A SIMPLE APPROACH TO GLOBAL REGIME OF THE RANDOM MATRIX THEORY

LEONID PASTUR

ABSTRACT. We discuss a method of the asymptotic computation of moments of the normalized eigenvalue counting measure of random matrices of large order. The method is based on the resolvent identity and on some formulas relating expectations of certain matrix functions and the expectations including their derivatives or, equivalently, on some simple formulas of the perturbation theory. In the framework of this unique approach we obtain functional equations for the Stieltjes transforms of the limiting normalized eigenvalue counting measure and the bounds for the rate of convergence for the majority known random matrix ensembles.

1. INTRODUCTION

Random matrix theory is actively developing. Among numerous topics of the theory and its various applications those related to the asymptotic eigenvalue distribution of random matrices of large order are of considerable interest. An important role in this branch of the theory plays the eigenvalue counting measure defined for any Hermitian or real symmetric matrix

$$M_n = \left\{ M_{jk}^{(n)} \right\}_{j,k=1}^{n} \tag{1.1}$$

as follows

$$N_n(\Delta) = \frac{1}{n} \natural \left\{ \lambda_i^{(n)} \in \Delta \right\} \tag{1.2}$$

where Δ is a Borel set of the real axis $\mathbf{R}$ and $\{\lambda_i^{(n)}\}_{i=1}^{n}$ are eigenvalues of M_n.

One distinguishes several large-n asymptotic regimes for the probability properties of eigenvalues (see e.g. [1]). In this paper we deal with the global regime, defined by the requirement that the expectation $\mathbf{E}(N_n(\Delta))$ has well defined (i.e. not zero and not infinite) weak limit

$$N(\Delta) = \lim_{n\to\infty} \mathbf{E}(N_n(\Delta)). \tag{1.3}$$

This limit is called the Integrated Density of States (IDS). We shall see below that explicit conditions to be in the global regime may look differently in different cases.

The IDS is a quantity to be found and analyzed first in any random matrix study, because it enters in practically any problem and result of the theory. A new wave of interest to the global regime is motivated by recent studies of the free group factors of operator algebras known now as free probability theory [2].

Research at MSRI is supported in part by NSF grant DMS-9701755.

In this paper we present a simple approach of the study of the random measure (1.2) in the global regime. The approach allows us to find the limit (1.3) in many interesting cases, to show that the sequence of random measures (1.2) converges to this nonrandom limit either in probability or even with probability 1 and to find bounds on the rate of the convergence.

The paper is organized as follows. In Section 2 we consider most studied random matrix ensembles, in particular, the Gaussian and the Circular Ensembles. This allows us to explain the method by using the simple and known setting of these ensembles. We call these ensembles classical because, first, they are indeed classic objects of the theory, and second, because of the role of the classical orthogonal polynomials in their studies (although, we almost do not use this technique in the paper). Section 3 is devoted to ensembles whose probability distribution is invariant with respect to unitary or orthogonal transformations and whose study is motivated by the Quantum Field Theory. In Section 4 we present new results on the form of the limiting normalized eigenvalue counting measure of the sum of two Hermitian or real symmetric matrices randomly rotated one with respect to another. In Section 5 we first discuss ensembles with independent but not necessary Gaussian entries. These ensembles are known as the Wigner Ensembles. Then we consider the ensembles that can be represented as the sum of the rank one independent operators. This form generalizes the sample covariance matrices widely used in multivariate analysis and is also motivated by statistical mechanics.

Most of results, presented in the paper are known, sometimes for decades. However they were obtained by different and often rather complicated methods while in this paper we derive them in the framework of an unique approach, that we present in three slightly different versions, according to the case considered. We do not give here complete proofs of all presented results, but only outline basic moments of their proofs. The complete versions of the proofs will be published in [3].

2. Classical Ensembles

2.1. **Gaussian Ensembles.** We start from the well known ensembles among which the Gaussian Ensembles (GE) are most known. We restrict ourself by the technically simplest of Gaussian Ensembles, consisting of Hermitian matrices and known as the Gaussian Unitary Ensemble (GUE) because its probability distribution

$$P_n(dM) = Z_n^{-1} \exp\left(-\frac{n}{4w^2} \operatorname{Tr} M^2\right) dM \tag{2.1}$$

is unitary invariant. Here

$$dM = \prod_{1 \le j \le n} dM_{jj} \prod_{1 \le j < k \le n} d\mathrm{Re} M_{jk} d\mathrm{Im} M_{jk} \tag{2.2}$$

is the "Lebesgue" measure on the space of $n \times n$ Hermitian matrices. It is often convenient to write

$$M_n = n^{-1/2} W_n \tag{2.3}$$

where now $W_n = \{W_{jk}\}_{j,k=1}^n$ can be considered as the left upper corner of the semi-infinite Hermitian matrix $W = \{W_{jk}\}_{j,k=1}^\infty$, whose entries are complex Gaussian random variables defined by the relations

$$\mathbf{E}(W_{jk}) = 0, \quad \mathbf{E}(W_{jk}\overline{W_{lm}}) = 2w^2\delta_{jl}\delta_{km}. \tag{2.4}$$

Thus the probability space in this case consists of these matrices and has as a probability measure the infinite product of the Gaussian measures defined by relations (2.4).

The relations (2.3) and (2.4) define the global regime in this case.

Theorem 2.1. *For the Gaussian Unitary Ensemble defined above the sequence of eigenvalue counting measures (1.2) converges with probability 1 to the nonrandom measure*

$$N_{sc}(\Delta) = \frac{1}{4\pi w^2}\int_{\Delta\cap[-2\sqrt{2}w,2\sqrt{2}w]} \sqrt{8w^2-\lambda^2}\,d\lambda \tag{2.5}$$

i.e. N_{sc} *has the density*

$$\rho_{sc}(\lambda) = (4\pi w^2)^{-1}\sqrt{8w^2-\lambda^2} \tag{2.6}$$

concentrated on the interval $[-2\sqrt{2}w, 2\sqrt{2}w]$. *The convergence* N_n *to* N_{sc} *has to be understood as the weak convergence of measures.*

The theorem dates back in fact to Wigner (see [1]). We give below the two proofs of the theorem to illustrate two methods that can be used in rather general situation of dependent and not necessary Gaussian entries. Both proofs as well as other proofs in this paper are based on the study the *Stieltjes transforms* of measures instead of measures themselves. In the random matrix theory the Stieltjes transform was used for the first time in paper [14] and since then is proved to be a rather efficient tool of the study of the global regime.

Recall that the Stieltjes transform $f(z)$ of a non-negative measure $m(d\lambda)$, $m(\mathbf{R}) = 1$, is the function of the complex variable z defined for all non-real z by the integral

$$f(z) = \int \frac{m(d\lambda)}{\lambda - z}, \quad \mathrm{Im} z \neq 0. \tag{2.7}$$

Here and below we use integrals without indicated limits denote integrals over whole real axis. $f(z)$ is obviously analytic for non-real z and satisfies the conditions

$$\mathrm{Im} f \cdot \mathrm{Im} z > 0, \mathrm{Im} z \neq 0, \quad \sup_{y\geq 1} y|f(iy)| = 1. \tag{2.8}$$

It can be shown [4] that any function $f(z)$ defined and analytic for non-real z and satisfying conditions (2.8) is the Stieltjes transform of a unique nonnegative and normalized to 1 measure $m(d\lambda)$ and that for any continuous function $\varphi(\lambda)$ with a compact support

$$\int \varphi(\lambda)m(d\lambda) = \lim_{\varepsilon\to 0}\frac{1}{\pi}\int \varphi(\lambda)\mathrm{Im} f(\lambda + i\varepsilon)d\lambda \tag{2.9}$$

Besides the one-to-one correspondence between measures and their Stieltjes transforms is continuous if one will consider the weak convergence of measures and the convergence of the Stieltjes transforms that is uniform on all compacts in $\mathbf{C}\backslash\mathbf{R}$.

The use of the Stieltjes transform in this context is based on the spectral theorem expressing the Stieltjes transform

$$g_n(z) = \int \frac{N_n(d\lambda)}{\lambda - z} \tag{2.10}$$

of the normalized eigenvalue counting measure (1.2) of a matrix M_n via its resolvent

$$G(z) = (M_n - z)^{-1}, \quad \mathrm{Im} z \neq 0 \tag{2.11}$$

by the formula

$$g_n(z) = \frac{1}{n} \operatorname{Tr} G(z) \tag{2.12}$$

Our proof has as a basic ingredient the following

Proposition 2.1. *The Stieltjes transform $g_n(z)$ of the eigenvalue counting measure (1.2) for the Gaussian Unitary Ensemble defined by (2.1) or by (2.3) and (2.4) has the asymptotic properties for $Imz \geq y_0 = 5w$:*

$$\lim_{n\to\infty} \mathbf{E}(g_n(z)) = f_{sc}(z), \tag{2.13}$$

$$\mathbf{E}(|\gamma_n(z)|^2) \leq \frac{C}{w^2 n^2} \tag{2.14}$$

where

$$\gamma_n(z) = g_n(z) - \mathbf{E}(g_n(z)), \tag{2.15}$$

$f_{sc}(z)$ is the unique solution of the equation

$$2w^2 f^2 + zf + 1 = 0 \tag{2.16}$$

verifying condition (2.8), and we denote here and below by C numerical constants that may be different in different formulas.

