Random Polynomials and the Kac-Rice Formula

Q: Let $F(\sigma)$ be a random function on a high-dimensional space $\sigma \in \Omega \subseteq \mathbb{R}^n$. σ is a critical point of F if $\nabla F(\sigma) = 0$: This can be a local minimum, local maximum, or saddle point of F. What is the "typical number" of critical points $\{\sigma \in \Omega : \nabla F(\sigma) = 0,\ F(\sigma) \in B\}$ with function value in some set B?

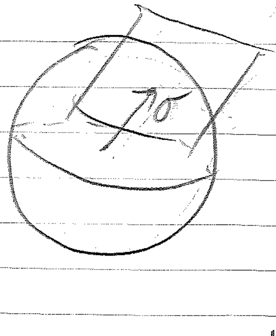
Consider a random homogeneous polynomial of degree p:

$$F(\sigma) = \frac{1}{\sqrt{n}} \sum_{i_1, \dots, i_p = 1}^{n} W_{i_1 \dots i_p} \sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_p}$$

where $\{W_{i_1 \dots i_p} : 1 \le i_1, \dots, i_p \le n\}$ are n^p iid $N(0,1)$ random variables. E.g. for $p=3$, W_{222} is the coefficient of σ_2^3. F is scale-invariant: $F(c\sigma) = c^p F(\sigma)$.

We'll view F as a function on the sphere $S^{n-1} = \{\sigma \in \mathbb{R}^n : \|\sigma\|_2^2 = 1\}$.

Then σ is a <u>critical point</u> of F on the sphere if $\nabla F(\sigma)$ is parallel to σ,



i.e. orthogonal to the hyperplane tangent to S^{n-1} at σ. We may write this as $(I - \sigma\sigma^T)\nabla F(\sigma) = 0$, where $I - \sigma\sigma^T$ is the orthogonal projection onto this hyperplane.

Example: Consider $p = 2$. Then

$$F(\sigma) = \frac{1}{\sqrt{n}} \sum_{i,j=1}^{n} W_{ij} \sigma_i \sigma_j .$$

Note that for each $i \ne j$, the term $\sigma_i \sigma_j$ appears twice. We can symmetrize and write

$$F(\sigma) = \frac{1}{\sqrt{n}} \cdot \sum_{i,j=1}^{n} \frac{W_{ij} + W_{ji}}{2} \sigma_i \sigma_j$$

$$= \frac{1}{\sqrt{2}} \cdot \sum_{i,j=1}^{n} \widetilde{W}_{ij} \sigma_i \sigma_j, \quad \text{where } \widetilde{W} \sim GOE(n).$$

$$\left(\mathrm{Var}[\widetilde{W}_{ij}] = \tfrac{1}{n} \text{ for } i \neq j, \ \tfrac{2}{n} \text{ for } i = j\right).$$

$$= \frac{1}{\sqrt{2}} \sigma^T \widetilde{W} \sigma.$$

The gradient is $\nabla F(\sigma) = \sqrt{2} \cdot \widetilde{W}\sigma$. This is parallel to σ iff σ is an eigenvector of $\widetilde{W}$, with eigenvalue $\sigma^T \widetilde{W} \sigma$.

$$\left((I - \sigma\sigma^T)\widetilde{W}\sigma = 0 \Leftrightarrow \widetilde{W}\sigma = (\sigma^T \widetilde{W}\sigma) \cdot \sigma.\right)$$

With probability 1, F has $2n$ critical points, given by $\pm v_1, \dots, \pm v_n$ for the eigenvectors of $\widetilde{W}$ corresponding to $\lambda_1 \geq \dots \geq \lambda_n$.

Here, $\pm v_1$ are local maxima, $\pm v_n$ are local minima, rest are saddles.

We've normalized F so that $\max_\sigma F(\sigma)$, $\min_\sigma F(\sigma)$ are typically of constant order.

$$|\{\sigma \text{ critical point}: F(\sigma) \in B\}| = 2 \cdot |\{\lambda_i(\widetilde{W}) \in \sqrt{2} \cdot B\}|$$

$$\approx 2n \int_{\sqrt{2}B} \rho_{sc}(x)\,dx, \quad \text{for large } n.$$

The case $p=2$ is special. What happens for $p \geq 3$?

