On the Rate of Convergence in the Central Limit
Theorem for Linear Statistics of
Gaussian, Laguerre, and Jacobi Ensembles

Sergey Berezin  Alexander I. Bufetov

Abstract

Under the Kolmogorov-Smirnov metric, an upper bound for the rate of convergence
to the Gaussian law is obtained for linear statistics of matrix ensembles correspond-
ing to Gaussian, Laguerre, and Jacobi weights. The main lemma, obtained by an
analysis of the corresponding Riemann-Hilbert problem, is a uniform estimate for the
characteristic function of our linear statistic over a growing interval.

Contents

1 Introduction and the formulation of the main result 2
2 Analysis of the Riemann–Hilbert problem 6  
2.1 First transformation. Normalisation at $z = \infty$ 8  
2.2 Second transformation. Deformation of the contour 10  
2.3 Parametrix at $z = \infty$ 12  
2.4 Local parametrix at $z = 1$ 14  
2.4.1 The case of GUE and LUE. Soft edge 14  
2.4.2 The case of JUE. Hard edge 18  
2.5 Local parametrix at $z = -1$ 21  
2.5.1 The case of GUE. Soft edge 21  
2.5.2 The case of LUE and JUE. Hard edge 23  
2.6 Final transformation. Small norm problem 26  
3 Proof of Lemma 2 27  
3.1 Deformation of the weight 27  
3.2 Integration of the differential identity 28  
4 Proof of Lemma 1 33  
5 Proof of Theorem 1 35
1 Introduction and the formulation of the main result

The main result of this paper is an upper bound, under the Kolmogorov-Smirnov metric, for the rate of convergence in the Central Limit Theorem for linear statistics of unitary matrix ensembles that correspond to Gaussian (GUE), Laguerre (LUE), and Jacobi (JUE) weights. Consider the set of \( n \times n \) Hermitian matrices \( M = M^\dagger = \{M_{j,k}\}_{j,k=1}^n, M_{j,k} \in \mathbb{C} \) endowed with the probability measure

\[
\mathbb{P}_n(dM) = \frac{1}{Z_n} e^{-\text{Tr} Q_n(M)} dM, \quad dM = \prod_j dM_{j,j} \prod_{j<k} d(\text{Re} M_{j,k}) d(\text{Im} M_{j,k}),
\]

supported on the interval \( \mathcal{I} \), where

\[
Q_n(x) := nV(x) - \omega(x),
\]

the interval \( \mathcal{I} \), the potential \( V(x) \), and the additional term \( \omega(x) \) are given by

\[
\mathcal{I} = \begin{cases} 
(-\infty, +\infty), & \text{for GUE}, \\
[-1, +\infty), & \text{for LUE}, \\
[-1, 1], & \text{for JUE},
\end{cases} \quad
V(x) = \begin{cases} 
2x^2, & \text{for GUE}, \\
2(x + 1), & \text{for LUE}, \\
0, & \text{for JUE},
\end{cases} \quad
\omega(x) = \begin{cases} 
0, & \text{for GUE}, \\
\alpha \log (1 + x), & \text{for LUE}, \\
\alpha \log (1 + x) + \beta \log (1 - x), & \text{for JUE},
\end{cases}
\]

\( \alpha, \beta > -1 \), and \( Z_n \) is the normalising constant. The scaling of all the weights is chosen in such a way that the corresponding equilibrium measures (see Section 2.1) are supported on the interval \([-1, 1]\); a similar convention is adopted, e.g., by Charlier and Gharakhloo in [5].

For a continuous real-valued function \( f \) on \( \mathcal{I} \), introduce linear functionals \( \kappa[f] \) and \( \mu[f] \) by the formulas

\[
\kappa[f] = \begin{cases} 
\frac{2}{\pi} \int_{-1}^{1} f(x) \sqrt{1 - x^2} \, dx, & \text{for GUE}, \\
\frac{1}{\pi} \int_{-1}^{1} f(x) \sqrt{\frac{1 - x}{1 + x}} \, dx, & \text{for LUE}, \\
\frac{1}{\pi} \int_{-1}^{1} f(x) \, dx, & \text{for JUE},
\end{cases}
\]


\[
\mu[f] = \begin{cases} 
0, & \text{for GUE}, \\
\frac{\alpha}{2\pi} \int_{-1}^{1} \frac{f(x) - f(-1)}{\sqrt{1 - x^2}} \, dx, & \text{for LUE}, \\
\frac{1}{2\pi} \left( \alpha \int_{-1}^{1} \frac{f(x) - f(-1)}{\sqrt{1 - x^2}} \, dx + \beta \int_{-1}^{1} \frac{f(x) - f(1)}{\sqrt{1 - x^2}} \, dx \right), & \text{for JUE},
\end{cases}
\]  

(5)

and a non-negative quadratic functional \( K[f] \) by the formula

\[
K[f] = \frac{1}{2\pi^2} \int_{-1}^{1} \frac{f(x)}{\sqrt{1 - x^2}} \, dx \, dy \, dx.
\]  

(6)

Let \( F_{f,n}(x) = \mathbb{P}_n\{S_{f,n} \leq x\} \) be the cumulative distribution function, under the measure \( \mathbb{P} \), of the random variable \( S_{f,n} \) defined by

\[
S_{f,n} = \frac{\text{Tr} f(M) - n \mathbb{E}[f] - \mu[f]}{\sqrt{K[f]}},
\]  

(7)

and let \( F_N \) stand for the cumulative distribution function of the standard Gaussian law of expectation zero and variance one:

\[
F_N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-s^2/2} \, ds.
\]  

(8)

In this setting the following theorem holds for the Kolmogorov–Smirnov distance \( \sup_x |F_{f,n}(x) - F_N(x)| \).

**Theorem 1.** Let \( f: \mathcal{I} \to \mathbb{R} \) be a locally Hölder continuous function admitting an analytic continuation to a complex neighbourhood of \([-1, 1]\), and, for GUE and LUE, such that \( f(x) = O(e^{AV(x)}) \) as \(|x| \to +\infty, x \in \mathcal{I}\), for some \( A > 0 \). Then

\[
\sup_{n,x} (n^{1/d}|F_{f,n}(x) - F_N(x)|) < +\infty,
\]  

(9)

where \( d = 5 \) for GUE and LUE, and \( d = 3 \) for JUE.

**Remark.** We point out that for LUE and JUE the rate of convergence prescribed by this theorem is \( 1/d \), and this quantity is uniform with respect to \( \alpha, \beta > -1 \). However, in all the three cases (GUE, LUE, JUE) the question remains to find optimal estimates for such rate.

The rate of convergence in the Central Limit Theorem for the traces of powers of random matrices from compact classical groups has been studied by Stein [26] and Johansson [19].
In [26] the super-polynomial convergence was proven for the circular real ensemble, corresponding to the normalized Haar measure on the orthogonal group. Johansson [19] obtained exponential convergence in the Central Limit Theorem for the circular unitary, real, and quaternion ensembles, corresponding to the normalized Haar measure on the unitary, orthogonal, and symplectic groups, respectively. In [19] the estimate of the Kolmogorov–Smirnov distance is obtained from a super-exponential estimate for the characteristic functions (see [19, Proposition 2.10]). The corresponding proofs are based on explicit representation of the moments and combinatorial identities for Toeplitz determinants.

Let \( \varphi_{f,n}(h) = \mathbb{E}_n[e^{ihS_{f,n}}] \) be the characteristic function of our (centered and normalized) linear statistic, and let \( \varphi_N(h) = e^{-h^2/2} \) be the characteristic function of the standard Gaussian distribution. In order to proof Theorem 1 we need the following lemma.

**Lemma 1.** Let \( f \) satisfying the assumption of Theorem 1 be fixed. Then, there exist \( \varepsilon > 0 \) such that the following takes place for any \( \gamma \in [0, 1/d] \):

\[
\sup_n \sup_{|h| < \varepsilon n^{\gamma}} \left( n^{1-(d-1)\gamma} \left| \frac{\varphi_{f,n}(h) - \varphi_N(h)}{h\varphi_N(h)} \right| \right) < +\infty, \tag{10}
\]

where \( d = 5 \) for GUE and LUE, \( d = 3 \) for JUE.

Also note that an equivalent expression for the functional \( K[f] \) in (6) is given by the formula

\[
K[f] = \frac{1}{4} \sum_{j=1}^{\infty} ja_j^2, \quad a_j = \frac{2}{\pi} \int_0^\pi f(\cos s) \cos js \, ds, \tag{11}
\]

from which the non-negativity of \( K[f] \) follows. The coefficients \( a_j \) are the generalised Fourier coefficients with respect to the orthogonal system of the Chebyshev polynomials of the first kind \( \{T_j(x)\}_{j=0}^{\infty} \):

\[
f(x) = \sum_{j=0}^{\infty} a_j T_j(x), \quad x \in [-1, 1]. \tag{12}
\]

The key element of the proof of Lemma 1 is the local asymptotic analysis of the corresponding Riemann–Hilbert problem (see Section 2.4 and 2.5) in the neighbourhoods of the edges \( \{-1, 1\} \). Our estimates determined by the edge: indeed, the results are the same for GUE and LUE, and the absence of the soft edge makes the convergence faster for JUE.

Of special interest is the case in which \( \gamma = 0 \) and \( h \in \mathbb{R} \) is fixed. The condition \( f(x) = O(e^{AV(x)}) \) as \( x \to \infty, x \in \mathcal{I} \), turns out to be unnecessary, and we have

**Lemma 2.** Fix \( h \in \mathbb{R} \) and let \( f : \mathcal{I} \to \mathbb{R} \) be a fixed locally Hölder continuous function admitting an analytic continuation onto a complex neighbourhood of \([-1, 1]\). Then

\[
\sup_n (n|\varphi_{f,n}(h) - \varphi_N(h)|) < +\infty. \tag{13}
\]
The asymptotic for the real exponential moments $E_n[e^{\text{Tr} f(M)}]$ is due to Charlier and Gharakhloo [5], who also consider more general weights. The Central Limit Theorem follows from the convergence of the real moments, but we do not see how to estimate the rate of convergence to the Gaussian distribution effectively from the asymptotics for the real moments alone. The presence of imaginary exponents (characteristic functions) in (10) and (13) creates additional difficulties: indeed, as we will see in greater detail below, the weight (18) corresponding to the imaginary exponent can have zeros, and therefore the function $\chi(z)$ in (34) can have zeros in a neighbourhood of $[-1, 1]$. To overcome this difficulty, the deformations (111) and (138) of the weight (18) are used to prove Lemma 1 and 2, respectively.

Note also that Lemma 2 holds for a broader class of test functions: existence of exponential moments is not required. In particular, let $T_k$ be the Chebyshev polynomials of the first kind. For $k \geq 1$ set

$$\begin{align*}
\kappa_k &= \begin{cases} 
-\delta_{k,2}/2, & \text{for GUE}, \\
-\delta_{k,1}/2, & \text{for LUE}, \\
0, & \text{for JUE},
\end{cases} \\
\mu_k &= \begin{cases} 
0, & \text{for GUE}, \\
(-1)^{k-1} \alpha, & \text{for LUE}, \\
((-1)^{k-1} \alpha - \beta)/2, & \text{for JUE},
\end{cases}
\end{align*}$$

Introduce a diagonal matrix $\Sigma = \frac{1}{4} \text{diag}\{1, \ldots, l\}$ and the corresponding centred Gaussian distribution $N(0, \Sigma)$.

**Corollary 1.** For the random variables $Y_k = \text{Tr} T_k(M) - n\kappa_k - \mu_k$ we have the following convergence in distribution

$$(Y_1, \ldots, Y_l) \xrightarrow{d} N(0, \Sigma), \quad n \to \infty. \quad (15)$$

Our proof of the upper bound (9) for the rate of convergence to the Gaussian law in Theorem 1 uses the smoothing inequality of Feller [16], that is, an estimate of the Kolmogorov–Smirnov distance via the integral of the distance between the characteristic functions. In order to bound the latter from above one needs to know the behaviour of the characteristic function for large (as $n \to \infty$) arguments, given by Lemma 1. This behaviour does not follow from the results in [5], and that is why this lemma is essential.

Consider a complex-valued function $\tilde{f}$. Passing to the radial part, write

$$E_n\left[e^{\text{Tr} \tilde{f}(M)}\right] = \frac{1}{Z_n} \int_{I^n} e^{\sum_j (\tilde{f}(\lambda_j) - Q_n(\lambda_j))} \prod_{j<k} (\lambda_k - \lambda_j)^2 d\lambda_1 \cdots d\lambda_n. \quad (16)$$

The Andreieff identity gives

$$E_n\left[e^{\text{Tr} \tilde{f}(M)}\right] = \frac{H_{n,n}[\tilde{f}]}{H_{n,n}[0]}, \quad (17)$$

where $H_{n,m}[\tilde{f}] = \det\{\tilde{f}_{j+k-2}^{(m)}\}_{j,k=1}^n$ is the Hankel determinant with the symbol

$$w_m(x) = e^{\tilde{f}(x) - Q_m(x)}, \quad x \in I, \quad (18)$$
and the moments $\mu_j^{(m)}$ are defined by

$$\mu_j^{(m)} = \int_I x^j w_m(x) \, dx. \quad (19)$$

In the sections to come we analyze the large $n$ behaviour of the Hankel determinant in (17). This analysis follows the steepest descent method of Deift and Zhou [15] applied to the Riemann–Hilbert problem for the corresponding system of the orthogonal polynomials. This approach has a long standing history and many applications, e.g., see [1 4 5 8 9 11, 12, 13, 14, 18, 21, 22, 28, 29, 30]. The asymptotic of the Hankel determinant with fixed symbol is well-known. A special feature of our case is that the symbol of the Hankel determinant depends on $n$ in a certain way, which requires new analysis. As a result we obtain Lemma 1.

Johansson [20] deals with the matrix models with the continuous weight on the whole real line. Though the Laguerre and Jacobi ensembles do not satisfy the assumptions of his theorem, the final asymptotic formula (13) is still closely related to Johansson’s. Vanlessen [28] studies the Plancherel–Rotach asymptotics for the orthogonal polynomials with a Laguerre-type weight. Zhao, Cao and Dai [30] obtained the asymptotic expansion of the partition function of a Laguerre-type model. The case of a Laguerre-type singularly perturbed weight was studied by Xu, Dai and Zhao in [29], where the connection was found between the Painlevé III transcendent and the behaviour of the leading and recurrence coefficients of the corresponding orthogonal polynomials. Lyu and Chen [23] studied the distribution of the largest eigenvalue in the Laguerre unitary ensembles. In the physical literature, the connection of eigenvalue statistics to wireless relaying has been studied by Chen and Lawrence [6], Chen, Haq and McKay [7].

Acknowledgements. We are deeply grateful to Christophe Charlier, Tom Claeys, Tamara Grava, Igor Krasovsky, Arno Kuijlaars and Oleg Lisovyi for helpful discussions. Our research is supported by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme, grant 647133 (ICHAOS). A. B. is also supported by Agence Nationale de Recherche, project ANR-18-CE40-0035, and by the Russian Foundation for Basic Research, grant 18-31-20031.

