Some Elementary Results around the Wigner Semicircle Law

Anne Boutet de Monvel and Alexei Khorunzhy

Contents

Lecture 0. Introduction. Motivations and Generalities 3

Lecture 1. GOE and the Semicircle Law 7

Lecture 2. Proof of the Semicircle Law 15

Lecture 3. Smoothed Eigenvalue Density 23

Generalizations of Theorems 3.1-3.3 and universality conjecture 29

Lecture 4. Eigenvalues outside of the limiting spectrum 33

Bibliography 41

In these lecture notes we give an accessible introduction to the spectral theory of random matrices. We consider Gaussian Orthogonal Ensemble as the main subject to present and prove the semicircle (or Wigner) law. This is the fundamental statement in the spectral theory of large random matrices.

We deal in frameworks of the resolvent and moments approaches and give two proofs of the semicircle law. Then we formulate the theorems that can be regarded as generalizations and improvements of this statement. In particular, we show the relevance of these two techniques in the studies of local properties of the eigenvalue distribution inside and outside of the limiting spectra.

We try not to overload these notes with technical details; our main task is to make the reader familiar with key points of the reasonings. Therefore we do not present the complete proofs of the improvements of the Wigner semicircle law.

Lecture 0

Introduction. Motivations and Generalities

Random Matrices and their Use. Random matrices are in extensive use in various fields of theoretical physics (in particular, in models of disordered solid state and chaotic systems, statistical mechanics, quantum field theory). The mathematical contents of random matrices is rich and provides structures of fairly general type (for example, see the book-length review [**22**]). Graph theory, classical compact groups, orthogonal polynomials, integral equations, non-commutative probability theory, combinatorics are enriched due to the studies of random matrix properties. We refer the reader to recent papers and reviews, for example [**9, 20, 24, 29**], to get acquainted with references to recent results and various applications of random matrices.

What is important that under "random matrices" we mean here matrices whose entries are of the same order of magnitude. One of the examples is given by random matrices

$$A_N(x, y) = a(x, y), \quad x, y = 1, \dots, N \tag{0.1}$$

with independent identically distributed random variables $a(x, y)$.

The family of random matrices is vast and incorporates different ensembles whose probability distribution is chosen according to the model to be described.

To get some examples, one can can consider such representatives of (0.1) as the set $\mathcal{U}_N$ of all unitary matrices U_N. The compact group $\mathcal{U}_N$ supplied with the Haar measure becomes the probability space.

Another example of extensively studied family of ensembles is given by random Hermitian $N \times N$ matrices M_N whose probability distribution P_N is invariant under unitary transformations of $\mathbb{C}^N$. In particular, one can consider P_N with the density

$$p(M_N) = Z_N^{-1} \exp\{-N \operatorname{Tr} V(M_N)\}, \tag{0.2}$$

where Z_N is the normalization constant and $V(t)$ is a function from a suitable class.

In the last two examples the matrices U_N and M_N have entries that are strongly correlated between themselves. Nevertheless, for these classes there exist explicit form of the joint probability distribution of eigenvalues of these matrices (see, e.g., [**22**]). This allows one to get into deep details in their study.

Our goals. In present lectures we would like to present the tools that can be used when the explicit form of the eigenvalue distribution of matrices (0.1) is unknown. For example, this takes place when $\{a(x,y)\}$ are independent arbitrarily distributed random variables. This family of random matrices was the first under consideration (see [**30**]) and the *semicircle law* was first established for it in the limit $N \to \infty$. It concerns the limiting distribution of eigenvalues that is given, broadly speaking, by a rather massive (comparing with N) part of the whole collection of eigenvalues. This asymptotic regime is known as *the global* one.

More detailed properties of the spectrum (in other words, those that are determined by more local regimes than that given by the semicircle law) have not been studied in this case of arbitrarily distributed $\{a(x,y)\}$.

Our aim is to present here several results on the distribution of eigenvalues of A_N in the limit $N \to \infty$. We describe two main techniques to prove the semicircle law. They are based on the classical *resolvent* and the *moment approaches* of the spectral theory of operators. In the global asymptotic regime these two approaches are equivalent.

However, their use in local regimes is no more classical and require essential modifications. We develop necessary modifications and show that these two approach are complementary in the studies of the inner and outer parts of the limiting spectra (i.e., the support of the semicircle distribution), respectively.

Organization of these Lecture Notes. To give more clear account on our ideas, we consider the ensemble of *gaussian random matrices* as our main subject. Namely, we consider the case of real symmetric random matrices with independent entries that have joint Gaussian distribution; the ensemble we are based on is known as the *Gaussian Orthogonal Ensemble* of random matrices (GOE). Together with its Hermitian analogue abbreviated by GUE, these two ensembles represent the principal subject of random matrix theory. They are the most deeply studied and they are the easiest to verify one or another conjecture about the random matrix properties.

We prove the semicircle law by using the moment and the resolvent approaches in lectures 1 and 2, respectively. Then we turn to improvements of the semicircle law inside of the limiting spectrum and outside of it. Lectures 3 and 4 are devoted to these considerations.

We give a brief account on results that are similar to the semicircle law but are derived for other ensembles of random matrices.

Although our main attention in this addition is devoted to random matrices with independent entries, we formulate analogues of several statements valid also in the case of Gaussian correlated random variables.

Acknowledgements. These lecture notes represent amended and enlarged version of the notes prepared for the MSRI Program "Random Matrix Models and Their Applications" (Spring-Summer 1999). The authors are grateful to A. Its, P. Bleher, and H. Widom and other organizers for their kind invitation to participate the workshops and for financial support.

Lecture 1

GOE and the Semicircle Law

In this lecture we prove the semicircle law for random symmetric matrices A_N whose entries are jointly independent (excepting the symmetry) real random variables. This statement concerns the eigenvalue distribution of the ensemble $\{A_N\}$ in the limit $N \to \infty$.

We consider real symmetric matrices in order to simplify computations. The same result is valid for Hermitian random matrices. Also for simplicity, we start with the case when the entries of A_N have joint Gaussian distribution.

Our aim is to describe two general approaches of the proof in the shortest and simplest way that makes the ideas clear. That is why we are related in this lecture only with the Gaussian ensemble $\{A_N\}$. Generalizations of $\{A_N\}$ and their properties will be considered further on.

Definition of GOE. Thus, we consider an $N \times N$ matrix with entries

$$\mathscr{A}_N(x,y) = a(x,y), \quad 1 \le x \le y \le N$$

that is real and symmetric

$$\mathscr{A}_N(x,y) = \mathscr{A}_N(y,x).$$

We assume that any collection $\{a(x,y)\}_{1\le x\le y\le N}$ is a family of random variables whose joint distribution is the one of Gaussian independent random variables.

We also assume $a(x,y)$, $x<y$ to be identically distributed. The same concerns random variables $a(x,x)$. More precisely, we write that

$$\mathsf{E}\,a(x,y)=0,\quad \mathsf{E}\,a(x,y)^2=\begin{cases} v^2, & \text{if } x\neq y;\\ 2v^2, & \text{if } x=y,\end{cases} \tag{1.1}$$

where E denotes the mathematical expectation with respect to the measure generated by the family $\{a(x,y),\ x\le y\}_{x,y=1}^N$. In fact, one can determine all random variables $a(x,y)$, $x,y\in\mathbb{N}$ on the same probability space. In this case E also denotes corresponding mathematical expectation.

One can rewrite the last condition of (1.1) in the form

$$\mathsf{E}\,a(x,y)a(s,t)=v^2(\delta_{xs}\delta_{yt}+\delta_{xt}\delta_{ys}), \tag{1.2}$$

where δ is the Kronecker δ-symbol:

$$\delta_{xy}=\begin{cases}1, & \text{if } x=y,\\ 0, & \text{if } x\neq y.\end{cases}$$

Given (1.2), it is convenient to write the distribution of $\{\mathcal{A}_N\}$ in a compact form:

$$P(\mathcal{A}_N)=\frac{1}{Z_N}\exp\Big\{-\frac{1}{4v^2}\operatorname{Tr}\mathcal{A}_N^2\Big\}, \tag{1.3}$$

where Z_N is the normalization constant:

$$Z_N=\int\exp\Big\{-\frac{1}{4v^2}\operatorname{Tr}\mathcal{A}_N^2\Big\}\prod_{1\le x\le y\le N}\mathrm{d}\,a(x,y).$$

Indeed, one can easily observe that

$$\begin{aligned}P(\mathcal{A}_N)&=\frac{1}{Z_N}\prod_{1\le x<y\le N}\exp\Big\{-\frac{1}{2v^2}a(x,y)^2\Big\}\prod_{1\le x\le N}\exp\Big\{-\frac{1}{4v^2}a(x,x)^2\Big\}\\ &=\frac{1}{Z_N}\exp\Big\{-\frac{1}{4v^2}\Big(2\sum_{1\le x<y\le N}a(x,y)^2+\sum_{1\le x=y\le N}a(x,y)^2\Big)\Big\}\end{aligned}$$

$(a(x,y)=a(y,x))$

$$\begin{aligned}&=\frac{1}{Z_N}\exp\Big\{-\frac{1}{4v^2}\sum_{x,y=1}^N a(x,y)^2\Big\}\\ &=\frac{1}{Z_N}\exp\Big\{-\frac{1}{4v^2}\operatorname{Tr}\mathcal{A}_N^2\Big\}\end{aligned}$$

which is (1.3).

Definition (1.3) shows that the distribution $P(\mathcal{A}_N)$ is invariant under the orthogonal transformations of $\mathbb{R}^N$. Therefore the ensemble described is known as the Gaussian Orthogonal Ensemble (GOE) of random matrices.

See the book [**22**] for the history and basic results on eigenvalue distribution of this and other ensembles of random matrices. The introductory article [**31**] is recommended for those who are interested in combinatorics of matrix integrals.

Eigenvalue distribution function. The eigenvalue distribution of real symmetric (or complex hermitian) $N \times N$ matrix H_N is described by the function

$$\sigma(\lambda; H_N) = \frac{1}{N}\, \#\{\lambda_j^{(N)} \le \lambda\}, \tag{1.4}$$

where

$$\lambda_1^{(N)} \le \cdots \le \lambda_N^{(N)}$$

are the eigenvalues of H_N. This function is called the *normalized eigenvalue counting function* (NCF) of the matrix H_N. It is clearly a step-like function increasing from 0 to 1.