Define $\quad \mathrm{Crit}_n(B) = |\{\sigma \in S^{n-1} : (I - \sigma\sigma^T)\nabla F(\sigma) = 0, \ F(\sigma) \in B\}|.$

Thm (Auffinger, Ben Arous, Černý '13): For any open interval $B \subseteq \mathbb{R}$,

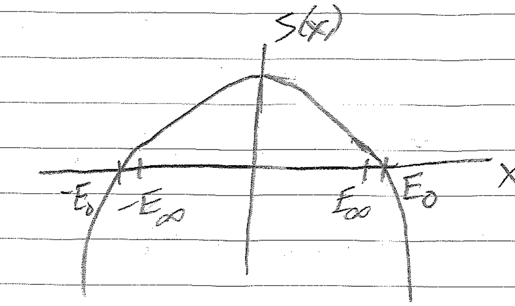
$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[\mathrm{Crit}_n(B)] = \sup_{x \in B} S(x), \quad \text{where}$$

$$S(x) = \frac{1}{2}\log(p-1) - \frac{p-2}{4(p-1)} x^2$$

$$- \mathbb{1}\{|x| > E_\infty\}\left(\frac{|x|}{E_\infty^2}\sqrt{x^2 - E_\infty^2} - \log\left(|x| + \sqrt{x^2 - E_\infty^2}\right) + \log E_\infty\right),$$

$$E_\infty = 2\sqrt{\frac{p-1}{p}}.$$

For $p=3$:



There is a value $E_0 \approx 1.657$ where $S(\pm E_0) = 0$. For $B \subset [-E_0, E_0]$, $S(x)$ is positive on B, and $\mathbb{E}[\mathrm{Crit}_n(B)]$ is exponentially large in n. Conversely, for $B \subseteq (-\infty, -E_0 - \varepsilon)$ or $B \subseteq (E_0 + \varepsilon, \infty)$, $S(x)$ is negative on B, and $\mathbb{E}[\mathrm{Crit}_n(B)]$ is exponentially small in n.

Corollary: For any $\varepsilon > 0$, there exists $c \equiv c(\varepsilon) > 0$ s.t. for all large n,

$$\mathbb{P}\left[\min_{\sigma \in S^{n-1}} F(\sigma) > -E_0 - \varepsilon, \ \max_{\sigma \in S^{n-1}} F(\sigma) < E_0 + \varepsilon\right] \geq 1 - e^{-cn}.$$

Proof: Let $B = (-\infty, -E_0 - \varepsilon)$. Then, letting $c = -\frac{1}{2}\sup_{x \in B} S(x)$, the theorem implies $\mathbb{E}[\mathrm{Crit}_n(B)] \leq e^{-cn}$ for all large n. Then

$$\mathbb{P}[\mathrm{Crit}_n(B) \geq 1] \leq \mathbb{E}[\mathrm{Crit}_n(B)] \leq e^{-cn}.$$ Similarly for $B = (E_0 + \varepsilon, \infty)$.

Warning: This does not show $\min_{\sigma \in S^{n-1}} F(\sigma) < -E_0 + \varepsilon$ whp. The <u>expected</u> number of critical points where $F(\sigma) \in [-E_0, -E_0 + \varepsilon]$ is exponentially large, but this might be due to an exceptional event of vanishing probability. For this particular model, this does not happen, and $\min_{\sigma \in S^{n-1}} F(\sigma) \to -E_0$ is correct (ABC '13, Subag '17). But this heuristic may be wrong in other models.

