathbf{E}Z + \mathbf{E}T\langle Z \rangle \mathbf{E}Y + \mathbf{E}T\langle Y \rangle \langle Z \rangle - \mathbf{E}T\mathbf{E}\langle Y \rangle \langle Z \rangle.$$

Using this identity and applying (4.5) to the third term of the right-hand side of (4.7), we obtain our second main inequality,

$$D_{2k}^{(m,g)}(b_1, \dots, b_m) \le \sum_{i=1}^{7} R_i, \tag{4.8}$$

where

$$R_1 = \sum_{s=0}^{b_m-1} \sum_{t=0}^{b_m-s-1} \sum_{j=0}^{k-1} D_{2k-2-2j}^{(m,0)}(b_1, \dots, b_{m-1}, b_m - s - t) \mathbf{E}\overline{L_{2j}^{(t+1)} V^{q+s+2}},$$

$$R_2 = \sum_{s=0}^{b_m-1} \sum_{t=0}^{b_m-s-1} \sum_{j=0}^{k-1} D_{2k-2-2j}^{(m,q+s+1)}(b_1, \dots, b_{m-1}, t+1) \mathbf{E}\overline{L_{2j}^{(b_m-s-t)} V},$$

$$R_3 = \sum_{s=0}^{b_m-1} \sum_{t=0}^{b_m-s-1} D_{2k-2}^{(m+1,q+s+1)}(b_1, \dots, b_{m-1}, b_m - s - t, t+1),$$

$$R_4 = \sum_{s=0}^{b_m-1} \sum_{j=0}^{k-1} D_{2k-2-2j}^{(m-1,0)}(b_1, \dots, b_{m-1}) \sum_{t=0}^{b_m-s-1} D_{2j}^{(2,q+s+1)}(b_m - s - t, t+1),$$

$$R_5 = \sum_{s=0}^{b_m-1} \frac{b_m - s}{N} D_{2k-2}^{(m,q+s+1)}(b_1, \dots, b_{m-1}, b_m - s + 1),$$

$$R_6 = \sum_{r=1}^{m-1} \frac{2b_r}{N^2} D_{2k-2}^{(m-1,\,q+s+1)}(b_1,\dots,b_{r-1},b_{r+1},\dots,b_{m-1},b_m+b_r-s+1),$$

$$R_7 = \sum_{r=1}^{m-1} \frac{2b_r}{N^2} \times \sum_{j=0}^{k-1} D_{2k-2-2j}^{(m-2,0)}(b_1,\dots,b_{r-1},b_{r+1},\dots,b_{m-1}) \times (2j+1)\mathbf{E}\overline{L_{2j}^{(b_r+b_m-s+1)}V^{q+s+1}}.$$

It should be noted that in Section 3 we defined the variables $D_{2k}^{(m)}$ only for $2 \le m \le 2k$, $k \ge 1$. This means that for $m = 2k$, $m = 2k-1$, or $m = 2k-2$, certain terms in the right-hand side of (4.8) should be omitted. However, if we set $D_{2k}^{(m)} = 0$ for $m > 2k$, then (4.8) will be valid for all $k \ge 1$, $m \ge 2$.

Now we describe the induction procedure that we use to prove Theorem 3.1. Let us call "D-plane" the set of pairs (k, m), $m \ge 2$, $k, m \in \mathbb{N}$ and "L-line" the set of points $j, j \in \mathbb{N}$. Suppose that for some integer $J \ge 3$ the estimates (3.5) and (3.6) are valid for all $D_{2k'}^{(m')}$ and $L_{2j'}$ such that $k' + m' \le J$, $j' \le J - 2$. We will say that such points on our D-plane and L-line are positive.

Let us introduce the set $D(J+1)$ of points (k, m) determined by the relation $k + m = J + 1$. The step of the induction is to prove that $D(J+1)$ consists of positive points and that the point $j = J - 1$ is also positive. Then, according to the induction principle, (3.5) and (3.6) will be proved for all fixed k and m.

To add $D(J+1)$ to the set of positive points, we start from the point $k' = 1$, $m' = J$, which is apparently positive because (3.6) is true for these k' and m'. Now let us assume that for some integer $k \ge 2$ all points of $D(J+1)$ satisfying $k' \le k - 1$ are positive. Then all terms D and L involved in the right-hand side of (4.8) correspond to positive points of the D-plane and L-line and therefore satisfy (3.5) and (3.6). Our main goal is to derive that this implies (3.5) for $D_{2k}^{(m,q)}$. When this statement is proved, the point $k' = k$, $m' = J + 1 - k$ of the D-plane is shown to be positive so $D(J+1)$ consists of positive points. When this is proved, we will show that the point $j = J - 1$ in the L-line is also positive. We do this with the help of (4.4).

Following the induction principle, let us first ensure that (3.5) and (3.6) are valid for the initial points $\overline{L_2^{(b)}V^q}$ and $D_2^{(2,q)}(b_1, b_2)$. Let us first consider

$$\mathbf{E}\overline{L_2^{(b)}V^q} = \sum_{\substack{\alpha_i \ge 0,\\ \alpha_1+\cdots+\alpha_b=2}} \mathbf{E}\overline{H^{\alpha_1}V \cdots VH^{\alpha_b}V^q} = \sum_{\substack{\pi+\rho+\sigma=b-1,\\ \pi,\rho,\sigma\ge 0}} \mathbf{E}\overline{V^{\pi}HV^{\rho}HV^{\sigma+q}}.$$

It is not hard to see that in the latter sum there are $b(b+1)/2$ terms corresponding to $\rho > 0$ and these can be estimated by $\tilde{V}^{b-1}(\overline{V^{q+1}}\overline{V} +$

$N^{-1}\overline{V^{q+2}}$). The b remaining terms corresponding to $\rho = 0$ have the form $\mathbf{E}\overline{H^2 V^{q+b-1}} = \overline{V^{q+b}}\overline{V} + N^{-1}\overline{V^{q+b+1}}$, and (3.5) obviously follows. Next, we find that $D_2^{(2,q)}(b_1, b_2) = 2b_1 b_2 N^{-1}\overline{V^{b_1+b_2+q+2}}$ and that (3.6) is true for $k = 1$. We emphasize that (3.5) and (3.6) are proved for $k = 1$ when q, b, b_1 and b_2 are arbitrarily fixed. This implies that we do not need to care about changes of these parameters in (4.4) and (4.8).