To prove the proposition we need the following elementary facts:

(i) for any two matrices A and B

$$(B - z)^{-1} = (A - z)^{-1} - (B - z)^{-1}(B - A)(A - z)^{-1}, \tag{2.17}$$

(the resolvent identity);

(ii) if ζ is a complex valued Gaussian random variable defined by

$$\mathbf{E}(\zeta) = \mathbf{E}(\zeta^2) = 0, \quad \mathbf{E}(|\zeta|^2) = 2w^2$$

and $\varphi(z, \bar{z})$ is a differentiable function polynomially growing at infinity and having the same property of its derivatives, then

$$\mathbf{E}(\zeta\varphi(\zeta, \bar{\zeta})) = \mathbf{E}(|\zeta|^2)\mathbf{E}\left(\frac{\partial\varphi}{\partial\bar{\zeta}}\right) = 2w^2\mathbf{E}\left(\frac{\partial\varphi}{\partial\bar{\zeta}}\right); \tag{2.18}$$

(iii) for the resolvent $G(z) = (A - z)^{-1}$ of any Hermitian or real symmetric matrix A we have

$$||G(z)|| \leq |\mathrm{Im} z|^{-1}, \quad |G_{jk}(z)| \leq |\mathrm{Im} z|^{-1} \tag{2.19}$$

where $G_{jk}(z), j, k = 1...n$ are the matrix elements of the resolvent.

Now using (2.17) for the pair $B = M, A = 0$ and applying (2.18) we obtain the system of identities for the moments

$$m_p(z_1, ..., z_p) = \mathbf{E}(g_n(z_1)...g_n(z_p)) \tag{2.20}$$

of the random function g_n with non-real arguments $z_1, ..., z_p$

$$m_1(z_1) = -\frac{1}{z_1} - \frac{2w^2}{z_1} m_2(z_{1,} z_1), \tag{2.21}$$

$$m_p(z_1, ..., z_p) = -\frac{1}{z_1} m_{p-1}(z_2, ..., z_p) - \frac{2w^2}{z_1} m_{p+1}(z_1, z_{1,} z_{2,} ..., z_p) + r_p(z_1, ..., z_p), p \geq 2.$$

where

$$r_p(z_1, ..., z_p) = \frac{2w^2}{n^2 z_1} \sum_{q=2}^{p} \mathbf{E}(\frac{1}{n} \text{ Tr } (G_n(z_1) G_n^2(z_q)) g_n(z_2)...g_n(z_{q-1}) g_n(z_{q+1})...g_n(z_p),)$$

Assume now that $|\mathrm{Im} z_q| \geq y, q = 1, ...$ for some $y > 0$. Then relations (2.12) and (2.19) imply the bounds

$$|g_n(z)| \leq \frac{1}{|\mathrm{Im} z|} \leq \frac{1}{y}, \tag{2.22}$$

$$|m_p| \leq \frac{1}{y^p}, \quad |r_p| \leq \frac{2w^2 p}{n^2 y^{p+1}}.$$

Following statistical mechanics (see e.g.[5]) we can treat system of identities (2.21) as a linear equation in the Banach space $\mathcal{B}$ of complex valued sequences m of functions $m = \{m_p(z_1, ..., z_p)\}_{p=1}^{\infty}$ equipped with the norm

$$||m|| = \sup_{p \geq 1} \eta^p \sup_{|\mathrm{Im} z_q| \geq y, q=1,..p} |m_p(z_1, ..., z_p)| \tag{2.23}$$

for some $\eta > 0$. The equation has the form

$$m = Am + b + r \tag{2.24}$$

where

$$\begin{aligned} (Am)_1(z_1) &= -\frac{2w^2}{z_1} m_2, \\ (Am)_p(z_1, ..., z_p) &= -\frac{2w^2}{z_1} m_{p-1}(z_2, ..., z_{p-1}) - \frac{2w^2}{z_1} m_{p+1}(z_1, z_1, z_2, ..., z_{p-1}), p \geq 2, \end{aligned}$$

and $b = (-z_1^{-1}, 0, ...)$, $\quad r = \{r_p\}_{p=1}^{\infty}$. It is easy to show that optimal with respect to η bound for the norm of A is $||A|| \leq 2^{\frac{3}{2}}\ wy^{-1}$ for $\eta = \sqrt{2}w$. Choosing say $y \geq 5w$ we have uniformly in n that the norm of A is strictly less than 1, the vectors m, b and r belong to $\mathcal{B}$ and

$$||r|| \leq \frac{C}{n^2}. \tag{2.25}$$

Thus equation (2.24) is uniquely soluble uniformly in n. It is easy to check that this equation with r replaced by zero has the factorized solution

$$m_p^{(0)}(z_1, ..., z_p) = \prod_{q=1}^{p} f(z_q)$$

where f verifies (2.16). Thus, in view of (2.25) we have

$$|m_p(z_1, ..., z_p) - \prod_{q=1}^{p} f(z_q)| \leq \frac{C}{w^p n^2} \tag{2.26}$$

uniformly in

$$|\mathrm{Im} z| \geq 5w. \tag{2.27}$$

In particular we obtain that

$$\begin{aligned} \lim_{n\to\infty} m_1(z) &= f(z), \\ |m_2(z, \overline{z}) - m_1(z)m_1(\overline{z})| &\leq \frac{C}{n^2 w^2} \end{aligned}$$

Recalling definition (2.20) of m_p we see that the first relation implies that $f(z)$ satisfies (2.8). The unique solubility of (2.16) in this class can be easily verified. The second relation is just another form of (2.14). The proposition is proved.

Remark. As was mentioned the technique of the proof is similar to the technique of the correlation equations of the statistical mechanics (the Kirkwood-Salzburg equations, the Montroll-Mayer equations, etc. [5]) combined with the mean field approximation also widely used in the statistical mechanics.The reason to have here an analogue of the mean field approximation regime is again similar to that of statistical mechanics: the entries of the GUE matrices are all of the same order of magnitude (see (2.3) - (2.4)), like interactions in the Curie-Weiss model. Thus to obtain a nontrivial (not zero and not infinite) limit (1.3) (an analogue of an extensive quantity per unit volume in the statistical mechanics) we have to introduce the n-dependent normalizing factor $n^{-1/2}$ in (2.3). This fixes the global regime scaling but also leads to vanishing of the statistical correlations and to the factorization of the moments (see (2.14) and (2.26)) and to a nonlinear self-consistent equation determining the first moment that can be regarded as an analogue of the Curie-Weiss equation for the magnetization.

To prove Theorem 2.1 denote by N_{sc} the measure (1.2) corresponding to $f_{sc}(z)$ via (2.9). In view of Proposition 2.1 and the Borel-Cantelli lemma we have for any fixed z with $|\mathrm{Im} z| \geq 5w$ the convergence of $g_n(z)$ to $f_{sc}(z)$ with probability 1 for any fixed z. Since any analytic function is uniquely determined by its values on a countable set having at least one accumulation point we have the convergence of g_n to f with probability 1 on any compact of the domain $|\mathrm{Im} z| \geq 5w$. Now by continuity of the correspondence between measures and their Stieltjes transforms we obtain that the measure N_n converges weakly to N_{sc} with probability 1. Solving explicitly equation (2.16) in the class (2.8) we find that

$$f_{sc}(z) = \frac{1}{4w^2}(\sqrt{z^2 - 8w^2} - z) \tag{2.28}$$

where the radical is defined by the condition that it behaves as z as $z \to \infty$. By using the inversion formula (2.9) we obtain (2.5). Theorem 2.1 is proved.

The method was proposed in [6] and subsequently used in [7, 8, 9, 10] to study a variety of problem of random matrix theory and its applications. A certain disadvantage of the method is that it is rather tedious. An important simplification of the method was proposed by A.Khorunzhy [11]. We describe now this simpler version by giving another proof of the previous proposition.

