Largest Eigenvalue of Rank one Perturbation of Gaussian ensemble

Kevin Kögler

Rotation project in the Erdös group
February 2021

Supervisor: Prof. László Erdös

Chapter 1

Introduction

The purpose of this work is to present a proof of a special case of Theorem 2.1 in [1] (see Theorem 1 below). Our proof will follow closely the proof of Theorem 2.1 in [1] but we try to present it in a more accessible way. Theorem 2.1 in [1] characterizes the high dimensional limit of the extremal eigenvalues of some random matrix X which is perturbed by some deterministic matrix A. It states that in the high dimensional limit the matrix $M = X + A$ has eigenvalues outside the spectrum of X iff A is "big enough". Further it gives the explicit values of these eigenvalues outside the spectrum of X.

For simplicity we will focus on the Gaussian case namely X being a Gaussian unitary ensemble (GUE) in the complex case and a Gaussian orthogonal ensemble (GOE) in the real case instead of a general Wigner matrix. Further we will also restrict ourselves to the case of A being a rank one projection instead of a finite rank projection. Our convention for the definition of a GUE (resp. GOE) is the following. A GUE (resp. GOE) is a collection of Hermitian matrices $W = W(N) \in \mathbb{C}^{N\times N}$(resp. $W \in \mathbb{R}^{N\times N}$) such that W_{ii} for $1 \leq i \leq N$ and $\sqrt{2}\mathfrak{Re}(W_{ij})$, $\sqrt{2}\mathfrak{Im}(W_{ij})$ for $1 \leq i < j \leq N$ (resp. $\frac{1}{\sqrt{2}}W_{ii}$ for $1 \leq i \leq N$ and W_{ij} for $1 \leq i < j \leq N$) have centered Gaussian distribution with variance σ^2 and are all independent. We write the rank one projection as $A = \theta aa^*$ with $a = a(N) \in \mathbb{C}^N$, $\|a\| = 1$ (resp. $a \in \mathbb{R}^N$) and $\theta \in \mathbb{R}$. For ease of notation we will also write $X = \frac{1}{\sqrt{N}}W$.

Our goal is to prove the following.

Theorem 1 (Theorem 2.1 in [1])**.** *Let $M = M(N) = X + A$ and denote by λ_k the k-th largest eigenvalue. Then for any joint realization of $(M(N))_{N\in\mathbb{N}}$ the following statements hold almost surely*

$$\lim_{N\to\infty} \lambda_1(M) = \begin{cases} 2\sigma & \text{if } \theta \leq \sigma \\ \theta + \frac{\sigma^2}{\theta} & \text{if } \theta > \sigma \end{cases} \tag{1.1}$$

and for any fixed $k > 1$

$$\lim_{N\to\infty} \lambda_k(M) = 2\sigma. \tag{1.2}$$

Similarly

$$\lim_{N\to\infty} \lambda_N(M) = \begin{cases} 2\sigma & \text{if } \theta \geq -\sigma \\ \theta + \frac{\sigma^2}{\theta} & \text{if } \theta < -\sigma \end{cases} \tag{1.3}$$

and for any fixed $k > 1$

$$\lim_{N\to\infty} \lambda_{N-k}(M) = -2\sigma. \tag{1.4}$$

We want to briefly explain some intuition about the result. The semicircle law states that in the Wigner case ($A = 0$) the empirical distribution of the eigenvalues of M converges almost surely to the semicircle distribution μ_{sc} given by

$$\mu_{sc} \sim \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\mathbb{1}_{[-2\sigma,2\sigma]}(x).$$

A pedagogical reference is [2]. It can be further shown (Theorem 2.12 in [6]) that in the Wigner case for any fixed $k \in \mathbb{N}$ we have that almost surely the extremal eigenvalues $\lambda_1(M), ..., \lambda_k(M)$ and $\lambda_N(M), ..., \lambda_{N-k}(M)$ converge to 2σ and -2σ respectively as $N \to \infty$. Thus the result of Theorem 1 can be understood in the following way. If $|\theta| \leq \sigma$ then in the high dimensional limit the extremal eigenvalues of M are the same as in the case $A = 0$. If $|\theta| > \sigma$ then we get exactly one eigenvalue outside the support of the semicircle distribution. It is also useful to note that $\left|\theta + \frac{\sigma^2}{\theta}\right| > 2\sigma$ iff $|\theta| > \sigma$.

Our main tool in proving Theorem 1 will be the Stieltjes transform. For $z \in \mathbb{C} \setminus \mathbb{R}$ we denote the resolvent of M by

$$G(z) = (zI - M)^{-1}$$

and define

$$g(z) = \mathbb{E}[\operatorname{tr} G(z))],$$

where $\operatorname{tr} = \frac{1}{N}\operatorname{Tr}$ is the normalized trace.

We also denote by

$$g_\sigma(z) = \mathbb{E}[(z - s)^{-1}] = \int_{\mathbb{R}} \frac{1}{z - x}\, d\mu_{sc}(x) \tag{1.5}$$

the Stieltjes transform of a random variable s with semicircle distribution.

We will use C to denote a generic constant the value of which may change from line to line. Similarly we let P denote a generic polynomial that may change from line to line. We will also say that a quantity $\Delta(N, z), z \in \mathbb{C} \setminus \mathbb{R}$ is $O_z(N^{-p})$ for some $p \in \mathbb{N}$ if for some C, l

$$|\Delta(z)| \leq C\frac{1}{N^p}(1 + |z|)^l(1 + |\mathfrak{Im}(z)|^{-1})^l.$$

To make notation simpler we will be writing the RHS as $\frac{1}{N^p}(1+|z|)^l P(|\mathfrak{Im}(z)|^{-1})$. In our convention always sure convergence refers to a probability space where all $(M_N)_{N\in\mathbb{N}}$ are realized jointly. Since we will be deriving the almost sure convergence through the Borel-Cantelli Lemma, the convergence results holds for any joint realization, so we do not specify an explicit realization.