2 Analysis of the Riemann–Hilbert problem

In this section let $\tilde{f}$ be a complex-valued function on $I$ satisfying the following:

A1. $\tilde{f}$ is locally Hölder continuous on $I$;

A2. for GUE and LUE, $\tilde{f}$ satisfies $\max\{\Re \tilde{f}(x), 0\} = O(V(x))$ as $|x| \to +\infty$, $x \in I$;

A3. there exists $n_0 \in \mathbb{N}$ such that the Hankel determinant $H_{k,n}[\tilde{f}]$ is non-zero for all $k = 1, \ldots, n$ and all $n > n_0$;

A4. the function $\tilde{f}$ admits an analytic continuation to a complex neighbourhood of $[-1, 1]$. 

6
The assumptions (A1) and (A2) guarantee that the integral in (19) exists and reduction of (16) to (17) is legitimate. The assumption (A3) implies that the orthogonal polynomials \( p_n^{(k)}(x) = x^k + \ldots \) with respect to the weight (18) are well-defined for \( n > n_0 \).

Now let us consider the matrix function

\[
Y_n(z) = \begin{bmatrix}
\pi_n^{(n)}(z) & C(\pi_n^{(m)} w_n)(z) \\
\beta_{n-1,n} \pi_n^{(n-1)}(z) & \beta_{n-1,n} C(\pi_n^{(n-1)} w_n)(z)
\end{bmatrix}, \quad z \in \mathbb{C} \setminus \mathcal{I},
\]

(20)

where \( \beta_{n,m} = -2\pi i \gamma_{n,m}^2, \gamma_{n,m}^2 = \frac{H_{n,m}[f]}{H_{n+1,m}[f]} \) and \( C \) is the Cauchy-type integral:

\[
C(g)(z) = \frac{1}{2\pi i} \int_{\mathcal{I}} \frac{g(s)}{s - z} \, ds, \quad z \in \mathbb{C} \setminus \mathcal{I}.
\]

(21)

When it brings no ambiguity, we will drop the index \( n \) to make the notation lighter.

Let \( \tilde{\mathcal{I}} \) be the set of interior points of \( \mathcal{I} \). Because of the assumption (A1) upper and lower limits \( Y^\pm(x) \) of \( Y(z) \), as \( z \to x \pm i0, x \in \tilde{\mathcal{I}} \), are well defined. And due to observation by Fokas, Its, and Kitaev [17], it turns out that \( Y(z) \) satisfies the following Riemann–Hilbert problem (Y-RH):

1. \( Y(z) \) is analytic in \( \mathbb{C} \setminus \tilde{\mathcal{I}} \);
2. \( Y^+(x) = Y^-(x) J_Y(x), x \in \tilde{\mathcal{I}}, \) where \( J_Y(x) = \begin{bmatrix} 1 & w_n(x) \\ 0 & 1 \end{bmatrix} \);
3. \( Y(z) = (I + \mathcal{O}(1/z)) z^{n\sigma_3} \) as \( z \to \infty \), where \( \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \), and \( \mathcal{O}(1/z) := \begin{bmatrix} O(1/z) & O(1/z) \\ O(1/z) & O(1/z) \end{bmatrix} \);

4.

\[
Y(z) = \begin{cases}
\mathcal{O}(1), & \text{for GUE}, \\
\begin{bmatrix} O(1) & O(1) + O(|x + 1|^\alpha) \\ O(1) & O(1) + O(|x + 1|^\alpha) \end{bmatrix}, & \alpha \neq 0,
\end{cases}
\]

\[
\begin{cases}
\begin{bmatrix} O(1) & O(1) + O(|x - 1|^\beta) \\ O(1) & O(1) + O(|x - 1|^\beta) \end{bmatrix}, & \beta \neq 0,
\end{cases}
\]

\[
\begin{cases}
O(1) & O(1) + O(|x - 1|^\beta) \\
O(1) & O(1) , & \beta = 0,
\end{cases}
\]

(22)

as \( z \to -1, z \in \mathbb{C} \setminus \tilde{\mathcal{I}} \),

\[
Y(z) = \begin{cases}
\mathcal{O}(1), & \text{for GUE, LUE}, \\
\begin{bmatrix} O(1) & O(1) + O(|x - 1|^\beta) \\ O(1) & O(1) + O(|x - 1|^\beta) \end{bmatrix}, & \beta \neq 0,
\end{cases}
\]

\[
\begin{cases}
\begin{bmatrix} O(1) & O(1) \\ O(1) & O(1) \end{bmatrix}, & \beta = 0,
\end{cases}
\]

(23)

as \( z \to 1, z \in \mathbb{C} \setminus \tilde{\mathcal{I}} \).
Remark. We note that the constant in $O$’s and thus in $O$’s can depend on $n$.

The solution of this problem is unique, and $\det Y(z) = 1$. Indeed, $\det Y(z)$ is an analytic function in $\mathbb{C} \setminus \mathcal{I}$, and from the second condition of (Y-RH) it follows that $\det Y(z)$ has no jumps over $\mathcal{I}$. For LUE and JUE, it is possible that $\det Y(z)$ has isolated singularities at $z = \pm 1$, but from the 4th condition one sees that these singularities are removable. So, $\det Y(z)$ turns to be an entire function in all three cases. Finally, from the 3rd condition and Liouville’s theorem we have that $\det Y(z) = 1$ for all $z \in \mathbb{C}$. Therefore, the matrix $Y(z)$ is invertable, i.e., $(Y(z))^{-1}$ is well-defined. Now, suppose that there are two solutions of (Y-RH), $Y_1(z)$ and $Y_2(z)$. Using similar reasoning as above, $Y_1(z)(Y_2(z))^{-1}$ can be shown to be the identity matrix, and thus the solution is unique.

Asymptotic analysis of Riemann–Hilbert problems includes several steps, which are equivalent transformations of the initial problem (Y-RH). We now describe these steps, starting with the normalisation at $z = \infty$.

2.1 First transformation. Normalisation at $z = \infty$

Consider the equilibrium measure $\nu(dx)$ corresponding to the potential $V(x)$. This measure is a unique solution to the variational problem

$$\int \int \log \frac{1}{|x-y|} \mu(dx)\mu(dy) + \int V(x)\mu(dx) \to \min,$$  \hspace{1cm} (24)

where the minimisation is done over the (convex) set of probability measures $\mu(dx)$ supported on $\Sigma \subset \mathcal{I}$. The optimality conditions following from the corresponding variational inequality can be written as

$$2 \int \log \frac{1}{|x-y|} \mu(dy) + V(x) = l, \quad x \in \Sigma,$$

$$2 \int \log \frac{1}{|x-y|} \mu(dy) + V(x) \geq l, \quad x \in \mathcal{I} \setminus \Sigma,$$ \hspace{1cm} (25)

where $l$ is a real number called the modified Robin constant (see [25]).

The equilibrium measure $\mu(dx) = \psi(x) \, dx$, solution of (24) and (25) for the potential $V(x)$, is known explicitly (e.g., see [5, 14]) for all three potentials in (3) and is supported on $[-1, 1]$

$$\psi(x) = \begin{cases} \frac{2}{\pi} \sqrt{1-x^2}, & \text{for GUE,} \\ \frac{1}{\pi} \sqrt{\frac{1-x}{1+x}}, & \text{for LUE}, \\ \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}}, & \text{for JUE}, \end{cases}$$

$$l = \begin{cases} 1 + 2 \log 2, & \text{for GUE,} \\ 2 + 2 \log 2, & \text{for LUE}, \\ 2 \log 2, & \text{for JUE.} \end{cases}$$ \hspace{1cm} (26)
Now let us introduce the logarithmic potential
\[ g(z) = \int_{-1}^{1} \log(z - s)\psi(s) \, ds, \quad z \in \mathbb{C} \setminus (-\infty, 1], \] (27)
and the auxiliary function
\[ \phi(z) = \begin{cases} 4 \int_{1}^{z} \sqrt{s^2 - 1} \, ds, & \text{for GUE,} \\ 2 \int_{1}^{z} \sqrt{s - 1} \, ds, & \text{for LUE,} \\ -2 \int_{1}^{z} \frac{1}{\sqrt{s^2 - 1}} \, ds, & \text{for JUE,} \end{cases} \] (28)
where the principal branches of log and \(\sqrt{\cdot}\) are used. From these definitions and (25), one can easily show that \(g(z)\) and \(\phi(z)\) are analytic in \(\mathbb{C} \setminus (-\infty, 1]\) and that the following identities hold:
\[ 2g(z) - V(x) + l = -\phi(z), \quad z \in \mathbb{C} \setminus (-\infty, 1], \]
\[ g^+(x) + g^-(x) - V(x) + l = 0, \quad x \in [-1, 1], \]
\[ g^+(x) - g^-(x) = -\phi^+(x) = \phi^-(x), \quad x \in [-1, 1], \]
\[ g^+(x) - g^-(x) = 0, \quad x \in [1, +\infty), \]
\[ g^+(x) - g^-(x) = 2\pi i, \quad x \in (-\infty, 1], \] (29)
where by the subscript indexes + and − we denote upper and lower half-plane limits.

Now, we are ready to perform the first step of the steepest descent analysis and to change the variables in the Riemann–Hilbert problem (Y-RH):
\[ U(z) = e^{\frac{nl}{2} \sigma_3} Y(z) e^{-n(l + g(z)) \sigma_3}. \] (30)
It follows from (29) that \(U(z)\) is analytic in \(\mathbb{C} \setminus \mathcal{I}\). One also has \(U(z) = I + O(1/z)\) as \(z \to \infty\), which follows from the asymptotics \(g(z) = \log(z) + O(1/z)\) as \(z \to \infty\). For the sake of convenience let us introduce the short notation
\[ \chi(x) = e^{\tilde{f}(x) + \omega(x)}. \] (31)
The Riemann–Hilbert problem (U-RH) for \(U(z)\) reads as follows.
1. \(U(z)\) is analytic in \(\mathbb{C} \setminus \mathcal{I}\);
2. \(U^+(x) = U^-(x)J_U(x), \quad x \in \mathcal{I}, \)
\[ J_U(x) = \begin{cases} \begin{bmatrix} 1 & \chi(x)e^{-n\phi(x)} \\ 0 & 1 \end{bmatrix}, & x \in \mathcal{I} \setminus [-1, 1], \\ \begin{bmatrix} e^{n\phi^+(x)} & \chi(x) \\ 0 & e^{n\phi^-(x)} \end{bmatrix}, & x \in (-1, 1); \end{cases} \] (32)
3. $U(z) = (I + \mathcal{O}(1/z))$ as $z \to \infty$;

4. The behaviour of $U(z)$ as $z \to \pm 1$ is the same as that of $Y(z)$ in (Y-RH).

It is not difficult to check that the formulas (32) for the jump $J_U(x)$ follow from the direct computations:

$$\begin{align*}
J_U(x) &= e^{n(\frac{i}{2} + g^{-}(x))\sigma_3}J_Y(z)e^{-n(\frac{i}{2} + g^{+}(x))\sigma_3} = \begin{bmatrix} e^{-n(g^{+}(x)+g^{-}(x))} & \chi(x)e^{n(\gamma^{+}(x)+g^{-}(x)-V(x)+i)} \\ 0 & e^{n(g^{+}(x)-g^{-}(x))} \end{bmatrix} = \\
&= \begin{cases} 
\begin{bmatrix} 1 & \chi(x)e^{-n\phi(x)} \\
0 & 1 \end{bmatrix}, & x \in I \setminus [-1,1], \\
\begin{bmatrix} e^{n\phi^{+}(x)} & \chi(x) \\
0 & e^{n\phi^{-}(x)} \end{bmatrix}, & x \in (-1,1), 
\end{cases}
\end{align*}$$

(33)

where we used the definition of $J_Y(x)$ from (Y-RH) and the formulas (29).

We highlight that $\phi(z)$ is defined in such a way that $\text{Re} \phi(x) > 0$ for $x \in I \setminus [-1,1]$, and therefore for such $x$ we have $J_U(x) \to I$ as $n \to \infty$. What is more, this convergence is exponentially fast but not uniform because of the edges $\{-1,1\}$. On the other hand, $\phi^{\pm}(x)$ are imaginary for $x \in (-1,1)$, and thus $e^{n\phi^{\pm}(x)}$ are oscillating as $n \to \infty$. The latter brings us to the next step of the steepest descent analysis, namely the contour deformation.

### 2.2 Second transformation. Deformation of the contour

Observe that there is a simple matrix identity:

$$\begin{align*}
\begin{bmatrix} e^{n\phi^{+}(x)} & \chi(x) \\
0 & e^{n\phi^{-}(x)} \end{bmatrix} &= \begin{bmatrix} 1 & \frac{1}{\chi(x)}e^{n\phi^{-}(x)} \\
0 & \frac{1}{\chi(x)}e^{n\phi^{+}(x)} \end{bmatrix} \cdot \begin{bmatrix} \chi(x) & 0 \\
-\frac{1}{\chi(x)} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \frac{1}{\chi(x)}e^{n\phi^{+}(x)} \\
0 & 1 \end{bmatrix}^{-1} \\
&= : J_T^{-}(x)J_T^{0}(x)J_T^{+}(x).
\end{align*}$$

(34)

We deform the positive half-line into the lens-shaped contour as in Fig. 1. The idea is to find the function $T(z)$ such that it has the same jump as $Y(z)$, but the jump is spread over the contour $L = L_{+} \cup L_{-} \cup I$; note that the lips $L_{\pm}$ do not include the edges $\{-1,1\}$. More precisely, we specify $T(z)$ by the following formula:

$$\begin{align*}
T(z) &= \begin{cases} 
U(z), & z \in \Omega = \mathbb{C} \setminus (I \cup \Omega^{+} \cup \Omega^{-}), \\
U(z)(J_T^{+}(z))^{-1}, & z \in \Omega^{+}, \\
U(z)J_T^{-}(z), & z \in \Omega^{-}.
\end{cases}
\end{align*}$$

(35)

At this moment we need the assumption (A4) that says that there is the analytic continuation of the function $\hat{f}(x)$ to some neighbourhood of $[-1,1]$. Without loss of generality, the lens is embedded in this neighbourhood, and one realises that $J_T^{+}(x)$ and $J_T^{-}(x)$ have analytic continuous-up-to-the-boundary extensions from $(-1,1)$ to $\Omega^{+}$ and $\Omega^{-}$, respectively, so the formula (35) makes sense. Note that in $\omega(z)$ (see (3)) we use the principal branch
of log\( z \) defined for \( z \in \mathbb{C} \setminus (-\infty,0] \). For the sake of readability, all the analytic extensions are denoted by the same symbols as their counterparts.