Rewrite the first two equations of the system (2.21) for $z_1 = z, z_2 = \bar{z}$ in the form

$$\mathbf{E}(g_n(z)) = -\frac{1}{z} - \frac{2w^2}{z}\mathbf{E}(g_n^2(z)), \tag{2.29}$$

$$\mathbf{E}(|g_n(z)|^2) = -\frac{1}{z}\mathbf{E}(\overline{g_n(z)}) - \frac{2w^2}{z}\mathbf{E}(g_n^2(z)\overline{g_n(z)}) + r_2(z). \tag{2.30}$$

Expressing $-1/z$ in the first term of the r.h.s. of (2.30) from (2.29) we obtain

$$\mathbf{E}(|\gamma_n(z)|^2) = -\frac{2w^2}{z}\mathbf{E}(g_n^2(z)\overline{\gamma_n(z)}) + r_2(z) \tag{2.31}$$

where $\gamma_n(z)$ is defined in (2.15). By using this definition we can rewrite the expectation in the first term in the r.h.s. of this relation as $\mathbf{E}(g_n(z)|\gamma_n(z)|^2)+\mathbf{E}(g_n(z))\mathbf{E}(|\gamma_n(z)|^2)$. Hence in view of (2.19)

$$|\mathbf{E}(g_n^2(z)\overline{\gamma_n(z)})| \leq \frac{2}{|\mathrm{Im} z|}\mathbf{E}(|\gamma_n(z)|^2). \tag{2.32}$$

By using this inequality and (2.22) we get from (2.31)

$$(1 - \frac{4w^2}{|\mathrm{Im} z|^2})\mathbf{E}(|\gamma_n(z)|^2) \leq \frac{2w^2}{|\mathrm{Im} z|^4 n^2}. \tag{2.33}$$

Thus, we obtain the bound (2.14) on the variance of $g_n(z)$ under condition (2.27). By using this bound we replace $\mathbf{E}(g_n^2(z))$ in (2.29) by $\mathbf{E}^2(g_n(z))$ with an error of the order $O(1/n^2)$ uniformly in the domain (2.27). Now by using standard compactness arguments we can prove that any subsequence of the sequence $\mathbf{E}(g_n(z))$ converges uniformly on compacts of the domain (2.27) to a solution of equation (2.16) satisfying (2.8). Since this solution is unique we obtain other assertions of the Proposition 2.1.

Remarks. (i). Analogous results are valid for two other widely used ensembles, the Gaussian Orthogonal Ensemble (GOE) consisting of real symmetric matrices distributed according the real analogue of (2.1), and for the Gaussian Symplectic Ensemble (GSE), consisting of self-dual Hermitian matrices having also the Gaussian probability distribution (see [1] for definitions and properties). This allows us to write that the density of the semicircle law of all three cases is concentrated on the interval $[-2\sqrt{\beta}w, 2\sqrt{\beta}w]$ and has the form

$$\rho_{sc\beta}(\lambda) = \frac{1}{2\beta\pi w^2}\sqrt{4\beta w^2 - \lambda^2}, \quad |\lambda| \le 2\sqrt{\beta}w. \tag{2.34}$$

for the GOE ($\beta = 1$), the GUE ($\beta = 2$), and the GSE ($\beta = 4$) cases respectively.

(ii) We can consider an ensemble of the more general form

$$H_n = H_n^{(0)} + M_n \tag{2.35}$$

where M_n is as before and $H_n^{(0)}$ is a matrix whose eigenvalue counting measure $N_n^{(0)}$ has a weak limit N_0 as $n \to \infty$. We denote by $f_0(z)$ the Stieltjes transform of N_0. In this case we use the resolvent formula (2.17) for $B = H_n$ and $A = H_n^{(0)}$ and a natural extension of the above arguments. We obtain an analogue of Theorem 2.1 in which the Stieltjes transform $f(z)$ of the limiting counting measure N is a unique solution of the functional equation

$$f(z) = f_0(z + 2w^2 f(z)) \tag{2.36}$$

belonging to the class (2.8) . The equation defines the *the deformed semicircle law* (see [9] for its properties).

2.2. **Laguerre ensemble.**

The ensemble is defined as

$$M_n = \frac{1}{n}A_n A_n^* \tag{2.37}$$

where the $n \times n$ matrix A_n has the probability distribution (cf.(2.1))

$$P(dA) = Z_n^{-1} \exp\left(-\frac{1}{2a^2}\ \mathrm{Tr}\ AA^*\right) dA \tag{2.38}$$

$$dA = \prod_{j,k=1}^{n} d\mathrm{Re}A_{jk}\mathrm{Im}A_{jk}.$$

In other words the entries $A_{jk}, j, k = 1, ..., n$ of A_n are independent complex Gaussian random variables defined by

$$\mathbf{E}(A_{jk}) = \mathbf{E}(A_{jk}^2) = 0, \quad \mathbf{E}(|A_{jk}|^2) = 2a^2.$$

Note that the matrix A is not Hermitian. The name of the ensemble is recent (see e.g. [12]) and is related to the fact that in the orthogonal polynomial approach [1] one uses in the case of this ensemble the Laguerre polynomials (recall that in the case of the

Gaussian Unitary Ensemble one uses the Hermite polynomials). The ensemble models a generic positively defined matrix. The real symmetric version of the ensemble in which A is $n\times m$ matrix with statistically independent Gaussian entries is well known since the 30's in the multivariate analysis as the Wishart distribution and describes the sample covariance matrix of m random Gaussian n-dimensional vectors [13].

Theorem 2.2. *Let the Laguerre ensemble of random matrices be defined as above. Then its eigenvalue counting measure converges in probability 1 to the nonrandom measure of the form*

$$N_L(\Delta) = \frac{1}{4\pi a^2}\int_{\Delta\cap[0,8w^2]}\sqrt{\frac{8a^2-\lambda}{\lambda}}d\lambda \tag{2.39}$$

The proof of the theorem follows the scheme of that of Theorem 2.1, that is it is based on an analogue of Proposition 2.1. In particular, we have here an analogue of the important inequality (2.14). However, the analogue of (2.29) has the form

$$\mathbf{E}(g_n(z)) = -\frac{1}{z} - 2a^2\mathbf{E}(g_n^2).$$

As a result, the corresponding quadratic equation is $2a^2zf^2+zf+1=0$ (cf.(2.16)). This leads to (2.39).

Remarks. (i). Analogous results are valid for the cases when the matrix A_n is real or quaternion. Thus we obtain the general formula for the density of the limiting measure

$$\rho_{L\beta}(\lambda) = \frac{1}{2\beta\pi a^2}\sqrt{\frac{4\beta a^2-\lambda}{\lambda}}, \quad 0\le\lambda\le 4\beta a^2, \tag{2.40}$$

that is concentrated on the interval $[0,4\beta a^2]$ for the orthogonal ($\beta=1$), complex ($\beta=2$) and quaternion ($\beta=4$) cases respectively.

(ii). We can consider an ensemble of more general form

$$H_n = H_n^{(0)} + M_n \tag{2.41}$$

where M_n is as in (2.37) and $H_n^{(0)}$ is a matrix whose eigenvalue counting measure $N_n^{(0)}$ has a weak limit N_0 as $n\to\infty$. We denote by $f_0(z)$ the Stieltjes transform of N_0. In this case we use the resolvent formula (2.17) for $B=H_n$ and $A=H_n^{(0)}$ and a natural extension of the above arguments. We obtain an analogue of (2.36) in which the Stieltjes transform $f(z)$ of the limiting counting measure N is a unique solution of the functional equation

$$f(z) = f_0\left(z - \frac{2a^2}{1+2a^2f}\right) \tag{2.42}$$

belonging to the class (2.8).

(iii) We can also consider a more general case when the random matrix A_n is the $n\times m$ matrix with Gaussian i.i.d. entries. This case can be treated similarly.

We discuss this case in more detail in Section 5 considering arbitrary distributed independent entries.

2.3. **Circular Ensemble.** The ensemble consists of $n \times n$ unitary matrices whose probability distribution is given by the normalized Haar measure on $U(n)$. The ensemble was introduced by Dyson together with its orthogonal and symplectic analogues (see [1] for references and results). We discuss below the simplest Circular Unitary Ensemble (CUE) but one can obtain similar results for two other ensembles.