This work is structured as follows. In chapter 2 we state some properties of the resolvent G and the Stieltjes transform of the semicircle law g_σ. We also introduce two important tools, namely multivariate versions of the Poincaré inequality and Stein's Lemma. Our goal in chapter 3 is to obtain the "Master equation"

$$g(z) = g_\sigma(z) + \frac{1}{N}L_\sigma(z) + O_z(\frac{1}{N^2})$$

with an explicit L_σ given in (3.15). This gives us an explicit expression for the deviation of g from g_σ. In chapter 4 we can then use this to show that almost surely

$$\overline{\lim_{N\to\infty}} \operatorname{Spect}(M) \subseteq K_\sigma = [-2\sigma, 2\sigma] \cup \{\theta + \frac{\sigma^2}{\theta}\}. \tag{1.6}$$

In chapter 5 we deduce Theorem 1 from (1.6) by first considering the case of small σ and then obtaining the general case by a rescaling argument.

Chapter 2

A few tools

We have the following useful properties of $G(z)$ that are immediate to check.

Lemma 2. *Let $G = (zI - M)^{-1}$ be the resolvent of a Hermitian (resp. symmetric) matrix M and $z \in \mathbb{C} \setminus \mathbb{R}$. Then*

1. $$\|G(z)\| \le |\mathfrak{Im}(z)|^{-1}, \tag{2.1}$$
   where $\|\cdot\|$ denotes the operator norm. In particular it holds that $|G_{ij}(z)| \le |\mathfrak{Im}(z)|^{-1}$.

2. *We have*
   $$\frac{1}{N}\sum_{i,j=1}^{N} |(GG)_{ij}|^2 = \frac{1}{N}\operatorname{Tr}(G^*G^*GG) \le |\mathfrak{Im}(z)|^{-4}. \tag{2.2}$$

3. *The derivative of G with respect to M is given by*
   $$G'B = GBG \tag{2.3}$$
   for any matrix B.

4. *For any $z \in \mathbb{C}$ such that $|z| > \|M\|$*
   $$\|G\| \le \frac{1}{|z| - \|M\|}. \tag{2.4}$$

We will also need the following properties of $g_\sigma(z)$. All of the statements apart from (2.6) can be easily deduced from (2.6) and the definition of the Stieltjes transform. One way to obtain (2.6) is to explicitly compute $g_\sigma(z)$ for $z \in \mathbb{C} \setminus \mathbb{R}$ using the explicit definition in (1.5).

Lemma 3. *The following holds*

1. $$g_\sigma \quad \text{is analytic on} \quad \mathbb{C} \setminus [-2\sigma, 2\sigma]. \tag{2.5}$$

2. *For all z such that $\mathfrak{Im}(z) \neq 0$*

   (a) $$\sigma^2 g_\sigma^2(z) - z g_\sigma(z) + 1 = 0, \tag{2.6}$$

   (b) $$|g_\sigma(z)| \le |\mathfrak{Im}(z)|^{-1}, \tag{2.7}$$

   (c) $$\left|g_\sigma(z)^{-1}\right| \le |z| + \sigma^2 |\mathfrak{Im}(z)|^{-1}, \tag{2.8}$$

   (d) $$|g'_\sigma(z)| = \left|\int \frac{1}{(z-x)^2}\, d\mu_{sc}(x)\right| \le |\mathfrak{Im} z|^{-2}, \tag{2.9}$$

(e)
$$\mathfrak{Im}(g_\sigma(z))\mathfrak{Im}(z) < 0, \tag{2.10}$$

(f)
$$\left|\frac{1}{ag_\sigma - z + \theta}\right| \leq |\mathfrak{Im} z|^{-1} \text{ for all } a > 0, \theta \in \mathbb{R}. \tag{2.11}$$

3. For all z such that $|z| > 2\sigma$

(a)
$$|g_\sigma(z)| \leq \frac{1}{|z| - 2\sigma}, \tag{2.12}$$

(b)
$$|g'_\sigma(z)| = \left|\int \frac{1}{(z-x)^2}\, d\mu_{sc}(x)\right| \leq \frac{1}{(|z| - 2\sigma)^2}, \tag{2.13}$$

(c)
$$\left|g_\sigma(z)^{-1}\right| \leq |z| + \frac{\sigma^2}{|z| - 2\sigma}. \tag{2.14}$$

The Gaussian measure satisfies the following Poincaré inequality. Let μ denote the Gaussian measure of the entries of W. Then for any $f \in H^1(\mathbb{R}, \mu) \cap C^1(\mathbb{R})$ we have

$$\mathrm{Var}(f) \leq C\sigma^2 \int |f'|^2\, d\mu, \tag{2.15}$$

where $\mathrm{Var}(f) = \mathrm{Var}(f(X))$ and $X \sim \mathcal{N}(0, \sigma^2)$. Since we do not need the dependence on σ in this inequality we will suppress it in the constant. The Poincaré inequality generalizes to

Lemma 4. *We have for any complex valued function f on $\mathbb{R}^{N^2}$ (resp. $\mathbb{R}^{\frac{N(N+1)}{2}}$) such that both f and ∇f are polynomially bounded*

$$\mathrm{Var}(f(M)) \leq \frac{C}{N}\,\mathbb{E}[\|\nabla f(M)\|_2^2],$$

where $\|M\|_2$ denotes the Frobenius norm of a matrix.

We refer to Theorem 3.20 in [4] for a proof of Lemma 4 (and (2.15)) in the case $A = 0$. It is straightforward to deduce Lemma 4 from Theorem 3.20 in [4] by applying Theorem 3.20 for $\tilde{f}$ defined by $f(M) = \tilde{f}(X)$.