It is readily verified that the function \( T(z) \) satisfies the following Riemann–Hilbert problem (T-RH):

1. \( T(z) \) is analytic in \( \mathbb{C} \setminus L \);
2. \( T^+(z) = T^-(z)J_T(z) \), \( z \in L \),

\[
J_T(z) = \begin{cases} 
J^+_T(z), & z \in L_+, \\
J^0_T(z), & z \in (-1,1), \\
J^-_T(z), & z \in L_-, \\
J_U(z), & z \in \mathcal{I} \setminus [-1,1];
\end{cases}
\]  

(36)

3. \( T(z) = (I + \mathcal{O}(1/z)) \) as \( z \to \infty \);

4. The behaviour of \( T(z) \) in the neighbourhoods of points \( z = -1 \) and \( z = 1 \) is the same as that of \( U(z) \), if approaching from \( \Omega \). If approaching from \( \Omega^+ \) and \( \Omega^- \), the result can be obtained by multiplying by the corresponding jump matrix (see (35)).

Employing the definition (28), one can check directly that \( \text{Re} \phi(z) < 0 \) on the lips (excluding the endpoints). Consequently, one has \( J_T(z) \to I \) exponentially fast, however again this convergence is not uniform because of the edges \( \{-1,1\} \).

Now we summarise what we have done so far. The initial problem for \( Y(z) \) is reduced to the problem for \( T(z) \) in such a way that \( J_T(z) \to I \) on \( L_+ \cup L_- \cup (\mathcal{I} \setminus [-1,1]) \), exponentially fast but not uniformly. In view of the small norm theorem, one may want to consider the limit problem with the jump matrix \( J^x_T(x) \) on the contour \((-1,1)\). We will explicitly solve

---

**Figure 1:** The deformed contour \( L = L_+ \cup L_- \cup \mathcal{I} \) of the Riemann–Hilbert problem (T-RH).
such a problem in the next section and construct the so-called global parametrix. However, we underline that in order to apply the small norm theorem, the uniform convergence is crucial. And to make up for lack of it—as it is usually done—in the sections to come we perform the local analysis of the problem and construct the so-called local parametrices in the neighbourhoods of \( z = -1 \) and \( z = 1 \).

### 2.3 Parametrix at \( z = \infty \)

Consider the following Riemann–Hilbert problem (N-RH):

1. \( N(z) \) is analytic in \( \mathbb{C} \setminus [-1, 1] \);
2. \( N^+(x) = N^-(x) J^o_T(x) \), \( x \in (-1, 1) \);
3. \( N(z) = (I + \mathcal{O}(1/z)) \) as \( z \to \infty \).

In order to find its solution, first we notice that the following identity takes place:

\[
J^o_T(x) = (\chi(x))^a \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.
\]

Next, we introduce the Szegő function \( D(z) \) by the formula

\[
D(z) = \exp \left( \frac{1}{2\pi} \sqrt{z^2 - 1} \int_{-1}^{1} \frac{\omega(s) + \tilde{f}(s)}{\sqrt{1-s^2}} \cdot \frac{ds}{z-s} \right), \quad z \in \mathbb{C} \setminus [-1, 1],
\]

where \( \omega(x) \) is given by (3), and the principal branch of \( \sqrt{\cdot} \) is used. This function is clearly analytic in \( \mathbb{C} \setminus [-1, 1] \), and from the Sokhotskii–Plemelj formulas it is immediate to see that

\[
D^+(x)D^-(x) = \chi(x), \quad x \in (-1, 1).
\]

The expression (38) can be easily factorised \( D(z) = D_1(z)D_2(z) \), where

\[
D_1(z) = \exp \left( \frac{1}{2\pi} \sqrt{z^2 - 1} \int_{-1}^{1} \frac{\omega(s)}{\sqrt{1-s^2}} \cdot \frac{ds}{z-s} \right),
\]

\[
D_2(z) = \exp \left( \frac{1}{2\pi} \sqrt{z^2 - 1} \int_{-1}^{1} \frac{\tilde{f}(s)}{\sqrt{1-s^2}} \cdot \frac{ds}{z-s} \right).
\]

It is known that \( D_1(z) \) can be written explicitly (e.g., see \[5, 28\]). However, for the sake of completeness we obtain the corresponding expression the case of LUE using the residue
theory. First, we make the change of variables: \( s \mapsto \frac{1-s^2}{1+s^2} \), and then the integral in \( D_1(z) \) takes the form

\[
I = \frac{1}{2\pi} \int_{-1}^{1} \frac{\log (1+s)}{z-s} \frac{ds}{\sqrt{1-s^2}} = -\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\log((s^2+1)/2)}{(z+1)s^2+z-1} ds. \tag{41}
\]

Calculating the corresponding residues one obtains

\[
I = \frac{\log(z+1)}{2\sqrt{z^2-1}} - \int_{1}^{+\infty} \frac{ds}{(z+1)s^2-z+1}
= \frac{1}{2\sqrt{z^2-1}} \left( \log(z+1) + \log\frac{\sqrt{z+1} + \sqrt{z-1}}{\sqrt{z+1} + 1 + \sqrt{z-1}} \right), \tag{42}
\]

from which it follows that

\[
D_1(z) = (z+1)^{\alpha/2} \left( \frac{\sqrt{z+1} + \sqrt{z-1}}{\sqrt{z+1} + 1 + \sqrt{z-1}} \right)^{\alpha/2}. \tag{43}
\]

It is interesting to point out that the function

\[
z \mapsto \frac{\sqrt{z+1} - \sqrt{z-1}}{\sqrt{z+1} + \sqrt{z-1}} \tag{44}
\]

is a conformal map of the complex plane with the slit \([-1, 1]\) onto the interior of the standard unit disk without \(\{0\}\). Thereby, this function maps any loop going around \([-1, 1]\) counterclockwise into a loop inside the disk going around zero clockwise. In this way the jump of \((z+1)^{\alpha/2}\) is compensated, and the resulting function \(D_1(z)\) is indeed analytic in \(\mathbb{C} \setminus [-1, 1]\).

In the similar way one can handle the case of JUE, and finally we arrive at

\[
D_1(z) = \begin{cases} 
1, & \text{for GUE,} \\
(z+1)^{\alpha/2} \left( \frac{\sqrt{z+1} - \sqrt{z-1}}{\sqrt{z+1} + 1 + \sqrt{z-1}} \right)^{\alpha/2}, & \text{for LUE,} \\
(z+1)^{\alpha/2}(z-1)^{\beta/2} \left( \frac{\sqrt{z+1} - \sqrt{z-1}}{\sqrt{z+1} + \sqrt{z-1}} \right)^{(\alpha+\beta)/2}, & \text{for JUE.}
\end{cases} \tag{45}
\]

We also need to know the formula for \(D(\infty)\), which clearly is

\[
D(\infty) = D_1(\infty)D_2(\infty), \tag{46}
\]

where

\[
D_1(\infty) = \begin{cases} 
1, & \text{for GUE,} \\
2^{-\alpha/2}, & \text{for LUE,} \\
2^{-(\alpha+\beta)/2}, & \text{for JUE,}
\end{cases} \quad D_2(\infty) = \exp \left( \frac{1}{2\pi} \int_{-1}^{1} \frac{\tilde{f}(s)}{\sqrt{1-s^2}} ds \right). \tag{47}
\]

13
Now, we make the change of variables in (N-RH):

\[ C(z) = (D(\infty))^{-\sigma_3}N(z)(D(z))^{i\sigma_3} \tag{48} \]

and find that \( C(z) \) satisfies the following Riemann–Hilbert problem (C-RH):

1. \( C(z) \) is analytic in \( \mathbb{C} \setminus [-1, 1] \);
2. \( C^+(x) = C^-(x) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \ x \in (-1, 1); \)
3. \( C(z) = (I + O(1/z)) \) as \( z \to \infty \).

Diagonalizing the constant jump matrix one can find the the solution in the following form:

\[ C(z) = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} (q(z))^{\sigma_3} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2i}(q(z) + q^{-1}(z)) & \frac{1}{2i}(q(z) - q^{-1}(z)) \\ -\frac{1}{2i}(q(z) - q^{-1}(z)) & \frac{1}{2i}(q(z) + q^{-1}(z)) \end{bmatrix}, \tag{49} \]

where \( q(z) = (\frac{z-1}{z+1})^{1/4} \) with the principal branch of the root used.

Expressing \( N(z) \) in terms of \( C(z) \) using \( [18] \), we arrive at the formula for the global parametrix:

\[ N(z) = (D(\infty))^{\sigma_3}C(z)(D(z))^{-\sigma_3} \]

\[ = \begin{bmatrix} \frac{D(\infty)}{2D(z)}(q(z) + q^{-1}(z)) & \frac{D(z)D(\infty)}{2i}(q(z) - q^{-1}(z)) \\ -\frac{1}{2iD(z)D(\infty)}(q(z) - q^{-1}(z)) & \frac{D(z)}{2D(\infty)}(q(z) + q^{-1}(z)) \end{bmatrix}. \tag{50} \]

It is worth noticing that the Szegö function is not a unique solution to the multiplicative Riemann–Hilbert problem \([39]\) due to the fact that the asymptotic behaviour at \( z = -1 \) and \( z = 1 \) is not specified. However, as we will see later, this solution is exactly what one needs. Also we indicate that \( \det N(z) = 1 \), thus the matrix \( N(z) \) is invertible, and one can write

\[ N^{-1}(z)N'(z) = \begin{bmatrix} \frac{D'(z)}{D(z)} & \frac{D^2(z)q'(z)}{q(z)} \\ -\frac{D'(z)}{D(z)} & \frac{D(z)q(z)}{q(z)} \end{bmatrix} = \begin{bmatrix} \frac{D'(z)}{D(z)} & \frac{D^2(z)}{2(z^2-1)} \\ -\frac{D'(z)}{D(z)} & \frac{D(z)}{2(z^2-1)} \end{bmatrix}, \tag{51} \]

which we need for future reference.

### 2.4 Local parametrix at \( z = 1 \)

#### 2.4.1 The case of GUE and LUE. Soft edge

We start off by considering the situation where \( z = 1 \) is the so-called soft edge, that is, the equilibrium measure vanishes like \( |z-1|^{1/2} \) as \( z \to 1 \), or equivalently \( z = 1 \in T \) and particles can escape to the right of 1. This corresponds to the case of GUE and LUE.

Let \( \Omega_1 \) be a small neighbourhood of the point \( z = 1 \), and consider the Riemann–Hilbert problem (P1-RH):
1. $P_1(z)$ is analytic in $\Omega_1 \setminus L$;
2. $P_1^+(z) = P_1^-(z)J_T(z)$, $z \in L \cap \Omega_1$, where $J_T(z)$ is defined in (36);
3. $P_1(z)(N(z))^{-1} = I + O(1/n)$ as $n \to \infty$, uniformly on $\partial \Omega_1$;
4. $P_1(z)$ is bounded at $z = 1$.

The function $P_1(z)$ satisfies the same jump conditions and has the same local behaviour as $T(z)$, and one can see that the idea is to match $P_1(z)$ with the global parametrix $N(z)$ on the boundary $\partial \Omega_1$ as $n \to \infty$. We note that although $N(z)$ is discontinuous at $\partial \Omega_1 \cap (-1, 1)$, the 3rd condition implies that $P_1(z)(N(z))^{-1}$ is not. Also the 4th condition of (P1-RH) is the same as that of (T-RH).

The problem (P1-RH) is local, and making use of the fact that one can choose $\Omega_1$ to be sufficiently small so that $\chi(x)$ is analytic in $\Omega_1 \setminus L$, we can perform the simple change of variables that turns the jump matrix $J_T(z)$ into a piecewise constant one:

$$\tilde{P}_1(z) = P_1(z)e^{-\frac{\pi}{2}\phi(z)\sigma_3(\chi(z))^{\sigma_3/2}}, \quad z \in \Omega_1 \setminus L. \quad (52)$$

The function $\tilde{P}_1(z)$ satisfies the following Riemann–Hilbert problem ($\tilde{P}_1$-RH):
1. $\tilde{P}_1(z)$ is analytic in $\Omega_1 \setminus L$;
2. $\tilde{P}_1^+(z) = \tilde{P}_1^-(z)J_1(z)$, $z \in L \cap \Omega_1$,

$$J_1(z) = \begin{cases} 
\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, & z \in (1, +\infty) \cap \Omega_1, \\
\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, & z \in (L_+ \cup L_-) \cap \Omega_1, \\
\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, & z \in (-1, 1) \cap \Omega_1;
\end{cases} \quad (53)$$
3. $\tilde{P}_1(z)(\chi(z))^{-\sigma_3/2}e^{\frac{\pi}{2}\phi(z)\sigma_3(N(z))^{-1}} = I + O(1/n)$, uniformly on $\partial \Omega_1$;
4. $\tilde{P}_1(z)$ is bounded at $z = 1$.

In order to find $\tilde{P}_1(z)$, we study a close related Riemann–Hilbert problem (A-RH) with the same jumps as ($\tilde{P}_1$-RH) but on the other contour (see Fig.2): We need to find a function $A(\zeta)$ analytic in $\mathbb{C} \setminus \tilde{L}$, bounded at $\zeta = 0$, and such that $A^+(\zeta) = A^-(\zeta)J_A(\zeta)$, $\zeta \in \mathbb{C} \setminus \tilde{L}$ with the jump matrix $J_A(\zeta)$ given by

$$J_A(\zeta) = \begin{cases} 
\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, & \zeta \in (0, +\infty), \\
\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, & \zeta \in (\tilde{L}_+ \cup \tilde{L}_-), \\
\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, & \zeta \in (-\infty, 0).
\end{cases} \quad (54)$$
Although, this problem does not have a unique solution, as the behaviour at $\zeta = \infty$ is not specified, the useful choice of $A(z)$ turns out to be the one in terms of the Airy functions (see [13]):

$$A(\zeta) = \sqrt{2\pi} \begin{bmatrix} \operatorname{Ai}(\zeta) & -\omega^2 \operatorname{Ai}(\omega^2 \zeta) \\ -i \operatorname{Ai}'(\zeta) & i\omega \operatorname{Ai}'(\omega^2 \zeta) \end{bmatrix}, \quad \arg \zeta \in (0, 2\pi/3),$$

(55)

where $\omega = e^{2\pi i/3}$. And for the other sectors the similar formulas can be obtained by using the jump matrix $J_A(\zeta)$. We also draw our attention to the fact that $\det A(\zeta) = 1$, particularly, the matrix $A(\zeta)$ is invertible.

Recalling the asymptotic behaviour of the Airy function as $\zeta \to \infty$, one can find that

$$A(\zeta) = \zeta^{-\sigma_3/4} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} (I + O(\zeta^{-3/2})) e^{-\frac{2}{3} \zeta^{3/2} \sigma_3}$$

(56)

as $\zeta \to \infty$ for $\zeta \in \mathbb{C} \setminus \tilde{L}$, where the principal branches of the roots are used.

Next, we find the conformal map $\xi_n(z)$ from $\Omega_1$ onto some neighbourhood of $\zeta = 0$ and construct the solution to $(\tilde{P}_1\text{-RH})$ in the form

$$\tilde{P}_1(z) = E_n(z) A(\xi_n(z)),$$

(57)

where $E_n(z)$ is some analytic function in $\Omega_1$. It is clear that if the conformal map $\xi_n(z)$ is found, then (57) satisfies 1st, 2nd, and 4th conditions of $(\tilde{P}_1\text{-RH})$ automatically. We choose $E_n(z)$ so that the 3rd condition is also satisfied.