It is useful to write eigenvalues λ_j of the ensemble in the form $\lambda_j = e^{i\theta_j}, 0 \le \theta_j < 2\pi, j = 1, ..., n$ and to introduce the normalized counting measure (cf.(1.2))

$$N_n(\Delta) = \frac{1}{n}\natural\{\theta_j \in \Delta\} \tag{2.43}$$

where Δ is Borel set of $[0, 2\pi)$. An analogue of the Stieltjes transform (2.7) for measures on the unit circle is the *Herglotz transform* [4]

$$h(z) = \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} m(d\theta), \quad |z| \ne 1. \tag{2.44}$$

Respective inversion formula is (cf. (2.9))

$$\int_0^{2\pi} \varphi(\theta) m(d\theta) = \lim_{r \to 1-0} \frac{1}{2\pi} \int_0^{2\pi} \varphi(\theta) \mathrm{Re} h(re^{i\theta}) d\theta.$$

We use here instead of (2.18) the differentiation formula

$$\mathbf{E}(\varphi'(M)AM) = \mathbf{E}(\varphi'(M)MA) = 0 \tag{2.45}$$

valid for any C^1 function $\varphi : U(n) \to \mathbf{C}$ and any Hermitian matrix A. This formula and the spectral theorem for unitary matrices according to which the Herglotz transform $h_n(z)$ of the eigenvalue counting measure can be written as (cf. (2.12))

$$h_n(z) = \frac{1}{n} \operatorname{Tr} \frac{U + z}{U - z} \tag{2.46}$$

allow us to write the following relations for the moments of $h_n(z)$:

$$\mathbf{E}(h_n(z)) = -1, \quad |z| \ne 1, \tag{2.47}$$

$$\mathbf{E}(|h_n(z)|^2) - |\mathbf{E}(h_n(z))|^2 \le \frac{C}{n^2}, \quad |z| \le \frac{1}{4}. \tag{2.48}$$

The first relation shows that $\mathbf{E}(N_n(\Delta)) = |\Delta|/2\pi$ for all n. This is easy to understand because the Haar measure is shift invariant. The second relation plays the role of (2.14). By using these relations and following the scheme of proof of Theorem 2.1 one obtains

Theorem 2.3. *Consider the ensemble of unitary matrices distributed according to the Haar measure on $U(n)$ (the CUE). Then the eigenvalue counting measure (2.43) of the ensemble converge in probability to the uniform measure on the unit circle.*

Remarks. (i). Analogous statements are valid also for the Circular Orthogonal Ensemble and for the Circular Symplectic Ensemble (see [1] for their definitions and properties).

(ii). Note that unlike Theorems 2.1 and 2.2 where we have the convergence with probability 1, Theorem 2.3 asserts only the convergence in probability, despite the bound (2.48). The reason is that in Theorems 2.1 and 2.2 we can consider W_n and A_n for all n as defined on the same probability space of realizations of the semi-infinite matrices $W = \{W_{jk}\}_{j,k=1}^{\infty}$ and $A = \{A_{jk}\}_{j,k=1}^{\infty}$ equipped with the infinite product Gaussian measure. It is clear that similar natural and simple embedding does not exist for unitary matrices.[1] This case can be regarded as an analogue of the triangular array scheme of probability and the Theorem 2.3 is an analogue of the (Tchebyshev) law of large numbers, while the case of the Gaussian and the Laguerre ensembles is analogous to the scheme of infinite number of i.i.d. random variables and Theorems 2.1 and 2.2 are analogues of the strong law of large numbers. To deduce the convergence in probability of $N_n(\Delta)$ for a fixed Borelien Δ from the convergence in probability its Herglotz transforms for a fixed $z, |z| \leq \frac{1}{4}$ one has to use the argument of Section 4 of [14].

(iii). The simple differentiation formula (2.45) as well as its version (4.5) below allows one to give a direct proof of the asymptotic freeness of unitary and diagonal matrices as $n \to \infty$ (see [2] for definitions and results and [16, 26, 17] for some related recent results). Existing proofs are based on the representation of the Haar distributed unitary random matrices U as the phase in the polar decomposition of the Gaussian distributed random matrix X with complex i.i.d. entries and on the approximation of the phase by polynomials in X. Because of singularities of the polar decomposition representation $U = X(XX^*)^{-1/2}$ these proofs are not simple to implement in all details. The approach based on the formula (4.5), i.e. on the shift invariance of the Haar measure, seems more direct and simple (see [23] and Remark (iv) of Section 4).

3. Invariant Ensembles

In this Section we discuss the random matrix ensembles defined by the probability law

$$P(dM) = Z_n^{-1} \exp(-n \text{ Tr } V(M))dM \tag{3.1}$$

where M is a real symmetric or a Hermitian or a quaternion self-dual Hermitian matrix and V is a bounded below and growing sufficiently fast at infinity function of respective matrix. For $V(\lambda) = \lambda^2/4w^2$ we obtain the Gaussian Ensembles that were considered in the previous Section. In this paper we restrict ourselves by polynomial V's. As it was in the case of the Gaussian Ensembles we discuss here the technically simplest Hermitian matrices. This subclass of ensembles (3.1) is motivated by Quantum Field Theory (see e.g. review [18]). Following Quantum Field Theory we will

[1] Although one can always use the probability space that is the product over all n of the probability spaces consisting of the groups $U(n)$ with the normalized Haar measure as the probability measure.

call V the potential. We will give below a result for convex V's. More general case see in [3].

Theorem 3.1. *Consider the random matrix ensemble consisting of Hermitian $n \times n$ matrices distributed according to (3.1) in which the potential V is a convex polynomial of an even degree $2p$ and growing at infinity. Then the eigenvalue counting measure of the ensemble converges in probability* [2] *to the nonrandom measure whose density is concentrated on the interval (a,b) and has the form*

$$\rho(\lambda) = p_{2p-2}(\lambda)\sqrt{(b-\lambda)(\lambda-a)}, \quad a \le \lambda \le b \tag{3.2}$$

where

$$p_{2p-2}(\lambda) = \frac{1}{\pi^2}\int_a^b \frac{V'(\lambda)-V'(\mu)}{\lambda-\mu}\frac{d\mu}{\sqrt{(b-\mu)(\mu-a)}} \tag{3.3}$$

is a positive on (a,b) polynomial of the degree $2p-2$ and a and b are uniquely defined by the equations

$$\int_a^b \frac{\mu^q V'(\mu)d\mu}{\sqrt{(b-\mu)(\mu-a)}} = 2\pi\delta_{1q}, \quad q = 0,1. \tag{3.4}$$

The proof of the theorem follows again the scheme of the proof of Theorem 2.1. In particular we have the analogue of Proposition 2.1

Proposition 3.1. *Under the conditions of the preceding theorem the Stieltjes transform $g_n(z)$ of the eigenvalue counting measure of the ensemble (3.1) has the following properties for $|Imz| \ge y$ and a certain y depending on V*

$$\lim_{n\to\infty} \mathbf{E}(g(z)) = f(z), \tag{3.5}$$

$$\mathbf{E}(|g_n^2(z)|) - |\mathbf{E}(g_n(z))| \le \frac{const}{n^2} \tag{3.6}$$

where $f(z)$ is a unique solution of the quadratic equation

$$f^2 + V'(z)f + Q(z) = 0 \tag{3.7}$$

satisfying (2.8) and

$$Q(z) = \int \frac{V'(z)-V'(\mu)}{z-\mu}N(d\mu) \tag{3.8}$$

in which $N(d\mu)$ is the measure corresponding to f.

To prove the proposition we use the differentiation formula [15]

$$\mathbf{E}(\varphi'(M)\cdot B) - n\mathbf{E}(\varphi(M) \text{ Tr } V'(M)B) = 0 \tag{3.9}$$

valid for the matrix distribution (3.1), any function $\varphi : \mathbf{R} \to \mathbf{C}$ whose derivative is polynomially bounded on the whole real line and any Hermitian matrix B. By

[2] See Remark (ii) after Theorem 2.3.

applying this formula to the matrix element of the resolvent $G(z) = (M - z)^{-1}$ we obtain the relation

$$\mathbf{E}(g_n^2(z)) + \mathbf{E}(\frac{1}{n} \text{ Tr } G(z)V'(M)) = 0 \tag{3.10}$$

where $g_n(z)$ is defined in (2.12). The identity

$$G(z)V'(M) = G(z)V'(z) + G(z)(V'(M) - V'(z)) \tag{3.11}$$

allows us to rewrite this relation in the form

$$\mathbf{E}(g_n^2(z)) + V'(z)\mathbf{E}(g_n(z)) + Q_n(z) = 0 \tag{3.12}$$

where $Q_n(z)$ is defined as

$$Q_n(z) = \mathbf{E}(\frac{1}{n} \text{ Tr } Q_n(z, M)), \quad Q_n(z, M) = G(z)(V'(M) - V'(z)) \tag{3.13}$$

and is a polynomial of the degree $2p - 2$ if $V(z)$ is a polynomial of the degree $2p$. It is easy to see that for $V(z) = z^2/4w^2$ (3.12) coincides with (2.16). Thus (3.12) is an extension of (2.16) to the more general case of distribution (3.1), where $V(\lambda)$ is a polynomial of an even degree bigger than 2. An analogue of (2.33) has the form

$$(1 - \frac{2}{|\text{Im}z \cdot V'(z)|})\mathbf{E}(|\gamma_n^2(z)|) \leq \frac{1}{|V'(z)|}\mathbf{E}(\frac{1}{n} \text{ Tr } \gamma_n(z)q_n(z, M)) + \frac{1}{n^2|(\text{Im}z)^3 V'(z)|} \tag{3.14}$$

where

$$q_n(z, M) = Q_n(z, M) - \mathbf{E}(Q_n(z, M)).$$

In the Gaussian case the first term in the r.h.s. of (3.14) is absent.