We also have the following generalization of Stein's Lemma

Lemma 5. *Let Φ be a C^1 function on the space of Hermitian (resp. symmetric) matrices. Then for any deterministic Hermitian (resp. symmetric) matrix H*

$$\mathbb{E}[\Phi'(X) \cdot H] = \frac{N}{\sigma^2}\,\mathbb{E}[\Phi(X)\,\mathrm{Tr}(XH)]$$

as long as both sides are well defined.

Note that Lemma 5 is just a reformulation of the multivariate version of Stein's Lemma. The multivariate version is obtained by coordinate wise integration by parts as in the one dimensional version of Stein's Lemma.

Chapter 3

Master equation

The goal of this section is to establish Lemma 9 which states that

$$g(z) = g_\sigma(z) + \frac{1}{N} L_\sigma(z) + O_z(\frac{1}{N^2}) \tag{3.1}$$

for an explicit L_σ (defined in (3.15)). This will then allow us to obtain information on $\mathrm{Spect}(M)$ in the next chapter. We now briefly explain the strategy. The main idea is to first prove that $g(z)$ satisfies

$$\sigma^2 g(z) - zg(z) + 1 + \Delta(z) = O_z(\frac{1}{N^2}) \tag{3.2}$$

for some appropriate quantity $\Delta(z)$ that is $O_z(N^{-1})$. Then we use that g_σ satisfies an equation similar to (3.2) (namely (2.6)) to obtain an asymptotic estimate of the form

$$g(z) = g_\sigma(z) + \tilde{\Delta}(z) + O_z(\frac{1}{N^p}) \quad \text{for } p \in \{1, 2\}. \tag{3.3}$$

We will use a two step bootstrap argument. In the first step we will prove a crude version of (3.2) in Lemma 6 that will allow us to obtain (3.3) for some $\tilde{\Delta}(z)$ and $p = 1$ in Lemma 7. We can then obtain a stronger version of (3.2) in Lemma 8, which then allows us to prove (3.1) in Lemma 9.

Lemma 6. *It holds that*

$$\left| \sigma^2 g^2(z) - zg(z) + 1 + \frac{1}{N} \mathbb{E}[\mathrm{Tr}(G(z)A)] \right| \leq \frac{P(|\mathfrak{Im}(z)|^{-1})}{N^2}. \tag{3.4}$$

Proof. Let $\{E_{ij}\}_{1 \leq i,j \leq N}$ be the canonical basis of the space of $N \times N$ matrices. Applying Lemma 5 to $\Phi(X) = G_{ij} = (zI - X - A)^{-1}_{ij}$ and $H = E_{ij}$ for any $1 \leq i, j \leq N$ we obtain

$$\mathbb{E}[G_{ii} G_{jj}] = \frac{N}{\sigma^2} \mathbb{E}[G_{ij} X_{ij}].$$

Taking the normalized sum $\frac{1}{N^2} \sum_{i,j}$ gives

$$\mathbb{E}[\mathrm{tr}(G)^2] = \frac{1}{\sigma^2} \mathbb{E}[\mathrm{tr}(GX)]. \tag{3.5}$$

Writing

$$GX = (zI - X - A)^{-1}(X + A - zI - A + zI) = -I - GA + zG,$$

(3.5) implies

$$\mathbb{E}[\mathrm{tr}(G)^2] + \frac{1}{\sigma^2}(1 + \mathbb{E}[\mathrm{tr}(GA)] - z\,\mathbb{E}[\mathrm{tr}(G)]) = 0.$$

To conclude (3.4) it now suffices to prove that

$$\mathrm{Var}(\mathrm{tr}(G)) \leq C\frac{|\mathfrak{Im}(z)|^{-4}}{N^2}. \tag{3.6}$$

Applying Lemma 4 with $f(\Phi(M)) = \mathrm{tr}(G)$ and using the identity (2.3) for $B = E_{ij}$ we get

$$\mathrm{Var}(\mathrm{tr}(G)) \leq \frac{C}{N^3}\,\mathbb{E}\left[\sum_{i,j} |(GG)_{ij}|^2\right] \leq \frac{C}{N^2}\,|\mathfrak{Im}(z)|^{-4}$$

so (3.6) holds.

□

Lemma 7. *For any $z \in \mathbb{C}$ with $\mathfrak{Im}(z) > 0$*

$$|g(z) - g_\sigma(z)| \leq (|z| + C)\frac{P(|\mathfrak{Im}(z)|^{-1})}{N}. \tag{3.7}$$

Proof. We first note that since A is rank one we have

$$|\mathbb{E}[\mathrm{Tr}(G(z)A)]| \leq \|G(z)\| \leq |\mathfrak{Im}(z)|^{-1},$$

so it follows from Lemma 6 that

$$\left|\sigma^2 g^2 - zg + 1\right| \leq \frac{P(|\mathfrak{Im}(z)|^{-1})}{N}. \tag{3.8}$$

We also define

$$\mathcal{O} = \left\{ z \in \mathbb{C} | \mathfrak{Im}(z) > 0, \frac{P_{\mathcal{O}}(|\mathfrak{Im}(z)|^{-1})}{N}(|z| + \sigma^2\,|\mathfrak{Im}(z)|^{-1})\,|\mathfrak{Im}(z)|^{-1} < \frac{1}{4} \right\},$$

where $P_{\mathcal{O}}$ chosen such that (3.4) and (3.8) hold for $P = P_{\mathcal{O}}$. It is easily seen that $\mathcal{O}$ is nonempty for large enough N. Now let $z \in \mathcal{O}$. Noticing that

$$\frac{P_{\mathcal{O}}(|\mathfrak{Im}(z)|^{-1})}{N} < \frac{1}{4},$$

it follows from (3.8) that

$$|g(z)|\left|\sigma^2 g(z) - z\right| \geq \frac{1}{2}$$

which, using (2.1), implies

$$\frac{1}{|g(z)|} \leq 2(|z| + \sigma^2\,|\mathfrak{Im}(z)|^{-1}). \tag{3.9}$$

Thus we can define

$$\Lambda(z) = \sigma^2 g(z) + \frac{1}{g(z)}.$$

Then from (3.8) and (3.9) it follows that

$$|\Lambda(z) - z| \leq \frac{P(|\mathfrak{Im}(z)|^{-1})}{N} 2(|z| + \sigma^2 |\mathfrak{Im}(z)|^{-1}). \tag{3.10}$$