In mathematical literature, one can meet also the term *empirical eigenvalue distribution function*. This term seems somewhat misleading (because we consider σ determined for a matrix A_N but not for the sum over N samples). In our notes we keep the term NCF common for the spectral theory.

Given a random matrix $\mathcal{A}_N$, the corresponding function $\sigma(\lambda; \mathcal{A}_N)$ is random. The semicircle law first proved by Wigner [**30**] states that the NCF of the matrix

$$A_N(x, y) = \frac{1}{\sqrt{N}} \mathcal{A}_N(x, y)$$

weakly converges in average as $N \to \infty$ to a nonrandom distribution:

$$\lim_{N\to\infty} \sigma(\lambda; A_N) = \sigma_{\mathrm{W}}(\lambda), \tag{1.5a}$$

whose density is given by

$$\sigma'_{\mathrm{W}}(\lambda) \equiv \rho_{\mathrm{W}}(\lambda) = \begin{cases} \sqrt{4v^2 - \lambda^2}, & \text{if } |\lambda| \le 2v, \\ 0, & \text{if } |\lambda| > 2v. \end{cases} \tag{1.5b}$$

Weak convergence in average means here that for any non-random function $\phi(\lambda) \in C_0^\infty(\mathbb{R})$,

$$\mathsf{E} \lim_{N\to\infty} \int_{\mathbb{R}} \phi(\lambda)\, \mathrm{d}\, \sigma(\lambda; A_N) = \int_{\mathbb{R}} \phi(\lambda)\, \mathrm{d}\, \sigma_{\mathrm{W}}(\lambda).$$

As it was mentioned above, we will prove this statement twice by two different approaches.

❶ **Moment relations approach.** We describe first the method used by Wigner in the proof of the semicircle law (see [**30**] and [**7, 26**] for the source and for improvements of the method, respectively). Here one is interested in the asymptotic behavior of the moments

$$M_j^{(N)} = \mathsf{E} \int_{\mathbb{R}} \lambda^j \, \mathrm{d}\,\sigma_N(\lambda)$$

of the measure

$$\mathrm{d}\,\sigma_N(\lambda) \equiv \mathrm{d}\,\sigma(\lambda; A_N).$$

Let us note that due to the definition (1.4) of the NCF, we simply have that

$$M_j^{(N)} = \mathsf{E}\, \frac{1}{N} \operatorname{Tr} A_N^j \equiv \mathsf{E} \left\langle A_N^j \right\rangle,$$

where we denote the normalized trace of a matrix A_N by angle brackets:

$$\langle A_N \rangle \equiv \frac{1}{N} \operatorname{Tr} A_N.$$

Basing on computations that are somewhat different from the original technique by Wigner, we will derive the relations

$$\lim_{N\to\infty} M_j^{(N)} = m_j = \begin{cases} t_k v^{2k}, & \text{if } j = 2k, \\ 0, & \text{if } j = 2k+1, \end{cases} \tag{1.6a}$$

where t_k, $k \in \mathbb{N}$ are given by the recurrence relations

$$\begin{aligned} t_0 &= 1 \\ t_k &= \sum_{j=0}^{k-1} t_{k-1-j} t_j. \end{aligned} \tag{1.6b}$$

Then we will show that (1.6) is equivalent to (1.5).

❷ **Resolvent approach.** Another method to study the limiting NCF is related to the resolvent $G_N(z) = (A_N - z)^{-1}$. It is not hard to see that its normalized trace is simply the Stieltjes transform of $\sigma_N(\lambda)$:

$$g_N(z) \equiv \frac{1}{N} \operatorname{Tr} G_N(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\,\sigma_N(\lambda)}{\lambda - z}.$$

In these terms, convergence (1.5) means that for all $z \in \mathbb{C}_\pm = \mathbb{C} \setminus \mathbb{R}$,

$$\lim_{N\to\infty} \mathsf{E}\, g_N(z) = f_{\mathrm{W}}(z), \tag{1.7a}$$

where $f_{\mathrm{W}}(z)$ is the Stieltjes transform of $\sigma_{\mathrm{W}}(\lambda)$;

$$f_{\mathrm{W}}(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\,\sigma_{\mathrm{W}}(\lambda)}{\lambda - z}. \tag{1.7b}$$

We will derive shortly that $f_{\mathrm{W}}(z)$ satisfies the equation

$$f_{\mathrm{W}}(z) = \frac{1}{-z - v^2 f_{\mathrm{W}}(z)}. \tag{1.8}$$

(1.8) is equivalent to (1.6b). This can be easily derived from the definition (1.7b) of $f_{\mathrm{W}}(z)$ as Stieltjes transform that implies the relation

$$f_{\mathrm{W}}(z) = \frac{1}{-z} \sum_{k=0}^{\infty} \frac{t_k v^{2k}}{z^k}.$$

The proposition to consider $g_N(z)$, that is the generating function of the moments $M_j^{(N)}$, instead of the moments by themselves is due to V.Marchenko and L.Pastur who derived (1.8) as a by-product of their more general results [**21**]. The resolvent approach in random matrix theory was further developed by A.Khorunzhy and L.Pastur (see for example [**17**] and [**18**]).

The crucial step here is to consider the moments

$$P_l^{(N)} = \mathsf{E}[g_N(z)]^l, \quad l \geq 1$$

that are shown to satisfy an infinite system of relations resembling the system of equations for correlation functions of statistical mechanics. The idea to write such equations for random matrices dates back to F.Berezin [**1**]. Broadly speaking, Berezin showed that the moments $P_l^{(N)}$ factorize to the powers of $f_{\mathrm{W}}(z)$. This fact leads to statements like (1.7).

Recently it was shown that for convergence (1.7) it is sufficient to consider the two first relations from this infinite hierarchy. This leads to a rather short proof of the semicircle law.

Derivation of the moment relations. We start with the moment approach to derive relations (1.6b). Gaussian random variables are rather convenient to deal with. One of the reasons is that if one has a centered Gaussian random variable γ, then

$$\mathsf{E}\,\gamma\phi(\gamma) = \mathsf{E}\,\gamma^2\,\mathsf{E}\,\phi'(\gamma) \tag{1.9}$$

for all non-random functions ϕ such that the integrals in (1.9) exist. In the more general case of a vector $\vec{\gamma} = (\gamma_1, \dots, \gamma_m)$ of Gaussian random variables with zero average one has

$$\mathsf{E}\,\gamma_j\phi(\vec{\gamma}) = \sum_{l=1}^{m} \mathsf{E}\,\gamma_j\gamma_l\,\mathsf{E}\,\frac{\partial\phi(\vec{\gamma})}{\partial\gamma_l}\ . \tag{1.10}$$

As the simplest application, one can easily derive from (1.9) that

$$\mathsf{E}\,\gamma^{2k} = \mathsf{E}\,\gamma \times \gamma^{2k-1} = v_2^k(2k-1)!!,$$

where $v_2 = \mathsf{E}\,\gamma^2$.

Let us now consider the moments

$$M_{2k}^{(N)} = N^{-1} \sum_{x,y=1}^{N} \mathsf{E}\, A_N^{2k-1}(x,y) A_N(y,x). \tag{1.11}$$

Using (1.10), we can write that

$$\begin{aligned} &\mathsf{E} A_N^{2k-1}(x,y) A_N(y,x) \\ &= \sum_{s,t=1}^{N} \mathsf{E}\, A_N(y,x) A_N(s,t)\ \mathsf{E}\, \frac{\partial A_N^{2k-1}(x,y)}{\partial A_N(s,t)}\ . \end{aligned} \tag{1.12}$$

It is obvious that

$$\frac{\partial A_N^{2k-1}(x,y)}{\partial A_N(s,t)} = \sum_{l=0}^{2k-2} A_N^{2k-2-l}(x,s) A_N^l(t,y).$$

Using (1.2) and substituting these relations into the right-hand side of (1.12), we obtain that

$$\begin{aligned} &\mathsf{E}\, A_N^{2k-1}(x,y) A_N(y,x) \\ &= \frac{v^2}{N} \sum_{l=0}^{2k-2} \mathsf{E} \left[A_N^{2k-2-l}(x,x) A_N^l(y,y) + A_N^{2k-2-l}(x,y) A_N^l(x,y) \right]. \end{aligned}$$

Regarding the sum over y in (1.11) and taking into account the symmetry condition $A_N(x,y) = A_N(y,x)$, we can write that

$$\sum_{y=1}^{N} \sum_{l=0}^{2k-2} A_N^{2k-2-l}(x,y) A_N^l(x,y) = (2k-1) A_N^{2k-2}(x,x).$$

Thus, we derive relation

$$M_{2k}^{(N)} = v^2 \sum_{l=0}^{2k-2} \mathsf{E} \langle A_N^{2k-2-l} \rangle \, \langle A_N^l \rangle + v^2 \frac{2k-1}{N} M_{2k-2}^{(N)}.$$

One can rewrite this equality in the form

$$M_{2k}^{(N)} = v^2 \sum_{l=0}^{2k-2} M_{2k-2-l}^{(N)} M_l^{(N)} + v^2 B_{2k-2}^{(N)} + v^2 \frac{2k-1}{N} M_{2k-2}^{(N)}, \tag{1.13}$$

where

$$B_{2k-2}^{(N)} = \sum_{l=0}^{2k-2} \left[\mathsf{E} \langle A_N^{2k-2-l} \rangle \, \langle A_N^l \rangle - \mathsf{E} \langle A_N^{2k-2-l} \rangle \, \mathsf{E} \langle A_N^l \rangle \right]. \tag{1.14}$$

Now, if one assumes that

$$B_{2k-2}^{(N)} = o(M_{2k-2}^{(N)}) \text{ as } N \to \infty, \tag{1.15}$$

and accept that

$$M_{2k+1}^{(N)} = 0, \tag{1.16}$$

then one can easily derive (1.6) from (1.13).

(1.16) follows immediately from the observation that for any particular values of the variables $\{y_i\}$

$$\mathsf{E}\, a(x, y_1) \cdots a(y_{2k}, x) = 0$$

as the average of the product of an odd number of gaussian random variables.

Relation (1.15) reflects the property of selfaverageness of the moments $M_j^{(N)}$. In Lecture 4 we will show that the much more powerful estimate

$$B_{2k-2}^{(N)} \leq \frac{(2k+2)^{2\alpha}}{N^2} M_{2k-2}^{(N)}, \quad \alpha > 1 \tag{1.17}$$

holds for all $k \ll N^{2/3}$ as $N \to \infty$. This will lead to estimates of the norm of A_N and other important consequences.