Proposition (Kac-Rice formula): Let ∇_s and ∇_s^2 be the gradient and Hessian on the sphere S^{n-1}. (Explicitly, letting $P_\sigma = I - \sigma\sigma^T$, these are defined as

$$\nabla_s F(\sigma) = P_\sigma \nabla F(\sigma)$$
$$\nabla_s^2 F(\sigma) = P_\sigma[\nabla^2 F(\sigma)]P_\sigma - (\sigma^T \nabla F(\sigma)) I$$

where ∇, ∇^2 are the usual gradient and Hessian in $\mathbb{R}^n$.) Let $p_{\nabla_s F(\sigma)}(0)$ be the density of $\nabla_s F(\sigma)$ (wrt. Lebesgue measure on $\mathbb{R}^{n-1}$) at the point 0, and $\det \nabla_s^2 F(\sigma)$ be the determinant as an $\mathbb{R}^{(n-1)\times(n-1)}$ matrix. Then

$$\mathbb{E}[\mathrm{Crit}_n(B)] = \int_{S^{n-1}} \mathbb{E}\left[|\det \nabla_s^2 F(\sigma)| \cdot \mathbb{1}\{F(\sigma)\in B\} \,\middle|\, \nabla_s F(\sigma) = 0\right] \cdot p_{\nabla_s F(\sigma)}(0)\, d\sigma.$$

Some preliminary observations:

① The model is rotationally invariant: For any fixed orthogonal matrix Q,
$$(F(\sigma) : \sigma \in S^{n-1}) \overset{L}{=} (F(Q\sigma) : \sigma\in S^{n-1}).$$
For $p=2$ this comes from the rotational invariance of GOE. The proof is the same for $p \geq 3$, as that for the GOE.

② The gradient and Hessian of F (in $\mathbb{R}^n$) are

$$(\nabla F(\sigma))_i = \frac{\partial}{\partial\sigma_i}\left[\frac{1}{\sqrt n}\sum_{i_1,\dots,i_p=1}^n W_{i_1\dots i_p}\sigma_{i_1}\cdots\sigma_{i_p}\right]$$
$$= \frac{1}{\sqrt n}\sum_{i_1,\dots,i_p=1}^n W_{i_1\dots i_p}\left(\frac{\partial}{\partial\sigma_i}\sigma_{i_1}\cdot\sigma_{i_2}\cdots\sigma_{i_p} + \sigma_{i_1}\frac{\partial}{\partial\sigma_i}\sigma_{i_2}\cdot\sigma_{i_3}\cdots\sigma_{i_p} + \dots + \sigma_{i_1}\cdots\sigma_{i_{p-1}}\frac{\partial}{\partial\sigma_i}\sigma_{i_p}\right)$$
$$= \frac{1}{\sqrt n}\sum_{i_1,\dots,i_{p-1}=1}^n \left(W_{i,i_1,\dots,i_{p-1}} + W_{i_1,i,i_2,\dots,i_{p-1}} + \dots + W_{i_1,\dots,i_{p-1},i}\right)\sigma_{i_1}\cdots\sigma_{i_{p-1}}$$

$$(\nabla^2 F(\sigma))_{ij} = \frac{1}{\sqrt n}\sum_{i_1,\dots,i_{p-2}=1}^n \Big(\underbrace{W_{i,j,i_1,\dots,i_{p-2}} + W_{i,i_1,j,i_2,\dots,i_{p-2}} + \dots + W_{i_1,\dots,i_{p-2},j,i}}_{\text{all } p(p-1) \text{ pairs of positions for } (i,j)}\Big)\cdot\sigma_{i_1}\cdots\sigma_{i_{p-2}}$$

③ By ①, the joint law of $(F(\sigma), \nabla_s F(\sigma), \nabla_s^2 F(\sigma))$ is the same at every $\sigma\in S^{n-1}$. Taking $\sigma = e_1$, and parametrizing $\nabla_s F(\sigma)$ and $\nabla_s^2 F(\sigma)$ by $e_2,\dots,e_n$, we get

$$F(e_1) = \frac{1}{\sqrt n} W_{1,\dots,1}$$
$$(\nabla_s F(e_1))_i = \frac{1}{\sqrt n}(W_{i,1,\dots,1} + W_{1,i,1,\dots,1} + \dots + W_{1,\dots,1,i}) \quad \text{for } i=2,\dots,n.$$
$$(\nabla_s^2 F(e_1))_{ij} = \frac{1}{\sqrt n}(W_{i,j,1,\dots,1} + \dots + W_{1,\dots,1,j,i}) - (e_1^T\nabla F(e_1))\mathbb{1}\{i=j\}$$
$$= \frac{1}{\sqrt n}(W_{i,j,1,\dots,1} + \dots + W_{1,\dots,1,j,i}) - \frac{p}{\sqrt n}W_{1,\dots,1}\mathbb{1}\{i=j\}. \quad \text{for } i,j=2,\dots,n.$$