Combining (3.10) with the fact that by construction of $\mathcal{O}$

$$\frac{P(|\mathfrak{Im}(z)|^{-1})}{N} 2(|z| + \sigma^2 |\mathfrak{Im}(z)|^{-1}) \leq \frac{|\mathfrak{Im}(z)|}{2}$$

we obtain

$$|\mathfrak{Im}(\Lambda(z)) - \mathfrak{Im}(z)| \leq |\Lambda(z) - z| \leq \frac{|\mathfrak{Im}(z)|}{2},$$

so

$$\mathfrak{Im}(\Lambda(z)) > \frac{\mathfrak{Im}(z)}{2} > 0. \tag{3.11}$$

Lastly we want to show that

$$g(z) = g_\sigma(\Lambda(z)). \tag{3.12}$$

Let $\mathcal{O}' = \{z \in \mathcal{O} | \, |\mathfrak{Im}(z)| > 2\sigma\}$, then for any $z \in \mathcal{O}'$ we have from (3.11) and (2.8) that $g_\sigma(\Lambda(z)) \neq 0$. Therefore we obtain from (2.6)

$$\sigma^2 g_\sigma(\Lambda(z)) + \frac{1}{g_\sigma(\Lambda(z))} = \Lambda(z) = \sigma^2 g(z) + \frac{1}{g(z)}.$$

Rearranging and multiplying both sides by $g_\sigma(\Lambda(z))g(z)$ we obtain

$$\sigma^2 g_\sigma(\Lambda(z))g(z) \left(g_\sigma(\Lambda(z)) - g(z)\right) = g_\sigma(\Lambda(z)) - g(z).$$

We now show that $\sigma^2 |g_\sigma(\Lambda(z))g(z)| < 1$. We have $|g(z)| \leq |\mathfrak{Im}(z)|^{-1} \leq \frac{1}{2\sigma}$ and due to (2.7), (3.11) also $|g_\sigma(\Lambda(z))| \leq |\mathfrak{Im}(\Lambda(z))|^{-1} < \frac{1}{\sigma}$ so (3.12) holds on $\mathcal{O}'$. Since $\mathcal{O}'$ is open and $\mathcal{O}$ is connected by analytic continuation (3.12) holds on all of $\mathcal{O}$.

Now (3.12)

$$\begin{aligned} |g(z) - g_\sigma(z)| &= \left|\mathbb{E}[(z-s)^{-1}(\Lambda(z)-s)^{-1}](\Lambda(z) - z)\right| \\ &\leq |\mathfrak{Im}(z)|^{-1} |\mathfrak{Im}(\Lambda(z))|^{-1} |\Lambda(z) - z| , \end{aligned}$$

where s is distributed according to the semicircle law with variance σ^2. Combining (3.10),(3.11) now shows that (3.7) holds true on $\mathcal{O}$.

It remains to consider the case $z \notin \mathcal{O}$. For such a z we have by definition of $\mathcal{O}$ that

$$|\mathfrak{Im}(z)|^{-1} \leq 4 \frac{P_{\mathcal{O}}(|\mathfrak{Im}(z)|^{-1})}{N} (|z| + \sigma^2 |\mathfrak{Im}(z)|^{-1}) |\mathfrak{Im}(z)|^{-2} . \tag{3.13}$$

Since $|g(z)|, |g_\sigma(z)| \leq |\mathfrak{Im}(z)|^{-1}$ due to (2.1) and (2.7) respectively we observe that (3.7) holds also for $z \notin \mathcal{O}$.

$\square$

Lemma 8. *For any $z \in \mathbb{C}$ with $\mathfrak{Im}(z) > 0$*

$$\left|\sigma^2 g(z) - z g(z) + 1 + \frac{1}{N} E_\sigma(z)\right| \leq (|z| + C) \frac{P(|\mathfrak{Im}(z)|^{-1})}{N^2},$$

where $E_\sigma(z) := \frac{\theta}{z - \sigma^2 g_\sigma(z) - \theta}$ and θ is the nonzero eigenvalue of A.

Proof. We argue similarly as in Lemma 6. Applying Lemma 5 with $\Phi = G_{il}$ and $H = E_{jl}$ we obtain

$$\mathbb{E}[G_{ij}G_{ll}] = \frac{N}{\sigma^2}\,\mathbb{E}[G_{il}X_{lj}],$$

so by taking the normalized sum over l we get

$$\frac{1}{N}\sum_{l}\mathbb{E}[G_{ij}G_{ll}] = \frac{1}{\sigma^2}\,\mathbb{E}[(GX)_{ij}],$$

which can be rearranged as

$$\sigma^2\,\mathbb{E}[G_{ij}\operatorname{tr}(G)] = \mathbb{E}[(GX)_{ij}].$$

Writing

$$GX = (zI - X - A)^{-1}(X + A - zI - A + zI) = -I - GA + zG,$$

we have for any i, j

$$h_{ij} := \sigma^2\,\mathbb{E}[G_{ij}\operatorname{tr}(G)] + \delta_{ij} - z\,\mathbb{E}[G_{ij}] + \mathbb{E}[(GA)_{ij}] = 0.$$

