

Lectures on Random Matrices (Spring 2025)

Lecture 6: Double contour integral kernel. Steepest descent and local statistics

Leonid Petrov

Wednesday, February 19, 2025*

Contents

1	Recap: Determinantal structure of the GUE	2
2	Double Contour Integral Representation for the GUE Kernel	3
2.1	One contour integral representation for Hermite polynomials	3
2.2	Another contour integral representation for Hermite polynomials	4
2.3	Normalization of Hermite polynomials	5
2.4	Double contour integral representation for the GUE kernel	6
2.5	Conjugation of the kernel	7
2.6	Extensions	8
3	Steepest descent — generalities for single integrals	8
3.1	Setup	8
3.2	Saddle points and steepest descent paths	9
3.3	Local asymptotic evaluation near a saddle point	10
4	Steepest descent for the GUE kernel	10
4.1	Scaling	10
4.2	Critical points	11
4.3	Imaginary critical points: $ X < 2$, “bulk”	12
4.4	Real critical points: $ X > 2$, “large deviations”	13
4.5	Double critical point: $ X = 2$, “edge”	14
4.6	Airy kernel, Tracy–Widom distribution, and convergence of the maximal eigenvalue	16
F	Problems (due 2025-03-19)	16
F.1	Different global positions	16
F.2	Sine kernel	16
F.3	General bulk case	17

*Course webpage • Live simulations • TeX Source • Updated at 13:02, Wednesday 12th February, 2025

F.4	Discrete sine process	17
F.5	Large deviations	17
F.6	Airy kernel	18

1 Recap: Determinantal structure of the GUE

Last time, we proved the following result:

Theorem 1.1. *The GUE correlation functions are given by*

$$\rho_k(x_1, \dots, x_k) = \det \left[K_n(x_i, x_j) \right]_{i,j=1}^k,$$

with the correlation kernel

$$K_n(x, y) = \sum_{j=0}^{n-1} \psi_j(x) \psi_j(y).$$

Here

$$\psi_j(x) = \frac{1}{\sqrt{h_j}} p_j(x) e^{-x^2/4},$$

where $p_j(x)$ are the monic Hermite polynomials, and h_j are the normalization constants so that $\psi_j(x)$ are orthonormal in $L^2(\mathbb{R})$.

For this theorem, we need Cauchy–Binet summation formula and Andreief identity (which is essentially the same as Cauchy–Binet, but when summation is replaced by integration). Having these, we can write

$$\begin{aligned} \rho_k(x_1, \dots, x_k) &= \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} p(x_1, \dots, x_n) dx_{k+1} \cdots dx_n \\ &= \frac{1}{(n-k)! \widehat{Z}_{n,2}} \sum_{\substack{\sigma, \tau \in S_n \\ \sigma(k+1)=\tau(k+1), \dots, \sigma(n)=\tau(n)}} \text{sgn}(\sigma) \text{sgn}(\tau) \prod_{i=1}^k \psi_{\sigma(i)-1}(x_i) \psi_{\tau(i)-1}(x_i) \\ &= \text{const}_n \sum_{I \subseteq [n], |I|=k} \sum_{\sigma', \tau' \in S(I)} \text{sgn}(\sigma') \text{sgn}(\tau') \prod_{i=1}^k \psi_{\sigma'(i)-1}(x_i) \psi_{\tau'(i)-1}(x_i) \\ &= \text{const}_n \sum_{I \subseteq [n], |I|=k} \det [\psi_{i_\alpha}(x_j)]_{\alpha,j=1}^k \det [\psi_{i_\alpha}(x_j)]_{\alpha,j=1}^k, \end{aligned}$$

where $I = \{i_1, \dots, i_k\}$ is a subset of $[n]$ of size k , and $S(I)$ is the set of permutations of I . The last sum of products of two determinants is written by the Cauchy–Binet formula as

$$\text{const}_n \cdot \det \left[\sum_{j=0}^{n-1} \psi_j(x_\alpha) \psi_j(x_\beta) \right]_{\alpha,\beta=1}^k,$$

and finally the constant is equal to 1 by Andreief identity.

2 Double Contour Integral Representation for the GUE Kernel

2.1 One contour integral representation for Hermite polynomials

Recall that the GUE kernel is defined by

$$K_N(x, y) = \sum_{n=0}^{N-1} \psi_n(x) \psi_n(y),$$

with the orthonormal functions

$$\psi_n(x) = \frac{1}{\sqrt{h_n}} p_n(x) e^{-x^2/4},$$

where the (monic, probabilists') Hermite polynomials are given by

$$p_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}. \quad (2.1)$$

Note that the monic Hermite polynomials are uniquely defined by the orthogonality property. We are not proving (2.1) here, it is an exercise.

Lemma 2.1 (Generator function for Hermite polynomials). *We have*

$$\exp\left(xt - \frac{t^2}{2}\right) = \sum_{n \geq 0} p_n(x) \frac{t^n}{n!}.$$

The series converges for all t since the left-hand side is an entire function of t .

Proof. Write the generating function as

$$\sum_{n \geq 0} p_n(x) \frac{t^n}{n!} = \sum_{n \geq 0} \frac{(-1)^n t^n}{n!} e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}.$$

Since the factor $e^{x^2/2}$ does not depend on n , we can factor it out:

$$\sum_{n \geq 0} p_n(x) \frac{t^n}{n!} = e^{x^2/2} \sum_{n \geq 0} \frac{(-t)^n}{n!} \frac{d^n}{dx^n} e^{-x^2/2}.$$

Now, recall Taylor's theorem: for any holomorphic function f we have

$$f(x - t) = \sum_{n \geq 0} \frac{(-t)^n}{n!} f^{(n)}(x).$$

Applying this with $f(x) = e^{-x^2/2}$, we deduce that

$$\sum_{n \geq 0} \frac{(-t)^n}{n!} \frac{d^n}{dx^n} e^{-x^2/2} = e^{-(x-t)^2/2}.$$

Thus, our generating function becomes

$$\sum_{n \geq 0} p_n(x) \frac{t^n}{n!} = e^{x^2/2} e^{-(x-t)^2/2},$$

as desired. □

By Cauchy's integral formula we can write using Lemma 2.1:

$$p_n(x) = \frac{n!}{2\pi i} \oint_C \frac{\exp\left(xt - \frac{t^2}{2}\right)}{t^{n+1}} dt, \quad (2.2)$$

where the contour C is a simple closed curve encircling the origin. Indeed, here we use the complex analysis property

$$\frac{1}{2\pi i} \oint_C \frac{1}{z^{k+1}} dz = \begin{cases} 1, & \text{if } k = 0, \\ 0, & \text{if } k \neq 0, \end{cases}$$

so (2.2) is simply a complex analysis version of the operation of extracting the coefficient of t^n in the Taylor expansion.