Let us find $\xi_n(z)$ in such a way that the exponents in the 3rd condition of $(\tilde{P}_1\text{-RH})$ and in (56) cancel out:

$$\frac{2}{3} (\xi_n(z))^{3/2} = \frac{n}{2} \phi(z).$$

(58)

To solve this equation we let $\xi_n(z) = (3n/4)^{2/3}(\phi(z))^{2/3}$, where the right-hand side is analytically continued to the analytic function in $\Omega_1$ using the principal branch of the power function. Expanding the function $\phi(z)$ given by (28) in the series in the neighbourhood of the branching point $z = 1$, we immediately see that

$$\xi_n(z) = n^{2/3} (z - 1) G(z),$$

(59)
for some analytic in this neighbourhood function $G(z)$ such that $G(1) \neq 0$. Note that asymptotic behaviour of $\phi(z)$ at $z = 1$ for GUE and LUE is the same. In fact, that is the reason why these situations are identified as the soft-edge ones.

The identity (69) shows that $\xi_n(z)$ is indeed a conformal map of the neighbourhood of $z = 1$, which without loss of generality is $\Omega_1$, to some neighbourhood $\tilde{\Omega}$ of $\zeta = 0$. Also, we note that $\xi_n(z)$ maps $(1, +\infty) \cap \Omega_1$ and $(-1, 1) \cap \Omega_1$ in the $z$-plane into $(0, +\infty) \cap \Omega$ and $(-\infty, 0) \cap \Omega$ in the $\zeta$-plane, preserving the orientation. Besides, we can always use the freedom to deform the lens so that its lips are mapped into $\tilde{L}_+$ and $\tilde{L}_-$.

The next step is to match the whole asymptotics as $n \to \infty$ and find the analytic factor $E_n(z)$. We see that if we fix $z$ and let $n \to \infty$, then $\xi_n(z) \to \infty$, consequently, for $A(\xi_n(z))$ the asymptotics (56) is relevant. Due to this fact and the formulas (52), (56), (57), and (59), the left part of the matching condition reads:

$$P_1(z)(N(z))^{-1} = \frac{E_n(z)}{\sqrt{2}} (3n\phi(z)/4)^{-\sigma_3/6} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \left( I + \mathcal{O}(\xi_n(z))^{-3/2} \right) (\chi(z))^{-\sigma_3/2} (N(z))^{-1}.$$  \hfill (60)

In order to satisfy the condition we choose

$$E_n(z) = \sqrt{2} N(z) (\chi(z))^{\sigma_3/2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}^{-1} (3n\phi(z)/4)^{\sigma_3/6}. \hfill (61)$$

The formula (60) becomes

$$P_1(z)(N(z))^{-1} = I + N(z)(\chi(z))^{\sigma_3/2} \mathcal{O}(\xi_n(z))^{-3/2} (\chi(z))^{-\sigma_3/2} (N(z))^{-1}. \hfill (62)$$

And since $N(z)$ and $\chi(z)$ are uniformly bounded and

$$\mathcal{O}(\xi_n(z))^{-3/2} = \mathcal{O}(1/n) \hfill (63)$$

uniformly in $z \in \partial \Omega_1$, the matching condition $P_1(z)(N(z))^{-1} = I + \mathcal{O}(1/n)$ is satisfied, as desired.

Now it is left to check that $E_n(z)$ is analytic in $\Omega_1$. Clearly, $E_n(z)$ is analytic in $\Omega_1 \setminus (-1, 1)$, therefore it is sufficient to verify that there is no jumps over $(-1, 1) \cap \Omega_1$ and that there is no singularity at $z = 1$. To check the first claim, we use the jump condition for $N(z)$ and compare the limits $E_n^+(z)$ and $E_n^-(z)$ from above and below of $(-1, 1) \cap \Omega_1$:

$$E_n^+(z) = \sqrt{2} N^+(z) (\chi(z))^{\sigma_3/2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}^{-1} [3n\phi(z)/4]^{\sigma_3/6}^+$$

$$= \sqrt{2} N^+(z) \begin{bmatrix} 0 & \chi(z) \\ -\chi(z) & 0 \end{bmatrix} (\chi(z))^{\sigma_3/2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}^{-1} e^{\frac{i\sigma_3}{12}} [(3n\phi(z)/4)^{\sigma_3/6}]^- \hfill (64)$$

$$= \sqrt{2} N^-(z) (\chi(z))^{\sigma_3/2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}^{-1} [(3n\phi(z)/4)^{\sigma_3/6}]^- = E_n^-(z).$$
One sees that \( E_n(z) \) is indeed analytic in \( \Omega_1 \setminus \{1\} \) and thus can only have an isolated singularity at \( z = 1 \). Yet, from the explicit formula (61), the order of this singularity is at most \( 1/2 \), and hence the singularity is bound to be removable. This justifies the second claim. Our final observation is that since \( \det N(z) = 1 \), one has \( \det E_n(z) = 1 \), and thus \( \tilde{P}_1(z) = det P_1(z) = 1 \). Particularly, all of these matrices are non-singular.

2.4.2 The case of JUE. Hard edge

Now we proceed with the situation where \( z = 1 \) is the so-called hard edge, that is, the equilibrium measure blows up like \( |z - 1|^{-1/2} \) as \( z \to 1 \), or equivalently \( z = 1 \in \partial I \) and particles cannot escape from \([-1, 1]\). This corresponds to the case of JUE.

Let \( \Omega_1 \) be a neighbourhood of \( z = 1 \) as earlier. The construction of the hard-edge parametrix is very similar to the construction in the previous section. The key difference, though, is that the contour of the Riemann–Hilbert problem is different, and that the solution is supposed to have different behaviour at \( z = 1 \) (see the 4th condition of (Y-RH)).

Consider the Riemann–Hilbert problem \((P_1-RH)\) from the previous section, but instead of the 4th condition we have

4. The behaviour of \( P_1(z) \) as \( z \to 1 \) is

\[
P_1(z) = \begin{cases} 
O(1) & O(1) + O(|z - 1|^\beta), \quad \beta \neq 0, \\
O(1) & O(1) + O(|z - 1|^\beta), \\
O(1) & O(\log |z - 1|), \\
O(1) & O(\log |z - 1|)
\end{cases}
\]

(65)

It is convenient to introduce the function \( \tilde{\omega}(z) = \alpha \log (z + 1) + \beta \log (z - 1), z \in \mathbb{C} \setminus (-\infty, 1] \), (cf. (3) for JUE) with the principal branch of log used and define \( \tilde{\chi}(z) = e^{\tilde{\phi}(z) + \tilde{\omega}(z)} \).

The latter is clearly analytic in \( \Omega_1 \setminus (-1, 1) \). Performing the change of variables

\[
\tilde{P}_1(z) = P_1(z)e^{-\frac{n}{2} \phi(z)\sigma_3 (\tilde{\chi}(z))^{\sigma_3/2}}, \quad z \in \Omega_1 \setminus L,
\]

(66)

we reduce \((P_1-RH)\) to the problem \((\tilde{P}_1-RH)\) with the piecewise constant jump matrix:

1. \( \tilde{P}_1(z) \) is analytic in \( \Omega_1 \setminus L \);
2. \( \tilde{P}_1^+(z) = \tilde{P}_1^-(z)J_1(z), z \in L \cap \Omega_1 \),

\[
J_1(z) = \begin{cases} 
1 & 0, \quad z \in L_+ \cap \Omega_1, \\
0 & 1, \quad z \in L_- \cap \Omega_1, \\
1 & 0, \quad z \in (-1, 1) \cap \Omega_1;
\end{cases}
\]

(67)
3. \( \tilde{P}_1 (z) (\tilde{\chi} (z))^{-\sigma_3 / 2} e^{\frac{\pi}{2} \phi (z) \sigma_3} (N (z))^{-1} = I + O (1/n) \) as \( n \to \infty \), uniformly on \( \partial \Omega_1 \);

4. The behaviour of \( \tilde{P}_1 (z) \) as \( z \to 1 \):

\[
\tilde{P}_1 (z) = \begin{cases}
O (1) |z - 1|^{\beta \sigma_3 / 2}, & \beta > 0, \\
O (|z - 1|^{\beta / 2}), & \beta < 0, \\
\begin{bmatrix}
O (1) & O (\log |z - 1|) \\
O (1) & O (\log |z - 1|)
\end{bmatrix}, & \beta = 0.
\end{cases}
\]  

Similarly as in the previous section, we consider an auxiliary Riemann–Hilbert problem \( (\Psi, \text{RH}) \) in the \( \zeta \)-complex plane on the special contour (see Fig. 3). The jumps over \( \tilde{L}_+, \tilde{L}_- \), and \((-\infty, 1)\) are the same as those over \( L_+, L_- \), and \((0, 1)\) in (67).

![Figure 3: The contour \( \tilde{L} = \tilde{L}_+ \cup \tilde{L}_- \cup (-\infty, 0) \) of the auxiliary problem \( (\Psi, \text{RH}) \) in the \( \zeta \)-plane.](image)

The matrix function \( \Psi (\zeta) \) that satisfies the conditions mentioned is known (although not uniquely) and is given in terms of the modified Bessel functions \( I_\alpha \) and \( K_\alpha \) of order \( \alpha \) (see [5, 28]):

\[
\Psi (\zeta) = \begin{bmatrix}
I_\alpha (2\sqrt{\zeta}) & i K_\alpha (2\sqrt{\zeta}) \\
2\pi i \sqrt{\zeta} I'_\alpha (2\sqrt{\zeta}) & -2\sqrt{\zeta} K'_\alpha (2\sqrt{\zeta})
\end{bmatrix},
\]  

for \( \text{arg} \zeta \in (-2\pi/3, 2\pi/3) \). To define \( \Psi (\zeta) \) in the other sectors, one uses the corresponding jumps. It is worth noticing that the angle we used in Fig. 3 is somewhat arbitrary. In fact the solution is known in a much more general case (for details see [22]). Also, as it was before, \( \det \Psi (\zeta) = 1 \), and \( \Psi (\zeta) \) turns out to be non-singular.

The asymptotic behaviour of \( \Psi (\zeta) \) can be recovered from the known properties of the modified Bessel functions and is given by

\[
\Psi (\zeta) = (2\pi)^{-\sigma_3 / 2} \zeta^{-\sigma_3 / 4} \frac{1}{\sqrt{2}} \begin{bmatrix}
1 & i \\
i & 1
\end{bmatrix} (I + O (\zeta^{-1/2})) e^{2\zeta^{1/2} \sigma_3}
\]  

for \( \zeta \in \mathbb{C} \setminus \tilde{L} \) as \( \zeta \to \infty \).

Again, we look for the parametrix in the form

\[
\tilde{P}_1 (z) = E_n (z) \Psi (\eta_n (z)),
\]  

(71)
where $\eta_n(z)$ is a conformal map from $\Omega_1$ to some neighbourhood of $\zeta = 1$, and $E_n(z)$ is an analytic factor. Clearly, the 1st and the 2nd conditions of $(\tilde{P}_1\text{-RH})$ are satisfied. We find $E_n(z)$ and $\eta_n(z)$ such that the 3rd condition is satisfied as well.

For the exponents in (70) and in the 3rd condition of $(\tilde{P}_1\text{-RH})$ to cancel out, we need to solve the equation for $\eta_n(z)$:

$$e^{2\sqrt{\eta_n(z)}} = e^{-\frac{n}{2}\phi(z)},$$

which gives

$$\eta_n(z) = \frac{n^2}{16} \left( \phi(z) \right)^2.\quad (73)$$

Expanding $\phi(z)$ in the series in the neighbourhood of $z = 1$, one can find that

$$\eta_n(z) = \frac{n^2}{16} (z - 1)G(z)\quad (74)$$

for some analytic function $G(z)$ which satisfies $G(1) \neq 0$. Thus, $\eta_n(z)$ is a conformal map of the neighbourhood of $z = 1$, which without loss of generality is $\Omega_1$, to some neighbourhood $\tilde{\Omega}$ of $\zeta = 0$. We also note that $\eta_n(z)$ maps $(-1, 1)\cap \Omega_1$ into $(-\infty, 0)\cap \tilde{\Omega}$ in the $\zeta$-plane, preserving the orientation. Moreover, due to the freedom to deform the lens, we can always think that its lips are mapped into $\tilde{L}_+$ and $\tilde{L}_-$.  

Now we find the analytic factor $E_n(z)$ such that the 3rd condition of $(\tilde{P}_1\text{-RH})$ is satisfied fully. Since $\eta_n(z) \to \infty$ as $n \to \infty$, the asymptotics for $\Psi(\eta_n(z))$ as $z \to \infty$ is of relevance. We write the left-hand side of the 3rd (matching) condition, taking into account (60), (71), (70), and (73):

$$P_1(z)(N(z))^{-1} = \frac{E_n(z)}{\sqrt{2}} \left( \frac{\pi n \phi(z)}{2} \right)^{-\sigma_3/4} \left[ \begin{array}{cc} 1 & i \\ i & 1 \end{array} \right] \times \left( I + \mathcal{O}(\eta_n(z))^{-1/2} \right)(\tilde{\chi}(z))^{-\sigma_3/2}(N(z))^{-1}.$$  

(75)

In order to satisfy the matching condition we define $E_n(z)$ as follows

$$E_n(z) = N(z)(\tilde{\chi}(z))^\sigma_3/2 \frac{1}{\sqrt{2}} \left[ \begin{array}{cc} 1 & -i \\ -i & 1 \end{array} \right] (\pi n \phi(z)/2)^{\sigma_3/2}. \quad (76)$$

The formula (75) becomes

$$P_1(z)(N(z))^{-1} = I + N(z)(\tilde{\chi}(z))^\sigma_3/2 \mathcal{O}(\eta_n(z))^{-1/2}(\tilde{\chi}(z))^{-\sigma_3/2}(N(z))^{-1}. \quad (77)$$

(77)

Remembering that $N(z)$ and $\tilde{\chi}(z)$ are uniformly bounded on $\partial \Omega_1$ and that

$$\mathcal{O}(\eta_n(z))^{-1/2} = \mathcal{O}(1/n) \quad (78)$$

uniformly in $z \in \partial \Omega_1$, one finally arrives at $P_1(z)(N(z))^{-1} = I + \mathcal{O}(1/n)$ uniformly in $z \in \partial \Omega_1$ as desired.
Now, we check that $E_n(z)$ is analytic in $\Omega_1$. By construction, $E_n(z)$ is analytic in $\Omega_1 \setminus (-1, 1)$. We verify that $E_n(z)$ has no jumps over $(-1, 1)$ approaching from the upper and lower half planes:

\[
E_n^+(z) = \frac{1}{\sqrt{2}} N^+(z) [\tilde{\chi}(z)^{\sigma_1/2}]^+ \cdot \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} [(\pi n \phi(z)/2)^{\sigma_1/2}]^+ \\
E_n^-(z) = \frac{1}{\sqrt{2}} N^-(z) \begin{bmatrix} 0 & \chi(z) \\ -\frac{1}{\chi(z)} & 0 \end{bmatrix} (\chi(z)e^{\pi i \beta})^{\sigma_3/2} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} e^{\frac{\pi i \sigma_3}{2}} [(\pi n \phi(z)/2)^{\sigma_3/2}]^- \\
E_n(z) = \frac{1}{\sqrt{2}} N^-(z) [\tilde{\chi}(z)^{\sigma_3/2}]^- \cdot \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} [(\pi n \phi(z)/2)^{\sigma_3/2}]^- = E_n^-(z).
\]

Consequently, $E_n(z)$ is analytic in $\Omega_1 \setminus \{1\}$ and can only have an isolated singularity at $z = 1$. The explicit formula (76) shows that the order of the singularity is no more than 1/2, and once again one concludes that this singularity is in fact removable, so $E_n(z)$ is analytic in $\Omega_1$. In the usual manner we observe that $\det E_n(z) = 1$, and thus $\det \tilde{P}_1(z) = \det P_1(z) = 1$, which means in particular that all these matrices are non-singular.