Let us turn now to the general case $k > 1$ for $L_{2k}^{(b)}$. According to (3.5), we assume that the first term in the right-hand side of (4.4) is less than

$$\begin{aligned} &\sum_{s=0}^{b-1} \tilde{V}^s \sum_{t=0}^{b-s-1} \sum_{j=0}^{k-1} \left(4(k-j)^2(b-s-t-1)\right)^{b-s-t-1} \\ &\qquad\qquad \times (4j+4)^{2t} t^t \tilde{V}^{b-s-1} \mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}} \\ &\le F_{k,b} \sum_{s=0}^{b-1} \frac{1}{(4k+4)^{2s}(b-1)^{2s}} \\ &\quad \sum_{j=0}^{k-1} \sum_{t=0}^{b-s-1} \left(1 - \frac{j+1}{k+1}\right)^{2(b-s-1-t)} \left(\frac{j+1}{k+1}\right)^{2t} \\ &\qquad \times \mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}}, \end{aligned} \tag{4.9}$$

with

$$F_{k,b} = (4k+4)^{2(b-1)}\left[(b-1)\tilde{V}\right]^{b-1}.$$

For $s \le b-2$ we can estimate the last sum over t using the inequality

$$\begin{aligned} \pi_k(d) &= \sup_{j=0,\dots,k-1} \sum_{t=0}^{d} \left(1 - \frac{j+1}{k+1}\right)^{2(d-t)} \left(\frac{j+1}{k+1}\right)^{2t} \\ &\le 1 - \frac{1}{2(k+1)}, \end{aligned} \tag{4.10}$$

which is proved in Lemma 5.3 (see Section 5). Then we find that the right-hand side of (4.9) is less than

$$\begin{aligned} &F_{k,b}\left[\left(1 - \frac{1}{2(k+1)}\right) \sum_{s=0}^{b-2} \frac{1}{(4k+4)^{2s}(b-1)^{2s}} \right. \\ &\qquad\qquad \left. + \frac{1}{(4k+4)^{b-1}(b-1)^{b-1}}\right] X_{2k,q} \\ &\quad \le F_{k,b}\left(1 - \frac{1}{4(k+1)}\right) X_{2k,q}, \end{aligned}$$

where

$$X_{2k,q} \equiv \sum_{j=0}^{k-1} \mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}}.$$

Using (3.5) again, we conclude that the second term of the right-hand side of (4.4) is less than

$$\begin{aligned}
&\frac{b-1}{N}\sum_{s=0}^{b-1}(4k)^{2(b-s)}(b-s)^{b-s}\tilde{V}^{b}\mathbf{E}\overline{L_{2k-2}^{(1)}V^{q+1}}\\
&\quad\le F_{k,b}\theta\frac{b^2}{N}(4k+4)^2\left[\left(1+\frac{1}{b-1}\right)^{b-1}+\left(1-\frac{1}{(4k+4)^2(b-1)}\right)^{-1}\right]\\
&\qquad\times\overline{V}\mathbf{E}\overline{L_{2k-2}^{(1)}V^{q+1}}\\
&\quad\le 2\theta eF_{k,b}(4k+4)^2(b+1)^2N^{-1}X_{2k,q}.
\end{aligned}$$

In the last inequality, we used the elementary relation

$$\left(1+\frac{1}{b-1}\right)^{b-1}\le e\quad\text{for all } b\ge 2 \tag{4.11}$$

and the obvious estimate

$$\overline{V}\mathbf{E}\overline{L_{2k-2}^{(1)}V^{q+1}}\le X_{2k,q}.$$

Assuming that (3.6) is valid for the third term of the right-hand side of (4.4), we see that it is less than

$$\begin{aligned}
&\frac{4k^2}{N^2}\sum_{s=0}^{b-1}\sum_{t=0}^{b-s-1}(4k)^{2(b-s+1)}(b-s+1)^{b-s+1}\tilde{V}^{b}\mathbf{E}\overline{L_{2k-2}^{(1)}V^{q+1}}\\
&\quad\le F_{k,b}\frac{(4k)^6(b+1)^3}{N^2}\sum_{s=0}^{b-1}\left(1+\frac{2}{b-1}\right)^{b-1}\\
&\qquad\sum_{s=0}^{b-1}\frac{1}{(4k+4)^{2s}(b+1)^s}\theta\overline{V}\mathbf{E}\overline{L_{2k-2}^{(1)}V^{q+1}}\\
&\quad\le F_{k,b}2\theta e^2(4k)^6(b+1)^3N^{-2}X_{2k,q}.
\end{aligned}$$

Now we can conclude that

$$\begin{aligned}
\mathbf{E}\overline{L_{2k}^{(b)}V^{q}}&\le F_{k,b}X_{2k,q}\\
&\quad\times\left(1-\frac{1}{4(k+1)}+\frac{2\theta e(4k+4)^2(b+1)^2}{N}\right.\\
&\qquad\left.+\frac{2\theta e^2(4k+4)^6(b+1)^3}{N^2}\right).
\end{aligned}\tag{4.12}$$

It follows from (4.1) that

$$X_{2k,q}\le\mathbf{E}\overline{L_{2k}^{(1)}V^{q}}+\frac{1}{N}\mathbf{E}\overline{L_{2k-2}^{(2)}V^{q+1}}+D_{2k-2}^{(2,q)}(1,1).$$

Applying (3.5) and (3.6) to the last two terms of this inequality, we obtain

$$X_{2k,q} \le \mathbf{E}\overline{L_{2k}^{(1)}V^q} + \frac{\theta(4k)^2}{N}X_{2k,q} + \frac{4\theta(4k)^6}{N^2}X_{2k,q}$$

and finally that

$$X_{2k,q} \le \left[1 - \frac{\theta(4k)^2}{N} - \frac{4\theta(4k)^6}{N^2}\right]^{-1}\mathbf{E}\overline{L_{2k}^{(1)}V^q}. \tag{4.13}$$

Elementary computations show that inequalities (4.12) and (4.13) imply (3.5) for all N, k and b satisfying the conditions of Theorem 3.1.

Now we turn to the proof of (3.6). Let us denote

$$G_{k,m,B} = (4k+4)^{2B_m+m}(B_m)^{B_m}\tilde{V}^{B_m-1}$$

and assume that (3.6) is true for all terms $R_1, \dots, R_7$ involved in the right-hand side of (4.8). Then for the first term we have

$$\begin{aligned} R_1 \le \sum_{s=0}^{b_m}\sum_{t=0}^{b_m-s-1}\sum_{j=0}^{k-1} &(4k-4j)^{m+2(B_m-s-t)}(B_m-s-t)^{B_m-s-t} \\ &\times(4j+4)^{2t}t^t\tilde{V}^{B_m-1}\mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}} \\ \le G_{k,m,B_m}&\sum_{j=0}^{k-1}\sum_{s=0}^{b_m}\frac{1}{(4k+4)^{2s}B_m^s} \\ &\sum_{t=0}^{b_m-s-1}\frac{(k-j)^{2(b_m-t)}(j+1)^{2t}}{(k+1)^{2b_m}}\frac{(k-j)^m}{(k+1)^m}\mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}} \\ \le G_{k,m,B_m}&\left(1-\frac{1}{(4k+4)^2B_m}\right)^{-1}\sum_{j=0}^{k-1}\frac{(k-j)^m}{(k+1)^m}\mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}}. \end{aligned}$$