Lecture 2

Proof of the Semicircle Law

Now let us turn to the proof of (1.7) for GOE. We follow the scheme developed in [**17, 18**] and modified in [**11**].

We will use twice the resolvent identity

$$G - G' = -G(H - H')G', \qquad G = (H - z)^{-1}, \quad G' = (H' - z)^{-1}, \tag{2.1}$$

that is true for hermitian matrices H and H' of the same dimension and $z \in \mathbb{C}_\pm$.

Regarding (2.1) with $H = A_N$ and $H' = 0$, we obtain the relation

$$G_N(x, x') = \zeta\delta_{x,x'} - \zeta \sum_{y=1}^{N} G_N(x, y)A_N(y, x'),$$

where $G_N = (A_N - z)^{-1}$, $\zeta \equiv (-z)^{-1}$ and $\delta_{x,y}$ is the Kronecker δ-symbol. We are interested in the average value of the normalized trace $g_N(z) = N^{-1} \operatorname{Tr} G_N$. It is clear that

$$\mathsf{E}\, g_N(z) = \zeta - \zeta \frac{1}{N} \sum_{x,y} \mathsf{E}\, G_N(x, y)A_N(y, x). \tag{2.2}$$

Now we can apply (1.10) to the last average from (2.2) and obtain the relation

$$\mathsf{E}\, G_N(x, y)A_N(y, x) = \sum_{s,t=1}^{N} \mathsf{E}\, A_N(y, x)A_N(s, t)\, \mathsf{E}\, \frac{\partial G_N(x, y)}{\partial A_N(s, t)}.$$

One can easily deduce from (2.1) that

$$\frac{\partial G_N(x,y)}{\partial A_N(s,t)} = -G_N(x,s)G_N(t,y). \tag{2.3}$$

Indeed, it is sufficient to consider (2.1) with $H' = A_N$ and

$$H(x',y') = A_N(x',y') + \Delta\delta_{x',s}\delta_{y',t}$$

and to find the ratio $(G - G')/\Delta$ in the limit $\Delta \to 0$.

Remembering definition (1.2) and using (2.3), we obtain that

$$\mathsf{E}\, g_N(z) = \zeta + \zeta\frac{v^2}{N^2}\sum_{x,y}\mathsf{E}\big[G_N(x,x)G_N(y,y) + G_N(x,y)G_N(x,y)\big]. \tag{2.4}$$

One can rewrite this relation in the form

$$\mathsf{E}\, g_N(z) = \zeta + \zeta v^2\, \mathsf{E}[g_N(z)]^2 + \frac{v^2}{N}\,\mathsf{E}\,\Phi_N^{(1)}(z), \tag{2.5}$$

where $\Phi_N^{(1)}(z) = \langle G_N^2(z)\rangle \equiv N^{-1}\operatorname{Tr} G_N^2(z)$. We see that the first moment of $g_N(z)$ is expressed via the second moment of this variable added by the terms vanishing in the limit $N \to \infty$.

Indeed, elementary estimates

$$\langle G_N^2(z)\rangle \le \big\|G_N^2(z)\big\| \le \|G_N(z)\|^2 \le \frac{1}{|\operatorname{Im} z|^2} \tag{2.6}$$

show that $\left|\Phi_N^{(1)}(z)\right| = \mathrm{O}(1)$ as $N \to \infty$.

Having (2.5), we can proceed by two ways.

The first approach inspired by the work of Berezin [**1**] is to derive an infinite system of recurrence relations for the moments $L_k^{(N)} = \mathsf{E}[g_N(z)]^k$, $k \ge 2$. This method has been developed in [**17, 18**] and extensively used for various ensembles of random matrices and random operators (see, for example [**14, 15**]).

The second approach proposed in [**11**] represents a shortened version of the method of infinite system of relations. Loosely speaking, it uses only the two first relations and lead to a fairly short proof of the semicircle law. This shortened version has been widely applied in the studies of random matrix eigenvalue distribution [**5, 16, 19, 25**]. This approach seems to be unavoidable in the studies of smoothed eigenvalue density and its fluctuations [**2**]. However, certain passages can appear as somewhat tricky things in this shortened version. Thus, we start with the discussion of the infinite system method.

Infinite System Approach [17, 18]. One can derive a system of relations for the moments $L_k^{(N)} = \mathsf{E}[g_N(z)]^k$ subsequently applying to the last factor in L_k relations (2.2), (1.10), and (2.3). Then one obtains for $k \geq 2$

$$\mathsf{E}[g_N(z)]^k = \zeta\,\mathsf{E}[g_N(z)]^{k-1} + \zeta v^2\,\mathsf{E}[g_N(z)]^{k+1} + \frac{v^2}{N}\,\mathsf{E}\,\Phi_N^{(k)}(z) + \frac{v^2}{N}\,\mathsf{E}\,\tilde{\Phi}_N^{(k)}(z),$$

where

$$\Phi_N^{(k)}(z) = [g_N(z)]^{k-1}\langle G_N^2(z)\rangle,$$
$$\tilde{\Phi}_N^{(k)}(z) = \frac{k}{N}[g_N(z)]^{k-2}\langle G_N^3(z)\rangle.$$

Thus for the moments L_k we have the relations

$$L_k^{(N)} = \zeta\delta_{k,1} + (1-\delta_{k,1})\zeta L_{k-1}^{(N)} + \zeta v^2 L_{k+1}^{(N)} + \Psi_N^{(k)}(z), \tag{2.7}$$

where, according to estimates (2.6),

$$\left|\Psi_N^{(k)}(z)\right| \leq \frac{v^2}{\eta^k N}\left(1 + \frac{k}{N}\right). \tag{2.8}$$

It is not hard to see that (2.7) can be rewritten in a vector form

$$\vec{L}^{(N)} = \vec{l} + T_z\vec{L}^{(N)} + \vec{\Psi}^{(N)}, \tag{2.9}$$

where $\vec{l_k} = \delta_{k,1}\zeta$ and

$$[T_z\vec{e}]_k = (1-\delta_{k,1})\zeta e_{k-1} + v^2\zeta e_{k+1}.$$

Now it is not hard to show that

$$\|T_z\| \leq \eta + \frac{v^2}{\eta}.$$

Therefore for $\eta > 2v$ one has $\|T_Z\| < 1$. Introducing the equation

$$\vec{L}' = \vec{l} + T_z\vec{L}' \tag{2.10}$$

that obviously has one solution, one can easily deduce from (2.9) and (2.10) that

$$\left\|\vec{L}^{(N)} - \vec{L}'\right\| = \mathrm{O}(N^{-1}).$$

This proves convergence (1.7).

Short Proof of the Semicircle Law [10].

Proof. Denoting $\mathsf{E}\,g_N(z) \equiv f_N(z)$ and regarding that $\sum_y G_N(x,y)G_N(x,y) = G_N^2(x,y)$, we derive our first main relation

$$f_N(z) = \zeta + \zeta v^2 f_N^2(z) + \Phi_N(z) + \Psi_N(z), \tag{2.11}$$

where

$$\Phi_N(z) = \frac{v^2}{N}\,\mathsf{E}\langle G_N^2(z)\rangle$$

and

$$\Psi_N(z) = \mathsf{E}\, g_N(z) g_N(z) - \mathsf{E}\, g_N(z)\, \mathsf{E}\, g_N(z).$$

We see that (2.11) has a form close to (1.8) and all that we need is to show that $\Phi_N(z)$ and $\Psi_N(z)$ vanish as $N \to \infty$.

The first condition is fulfilled for $z \in \mathbb{C}_\pm$ because

$$\langle G_N^2(z)\rangle \le \left\|G_N^2(z)\right\| \le \|G_N(z)\|^2 \le \frac{1}{|\operatorname{Im} z|^2}. \tag{2.12}$$

Thus,

$$|\Phi_N(z)| \le \frac{v^2}{N\eta^2}, \tag{2.13}$$

where we have denoted $\eta := |\operatorname{Im} z|$.

The second condition reflects the selfaveraging property of $g_N(z)$. We are going to prove below that

$$\mathsf{E}\,|g_N - \mathsf{E}\, g_N(z)|^2 \le \frac{4v^2}{\eta^4 N^2} \tag{2.14}$$

provided $\eta \ge 4v^2$. Then (1.7) will be proved.

Let us introduce the centered random variable

$$g^\circ(z) = g(z) - \mathsf{E}\, g(z)$$

(we omit subscript N when no confusion can arise). It is easy to see that

$$\mathsf{E}\, g^\circ(z_1) g^\circ(z_2) = \mathsf{E}\, g^\circ(z_1) g(z_2).$$

Slight modification of (2.2) reads as

$$\mathsf{E}\, g^\circ(\hat z)\, g(z) = -\zeta \frac{1}{N} \sum_{x,y} \mathsf{E}\, g^\circ(\hat z)\, G(x,y) A_N(x,y).$$

The first term of the right-hand side vanishes because $\mathsf{E}\, g^\circ = 0$.

Using once more (1.10), one can write that

$$\begin{aligned}\mathsf{E}\, g^\circ(\hat z)\, G(x,y) A_N(x,y) = &\sum_{s,t=1}^{N} \mathsf{E}\, A_N(y,x) A_N(s,t) \\ &\times \left\{ \mathsf{E}\, g^\circ(\hat z) \frac{\partial G_N(x,y)}{\partial A_N(s,t)} + \mathsf{E}\, G_N(x,y) \frac{\partial g^\circ(\hat z)}{\partial A_N(s,t)} \right\}.\end{aligned}$$

The first derivative in the curly brackets is already computed. The second gives

$$\frac{\partial g^\circ(\hat z)}{\partial A_N(s,t)} = \frac{\partial g(\hat z)}{\partial A_N(s,t)} = -\frac{1}{N} \sum_{u=1}^{N} G(u,s) G(t,u) = -\frac{1}{N} \hat G^2(t,s),$$

where $\hat G \equiv G_N(\hat z)$.