Finally, we notice that the 4th condition is satisfied due to the known behaviour of the modified Bessel functions as $\zeta \to 0$ and the formulas (69), (71), and (74).

### 2.5 Local parametrix at $z = -1$

#### 2.5.1 The case of GUE. Soft edge

The treatment of the edge $z = -1$ virtually copies that of $z = 1$, we are only draw attention to important details. Consider the situation in which $z = -1$ is the soft edge. The equilibrium measure vanishes like $|z + 1|^{1/2}$ as $z \to -1$, and particles can escape to the left of $-1$. This corresponds to the case of GUE.

Let $\Omega_{-1}$ be a small neighbourhood of $z = -1$ such that $\chi(x)$ is analytic in $\Omega_{-1} \setminus L$. Consider the Riemann–Hilbert problem (P$_{-1}$-RH):

1. $P_{-1}(z)$ is analytic in $\Omega_{-1} \setminus L$;
2. $P^+_{-1}(z) = P^-_{-1}(z) J_T(z)$, $z \in L \cap \Omega_{-1}$, where $J_T(z)$ is defined by (36);
3. $P_{-1}(z)(N(z))^{-1} = I + O(1/n)$ as $n \to \infty$, uniformly on $\partial \Omega_{-1}$;
4. $P_1(z)$ is bounded at $z = -1$.

The function $P_{-1}(z)$ satisfies the same jump conditions and has the same local behaviour as $T(z)$, and we again are going to match $P_{-1}(z)$ with the global parametrix $N(z)$ on the boundary $\partial \Omega_{-1}$ as $n \to \infty$.

Perform the simple change of variables that turns the jump matrix $J_T(z)$ into a piecewise constant one:

\[
\tilde{P}_{-1}(z) = P_{-1}(z)e^{-\frac{x}{2}\phi(z)\sigma_3}(\chi(z))^{\sigma_3/2}, \quad z \in \Omega_{-1} \setminus L.
\]

The function $\tilde{P}_{-1}(z)$ satisfies the following Riemann–Hilbert problem ($\tilde{P}_{-1}$-RH):
1. \( \tilde{P}_{-1}(z) \) is analytic in \( \Omega_{-1} \setminus L \);

2. \( \tilde{P}^+_{-1}(z) = \tilde{P}^-_{-1}(z)J_{-1}(z) \), \( z \in L \cap \Omega_{-1} \),

\[
J_{-1}(z) = \begin{cases}
\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ \\ 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ -1 & 0 \\
\end{bmatrix}, & z \in (-\infty, -1) \cap \Omega_{-1}, \\
\end{cases}
\] (81)

3. \( \tilde{P}_{-1}(z)(\chi(z))^{-\sigma_3/2}e^{\frac{n}{2}\phi(z)}(N(z))^{-1} = I + \mathcal{O}(1/n) \), uniformly on \( \partial\Omega_{-1} \);

4. \( \tilde{P}_{-1}(z) \) is bounded at \( z = -1 \).

We construct \( \tilde{P}_{-1}(z) \) in terms of the same Airy parametrix (55) as in Section 2.4.1, however since the contour in the vicinity of \( z = 1 \) is different from that in the vicinity of \( z = -1 \), we need to transform the parametrix. Let us inverse the orientation on the contour in Fig. 2. It is easy to check that

\[
\tilde{A}(\zeta) := \sigma_3 A(\zeta) \sigma_3
\] (82)

satisfies the same jump conditions (54) on such a contour.

In the same way as before we find the conformal map \( \xi_n(z) \) from \( \Omega_{-1} \) onto some neighbourhood of \( \zeta = 0 \) (see Fig. 2) and construct the solution to \( \tilde{P}_{-1} \)-RH in the form

\[
\tilde{P}_{-1}(z) = E_n(z)\tilde{A}(\xi_n(z)),
\] (83)

where \( E_n(z) \) is some analytic function in \( \Omega_{-1} \). It is clear that if the conformal map \( \xi_n(z) \) is found, then (83) satisfies 1st, 2nd, and 4th conditions of (\( \tilde{P}_{-1} \)-RH) automatically. We choose \( E_n(z) \) so that the 3rd condition is also satisfied.

It is convenient to introduce the function

\[
\tilde{\phi}(z) = -4 \int_{-1}^{z} \sqrt{s^2 - 1} \, ds = \begin{cases}
\phi(z) + 2\pi i, & \text{Im} \, z > 0, \\
\phi(z) - 2\pi i, & \text{Im} \, z < 0,
\end{cases}
\] (84)

where the branch of \( \sqrt{s} \) is chosen so that \( \sqrt{s^2 - 1} > 0 \) when \( s = -2 \). Let

\[
\xi_n(z) = (3n/4)^{2/3}(\tilde{\phi}(z))^{2/3},
\] (85)

where the right-hand side is analytically continued to the analytic function in \( \Omega_{-1} \) using the principal branch of the power function. Expanding the function \( \tilde{\phi}(z) \) given by (84) in a series in the neighbourhood of the branching point \( z = -1 \), we immediately see that

\[
\xi_n(z) = n^{2/3}(z + 1)G(z),
\] (86)
for some analytic in this neighbourhood function $G(z)$ such that $G(-1) \neq 0$.

The identity (80) shows that $\xi_n(z)$ is indeed a conformal map of the neighbourhood of $z = -1$, which without loss of generality is $\Omega_{-1}$, to some neighbourhood $\tilde{\Omega}$ of $z = 0$. Also, we note that $\xi_n(z)$ maps $(-\infty, -1) \cap \Omega_{-1}$ and $(-1, 1) \cap \Omega_{-1}$ in the $z$-plane into $(0, +\infty) \cap \tilde{\Omega}$ and $(\infty, 0) \cap \tilde{\Omega}$ in the $\zeta$-plane, preserving the orientation. Taking into account the angle-preserving property of a conformal map and freedom to deform the lips, one sees that the resulting contour looks like the one in Fig. 2 with the opposite orientation. That is the reason why we introduced $\tilde{A}(\zeta)$ instead of $A(\zeta)$!

Now we notice that the following identity holds

$$e^{\frac{n}{2}\phi(z)} = (-1)^n e^{2/3(n(z))^{3/2}}. \quad (87)$$

Clearly, if we fix $z$ and let $n \to \infty$, then $\xi_n(z) \to \infty$. Therefore, using the formulas (80), (56), (83), and (85), the left part of the matching condition becomes

$$P_{-1}(z)(N(z))^{-1} = \frac{(-1)^n E_n(z)}{\sqrt{2}} (3n\bar{\phi}(z)/4)^{-\sigma_3/6} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \times \left( I + O((\xi_n(z))^{-3/2}) \right) (\chi(z))^{-\sigma_3/2}(N(z))^{-1}. \quad (88)$$

We choose

$$E_n(z) = \frac{(-1)^n N(z)}{\sqrt{2}} (\chi(z))^{\sigma_3/2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} (3n\bar{\phi}(z)/4)^{\sigma_3/6}. \quad (89)$$

The formula (83) becomes

$$P_{-1}(z)(N(z))^{-1} = I + N(z)(\chi(z))^{\sigma_3/2}O((\xi_n(z))^{-3/2})(\chi(z))^{-\sigma_3/2}(N(z))^{-1}. \quad (90)$$

And since $N(z)$ and $\chi(z)$ are uniformly bounded and

$$O((\xi_n(z))^{-3/2}) = O(1/n) \quad (91)$$

uniformly in $z \in \partial \Omega_1$, the matching condition $P_{-1}(z)(N(z))^{-1} = I + O(1/n)$ is satisfied, as desired.

The analyticity of $E_n(z)$ in $\Omega_{-1}$ follows from the argument in Section 2.4.1 mutatis mutandis. Our final observation is that since $\det N(z) = 1$, one has $\det E_n(z) = 1$, and thus $\det \tilde{\bar{P}}_{-1}(z) = \det P_{-1}(z) = 1$. Particularly, all of these matrices are non-singular.

### 2.5.2 The case of LUE and JUE. Hard edge

Consider the situation in which $z = -1$ is the hard edge. The equilibrium measure blows up like $|z + 1|^{-1/2}$ as $z \to -1$, and particles cannot escape from to the left of $-1$. This corresponds to the case of LUE and JUE.

Let $\Omega_{-1}$ be a neighbourhood of $z = -1$ as before. Now we give a construction of the local parametrix at $z = -1$, which is very similar to that in Section 2.4.2.

Consider the Riemann–Hilbert problem (P-1-RH) from the previous section, however we change the 4th condition:
4. The behaviour of $P_{-1}(z)$ as $z \to -1$ is

\[
P_{-1}(z) = \begin{bmatrix}
O(1) & O(1) + O(|z + 1|^\alpha) \\
O(1) & O(1) + O(|z + 1|^\alpha)
\end{bmatrix}, \quad \alpha \neq 0,
\]

\[
\begin{bmatrix}
O(1) & O(\log |z + 1|) \\
O(1) & O(\log |z + 1|)
\end{bmatrix}, \quad \alpha = 0.
\]  

(92)

It is convenient to introduce the function (cf. (3) for LUE and JUE)

\[
\tilde{\omega}(z) = \begin{cases}
\alpha \log (-z - 1), & \text{for LUE, } \ z \in \mathbb{C} \setminus [-1, +\infty)
\end{cases},
\]

with the principal branch of log used and to define $\tilde{\chi}(z) = e^{i\tilde{f}(z)+\tilde{\omega}(z)}$. The latter is clearly analytic in $\Omega_{-1} \setminus (-1, 1)$. Then, we perform the change of variables

$$
\tilde{P}_{-1}(z) = P_{-1}(z)e^{-\frac{2}{n}\phi(z)\sigma_3}(\tilde{\chi}(z))^{\sigma_3/2}, \quad z \in \Omega_{-1} \setminus L,
$$

(94)

to reduce $(P_{-1}-\text{RH})$ to the problem $(\tilde{P}_{-1}-\text{RH})$ with the piecewise continuous jump matrix:

1. $\tilde{P}_{-1}(z)$ is analytic in $\Omega_{-1} \setminus L$;

2. $\tilde{P}_{-1}^+(z) = \tilde{P}_{-1}^-(z)J_{-1}(z)$, $z \in L \cap \Omega_{-1}$,

\[
J_{-1}(z) = \begin{cases}
\begin{bmatrix}
1 & 0 \\
e^{-\pi i \alpha} & 1
\end{bmatrix}, & z \in L_+ \cap \Omega_{-1},
\end{cases}
\]

\[
\begin{cases}
\begin{bmatrix}
1 & 0 \\
e^{\pi i \alpha} & 1
\end{bmatrix}, & z \in L_- \cap \Omega_{-1},
\end{cases}
\]

\[
\begin{cases}
0 & 1 \\
-1 & 0
\end{cases}, & z \in (-1, 1) \cap \Omega_{-1};
\]

(95)

3. $\tilde{P}_{-1}(z)(\tilde{\chi}(z))^{-\sigma_3/2}e^{\frac{2}{n}\phi(z)\sigma_3}(N(z))^{-1} = I + O(1/n)$ as $n \to \infty$, uniformly on $\partial \Omega_{-1}$;

4. The behaviour of $\tilde{P}_{-1}(z)$ as $z \to -1$:

\[
\tilde{P}_{-1}(z) = \begin{cases}
O(1)|z + 1|^\alpha, & \alpha > 0,
O(|z + 1|^\alpha), & \alpha < 0,
\begin{bmatrix}
O(1) & O(\log |z + 1|) \\
O(1) & O(\log |z + 1|)
\end{bmatrix}, & \alpha = 0.
\end{cases}
\]  

(96)

We construct $\tilde{P}_{-1}(z)$ in terms of the same Bessel parametrix (69) as in Section 2.4.2, however since the contour in the vicinity of $z = 1$ is different from that in the vicinity
of $z = -1$, we need to do the transformation of the parametrix. Let us inverse the orientation on the contour in Fig. 3. It is easy to check that
\[
\tilde{\Psi}(\zeta) := \sigma_3 \Psi(\zeta) \sigma_3
\]  
(97)
satisfies the same jump conditions as $\Psi(\zeta)$ on such a contour.

As it was in the previous section we find the conformal map $\eta_n(z)$ from $\Omega_{-1}$ onto some neighbourhood of $\zeta = 0$ (see Fig. 3) and construct the solution to $\tilde{P}_{-1}$-RH in the form
\[
\tilde{P}_{-1}(z) = E_n(z) \tilde{\Psi}(\eta_n(z)),
\]  
(98)
where $E_n(z)$ is some analytic function in $\Omega_{-1}$.

It is convenient to introduce the function
\[
\tilde{\phi}(z) = \begin{cases} 
-2 \int_{-1}^{z} \sqrt{\frac{s-1}{s+1}} \, ds, & \text{for LUE}, \\
-2 \int_{-1}^{z} \frac{1}{\sqrt{s^2-1}} \, ds, & \text{for JUE},
\end{cases}
\]  
(99)
where the branch of $\sqrt{\cdot}$ is chosen so that the integrands are positive when $s = -2$. Let
\[
\eta_n(z) = \frac{n^2}{16} (\tilde{\phi}(z))^2.
\]  
(100)
Expanding $\tilde{\phi}(z)$ in a series in the neighbourhood of $z = -1$, one can find that
\[
\eta_n(z) = \frac{n^2}{16} (z+1) G(z)
\]  
(101)
for some analytic function $G(z)$ which satisfies $G(-1) \neq 0$. Thus, $\eta_n(z)$ is a conformal map of the neighbourhood of $z = -1$, which without loss of generality is $\Omega_{-1}$, to some neighbourhood $\tilde{\Omega}$ of $\zeta = 0$. We also note that $\eta_n(z)$ maps $(-1, 1) \cap \Omega_{-1}$ into $(-\infty, 0) \cap \tilde{\Omega}$ in the $\zeta$-plane, preserving orientation. Taking into account the angle-preserving property of a conformal map and freedom to deform the lips, one sees that the resulting contour looks like the one in Fig. 3 with the opposite orientation. That is the reason why we introduced $\tilde{\Psi}(\zeta)$ instead of $\Psi(\zeta)$! Moreover $L_+$ and $L_-$ are mapped into $\tilde{L}_+$ and $\tilde{L}_-$, respectively, and thus the jumps over $L_{\pm}$ of $\Psi(\eta_n(z))$ are consistent with the jumps of $\tilde{P}_{-1}(z)$ (see (95)).