Therefore,

$$\psi_n(x) = \frac{1}{\sqrt{h_n}} p_n(x) e^{-x^2/4} = \frac{e^{-x^2/4}}{\sqrt{h_n}} \frac{n!}{2\pi i} \oint_C \frac{\exp\left(xt - \frac{t^2}{2}\right)}{t^{n+1}} dt.$$

2.2 Another contour integral representation for Hermite polynomials

We start with the Fourier transform identity

$$\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2} + i t x\right) dt = \sqrt{2\pi} e^{-x^2/2}.$$

Differentiating both sides n times with respect to x yields

$$\frac{d^n}{dx^n} \left(e^{-x^2/2} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (i t)^n e^{-t^2/2 + i t x} dt.$$

Recalling the definition

$$p_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} \left(e^{-x^2/2} \right),$$

we obtain

$$p_n(x) = \frac{(-1)^n e^{x^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (i t)^n e^{-t^2/2 + i t x} dt.$$

Next, perform the change of variable

$$s = i t, \quad \text{so that} \quad t = -i s, \quad dt = -i ds.$$

Under this substitution the factors transform as follows:

$$(i t)^n = s^n,$$

and the exponent becomes

$$-\frac{t^2}{2} + i t x = -\frac{(-i s)^2}{2} + i (-i s) x = \frac{s^2}{2} + s x.$$

Thus, the integral transforms into

$$\int_{-\infty}^{\infty} (it)^n e^{-t^2/2+itx} dt = -i \int_{-i\infty}^{i\infty} s^n e^{s^2/2+sx} ds.$$

Substituting back we have

$$p_n(x) = \frac{(-1)^n e^{x^2/2}}{\sqrt{2\pi}} (-i) \int_{-i\infty}^{i\infty} s^n e^{s^2/2+sx} ds.$$

That is,

$$p_n(x) = \frac{i(-1)^{n+1} e^{x^2/2}}{\sqrt{2\pi}} \int_{-i\infty}^{i\infty} s^n e^{s^2/2+sx} ds.$$

Finally, change the sign of s , and we get:

$$p_n(x) = \frac{i e^{x^2/2}}{\sqrt{2\pi}} \int_{-i\infty}^{i\infty} s^n e^{s^2/2-sx} ds.$$

Therefore,

$$\psi_n(x) = \frac{i e^{x^2/4}}{\sqrt{2\pi} h_n} \int_{-i\infty}^{i\infty} s^n e^{s^2/2-sx} ds.$$

2.3 Normalization of Hermite polynomials

Lemma 2.2. *We have*

$$h_n = \int_{-\infty}^{\infty} p_n(x)^2 e^{-x^2/2} dx = n! \sqrt{2\pi}.$$

Proof. Multiply the generating function

$$\exp\left(xt - \frac{t^2}{2}\right) = \sum_{n \geq 0} p_n(x) \frac{t^n}{n!}$$

with a second copy (with parameter s):

$$\exp\left(xs - \frac{s^2}{2}\right) = \sum_{m \geq 0} p_m(x) \frac{s^m}{m!}.$$

Then,

$$\exp\left(xt - \frac{t^2}{2}\right) \exp\left(xs - \frac{s^2}{2}\right) = \sum_{n,m \geq 0} p_n(x) p_m(x) \frac{t^n s^m}{n! m!}.$$

Integrate both sides against $e^{-x^2/2} dx$. Using the orthogonality

$$\int_{-\infty}^{\infty} p_n(x) p_m(x) e^{-x^2/2} dx = h_n \delta_{nm},$$

the right-hand side becomes

$$\sum_{n \geq 0} \frac{h_n}{(n!)^2} (ts)^n.$$

On the left-hand side, we have

$$\int_{-\infty}^{\infty} e^{-x^2/2} \exp\left(x(t+s) - \frac{t^2 + s^2}{2}\right) dx.$$

Completing the square in x or recalling the standard Gaussian integral yields

$$\sqrt{2\pi} \exp\left(\frac{(t+s)^2 - (t^2 + s^2)}{2}\right) = \sqrt{2\pi} \exp(ts).$$

Thus, we obtain

$$\sqrt{2\pi} \exp(ts) = \sum_{n \geq 0} \frac{h_n}{(n!)^2} (ts)^n.$$

Expanding the left side as

$$\sqrt{2\pi} \sum_{n \geq 0} \frac{(ts)^n}{n!},$$

and comparing coefficients, we conclude that

$$\frac{h_n}{(n!)^2} = \frac{\sqrt{2\pi}}{n!} \implies h_n = n! \sqrt{2\pi}.$$

This completes the proof. □

2.4 Double contour integral representation for the GUE kernel

We can sum up the kernel (essentially, this is another proof of the Christoffel–Darboux formula):

$$\begin{aligned} K_n(x, y) &= \sum_{k=0}^{n-1} \psi_k(x) \psi_k(y) \\ &= \frac{e^{\frac{x^2 - y^2}{4}}}{(2\pi)^2} \oint_C dt \int_{-i\infty}^{i\infty} ds \exp\left\{-\frac{t^2}{2} + xt + \frac{s^2}{2} - ys\right\} \underbrace{\sum_{k=0}^{n-1} s^k t^{-k-1}}_{\frac{1 - (s/t)^n}{t-s}}. \end{aligned} \quad (2.3)$$

Here we used the two contour integral representations for Hermite polynomials, and the explicit norm (Lemma 2.2). At this point, the t contour is a small circle around 0, and the s contour is a vertical line in the complex plane. Their mutual position can be arbitrary at this point — the s contour goes along the imaginary line. Indeed, the fraction $\frac{1 - (s/t)^n}{t-s}$ does not have a singularity at $s = t$ due to the cancellation.