After some simple manipulations, we derive our second main relation

$$\mathsf{E}\, g^\circ(\hat z)g(z) = \zeta v^2\, \mathsf{E}\, g^\circ(\hat z)g(z)g(z) + \zeta\frac{v^2}{N}\, \mathsf{E}\, g^\circ(\hat z)\left\langle G^2\right\rangle + \zeta\frac{2v^2}{N^2}\, \mathsf{E}\left\langle \hat G^2 G\right\rangle . \tag{2.15}$$

All that we need now is the identity

$$\mathsf{E}\, g^\circ(\hat z)g(z)g(z) = \mathsf{E}\, g^\circ(\hat z)g(z)\, \mathsf{E}\, g(z) + \mathsf{E}\, g^\circ(\hat z)g^\circ(z)g(z)$$

and the fact that $g(\bar z) = \overline{g(z)}$.

Regarding (2.15) with $\hat z = \bar z$, we derive that

$$\begin{aligned}\mathsf{E}\,|g^\circ(z)|^2 &\equiv \mathsf{E}\, g^\circ(\hat z)g(z)\\ &\le |\zeta|\, v^2\, \mathsf{E}\, g^\circ(\hat z)g(z)\, |\mathsf{E}\, g(z)| + |\zeta|\, v^2\, \mathsf{E}\, g^\circ(\hat z)g^\circ(z)\, |g(z)|\\ &\quad + \frac{|\zeta|\, v^2}{N}\, \mathsf{E}\, |g^\circ(\hat z)|\, \left|\left\langle G^2\right\rangle\right| + \frac{2\,|\zeta|\, v^2}{N^2}\, \mathsf{E}\left|\left\langle \hat G^2 G\right\rangle\right| .\end{aligned} \tag{2.16}$$

Using estimates similar to (2.12), we obtain that

$$\mathsf{E}\,|g^\circ(z)|^2 \le 2v^2\eta^{-1}\, \mathsf{E}\,|g^\circ(z)|^2 + \frac{v^2}{N\eta^3}\left[\mathsf{E}\,|g^\circ(z)|^2\right]^{1/2} + \frac{2v^2}{N^2\eta^4}\ . \tag{2.17}$$

This implies (2.14). The semicircle law (1.7) is proved for GOE. □

Let us make several important remarks here.

Remark. The first observation is that the estimate (2.6) indicates fairly fast decreasing of the variance of the random variable

$$g_N(z) = \frac{1}{N}\sum_{x=1}^{N} G_N(x,x;z) \tag{2.9}$$

as $N \to \infty$. It follows from (2.6) and the Borel-Cantelli lemma that $g_N(z)$ converges to a non-random limit (actually, $f_{\mathrm{W}}(z)$) with probability 1. Let us stress that in the classical probability theory the variance of the sum of independent random variables $S_N = (\xi_1 + \cdots + \xi_N)N^{-1}$ that is analogous to (2.17) is of the order N^{-1}. The difference is that in (2.17) we have a sum of dependent random variables $G_N(x,x;z)$. Let us note that (2.13) is a consequence of the more powerful statement that the centered random variable

$$\gamma_N(z) = \mathrm{Tr}\, G_N(z) - \mathsf{E}\,\mathrm{Tr}\, G_N(z)$$

converges in distribution to a gaussian random variable as $N \to \infty$. We discuss this property in more details in Lecture 4.

Remark. The next remark is related to the observation that estimates (2.13) and (2.14) show that in (2.11) terms $\Phi_N(z)$ and $\Psi_N(z)$ vanish not only for $\eta > \eta_0$ but also for z with imaginary part vanishing at the same time as N increases. This implies serious consequences concerning the smoothed eigenvalue density of large random matrices. In particular, one can trace out the proof of a version of the famous universality conjecture for local properties of random matrix spectra. We address this topics also in Lecture 4.

Now let us discuss generalization of the semicircle law to the case of random matrices with arbitrary distributed random entries $a(x,y)$.

The first ensemble generalizing GOE is the *Wigner ensemble* of random real symmetric matrices

$$W_N(x,y) = \frac{1}{\sqrt{N}} w(x,y), \quad x,y = 1,\dots,N$$

whose entries are jointly independent random variables satisfying conditions (1.1). We do not assume the probability distribution functions

$$P_{(x,y)}(\xi) = \mathsf{Prob}\{w(x,y) \le \xi\}$$

are the same and have a special form. In fact, one can consider here the more general case when the distributions $P^{(N)}_{(x,y)}$ can be dependent on N. In this case it should be pointed out that the set of random variables $\{w_N(x,y)\}_{1\le x\le y\le N}$ is determined on the same probability space.

To complete preparations, let us introduce notations for the moments of $w_N(x,y)$

$$V_j^{(N)}(x,y) = \int_{-\infty}^{\infty} \xi^j \,\mathrm{d}\, P_N(x,y).$$

Theorem 2.1. *Let us consider the ensemble*

$$H_N = h_N + W_N, \tag{2.18}$$

where h_N is a sequence of non-random matrices such that there exists the limit

$$\mu(\lambda) = \lim_{N\to\infty} \sigma(\lambda; h_N).$$

Let $P^{(N)}_{(x,y)}$ satisfy the Lindeberg condition

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{x,y=1}^{N} \int_{\tau\sqrt{N}} \xi^2 \,\mathrm{d}\, P^{(N)}_{(x,y)}(\xi) = 0 \quad \forall \tau > 0. \tag{2.19}$$

Then the Stieltjes transform

$$g_N^{(1)}(z) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\,\sigma(\lambda; H_N)}{\lambda - z}$$

converges in probability as $N \to \infty$ to a nonrandom function $f_h(z)$ that satisfies the equation

$$f_h(z) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\,\mu(\lambda)}{\lambda - z - v^2 f_h(z)}. \tag{2.20}$$

This statement is somewhat more general than the theorem proved by Pastur [**23**]. He considered H_N with a diagonal non-random part h_N and the conditions imposed on P_N were a bit restrictive than (2.18) but also close to the Lindeberg conditions.

Another generalization of GOE is given by the real symmetric matrices Γ_N

$$\Gamma_N(x,y) = \frac{1}{\sqrt{N}}\gamma(x,y), \tag{2.21}$$

where random variables $\gamma(x,y)$, $x \le y$ have joint Gaussian distribution with zero average and covariance

$$\mathsf{E}\,\gamma(x,y)\gamma(s,t) = V(x,s)V(y,t) + V(x,t)V(y,s),$$

where V is a symmetric and non-negatively defined matrix.

Theorem 2.2. [**3**] *Let the matrices*

$$V_N(x,y) = \begin{cases} V(x,y), & \textit{if } x \le N \textit{ and } y \le N, \\ 0, & \textit{otherwise} \end{cases}$$

be bounded

$$\|V_N\| \le v \tag{2.22}$$

and satisfy condition

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

Then the Stieltjes transform

$$g_N^{(2)} = \int_{-\infty}^{\infty} \mathrm{d}\,\sigma(\lambda;\Gamma_N)\lambda - z$$

converges with probability 1 *to a nonrandom function* $f_2(z)$ *given by*

$$f_2(z) = \int_0^v \frac{\mathrm{d}\,\nu(\lambda)}{-z-\lambda\phi(z)}$$

and $\phi(z)$ *satisfies the equation*

$$\phi(z) = \int_0^v \frac{\lambda\,\mathrm{d}\,\nu(\lambda)}{-z-\lambda\phi(z)}\ . \tag{2.23}$$

Lecture 3

Smoothed Eigenvalue Density

In the spectral theory of random matrices, the universality conjecture can be regarded as the most challenging problem. It concerns the local spectral characteristics of large random matrices.

In paper [**11**] we developed an approach to study the asymptotic regime that can be called *semi-local*, or *mesoscopic*. The subject under consideration is the eigenvalue distribution function smoothed over the intervals Δ_N of the length $1 \ll |\Delta_N| \ll N$, $N \to \infty$. In papers [**2, 4**] we proved limiting theorems that reflect the universality property of the smoothed eigenvalue density of large random matrices.

In this lecture we present theorems of papers [**11**] and [**2**] and describe briefly the scheme of their proofs.

It is not hard to see that the Stieltjes transform $f_N(z)$ with $\operatorname{Im} z = \varepsilon > 0$ effects the control of the eigenvalues that are situated in the vicinity of the interval $(\lambda - \varepsilon, \lambda + \varepsilon)$. It becomes clear if one consider $\operatorname{Im} f_N(\lambda + \mathrm{i}\,\varepsilon)$ as a smoothing of the measure $\mathrm{d}\,\sigma_N(\lambda)$:

$$\operatorname{Im} f_N(\lambda + \mathrm{i}\,\varepsilon) = \int \frac{\varepsilon}{(\lambda - \mu)^2 + \varepsilon^2}\, \mathrm{d}\,\sigma_N(\mu) \equiv \int \varphi_{\varepsilon,\lambda}(\mu)\, \mathrm{d}\,\sigma_N(\mu).$$

In these terms, the limiting transition $N \to \infty$ for random variable $f_N(\lambda + \mathrm{i}\,\varepsilon)$ with given positive ε can be regarded as the global spectral characteristics, i.e. as the variable related with $\mathrm{O}(N)$ eigenvalues of A_N.

Indeed, the variable

$$\sigma(\lambda+\varepsilon; A_N) - \sigma(\lambda-\varepsilon; A_N),$$
$$|\lambda| < 2v$$

is also global because it involves $\mathrm{O}(N)$ eigenvalues of A_N.

If one is interested in more detailed description of the limiting eigenvalue distribution of A_N, it is natural to study the limit of $f_N(\lambda + \mathrm{i}\,\varepsilon_N)$, where $\varepsilon_N \to 0$ at the same time as N infinitely increases. We call the variable

$$\xi_N(\lambda) := \operatorname{Im} f_N(\lambda + \mathrm{i}\,\varepsilon_N)$$

the *smoothed* or *regularized eigenvalue density*.

In these studies, one can separate two regimes that are fairly different. The asymptotic regime that correspond to the case of $\varepsilon_N = \mathrm{O}(N^{-1})$ is known as the local one. The regime intermediate between the global and local ones can be described as $\varepsilon_N = N^{-\alpha}$ with $0 < \alpha < 1$. According to the theoretical physics terminology, it can be called the *mesoscopic regime*.

In this lecture we are going to discuss the proof and several consequences of the following statements.

Theorem 3.1 ([**11**]). *Consider the* GOE A_N *and the resolvent* $G_N(z) = (A_N - z)^{-1}$. *Then convergence in average*

$$\lim_{N\to\infty} \mathsf{E}\,\frac{1}{N}\operatorname{Tr} G_N(\lambda + \mathrm{i}\,N^{-\alpha}) = -\frac{\lambda}{2v^2} + \mathrm{i}\,\frac{\sqrt{4v^2-\lambda^2}}{2v^2} \tag{3.1}$$

holds provided $0 < \alpha < 1$ *and* $|\lambda| < 2v$.