It is easy to derive that R_2 is bounded by a similar expression, with $\mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\,\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}}$ replaced by $\mathbf{E}\overline{L_{2k-2-2j}^{(1)}V^{q+1}}\mathbf{E}\overline{L_{2j}^{(1)}V}$. Then we can write the inequality

$$R_1 + R_2 \le G_{k,m,B_m}\left(1-\frac{1}{(4k+4)^2B_m}\right)^{-1}\sum_{j=0}^{k-1}\frac{(k-j)^m}{(k+1)^m}W_k(j),$$

where we use that $B_m \ge 2m$ and denote

$$W_k(j) \equiv \mathbf{E}\overline{L_{2k-2-2j}^{(1)}V}\mathbf{E}\overline{L_{2j}^{(1)}V^{q+1}} + \mathbf{E}\overline{L_{2k-2-2j}^{(1)}V^{q+1}}\mathbf{E}\overline{L_{2j}^{(1)}V}.$$

The function $W_k(j)$, $j = 1, \dots, k-1$ is symmetric with respect to $(k-1)/2$. Since $(k-j)^m$ is convex for $m \ge 2$, we can write

$$\sum_{j=0}^{k-1}(k-j)^m W_k(j) \le \tfrac{1}{2}(k^m+1)\sum_{j=0}^{k-1}W_k(j) = (k^m+1)X_{2k,q}.$$

Thus

$$
\begin{aligned}
R_1 + R_2 &\le G_{k,m,B_m} X_{2k,q} \left(1 - \frac{1}{32m(k+1)^2}\right)^{-1} \\
&\quad \times \left[\left(1 - \frac{1}{k+1}\right)^m + \frac{1}{(k+1)^m}\right].
\end{aligned}
\tag{4.14}
$$

For the next term R_3 of (4.8) we find easily that

$$
\begin{aligned}
R_3 &\le G_{k,m,B_m} \frac{(4k)^3}{N} \sum_{s=0}^{b_m - 1} \\
&\quad \sum_{t=0}^{b_m - s - 1} \frac{k^m k^{2(B_m - s)} B_m}{(k+1)^m (k+1)^{2(B_m - s)} (4k+4)^{2s} B_m^s} \tilde{V}\, \mathbf{E} \overline{L_{2k-2}^{(1)} V^{q+1}} \\
&\le G_{k,m,B_m} \frac{(4k+4)^3 B_m^2 \theta}{N} \left(1 - \frac{1}{32m(k+1)^2}\right)^{-1} X_{2k,q}.
\end{aligned}
$$

It is not hard to derive that R_4 is bounded by the same expression and that

$$
R_5 \le G_{k,m,B_m} \frac{(4k+4)^2 B_m^2 \theta}{N} \left(1 - \frac{1}{32m(k+1)^2}\right)^{-1} X_{2k,q}
$$

and

$$
R_6 \le G_{k,m,B_m} \frac{(4k+4) B_m^2 \theta}{N} \left(1 - \frac{1}{32m(k+1)^2}\right)^{-1} X_{2k,q}.
$$

Similar computations show that

$$
\begin{aligned}
R_7 \le G_{k,m,B_m} &\sum_{j=0}^{k-1} \frac{4k}{(4k+4)^2} \\
&\times \sum_{s=0}^{b_m - 1} \frac{1}{(4k+4)^s B_m^s} \\
&\times \sum_{r=1}^{m-1} b_r \left(1 - \frac{j+1}{k+1}\right)^{2(B_{m-1} - b_r)} \\
&\times \left(\frac{j+1}{k+1}\right)^{2b_r} \frac{(B_{m-1} - b_r)^{B_{m-1} - b_r}}{(B_m)^{B_{m-1} - b_r}} \\
&\times \mathbf{E} \overline{L_{2k-2-2j}^{(1)} V}\, \mathbf{E} \overline{L_{2j}^{(1)} V^{q+1}}.
\end{aligned}
$$

Denoting $B' = B_{m-1} - b_r$, we observe that $b_r(B'/(B' + b_r))^{B'} \le 1$. Taking into account that R_7 is positive only when $m \ge 3$, we can estimate

$$
\sum_{r=1}^{m} (1 - P)^{2B'} P^{2b_r} \le m(1 - P)^{4(m-2)} P^4 < 1, \qquad P = \frac{j+1}{k+1}.
$$

In view of these inequalities, we obtain

$$R_7 \le G_{k,m,B_m} \frac{k}{(k+1)^2} \left(1 - \frac{1}{32m(k+1)^2}\right)^{-1} X_{2k,q}.$$

Using (4.12), we finally derive from (3.8) the estimate

$$\begin{aligned} D_{2k}^{(m,q)}&(b_1,\dots,b_m) \\ &\le G_{k,m,B_m}\left[\left(1 - \frac{1}{k+1}\right)^m + \frac{1}{(k+1)^m}\right. \\ &\qquad \left. + \frac{k}{(k+1)^2} + \frac{4\theta B_m^2 (4k+4)^3}{N}\right] \\ &\quad \times \left(1 - \frac{1}{32m(k+1)^2}\right)^{-1} \\ &\quad \times \left(1 - \frac{\theta(4k)^2}{N} - \frac{4\theta(4k)^6}{N^2}\right)^{-1} \mathbf{E}\overline{L_{2k}^{(1)} V^q}. \end{aligned} \tag{4.15}$$

It remains to show that for all N, k, m, and B_m satisfying the conditions of Theorem 3.1, the inequality

$$D_{2k}^{(m,q)}(b_1,\dots,b_m) \le G_{k,m,B_m} \mathbf{E}\overline{L_{2k}^{(1)} V^q}$$

holds. This can be done easily by simple computations. Indeed, the maximum value of the expression in the square brackets of (4.15) is obtained when $m = 2$ and $B_m = (N/3\theta)^{1/6}$. The maximum value of the next two factors from the right-hand side of (4.15) is obtained in the cases when $m = 2$ and $B_m = (N/3\theta)^{1/6}$, respectively. Then after simple computations, we see that the product of these three factors is bounded by the expression

$$\left[1 - \frac{1}{2(k+1)} + \frac{4\theta(4k+4)^3}{(3\theta)^{1/3} N^{2/3}}\right]\left(1 + \frac{1}{32(k+1)^2}\right)\left(1 + \frac{4\theta(4k)^2}{N}\right).$$

Now it is not hard to observe that this product is strictly less than 1 when the inequality

$$\frac{4\theta^{2/3}(4k+4)^3}{N^{2/3}} < \frac{1}{3(k+1)}$$

holds. This inequality is valid for all k satisfying the conditions of Theorem 3.1. Theorem 3.1 is proved. □