Let us now move the s contour to be to the left of the t contour, as in Figure 1. On the new contours, we have $|s| > |t|$. Now we can add the summands $s^k t^{-k-1}$ for all $k \leq -1$ into the sum

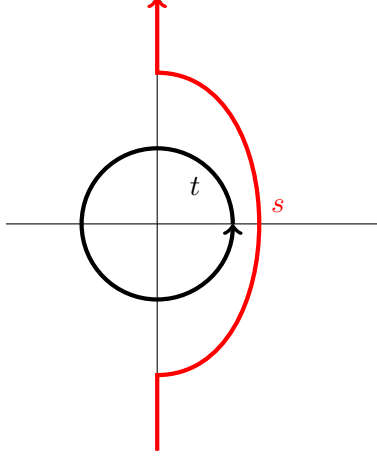


Figure 1: Integration contours for the GUE kernel (2.4).

in (2.3). Indeed, for $|s| > |t|$, the series in k converges, while the summand $s^k t^{-k-1}$ has zero residue at 0 and thus adding the summands does not change the value of the integral.

With this extension of the sum, formula (2.3) becomes

$$K_n(x, y) = \frac{e^{(y^2-x^2)/4}}{(2\pi)^2} \oint_C dt \int_{-i\infty}^{i\infty} ds \frac{\exp\left\{\frac{s^2}{2} - sy - \frac{t^2}{2} + tx\right\}}{s-t} \left(\frac{s}{t}\right)^n. \quad (2.4)$$

Remark 2.3. The s contour passes to the right of the t contour, but it might as well pass to the left of it. Indeed, one can deform the s contour to the left while picking the residue at $s = t$:

$$2\pi i \operatorname{Res}_{s=t} \frac{\exp\left\{\frac{s^2}{2} - sy - \frac{t^2}{2} + tx\right\}}{s-t} \left(\frac{s}{t}\right)^n = -e^{t(x-y)}.$$

This function is entire in t , and its integral over the t contour is zero. Therefore, there is no difference where the s contour passes with respect to the t contour.

2.5 Conjugation of the kernel

The kernel $K_n(x, y)$ contains a factor $e^{\frac{y^2-x^2}{4}} = g(x)/g(y)$, where $g(\cdot)$ is a nonvanishing function. This factor can be safely removed, since in all determinants $\det[K_n(x_i, x_j)]_{i,j=1}^k$ representing the correlation functions, the conjugation factors $g(x_i)/g(x_j)$ do not affect the value of the determinant. Thus, we can and will deal with the correlation kernel

$$K_n(x, y) = \frac{1}{(2\pi)^2} \oint_C dt \int_{-i\infty}^{i\infty} ds \frac{\exp\left\{\frac{s^2}{2} - sy - \frac{t^2}{2} + tx\right\}}{s-t} \left(\frac{s}{t}\right)^n, \quad (2.5)$$

and will use the same notation for it. Throughout the asymptotic analysis in Section 4 below, other conjugation factors may appear, but we can similarly remove them.

2.6 Extensions

Many other versions of the GUE / unitary invariant ensembles admit determinantal structure:

1. The GUE corners process [JN06]
2. The Dyson Brownian motion (e.g., add a GUE to a diagonal matrix) [NF98]
3. GUE corners plus a fixed matrix [FF14]
4. Corners invariant ensembles with fixed eigenvalues UDU^\dagger , where D is a fixed diagonal matrix and U is Haar distributed on the unitary group [Met13]

We will discuss the corners process structure in the next [Lecture 7](#).

3 Steepest descent — generalities for single integrals

3.1 Setup

In many problems arising in random matrix theory—as well as in asymptotic analysis more generally—it is necessary to evaluate integrals of the form

$$I(\Lambda) = \int_{\gamma} e^{\Lambda f(z)} \phi(z) dz, \quad (3.1)$$

where

- $\Lambda > 0$ is a large parameter,
- $f(z)$ and $\phi(z)$ are holomorphic functions in a neighborhood of the contour $\gamma \subset \mathbb{C}$,
- and the contour γ is chosen in such a way that the integral converges.

The *method of steepest descent* (also known as the *saddle point method*) provides a systematic procedure for obtaining the asymptotic behavior of $I(\Lambda)$ as $\Lambda \rightarrow +\infty$.

The key observation is that for large Λ , the exponential term $e^{\Lambda f(z)}$ is highly oscillatory or decaying, so that the main contributions to the integral come from small neighborhoods of points where the real part of $f(z)$ is maximal. Moreover, since we can deform the integration contour γ to pick points where $\operatorname{Re} f(z)$ is even bigger, it makes sense to find points *not only on the original contour* where $\operatorname{Re} f(z)$ is maximal. Such *critical* (or *saddle*) points are found from the equation with the complex derivative:

$$f'(z) = 0$$

Indeed, since $\operatorname{Re} f(z)$ is harmonic and $f(z)$ satisfies the Cauchy–Riemann equations, the condition $f'(z) = 0$ is equivalent to the condition that $\operatorname{Re} f(z)$ has zero gradient. Moreover, by harmonicity, all critical points of $\operatorname{Re} f(z)$ are saddle-like.