Let us look once more at the scheme presented in Lecture 2. It is clear that it does not work directly because one does not have any more estimates of the type (2.5). Then the relation (2.8) cannot be reduced to inequality (2.9) that gives the estimate of the variance $\mathsf{E}\,|g^\circ|^2$ by itself multiplied by $\mathsf{E}\,|g|$ is out of use. As a consequence, one cannot derive (2.7) from (2.9). Therefore the proof of Theorem 3.1 requires essential modifications of the approach.

To do this, we have to pass back from the short scheme described in Lecture 2 to the infinite system of relations inspired by the idea of Berezin. Let us explain now how it has to be modified.

The estimate (2.6) and its consequence (2.8) allows one to consider only first $k \le k_0$ relations from the infinite system (2.7). The matter is that the infinite system of relations (2.7) and its finite counterpart are related via the term L_{k_0}. If one consider a new system of k_0 relations of the form (2.7) but with the term L_{k_0+1} removed in the relation number k_0, this will change a little the first components of the solution of this equation with respect to the first several moments $L_k^{(N)}$. This is due to a priori estimate $|L_k| \le \eta^{-k}$. The truncated system of k_0 relations is closed and can be solved uniquely. This procedure is in fact equivalent to that one described in Lecture 2 in the long scheme.

Scheme of the proof of Theorem 3.1. The first principal modification of the scheme of Lecture 2 is that we consider the part of (2.7) with $k \ge k_0 > 1$ and rewrite it in the form

$$v^2 L_{k+1}^{(N)} = L_{k-1} - zL_k - \Psi_N^{(k)}. \tag{3.2}$$

To simplify the description of the proof, let us assume that $\operatorname{Re} z = 0$. Then under conditions of Theorem 3.1 relation (3.2) will have the form

$$L_{k+1}^{(N)} = \frac{1}{v^2} L_{k-1} - \frac{\mathrm{i}}{v^2 N^\alpha} L_k - \Psi_N^{(k)}. \tag{3.3}$$

Since $\operatorname{Im} z = N^{-\alpha}$, we cannot use the absolute estimates as it is done in Lecture 2 (see estimate (2.8)). One should use the estimates with respect to L_k.

Loosely speaking, the term $\Psi_N^{(k)}$ can be estimated by $N^{-\gamma} L_k^{(N)}$ with some $\gamma > 0$ (to make possible this estimate, we will need our second principal modification).

Using the fact that L_k enters into relation (3.3) with factor $N^{-\gamma}$, one can reduce it by subsequent substitutions to the form

$$L_{k+1}^{(N)} = \frac{1}{v^2} L_{k-1} + \sum_{l=1}^{k} \left(\frac{\mathrm{i}}{v^2 N^\gamma} \right)^l L_{k-l+1} + \hat{\Psi}_N^{(k)}(z). \tag{3.4}$$

Inequality (2.6) shows that $L_1 \le N^\alpha$, but it enters (3.4) with the factor $N^{-k\gamma}$. This allows us to neglect the terms that increase provided $k > k_0$, where k_0 is sufficiently large.

The term $\Psi_N^{(k)}$ involves the factors $\mathsf{E}\, g_N^k(N)$ and cannot be directly estimated in terms of $L_k^{(N)}$. To do this, we pass from complex variables $g_N(z)$ to the real variables $\xi = \lambda g_N(\lambda + \mathrm{i}\, N^{-\alpha}) \ge 0$ and $\mu = \operatorname{Re} g_N(\lambda + \mathrm{i}\, N^{-\alpha})$. This is the second modification of the general scheme. It seems to be a trivial one, but this is not completely right.

The matter is that consideration of the moments $\mathsf{E}\, g_N^k$ leads to the necessity of consideration of the family of moments $\mathsf{E}\, \xi^p \mu^q$. Those with large number of q are difficult to estimate. Fortunately, we will need the family of three types of moments

$$W_k^{(N)} = \mathsf{E}\,\xi_N^k,$$
$$V_k^{(N)} = \mathsf{E}\,\mu_N \xi_N^k,$$
$$U_k^{(N)} = \mathsf{E}\,\mu_N^2 \xi_N^k.$$

This family is closed and satisfies our conditions.

Proof of Theorem 3.1. Let us introduce the matrices

$$B_\lambda = N^\alpha (A_N - \lambda I), \quad B = N^\alpha A_N,$$

and

$$P = \frac{1}{1 + B_\lambda^2}, \quad Q_\lambda = \frac{B_\lambda}{1 + B_\lambda^2}.$$

Then

$$\xi_N(\lambda) = \overline{P_\lambda}^\alpha, \quad \mu = \overline{Q_\lambda}^\alpha,$$

where for we denoted for N-dimensional matrix H_N

$$\overline{H_N}^\alpha = N^{\alpha-1}\,\mathrm{Tr}\, H_N.$$

It is not hard to derive that

$$\frac{\partial P(x,y)}{\partial B(s,t)} = -P(x,s)Q(t,y) - P(x,t)Q(s,y) - Q(x,s)P(t,y) - Q(x,t)P(s,y)$$

and

$$\frac{\partial Q(x,y)}{\partial B(s,t)} = P(x,s)P(t,y) + P(x,t)P(s,y) - Q(x,s)Q(t,y) - Q(x,t)Q(s,y)$$

These relations represent (2.3) rewritten for real and imaginary parts separately.

Using these relations and the formula (1.10) and regarding the identity

$$\begin{aligned} P_\lambda &= I - P_\lambda B_\lambda^2 \\ &= I + \lambda N^\alpha Q_\lambda - Q_\lambda B, \end{aligned}$$

one can derive the system of relations

$$W_{k+1} = \frac{1}{v^2} W_{k-1} + \frac{1}{v^2} V_{k-1} + U_{k-1} + \Gamma_1^{(N)}(k), \tag{3.5a}$$

$$V_{k+1} = -\frac{\lambda}{2v^2} W_{k+1} + \frac{1}{2v^2 N^\alpha} V_k + \Gamma_2^{(N)}(k), \tag{3.5b}$$

and

$$U_{k+1} = -\frac{\lambda}{2v^2} V_{k+1} + \frac{1}{2v^2 N^\alpha} U_k + \Gamma_3^{(N)}(k), \tag{3.5c}$$

where

$$\Gamma_1^{(N)}(k) = \frac{v^2}{N^{1-\alpha}} \mathsf{E}\,\xi^{k-1}(\overline{P_\lambda^2}^\alpha + \overline{Q_\lambda^2}^\alpha)\frac{4v^2(k-1)}{N^{2-2\alpha}} \mathsf{E}\,\xi^{k-2}\overline{P_\lambda Q_\lambda^2}^\alpha, \tag{3.5}$$

$$\Gamma_2^{(N)}(k) = -\frac{2v^2}{N^{1-\alpha}} \mathsf{E}\,\xi^k\overline{P_\lambda Q_\lambda}^\alpha - \frac{4v^2k}{N^{2-2\alpha}} \mathsf{E}\,\xi^{k-1}\overline{P_\lambda^2 Q_\lambda}^\alpha, \tag{3.6}$$

$$\Gamma_3^{(N)}(k) = -\frac{2v^2}{N^{1-\alpha}} \mathsf{E}\,\mu\xi^k\overline{P_\lambda Q_\lambda}^\alpha - \frac{2v^2}{N^{1-\alpha}} \mathsf{E}\,\xi^k(\overline{P_\lambda^3}^\alpha - \overline{P_\lambda^2 Q_\lambda}^\alpha). \tag{3.7}$$

It is easy to see that

$$\left|\Gamma_1^{(N)}(k)\right| \le \frac{8(k-1)}{N^{2-2\alpha}}W_{k-1} + \frac{2}{N^{1-\alpha}}W_k + \frac{1}{N^\alpha}W_k,$$

and $\Gamma_2^{(N)}(k)$ and $\Gamma_3^{(N)}(k)$ can be estimated similarly. This estimate looks appropriate excepting the fact that we have in the right-hand side the term W_{k-1}. To ensure that W_{k-1} can be estimated via W_k, we prove the following simple statement.

Lemma 3.2. *Under conditions of Theorem* 3.1

$$\mathsf{E}\,\xi^2 \ge \frac{1}{2}\frac{4v^2-\lambda^2}{4v^4} \equiv \frac{\pi\rho(\lambda)}{2}$$

for large enough N.

Proof. Let consider (3.5a) with $k=1$;

$$\mathsf{E}\,\xi^2 = v^{-2} + \lambda v^{-2}\,\mathsf{E}\,\mu + \mathsf{E}\,\mu^2 + \Gamma_1^{(N)}(1),$$

where

$$\left|\Gamma_1^{(N)}(1)\right| \le 3N^{-\chi}\,\mathsf{E}\,\xi \text{ with } \chi = \min\{\alpha; 1-\alpha\}.$$

Since $|\mathsf{E}\,\mu| \le \sqrt{\mathsf{E}\,\mu^2}$, then we can write that

$$\begin{aligned}\mathsf{E}\,\xi^2 &\ge \left(\sqrt{\mathsf{E}\,\mu^2} - \frac{\lambda}{2v^2}\right)^2 + \frac{4v^2-\lambda^2}{4v^4} - \left|\Gamma_1^{(N)}(1)\right| \\ &\ge (\pi\rho)^2 - 3N^{-\chi}\sqrt{\mathsf{E}\,\xi^2}.\end{aligned}$$

□

An important consequence of Lemma 3.1 is that

$$w_k := (W_k)^{1/k} \ge \frac{\pi\rho}{2} \text{ and } W_k \le W_{k+m}\left(\frac{2}{\pi\rho}\right)^m.$$

Now we can derive from (3.5b) the relation

$$V_k = -\frac{\lambda}{2v^2}W_{k-1} + \tilde{\Gamma}_2^{(N)}(k),$$

where

$$\left|\tilde{\Gamma}_2^{(N)}(k)\right| \le \left(1+\frac{\lambda}{2v^2}\right)\sum_{p=1}^{k-1}\left(\frac{1}{2v^2N^\alpha}\right)^p W_{k-1-p} + \frac{1}{(2v^2N^\alpha)^{k-1}}U_0$$
$$\le \frac{1}{\pi\rho v^2N^\alpha(1-[\pi\rho v^2N^\alpha]^{-1})}W_{k-1} + \frac{N^\alpha}{(2v^2N^\alpha)^{k-1}}.$$