5. Auxiliary statements.

LEMMA 5.1. *The relation* (3.2) *is true, so is the equality*

$$\begin{aligned}\mathbf{E}L_p^{(b)}(x,y) &= N^{-1}\sum_{t=1}^{N}\sum_{s=0}^{b-1}\sum_{r=1}^{b-1}\sum_{j=0}^{p-2}\mathbf{E}\Big\{\big(L_{p-2-j}^{(b-s-r)}V\big)(x,t)\big(VL_j^{(r+1)}V^s\big)(y,t)\Big\} \\ &\quad + \mathbf{E}\Big\{\big(L_{p-2-j}^{(b-s-r)}V\big)(x,y)\big(VL_j^{(r+1)}V^s\big)(t,t)\Big\}.\end{aligned} \tag{5.1}$$

PROOF. Let us assume that $x \le y$ in (3.2). Then according to (3.1) we can write

$$\begin{aligned}&\mathbf{E}H^p(a,b)H(x,y)\\ &\quad = \frac{1}{\sqrt{N}}\sum_{s\le t}\mathbf{E}\left(\frac{\partial H^p(a,b)}{\partial h(s,t)}\right)[V(s,x)V(t,y)+V(s,y)V(t,x)].\end{aligned}$$

Regarding the equality

$$\frac{\partial H^p(a,b)}{\partial h(s,t)} = \sum_{l=0}^{p-1}\sum_{u,v=1}^{N}H^{p-1-l}(a,u)\frac{\partial H(u,v)}{\partial h(s,t)}H^l(v,b), \qquad p\ge 1$$

and the consequence of the definition (2.2)

$$\frac{\partial H(u,v)}{\partial h(s,t)} = \frac{1}{\sqrt{N}}\begin{cases}\delta(u-s)\delta(v-t), & \text{if } u\le v,\\ \delta(u-t)\delta(v-s), & \text{if } u>v,\end{cases}$$

we obtain that

$$\begin{aligned}&\mathbf{E}H^p(a,b)H(x,y)\\ &\quad = N^{-1}\sum_{l=0}^{p-1}\sum_{s\le t}\mathbf{E}\{H^{p-1-l}(a,s)H^l(t,b)+H^{p-1-l}(a,t)H^l(s,b)\}\\ &\qquad\qquad \times[V(s,x)V(t,y)+V(s,y)V(t,x)](1+\delta(s-t))^{-1}.\end{aligned}$$

This equality implies (3.2).

To prove (5.1), we start with the terms of $L_p^{(b)}$ such that $\alpha_b \ge 1$. Given fixed numbers $(\alpha_1,\dots,\alpha_b)$, let us compute the expectation

$$\mathbf{E}(H^{\alpha_1}V\cdots VH^{\alpha_b-1})(x,t)H(t,y) \equiv S$$

with the help of (3.1). We obtain

$$\begin{aligned}S &= N^{-1}\sum_{r=1}^{b}\sum_{j=0}^{\alpha_r-1}(H^{\alpha_1}V\cdots VH^{\alpha_r-1-j}V)(x,t)(VH^jV\cdots VH^{\alpha_b-1})(y,t)\\ &\quad + N^{-1}\sum_{r=1}^{b}\sum_{j=0}^{\alpha_r-1}(H^{\alpha_1}V\cdots VH^{\alpha_r-1-j}V)(y,t)(VH^jV\cdots VH^{\alpha_b-1})(x,t).\end{aligned}$$

Note that if $\alpha_b = 1$ or $\alpha_i = 0$ for some i, $1 \le i \le b - 1$, then the terms corresponding to $r = b$ or $r = i$ are absent.

Thus, we have

$$\sum_{\substack{\alpha_i \ge 0,\, \alpha_b \ge 1 \\ \alpha_1 + \cdots + \alpha_b = p}} \mathbf{E}(H^{\alpha_1} V \cdots V H^{\alpha_b})(x, y)$$

$$= N^{-1} \sum_{t=1}^{N} \sum_{r=1}^{b} \sum_{j=0}^{p-2} \mathbf{E}\Big\{ \big(L^{(r)}_{p-2-j} V\big)(x, t) \big(V L^{(b-r+1)}_j\big)(y, t) + \big(L^{(r)}_{p-2-j} V\big)(x, y) \big(V L^{(b-r+1)}_j\big)(t, t) \Big\}.$$

Similarly,

$$\sum_{\substack{\alpha_i \ge 0,\, \alpha_{b-1} \ge 1,\, \alpha_b = 0 \\ \alpha_1 + \cdots + \alpha_b = p}} \mathbf{E}(H^{\alpha_1} V \cdots V H^{\alpha_b})(x, y)$$

$$= N^{-1} \sum_{t=1}^{N} \sum_{r=1}^{b-1} \sum_{j=0}^{p-2} \mathbf{E}\Big\{ \big(L^{(r)}_{p-2-j} V^2\big)(x, t) \big(V L^{(b-r+1)}_j V\big)(y, t) + \big(L^{(r)}_{p-2-j} V^2\big)(x, y) \big(V L^{(b-r+1)}_j V\big)(t, t) \Big\}.$$

Denoting by l a number such that $\alpha_l \ge 1$ and $\alpha_{l+1} = \cdots = \alpha_b = 0$, we arrive at (5.1). □

Lemma 5.2. *Let us consider the random variables*

$$M_N(\mathbf{a}, \mathbf{b})(x, y) = [H^{\alpha_1} V^{\beta_1} \cdots V^{\beta_{k-1}} H^{\alpha_k} V^{\beta_k}](x, y), \tag{5.2}$$

where the vectors $\mathbf{a} = (\alpha_1, \dots, \alpha_k)$, $\mathbf{b} = (\beta_1, \dots, \beta_k)$ *have positive integer components. Given vectors* $\mathbf{a}_1, \dots, \mathbf{a}_m$, *let us denote* $\zeta \equiv \sum_{l=1}^{m} |\mathbf{a}_l|$, *where* $|\mathbf{a}| = \sum_{s=1}^{k} \alpha_s$ *and*

$$I_N \equiv \mathbf{E}\big\{\overline{M_N(\mathbf{a}_1, \mathbf{b}_1)} \cdots \overline{M_N(\mathbf{a}_m, \mathbf{b}_m)}\big\}, \qquad \overline{M_N(\mathbf{a}, \mathbf{b})} \equiv N^{-1} \operatorname{Tr} M_N(\mathbf{a}, \mathbf{b}).$$

Then:

(i) *If* ζ *is even, then* I_N *is strictly positive.*
(ii) *If* ζ *is odd, then* I_N *is equal to zero.*

Proof. (i) Using (3.2), it is easy to show that I_N with $\zeta = 2q$ is a linear combination of terms of the form (5.2) but with new $\mathbf{a}'_l$, $\mathbf{b}'_l$ such that $\zeta' = 2q - 2$. The coefficients in this linear combination are of the form $\overline{V_N^q} > 0$, $q \ge 0$ multiplied by 1 or $1/N$. Repeating this procedure, we arrive at a finite number of summands that include factors

$$\mathbf{E}\big\{\overline{H_N V_N^a}\, \overline{H_N V_N^b}\big\}, \qquad \mathbf{E}\big\{\overline{H_N^2 V_N^a}\big\} \quad \text{or} \quad \mathbf{E}\big\{\overline{H_N V_N^a H_N V_N^b}\big\}.$$

Easy computations based on (3.1) show that these factors are also positive because of the positivity of V_N. Thus, (i) is proved.