Now recalling that $A = \theta aa^*$ with $\|a\| = 1$ we define $\alpha = \sum_{i,j}\bar{a}_i a_j G_{ij}$. Noting that

$$\sum_{i,j}\bar{a}_i a_j (GA)_{ij} = \langle a, GAa\rangle = \theta\alpha,$$

we have

$$0 = \sum_{ij}\bar{a}_i a_j h_{ij} = \sigma^2\,\mathbb{E}[\alpha\operatorname{tr}(G)] + 1 + (\theta - z)\,\mathbb{E}[\alpha]. \tag{3.14}$$

By Jensen's inequality we have

$$|\mathbb{E}[\alpha(\operatorname{tr}(G) - g)]| \le \mathbb{E}[|\alpha(\operatorname{tr}(G) - g)|^2]]^{\frac{1}{2}} = O_z(\frac{1}{N}),$$

since α is bounded and $\operatorname{Var}(\operatorname{tr}(G)) = O_z(\frac{1}{N^2})$ due to (3.6). We can combine this with (3.14) to obtain

$$\mathbb{E}[\alpha](\sigma^2 g(z) + \theta - z) + 1 = O_z(\frac{1}{N}).$$

Now it follows from Lemma 7 that $\mathbb{E}[\alpha](\sigma^2 g_\sigma(z) + \theta - z) + 1 = O_z(N^{-1})$ so due to (2.11)

$$\operatorname{Tr}(GA) = \theta\,\mathbb{E}[\alpha] = \frac{\theta}{z - \sigma^2 g_\sigma(z) - \theta} + O_z(\frac{1}{N}).$$

Combining this with Lemma 7 finishes the proof. □

Lemma 9. *For any* $z \in \mathbb{C}\setminus\mathbb{R}$

$$\left|g_\sigma(z) - g(z) + \frac{1}{N}L_\sigma(z)\right| = O_z(\frac{1}{N^2}),$$

where

$$L_\sigma(z) = g_\sigma(z)^{-1}g'_\sigma(z)E_\sigma(z) = g_\sigma(z)^{-1}\,\mathbb{E}[(z-s)^{-2}]E_\sigma(z) \tag{3.15}$$

and s is distributed according to the semicircle law with variance σ^2.

Proof. We may assume that $\mathfrak{Im}(z) > 0$, since writing out the definitions of $g_\sigma(z)$, $g(z)$, $L_\sigma(z)$ we observe that $g_\sigma(\bar{z}) = \bar{g}_\sigma(z), g(\bar{z}) = \bar{g}(z), L_\sigma(\bar{z}) = \bar{L}_\sigma(z)$. If $z \in \mathcal{O}$

$$\begin{aligned}\left|g_\sigma(z) - g(z) + \frac{1}{N}L_\sigma(z)\right| &= \left|g_\sigma(z) - g_\sigma(\Lambda(z)) + \frac{1}{N}L_\sigma(z)\right| \\ &= \left|\mathbb{E}[(z-s)^{-1}(\Lambda(z)-s)^{-1}(\Lambda(z)-z) + g_\sigma(z)^{-1}(z-s)^{-2}E_\sigma(z)]\right| \\ &\leq \left|\mathbb{E}[(z-s)^{-1}(\Lambda(z)-s)^{-1}\left(\Lambda(z) - z + \frac{1}{N}g_\sigma^{-1}(z)E_\sigma(z)\right)\right| \\ &+ \mathbb{E}\left[\left|(z-s)^{-1}((z-s)^{-1} - (\Lambda(z)-s)^{-1})\right|\right]\left|\frac{1}{N}g_\sigma^{-1}(z)E_\sigma(z)\right| \\ &\leq 2\,|\mathfrak{Im}(z)|^{-2}\left|\Lambda(z) - z + \frac{1}{N}g_\sigma^{-1}(z)E_\sigma(z)\right| \qquad (3.16) \\ &+ \frac{P(|\mathfrak{Im}(z)|^{-1})}{N}\,|\Lambda(z) - z|\,(|z| + C). \qquad (3.17)\end{aligned}$$

Here we have used (2.7), (3.11), (2.8) and

$$|E_\sigma(z)| \leq P(|\mathfrak{Im}(z)|^{-1}), \tag{3.18}$$

which follows from (2.11). By (3.10) it follows that the term in (3.17) is $O_z(N^{-2})$. For the term in (3.16) we have

$$\begin{aligned}\left|\Lambda(z) - z + \frac{1}{N}g_\sigma^{-1}(z)E_\sigma(z)\right| &= \frac{1}{g(z)}\left(\sigma^2 g^2(z) - zg(z) + 1 + \frac{E_\sigma(z)}{N}\right) \\ &+ \frac{1}{N}\frac{E_\sigma(z)}{g(z)g_\sigma(z)}(g(z) - g_\sigma(z)).\end{aligned}$$

Using Lemma 8, (3.9), Lemma 7, (3.18) and (2.8) we obtain

$$\left|\Lambda(z) - z + \frac{1}{N}g_\sigma^{-1}(z)E_\sigma(z)\right| = O_z(N^{-2}).$$

It remains to consider the case $z \notin \mathcal{O}$. By definition of $\mathcal{O}$ we have that $z \notin \mathcal{O}$ implies

$$1 < (|z| + C)\frac{P(|\mathfrak{Im}(z)|^{-1})}{N},$$

so it is enough to show that

$$\left|g_\sigma(z) - g(z) + \frac{1}{N}L_\sigma(z)\right| = O_z(1).$$

But this holds since Lemma 7 implies that $|g_\sigma(z) - g(z)| = O_z(N^{-1})$ and (2.8) and (3.18) imply that $L_\sigma(z) = O_z(1)$.