In particular, $\nabla_s F(\sigma)$ and $\nabla_s^2 F(\sigma)$ are independent, for any fixed $\sigma\in S^{n-1}$.

Proof sketch of Kac-Rice: With probability 1, $\nabla_s^2 F(\sigma)$ is invertible at every critical point (where $\nabla_s F(\sigma)=0$), by the above. Then there is a (random, n-dependent) $\varepsilon>0$ s.t.

$$\Omega_\varepsilon \equiv \{\sigma\in S^{n-1} : \|\nabla_s F(\sigma)\| \leq \varepsilon,\ F(\sigma)\in B\}$$

is a union of $\mathrm{Crit}_n(B)$ disjoint sets $\Omega_\varepsilon^{(i)}$, each containing exactly one critical point σ_i where $\nabla_s F(\sigma) = 0$, $F(\sigma)\in B$. The map $\sigma \mapsto \nabla_s F(\sigma)$ is 1-to-1 on each $\Omega_\varepsilon^{(i)}$, with Jacobian $\nabla_s^2 F(\sigma)$. Thus for small ε,

$$\mathrm{Vol}(\varepsilon\text{-ball in } n-1 \text{ dimensions}) / \mathrm{Vol}(\Omega_\varepsilon^{(i)}) \approx |\det \nabla_s^2 F(\sigma)|$$

for all $\sigma\in\Omega_\varepsilon$. So

$$\mathrm{Crit}_n(B) = \sum_{i=1}^{\mathrm{Crit}_n(B)} \frac{1}{\mathrm{Vol}(\Omega_\varepsilon^{(i)})}\int_{\Omega_\varepsilon^{(i)}} d\sigma$$
$$= \lim_{\varepsilon\to 0}\sum_{i=1}^{\mathrm{Crit}_n(B)}\int_{\Omega_\varepsilon^{(i)}}\frac{|\det\nabla_s^2 F(\sigma)|}{\mathrm{Vol}(\varepsilon\text{-ball})}d\sigma = \int_{S^{n-1}}\mathbb{1}\{\sigma\in\Omega_\varepsilon\}\cdot\frac{|\det\nabla_s^2 F(\sigma)|}{\mathrm{Vol}(\varepsilon\text{-ball})}d\sigma.$$

Now take expectations:

$$\mathbb{E}[\mathrm{Crit}_n(B)] = \mathbb{E}\left[\lim_{\varepsilon\to 0}\int_{S^{n-1}} \mathbb{1}\{\sigma\in\Omega_\varepsilon\}\cdot\frac{|\det\nabla_s^2F(\sigma)|}{\mathrm{Vol}(\varepsilon\text{-ball})}\,d\sigma\right]$$

Requires checking some technical conditions $\longrightarrow$

$$= \int_{S^{n-1}}\left(\lim_{\varepsilon\to 0}\mathbb{E}\left[\mathbb{1}\{\sigma\in\Omega_\varepsilon\}\cdot|\det\nabla_s^2F(\sigma)|\right]/\mathrm{Vol}(\varepsilon\text{-ball})\right)d\sigma$$

$$= \int_{S^{n-1}}\left(\lim_{\varepsilon\to 0}\mathbb{E}\left[\mathbb{1}\{F(\sigma)\in B\}\cdot|\det\nabla_s^2F(\sigma)| \,\middle|\, \|\nabla_sF(\sigma)\|\le\varepsilon\right]\cdot\frac{\mathbb{P}[\|\nabla_sF(\sigma)\|\le\varepsilon]}{\mathrm{Vol}(\varepsilon\text{-ball})}\right)d\sigma$$

$$= \int_{S^{n-1}}\mathbb{E}\left[\mathbb{1}\{F(\sigma)\in B\}|\det\nabla_s^2F(\sigma)| \,\middle|\, \nabla_sF(\sigma)=0\right]\cdot p_{\nabla_sF(\sigma)}(0)\,d\sigma.$$

Remark: Argument is generic, and holds for any random smooth function on a manifold under some regularity conditions. See Adler, Taylor "Random Fields and Geometry" Chapters 11-12.