Once the saddle points are identified, one deforms the contour γ to Γ so that Γ passes through the saddle point(s) with the maximal value of $\operatorname{Re} f(z)$, and, moreover, such that on the rest of the new contour Γ the real part of $f(z)$ is strictly less than the value(s) at the saddle point(s). The decrease of $\operatorname{Re} f(z)$ along Γ may be ensured if one picks Γ to be *steepest descent* for $\operatorname{Re} f(z)$. By holomorphicity of $f(z)$, the steepest descent of Re is equivalent to the condition that the imaginary part of $f(z)$ is constant along Γ .

Remark 3.1. In practical applications, one does not need Γ to be fully steepest descent (it is usually hard to control). One can either choose Γ to be steepest descent in a neighborhood of the critical point and estimate the real part outside, or simply estimate the change of $\operatorname{Re} f(z)$ directly along a given contour.

Remark 3.2. The function $\phi(z)$ might not be holomorphic, and might have poles. The deformation of the contour from γ to Γ might pick residues at these poles. These residues can be harmless (easy to account for) or not (hard to account for; or affect the asymptotics of the integral), and one has to be careful with the contour deformation.

Despite the caveats in Remarks 3.1 and 3.2, in what follows in this section we will discuss the easiest case of steepest descent analysis. We also assume that there is only one saddle point z_0 to take care of.

3.2 Saddle points and steepest descent paths

Definition 3.3 (Saddle point). A point $z_0 \in \mathbb{C}$ is called a *saddle point* of $f(z)$ if

$$f'(z_0) = 0.$$

We shall assume in what follows that at every saddle point under consideration the second derivative satisfies

$$f''(z_0) \neq 0.$$

Definition 3.4 (Steepest descent path). Let z_0 be a saddle point of $f(z)$. A curve $\Gamma \subset \mathbb{C}$ passing through z_0 is called a *steepest descent path* for $f(z)$ if along Γ the imaginary part of $f(z)$ is constant (i.e., $\operatorname{Im}(f(z)) = \operatorname{Im}(f(z_0))$ for all $z \in \Gamma$), which implies that the real part $\operatorname{Re}(f(z))$ decreases away from z_0 .

In a neighborhood of a saddle point z_0 ,

$$z = z_0 + w, \quad f(z) = f(z_0) + \frac{1}{2}f''(z_0)w^2 + O(w^3).$$

If we denote

$$f''(z_0) = |f''(z_0)|e^{i\theta_0},$$

then writing $w = r e^{i\varphi}$, we obtain

$$f(z) = f(z_0) + \frac{1}{2}|f''(z_0)|r^2 e^{i(2\varphi+\theta_0)} + O(r^3).$$

For the imaginary part to remain constant in a neighborhood of z_0 , and, moreover, for the phase of the quadratic term to be π modulo 2π , one must choose φ so that

$$2\varphi + \theta_0 = \pi \pmod{2\pi}. \quad (3.2)$$

We need the phase π so that the exponent is negative, for the integral to converge.

There are two directions satisfying (3.2) through z_0 , and we use both of them for our contour Γ . Along these directions, one finds that

$$\operatorname{Re}(f(z)) = \operatorname{Re}(f(z_0)) - \frac{1}{2}|f''(z_0)|r^2 + O(r^3),$$

so that $\operatorname{Re}(f(z))$ is maximal at $z = z_0$ and decays quadratically as one moves away from z_0 along the steepest descent paths.

3.3 Local asymptotic evaluation near a saddle point

Assume now that the contour γ in (3.1) has been deformed so that it passes through a saddle point z_0 along a steepest descent path. In a small neighborhood of z_0 , we write

$$z = z_0 + w,$$

so the local contribution of a neighborhood of z_0 to the integral is

$$I_{z_0}(\Lambda) = e^{\Lambda f(z_0)} \phi(z_0) \int_{-\infty}^{\infty} e^{\Lambda \frac{1}{2} f''(z_0) w^2} dw \left(1 + O\left(\frac{1}{\Lambda}\right)\right). \quad (3.3)$$

Here the integration is taken along the steepest descent direction, so that the quadratic term in the exponent is real and negative. (That is, by the choice (3.2), we have $\operatorname{Re}(f''(z_0)w^2) = -|f''(z_0)|r^2$.) Then the Gaussian integral evaluates to

$$\int_{-\infty}^{\infty} e^{-\Lambda \frac{|f''(z_0)|}{2} r^2} dr = \sqrt{\frac{2\pi}{\Lambda |f''(z_0)|}}.$$

Hence, we arrive at the following fundamental result.

Theorem 3.5 (Local asymptotics via steepest descent). *Let z_0 be a saddle point of $f(z)$ with $f'(z_0) = 0$ and $f''(z_0) \neq 0$, and assume that $\phi(z)$ is holomorphic in a neighborhood of z_0 . Then, as $\Lambda \rightarrow +\infty$, the contribution of a small neighborhood of z_0 to the integral (3.1) is given by*

$$I_{z_0}(\Lambda) \sim e^{\Lambda f(z_0)} \phi(z_0) \sqrt{\frac{2\pi}{\Lambda |f''(z_0)|}}, \quad \Lambda \rightarrow +\infty. \quad (3.4)$$

Moreover, the behavior (3.4) captures the full asymptotic behavior of the integral (3.1) as long as on the new contour Γ , the real part of $f(z)$ is maximized at z_0 and is separated from $\operatorname{Re} f(z_0)$ everywhere else on Γ outside of a small neighborhood of z_0 .

Under appropriate assumptions (typically, if f and ϕ are holomorphic on a neighborhood that can be reached by the deformed contour and if the contributions away from the saddle points are exponentially small), one may show that the error in approximating the full integral by the sum of the local contributions is itself exponentially small relative to the leading order terms. In many cases, the next-order corrections can be computed by carrying the expansion in (3.3) to higher order in w . (See, e.g., [Olv74] for a systematic treatment.)

4 Steepest descent for the GUE kernel

4.1 Scaling

Let us now consider the GUE kernel (2.5),

$$K_n(x, y) = \frac{1}{(2\pi)^2} \oint_C dt \int_{-i\infty}^{i\infty} ds \frac{\exp\left\{\frac{s^2}{2} - sy - \frac{t^2}{2} + tx\right\}}{s - t} \left(\frac{s}{t}\right)^n,$$

where the integration contours are as in Figure 1.