$\square$

Chapter 4

The spectrum of M

Using the bounds from the last chapter we are now able to prove the following

Theorem 10. *Let $\rho_\sigma := \theta + \frac{\sigma^2}{\theta}$ where θ is the nonzero eigenvalue of A and*

$$K_\sigma := [-2\sigma, 2\sigma] \cup \{\rho_\sigma\}.$$

Then almost surely

$$\overline{\lim_{N\to\infty}} \mathit{Spect}(M) \subseteq K_\sigma.$$

This theorem will be the main ingredient for the proof of Theorem 1 in the next chapter. To prove Theorem 10 we will first show that $L_\sigma(z)$ is the Stieltjes transform of a distribution μ_σ, the support of which is contained in K_σ. Then Lemma 9 will give us a bound on the expected number of eigenvalues outside of K_σ. Finally a bound on the variance of this number will allow us to deduce Theorem 10.

For the first step will we use the following characterization which we do not prove.

Theorem 11 ([5])**.** *Let μ be a distribution (in the sense of generalized functions) on $\mathbb{R}$ with compact support K and denote by $l(z) = \mu(\frac{1}{z-x})$ its Stieltjes transform. Then l is analytic on $\mathbb{C}\setminus\mathbb{R}$ and can be extended to $\mathbb{C}\setminus K$. Moreover l satisfies*

i) $\lim_{|z|\to\infty} l(z) = 0$.

ii) There exists a compact set K and $n \in \mathbb{N}$, such that for any $z \in \mathbb{C}\setminus\mathbb{R}$

$$|l(z)| \leq C \max\{1, \mathit{dist}(z,K)^{-n}\}.$$

iii) For any test function $\varphi \in C_c^\infty(\mathbb{R},\mathbb{R})$

$$\mu(\varphi) = -\frac{1}{\pi} \lim_{y\to 0^+} \mathfrak{Im} \int_{\mathbb{R}} \varphi(x) l(x+iy)\, dx.$$

Conversely if $K \subset \mathbb{R}$ is compact and l is an analytic function on $\mathbb{C}\setminus K$ satisfying $i), ii)$, then l is the Stieltjes transform of a distribution μ on $\mathbb{R}$ with $\operatorname{supp}(\mu) \subset K$*. Moreover* $\operatorname{supp}\mu$ *is precisely the set of singular points of l.*

This allows us to prove

Lemma 12. *We have that $L_\sigma(z)$ defined in* (3.15) *is the Stieltjes transform of a distribution μ_σ, the support of which is contained in K_σ*

Proof. We will first show that L_σ satisfies $i)$ and $ii)$ in Theorem 11. Recall that

$$L_\sigma(z) = g_\sigma(z)^{-1} g'_\sigma(z) \frac{\theta}{z - \sigma^2 g_\sigma - \theta}.$$

It follows from (2.12), (2.13), (2.14) that L_σ satisfies $i)$. From (2.8), (2.9), (2.11) it follows that $ii)$ holds locally with $n = 4$. Since L_σ vanishes at ∞ it follows that $ii)$ holds (globally).

Due to Theorem 10 we now only need to show that L_σ has no singular points in K_σ^c. By (2.6) we can rewrite

$$L_\sigma(z) = g_\sigma(z)^{-1} g'_\sigma(z) \frac{\theta}{1/g_\sigma(z) - \theta}.$$

Using that for $x \in \mathbb{R} \setminus [-2\sigma, 2\sigma]$

$$g_\sigma(x) = \frac{x}{2\sigma^2}(1 - \sqrt{1 - 4\sigma^2/x^2}),$$

we see that

$$\frac{1}{g_\sigma(x)} - \theta = 0 \quad \iff \quad x = \rho_\sigma.$$

Thus the statement of the Lemma is immediate from the explicit form of $L_\sigma(x)$. $\square$

We will now show that

Lemma 13. *For any test function $\varphi \in C_c^\infty(\mathbb{R}, \mathbb{R})$ we have*

$$\mathbb{E}[\mathrm{tr}(\varphi(M))] = \int \varphi \, d\mu_{sc} - \frac{1}{N}\mu_\sigma(\varphi) + O(\frac{1}{N^2}). \tag{4.1}$$

Thus for any real valued smooth φ such that φ is constant outside of a compact set and $\operatorname{supp} \varphi \cap K_\sigma = \emptyset$ *there is a constant C_φ such that almost surely*

$$|\mathrm{tr}(\varphi(M))| \leq C_\varphi N^{-\frac{4}{3}} \quad \textit{for almost every} \, N. \tag{4.2}$$

Proof. Let $r(z) = g(z) - g_\sigma(z) + \frac{1}{N} L_\sigma(z)$. Since $g(z), -g_\sigma(z), \frac{1}{N} L_\sigma(z)$ are all the Stieltjes transform of some distribution ($g(z), -g_\sigma(z)$ by definition and $\frac{1}{N} L_\sigma(z)$ due to Lemma 12) , r is the Stieltjes transform of the sum of these distributions. Observing that the definitions of g, g_σ, L_σ imply $r(\overline{z}) = \overline{r(z)}$ we can write the inverse Stieltjes transform of r as

$$\mathbb{E}[\mathrm{tr}(\varphi(M))] - \int \varphi \, \mu_{sc} + \frac{1}{N}\mu_\sigma(\varphi) = -\frac{1}{\pi} \lim_{y \to 0^+} \mathfrak{Im} \int_{\mathbb{R}} \varphi(x) r(x + iy) \, dx.$$

We know from Lemma 9 that

$$|r(z)| \leq \frac{1}{N^2}(|z| + C)^\alpha P(|\mathfrak{Im}(z)|^{-1}) \tag{4.3}$$

for some $\alpha \in \mathbb{R}_+$. It can be shown that this implies

$$\limsup_{y \to 0^+} \left| \int_{\mathbb{R}} \varphi(x) r(x + iy) \, dx \right| \leq \frac{C}{N^2}.$$