To prove (ii), we use a similar procedure. Starting from I_N with $\zeta = 2q + 1$, we obtain at each step a linear combination of terms of the same form but with σ diminished by 2. At the final stage with $\zeta'' = 1$, we obtain zero due to (2.1a).

Having proved (i) and (ii), it is easy to obtain the relations (4.2).

Next, let us derive (4.3). According to the definition, $D^{(2,q)}_{2j+1}(b_1, b_2)$ is a sum of terms of the form

$$\mathbf{E}\Big\{\overline{L^{(b_1)}_{2q}V}\,\overline{L^{(b_2)}_{2r+1}V}\Big\} \quad \text{or} \quad \mathbf{E}\overline{L^{(b_1)}_{2q}V}\mathbf{E}\overline{L^{(b_2)}_{2r+1}V},$$

that are equal to zero according to (ii).

In conclusion, let us note that

$$\mathbf{E}\overline{L^{(b)}_{2q,N}V_N^{r+s}} \le \tilde{V}_N^s\mathbf{E}\overline{L^{(b)}_{2q,N}V_N^r}, \qquad \tilde{V}_N \equiv \|V_N\|. \tag{5.3}$$

One can easily prove this inequality using the previous argument.

LEMMA 5.3. *Inequality* (4.10) *is valid for arbitrary* $d \in \mathbb{N}$.

PROOF. Let us check (4.10) for $d = 1$. We denote $P = (j+1)/(k+1)$ and observe that the function $(1-P)^2 + P^2$, $1/(k+1) \le P \le k/(k+1)$ takes its maxima at the boundary points. Since $k \ge 1$, then

$$\left(1 - \frac{1}{k+1}\right)^2 + \left(\frac{1}{k+1}\right)^2 = 1 - \frac{2}{k+1} + \frac{2}{(k+1)^2} \le 1 - \frac{1}{2(k+1)}.$$

Now assume that $d \ge 2$ and consider the case $P \le 1/2$. Then

$$\sum_{t=0}^{d}(1-P)^{2(d-t)}P^{2t} \le (1-P)^{2d} + dP^2(1-P)^{2d-2} \equiv \psi(P).$$

The function $\psi(P)$ is strictly decreasing for $1/(k+1) \le P \le 1/2$. Therefore,

$$\pi_k(d) \le \left(1 - \frac{1}{k+1}\right)^{2d-2}\left[\frac{d}{(k+1)^2} + \frac{k^2}{(k+1)^2}\right]. \tag{5.4}$$

Introduce the variable $\tau = (d-1)/(k+1)$. Then (5.4) together with the elementary inequality (4.11) imply that

$$\pi_k(d) \le \frac{e^{-2\tau}}{(k+1)^2}(\tau + 1 + k^2).$$

Since the function $e^{-2\tau}\tau$, $\tau > 0$ is strictly decreasing, then

$$\pi_k(d) \le \frac{1+k^2}{(k+1)^2} < 1 - \frac{1}{2(k+1)}. \qquad \square$$

LEMMA 5.4. *Let* $n^{(q)}_k$ *and* $\hat{n}^{(q)}_k$ *be determined by the relations*

$$n^{(q)}_k = \sum_{j=0}^{k-1} n^{(1)}_{k-1-j}n^{(q+1)}_j, \qquad n^{(q)}_0 = Z_q > 0,\ k, q \in \mathbb{Z}_+,$$

$$\hat{n}^{(q)}_k = \sum_{j=0}^{k-1} \hat{n}^{(1)}_{k-1-j}\hat{n}^{(q+1)}_j, \qquad \hat{n}^{(q)}_0 = \hat{Z}_q > 0,\ k, q \in \mathbb{Z}_+,$$

where $m_0^{(0)} = \hat{m}_0^{(0)} = 1$ *and* Z_q *and* $\hat{Z}_q$ *satisfy*

$$Z_q \le \tilde{V}^q, \qquad \hat{Z}_q \le \tilde{V}^q.$$

Suppose that there exists a positive number ε *such that*

$$|Z_q - \hat{Z}_q| \le \varepsilon \tilde{V}^q.$$

Then

$$\left| n_k^{(0)} - \hat{n}_k^{(0)} \right| \le \varepsilon 8^k \tilde{V}^{2k}. \tag{5.5}$$

PROOF. Obviously we have that

$$n_k^{(q)} \le \tilde{n}_k^{(q)}, \qquad \hat{n}_k^{(q)} \le \tilde{n}_k^{(q)}, \tag{5.6}$$

where

$$\tilde{n}_k^{(q)} = \sum_{j=0}^{k-1} \tilde{n}_{k-1-j}^{(1)} \tilde{n}_j^{(q+1)}, \qquad \tilde{n}_0^{(q)} = \tilde{V}^q,\ k, q \in \mathbb{Z}_+,\ \tilde{n}_0^{(0)} = 1.$$

It is easy to observe that $\tilde{n}_k^{(q)} = \tilde{V}^q n_k$, where the moments n_k are determined by the recurrence relations

$$n_k = \tilde{V}^2 \sum_{j=0}^{k-1} n_{k-1-j} n_j, \qquad n_0 = 1.$$

Denoting $\Delta_k^{(q)} = n_k^{(q)} - \hat{n}_k^{(q)}$, we deduce

$$\Delta_k^{(q)} = \sum_{j=0}^{k-1} \left[\Delta_{k-1-j}^{(1)} \hat{n}_j^{(q+1)} + \hat{n}_{k-1-j}^{(1)} \Delta_j^{(q+1)} \right].$$

Then it is not hard to show that the inequalities $\Delta_0^{(q)} \le \varepsilon \tilde{V}^q$ imply the estimates $\Delta_k^{(q)} \le \varepsilon 2^k (2\tilde{V})^{2k} \tilde{V}^q$. One can check this directly with the help of the last equality and estimates $n_k^{(q)} \le (2\tilde{V})^k \tilde{V}^q$ and $\hat{n}_k^{(q)} \le (2\tilde{V})^k \tilde{V}^q$. These are the consequences of inequalities (5.6) and

$$n_k \le (2\tilde{V})^{2k}.$$

This estimate is valid because, was as proved by Wigner (1955), the n_k are the even moments of the semicircle distribution (1.4) with $v = \tilde{V}$:

$$n_k = \int \lambda^{2k}\, d\sigma_{\tilde{V}}(\lambda). \qquad \square$$

LEMMA 5.5. *Let* V *be given by* (2.22) *with* $u(x) \ge 0$. *Then* (2.20) *is true*.