We know from the Wigner semicircle law (established for real symmetric matrices with general iid entries in in [Lecture 2](#), and for the GUE in [Lecture 4](#)) that the eigenvalues live on the scare \sqrt{n} . This means that to capture the local asymptotics, we need to scale

$$x = X\sqrt{n} + \frac{\Delta x}{\sqrt{n}}, \quad y = Y\sqrt{n} + \frac{\Delta y}{\sqrt{n}}, \quad \Delta x, \Delta y \in \mathbb{R}. \quad (4.1)$$

Moreover, if $X \neq Y$ (i.e., different global positions), one can check that the kernel vanishes. In other words, the local behaviors at different global positions are independent. See Problem [F.1](#). In what follows, we take $Y = X$.

Let us also make a change of the integration variables:

$$t = z\sqrt{n}, \quad s = w\sqrt{n}.$$

The integration contours for z and w look the same as in Figure 1, up to a rescaling. However, as 0 and $t = s$ are the only singularities in the integrand, we can deform the z, w contours as we wish, while keeping $|z| < |w|$ and the general shape as in Figure 1.

We thus have:

$$\begin{aligned} & K_n(X\sqrt{n} + \Delta x/\sqrt{n}, X\sqrt{n} + \Delta y/\sqrt{n}) \\ &= \frac{\sqrt{n}}{(2\pi)^2} \oint_C dz \int_{-i\infty}^{i\infty} dw \frac{\exp \left\{ n \left(\log w - \log z + \frac{w^2}{2} - \frac{z^2}{2} + X(z - w) + \frac{z\Delta x - w\Delta y}{n} \right) \right\}}{w - z}. \end{aligned} \quad (4.2)$$

Remark 4.1. The logarithms in the exponent are harmless, since for the estimates we only need the real parts of the logarithms, and for the main contributions, we will have $z \approx w$, so any phases of the logarithms would cancel.

The asymptotic analysis of double contour integrals like (4.2) in the context of determinantal point processes was pioneered in [[Oko02](#), Section 3].

4.2 Critical points

Let us define

$$S(z) := \frac{z^2}{2} + \log z - Xz.$$

Then the exponent contains $n(S(w) - S(z))$. According to the steepest descent ideology, we should deform the integration contours to pass through the critical point(s) z_{cr} of $S(z)$. Moreover, the new w contour should maximize the real part of $S(z)$ at z_{cr} , and the new z contour should minimize it. If $S''(z_{cr}) \neq 0$, it is possible to locally choose such contours, they will be perpendicular to each other at z_{cr} .

Thus, we need to find the critical points of $S(z)$. They are found from the quadratic equation:

$$S'(z) = z + \frac{1}{z} - X = 0, \quad z_{cr} = \frac{X \pm \sqrt{X^2 - 4}}{2}. \quad (4.3)$$

Depending on whether $|X| < 2$, there are three cases. Unless $|X| = 2$, equation (4.3) has a single root, and thus $S''(z_{cr}) \neq 0$. We will consider the three cases in Sections [4.3](#) to [4.5](#) below.

4.3 Imaginary critical points: $|X| < 2$, “bulk”

When $|X| < 2$, the critical points are complex conjugate. Denote them by z_{cr} and $\overline{z_{cr}}$. Since $S(z)$ has real coefficients, we have

$$\operatorname{Re} S(z_{cr}) = \operatorname{Re} S(\overline{z_{cr}}).$$

Thus, we need to consider the contribution from both points. For simplicity of the computations, let us consider only the case $X = 0$. See Problem F.3. We have

$$z_{cr} = i, \quad S''(z_{cr}) = 2.$$

The behavior of $\operatorname{Re} S(z)$ on the complex plane can be illustrated by a 3D plot or by a region plot of the regions where $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr})$ has constant sign. See Figure 2 for an illustration in the case $X = \frac{1}{2}$. (We take $X \neq 0$ to break symmetry, for a better intuition.)

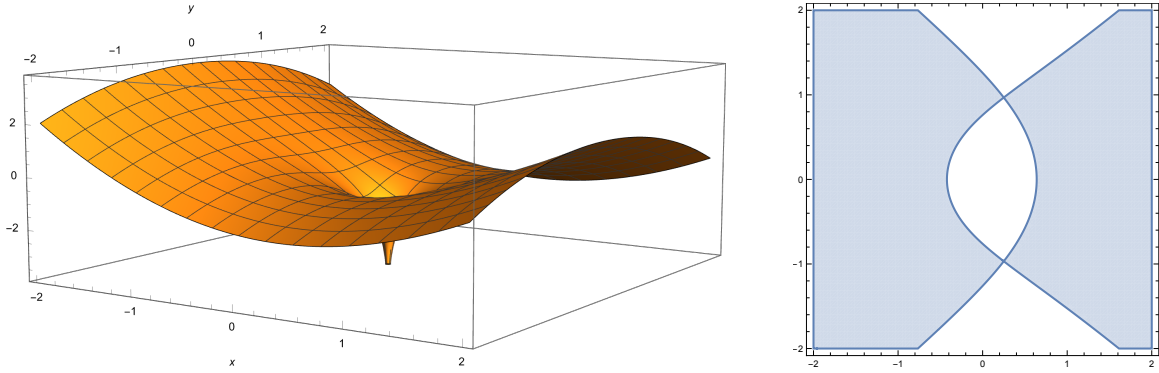


Figure 2: A 3D plot and a region plot of the regions where $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr})$ is positive (highlighted) or negative, in the case $X = \frac{1}{2}$. In this case, $z_{cr} \approx 0.25 + 0.96i$.

From the region plot, we see that the new z contour should pass through the shaded region $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr}) > 0$, and the new w contour should pass through the unshaded region $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr}) < 0$.