In view of the inequality

$$|\mathbf{E}(\frac{1}{n} \text{ Tr } \gamma_n(z)q_n(z, M))| \leq \mathbf{E}(|\gamma_n(z)|^2)^{1/2}\mathbf{E}(\frac{1}{n} \text{ Tr } q_n(z, M)q_n^*(z, M))^{1/2}$$

where $\gamma_n(z)$ is defined in (2.15), it seems that the most natural way to obtain (3.6) is to prove the estimate

$$\mathbf{E}(\frac{1}{n} \text{ Tr } q_n(z, M)q_n^*(z, M)) \leq \frac{const}{n^2}$$

thereby reducing the estimation of the variance of $g_n(z)$ to that of $q_n(z, M)$. Unfortunately, we do not know the proof of the last estimate based on the differentiation formula (3.9) and similar to that in the second proof of Proposition 2.1. Thus we refer the reader to works [19],[20], where the bound (3.6) is proven for all $\text{Im}z| > 0$ by using a combination of the orthogonal polynomials and variational techniques. A simple proof of a weaker version of (3.6) with n instead n^2 in the r.h.s. will be given in [3] also by using the orthogonal polynomial technique. Any of these bounds allows us to replace $\mathbf{E}(g_n(z)^2)$ in (3.5) by $\mathbf{E}(g_n(z)^2) = \mathbf{E}(g_n(z))^2 \equiv f_n^2(z)$. Besides, by applying (3.9) to $\varphi(M) = M$, we obtain the equality

$$\mathbf{E}(\frac{1}{n} \text{ Tr } MV'(M)) = 1 \tag{3.15}$$

that allows us to prove that all coefficients of the polynomial $Q_n(z)$ are uniformly bounded in n . After that simple compactness arguments yield that the limit of any convergent subsequence $f_{n_j}(z)$ satisfies (3.7).

Having Proposition 3.1 we can prove Theorem 3.1 by the following arguments. By solving the quadratic equation (3.7) we find that the measure N has the bounded Hölder density ρ, that for a convex V the support of the measure N corresponding to f is a finite interval (a, b) and that (see [3])

$$v.p. \int_a^b \frac{\rho(\mu) d\mu}{\mu - \lambda} = -\frac{V'(\lambda)}{2}, \quad \lambda \in (a, b) \tag{3.16}$$

where the symbol $v.p. \int$ denotes the singular Cauchy integral. Regarding this relation as a singular integral equation for $\rho(\lambda)$ and using standard facts of the theory of singular integral equations [21] we find that the bounded solution of the equation has the form

$$\rho(\lambda) = \frac{1}{\pi^2} \sqrt{R(\lambda)} \int_a^b \frac{V'(\mu) - V'(\mu)}{\mu - \lambda} \frac{d\mu}{\sqrt{R(\mu)}}, \quad R(\lambda) = (b - \lambda)(\lambda - a), \tag{3.17}$$

provided that

$$\int_a^b \frac{V'(\mu) d\mu}{\sqrt{R(\mu)}} = 0. \tag{3.18}$$

This gives condition (3.4) for $q = 0$. Besides we have the normalization condition

$$\int_a^b \rho(\mu) d\mu = 1 \tag{3.19}$$

that can be rewritten in the form (3.4) for $q = 1$ by using (3.17). It is clear that the integral is positive for a convex V and that it is a polynomial of degree $2p - 2$, if V is a polynomial of the degree $2p$. The unique solubility of system (3.7) can be proved by using the implicit function theorem [3].

Remark. Consider the case of the monomial $V(\lambda) = |\lambda|^{2p}/2p$. In this case above formulas can be written in the form

$$\rho(\lambda) = \frac{1}{2\pi I_{2p-1}} \int_{|\lambda|}^a \frac{t^{2p-1} dt}{\sqrt{t^2 - \lambda^2}}, \quad a^{2p} = \frac{\pi}{I_{2p}} \tag{3.20}$$

where

$$I_\alpha = \int_0^1 \frac{t^\alpha dt}{\sqrt{1 - t^2}}$$

These formulas are also valid for non-integer p, i.e. for potentials of the form $V(\lambda) = |\lambda|^\alpha/\alpha$ provided that $\alpha \geq 2$. For this case the formulas were obtained in [22] by another method. They can also be obtained by a version of the method presented in this Section. In this version we use the identity

$$G(z)V'(M) = G(z)V'(\lambda) + G(z)(V'(M) - V'(\lambda)), \quad z = \lambda + i\varepsilon \tag{3.21}$$

instead of (3.11). It can be shown that this allows us to obtain the final formulas (3.2)-(3.4) for example for non-polynomial (and even non-analytic) V's provided that they are convex, even, grow faster than logarithmically at infinity and are of the class C^2 on any finite interval $(-L, L)$.

4. Law of Addition of Random Matrices

Consider Hermitian matrices of the form

$$V_n A_n V_n^* + U_n B_n U_n^* \tag{4.1}$$

where V_n and U_n are random independent unitary matrices distributed both according to the normalized Haar measure on $U(n)$, A_n and B_n are Hermitian matrices such that their normalized eigenvalue counting measures N_{A_n} and N_{B_n} converge weakly to the limits N_A and N_B respectively and satisfy the condition

$$\sup_n \int |\lambda| N_{A_n,B_n}(d\lambda) < \infty. \tag{4.2}$$

Theorem 4.1. *The normalized eigenvalue counting measure of the ensemble of random matrices defined above tends in probability as $n \to \infty$ to the nonrandom limit whose Stieltjes transform $f(z)$ is a unique solution of the system of functional equations*

$$\begin{aligned} f(z) &= f_A(z - \Delta_B(z) f^{-1}(z)) \\ f(z) &= f_B(z - \Delta_A(z) f^{-1}(z)) \\ z f(z) &= \Delta_A(z) + \Delta_B(z) - 1 \end{aligned} \tag{4.3}$$

where f belongs to the class (2.8) and Δ_A and Δ_B are analytic for non-real z and such that

$$\sup_{y \ge 1} y|\Delta_A(iy)| < \infty, \quad \sup_{y \ge 1} y|\Delta_B(iy)| < \infty \tag{4.4}$$

The theorem was proved in [25] for the case of uniformly bounded in n matrices A_n and B_n by computing asymptotic form of moments of the sum via the moments of summands. This requires rather involved combinatorial analysis and impose the boundedness condition on matrices.

We outline now the proof [23] based on the same ideas as above, i.e. on the resolvent identity and on a certain differentiation formula. The formula used in this case is

$$\mathbf{E}(\varphi'(UBU^*)[UBU^*, C]) = 0 \tag{4.5}$$

where $\varphi : \mathbf{R} \to \mathbf{C}$ is a C^1 function whose derivative is polynomially bounded on the real line, B and C are Hermitian matrices, $[B, C] = BC - CB$ and the symbol $\mathbf{E}(...)$ denotes the integration over $U(n)$ with respect to the Haar measure normalized to 1. The formula can be easily derived from the shift invariance of the Haar measure.

Assume first that the norms of the matrices A_n and B_n are bounded uniformly in n. By applying (4.5) to the resolvent identity (2.17) relating the resolvent G of matrix (4.8) and the resolvent G_1 of matrix A_n we obtain the matrix identity

$$\mathbf{E}(G\frac{1}{n} \text{ Tr } G) = G_1 \mathbf{E}(\frac{1}{n} \text{ Tr } G) - G_1 \mathbf{E}(G\frac{1}{n} \text{ Tr } GUBU^*).$$

Assuming now asymptotic vanishing of the fluctuations of normalized traces, that we had in all cases above, we can rewrite this matrix identity in the form

$$\mathbf{E}(G) = G_1\left(z - \frac{\Delta_{B_n}(z)}{f_n(z)}\right) + o(1), \quad n \to \infty \tag{4.6}$$

where

$$\Delta_{B_n} = \mathbf{E}(\frac{1}{n}\ \mathrm{Tr}\ U_n B_n U_n^* G), \quad f_n = \mathbf{E}(\frac{1}{n}\ \mathrm{Tr}\ G). \tag{4.7}$$

and $|\mathrm{Im}z|$ is large enough to guarantee inversibility of the argument of G_1 uniformly in n. Applying to (4.6) the operation n^{-1} Tr we obtain the prelimit form of the first equation of system (4.3). The second equation follows from the analogous procedure in which the roles of A_n and B_n are interchanged (recall that U_n and U_n^* have the same distribution). The third equation is the limiting form of the identity n^{-1} Tr $G(z - A_n - U_n B_n U_n^*) = 1$ and of (4.7). The unique solubility of (4.3) follows from the implicit function theorem applicable for large $|\mathrm{Im}z|$. The proof of the vanishing of the correlations, more precisely, a bound similar to (2.14), is based on the same idea (see [23]).