Now we notice that $\eta_n(z) \to \infty$ as $n \to \infty$, and thus the asymptotics for $\Psi(\eta_n(z))$ as $z \to \infty$ is of relevance. It is readily verified that
\[
e^{2\sqrt{\eta_n(z)}} = (-1)^n e^{-\frac{\pi i}{2} \tilde{\phi}(z)},
\]  
(102)
and taking into account (94), (70), (97), (98), (100), one can write the 3rd (matching) condition on $\partial \Omega_{-1}$ to be
\[
P_{-1}(z)(N(z))^{-1} = E_n(z)(\pi n \tilde{\phi}(z)/2)^{-\sigma_3/4}(-1)^n \sqrt{2} \left[ \begin{array}{cc} 1 & -i \\ -i & 1 \end{array} \right] \\
\times (1 + \mathcal{O}(\eta_n^{-1/2}(z)))(\tilde{\chi}(z))^{-\sigma_3/2}(N(z))^{-1}.
\]  
(103)
We choose
\[ E_n(z) = (-1)^n N(z)(\tilde{\chi}(z))^{\frac{\sigma_3}{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} (\pi n \tilde{\phi}(z)/2)^{\frac{\sigma_3}{2}} \] (104)
and rewrite (103) as
\[ P_{-1}(z)(N(z))^{-1} = I + (N(z))(\tilde{\chi}(z))^{\frac{\sigma_3}{2}} \mathcal{O}(\eta_n^{-1/2}(z))(\tilde{\chi}(z))^{-\frac{\sigma_3}{2}}(N(z))^{-1}. \] (105)
Since \( N(z) \) and \( \tilde{\chi}(z) \) are uniformly bounded and
\[ \mathcal{O}(\eta_n(z))^{-1/2) = \mathcal{O}(1/n) \] (106)
uniformly in \( z \in \partial \Omega_0 \), one finally arrives at \( P_{-1}(z)(N(z))^{-1} = I + \mathcal{O}(1/n) \) uniformly in \( z \in \partial \Omega_0 \), as desired.

The argument similar to the one in Section 2.4.2 shows that \( E_n(z) \) is analytic in \( \Omega_0 \).
Finally, we observe that since \( \det N(z) = 1 \), one has \( \det E_n(z) = 1 \), and thus \( \det \tilde{P}_{-1}(z) = \det P_{-1}(z) = 1 \). Particularly, all of these matrices are non-singular.

### 2.6 Final transformation. Small norm problem

Now we have everything ready to write the small norm problem. Let the function \( R(z) \) be
\[ R(z) = \begin{cases} 
T(z)(P_{-1}(z))^{-1}, & z \in \Omega_{-1}, \\
T(z)(P_{1}(z))^{-1}, & z \in \Omega_{1}, \\
T(z)(N(z))^{-1}, & z \in \mathbb{C} \setminus (\Omega_{-1} \cup \Omega_{1}).
\end{cases} \] (107)

By construction, \( R(z) \) only have jumps over the contour in Fig. 4. And it follows from the previous consideration that \( R(z) \) satisfies the following Riemann–Hilbert problem (R-RH):

![Contour Diagram](image)

Figure 4: The contour \( \Sigma_R \) of the small norm problem.
1. $R(z)$ is analytic in $\mathbb{C} \setminus \Sigma_R$;

2. $R^+(z) = R^-(z) J_R(z)$, where

$$J_R(x) = \begin{cases} 
P_1(z)(N(z))^{-1}, & z \in \partial \Omega_1, \\
P_{-1}(z)(N(z))^{-1}, & z \in \partial \Omega_{-1}, \\
N(z)J_T(z)(N(z))^{-1}, & z \in L_+ \cup L_-, \\
N(z)J_U(z)(N(z))^{-1}, & z \in \Sigma_R \setminus (\partial \Omega_{-1} \cup \partial \Omega_1 \cup L_+ \cup L_-),
\end{cases} \quad (108)$$

3. $R(z) = I + \mathcal{O}(1/z)$ as $z \to \infty$;

4. $R(z)$ is bounded if $z$ is approaching the points of the self-intersection of $\Sigma_R$.

By construction of $P_1(z)$, $P_{-1}(z)$ and since $N(z)$ is uniformly bounded, one has $J_R(z) = I + \mathcal{O}(1/n)$ uniformly on the contour $\Sigma_R$, as $n \to \infty$. Consequently, from the theory of small norm problems (for the details see [1]) it follows that (R-RH) has a unique solution for large enough $n > n_0$. Also the following takes place

$$R(z) = I + \mathcal{O}(1/n), \quad R'(z) = \mathcal{O}(1/n) \quad (109)$$

uniformly in $\mathbb{C} \setminus \Sigma_R$.

As the final remark we add that the fact itself that we were able to seam together the local and global parametrices justifies their choice, which, we remind, was not unique.

## 3 Proof of Lemma 2

### 3.1 Deformation of the weight

We prove the lemma for $f(x)$ such that it is locally Hölder continuous on $\mathcal{I}$ and admits the analytic continuation to some neighbourhood of $[-1, 1]$.

We start off by fixing a number $h \in \mathbb{R}$ and noticing that for all $t \in [0, 1]$ the function $(1-t) + te^{ih} f(x)$ is analytic with respect to $z$ in some (simply-connected) neighbourhood of $[-1, 1] \subset \mathbb{C}$. Therefore, one can choose large enough $q$ such that

$$((1-t) + te^{ih} f(x)) \neq 0 \quad (110)$$

in this neighbourhood. Now we follow an idea from [12] and apply the results of Section 2 to $\tilde{f}(x) := \tilde{f}_{l,t}(x)$, where

$$\tilde{f}_{l,t}(x) = \log((1-t) + te^{ih} f(x)) + \frac{ih(l-1)}{q} f(x), \quad t \in [0, 1], \ l = 1, \ldots, q. \quad (111)$$

We use the principal branch of log, and because of (110) the right-hand side of (111) is well-defined.
Notice that \( \tilde{f}_{l,t}(x) \) satisfies the assumption (A2) and (A4) of Section 2; however, the assumption (A3) is not necessarily true. Nevertheless, one notices that due to (16) the function \( H_{k,n}[\tilde{f}_{l,t}] \) is analytic in \( t \), consequently there is as many as a finite number of points \( t \) where this determinant vanishes for \( k = 1, \ldots, n \). If we denote this set by \( \mathcal{T}_0(n) \), then (A3) is satisfied for \( t \in [0, 1] \setminus \mathcal{T}_0 \).

The assumption (A1) is also not necessarily satisfied, since \((1 - t) + \text{ie} \frac{\theta}{\pi} f(x)\) can vanish at some \( x > 1 \). However, we notice that the latter is only possible if \( t = 1/2 \), and we overcome this problem by including \( t = 1/2 \) in \( \mathcal{T}_0(n) \). Finally we see that (A1) and (A4) are satisfied, if \( t \in [0, 1] \setminus \mathcal{T}_0(n) \).

Next ingredient of the proof is a special differential identity (see [1, 4, 5, 12, 21])

\[
\frac{\partial}{\partial t} \log H_{n,n}[\tilde{f}_{l,t}] = \frac{1}{2\pi i} \int_{\gamma} \left[ Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x) \right]_{2,1} \frac{\partial}{\partial t} \tilde{w}_{n,l,t}(x) \, dx
\]

following from the straightforward formula

\[
H_{n,n}[\tilde{f}_{l,t}] = \prod_{k=1}^{n-1} \gamma_{k,n,l,t}^{-2}
\]

From now on we will be adding the symbols \( l \) and \( t \) to all the relevant quantities of (Y-RH), since we substitute \( \tilde{f}_{l,t} \mapsto \tilde{f} \).

We point out that by integrating (112) over \([0, 1]\) with respect to \( t \), one would have \( \log H_{n,n}[\frac{\text{ie} \frac{\theta}{\pi} f(x)}{q}] - \log H_{n,n}[\frac{\text{ie} \frac{(\theta - 1)}{\pi} f(x)}{q}] \) in the left-hand side of (112). So, to find the asymptotics of \( \mathbb{E}_n[e^{\text{ie} \text{Tr} f(M)}] \) it is enough to sum up such quantities with respect to \( l \) from 1 to \( q \).

3.2 Integration of the differential identity

Now we fix \( n > n_0 \) so that (R-RH) is uniquely solvable, and suppose \( t \in [0, 1] \setminus \mathcal{T}_0(n) \). Our goal is to use the asymptotics of the Riemann–Hilbert problem and integrate the differential identity (112). A very similar problem has been studied in [1, 4, 5], and we adopt ideas from there for our needs. First, we have to break up the contour of the integration, since the asymptotics differ inside and outside the neighbourhood of \([-1, 1]\):

\[
\frac{\partial}{\partial t} \log H_{n,n}[\tilde{f}_{l,t}] = \frac{1}{2\pi i} \left( \int_{\mathcal{I} \setminus \mathcal{I}_\varepsilon} + \int_{\mathcal{I} \cap \mathcal{I}_\varepsilon} \right) \left[ Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x) \right]_{2,1} \frac{\partial}{\partial t} \tilde{w}_{n,l,t}(x) \, dx
\]

where \( \mathcal{I}_\varepsilon = [-1 + \varepsilon, 1 + \varepsilon] \).

Now, since \( \mathcal{I} \setminus \mathcal{I}_\varepsilon \) is away from \([-1, 1]\) (see Fig. 5), it is possible to use the global parametrix \( N(z) \) to calculate the integral along \( \mathcal{I} \setminus \mathcal{I}_\varepsilon \).

On the other hand, to calculate the integral along \( \mathcal{I} \cap \mathcal{I}_\varepsilon \) in an easier way, we would like to employ the idea of the contour deformation. First, we note that

\[
\left[ Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x) \right]_{2,1} \frac{\partial}{\partial t} \tilde{w}_{n,l,t}(x) = \left( \left[ Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x) \right]_{1,1} - \left[ Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x) \right]_{1,1}^+ \right) \frac{\partial}{\partial t} \tilde{f}_{l,t}(x),
\]
Figure 5: The contour $\Sigma_R$ of the small norm problem.

which easily follows from a direct calculation using the jump condition of the Riemann–Hilbert problem for $Y(z)$. And then we extend the contour $\mathcal{I} \cap \mathcal{I}_\varepsilon$ to $\mathcal{I}$ so that the deformed contour is far enough from $[-1,1]$. We underline that the right-hand side of (115) suits our needs better. Indeed, in the case LUE and JUE the term $\frac{\partial}{\partial t}\tilde{w}_{n,l,t}(x)$ cannot be analytically continued onto $\mathcal{I}_\varepsilon$, on the other hand $\frac{\partial}{\partial t}\tilde{f}_{l,t}(x)$ can be.

By the contour deformation argument one has:

$$\frac{1}{2\pi i} \int_{\mathcal{I}_\varepsilon} \left( [Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x)]_{1,1} - [Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x)]_{1,1}^+ \right) \frac{\partial}{\partial t} \tilde{f}_{l,t}(x) \, dx$$

$$= \frac{1}{2\pi i} \int_{\mathcal{I}_\varepsilon \setminus \mathcal{I}} \left( [Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x)]_{1,1}^+ - [Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x)]_{1,1}^- \right) \frac{\partial}{\partial t} \tilde{f}_{l,t}(x) \, dx$$

$$- \frac{1}{2\pi i} \left( \int_{\tau_+} \left[ Y_{n,l,t}^{-1}(z)Y'_{n,l,t}(z) \right]_{1,1}^+ \frac{\partial}{\partial t} \tilde{f}_{l,t}(z) \, dz - \int_{\tau_-} \left[ Y_{n,l,t}^{-1}(z)Y'_{n,l,t}(z) \right]_{1,1}^- \frac{\partial}{\partial t} \tilde{f}_{l,t}(z) \, dz \right),$$

(116)

where $\tau_\pm$ are defined in Fig. 5

Clearly, the integral along $\mathcal{I}_\varepsilon \setminus \mathcal{I}$ is zero for $[Y_{n,l,t}^{-1}(z)Y'_{n,l,t}(z)]_{1,1}$ is continuous over $\mathcal{I}_\varepsilon \setminus \mathcal{I}$. Consequently, we only need to find the integrals along $\tau_+$ and $\tau_-$. For which it is again possible to use the global parametrix $N(z)$, since the contour of integration is away from $[-1,1]$.

We proceed by considering the quantity $Y_{n,l,t}^{-1}(z)Y'_{n,l,t}(z)$ for which a direct calculation using (30), (35), and (107), shows that

$$Y_{n,l,t}^{-1}(z)Y'_{n,l,t}(z) = n'g(z)\sigma_3 + e^{-n(l/2+g(z))\sigma_3}N_{t,l}^{-1}(x)N'_{t,l}(z)e^{n(l/2+g(z))\sigma_3}$$

$$+ e^{-n(l/2+g(z))\sigma_3}N_{t,l}^{-1}(z)R_{l,t}^{-1}(z)R'_{l,t}(z)N_{t,l}(z)e^{n(l/2+g(z))\sigma_3}.$$

(117)

According to (30), the global parametrix $N_{t,l}(z)$ and its inverse $N_{t,l}^{-1}(z)$ are bounded uniformly in $z \in \mathcal{I} \setminus \mathcal{I}_\varepsilon$, in $l = 1, \ldots, q$, and in $t$ (see Fig. 5). Hence, by (109) we get

$$N_{t,l}^{-1}(z)R_{l,t}^{-1}(z)R'_{l,t}(z)N_{t,l}(z) = \mathcal{O}(1/n)$$

(118)

as $n \to \infty$, where the $\mathcal{O}$-term is uniform in all the parameters and $z$. 

29
Now in the usual way (see [8, 12]) we extend the differential identity (112) for all \( t \in [0, 1] \). Let us introduce the function

\[
S_{n,l}(t) = H_{n,n} \bar{f}_{l,t}(e^{-\int_0^\infty r_{n,l}(s) ds},
\]

where \( r_{n,l}(t) \) is the right-hand side of (112). Later it will be clear from (121) and (126) that for \( n > n_0 \) the function \( r_{n,l}(t) \) is continuous in \( t \), and thus \( S_{n,l}(t) \) is continuously differentiable. Then, the differential identity can be rewritten simply as \( \frac{\partial}{\partial t} S_{n,l}(t) = 0 \) for \( t \in [0, 1] \setminus T_0(n) \). But then \( \frac{\partial}{\partial t} S_{n,l}(t) = 0 \) for all \( t \in [0, 1] \). Therefore, \( S_{n,l}(t) \) is constant in \( t \).

