

Random Matrix Model with External Source and a Constrained Vector Equilibrium Problem

P. BLEHER

Indiana University–Purdue University Indianapolis

S. DELVAUX

Katholieke Universiteit Leuven

AND

A. B. J. KUIJLAARS

Katholieke Universiteit Leuven

Abstract

We consider the random matrix model with external source in the case where the potential $V(x)$ is an even polynomial and the external source has two eigenvalues $\pm a$ of equal multiplicity. We show that the limiting mean eigenvalue distribution of this model can be characterized as the first component of a pair of measures (μ_1, μ_2) that solve a constrained vector equilibrium problem. The proof is based on the steepest-descent analysis of the associated Riemann-Hilbert problem for multiple orthogonal polynomials.

We illustrate our results in detail for the case of a quartic double-well potential $V(x) = \frac{1}{4}x^4 - \frac{t}{2}x^2$. We are able to determine the precise location of the phase transitions in the ta -plane, where either the constraint becomes active or the two intervals in the support come together (or both). © 2010 Wiley Periodicals, Inc.

1 Introduction

The random matrix model with external source is the probability measure

$$(1.1) \quad \frac{1}{Z_n} \exp(-n \operatorname{Tr}(V(M) - AM)) dM$$

defined on the space of $n \times n$ Hermitian matrices M . Here A is a given Hermitian matrix (the external source), $V : \mathbb{R} \rightarrow \mathbb{R}$ is a function with sufficient increase at infinity (the potential), and Z_n is a normalization constant.

The model (1.1) was first studied by Brézin and Hikami [11, 12] and P. Zinn-Justin [39, 40], who showed that the eigenvalue correlations are determinantal. In [9] it was observed that the correlation kernel can be expressed in terms of multiple orthogonal polynomials. Due to the Riemann-Hilbert (RH) problem for multiple orthogonal polynomials [38], this opened up a new way for asymptotic analysis.

For the quadratic case

$$(1.2) \quad V(x) = \frac{1}{2} x^2$$

with external source

$$(1.3) \quad A = \text{diag}(\underbrace{a, \dots, a}_{n/2 \text{ times}}, \underbrace{-a, \dots, -a}_{n/2 \text{ times}})$$

with two eigenvalues $\pm a$ of equal multiplicity (thus n is even); this was done in great detail in the three papers [2, 8, 10].

The quadratic case is of special interest because it has an equivalent formulation in terms of nonintersecting Brownian motions that start at one value and end at certain prescribed values [2], which is a variation on Dyson's Brownian motion [23]. The quadratic model with external source (1.3) exhibits a phase transition, since for small $a > 0$ the eigenvalues accumulate on one interval while for larger a the eigenvalues accumulate on two disjoint intervals. At the critical value of a the local eigenvalue correlations are given in terms of Pearcey integrals [1, 10, 34, 37].

In this paper we study the external source model (1.1) with a more general potential V . We assume that V is an even polynomial

$$(1.4) \quad V(x) = \sum_{j=1}^d v_j x^{2j}, \quad v_d > 0,$$

of degree $2d$.

For $a = 0$ the external source model (1.1) reduces to the usual unitary matrix model

$$(1.5) \quad \frac{1}{Z_n} \exp(-n \text{Tr} V(M)),$$

which is one of the most studied models in random matrix theory in both mathematics and physics; see, e.g., [6, 7, 13, 14, 18] for rigorous study using the Riemann-Hilbert approach. A basic fact is that for $n \rightarrow \infty$, the limiting mean eigenvalue distribution of the matrix M in (1.5) minimizes the energy functional

$$(1.6) \quad E(\mu) = \iint \log \frac{1}{|x-y|} d\mu(x) d\mu(y) + \int V(x) d\mu(x)$$

over all probability measures μ on \mathbb{R} .

It is an open problem to find an analogue for the equilibrium problem (1.6) in the general context of the random matrix model with external source. This paper contains a first result in this direction. We consider the external source model (1.1) in the case where the potential V is an even polynomial (1.4). The external source A is again given by (1.3) with two eigenvalues $\pm a$ of equal multiplicity. We show that under these assumptions, the limiting mean eigenvalue distribution of the matrix M in (1.1) exists, and it arises as the first component of a pair of measures (μ_1, μ_2) solving a certain vector equilibrium problem; see Section 2.

We will illustrate our results in detail for a particular case of a nonconvex potential, namely the quartic double-well potential

$$(1.7) \quad V(x) = \frac{1}{4}x^4 - \frac{t}{2}x^2, \quad t > 0.$$

For the random matrix model (1.5) with quartic potential (1.7) it is known that the eigenvalues accumulate on either one or two intervals. The local eigenvalue correlations for the critical value of $t = t_{\text{cr}} = 2$ are given in terms of Ψ -functions associated with the Hastings-McLeod solution of the Painlevé II equation [7, 13].

So in the quartic model with external source there exist at least two mechanisms by which a transition from one to two intervals can occur: a Pearcey transition and a Painlevé II transition. It will be one of the outcomes of the present paper that we can determine precisely the location of the phase transitions in the ta -plane.