Deforming the contours from Figure 1 to the new contours is impossible without passing through the residue at $w = z$. Moreover, this residue appears only for certain values of z . Namely, for $X = 0$, let us first make the z contour to be the positively (counterclockwise) oriented unit circle. It passes through the critical points $z_{cr} = i$ and $\overline{z_{cr}} = -i$. Since the original w contour is to the right of the z contour, we only encounter the residue when z is in the right half of the circle.

Thus, we can write

$$\oint_{\text{old contours}} = \oint_{\text{new contours}} + \int_{-i}^i 2\pi i \operatorname{Res}_{w=z} dz, \quad (4.4)$$

where in the single integral, the z contour passes to the right of the origin, along the right half of the unit circle.

It remains to consider the two integrals in the right-hand side of (4.4). Recall that the correlation functions are defined relative to a reference measure, and the right object to scale is

$$K_n(x, y)dy = \frac{1}{\sqrt{n}} d(\Delta y).$$

The extra factor $n^{-1/2}$ compensates the prefactor \sqrt{n} in (4.2).

The single integral takes the form

$$\frac{-i}{2\pi} \int_{-i}^i e^{z(\Delta x - \Delta y)} dz = \frac{\sin(\Delta x - \Delta y)}{\pi(\Delta x - \Delta y)}, \quad \Delta x, \Delta y \in \mathbb{R}. \quad (4.5)$$

Definition 4.2. The *sine kernel* is defined as

$$K_{\text{sine}}(x, y) := \begin{cases} \frac{\sin(x - y)}{\pi(x - y)}, & x \neq 0, \\ \frac{1}{\pi}, & x = 0. \end{cases}$$

(The value at $x = y$ is defined by continuity.)

This kernel is translation invariant, and is often defined with a single argument, as $K_{\text{sine}}(x - y)$.

The double integral has both contours in the “steepest descent” regime, which means that the main contribution is

$$\text{const} \cdot \frac{e^{n(\text{Re } S(z_{cr}) - \text{Re } S(z_{cr}))}}{\sqrt{n}} \sim \frac{\text{const}}{\sqrt{n}}.$$

At this rate, the double integral over the new contours *does not* contribute to the asymptotics of the correlation functions. Recall that the correlation functions are expressed as finite-dimensional determinants of the kernel $K_n(x, y)$, and the error $O(n^{-1/2})$ is negligible in the limit $n \rightarrow +\infty$. This is because the main term comes from the single integral, which does not vanish.

We have established the following result:

Proposition 4.3 (Bulk asymptotics at $X = 0$). *The correlation kernel K_n of the GUE has the following asymptotics close to zero as $n \rightarrow +\infty$:*

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} K_n \left(\frac{\Delta x}{\sqrt{n}}, \frac{\Delta y}{\sqrt{n}} \right) = K_{\text{sine}}(\Delta x, \Delta y), \quad \Delta x, \Delta y \in \mathbb{R}.$$

Consequently, the eigenvalues of the GUE converge to the sine process determined by the sine kernel (Definition 4.2), in the sense of finite-dimensional distributions.

Remark 4.4. Beyond $X = 0$, the local correlations are essentially the same, up to rescaling of the real line by a constant factor (depending on the semicircle density). See Problem F.3.

4.4 Real critical points: $|X| > 2$, “large deviations”

For $X^2 > 4$, both solutions (4.3) are real. Let us assume $X > 2$, the case $X < -2$ is similar. For $X > 2$, both solutions are positive. Label these solutions as

$$z_+ = \frac{X + \sqrt{X^2 - 4}}{2}, \quad z_- = \frac{X - \sqrt{X^2 - 4}}{2}, \quad \text{so that} \quad z_+ z_- = 1.$$

A straightforward check reveals that $z_+ > 1$ and $z_- < 1$ (for $X > 2$). Note that $S''(z) = 1 - z^{-2}$, which is positive for $z_+ > 1$ and negative for $z_- < 1$. Thus, the critical points z_+ and z_- are a local minimum and a local maximum. A crucial observation is that

$$S(z_+) < S(z_-).$$

One can deform the z integration contour to pass through z_- and the w contour to pass through z_+ . Then, on these contours, one can show that

$$\operatorname{Re} S(w) - \operatorname{Re} S(z) < 0.$$

According to the steepest descent ideology, we see that the main exponential behavior of the double contour integral is

$$\exp \{n (\operatorname{Re} S(z_+) - \operatorname{Re} S(z_-))\} = O(e^{-\delta(X)n}), \quad |X| > 2. \quad (4.6)$$

Here $\delta(X) > 0$ for $|X| > 2$, and $\delta(X) \rightarrow 0$ when $|X| \rightarrow 2$.

The outcome (4.6) reflects the fact that the Wigner semicircle law places all eigenvalues inside the interval $|X| \leq 2$. The probability to see even a single eigenvalue outside $[-2, 2]$ is exponentially small.

This exponential decay corresponds to a large deviation regime. Indeed, if at least one of the diagonal entries of the matrix is unusually large, this corresponds to the maximal eigenvalue to get outside the interval $[-2, 2]$. See also Problem F.5.

4.5 Double critical point: $|X| = 2$, “edge”

Throughout the subsection, we assume that $X = 2$. The case $X = -2$ is symmetric.

When $X = 2$, the two solutions in (4.3) merge into a double critical point $z_{cr} = 1$. We have

$$S'(1) = 0, \quad S''(1) = 0, \quad S'''(1) = 2.$$

Thus, the usual quadratic approximation fails and one must expand to third order. Writing

$$z = 1 + u, \quad w = 1 + v,$$

with u, v small, we have

$$S(1 + u) = S(1) + \frac{S'''(1)}{6} u^3 + O(u^4) = S(1) + \frac{u^3}{3} + O(u^4),$$

and similarly for $S(1 + v)$. Hence, the difference in the exponents becomes

$$S(1 + v) - S(1 + u) = \frac{v^3 - u^3}{3} + O(u^4 + v^4).$$

To capture the correct asymptotics, we rescale the local variables by setting

$$u = \frac{U}{n^{1/3}}, \quad v = \frac{V}{n^{1/3}},$$

so that

$$n[S(1+v) - S(1+u)] = \frac{V^3 - U^3}{3} + O(n^{-1/3}).$$

Moreover, the correct edge scaling for the spatial variables is obtained by writing

$$x = 2\sqrt{n} + \frac{\xi}{n^{1/6}}, \quad y = 2\sqrt{n} + \frac{\eta}{n^{1/6}}, \quad \xi, \eta \in \mathbb{R}.$$

We have

$$n(S(w) - S(z)) = n^{1/3}(\xi - \eta) + \frac{V^3 - U^3}{3} + \xi U - \eta V + O(n^{-1/3}).$$

The terms $n^{1/3}(\xi - \eta)$ are harmless as they can be removed by conjugation.