Thus,

$$V_k = -\frac{\lambda}{2v^2}W_{k-1}(1+o(1)).$$

Similar computations lead to the relation

$$U_k = -\frac{\lambda}{2v^2}V_{k-1}(1+o(1)).$$

Substituting these two last equalities into (3.5a) and treating it in the same way, one can obtain easily that

$$W_{k+1} = (\pi\rho)^2W_{k-1} + 4N^{-\chi}W_k + \tilde{\Gamma}_1^{(N)}(k).$$

Then, an elementary procedure leads to the proof of convergences

$$\mathsf{E}\,\xi \to \pi\rho,$$
$$\mathsf{E}\,\mu \to -\frac{\lambda}{2v^2},$$
$$\mathsf{E}\,\mu^2 \to \frac{\lambda^2}{4v^4}.$$

Theorem 3.1 is proved. □

In this statement, the most important is the convergence of the smoothed density of eigenvalues

$$\mathsf{E}\,\xi_N(\lambda) \equiv \mathsf{E}\,\mathrm{Im}\, g_N(\lambda + \mathrm{i}\,N^{-\alpha}) \to \pi\rho(\lambda), \quad N\to\infty. \tag{3.6}$$

This relation plays a crucial role in the proof of the selfaveraging (or strong selfaveraging) property of the random variable $\xi_N(\lambda)$ and in the proof of the universal behaviour of the correlation function $\mathsf{E}\,\xi_N(\lambda_1)\xi_N(\lambda_2)$ as well. Let us formulate the corresponding results.

Theorem 3.3. *Under hypotheses of Theorem* 3.1*,*

$$\mathsf{E}\left|g^\circ(\lambda + \mathrm{i}\,N^{-\alpha})\right|^2 = \mathrm{O}(N^{2-2\alpha}). \tag{3.7}$$

Proof. To explain the proof of this statement, let us consider relation (2.15) and assume once more that $\lambda = 0$. Using identity

$$\mathsf{E}\,g^\circ gg = 2\,\mathsf{E}\,g^\circ g\,\mathsf{E}\,g + \mathsf{E}\,g^\circ g^\circ g^\circ,$$

we can rewrite (2.15) in the form

$$2v^2 \,\mathsf{E}\, g^\circ g \,\mathsf{E}\, g \le \frac{1}{N^\alpha} \,\mathsf{E}\, g^\circ g + \frac{1}{N^{1-\alpha}} \,\mathsf{E}\, |g^\circ| + \mathrm{O}\left(\frac{1}{N^{2-2\alpha}}\right).$$

If we have (3.6), it is not hard to derive from this inequality the estimate (3.7). □

Finally, let us formulate the theorem about the fluctuations of the smoothed eigenvalue density of GOE. This is an analog of the central limit theorem.

Theorem 3.4. [**2**] *Under hypotheses of Theorem* 3.1, *the random variable*

$$\gamma_N(\lambda) = N^{1-\alpha}[\xi_N(\lambda) - \mathsf{E}\,\xi_N(\lambda)]$$

converges in distribution as $N \to \infty$ to a centered Gaussian random variable with variance $1/4$. *If one consider two points $\lambda_1 \neq \lambda_2$ such that $\lambda_1, \lambda_2 \to \lambda \in (-2v, 2v)$, then*

$$\mathsf{E}\,\xi_N^o(\lambda_1)\xi_N^o(\lambda_2) = -\frac{1}{N^2(\lambda_1-\lambda_2)^2}(1+o(1)) \tag{3.8}$$

in the limit $N \to \infty$ provided

$$\frac{1}{N^\alpha} \ll |\lambda_1 - \lambda_2| \ll 1. \tag{3.9}$$

Generalizations of Theorems 3.1-3.3 and universality conjecture

Wigner random matrices. It should be noted that Theorems 3.1-3.3 considered for $0 < \alpha < \alpha_0$ are valid for the Wigner ensemble of random matrices (2.8) with jointly independent arbitrary distributed random variables $w(x, y)$ having several first moments finite

$$\mathsf{E}[w_N(x,y)]^{2k} = V_k < \infty.$$

In particular, Theorem 3.3 holds for $k = 4$ and $\alpha_0 = 1/8$ (see [**2**]-II). These results show that the universality conjecture holds for large random matrices with independent entries. They are far from being optimal, and it is interesting to check out the optimal bound for α_0 and its dependence on V_k.

Wishart-type random matrices. Let us consider the ensemble of random matrices

$$H_{m,N}(x,y) = \frac{1}{N}\sum_{\mu=1}^{m} \theta_\mu(x)\theta_\mu(y), \quad x, y = 1, \dots, N, \tag{3.10}$$

where the random variables $\{\theta_\mu(x)\}, x, \mu \in \mathbb{N}$ have joint Gaussian distribution with zero mathematical expectation and covariance

$$\mathsf{E}\{\theta_\mu(x)\theta_\nu(y)\} = u^2\delta_{xy}\delta_{\mu\nu}.$$

Here δ_{xy} denotes the Kronecker delta-symbol.

This ensemble introduced in mathematical statistics is known for $m = N$ as the Wishart ensemble.

The eigenvalue distribution of (3.10) in the limit $N, m \to \infty$ was considered first [**21**], where more general random matrix ensembles were also considered.

Random matrices of the form (3.10) are at present of extensive use in the statistical mechanics of disordered spin systems and in the models of memory in the theory of neural networks. The difference between the Wigner random matrices (2.8) and (3.10) is that in the second case the entries $H_{m,N}(x, y)$ are statistically dependent random variables.

Theorem 3.5. *Let $G_{m,N}(z) = (H_{m,N} - z)^{-1}$. Then, for $N, m \to \infty$, $m/N \to c > 0$, the random variable*

$$R^{(\alpha)}_{m,N}(\lambda) := \operatorname{Im} \operatorname{Tr} G_{m,N}(\lambda + \mathrm{i}\, N^{-\alpha})N^{-1}$$

converges with probability 1 *as $N \to \infty$ to the nonrandom limit*

$$\pi\varrho_c(\lambda) = \frac{1}{2\lambda u^2}\sqrt{4cu^4 - [\lambda - (1 + cu^2)]^2} \tag{3.11}$$

provided $0 < \alpha < 1$ and $\lambda \in \Lambda_{c,u} = \big(u^2(1 - \sqrt{c})^2, u^2(1 + \sqrt{c})^2\big)$.

Remark. The limiting expression (3.11) for the eigenvalue distribution was derived in the global regime in [**21**].

Theorem 3.6. *Consider k random variables, $i = 1, \dots, k$,*

$$\gamma^{(\alpha)}_{m,N}(i) \equiv N^{1-\alpha}\left[R^{(\alpha)}_{m,N}(\lambda_i) - \mathsf{E}\, R^{(\alpha)}_{m,N}(\lambda_i)\right],$$

where $\lambda_i = \lambda + \tau_i N^{-\alpha}$ with given τ_i. Then under hypotheses of Theorem 3.5 *the joint distribution of the vector $(\gamma_N(1), \dots, \gamma_N(k))$ converges to the centered Gaussian k-dimensional distribution with covariance*

$$C(\tau_i, \tau_j) = \frac{4 - (\tau_i - \tau_j)^2}{[4 + (\tau_i - \tau_j)^2]^2}. \tag{3.12}$$

Remark. It is easy to see that if $|\tau_1 - \tau_2| \to \infty$, then

$$C(\tau_1, \tau_2) = -(\tau_1 - \tau_2)^{-2}(1 + o(1)). \tag{3.13}$$

This coincides with the average value of the Dyson's 2-point correlation function for real symmetric matrices considered at large distances $|t_1 - t_2| \gg 1$.

Let us note that the limiting distribution of the random variables γ_N and $\gamma_{m,N}$ coincide and do not depend on particular values of λ and α. This shows that the fluctuations of the smoothed eigenvalue density ξ are universal in the mesoscopic regime. Thus, our results can be regarded as a support of the universality conjecture for local spectral statistics of large random matrices.

Lecture 4

Eigenvalues outside of the limiting spectrum

Our main goal in this section is to describe the proof of estimate (1.17) and discuss its consequence with respect to the norm of random matrices.

Moments and extreme eigenvalues. First of all, let us note that having proved (1.17) for all $k \le N^\beta$ with $\beta > 0$, we easily derive from (1.13) that inequality

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

with positive ε also holds for all $k \le N^\beta$ and $N > N_0(\varepsilon)$.

Regarding the numbers $m_{2k}^* \equiv m_{2k}^*(\varepsilon)$ determined by the following recurrence relations

$$m_{2k}^* = (1+\varepsilon)^2 v^2 \sum_{j=0}^{k-1} m_{2k-2-2j}^* m_{2j}^*, \quad m_0^* = 1,$$

we then derive that inequality

$$M_{2k}^{(N)} \le m_{2k}^*$$

holds for $k \le N^\beta$.

Now let us follow the reasoning that is usual in the norm estimates for random matrices (see, for example [**7, 8**]). Taking into account that the family $m_{2k}^*(\varepsilon), k \in \mathbb{N}$ represents the moments of the semicircle distribution (1.5) with v^2 replaced by $[(1+\varepsilon)v]^2$, we obtain the estimate

$$m_{2k}^*(\varepsilon) \le [(1+\varepsilon)v]^{2k}.$$

Thus, we have that

$$M_{2k}^{(N)} \le [(1+\varepsilon)v]^{2k} \quad \forall k \le N^{\beta}. \tag{4.1}$$

This implies the estimate with probability 1

$$\limsup_{N\to\infty} \|A_N\| \le 2v(1+2\varepsilon), \tag{4.2}$$

where the spectral norm $\|A_N\|$ is defined as the largest absolute value of an eigenvalue of A_N.