Proof of Thm: By Kac-Rice,

$$\mathbb{E}[\mathrm{Crit}_n(B)] = \int_{S^{n-1}}\mathbb{E}\left[|\det\nabla_s^2F(\sigma)|\cdot\mathbb{1}\{F(\sigma)\in B\} \,\middle|\, \nabla_sF(\sigma)=0\right]\cdot p_{\nabla_sF(\sigma)}(0)\,d\sigma.$$

Recall $(F(\sigma), \nabla_sF(\sigma), \nabla_s^2F(\sigma)) \overset{L}{=} (F(e_1), \nabla_sF(e_1), \nabla_s^2F(e_1))$ where

$$F(e_1) = \frac{1}{\sqrt{n}}W_{1,1,\dots,1} \sim N(0,\tfrac{1}{n})$$

$$\nabla_sF(e_1) = \frac{1}{\sqrt{n}}\left(W_{i,1,\dots,1}+\dots+W_{1,\dots,1,i}\right)_{i=2}^n \sim N(0,\tfrac{p}{n}I_{n-1})$$

$$\nabla_s^2F(e_1) = \underbrace{\frac{1}{\sqrt{n}}\left(W_{i,j,1,\dots,1}+\dots+W_{1,\dots,1,j,i}\right)_{i,j=2}^n}_{\sim\sqrt{p(p-1)}\sqrt{\frac{n-1}{n}}\,GOE(n-1)} - \frac{p}{\sqrt{n}}W_{1,\dots,1}\cdot I_{n-1}$$

Thus: $(F(\sigma), \nabla_sF(\sigma), \nabla_s^2F(\sigma)) \overset{L}{=} \left(\xi, \sqrt{p}\,g, \sqrt{\frac{p(p-1)(n-1)}{n}}\,G - p\xi I_{n-1}\right)$

where $\xi\sim N(0,\frac{1}{n})$, $g\sim N(0,\frac{1}{n}I)$, and $G\sim GOE(n-1)$ are independent.

Then $\mathbb{E}[\mathrm{Crit}_n(B)] = \int_{S^{n-1}}\mathbb{E}\left[\left|\det\left(\sqrt{\frac{p(p-1)(n-1)}{n}}G - p\xi I\right)\right|\cdot\mathbb{1}\{\xi\in B\}\right]\left(\frac{n}{2\pi p}\right)^{\frac{n-1}{2}}d\sigma$

$$= \mathrm{Vol}(S^{n-1})\cdot\left(\frac{n}{2\pi p}\right)^{\frac{n-1}{2}}\cdot\left(\frac{p(p-1)(n-1)}{n}\right)^{\frac{n-1}{2}}\cdot\mathbb{E}\left[\left|\det\left(G-\sqrt{\tfrac{pn}{(p-1)(n-1)}}\,\xi I\right)\right|\cdot\mathbb{1}\{\xi\in B\}\right].$$

We have $\mathbb{E}\left[|\det(G-t\xi I)|\cdot\mathbb{1}\{\xi\in B\}\right]$ (set $t=\sqrt{\frac{pn}{(p-1)(n-1)}}$)

$$= \int_B \mathbb{E}\left[|\det(G-txI)|\right]\sqrt{\frac{n}{2\pi}}\cdot e^{-\frac{nx^2}{2}}\,dx.$$

Let μ_n be the ESD of $GOE(n)$. Then

$$|\det(G-txI)| \overset{L}{=} \exp\left((n-1)\int\log|\lambda-tx|\,d\mu_{n-1}(\lambda)\right).$$

Let $L(y)=\int\log|\lambda-y|\,d\mu(y)$ where μ is the semicircle law on $[-2,2]$.