The region plot of $\operatorname{Re} S(z) - \operatorname{Re} S(1)$ (shown in Figure 3) makes sure that we can deform the z contour so that it passes through $z_{cr} = 1$ as the new U contour at the angles $\pm \frac{2\pi}{3}$ (where $\operatorname{Re} U^3 > 0$), we can deform the w contour so that it passes through $z_{cr} = 1$ as the new V contour at the angles $\pm \frac{\pi}{3}$ (where $\operatorname{Re} V^3 < 0$). This will ensure the convergence of the new double integral.

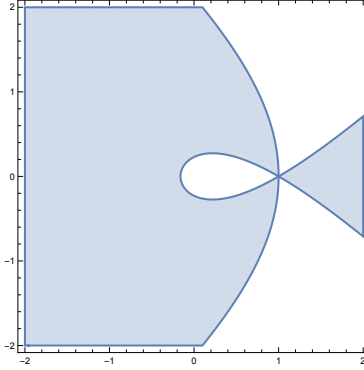


Figure 3: The plot of the region $\operatorname{Re} S(z) - \operatorname{Re} S(1) > 0$ for $X = 2$.

Thus, we have shown that under the rescaling, the GUE correlation kernel $K_n(x, y) dy$ converges to a new kernel.

Definition 4.5. Define the *Airy kernel* on \mathbb{R} by

$$K_{\text{Ai}}(\xi, \eta) = \frac{1}{(2\pi i)^2} \int_{e^{-\frac{\pi i}{3}} \infty}^{e^{\frac{\pi i}{3}} \infty} dV \int_{e^{-\frac{2\pi i}{3}} \infty}^{e^{\frac{2\pi i}{3}} \infty} dU \frac{\exp\left\{\frac{V^3 - U^3}{3} + U\xi - V\eta\right\}}{V - U}.$$

For another formula for the Airy kernel which does not involve integrals, see Problem F.6.

Proposition 4.6. *We have*

$$\lim_{n \rightarrow \infty} \frac{1}{n^{1/6}} K_n\left(2\sqrt{n} + \frac{\xi}{n^{1/6}}, 2\sqrt{n} + \frac{\eta}{n^{1/6}}\right) \rightarrow K_{\text{Ai}}(\xi, \eta).$$

Consequently, the eigenvalue statistics at the edge of the spectrum converge to the Airy point process, in the sense of fine-dimensional distributions.

4.6 Airy kernel, Tracy–Widom distribution, and convergence of the maximal eigenvalue

Let us make a few remarks on the asymptotic results of Propositions 4.3 and 4.6. First, a rigorous justification of convergence of contour integrals requires some estimates on the error terms in the steepest descent analysis, but these estimates are mild and not hard to obtain.

Second, the GUE has the maximal eigenvalue λ_{max} . It is reasonable to assume that the Airy process also (almost surely) admits a maximal point (usually denoted by \mathbf{a}_1), and that λ_{max} converges to \mathbf{a}_1 under appropriate rescaling:

$$\lim_{n \rightarrow \infty} n^{\frac{1}{6}} (\lambda_{max} - 2\sqrt{n}) = \mathbf{a}_1. \quad (4.7)$$

This is indeed the case, but to show (4.7), one needs to show the convergence in distribution:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(n^{1/6} (\lambda_{max} - 2\sqrt{n}) \leq x \right) \rightarrow \mathbb{P}(\mathbf{a}_1 \leq x). \quad (4.8)$$

Both events (4.8) are so-called *gap probabilities*, for example,

$$\mathbb{P}(\mathbf{a}_1 \leq x) = \mathbb{P}(\text{there are no eigenvalues in the interval } (x, \infty)),$$

which is expressed as the Fredholm determinant

$$\det(1 - K_{\text{Ai}})_{(x, \infty)} = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \int_x^{\infty} dy_1 \int_x^{\infty} dy_2 \cdots \int_x^{\infty} dy_m \det_{i,j=1}^m K_{\text{Ai}}(y_i, y_j). \quad (4.9)$$

Thus, to get (4.8), one needs to show the convergence of sums like this for the GUE kernel to the corresponding sums for the Airy kernel. This is doable, but tedious.

Moreover, to get convergence in distribution of random variables, one would also have to argue either *tightness*, or independently show that (4.9) defines a cumulative probability distribution function in x :

$$F_2(x) = \det(1 - K_{\text{Ai}})_{(x, \infty)}. \quad (4.10)$$

The distribution (4.10) is known as the *GUE Tracy–Widom distribution*. The subscript 2 indicates that $\beta = 2$. There are distributions F_β for all beta, most notably, the GOE and GSE distributions. The classical distributions F_1, F_2, F_4 also appear as fluctuation distributions in interacting particle systems, while other beta values do not quite appear in the particle systems domain.

More details may be found in the original papers [TW93], [For93], [TW94].

F Problems (due 2025-03-19)

F.1 Different global positions

Show that if in (4.1) we take $X \neq Y$, then $K_n(x, y)$ vanishes as $n \rightarrow +\infty$. Moreover, establish that the decay is exponential in n .

F.2 Sine kernel

Compute the integral (4.5).