We only sketch the proof here. Define for $p \in \mathbb{N}_+$

$$I_p(z) = \frac{1}{(p-1)!} \int_0^\infty r(z + t) t^{p-1} \exp(-t) \, dt.$$

It is easy to check that

$$I_1(z) - I_1'(z) = r(z)$$

and for $p \geq 2$

$$I_p(z) - I_p'(z) = I_{p-1}.$$

Let D denote the operator that takes the derivative of a function. Using these identities we can iteratively use integration by parts to obtain

$$\int_{\mathbb{R}} \varphi(x) r(x+iy)\, dx = \int_{\mathbb{R}} ((1+D)^p \varphi) I_p(x+iy)\, dx. \tag{4.4}$$

Then one can show that (4.3) implies

$$\lim_{r\to\infty} \int_{[r,r+ir]} r(z+\tilde{z})\tilde{z}^{p-1} \exp(-\tilde{z})\, d\tilde{z} = 0$$

so by Cauchy's integral Theorem along the contour $[0,r]\cup[r,r+ir]\cup[r+ir,0]$, where$[z_1,z_2]$ denotes the line segment going from z_1 to z_2

$$I_p(z) = \lim_{r\to\infty} \int_{[0,r+ir]} r(z+\tilde{z})\tilde{z}^{p-1} \exp(-\tilde{z})\, d\tilde{z}.$$

Recall that by our convention the bound in Lemma 9 is to be understood as

$$|r(z)| \leq (|z|+C)^l \frac{P(|\mathfrak{Im}(z)|^{-1})}{N^2}.$$

Now we can choose $p = k+1$ where k is the degree of P. It is then straightforward to check that $I_p(z)$ is bounded on any compact set, so (4.1) follows from (4.4).

To prove (4.2) below we will prove a bound on the variance of $\operatorname{tr}(\varphi(M))$ namely

$$\operatorname{Var}(\operatorname{tr}(\varphi(M)) = O(N^{-4}) \tag{4.5}$$

for any φ such that φ is constant outside of a compact set and $\operatorname{supp}\varphi \cap K_\sigma = \emptyset$. Assume that we have already proven this bound and let $Z_N = \operatorname{tr}(\varphi(M)), \Omega_N = \{Z_N > N^{-4/3}\}$. From (4.1) and (4.5) we obtain

$$\mathbb{E}[|Z_N|^2] = O(N^{-4}).$$

Now

$$P(\Omega_N) \leq \int_\Omega \left|N^{4/3} Z_N(\omega)\right|^2 dP(\omega) \leq N^{8/3}\, \mathbb{E}[|Z_N|^2] = O(N^{-4/3}),$$

so (4.2) follows from the Borel-Cantelli Lemma.

It remains to prove (4.5). We will deduce this from Lemma 4. As in Lemma 4 we will identify matrices with vectors in $\mathbb{R}^{N^2}$ (resp. $\mathbb{R}^{\frac{N(N+1)}{2}}$) and denote by $\|\cdot\|_2$ the Frobenius norm. Write $\varphi = c + \psi$ with $c \in \mathbb{R}$ and $\psi \in C_c^\infty$. For $g(M) := \operatorname{tr}(\varphi(M))$ we obtain from Lemma 4

$$\operatorname{Var}(g(M)) \leq \frac{C}{N} \mathbb{E}[\|\nabla g(M)\|_2^2].$$

We have for any Hermitian (resp. symmetric) matrix B

$$\operatorname{Tr}(\nabla\psi(M)\cdot B) = \operatorname{Tr}(\psi'(M)\cdot B).$$

This can be proven by showing this identity for polynomials and then extending it to all bounded functions by approximating them with polynomials. We have

$$\begin{aligned}
\|\nabla g(M)\|_2^2 &= \sup_{\|B\|_2=1} |\langle \nabla g(M), B\rangle|^2 \\
&= \sup_{\|B\|_2=1} \left| \frac{d}{dt}\bigg|_{t=0} g(M+tB) \right|^2 .
\end{aligned}$$

Since the trace is linear we get

$$\begin{aligned}
\left| \frac{d}{dt}\bigg|_{t=0} g(M+tB) \right|^2 &= |\mathrm{tr}(\nabla \psi(M) \cdot B)|^2 \\
&= \left|\mathrm{tr}(\psi'(M) \cdot B)\right|^2 \\
&= \frac{1}{N^2} \left|\mathrm{Tr}(\psi'(M) \cdot B)\right|^2 \\
&\leq \frac{1}{N^2} \mathrm{Tr}(\psi'(M) \cdot \psi'(M)) \,\mathrm{Tr}(B \cdot B) \\
&= \frac{1}{N} \mathrm{tr}((\psi')^2(M)).
\end{aligned}$$

Where we have used the Cauchy-Schwartz inequality for the trace and the fact that $\mathrm{Tr}(B \cdot B) = \|B\|_2^2 = 1$. Putting everything together we obtain

$$\mathrm{Var}(g(M)) \leq \frac{C}{N} \|\nabla g(M)\|_2^2 \leq \frac{C}{N^2} \mathrm{tr}((\psi')^2(M)),$$

so now (4.5) follows from (4.1) since ψ vanishes on K_σ. □

Now we can quickly deduce Theorem 10

Proof of Theorem 10. For $\varepsilon > 0$ let φ be a smooth function that vanishes on K_σ and is equal to 1 on the complement of $K_\sigma + (-\varepsilon, \varepsilon)$. Lemma 13 implies that almost surely $\lim_{N\to\infty} \mathrm{Tr}(\varphi(M)) = 0$. Since the number of eigenvalues outside of $K_\sigma + (-\varepsilon, \varepsilon)$ is an integer and bounded by $\mathrm{Tr}(\varphi(M))$ that number must be equal to 0 almost surely as $N \to \infty$. Since ε was arbitrary we are done. □