Can compute:

$$L(y) = \frac{y^2}{4}-\frac{1}{2}-\mathbb{1}\{|y|>2\}\left(\frac{|y|\sqrt{y^2-4}}{4}-\log\left(\frac{|y|}{2}+\sqrt{\frac{y^2}{4}-1}\right)\right).$$

Lemma: If $G\sim GOE(n)$, then $\mathbb{E}[|\det(G-yI)|]=e^{nL(y)+o(n)}$ uniformly over bounded sets of y.

Assuming this Lemma, for bounded B,

$$\mathbb{E}\left[|\det(G-t\xi I)|\cdot\mathbb{1}\{\xi\in B\}\right] = \int_B e^{n\overbrace{L\left(\sqrt{\frac{pn}{(p-1)(n-1)}}x\right)}^{L(\sqrt{\frac{p}{p-1}}x)+o(1)}-\frac{nx^2}{2}+o(n)}\,dx$$

$$= \exp\left(n\cdot\sup_{x\in B}\left(L\left(\sqrt{\tfrac{p}{p-1}}x\right)-\frac{x^2}{2}\right)+o(n)\right)$$

by Laplace's method for the integral. Since $L(\sqrt{\frac{p}{p-1}}x)-\frac{x^2}{2}$ decreases quadratically in $|x|$, this holds also for unbounded B.

Thus:

$$\frac{1}{n}\log \mathbb{E}[Crit_n(B)] = \frac{1}{n}\log\left[Vol(S^{n-1})\left(\frac{(p-1)(n-1)}{2\pi}\right)^{\frac{n-1}{2}}\right]$$
$$+ \sup_{x\in B}\left(L\left(\sqrt{\tfrac{p}{p-1}}x\right) - \frac{x^2}{2}\right) + o(1).$$

Using the formula for the surface volume $Vol(S^{n-1})$ and Stirling's approx,
$$Vol(S^{n-1}) = \frac{n\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} \sim \left(\frac{2\pi e}{n}\right)^{\frac{n}{2}} \cdot \sqrt{\frac{n}{\pi}},$$ this gives
$$\frac{1}{n}\log \mathbb{E}[Crit_n(B)] \to \sup_{x\in B}\Big(\underbrace{L\left(\sqrt{\tfrac{p}{p-1}}x\right) - \frac{x^2}{2} + \frac{1}{2} + \frac{1}{2}\log(p-1)}_{\text{This is exactly } S(x).}\Big)$$

Proof ideas of Lemma:

① Fix $\varepsilon > 0$, and let $l_\varepsilon(x)$ be a smooth approximation to $\log|y-x|$:
$l_\varepsilon(x) = \log|y-x|$ if $1/\varepsilon \geq |y-x| \geq \varepsilon$, and l_ε and derivatives of all orders are bounded by $C \equiv C(\varepsilon) > 0$.

From the local law and Helffer-Sjöstrand calculus (Lecture 5),
$$|m_n(z) - m(z)| \prec \frac{1}{n\eta}$$
and hence $\left|\frac{1}{n}\sum_{i=1}^n l_\varepsilon(\lambda_i(G)) - \int l_\varepsilon(\lambda)d\mu(\lambda)\right| \prec \frac{1}{n}.$

This means: $\mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^n l_\varepsilon(\lambda_i(G)) - \int l_\varepsilon(\lambda)d\mu(\lambda)\right| > n^{-1+\delta}\right] < n^{-D}$, for all $n \geq n_0(\delta, D, \varepsilon)$.