F.3 General bulk case

Perform the asymptotic analysis of the correlation kernel as in Section 4.3, but in the general case $-2 < X < 2$.

F.4 Discrete sine process

Define the discrete sine kernel on \mathbb{Z} by

$$K_{\text{dsine}}(x, y) := \begin{cases} \frac{\sin \rho(x - y)}{\pi(x - y)}, & x \neq y, \\ \frac{\rho}{\pi}, & x = y, \end{cases}$$

where $\rho \in [0, 1]$ is the density parameter.

Let $\rho = 1/2$. Compute (numerically) the asymptotics of the two events under the discrete sine process:

$$\mathbb{P}\left(\underbrace{\circ \circ \dots \circ}_{n \text{ times}} \underbrace{\bullet \bullet \dots \bullet}_{n \text{ times}}\right), \quad \mathbb{P}\left(\underbrace{\circ \bullet \bullet \dots \circ \bullet}_{2n \text{ points}}\right),$$

If the sine process was of independent random points (with the same density $1/2$), both events would have the same probability 2^{-2n} . Which event is more favored by the sine process?

F.5 Large deviations

Let W_n be an $n \times n$ Wigner real or Hermitian matrix with finite variance entries. Assume that the matrix is normalized so that the variance of each diagonal entry is 1.

Assumption [BBP05]. *If a Wigner matrix is normalized to have diagonal variance 1, then a rank 1 perturbation of magnitude $c > 0$ is sufficient to shoot the maximum eigenvalue outside the support of the Wigner semicircle law. (For a simulation of this phenomenon, see [here](#).)*

Consider the following large deviation event. For a fixed $\eta > 0$, let

$$E_{n,\eta} := \left\{ \exists i \in \{1, \dots, n\} \text{ such that } W_{ii} \geq \eta \right\}.$$

Under the above assumption, if for some i the diagonal entry W_{ii} is unusually large, it will push the maximal eigenvalue of W_n outside the bulk.

1. Assuming that the entries are Gaussian, *lower bound* the probability of the event $E_{n,\eta}$ for large n .
2. Assuming another tail behavior of the diagonal entries (exponential or power-law tails), use the limit theorems for maxima of independent random variables to generalize the *lower bound* of $\mathbb{P}(E_{n,\eta})$.

F.6 Airy kernel

Define the Airy function by

$$Ai(\xi) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iU^3/3 + i\xi U} dU = \frac{1}{\pi} \int_0^{\infty} \cos\left(\frac{U^3}{3} + \xi U\right) dU.$$

This integral converges, but only conditionally. To improve convergence, one should instead integrate along a complex contour, from $e^{\frac{5\pi i}{6}}\infty$ to 0 to $e^{\frac{\pi i}{6}}\infty$.

Show that

$$K_{Ai}(\xi, \eta) = \frac{Ai(\xi) Ai'(\eta) - Ai(\eta) Ai'(\xi)}{\xi - \eta}.$$

Note that this expression is parallel to the sine kernel,

$$\frac{\sin(x-y)}{\pi(x-y)} = \frac{\sin x \cos y - \cos x \sin y}{\pi(x-y)}, \quad \cos x = (\sin x)'. \quad \cos x = (\sin x)'.$$

These correlation kernels are called *integrable* [IJKS90].

Hint for the problem: observe that

$$\exp\{-izx + iwy\} = \frac{i}{x-y} \left(\frac{\partial}{\partial z} + \frac{\partial}{\partial w} \right) \exp\{-izx + iwy\},$$

and use integration by parts in $K_{Ai}(\xi, \eta)$ from Definition 4.5.

References

- [BBP05] J. Baik, G. Ben Arous, and S. Péché, *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices*, Ann. Probab. **33** (2005), no. 5, 1643–1697. arXiv:math/0403022 [math.PR]. [↑17](#)
- [FF14] P. Ferrari and R. Frings, *Perturbed GUE minor process and Warren’s process with drifts*, J. Stat. Phys **154** (2014), no. 1-2, 356–377. arXiv:1212.5534 [math-ph]. [↑8](#)
- [For93] P. J. Forrester, *The spectrum edge of random matrix ensembles*, Nuclear Physics B **402** (1993), 709–728. [↑16](#)
- [IJKS90] A.R. Its, A.G. Izergin, V.E. Korepin, and N.A. Slavnov, *Differential equations for quantum correlation functions*, Int. J. Mod. Phys. B **4** (1990), no. 5, 1003–1037. [↑18](#)
- [JN06] K. Johansson and E. Nordenstam, *Eigenvalues of GUE minors*, Electron. J. Probab. **11** (2006), no. 50, 1342–1371. arXiv:math/0606760 [math.PR]. [↑8](#)
- [Met13] A. Metcalfe, *Universality properties of Gelfand-Tsetlin patterns*, Probab. Theory Relat. Fields **155** (2013), no. 1-2, 303–346. arXiv:1105.1272 [math.PR]. [↑8](#)
- [NF98] T. Nagao and P.J. Forrester, *Multilevel dynamical correlation functions for Dyson’s Brownian motion model of random matrices*, Physics Letters A **247** (1998), no. 1-2, 42–46. [↑8](#)
- [Oko02] A. Okounkov, *Symmetric functions and random partitions*, Symmetric functions 2001: Surveys of developments and perspectives, 2002. arXiv:math/0309074 [math.CO]. [↑11](#)
- [Olv74] F. W. J. Olver, *Asymptotics and Special Functions*, Academic Press, 1974. [↑10](#)
- [TW93] C. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Physics Letters B **305** (1993), no. 1, 115–118. arXiv:hep-th/9210074. [↑16](#)
- [TW94] C. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159** (1994), no. 1, 151–174. arXiv:hep-th/9211141. [↑16](#)

L. PETROV, UNIVERSITY OF VIRGINIA, DEPARTMENT OF MATHEMATICS, 141 CABELL DRIVE, KERCHOF HALL, P.O. BOX 400137, CHARLOTTESVILLE, VA 22904, USA
E-mail: lenia.petrov@gmail.com