PROOF. It is easy to see that

$$\begin{aligned} \int_0^\infty \lambda^q\, d\nu(\lambda) &= N^{-1} \sum_{x=1}^{N} V^q(x, x) \\ &= N^{-1} \sum_{x=1}^{N} \sum_{s_i \in \mathbb{Z}} u(x - s_1) u(s_1 - s_2) \cdots u(s_{q-1} - x). \end{aligned}$$

Then

$$V^q(0) - N^{-1}\,\mathrm{Tr}\,V_N^q = N^{-1}\sum_{x=1}^{N}\sum_{s_i}{}' u(x-s_1)u(s_1-s_2)\cdots u(s_{q-1}-x)$$

where Σ' means that the sum is taken over the set of $(s_1,\dots,s_{q-1})$, such that $s_i > N$ for at least one variable s_i. It is obvious that the latter sum is nonnegative and (2.20) is true. □

LEMMA 5.6. *Inequality* (3.13) *is true.*

PROOF. Let us show first that the moments $m_k^{(q)}(N,\varepsilon)$ given by (3.11) satisfy

$$(5.7)\qquad m_k^{(q+1)}(N,\varepsilon) \le (1+\varepsilon)^{1/2}\tilde{V}_N m_k^{(q)}(N,\varepsilon),\qquad k,q\in\mathbb{N}\cup\{0\}.$$

If $k = 0$, then (5.7) obviously follows from definition (3.11b). As for (3.11a), one can easily prove that if (5.7) holds for $m_T^{(q)}(N,\varepsilon)$, $q\in\mathbb{Z}_+$, then the same is true for $m_{T+1}^{(q)}(n,\varepsilon)$, $q\in\mathbb{Z}_+$.

We derive from (5.7) that for $k,q\in\mathbb{N}\cup\{0\}$,

$$m_k^{(q)}(N,\varepsilon) \le (1+\varepsilon)^{3/2}\tilde{V}_N\sum_{j=0}^{k-1} m_{k-1-j}^{(1)}(N,\varepsilon)m_j^{(q)}(N,\varepsilon).$$

Let us introduce the numbers $\tilde{m}_k^{(q)}(N,\varepsilon)$ satisfying the relations

$$\tilde{m}_k^{(q)}(N,\varepsilon) = (1+\varepsilon)^{3/2}\tilde{V}_N\sum_{j=0}^{k-1}\tilde{m}_{k-1-j}^{(1)}(N,\varepsilon)\tilde{m}_j^{(q)}(N,\varepsilon),$$

$$\tilde{m}_0^{(q)} = (1+\varepsilon)^{q/2}\,\overline{V_N^q}.$$

Then

$$(5.8)\qquad m_k^{(q)}(N,\varepsilon) \le \tilde{m}_k^{(q)}(N,\varepsilon).$$

Let us introduce the functions

$$\tilde{f}_{N,\varepsilon}^{(q)}(z) = -\sum_{k=0}^{\infty}\tilde{m}_k^{(q)}(N,\varepsilon)z^{-2k-1}$$

that satisfy the relations

$$(5.9)\qquad \tilde{f}_{N,\varepsilon}^{(q)}(z) = \frac{\tilde{m}_0^{(q)}(N,\varepsilon)}{-z-(1+\varepsilon)^{3/2}\tilde{V}_N\tilde{f}_{N,\varepsilon}^{(1)}(z)},\qquad q\in\mathbb{Z}_+ = \mathbb{N}\cup\{0\}.$$

In Lemma 5.7 we prove that the equation given by (5.9) with $q = 1$ is uniquely solvable in $\mathbf{F}$ and the function $\tilde{f}_N(z)\equiv\tilde{f}_{N,\varepsilon}^{(0)}(z)$ is the Stieltjes transform of a measure $d\tilde{\sigma}_{N,\varepsilon}(\lambda)$ such that

$$\tilde{m}_k^{(0)}(N,\varepsilon) = \int_{-\infty}^{\infty}\lambda^{2k}\,d\tilde{\sigma}_{N,\varepsilon}(\lambda).$$

It is easy to see that $\tilde{f}_N(z)$ has the form

$$\tilde{f}_N(z) = \frac{1}{-z - (1+\varepsilon)^{3/2}\tilde{V}_N \tilde{g}_N(z)},$$

where the function $\tilde{g}_N(z) \in \mathbf{F}$ is the unique solution of the equation

$$\tilde{g}_N(z) = \frac{(1+\varepsilon)^{3/2} v_1^{(N)}}{-z - (1+\varepsilon)^{3/2}\tilde{V}_N \tilde{g}_N(z)}.$$

It follows from the proof of Lemma 5.8 that

$$\operatorname{supp} d\tilde{\sigma}_{N,\varepsilon}(\lambda) \subset (-2(1+\varepsilon)l_N, 2(1+\varepsilon)l_N), \qquad l_N = \left[v_1^{(N)}\tilde{V}_N\right]^{1/2}.$$

Thus,

$$\tilde{m}_k^{(0)}(N,\varepsilon) \le \left[4(1+\varepsilon)^2 l_N^2\right]^k.$$

This inequality combined with (5.8) proves that $m_k^{(0)}(N,\varepsilon) \le [4(1+\varepsilon)^2 l_N^2]^k$. We derive from (5.7) that

$$m_k^{(q)}(N,\varepsilon) \le (1+\varepsilon)^{q/2}\tilde{V}_N^q m_k^{(0)}(N,\varepsilon) \le (1+\varepsilon)^{q/2}\tilde{V}_N^q\left[4(1+\varepsilon)^2 l_N^2\right]^k.$$

This gives (3.13). □

LEMMA 5.7. *The system of equations* (3.20) *considered for* $z \in U$ (3.19) *has a unique solution. This solution satisfies the relations* (3.21). *The equation given by* (3.21) *for* $q = 1$ *is uniquely solvable in the class of functions* $\mathbf{F}$ *defined in Theorem* 2.2.

PROOF. Let us rewrite (3.20) in vector form. To do this, we introduce the linear space $\mathbf{B}$ of vectors $\mathbf{K}$ with components $K_q \in \mathbb{C}$, $q \in \mathbb{Z}_+$ and determine the norm

$$\|\mathbf{K}\|_{\mathbf{B}} = \sup_{q \in \mathbb{Z}_+} \tilde{V}^{-q}|K_q|.$$

Let us introduce a linear operator S_z such that

$$(S_z\mathbf{K})_q = -z^{-1}K_{q+1}, \qquad q \in \mathbb{Z}_+.$$

Then (3.20) can be rewritten in terms of the vector $\mathbf{K}$ with components $K_q \equiv f^{(q)}(z)$ as follows:

$$\mathbf{K} = -z^{-1}\mathbf{M} + K_1 S_z \mathbf{K}, \tag{5.10}$$

where

$$M_q = \hat{m}_0^{(q)} = \int \lambda\, d\nu(\lambda), \qquad q \in \mathbb{Z}_+, \hat{\mathbf{M}} \in \mathbf{B}.$$