To obtain the general case (4.2) we truncate eigenvalues of A_n and B_n by a large number T and use the minimax principle to control this procedure as $T \to \infty$, the compactness arguments and the unique solubility of (4.3) in the class (2.8), (4.4).

Remarks. (i). Since the normalized eigenvalue counting measure is unitary invariant and the Haar measure is shift invariant we can restrict ourselves without loss of generality to matrices of the simpler form

$$H_n = A_n + U_n B_n U_n^* \tag{4.8}$$

This form have, for example, the matrices (2.35) of the deformed GUE. Indeed, any matrix belonging to the GUE can be written in the form $\Psi_n \Lambda_n \Psi_n^*$, where Ψ_n is the matrix of its eigenvectors, distributed uniformly over the $U(n)$ according to the Haar measure, Λ_n is the random diagonal matrix of eigenvalues and Ψ_n and Λ_n are independent [1]. Besides, according to Theorem 2.1 the normalized eigenvalue counting measure of Λ_n converges with probability 1 to the semicircle law (2.5). Thus Ψ_n plays the role of U_n, Λ_n plays the role of B_n and N_B is given by (2.5). It can be easily checked that in this case the system (4.3) reduces to (2.36). Analogous fact is also valid for the deformed Laguerre ensemble (2.41) and also for certain classes of random operators acting in $l^2(\mathbf{Z}^d)$ [26].

(ii). It can be shown [23] that the theorem is also valid in the case when matrices A_n and B_n in (4.1) are also random, but independent of U_n and V_n and N_{A_n} and N_{B_n} converge weakly in probability to the nonrandom N_A and N_B. Then the ensemble of deformed covariance matrices (2.36) in which the random vectors a_l are uniformly distributed over the unit sphere in $\mathbf{C}^n$ also has form (4.8). As for the form (4.1), it is the case for the sum of two independent matrices distributed each according to the law (3.1) with possibly different polynomials $V_{1,2}$. In this case the condition (4.2) follows from (3.15). This case was considered in [24] by using formal perturbation theory around the Gaussian ensemble. Thus, we see that Theorem 4.1 describes in a rather general setting the result of deformation of a random matrix by another matrix randomly rotated with respect to the first and allows us to find the limiting eigenvalue counting measure of the the sum (of the deformed or of the perturbed ensemble) provided that we know these measure for the both terms of the sum.

(iii). For any function $f(z)$ satisfying (2.8) one can introduce the "selfenergy" $\Sigma(z)$ by the relation

$$f(z) = -(z + \Sigma(z))^{-1}. \tag{4.9}$$

It can be shown that $\Sigma(z)$ is also analytic for non-real z, and has the same property $\mathrm{Im}\Sigma(z)\mathrm{Im}z > 0, \quad \mathrm{Im}z \neq 0$ as f (see (2.8)). Denote by $z(f)$ the functional inverse of $f(z)$ and set $\Sigma(z) = R(f)$. Then it is easy to see that (4.3) is equivalent to the relation

$$R(f) = R_A(f) + R_B(f) \tag{4.10}$$

where $R_A(f)$ and $R_B(f)$ are the selfenergies corresponding to $f_A(z)$ and $f_B(z)$. The inverses of the Stieltjes transforms of limiting eigenvalue counting measures were used in [14] in the qualitative study of the support of of these measures in the case (5.16) below where

$$R(f) = -c\int \frac{t\sigma(dt)}{1+tf}. \tag{4.11}$$

Relation (4.10) was noted in [28] for the case when N_A and N_B are both the semicircle laws (2.5), when $R(f) = w^2 f$. The general form of this relation was proposed by D.Voiculescu in the context of the operator-algebras theory and its new branch known as the free probability theory (see [2] for results and references). In this theory the semicircle law plays the role of the Gaussian distribution and the measure defined by (2.40) (more generally, by formula (5.16) below) plays the role of the Poisson distribution.

(iv). Similar technique can be applied to multiplicative families of positive defined Hermitian and or unitary matrices and gives results [27] that generalize and simplify those of [14] and also gives a more direct proof of certain results obtained for these ensembles in the context of free probability [2].

5. Matrices With Independent and Weakly Dependent Entries. Tiny Perturbations

5.1. **Wigner Ensemble.** The proofs outlined in previous Sections for simplest archetypal ensembles, the GUE first of all, can be elaborated and used in rather general case of Hermitian, real symmetric or self-dual Hermitian random matrices whose entries are independent or weakly dependent modulo symmetry conditions. The matrices can be written in the form (cf.(2.3))

$$M_n = n^{-1/2} W_n \tag{5.1}$$

where matrix elements $W_{jk}^{(n)}$ of the matrix W_n still satisfy (2.4) but their probability laws $P_{jk}^{(n)}(dW)$ are not necessary Gaussian and may be n-dependent. We call these ensembles the Wigner Ensembles. In this general case we have to use instead of differentiation formula (2.18) the formula [31]

$$\mathbf{E}(\xi\varphi(\xi)) = \sum_{a=1}^{s} \frac{\kappa_{a+1}}{a!}\mathbf{E}(\varphi^{(a)}(\xi)) + \varepsilon_s \tag{5.2}$$

where κ_a are semi-invariants (cumulants) of a real-valued random variable ξ, $\varphi : \mathbf{R} \to \mathbf{C}$ is a function of the class C^{s+1} and $|\varepsilon_s| \leq C_s \sup_x |\varphi^{(s+1)}(x)|\mathbf{E}(|\xi|^{s+1})$.

Another version of the method is based on the perturbation expansion of matrix elements of the resolvent in a particular matrix element of the matrix M_n. Indeed, according to Section 2 an important moment of the method is the asymptotical computation of the expectation

$$\mathbf{E}(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} G_{jk}W_{kl}^{(n)}) \tag{5.3}$$

basing on various differentiation formulas (see formulas (2.18),(2.45),(3.9), and(5.2) above). However, in the case of independent entries satisfying (2.4), this requires the knowledge of dependence of G_{jk} on $W_{kl}^{(n)}$ up to linear terms only. Indeed, writing the resolvent identity (2.17) for $B = n^{-1/2}W_n$ and $A = n^{-1/2}W|_{W_{kl}^{(n)}=0}$ we obtain

$$G_{jk} = G_{jk}^{kl} - n^{-1/2}(W_{kl}^{(n)}G_{jk}^{kl}G_{lk}^{kl} + \overline{W_{kl}^{(n)}}G_{jl}^{kl}G_{kk}^{kl}) + r_n \tag{5.4}$$

where

$$G^{kl} = G|_{W_{kl}^{(n)}=0}, \quad |r_n| \leq \frac{|W_{kl}^{(n)}|^2}{n|\mathrm{Im}z|^3}. \tag{5.5}$$

Substituting (5.4) in (5.3) and taking into account that G^{kl} is independent of $W_{kl}^{(n)}$ we can perform explicitly the expectation with respect to $W_{kl}^{(n)}$ and obtain the relation

$$\mathbf{E}(g_n(z)) = -1/z - \frac{w^2}{zn^2}\mathbf{E}(\sum_{k=1}^{n} G_{jl}^{kl}G_{kk}^{kl}) + O(\max_{0\leq j,k\leq n} \mathbf{E}(|W_{jk}^{(n)}|^3)/n^{1/2}|\mathrm{Im}z|^4). \tag{5.6}$$

Now we can use (5.4) in the opposite direction to replace matrix element of G^{kl} by those of G. Thus, if

$$\sup_n \max_{0\leq j,k\leq n} \mathbf{E}(|W_{jk}^{(n)}|^3) \leq w_3 < \infty, \tag{5.7}$$

we obtain the analogue of (2.29) in the case of independent entries satisfying (2.4) and (5.7) with the error of the order $n^{-1/2}$. Similar arguments allows one to prove an analogue of (2.14) with the r.h.s. of the order $n^{-1/2}$. This is sufficient for the proof of an analogue of Theorem 2.1 for the independent entries satisfying (2.4) and (5.7) and with convergence in probability instead of convergence with probability 1 (see also Remark (ii) after Theorem 2.3).

We list now several recent results obtained by combinations of approaches based on formulas (5.2) and (5.4) (for an account of previous results see [32]).

(i) **Semicircle Law.** The normalized eigenvalue counting measure converges weakly in probability to the semicircle law (2.5) if and only if in addition to (2.4) for any $\tau > 0$ matrix elements (5.1) of satisfy the condition

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{1\le j\le k\le n}^{n} \int_{|W|\ge \tau n^{1/2}} |z|^2 P_{jk}^{(n)}(dW) = 0, \tag{5.8}$$

reminiscent the well known Lindeberg condition of the validity of the central limit theorem. This fact is known since the seventies, see [28] for the sufficiency of somewhat stronger version of (5.8) and [29, 30] for the necessity and sufficiency of (5.8). However these results were obtained by rather complicated method. In [3] we give a simple proof based on the approach of this paper.