By concentration of spectral measure (Lecture 2),
$$\mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^n l_\varepsilon(\lambda_i(G)) - \mathbb{E}\frac{1}{n}\sum_{i=1}^n l_\varepsilon(\lambda_i(G))\right| \geq t\right] \leq e^{-Ct^2n^2}.$$

Together, this means
$$\left|\int l_\varepsilon(\lambda)d\mu(\lambda) - \mathbb{E}\frac{1}{n}\sum_{i=1}^n l_\varepsilon(\lambda_i(G))\right| < n^{-1+\delta}$$
(deterministically) for all $n \geq n_0(\delta,\varepsilon)$ and hence
$$\mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^n l_\varepsilon(\lambda_i(G)) - \int l_\varepsilon(\lambda)d\mu(\lambda)\right| > n^{-1+\delta} + t\right] \leq e^{-Ct^2n^2}$$
for all $n \geq n_0(\delta,\varepsilon)$ and all $t \geq 0$. Thus
$$\mathbb{P}\left[e^{\sum_{i=1}^n l_\varepsilon(\lambda_i(G))} \geq e^{n\int l_\varepsilon(\lambda)d\mu(\lambda) + n^\delta + nt}\right] \leq e^{-Ct^2n^2}.$$

Can integrate this over t to get
$$\mathbb{E}\left[e^{\sum_{i=1}^n l_\varepsilon(\lambda_i(G))}\right] \leq e^{n\int l_\varepsilon(\lambda)d\mu(\lambda) + o(n)}$$ ← may depend on ε.

Similarly, $\mathbb{E}\left[e^{\sum_{i=1}^n l_\varepsilon(\lambda_i(G))}\right] \geq e^{n\int l_\varepsilon(\lambda)d\mu(\lambda) - o(n)}$.

② $$\frac{\mathbb{E}[|\det(G-yI)|]}{\mathbb{E}\left[e^{\sum_{i=1}^n l_\varepsilon(\lambda_i(G))}\right]} \in \left[e^{-c\varepsilon n}, e^{c\varepsilon n}\right]$$ for all large n. (*)

Here is some intuition for why:

Note that $$\frac{|\det(G-yI)|}{e^{\sum_{i=1}^n l_\varepsilon(\lambda_i(G))}} = \overbrace{\prod_{i:|\lambda_i(G)-y|>1/\varepsilon} \frac{|\lambda_i(G)-y|}{1/\varepsilon}}^{\mathrm{I}} \cdot \overbrace{\prod_{i:|\lambda_i(G)-y|<\varepsilon} \frac{|\lambda_i(G)-y|}{\varepsilon}}^{\mathrm{II}}.$$

For each i, $\mathbb{P}[|\lambda_i(G)-y| > t] \leq e^{-ctn}$ for t large enough.

Then $\mathbb{E}\left[\max\left(\frac{|\lambda_i(G)-y|}{1/\varepsilon}, 1\right)\right] \approx 1 + \frac{1}{\varepsilon}\cdot e^{-\frac{c}{\varepsilon}n}$, so we expect $\mathbb{E}[\mathrm{I}] \approx 1$.

For II, the i^{th} closest eigenvalue to y is at expected distance $\gtrsim \frac{ci}{n}$ from y. Letting $m \lesssim C'\varepsilon n$ be the number of eigenvalues within distance ε,

$$\mathbb{E}[\mathrm{II}] \gtrsim \prod_{i=1}^{m} \frac{ci}{n\varepsilon} = \left(\frac{c}{n\varepsilon}\right)^m \cdot m! = e^{m\log m - m + m\log c - m\log n\varepsilon + o(m)}$$

$$\gtrsim e^{m\log C' + m\log c - m + o(m)}$$

$$\gtrsim e^{-C''\varepsilon n + o(n)}.$$

Assuming $(*)$ is true, we get for any fixed $\varepsilon > 0$ and a constant $C > 0$ that

$$\lim_{n\to\infty} \left| \frac{1}{n} \log \mathbb{E}[|\det(G - yI)|] - \int \ell_\varepsilon(\lambda)\, d\mu(\lambda) \right| \leq C\varepsilon.$$

Taking now $\varepsilon \to 0$, we have $\int \ell_\varepsilon(\lambda)\, d\mu(\lambda) \to \int \log|\lambda - y|\, d\mu(\lambda)$, so

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}[|\det(G - yI)|] = \int \log|\lambda - y|\, d\mu(\lambda),$$ as desired.