Inequality (4.2) can be derived from (4.1) using elementary computations. Indeed, if one denotes by $n_N(s)$ the number of eigenvalues lying outside of the interval $(-s, s)$

$$n_N(s) \equiv \#\{j \mid \left|\lambda_j^{(N)}\right| \ge s\},$$

then one can write the sequence of inequalities

$$M_{2k}^{(N)} \ge \mathsf{E} \int_{\mathbb{R}\setminus(-s,s)} \ge s^{2k}\, \mathsf{E}\, n_N(s) \ge \frac{s^{2k}}{N} \mathsf{Prob}\{\|A_N\| \ge s\}.$$

Then for $P_N(\varepsilon) \equiv \mathsf{Prob}\{\|A_N\| \ge 2v(1+2\varepsilon)\}$ we have the estimate

$$P_N(\varepsilon) \le N \inf_k \frac{M_{2k}^{(N)}}{[2v(1+\varepsilon)]^{2k}} = N \exp\{-N^{\beta} \log(1+\varepsilon/2)\}$$

that implies (4.2).

Let us discuss two aspects of the results presented. Inequality (4.2) is valid for all positive ε. This means that (4.2) holds for $\varepsilon = 0$ and this implies that the maximal eigenvalue of A_N is bounded by $2v$ in the limit $N \to \infty$. ¿From the other hand, the semicircle law states that with probability 1 there exist eigenvalues of A_N falling into vicinity of $2v$ in this limit. Thus, the maximal eigenvalue (and also the minimal one, due to symmetry of the probability distribution of A_N (1.3)) converges to $2v$ $(-2v)$.

This fact is also valid for the Wigner ensemble of random matrices W_N (2.17) with arbitrary distributed entries $w(x,y)N^{-1/2}$ provided that several first moments $\mathsf{E}\, w(x,y)^{2p}$ are finite (see for example [**7, 13**] and [**26**]).

The second remark concerns the maximal power β_0 in (4.1). This (critical) exponent reflects the behaviour of the differences Δ_l between the eigenvalues from the vicinity of the extremal eigenvalue $\lambda_{\max}^{(N)}$. Indeed, one can write that

$$\begin{aligned} M_{2k}^{(N)} &= \frac{1}{N}\left([\lambda_1^{(N)}]^{2k} + [\lambda_2^{(N)}]^{2k} + \dots [\lambda_N^{(N)}]^{2k}\right) \\ &= [\lambda_{\max}^{(N)}]^{2k}\big(1 + [1+\Delta_2^{(N)}]^{2k} + \dots\big) N^{-1}. \end{aligned}$$

If $2k$ grows faster with N than the difference $\Delta_2^{(N)}$ vanishes, then one gets an asymptotic behaviour of $M_{2k}^{(N)}$ different from (4.1).

Our scheme shows that (4.1) is valid for $k \ll N^{2/3}$ when the GOE is considered. Therefore one can conclude that $\beta_0 \geq 2/3$. The early studies of GOE with the help of the orthogonal polynomials approach [**6**] show that $\beta_0 = 2/3$. This conclusion is confirmed and improved by Tracy and Widom in a series of papers (see for example [**28**]). The same conclusion follows from the works by Sinai and Soshnikov [**26**] and by Soshnikov [**27**], where the Wigner ensemble is considered (see also [**13**] for certain generalizations of [**27**]). Equality $\beta_0 = 2/3$ means that the average distance between eigenvalues at the border of the is of the order $N^{-2/3}$.

Scheme of the proof of (1.17). We follow the technique developed in [**3**] for random matrices with Gaussian correlated entries. For simplicity, we consider in details only the case of independent random variables.

We rewrite definition (1.14) for B_{2k} in the form

$$B_{2k}^{(2)}(N) = \sum_{\substack{p_1,p_2\geq 1\\ p_1+p_2=2k}} \mathsf{E}\langle A_N^{p_1}\rangle^\circ \langle A_N^{p_2}\rangle^\circ,$$

where $\xi^\circ = \xi - \mathsf{E}\,\xi$. The general idea is to use recurrent relations for $B_{2k}^{(2)}$ that can be derived again with the help of (1.9). These relations are similar those we have got for $M_{2k}^{(N)}$ (1.13) and therefore one can expect to obtain estimates of B_{2k} in terms of M_{2k}.

Let us apply identity

$$\mathsf{E}\,\xi_1^\circ\xi_2^\circ = \mathsf{E}\,\xi_1^\circ\xi_2,$$

to $B_{2k}^{(2)}$ and rewrite it in the following form

$$B_{2k}^{(2)}(N) = \sum_{\substack{p_1,p_2\geq 1\\ p_1+p_2=2k}} \frac{1}{N}\sum_{x,y=1}^{N} \mathsf{E}\langle A_N^{p_1}\rangle^\circ A_N^{p_2-1}(x,y)A_N(y,x).$$

Now we can apply to the last average (1.9) with $\gamma = A_N(y,x)$. After simple computations similar to the formula (1.12), we obtain equality

$$\begin{aligned} B_{2k}^{(2)}(N) &= 2v^2\sum_{j=0}^{k-1} B_{2k-2-2j}^{(2)}(N)M_{2j}^{(N)} + v^2 B_{2k-2}^{(3)}(N) \\ &\quad + v^2 \sum_{\substack{p_1,p_2\geq 1\\ p_1+p_2=2k-2}} \frac{p_2}{N}\,\mathsf{E}\langle A_N^{p_1}\rangle^\circ\,\langle A_N^{p_2}\rangle + \frac{v^2(2k-2)^2}{2N^2}M_{2k-2}^{(N)}. \end{aligned} \tag{4.3}$$

Here and below we assume that

$$B_{2k}^{(m)}(N) = \sum_{\substack{p_1,\dots,p_m\ge 1\\ p_1+\dots+p_m=2k}} \mathsf{E}\, L^\circ_{p_1} L^\circ_{p_2}\cdots L^\circ_{p_m}, \tag{4.4}$$

where we denote $L_p = \langle A_N^p\rangle$.

Now let us remark that if regarding (4.3), we could forget about the term $B_{2k-2}^{(3)}$, then (1.17) would follow as a simple consequence of the ordinary principle of mathematical induction. Namely, assuming that the estimate

$$\left| B_{2j}^{(m)}(N)\right| \le \frac{(2j)^{m\tau}}{N^m} M_{2j}^{(N)} \tag{4.5}$$

taken with $m = 2$ holds for all $j \le k-1$ and substituting these inequalities into (4.3) with $B_{2k}^{(3)}$ ejected, one can derive after certain amount of computations that (4.5) holds also for $j = k$.

Presence of the term $B_{2k-2}^{(3)}$ makes the scheme of the proof more complicated, but not too much. The observation is that $B_{2k}^{(2)}$ is depends on $B_{2k-2}^{(3)}$, where variable m has increased from 2 to 3, but $2k$ has decreased to $2k-2$. Thus, one has just to modify the reasoning based on the mathematical induction principle.

Thus, our aim is to prove (4.5). We proceed from (4.3) by deriving a recurrent relation for $B_{2k}^{(m)}(N)$. It has a similar form, where $B_{2k}^{(m)}(N)$ is expressed in terms of $B_{2j}^{(m)}(N)$, $B_{2j}^{(m-1)}(N)$, $B_{2j}^{(m+1)}(N)$ with $j \le k-1$.

Let us note that due to definition of $B_{2k}^{(m)}$, one has always $m \le 2k$. The case of equality $m = 2k$ corresponds to one term in (4.4) where $p_1 = p_2 = \cdots = p_m = 1$. In this case estimate (4.5) can be verified by direct computations, as well as in the case of $m = 2k-1$.

Now, the scheme of the ordinary mathematical induction (the "linear" one) of the proof of (4.5) can be replaced by a two-dimensional scheme, where one moves along the points (k,m) such that $m+2k = L$. On the lines $m = 2k$ and $m = 2k-1$ relation (4.5) is easy to be verified. Next, assuming that (4.5) holds for all (k,m) such that $m+2k \le L$, one moves along the line $m+2k = L+1$ from the point closest to $m = 2k$ to the point with $m = 2$. The structure of relations (4.3) is such that this procedure leads to the estimate (4.5) for $B_{2k}^{(m)}(N)$ on the line $m+2k = L+1$. □

Now let us carry out several key-point computations of this proof. Regarding

$$B_{2k}^{(m)} = \sum_{p_i} \mathsf{E}[L^\circ_{p_1} L^\circ_{p_2}\cdots L^\circ_{p_{m-1}}]^\circ L_{p_m}$$

and applying to the last factor our usual scheme, we derive relation

$$B_{2k}^{(m)} = v^2 \sum_{p_i} \mathsf{E}[L_{p_1}^\circ L_{p_2}^\circ \cdots L_{p_{m-1}}^\circ]^\circ \Big\{ \sum_{q_1+q_2=p_m-2} L_{q_m} L_{q_{m+1}} + \frac{p_m - 1}{N} L_{p_m-2} \Big\}$$
$$+ \frac{v^2}{N} \sum_{p_i} \mathsf{E}\, L_{p_1}^\circ \cdots L_{p_{j-1}}^\circ \left(\frac{2p_j}{N} L_{p_m+p_j-2} \right) L_{p_{j+1}}^\circ \cdots L_{p_{m-1}}^\circ .$$

Using identity

$$\mathsf{E}\, X^\circ Y Z = \mathsf{E}\, X Y^\circ Z + \mathsf{E}\, X Y Z^\circ + \mathsf{E}\, X Y^\circ Z^\circ - \mathsf{E}\, X\, \mathsf{E}\, Y^\circ Z^\circ ,$$

and denoting

$$D_{2k}^{(m)} = \sum_{\substack{p_1,\dots,p_m \geq 1 \\ p_1+\cdots+p_m=2k}} \mathsf{E} |L_{p_1}^\circ L_{p_2}^\circ \cdots L_{p_m}^\circ| ,$$

we derive inequality

$$D_{2k}^{(m)} \leq 2v^2 \sum_{j=0}^{k-1} D_{2k-2-2j}^{(m)} M_{2j} + v^2 D_{2k-2}^{(m+1)}$$
$$+ v^2 \sum_{j=0}^{k-1} D_{2k-2-j}^{(m-1)} D_j^{(2)} + \frac{v^2(2k-1)}{N} D_{2k-2}^{(m)}$$
$$+ \frac{v^2(2k-1)^2}{N^2} \sum_{j=0}^{k-1} D_{2k-2-2j}^{(m-2)} M_{2j} + \frac{2v^2(2k-1)}{N^2} D_{2k-2}^{(m)} .$$

This implies the following inequality

$$\left| D_{2k}^{(m)} \right| \leq 2v^2 \sum_{j=0}^{k-1} \left| D_{2k-2-2j}^{(m)} \right| \; M_{2j}$$
$$+ v^2 \sum_{j=0}^{k-1} \frac{2j-1}{N^2} \left| D_{2k-2-2j}^{(m-2)} \right| \; M_{2j} + \Psi_k(N), \tag{4.6}$$

where $\Psi_k(N)$ contains unimportant terms. Among these terms there is the term $\left| D_{2k-2}^{(m+1)} \right|$ that is also of order smaller than $\left| D_{2k}^{(m)} \right|$. This means that we can actually return back to the ordinary mathematical induction of the proof of (4.5).