Chapter 5

Proof of Theorem 1

We are now able to deduce Theorem 1 from Theorem 10. The main tool that we will use is the min-max-principle :

Lemma 14. *Let H be a self-adjoint (resp. symmetric) matrix on $V = \mathbb{C}^N$ (resp.$V = \mathbb{R}^N$). Denote by λ_k the $k-th$ largest eigenvalue of H. Then*

$$\lambda_k = \min_{\substack{U \subset V \\ \dim(U)=k-1}} \max_{\substack{v \perp U \\ \|v\|=1}} \langle v, Hv \rangle. \tag{5.1}$$

Consequently for any self-adjoint(resp. symmetric) matrices $H, \tilde{H}$ we have

$$\lambda_{j+k-1}(H + \tilde{H}) \leq \lambda_j(H) + \lambda_k(\tilde{H}) \tag{5.2}$$

and for $j + k \geq N + 1$

$$\lambda_j(H) + \lambda_k(\tilde{H}) \leq \lambda_{j+k-N}(H + \tilde{H}). \tag{5.3}$$

We refer to Section 12.1 in [3] for a proof of (5.1). Statements (5.2) and (5.3) are easily deduced from (5.1).

The structure of the proof of Theorem 1 is the following. For $\sigma \geq \theta$ Theorem 1 follows directly from Lemma 14, the semicircle law and Theorem 10. For $\sigma < \theta$ again combining Lemma 14, the semicircle law and Theorem 10 will imply that Theorem 1 is true for small σ. We will then argue by contradiction that if Theorem 1 holds for some σ it is also holds for some slightly bigger σ, which then allows us to conclude. Here the semicircle law refers to the statement that in the Wigner case ($A = 0$) the empirical distribution of the eigenvalues of M converges almost surely to the semicircle distribution μ_{sc}. We are now ready for the

Proof of Theorem 1. As an immediate consequence of Lemma 14 we obtain (1.2) and (1.4) from the semicircle law, by setting $H = A$ and $\tilde{H} = W$ and using that A has rank one. Thus we only need to prove (1.1) and (1.3). By replacing A with $-A$ and using that W has the same distribution as $-W$ we may w.l.o.g assume that $\theta > 0$ (recall that $A = \theta a a^*$). In this case (1.3) is also an immediate consequence from Lemma 14 and the semicircle law. Thus we only need to prove (1.1). We first show that we may assume that $\theta > \sigma$. If $\theta \leq \sigma$ then from Theorem 10 it follows that almost surely

$$\limsup_{N\to\infty} \lambda_1(M) \leq 2\sigma.$$

But since due to (1.2) we always have

$$2\sigma = \lim_{N\to\infty} \lambda_2(M) \leq \lim_{N\to\infty} \lambda_1(M),$$

it follows that (1.1) is true for $\theta \leq \sigma$.

Thus we can assume that $\theta > \sigma$. Due to Theorem 10 we only need to show that almost surely

$$\liminf_{N\to\infty} \lambda_1(M) > 2\sigma, \tag{5.4}$$

as if we are given such a lower bound then Theorem 10 implies that almost surely

$$\lim_{N\to\infty} \lambda_1(M) = \rho_\sigma$$

i.e. (1.1) holds.

Let now θ be fixed and define Σ be the set of all $\sigma < \theta$ for which (5.4) fails (M, W, A still depend on this σ). We need to show that $\Sigma = \emptyset$. Assume that $\Sigma \neq \emptyset$, then it makes sense to define

$$\sigma_0 := \inf \Sigma.$$

Since Lemma 14 implies that (5.4) holds if σ is small enough we have $\sigma_0 > 0$. Now let $\sigma = \sigma_0 + \frac{\varepsilon}{2}$ for some (small) $\varepsilon > 0$ let $M_\varepsilon = M - \varepsilon W$. From the definition of σ_0 it follows that almost surely

$$\lim_{N\to\infty} \lambda_1(M_\varepsilon) = \rho_{\sigma-\varepsilon}.$$

Since $M = M_\varepsilon + \varepsilon W$ it follows from the min-max principle and the semicircle law that almost surely

$$\liminf_{N\to\infty} \lambda_1(M) \geq \rho_{\sigma-\varepsilon} - 2\sigma\varepsilon = \theta + \frac{\sigma^2}{\theta} + \varepsilon\left(\frac{-2+\varepsilon}{\theta} - 2\sigma\right).$$

Letting $\varepsilon \to 0$ we obtain

$$\liminf_{N\to\infty} \lambda_1(M) \geq \theta + \frac{\sigma^2}{\theta} > 2\sigma,$$

where we have used that by assumption $\theta > \sigma$. Thus $\sigma = \sigma_0 + \frac{\varepsilon}{2}$ satisfies (5.4). But this contradicts the definition of σ_0 so we must have $\Sigma = \emptyset$ which completes the proof.

$\square$

Bibliography

[1] Capitaine, Mireille and Donati-Martin, Catherine and Féral, Delphine, *The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations*, Ann. Probab., 37 (2009), pp. 1-47.

[2] Tao, Terence, *Topics in random matrix theory*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2012.

[3] Lieb, Elliot and Loss, Michael, *Analysis*, Second edition, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2001.

[4] Boucheron, Stéphane and Lugos, Gábor and Massart, Pascal, *Concentration inequalities. A nonasymptotic theory of independence*, Oxford University Press, Oxford, 2013.

[5] Tillman, Heinz-Günther, *Randverteilungen analytischer Funktionen und Distributionen*, Math. Z., 59 (1953), pp. 61-83.

[6] Bai, Zhidong, *Methodologies in spectral analysis of large-dimensional random matrices, a review.*, Statist. Sinica., 9 (1999), pp. 611-677.