Now we provide the argument which shows that (112) holds for all \( t \in [0, 1] \) and \( l = 1, \ldots, q \). First, directly from (16) we notice that \( \tilde{f}_{l,0} = 0 \) and \( H_{n,n}[\tilde{f}_{l,0}] \neq 0 \). Hence, since \( S_{n,l}(t) \) is constant in \( t \), we have that \( S_{n,1}(t) \neq 0 \) for all \( t \in [0, 1] \). Thus, \( H_{n,n}[\tilde{f}_{l,1}] \neq 0 \), and (112) holds in fact for all \( t \in [0, 1] \) and \( l = 1 \). Particularly, since \( \tilde{f}_{l-1,1} = \tilde{f}_{l,0} \), we have \( H_{n,n}[\tilde{f}_{l,0}] \neq 0 \). Then we can repeat the whole procedure to find out that (112) holds for all \( t \in [0, 1] \) and for \( l = 2 \). The proof is concluded by induction.

We proceed with finding the terms of asymptotics. Substituting (51) into (117) one finds that

\[
[Y_{n,l,t}(z)Y'_{n,l,t}(z)]_{2,1} = -e^{n(l+2g(x))}\left(\frac{1}{4iD_l^2(z)z(z-1)} + O(1/n)\right)
\]

and

\[
[Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x)]_{2,1} \frac{\partial}{\partial t} \tilde{w}_{n,l,t}(x) = -e^{\omega(x)} \left( e^{i\omega(x)} f(x) - e^{i\omega(x)-i\theta} f(x) \right)
\]

\[
\times \left( \frac{1}{4iD_l^2(x)x(x-1)} + O(1/n) \right) e^{n(l+2g(x)-4x)}. \tag{121}
\]

Therefore, taking into account (29) and (28), one arrives at

\[
\int_{1+\epsilon}^{\infty} [Y_{n,l,t}^{-1}(x)Y'_{n,l,t}(x)]_{2,1} \frac{\partial}{\partial t} \tilde{w}_{n,l,t}(x) dx = O(e^{-Cn}) \tag{122}
\]

for some \( C > 0 \) as \( n \to \infty \). Besides, \( C \) does not depend on both \( l \) and \( t \), and \( O \)-term is uniform in \( l \) and \( t \).

Evaluating the integrals along \( \tau_+ \) and \( \tau_- \) takes somewhat more effort. First, from (117) we notice that

\[
[Y_{n,l,t}^{-1}(z)Y'_{n,l,t}(z)]_{1,1} = ng'(z) - \frac{D_l'(z)}{D_l(z)} + O(1/n). \tag{123}
\]

Then, integrating the first term and applying the contour deformation argument one sees that

\[
-\frac{1}{2\pi i} \left( \int_{\tau_+} - \int_{\tau_-} \right) g'(z) \frac{\partial}{\partial t} \tilde{f}_{l,t}(z) dz = \int_{-1}^{1} \psi(x) \frac{\partial}{\partial t} \tilde{f}_{l,t}(x) dx, \tag{124}
\]

30
which gives, after being integrated with respect to $t$ along $[0,1]$, the leading term of the asymptotics:

$$\int_0^1 \left[ \int_0^1 \psi(x) \frac{\partial}{\partial t} \tilde{f}_{l,t}(x) \, dx \right] \, dt = \frac{i\hbar}{q} \int_{-1}^1 f(x) \psi(x) \, dx = \frac{i\hbar}{q} \mathcal{Z}[f].$$

(125)

Now in order to find the next term of asymptotics, we write

$$\frac{1}{2\pi i} \left( \int_{\tau_+}^{\tau_-} - \int_{\tau_-}^{\tau_+} \right) \frac{D_{l,t}^1(z)}{D_1(z)} \frac{\partial}{\partial t} \tilde{f}_{l,t}(x) \, dz = \frac{1}{2\pi i} \left( \int_{\tau_+}^{\tau_-} - \int_{\tau_-}^{\tau_+} \right) \left( \frac{D_{l,t}^1(z)}{D_1(z)} + \frac{D_{2,l,t}^1(z)}{D_2(z)} \right) \frac{\partial}{\partial t} \tilde{f}_{l,t}(x) \, dz.

(126)

A direct calculation shows that

$$D_{l,t}^1(z) = \begin{cases} 0, & \text{for GUE}, \\ \frac{\alpha}{2} \left( \frac{1}{z+1} - \frac{1}{\sqrt{z^2-1}} \right), & \text{for LUE}, \\ \frac{\alpha}{2} \left( \frac{1}{z+1} - \frac{1}{\sqrt{z^2-1}} \right) + \frac{\beta}{2} \left( \frac{1}{z-1} - \frac{1}{\sqrt{z^2-1}} \right), & \text{for JUE}. \end{cases}

(127)

Then, the fact that $D_1(z)$ does not depend on $t$ and the standard argument of the contour deformation lead us to the identity

$$\int_0^1 \left[ \frac{1}{2\pi i} \left( \int_{\tau_+}^{\tau_-} - \int_{\tau_-}^{\tau_+} \right) \frac{D_{l,t}^1(z)}{D_1(z)} \frac{\partial}{\partial t} \tilde{f}_{l,t}(z) \, dz \right] \, dt = \frac{\hbar}{2\pi q} \left( \int_{\tau_+}^{\tau_-} - \int_{\tau_-}^{\tau_+} \right) \frac{D_{l,t}^1(z)}{D_1(z)} f(z) \, dz = \frac{i\hbar}{q} \mu[f].

(128)

Now we move on to calculate the other integrals in (126). To shorten the notation, define

$$\theta_{l,t}(z) = \theta_{l,t}^{(1)}(z) + \theta_{l,t}^{(2)}(z) = \frac{1}{2\pi} \sqrt{z^2-1} \left( 1 \int_{-1}^1 \frac{\log \left( (1-t) + te^{i\hbar f(z)} \right)}{\sqrt{1-x^2}} \, dx \right) \frac{1}{z-x}$$

$$+ \frac{i\hbar(l-1)}{2\pi q} \sqrt{z^2-1} \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} \, dx \frac{1}{z-x}.$$

(129)

The problem is to find the integral

$$\int_0^1 \left[ \frac{1}{2\pi i} \left( \int_{\tau_+}^{\tau_-} - \int_{\tau_-}^{\tau_+} \right) \frac{D_{2,l,t}^1(z)}{D_2(z)} \frac{\partial}{\partial t} \tilde{f}_{l,t}(z) \, dz \right] \, dt = I^{(1)} + I^{(2)},

(130)

where

$$I^{(j)} = \int_0^1 \left[ \frac{1}{2\pi i} \left( \int_{\tau_+}^{\tau_-} - \int_{\tau_-}^{\tau_+} \right) \theta_{l,t}^{(j)}(z) \frac{\partial}{\partial t} \tilde{f}_{l,t}(z) \, dz \right] \, dt.$$

(131)
The result for $I^{(1)}$ follows immediately from [11, Lemma 5.4], if one notice that the last term in $\tilde{f}_{l,t}(z)$ does not depend on $t$:

$$I^{(1)} = -\frac{h^2}{4\pi^2 q^2} \int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}} \text{v.p.} \int_{-1}^{1} \frac{f'(y)\sqrt{1-y^2}}{x-y} \, dy \, dx = -\frac{h^2}{2q^2} K[f]. \quad (132)$$

To find the other integral $I^{(2)}$, we use the fact that $\theta_{l,t}^{(2)}(z)$ not depend on $t$ and perform integration with respect to $t$ first.

$$I^{(2)} = \frac{h}{2\pi q} \left( \int_{\tau_+} - \int_{\tau_-} \right) (\theta_{l,t}^{(2)}(z))' f(z) \, dz. \quad (133)$$

After that, we notice that $(\theta_{l,t}^{(2)}(z))'$ has integrable singularities at $z = -1$ and $z = 1$ and use the contour deformation argument, which leads to

$$I^{(2)} = \frac{h}{2\pi q} \int_{-1}^{1} (\theta_{l,t}^{(2)+}(y) - \theta_{l,t}^{(2)-}(y))' f(y) \, dy. \quad (134)$$

Now integrating by parts, using the Sokhotski–Plemelj formulas, and switching the order of integration by properties of the Hilbert transform (e.g., see [27]), we find that

$$I^{(2)} = -\frac{h^2(l-1)}{2\pi^2 q^2} \int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}} \text{v.p.} \int_{-1}^{1} \frac{f'(y)\sqrt{1-y^2}}{x-y} \, dy \, dx = -\frac{h^2(l-1)}{q^2} K[f]. \quad (135)$$

Finally, integrating the last term $O(1/n)$ in (133), uniform in $z$, and collecting everything together, we get to the desired asymptotic formula:

$$\log \frac{H_{n,n} \left[ \frac{i\hbar f}{q} \right]}{H_{n,n} \left[ \frac{i\hbar(l-1)}{q} f \right]} = \frac{i\hbar}{q} (n\varepsilon[f] + \mu[f]) - \frac{h^2(2l-1)}{2q^2} K[f] + O \left( \frac{1}{n} \right). \quad (136)$$

Summing up with respect to $l$ from 1 to $q$, we arrive at the expected formula

$$\mathbb{E}_{n} [e^{i\hbar \text{Tr} f(M)}] = i\hbar (n\varepsilon[f] + \mu[f]) - \frac{h^2}{2} K[f] + O \left( \frac{1}{n} \right). \quad (137)$$

The final part of the proof is to validate that (3) can be written as (11). This has been done in [20, p. 172].
4 Proof of Lemma 1

The idea of the proof is to allow for large $l$ (depending on $n$) in (136) by employing a special deformation of $f(x)$, more general than (111). It will also turn out that the corresponding $O$-term is uniform in $l$ and $h$. We obtain the analogue of (136) and sum up for all $l = 1, 2, \ldots$. This will yield the formula (10).

We begin by choosing $\varepsilon > 0$ in such a way that

$$
\tilde{f}_{l,t}(x) = \log \left( (1 - t) + t e^{ih[|l| \leq n^\gamma + 1] f(x)} \right) + ih(l - 1) \mathbb{1}[|l| \leq n^\gamma + 1] f(x), \ l = 1, 2, \ldots \quad (138)
$$

is well-defined for all $h$ such that $|h| < \varepsilon$ and for all $t \in [0, 1]$.

Next, we repeat all the steps of Section 2. Note that $e^{\tilde{f}_{l,t}(x)}$ is bounded, uniformly in all the relevant parameters for $x \in \mathcal{I} \setminus [-1, 1]$. Therefore, $J_U(x)$ in (32) converges to $I$ exponentially and uniformly for $x$ away from $\{-1, 1\} \cap \mathcal{I}$. Due to the last term in the right-hand side of (138), $e^{\tilde{f}_{l,t}(z)}$ is not bounded on $L^\pm$ (see Fig. 1) and can grow to infinity as $l, n \to \infty$. However, from the following simple inequality

$$
|e^{ih(l-1)\mathbb{1}[|l| \leq n^\gamma + 1] f(z)}| \leq e^{\varepsilon(l-1)\mathbb{1}[|l| \leq n^\gamma + 1] |\text{Im} f(z)|} \leq e^{cn^\gamma |\text{Im} f(z)|}
$$

(139)

one understands that, since $\gamma < 1$, this growth is damped by $e^{n\phi^\pm(z)}$ (see (34)). The local parametrices can be constructed in the similar way as before, however, now one ought to be careful since $N_{l,t}(z)$ is no longer bounded in $l, n$ uniformly. Let us proceed with the local parametrix at $z = 1$ and write (62):

$$
P_1(z)(N_{l,t}(z))^{-1} = I + N_{l,t}(z)(\chi_{l,t}(z))^{\sigma_3/2}O(1/(n\phi(z)))(\chi_{l,t}(z))^{-\sigma_3/2}(N_{l,t}(z))^{-1}.
$$

(140)

Note that in the case of the hard edge this expression remains the same up to the change $\chi \mapsto \tilde{\chi}$ (cf. (77)).

Using (50), (38), (31), and (138), one sees (in both of the cases) that there is a factor

$$
\exp \left( \pm ih\sigma_3 (l-1)/2 \mathbb{1}[l \leq n^\gamma + 1] \left( f(z) - \frac{\sqrt{z^2 - 1}}{\pi} \int_{-1}^{1} \frac{f(x) \, dx}{\sqrt{1 - x^2(z-x)}} \right) \right)
$$

(141)

$$
= \exp \left( \pm ih\sigma_3 (l-1)/2 \mathbb{1}[l \leq n^\gamma + 1] \sqrt{z^2 - 1}/\pi \int_{-1}^{1} \frac{f(z) - f(x)}{\sqrt{1 - x^2(z-x)}} \, dx \right)
$$

in (140), which is not bounded in $l, n$. To account for this, we note that the integral is bounded and choose the neighbourhood $\Omega_1^{(n)}$ contracting as $n \to \infty$ at the rate $O(1/n^{2\gamma})$. Taking into account the asymptotics of $\phi(z)$ in (28) as $z \to 1$, one also obtains

$$
O(1/(n\phi(z))) = \begin{cases} 
O(1/n^{1-3\gamma}), & \text{for GUE, LUE}, \\
O(1/n^{1-\gamma}), & \text{for JUE},
\end{cases}
$$

(142)
uniformly in \( z \in \partial \Omega_{1}^{(n)} \).

Recalling the asymptotics \( C(z) = \mathcal{O}((z - 1)^{-1/4}) \) as \( z \to 1 \) from (49), one immediately arrives at

\[
P_{1}(z)(N_{l,t}(z))^{-1} = I + \begin{cases} \mathcal{O}(1/n^{1-4\gamma}), & \text{for GUE, LUE,} \\
\mathcal{O}(1/n^{1-2\gamma}), & \text{for JUE,} \end{cases}
\]

(143)

uniformly in \( z \in \Omega_{1}^{(n)} \), in \( l = 1, 2, \ldots \), and in \( t \in [0, 1] \).

In an analogous way one handles the local parametrix at \( z = -1 \) by introducing the neighbourhood \( \Omega_{-1}^{(n)} \) contracting as \( n \to \infty \) at the rate \( \mathcal{O}(1/n^{2\gamma}) \). Taking into account asymptotics of \( \phi(z) \) and \( C(z) = \mathcal{O}((z + 1)^{-1/4}) \) as \( z \to -1 \), one can show that

\[
P_{-1}(z)(N_{l,t}(z))^{-1} = I + \begin{cases} \mathcal{O}(1/n^{1-4\gamma}), & \text{for GUE,} \\
\mathcal{O}(1/n^{1-2\gamma}), & \text{for LUE, JUE,} \end{cases}
\]

(144)

uniformly in \( z \in \Omega_{-1}^{(n)} \), in \( l = 1, 2, \ldots \), and in \( t \in [0, 1] \).