Assuming that (4.5) holds, we derive from (4.6) inequality

$$\left| D_{2k}^{(m)} \right| \leq 2v^2 \sum_{j=0}^{k-1} \frac{(2k-2-2j)^{m\tau}}{N^m} M_{2k-2-2j} M_{2j}$$
$$+ v^2 \sum_{j=0}^{k-1} \frac{(2k-2-2j)^{(m-2)\tau}}{N^m} (2j-1) M_{2k-2-2j} M_{2j} + |\Psi_k(N)| . \tag{4.7}$$

The first problem is related with number 2 in front of the first terms in the right-hand side of (4.7). However, one can easily avoid it assuming that $\tau > 1/2$. Then for all $m \geq 2$ the function $(2k-2t)^{m\tau}$ is convex and we deduce that

$$\left| D_{2k}^{(m)} \right| \leq \left\{ (2k-2)^{m\tau} + (2k-1)^{(m-2)\tau+1} \right\} X_{2k-2}(N),$$

where

$$X_{2k-2}(N) \equiv v^2 \sum_{j=0}^{k-1} (2j-1) M_{2k-2-2j}\ M_{2j}.$$

It follows from (1.13) that

$$\begin{aligned} |X_{2k-2}(N)| &= \left| M_{2k} - v^2 \frac{2k-1}{N} M_{2k-2} - v^2 D_{2k-2}^{(2)} \right| \\ &\leq M_{2k} + \frac{2k-1}{N} X_{2k-2}(N) + \frac{(2k-2)^{2\tau}}{N^2} X_{2k-2}(N). \end{aligned}$$

Combining these estimates, we see that to prove (4.5), one has to determine parameter τ and the relation between k and N in such a way that

(4.8)
$$\left((2k-2)^{m\tau} + (2k-1)^{(m-2)\tau+1} \right) \left(1 - \frac{2k-1}{N} - \frac{(2k-2)^{2\tau}}{N^2} \right) \leq (2k)^{m\tau}$$

for all $m \geq 2$ and all possible k.

We divide both sides by $(2k)^{m\tau}$, take $m = 2$ and observe that the inequality

$$\left(1 - \frac{1}{k} \right)^{2\tau} + \frac{4k-2}{(2k)^{2\tau}} < 1$$

is valid for large enough values of k only when τ is greater than 1.

In fact, taking $m = 2$, we provide the maximal value for the first factor from (4.8). Regarding its product with the second factor, we obtain that inequality (4.8) holds only when

$$\frac{2k}{N} + \frac{(2k-2)^{2\tau}}{N^2} < \frac{3}{2k}\ .$$

This inequality is true under condition that

$$k^{2\tau+1} < N^2.$$

This means that (4.5) holds in the limit $N \to \infty$ under condition that $k \ll N^{2/3}$.

Norm estimates. Let us complete this lecture with estimates for the norm of random matrices Γ_N (2.21) with correlated entries.

Theorem 4.1 ([**3**])**.** *Under hypotheses of Theorem* 2.2*,*

$$\limsup \|\Gamma_N\| \le 2\sqrt{v\nu_1}.$$

If the matrix V is such that

$$\frac{1}{N}\operatorname{Tr} V_N^r \le \int_0^\infty \lambda^r \, \mathrm{d}\,\nu(\lambda), \quad r \in \mathbb{N}, \tag{4.9}$$

then the upper bound of the support Λ of the distribution

$$\sigma_V(\lambda) = \lim_{N\to\infty} \sigma(\lambda; \Gamma_N)$$

coincides with the estimate from above for the norm. This means that there are no eigenvalues of Γ_N in the limit $N \to \infty$ outside of Λ.

Condition (4.9) holds for $V(x,y) = u(x-y)$ with $u(x) \ge 0$. However, in general one cannot guarantee that all eigenvalues of Γ_N are inside of Λ for $N = \infty$.

To show this, it is sufficient to consider $V(x,y) = w(x)$ with

$$w(x) = \begin{cases} v', & \text{if } x = 1, \\ v, & \text{if } x \ne 1 \end{cases}$$

with $v' > 4v$. Then

$$\Lambda = (-2v, 2v)$$

but

$$\|\Gamma_N e_1\| \to \sqrt{vv'} > 2v.$$

Theorem 4.2 gives sufficient conditions to avoid the situation when there could be eigenvalues outside of the support of the limiting eigenvalue distribution. This is important in a series of applications of random matrices, in particular in the statistical mechanics of disordered spin systems.

Bibliography

[1] F.A. Berezin, Some remarks on the Wigner distribution, *Teoret. Mat. Fiz.*, **17** (1973) 305–318.

[2] A. Boutet de Monvel and A. Khorunzhy, Asymptotic distribution of smoothed eigenvalue density. I. Gaussian Random Matrices & II. Wigner Random Matrices, *Random Oper. Stochastic Equations*, **7** (1999) 1–22 & 149–168.

[3] A. Boutet de Monvel and A. Khorunzhy, On the Norm and Eigenvalue Distribution of Large Random Matrices, *Ann. Probability*, **27** (1999) 913–944.

[4] A. Boutet de Monvel and A. Khorunzhy, On universality of smoothed eigenvalue density of large random matrices, *J. Phys. A: Math. Gen.*, **32** (1999) L413–L417.

[5] A. Boutet de Monvel, A. Khorunzhy and V. Vasilchuk, Limiting eigenvalue distribution of random matrices with correlated entries, *Markov Process. Related Fields*, **2**, 4 (1996), 607–636.

[6] B.V. Bronk, Accuracy of the semicircle approximation for the density of eigenvalues of random matrices, *J. Mathematical Phys.*, **5** (1964) 215–220.

[7] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, *Combinatorica*, **1** (1981) 233–241.

[8] S. Geman, A limit theorem for the norm of random matrices, *Ann. Probab.*, **8** (1980) 252–261.

[9] T. Guhr, A. Müller-Groeling and H. Weidenmüller, Random matrix theories in quantum physics: common concepts, *Physics Reports*, **299** (1998) 189–425.

[10] A. Khorunzhy, Eigenvalue distribution of large random matrices with correlated entries, *Mat. Fiz. Anal. Geom.*, **3** (1996) 80–101.

[11] A. Khorunzhy, On smoothed density of states for Wigner random matrices, *Random Oper. Stochastic Equations*, **5**, 2 (1997) 147–162.

[12] A. Khorunzhy, On dilute unitary random matrices, *J. Phys. A: Math. Gen.*, **31** (1998) 4773–4784.

[13] A. Khorunzhy, Sparse random matrices: spectral edge and statistics of rooted trees, *Adv. in Appl. Probab.*, **33** (2001) 1–18.

[14] A. Khorunzhy, B. Khoruzhenko and L. Pastur, Asymptotic properties of large random matrices with independent entries, *J. Math. Phys.*, **37** (1996) 5033–5060.

[15] A. Khorunzhy, B. Khoruzhenko, L. Pastur and M. Shcherbina, The large-n limit in statistical mechanics and the spectral theory of disordered systems, In: *Phase Transitions and Critical Phenomena*, Vol. **15**, Eds.: C. Domb and J.L. Lebowitz (1992) 73–239.

[16] A. Khorunzhy and W. Kirsch, Limit of infinite band width for product of two random matrices, *Random Oper. Stochastic Equations*, **5** (1997) 325–336.

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

[18] A. Khorunzhy and L. Pastur, On the eigenvalue distribution of the deformed Wigner ensemble of random matrices, In: "Spectral operator theory and related topics", 97–127, *Adv. Soviet Math.*, **19**, Ed. V.A. Marchenko, Amer. Math. Soc., Providence, RI, 1994.

[19] A. Khorunzhy and G.J. Rodgers, On the Wigner law in dilute random matrices, *Rep. Math. Phys.*, **43** (1998) 297–319.

[20] H. Kunz, *Matrices aléatoires en physique*, Presses Polytechniques et Universitaires Romandes, Lausanne, 1998.

[21] V. Marchenko and L. Pastur, Eigenvalue distribution of certain ensembles of random matrices. *Mat. Sbornik* **72** 507–536 (1967) (English translation: *Math. USSR Sb.*, **1** (1967) 457–483).

[22] M.L. Mehta, *Random Matrices*, Academic Press, New York, 1991.

[23] L. Pastur, On the spectrum of random matrices, *Teor. Mat. Fiz.*, **10** (1973) 102–111 (in russian).

[24] L. Pastur, Random Matrices as Paradigm, *Mathematical Physics 2000*, Eds. A. Fokas, A. Grigoryan, T. Kibble and B. Zegarlinski, Imp. Coll. Press, London, 2000, 216–265.

[25] L. Pastur, A. Khorunzhy and V. Vasilchuk, On asymptotic property of spectra of sum of random independent one-dimensional operators, *Reports of Acad. Sci. Ukraine*, N2 (1994) 25–27 (in russian).

[26] Ya.G. Sinaĭ and A.B. Soshnikov, A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices, *Funktsional. Anal. i Prilozhen.*, **32**, no. 2 (1998) 56–79 (English translation: *Func. Anal. Appl.*, **32**, no. 2 (1988) 114–131).

[27] A. Soshnikov, Universality at the edge of the spectrum in Wigner random matrices, *Comm. Math. Phys.*, **207** (1999) 697–733.

[28] C.A. Tracy and H. Widom, Introduction to random matrices, in "Geometric and quantum aspects of integrable systems", *Lecture Notes in Physics*, **424**, Springer, Berlin, 1993, 103–130.

[29] P. van Moerbeke, Integrable Lattices: Random Matrices and Random Permutations, "Random matrices and their applications", Eds. P. Bleher and A. Its, MSRI Publications **40**, Cambridge, 2000, 1–98.

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

[31] A. Zvonkin, Matrix integrals and map enumeration: an accessible introduction, *Math. Comput. Modelling*, **26** (1997) 281–304.