We are going to show that (5.10) has a unique solution in $\mathbf{B}$. First let us note that if a vector $\mathbf{K}'$ belongs to $\mathbf{B}$, then the vector $S_z\mathbf{K}'$ also belongs to $\mathbf{B}$. It is easy to observe that for $z \in U$ (3.19) the operator S_z is bounded,

$$\|S_z\|_{\mathbf{B}} = \sup_{\substack{\mathbf{K}' \in \mathbf{B} \\ \|\mathbf{K}'\|_{\mathbf{B}} = 1}} \|S_z\mathbf{K}'\| \le V|z|^{-1} < 1/2.$$

Let us introduce the sequence of vectors $\mathbf{K}^{(m)}$, $m \in \mathbb{Z}_+$ given by the relations

$$\mathbf{K}^{(m+1)} = -z^{-1}\mathbf{M} + K_1^{(m)} S_z \mathbf{K}^{(m)}, \qquad \mathbf{K}^{(0)} = \mathbf{M}.$$

Taking into account that $\|\mathbf{M}\| \le 1$, it is easy to derive that $\|\mathbf{K}^{(m)}\| \le 1$ for all $z \in U$.

The difference

$$\boldsymbol{\psi}_{m+1} = \mathbf{K}^{(m+1)} - \mathbf{K}^{(m)}, \qquad m \in \mathbb{Z}_+$$

can be estimated by

$$\begin{aligned}\|\boldsymbol{\psi}_{m+1}\| &= \left\| \left[K_1^{(m)} - K_1^{(m-1)} \right] S_z \mathbf{K}^{(m)} + K_1^{(m-1)} S_z \left(\mathbf{K}^{(m)} - \mathbf{K}^{(m-1)}\right) \right\| \\ &\le \frac{2\tilde{V}^2}{|z|} \|\boldsymbol{\psi}_m\|.\end{aligned}$$

Now it is clear that the sequence $\mathbf{K}^{(m)}$ converges in $\mathbf{B}$, when $m \to \infty$, to a vector $\mathbf{K}$ satisfying (5.10). Obviously, this is the unique solution.

Now let us prove that (3.21) for $q = 1$ has also a unique solution $f^{(1)}(z)$ that belongs to the class $\mathbf{F}$. We do this with the help of the successive approximations procedure used by Khorunzhy and Pastur (1994). We consider the sequence of functions $f_m^{(1)}(z)$, $m \in \mathbb{Z}_+$ given by the relations

$$f_{m+1}^{(1)}(z) = \int_0^\infty \frac{\lambda \, d\nu(\lambda)}{-z - \lambda f_m^{(1)}(z)}, \qquad f_0^{(1)}(z) = -z^{-1} \int_0^\infty \lambda \, d\nu(\lambda).$$

It is easy to see that $f_m^{(1)}(z) \in \mathbf{F}$ for all $m \in \mathbb{Z}_+$.

We denote $\varphi_{m+1} = f_{m+1}^{(1)} - f_m^{(1)}$ and obtain

$$\varphi_{m+1}(z) = \varphi_m(z) \int \frac{\lambda^2}{\left[z + \lambda f_m^{(1)}(z)\right]\left[z + \lambda f_{m+1}^{(1)}\right]} \, d\nu(\lambda).$$

Hence,

$$|\varphi_{m+1}(z)| \le \tilde{V}^2 |\mathrm{Im}\, z|^{-2} |\varphi_m(z)|.$$

Thus, we have that for $z \in U$ the sequence $f_m^{(1)}(z)$ converges when $m \to \infty$ to a unique function $f^{(1)}(z)$ satisfying (3.21) for $q = 1$.

Since $f_m^{(1)}(z) \in \mathbf{F}$, then these functions represent the Stieltjes transforms of the measures $d\sigma_m(\lambda)$ that converge weakly as $m \to \infty$. Therefore, $f_m^{(1)}(z) \to f^{(1)}(z)$ for all fixed $z \in \mathbb{C} \setminus \mathbb{R}$ and $f^{(1)}(z) \in \mathbf{F}$. Let us note that a similar argument proves the existence of a unique solution in the class $\mathbf{F}$ of equation (5.9) for $q = 1$. □

LEMMA 5.8. *The measure $d\mu(\lambda)$ has a finite support Λ_μ that satisfies* (2.10). *Λ_μ has the property that, if $\lambda \in \Lambda_\mu$, then $-\lambda \in \Lambda_\mu$.*

PROOF. Repeating the arguments of the proof of (5.7), we derive that the moments $\hat{m}_k^{(q)}$ of the measure $d\mu(\lambda)$ satisfy the inequalities

$$\hat{m}_k^{(q)} \le \tilde{m}_k^{(q)}, \qquad k, q \in \mathbb{Z}_+,$$

where the $\tilde{m}_k^{(q)}$ are given by the relations

$$\tilde{m}_k^{(q)} = v_m \sum_{j=0}^{k-1} \tilde{m}_{k-1-j}^{(1)} \tilde{m}_j^{(q)}, \qquad \tilde{m}_0^{(q)} = \int_0^\infty \lambda^q \, d\nu(\lambda).$$

It is easy to see that the functions

$$\tilde{f}^{(q)}(z) = -\sum_{k=0}^{\infty} \tilde{m}_k^{(q)} z^{-2k-1}$$

satisfy the system of equations

$$\tilde{f}^{(q)}(z) = \frac{\tilde{m}_0^{(q)}}{-z - v_m \tilde{f}^{(1)}(z)}, \tag{5.11}$$

which is uniquely solvable (see Lemma 5.7). We see that if $\lambda \neq 0$, then $\operatorname{Im} \tilde{f}^{(0)}(\lambda + i0)$ is equal to zero if and only if $\operatorname{Im} \tilde{f}^{(1)}(\lambda + i0) = 0$.

It follows from (5.11) that

$$\tilde{f}^{(1)}(z) = \frac{-z \pm \sqrt{z^2 - 4v_1 v_m}}{2v_m},$$

where we choose the branch of the square root which satisfies

$$\operatorname{Im} \tilde{f}^{(1)}(z) \operatorname{Im} z > 0, \qquad z \in \mathbb{C} \setminus \mathbb{R}.$$

Thus, $\operatorname{Im} \tilde{f}^{(1)}(\lambda + i0) > 0$ if and only if $|\lambda| < 2\sqrt{v_1 v_m}$. The same statement is valid for $\operatorname{Im} \tilde{f}^{(0)}(\lambda + i0)$. Using (2.18), we obtain that $\tilde{m}_k^{(0)} \leq (2\sqrt{v_1 v_m})^k$ and therefore

$$\hat{m}_k^{(0)} \leq \left(2\sqrt{v_1 v_m}\right)^k \qquad \forall\, k \in \mathbb{N}.$$

Obviously, this implies (2.10).