(ii) **1/n expansion** [31]. By using (5.2) and assuming that $\{W_{jk}^{(n)}\}$ are identically distributed (modulo symmetry conditions as usually) and have $s+1$ finite moments one can construct $1/n$ - expansion of moments (2.20) in powers of $1/n^l, l \le s$, with the error of the order $1/n^{s+1/2}$ provided that the complex spectral parameter z verifies (2.27). We give here the two results for the real symmetric matrices $M_{jk}^n = n^{-1/2}(1+\delta_{i,j})W_{jk}^n$, $\mathbf{E}(W_{jk}^{(n)}) = 0$, $\mathbf{E}(W_{jk}^2) = w^2$ and

$$\sup_n \mathbf{E}(|W_{jk}^{(n)}|^5) < \infty.$$

We have then the following asymptotic formulas:

$$m_1(z) = f_{sc}(z)\left\{1 + \frac{1}{n}\left[\frac{w^2 f_{sc}^2(z)}{(1-w^2 f_{sc}^2(z))^2} + \frac{\sigma f_{sc}^4(z)}{1-w^2 f_{sc}^2(z)}\right]\right\} + O(n^{-\frac{3}{2}}),$$

where $\sigma = \mathbf{E}(|W_{jk}^{(n)}|^4) - 3\mathbf{E}(|W_{jk}^{(n)}|^2)$ is known as the excess of random variable $W_{jk}^{(n)}$ and is assumed to be independent of n, and f_{sc} is defined in (2.28);

$$m_2(z_1,z_2) - m_1(z_1)m_1(z_2) = n^{-2}c(z_1,z_2) + O(n^{-\frac{5}{2}})$$

where

$$c(z_1,z_2) = \frac{2w^2}{(1-w^2 f_{sc}^2(z_1))(1-w^2 f_{sc}^2(z_2))}\left\{w^2\left[\frac{f_{sc}(z_1)-f_{sc}(z_2)}{z_1-z_2}\right]^2 + \sigma f_{sc}^3(z_1) f_{sc}^3(z_2)\right\}. \tag{5.9}$$

(iii) **Central Limit Theorem** [31]. Assume in addition to (2.4) that the forth moments of $W_{jk}^{(n)}$ exist and are independent of j, k and n. Then for z from the domain (2.27) the random function $g_n(z) - \mathbf{E}(g_n(z))$ converges in distribution to the Gaussian random function with zero mean and the covariance (5.9).

5.2. **Sample Covariance Matrices.** In this Subsection we consider an ensemble of random matrices, generalizing the Laguerre ensemble of Subsection 2.2 and its real symmetric version known as the Wishart Ensemble of the sample covariance matrices. [13]. Respective matrices have the form

$$M_{m,n} = \frac{1}{n} A_{m,n} T_m A_{m,n}^* \tag{5.10}$$

where $A_{m,n}$ are $n \times m$ random matrices whose entries $A_{jk}^{(m,n)}$ are i.i.d. complex random variables satisfying conditions

$$\mathbf{E}(A_{jk}^{(m,n)}) = \mathbf{E}((A_{jk}^{(m,n)})^2) = 0, \quad \mathbf{E}(|A_{jk}^{(m,n)}|^2) = 1, \tag{5.11}$$

$$\sup_{m,n} \max_{1 \le j \le n, 1 \le k \le m} \mathbf{E}(|A_{jk}^{(m,n)}|^4) \le a_4 < \infty, \tag{5.12}$$

and T_n is a diagonal matrix. We assume that

$$m \to \infty, \quad n \to \infty, \quad \frac{m}{n} \to c < \infty, \tag{5.13}$$

and that the normalized counting measure

$$\sigma_m(\Delta) = \frac{1}{n} \natural \{t_l \in \Delta\} \tag{5.14}$$

of eigenvalues $t_l, l = 1, ..., m$ of T_n has a weak limit

$$\sigma_m(\Delta) \to \sigma(\Delta), \quad m \to \infty. \tag{5.15}$$

In particular, $t_l, l = 1, ..., m$ may be i.i.d. random variables independent of $A_{m,n}$.

Theorem 5.1. *Under conditions listed above the eigenvalue counting measure of matrices (5.10) converges weakly in probability to the nonrandom measure whose Stieltjes transform is a unique solution of the functional equation*

$$f(z) = -\left(z - c \int \frac{t\sigma(dt)}{1 + tf(z)}\right)^{-1} \tag{5.16}$$

in the class (2.8).

We outline the scheme of the proof, based on the same idea as above, i.e. on the careful analysis of the result of infinitesimal as $n \to \infty$ changes of respective matrices.

Start again from the resolvent identity (2.17) written for the pair $B = M_n$, $A = 0$. Applying to the identity the operation $\mathbf{E}(n^{-1} \text{ Tr } ...)$ we obtain

$$\mathbf{E}(g_n(z)) = -\frac{1}{z} - \frac{1}{zn^2} \sum_{l=1}^{m} t_l \mathbf{E}(\sum_{j,k=1}^{n} A_{jl}^{(m,n)} G_{jk} \overline{A_{kl}^{(m,n)}}) \tag{5.17}$$

We can use now the scheme of proof of Theorem 2.1 using (5.2) instead (2.18). It is more convenient however to apply here a somewhat different scheme. It is analogous to that based on relations (5.3), (5.4), and (5.5) in the case of the Wigner Ensembles of the previous subsection, however applied not to individual matrix elements but to the columns $a_l = \{n^{-1/2} A_{jl}^{(m,n)}\}_{j=1}^{n}, l = 1, ..., m$ of the random matrix $A^{(m,n)}$. Treating the columns as vectors of $\mathbf{C}^n$ we can rewrite (5.17) as follows

$$\mathbf{E}(g_n(z)) = -\frac{1}{z} - \frac{1}{zn} \sum_{l=1}^{m} t_l \mathbf{E}((Ga_l, a_l)) \tag{5.18}$$

where $(.,.)$ is the scalar product in $\mathbf{C}^n$. Since vectors a_l are independent we perform first the asymptotic computation of the expectation with respect to a_l in the l-th term of the sum like we did in the previous subsection for $W_{jk}^{(n)}$. To this end we use the formula, giving in the explicit form the result of perturbation of the resolvent G_C of an arbitrary matrix C by the rank one matrix $L_a, a \in \mathbf{C}^n$ defined by its action on any vector $x \in \mathbf{C}^n$ as $L_a x = (x, a)a$:

$$(C + L_a - z)^{-1} = G_C - G_C L_a G_C (1 + (G_C a, a))^{-1}. \tag{5.19}$$

The formula can be easily derived from the general resolvent identity (2.17). By applying the formula to $C = M_n|_{a_l=0}$ we obtain that

$$(Ga_l, a_l) = -\frac{t_l(G_l a_l, a_l)}{1 + t_l(G_l a_l, a_l)}$$

where $G_l = G|_{a_l=0}$. This relation will play the role of (5.4). Indeed, assume first that for some finite T and a_2

$$\sup_n \max_{1 \le l \le m} |t_l| \le T, \quad \sup_n \max_{1 \le l \le m} ||a_l|| \le a_2 \tag{5.20}$$

Since G_l does not depend on a_l and since random vectors $\{a_l\}_{l=1}^m$ are mutually independent one can find from (5.11) and (5.12) that

$$\begin{aligned} \mathbf{E}_l((G_l a_l, a_l)) &= \frac{1}{n} \text{ Tr } G_l \equiv g_n^{(l)} \\ |\mathbf{E}_l((G_l a_l, a_l)) - \frac{1}{n} \text{ Tr } G_l|^2 &\le \frac{Ca_4}{n|\text{Im}z|^2} \end{aligned}$$

where the symbol $\mathbf{E}_l(...)$ denotes the operation of the expectation with respect the vector a_l only.