The use of contracting contour also affects the rate of convergence of \( J_{R}(z) \) to the identity matrix (see (103) on \( L_{+} \cup L_{-} \cup \mathcal{I} \setminus (\Omega_{-1} \cup \Omega_{1}) \) (see Fig. 3). However, it is not difficult to show that this convergence is still exponentially fast \( \mathcal{O}(e^{-Cn^{1-\gamma}}) \), \( C > 0 \), for \( \gamma < 1/4 \) (GUE and LUE); and \( \mathcal{O}(e^{-Cn^{1-\gamma^\prime}}), \ C > 0 \), for \( \gamma < 1/2 \) (JUE). Consequently, the Riemann–Hilbert analysis of Section 2 can be followed through using the analogue of the small norm theory for the varying (contracting) contours (see Appendix in [3]). Finally, one arrives at the asymptotic identities uniform in all the relevant parameters:

\[
R(z) = \begin{cases} I + \mathcal{O}(1/n^{1-4\gamma}), & \text{for GUE, LUE,} \\
I + \mathcal{O}(1/n^{1-2\gamma}), & \text{for JUE,} \end{cases}
\]

\[
R'(z) = \begin{cases} \mathcal{O}(1/n^{1-4\gamma}), & \text{for GUE, LUE,} \\
\mathcal{O}(1/n^{1-2\gamma}), & \text{for JUE,} \end{cases}
\]

(145)

uniformly in \( z \) away from the contour.

Now we repeat all the steps of Section 3. Note that the global parametrix \( N_{l,t}(x) \) is uniformly bounded in \( l, t, \) and \( h \) for \( x \in \mathcal{I} \setminus [-1, 1] \), therefore, one can write the analogue of (118)

\[
N_{l,t}^{-1}(z)R_{l,t}^{-1}(z)R'_{l,t}(z)N_{l,t}(z) = \begin{cases} \mathcal{O}(1/n^{1-4\gamma}), & \text{for GUE, LUE,} \\
\mathcal{O}(1/n^{1-2\gamma}), & \text{for JUE,} \end{cases}
\]

(146)

and repeat the analogous integration as in (122). Unfortunately, the uniform boundedness in \( l \) is lost for \( N_{l,t}(z) \) on the fixed contour \( \tau_{\pm} \) (see Fig. 5), because of the Szegő function (38) involved. Nevertheless, the direct calculation of the left-hand side of (146) shows that (1, 1) element of this matrix is still of order \( \mathcal{O}(1/n^{1-4\gamma}) \) (GUE, LUE), or \( \mathcal{O}(1/n^{1-2\gamma}) \) (JUE), thanks to \( D_{l,t}(z) \) cancel out.

From the condition \( f(x) = \mathcal{O}(e^{Ax}) \) and due to the following straightforward estimate

\[
|e^{ihx}[1 \leq n^{\gamma + 1}]f(x) - e^{ih(x-l)}[1 \leq n^{\gamma + 1}]f(x)| \leq |h||f(x)|, \quad x \in [1 + \varepsilon, +\infty),
\]

(147)

it is immediate that the right hand side of (122) is \( \mathcal{O}(he^{-Cn}) \) uniformly in \( h \).
Due to (146), the last term in (123) becomes $O(1/n^{1-4\gamma})$ (GUE, LUE), or $O(1/n^{1-2\gamma})$ (JUE), uniformly in $h$, and the resulting error after integration is $O(h/n^{1-4\gamma})$ (GUE, LUE) or $O(h/n^{1-2\gamma})$ (JUE), uniformly in $h$. Finally, one arrives at the analogue of (136):

$$\log \frac{H_{n,n}[ihl][l \leq n^{\gamma} + 1]}{H_{n,n}[ih(l-1)][l \leq n^{\gamma} + 1]} = (ih(n\varepsilon[f] + \mu[f]) - \frac{h^2(2l-1)}{2} K[f]$$

$$+ O \left( \frac{h}{n^{1-(d-1)\gamma}} \right) \right) 1[l \leq n^{\gamma} + 1],$$

where $d = 5$ for GUE and LUE, $d = 3$ for JUE, and the $O$-term is uniform in $h$ and $l$. Summing up with respect to $l = 1, 2, \ldots$ and replacing $h \mapsto hn^{\gamma}/([n^{\gamma}] + 1)$, where $[n^{\gamma}]$ is the integer part of $n^{\gamma}$, we arrive at the desired formula

$$\log \frac{H_{n,n}[inh\gamma f]}{H_{n,n}[0]} = ihn^{\gamma}(n\varepsilon[f] + \mu[f]) - \frac{n^{2\gamma}h^2}{2} K[f] + O \left( \frac{hn^{\gamma}}{n^{1-(d-1)\gamma}} \right),$$

uniformly in $h$ for $|h| < \varepsilon$. If $\gamma \leq 1/d$, taking exponents of the both sides of (149) immediately yields the statement of the lemma.

As a final remark we note that if $\gamma \in (1/d, 1/(d - 1))$, the $O$-term in (149) is growing. However, using the inequality $|e^z - 1| \leq |z|e^{|z|}$, $z \in \mathbb{C}$, one can still get

$$\sup_n \sup_{|h| < \varepsilon n^{\gamma}} \left( \frac{n^{1-(d-1)\gamma}}{e^{C|h|}} \frac{|\varphi_{f,n}(h) - \varphi_{\mathcal{N}}(h)|}{h\varphi_{\mathcal{N}}(h)} \right) < +\infty$$

(150)

but with an additional factor $e^{C|h|}$, for some independent from $n$ and $h$ constant $C > 0$.

5 Proof of Theorem 1

It immediately follows from Lemma 1 for $\gamma = 0$ that

$$c_n = \mathbb{E}_n \left[ (\text{Tr} f(M) - n\varepsilon[f] - \mu[f]) / \sqrt{K[f]} \right] = O \left( \frac{1}{n} \right), \quad n \to \infty. \quad (151)$$

In order to estimate Kolmogorov's distance between $F_{f,n}(x)$ and $F_{\mathcal{N}}(x)$ we introduce the cumulative distribution function $\tilde{F}_{f,n}(x)$ of the centered random variable $(\text{Tr} f(M) - \mathbb{E}_n[\text{Tr} f(M)]) / \sqrt{K[f]}$. Clearly, $\tilde{F}_{f,n}(x) = F_{f,n}(x + c_n)$, and we can write

$$\sup_x |F_{f,n}(x) - F_{\mathcal{N}}(x)| \leq \sup_x |\tilde{F}_{f,n}(x) - F_{\mathcal{N}}(x)| + \sup_x |F_{\mathcal{N}}(x + c_n) - F_{\mathcal{N}}(x)|. \quad (152)$$

The last term is easy to estimate directly

$$\sup_x |F_{\mathcal{N}}(x + c_n) - F_{\mathcal{N}}(x)| = 2F_{\mathcal{N}}(|c_n|/2) - 1 = \frac{|c_n|}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{s^2}{2}} ds = O \left( \frac{1}{n} \right) \quad (153)$$

35
as $n \to \infty$.

To estimate $\sup x |\tilde{F}_{f,n}(x) - F_N(x)|$ we note that $\varphi_{f,n}(h)e^{-ic_nh}$ is the corresponding to $\tilde{F}_{f,n}(x)$ characteristic function and apply the smoothing inequality (see [16, p. 538])

$$\sup x |\tilde{F}_{f,n}(x) - F_N(x)| \leq \frac{1}{\pi} \int_{-T}^{T} \left| \frac{\varphi_{f,n}(h)e^{-ic_nh} - \varphi_N(h)}{h} \right| dh + \frac{24}{\sqrt{2\pi^3}T} \tag{154}$$

with $T = \varepsilon n^\gamma$, where $\varepsilon$ is chosen according to Lemma[1]. The use of this inequality is justified, since $\tilde{F}_{f,n}(x)$ corresponds to a random variable with zero mean.

The use of the triangle inequality in the integral leads to

$$\sup x |\tilde{F}_{f,n}(x) - F_N(x)| \leq \frac{1}{\pi} \int_{-\varepsilon n^\gamma}^{\varepsilon n^\gamma} \left| \varphi_{f,n}(h) - \varphi_N(h) \right| dh + \frac{1}{\pi} \int_{-\varepsilon n^\gamma}^{\varepsilon n^\gamma} \left| \varphi_N(h) \right| \left| \frac{e^{ic_nh} - 1}{h} \right| dh + \frac{24}{\varepsilon n^\gamma \sqrt{2\pi^3}} \tag{155}$$

Then, for sufficiently large $n$ and $\gamma \leq 1/d$ Lemma[1] yields

$$\int_{-\varepsilon n^\gamma}^{\varepsilon n^\gamma} \left| \frac{\varphi_{f,n}(h) - \varphi_N(h)}{h} \right| dh \leq \frac{\tilde{C}}{n^{1-(d-1)\gamma}} \int_{-\varepsilon n^\gamma}^{\varepsilon n^\gamma} \varphi_N(h) dh \leq \frac{\tilde{C}}{n^{1-(d-1)\gamma}} \tag{156}$$

where $\tilde{C}$ and $\tilde{\tilde{C}}$ are some non-negative independent of $n$ constants. Because of (150), the final inequality in (156) also holds for $\gamma \in (1/d, 1/(d-1))$.

Finally, collecting all the terms and choosing $\gamma = 1/d$ to attain the best available rate of convergence, we arrive at the desired asymptotic formula

$$\sup_{n,x} \left( n^{1/d} |F_{f,n}(x) - F_N(x)| \right) < +\infty, \tag{157}$$

which concludes the proof.

References

[1] Berestycki, N., Webb, C., Wong, M. D. (2017). Random Hermitian matrices and Gaussian multiplicative chaos. Probability Theory and Related Fields, 1-87.

[2] Billingsley, P. (2008). Probability and measure. John Wiley & Sons.

[3] Kuijlaars, A. B., Bleher, P. M. (2007). Large n limit of Gaussian random matrices with external source, Part III: Double scaling limit. Comm. Math. Phys. 270, 481-517.
[4] Charlier, C. (2018). Asymptotics of Hankel determinants with a one-cut regular potential and Fisher-Hartwig singularities, Int. Math. Res. Not. 2018, 62 pp.

[5] Charlier, C., Gharakhloo, R. (2019). Asymptotics of Hankel determinants with a Laguerre-type or Jacobi-type potential and Fisher-Hartwig singularities. Preprint arXiv:1902.08162.

[6] Chen, Y., Lawrence, N. (1998). On the linear statistics of Hermitian random matrices. J. Phys. A: Math. Gen., 31, 1141-1152.

[7] Chen, Y., Haq, N. S., McKay, M. R. (2013). Random matrix models, double-time Painlevé equations, and wireless relaying. J. of Math. Physics, 54, 063506.

[8] Claeys, T., Its, A., Krasovsky, I. (2011). Emergence of a singularity for Toeplitz determinants and Painlevé V. Duke Math. J. 160, no. 2, 207–262.

[9] Claeys, T., Krasovsky, I. (2015). Toeplitz determinants with merging singularities. Duke Math. J. 164, no. 15, 2897–2987.

[10] Deift, P. (2000). Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach. Courant Lecture Notes, Vol. 3.

[11] Deift, P., Its, A., Krasovsky, I. (2011). Asymptotics of Toeplitz, Hankel, and Toeplitz plus Hankel determinants with Fisher-Hartwig singularities. Annals of Mathematics, 174(2), 1243-1299.

[12] Deift, P., Its, A., Krasovsky, I. (2014). On the asymptotics of a Toeplitz determinant with singularities. Random Matrices MSRI Publications, 65, 93-146.

[13] Deift, P., Kriecherbauer, T., McLaughlin, K. T. R., Venakides, S., Zhou, X. (1999). Strong asymptotics of orthogonal polynomials with respect to exponential weights. Communications on Pure and Applied Mathematics, 52(12), 1491-1552.

[14] Deift, P., Kriecherbauer, T., McLaughlin, K. T. R., Venakides, S., Zhou, X. (1999). Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. Communications on Pure and Applied Mathematics, 52(11), 1335-1425.

[15] Deift, P., Zhou, X. (1993). A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation. Annals of Mathematics, 137(2), 295-368.

[16] Feller, F. (1966) An Introduction to Probability Theory and its Applications, Vol. II, John Wiley & Sons.

[17] Fokas, A. S., Its, A. R., Kitaev, A. V. (1992). The isomonodromy approach to matrix models in 2d quantum gravity. Communications in Mathematical Physics, 147(2), 395-430.
[18] Fyodorov, Y. V., Khoruzhenko, B. A., Simm, N. J. (2016). Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian Unitary Ensemble. Annals of Probability, 44(4), 2980-3031.

[19] Johansson, K. (1997). On Random Matrices from the Compact Classical Group. Annals of Mathematics, 145(3), 519-545.

[20] Johansson, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. Duke mathematical journal, 91(1), 151-204.

[21] Krasovsky, I. V. (2007). Correlations of the characteristic polynomials in the Gaussian unitary ensemble or a singular Hankel determinant. Duke mathematical journal, 139(3), 581-619.

[22] Kuijlaars, A. B., McLaughlin, K. R., Van Assche, W., Vanlessen, M. (2004). The Riemann-Hilbert approach to strong asymptotics for orthogonal polynomials on $[-1,1]$. Advances in mathematics, 188(2), 337-398.

[23] Lyu, S., Chen, Y. (2017). The largest eigenvalue distribution of the Laguerre unitary ensemble. Acta Mathematica Scientia, 37(2), 439-462.

[24] Pastur, L. A., Shcherbina, M. (2011). Eigenvalue distribution of large random matrices (No. 171). American Mathematical Soc.

[25] Saff, E. B., Totik, V. (2013). Logarithmic potentials with external fields (Vol. 316). Springer Science & Business Media.

[26] Stein, C. (1995). The accuracy of the normal approximation to the distribution of the traces of powers of random orthogonal matrices. Technical Report 470, Stanford University, Department of Statistics.

[27] Titchmarsh, E. C. (1986). Introduction to the theory of Fourier integrals. Third edition. Chelsea Publishing Co.: New York, 394 p.

[28] Vanlessen, M. (2007). Strong asymptotics of Laguerre-type orthogonal polynomials and applications in random matrix theory. Constructive approximation, 25(2), 125-175.

[29] Xu, S. X., Dai, D., Zhao, Y. Q. (2015). Painlevé III asymptotics of Hankel determinants for a singularly perturbed Laguerre weight. Journal of Approximation Theory, 192, 1-18.

[30] Zhao, Y., Cao, L. H., Dai, D. (2014). Asymptotics of the partition function of a Laguerre-type random matrix model. Journal of Approximation Theory, 178, 64-90.

Sergey Berezin
Aix-Marseille Université, Centrale Marseille, CNRS, Institut de Mathématiques de Marseille, UMR7373, 39 Rue F. Joliot Curie 13453, Marseille, France; St. Petersburg Department of V.A. Steklov Mathematical Institute of RAS, 27 Fontanka 191023, St. Petersburg, Russia
E-mail address: servberezin@yandex.ru, sergey.berezin@univ-amu.fr
ALEXANDER I. BUFETOV
Aix-Marseille Université, Centrale Marseille, CNRS, Institut de Mathématiques de Marseille, UMR7373, 39 Rue F. Joliot Curie 13453, Marseille, France; Steklov Mathematical Institute of RAS, 8 Gubkina 119991, Moscow, Russia; Institute for Information Transmission Problems, 19 Bolshoy Karetny 127994, Moscow, Russia

E-mail address: bufetov@mi.ras.ru, alexander.bufetov@univ-amu.fr