The symmetry property of Λ_μ follows from the observation that the solution of (3.21) is odd in z

$$f^{(q)}(-z) = -f^{(q)}(z), \qquad q \in \mathbb{Z}_+.$$

According to (2.18), this equality implies that the support Λ_μ is a symmetric set. □

LEMMA 5.9. *Let $V(x, y)$ be the matrix defined in Section 2. Then the bilinear form determined by the matrix $C(x, y; s, t) = V(x, s)V(y, t) + V(x, t)V(y, s)$ on vectors $\boldsymbol{\xi}$ with components $\xi(x, y) \in \mathbb{C}$, $x \leq y$, $x, y \in \mathbb{N}$ such that $\sum_{x \leq y} |\xi(x, y)|^2 < \infty$, is positive.*

PROOF. The covariance criterion [Loève (1978)] provides the existence of a family of random variables $\gamma(x)$, $x \in \mathbb{N}$ with joint Gaussian distribution of zero mathematical expectation and covariance matrix $V(x, y) = \mathbf{E}\gamma(x)\gamma(y)$. Using (3.1), it is easy to show that the random variables $\alpha(x, y) = \gamma(x)\gamma(y) - V(x, y)$, $x \leq y$ are correlated with covariance matrix C,

$$\mathbf{E}(\alpha(x, y)\alpha(s, t)) = V(x, s)V(y, t) + V(x, t)V(y, s).$$

Therefore for any vector $\boldsymbol{\xi} \neq 0$ with a finite number of nonzero components, we have

$$\sum_{x \le y,\, s \le t} C(x, y; s, t)\,\xi(x, y)\overline{\xi(s, t)} = \mathbf{E}\left|\sum_{x \le y} \alpha(x, y)\,\xi(x, y)\right|^2 > 0. \qquad \square$$

Acknowledgments. The authors are grateful to Professor L. Pastur for useful remarks. The authors also express gratitude to the referee who has carefully read the paper and pointed out to us a series of mistakes and misprints. We thank the referee for a number of useful suggestions improving the exposition.

REFERENCES

ANDERSON, T. W. (1984). *Introduction to Multivariate Statistical Analysis*. Wiley, New York.

BAI, Z.D. and YIN, Y. Q. (1988). Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab.* **16** 1729–1741.

BEREZIN, F. (1973). Several notes on the Wigner distribution. *Theoret. and Math. Phys.* **17** 1163–1176.

BOUGEROL, P. and LACROIX, J. (1985). *Products of Random Matrices with Applications to Schroedinger Operator*. Birkhäuser, Boston.

BOUTET DE MONVEL, A. and SHCHERBINA, M. (1995). On the norm of random matrices. *Math. Notes* **57** 475–484.

BOUTET DE MONVEL, A., PASTUR, L. and SHCHERBINA, M. (1995). On the statistical mechanics approach in the random matrix theory: integrated density of states. *J. Statist. Phys.* **79** 585–613.

BOVIER, A., GAYRARD, V. and PICCO, P. (1995). Large deviation principles for the Hopfield model and the Kac–Hopfield model. *Probab. Theory Related Fields* **101** 511–546.

BRONK, B. V. (1964). Accuracy of the semicircle approximation for the density of eigenvalues of random matrices. *J. Math. Phys.* **5** 215–220.

COHEN, J. E., KESTEN, H. and NEWMAN C. M., eds. (1986). *Random Matrices and Their Applications*. Amer. Math. Soc., Providence, RI.

CRISANTI, A., PALADIN, G. and VULPIANI, A. (1993). *Products of Random Matrices in Statistical Physics*. Springer, Berlin.

DI FRANCESCO, P., GINSPARG, P. and ZINN-JUSTIN, J. (1995). 2D gravity and random matrices. *Phys. Rep.* **254** 1–133.

DONOGHUE, W. (1974). *Monotone Matrix Functions and Analytic Continuation*. Springer, Berlin.

GEMAN, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 1318–1328.

GIRKO, V. L. (1975). *Random Matrices*. Naukova Dumka, Kiev (in Russian).

GIRKO, V. L. (1988). *Spectral Theory of Random Matrices*. Nauka, Moscow (in Russian).

GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz Forms and Their Applications*. Univ. California Press, Berkeley.

ISOPI, M. and NEWMAN, C. (1992). The triangle law for Lyapunov exponents of large random matrices. *Comm. Math. Phys.* **143** 591–598.

KHORUNZHY, A. (1996). Eigenvalue distribution of large random matrices with correlated entries. *Matematich. Fizika, Analiz i Geometria* **3** 80–101.

KHORUNZHY, A. and PASTUR, L. (1993). Limits of infinite interaction radius, dimensionality and number of components for random operators with off-diagonal randomness. *Comm. Math. Phys.* **153** 605–646.

KHORUNZHY, A. and PASTUR, L. (1994). On the eigenvalue distribution of the deformed Wigner ensemble of random matrices. *Adv. Soviet Math.* **19** 97–127.

KHORUNZHY, A., KHORUZHENKO, B. and PASTUR, L. (1996). Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.* **37** 5033–5060.

LIFSHITZ, I., GREDESKUL, S. and PASTUR, L. (1988). *Introduction in the Theory of Disordered Systems*. Wiley, New York.

LOÈVE, M. (1978). *Probability Theory* **2**. Springer, Berlin.

MARCHENKO, V. A. and PASTUR, L. A. (1967). Eigenvalue distribution of certain ensembles of random matrices. *Mat. Sbornik* **72** 507–536 (English translation in *Math. USSR Sb.* **1** 457–483).

MEHTA, M. L. (1991). *Random Matrices*. Academic Press, New York.

PASTUR, L. and FIGOTIN, A. (1992). *Spectra of Random and Almost Periodic Operators*. Springer, New York.

PORTER, C. E., ed. (1965). *Statistical Theory of Spectra: Fluctuations*. Academic Press, New York.

SINAI, YA. and SOSHNIKOV, A. (1998). Central limit theorem for traces of large symmetric matrices with independent matrix elements. *Bol. Soc. Bras. Mat.* **29** 1–24.

TRACY, C. A. and WIDOM, H. (1994). Level spacing distribution and the Airy kernel. *Comm. Math. Phys.* **161** 289–309.

VOICULESCU, D., DYKEMA, K. J. and NICA, A. (1992). *Free Random Variables*. Amer. Math. Soc., Providence, RI.

WEGNER, F. (1979). Disordered system with n orbitals per site: $n = \infty$ limit. *Phys. Rev. B* **19** 783–792.

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

INSTITUT DE MATHÉMATIQUES DE JUSSIEU
UNIVERSITÉ PARIS 7-DENIS DIDEROT
2, PLACE JUSSIEU
75251 PARIS CEDEX 05
FRANCE
E-MAIL: anne@math.jussieu.fr

MATHEMATICAL DIVISION
INSTITUTE FOR LOW TEMPERATURE PHYSICS
KHARKOV 310164
UKRAINE
E-MAIL: khorunjy@ilt.kharkov.ua