These relations allow us to present (5.18) in the form (cf. (5.4))

$$\mathbf{E}(g_n) = -\frac{1}{z} - \sum_{l=1}^m t_l \mathbf{E}\left(\frac{g_n^{(l)}}{1 + t_l g_n^{(l)}}\right) + r_n \tag{5.21}$$

where now

$$|r_n| \le \frac{C(1 + a_4)}{n|\text{Im}z|^2}, \quad |\text{Im}z| \ge y_0,$$

and y_0 depends on T and on a_2 of (5.20). Besides, applying to (5.19) the operation $\frac{1}{n}$ Tr ... we obtain that

$$g_n - g_n^{(l)} = -\frac{1}{n} \cdot \frac{t_l(G_l^2 a_l, a_l)}{1 + t_l(G_l a_l, a_l)}$$

and thus

$$|g_n(z) - g_n^{(l)}(z)| \le \frac{1}{n|\text{Im}z|}.$$

By using three last relations we can write instead (5.21) for $|\text{Im}z| \ge y_0$ and n$\to \infty$ (cf. (5.6))

$$\mathbf{E}(g_n(z)) = -\frac{1}{z} - \int \mathbf{E}\left(\frac{tg_n(z)}{1 + tg_n(z)}\right) \sigma_m(dt) + o(1) \tag{5.22}$$

where $\sigma_m(dt)$ is defined in (5.14). This is an analogue of (2.29). Similar arguments allow us to prove an analogue of (2.14). As a result we obtain (5.16) in the case of bounded t_l and a_l. General case can be obtained from the proven one by using the the analyticity of the Stieltjes transform up to the real axis, the truncation of t_l and a_l, the minimax principle to control the truncation procedure, the compactness arguments, and the unique solubility of (5.16) in the class (2.8). The latter results from the the implicit function theorem.

Remarks. (i). In the case when $m = n$ and $\sigma(dt)$ has only one atom at $t = 2a^2$ we obtain (2.42).

(ii). Similar arguments shows that the eigenvalue counting measure of deformed ensemble (5.10)

$$H_n = H_n^{(0)} + M_{m,n} \tag{5.23}$$

where $M_{m,n}$ is defined by (5.10) and $H_n^{(0)}$ has the limiting eigenvalue counting measure N_0 (like in (2.35) and in (2.41) also tends weakly in probability to the nonrandom limit whose Stieltjes transform is a unique solution of the functional equation

$$f(z) = f_0\left(z - c\int \frac{t\sigma(dt)}{1 + tf(z)}\right) \tag{5.24}$$

This functional equation was derived first in [14] by another and rather complicated method. The method was based on the study of the sequence of matrices $H_n^{(p)}, \quad p = 1, ..., m$ defined as

$$(H_n^{(p)})_{jk} = (H_n^{(0)})_{jk} + \sum_{l=1}^{p} t_l A_{lj}^{(m,n)} \overline{A_{lk}^{(m,n)}} \tag{5.25}$$

and "interpolating" between $H_n^{(0)}$ and $(H_n^{(n)}) = H_n$. Asymptotic computation of the differences $\text{Tr}\ (H_n^{(p+1)} - z)^{-1} - \text{Tr}\ (H_n^{(p)} - z)^{-1}$ based on formula (5.19) lead to the first order partial differential equation for

$$f(t, z) = \lim_{n\to\infty, p/n\to t} \frac{1}{n}\ \text{Tr}\ (H_n^{(p)} - z)^{-1}.$$

Solving the differential equation subject the conditions $f(0, z) = f_0(z), f(1, z) = f(z)$ one obtains (5.24).

The derivations of equation (5.24) given later in [29, 30, 6, 11] and as well as the proof outlined above are more simple and direct. On the other hand, the sequence $H_n^{(p)}$ in (5.25) can be regarded as a matrix version of the sum of independent random variables with varying upper limit used often in the study of limit theorems and processes with independent increments. Similar observation was used recently in [2] to construct free (non-commutative) analogues of these processes where, in particular, an analogous partial differential equation was obtained (called in [2] the complex Burgers equation).

Acknowledgment

I am thankful to A.Khorunzhy, B.Khoruzhenko, A.Stoyanovich and V.Vasilchuk with whom the results mentioned above were obtained. The final version of this paper was written when I was participating the semester "Random Matrices and their Applications" at the MSRI. I am grateful to Prof.D.Eisenbud for the hospitality at the Institute, and to organizers of the semester P.Bleher and A.Its for the invitation. Research at MSRI is supported in part by NSF grant DMS-9701755.

References

[1] M. Mehta, *Random matrices*, Academic Press, New York, 1991.

[2] D.V. Voiculescu, K.J. Dykema, A. Nika, *Free random variables*, CRM Monogr.Ser.**1**, AMS., Providence, 1992.

[3] B. Khoruzhenko, L. Pastur, A. Stojanovich, (in preparation).

[4] N. Akhiezer, I. Glazman, *Theory of linear operators in Hilbert space*, Frederick Ungar, New York, 1961.

[5] D. Ruelle, *Statistical mechanics*, Benjamin, New York, 1969.

[6] L. Pastur, A. Khorunzhy, *On the eigenvalue distribution of random matrices with independent entries*, Preprint N21 of the Institute of Low Temperature Physics, Kharkov, 1989.

[7] A. Khorunzhy, S. Molchanov, L. Pastur, *On the eigenvalue distribution of band random matrices*, Teor. Math. Phys., **90**, 108–118 (1992).

[8] A. Khorunzhy, B. Khoruzhenko, L. Pastur, M.Shcherbina, *Large-n limit in the statistical mechanics and the spectral theory of disordered systems*, In: Phase Transitions and Critical Phenomena,**15** C.Domb, J.Lebowitz (Eds.), Academic Press, New York, 1992.

[9] A. Khorunzhy, L. Pastur, *Limits of infinite interaction radius, dimensionality and the number of components for random operators with off-diagonal randomness*, Commun. Math. Phys., **153**, 605–646, (1993).

[10] B. Khoruzhenko, W. Kirsch, L. Pastur, *The interband light absorption coefficient in the weak disorder regime: an asymptotically exactly solvable model* , J.Phys.A: Math. Gen., **27**, 2527–2543 (1994).

[11] A. Khorunzhy, *Eigenvalue distribution of large random matrices with correlated entries*, Mathem.Phys., Analysis and Geometry, **3**, 80-101 (1996).

[12] J.-L. Pichard, *Random transfer matrix theory and conductance fluctuations* , In: Quantum Coherence in Mesoscopic Systems, B.Kramer (Ed.), Plenum, New York, 1991.

[13] R. Muirhead, *Aspects of multivariate statistical theory*, Wiley, New York, 1982.

[14] V. Marchenko, L. Pastur *The eigenvalue distribution in some ensembles of random matrices* Math.USSR Sbornik, **1**, 457-483 (1967).

[15] D. Bessis, C. Itzykson, J. Zuber, *Quantum field theory techniques in graphical enumeration*, Adv.Appl.Math., **1**, 109–157 (1980).

[16] F. Xu *A random matrix model from two-dimensiona Yang-Mills theory*, Commun.Math.Phys., **190**, 287-307 (1997).

[17] D. Voiculescu (Ed.) *Free probability theory*, AMS, Providence, 1997.

[18] P. DiFrancesco, P.Ginsparg, J.Zinn-Justin, *2D gravity and random matrices*, Phys.Reports, **254**, 1-133 (1995).

[19] A. Boutet de Monvel, L.Pastur, M. Shcherbina, *On the statistical mechanics approach to the random matrix theory: the integrated density of states*, J.Stat.Phys., **79**, 585-611 (1995).

[20] L.Pastur, M. Shcherbina, *Universality of the local eigenvalue statistics for a class of unitary invariant ensembles*, J.Stat.Phys., **86**, 109-147 (1997).

[21] N.I. Muskhelishvili, *Singular integral equations*, Nordhoff, Groningen, 1953.

[22] L. Pastur, *On the universality of the level spacing distribution for some ensembles of random matrices*, Lett.in Math.Phys., **25**, 259-263 (1992).

[23] L. Pastur, V. Vasilchuk, *On the law of addition of random matrices*, (in preparation)
[24] A. Zee, *Law of addition in random matrix theory*, Nucl.Phys., **B474**, 726-744 (1996).
[25] R. Speicher, *Free convolution and the random sum of matrices*, Publ.RIMS, **29**, 731-744 (1993).
[26] P.Neu, R. Speicher, *Rigorous mean-field model for coherent-potential approximation: Anderson model with free variables*, J.Stat.Phys., **80**, 1279-1308 (1995).
[27] V. Vasilchuk, *On the law of multiplication of random matrices*, (in preparation)
[28] L. Pastur, *On the spectrum of random matrices*, Teor.Math.Phys., **10**, 67-74 (1972).
[29] V. Girko, *Theory of random matrices*, Kiev, Vyshcha Shkola, 1975 (in Russian).
[30] V. Girko, *Theory of random matrices* Kluwer Academic Publishers (1990).
[31] A. Khorunzhy, B. Khoruzhenko, L. Pastur, *Random matrices with independent entries: asymptotic properties of the Green function*, J.Math.Phys.,**37**, 5033-5060 (1996).
[32] L. Pastur, *Eigenvalue distribution of random matrices*, Ann.l'Inst.H.Poincare,**64**, 325-337(1996).

Department of Mathematics, University Paris-VII, 2, place Jussieu, 75251, Paris, France

E-mail address: `pastur@math.jussieu.fr`

Mathematical Division, Institute for Low Temperatures, 47, Lenin's Ave.,310164, Kharkov, Ukraine

E-mail address: `lpastur@ilt.kharkov.ua`