2 Statement of Results

2.1 Equilibrium Problem

The main ingredient in our analysis is a new vector equilibrium problem associated with the random matrix model (1.1) with external source. We emphasize that it only applies in the setting we are considering, namely an even polynomial potential V as in (1.4) and an external source (1.3) with two eigenvalues of equal multiplicity. This setting gives a symmetry with respect to the origin, which we use in an essential way.

The equilibrium problem is as follows. We minimize the energy functional

$$(2.1) \quad \begin{aligned} E(\mu_1, \mu_2) &= \iint \log \frac{1}{|x-y|} d\mu_1(x)d\mu_1(y) + \iint \log \frac{1}{|x-y|} d\mu_2(x)d\mu_2(y) \\ &\quad - \iint \log \frac{1}{|x-y|} d\mu_1(x)d\mu_2(y) + \int (V(x) - a|x|)d\mu_1(x) \end{aligned}$$

with respect to all pairs of measures (μ_1, μ_2) satisfying

- μ_1 and μ_2 have finite logarithmic energy,
- μ_1 is a measure on \mathbb{R} with total mass 1,
- μ_2 is a measure on $i\mathbb{R}$ with total mass $\frac{1}{2}$ that satisfies the constraint

$$(2.2) \quad \mu_2 \leq \sigma$$

where σ is the measure on $i\mathbb{R}$ with constant density

$$(2.3) \quad \frac{d\sigma}{|dz|} = \frac{a}{\pi}, \quad z \in i\mathbb{R}.$$

Standard references on potential theory in the complex plane are [15, 33, 36].

The equilibrium problem (2.1) has both an external field $V(x) - a|x|$ acting on μ_1 and an upper constraint σ acting on μ_2 . The interaction between μ_1 and μ_2 is of Nikishin type [27]. This type of vector equilibrium problem also appeared

recently in a model of nonintersecting squared Bessel paths [29] and in the two-matrix model with quartic potential [21].

Our first result concerns the structure of the minimizer of the equilibrium problem.

THEOREM 2.1 *There is a unique minimizer (μ_1, μ_2) that satisfies the following:*

(i) *The support of μ_1 is bounded and consists of a finite union of intervals*

$$(2.4) \quad \text{supp}(\mu_1) = \bigcup_{j=1}^N [a_j, b_j].$$

The measure μ_1 is absolutely continuous with density

$$(2.5) \quad \frac{d\mu_1(x)}{dx} = h(x) \sqrt{\prod_{j=1}^N (b_j - x)(x - a_j)}, \quad x \in \bigcup_{j=1}^N [a_j, b_j],$$

where h is a nonnegative function on $\text{supp}(\mu_1) = \bigcup_{j=1}^N [a_j, b_j]$ that is real analytic except possibly at 0.

(ii) *The support of μ_2 is the full imaginary axis and there exists $c \geq 0$ such that*

$$(2.6) \quad \text{supp}(\sigma - \mu_2) = (-i\infty, -ic] \cup [ic, i\infty).$$

We have that $c = 0$ if and only if

$$(2.7) \quad \int \frac{d\mu_1(s)}{|s|} \leq 2a,$$

and in that case, μ_2 has the density

$$(2.8) \quad \frac{d\mu_2(z)}{|dz|} = \frac{1}{2\pi} \int \frac{|s|}{|z|^2 + s^2} d\mu_1(s), \quad z \in i\mathbb{R}.$$

If (2.7) is not satisfied, then $c > 0$ is determined by the condition

$$(2.9) \quad \int \frac{d\mu_1(s)}{\sqrt{s^2 + c^2}} = 2a,$$

and in that case

$$(2.10) \quad \frac{d\mu_2(z)}{|dz|} = \frac{a}{\pi} \quad \text{if } z \in [-ic, ic]$$

and

$$(2.11) \quad \frac{d\mu_2(z)}{|dz|} = \frac{a}{\pi} - \frac{1}{2\pi} \int \frac{|z|\sqrt{|z|^2 - c^2}}{(|z|^2 + s^2)\sqrt{s^2 + c^2}} d\mu_1(s)$$

$$(2.12) \quad = \frac{1}{2\pi} \int \frac{|z|^2 + s^2 - |z|\sqrt{|z|^2 - c^2}}{(|z|^2 + s^2)\sqrt{s^2 + c^2}} d\mu_1(s),$$

if $z \in (-i\infty, -ic] \cup [ic, i\infty)$.

(iii) Both μ_1 and μ_2 are symmetric with respect to the origin.

Theorem 2.1 will be proved in Sections 3.1 and 3.2. Note that (2.4)–(2.8) are similar to statements proved in [21], while (2.9) and (2.11)–(2.12) have apparently not been stated before.

2.2 Variational Conditions

The minimizer (μ_1, μ_2) to the equilibrium problem in Section 2.1 is characterized by the following Euler-Lagrange variational conditions. We write

$$U^\mu(x) = \int \log \frac{1}{|x-y|} d\mu(y)$$

for the logarithmic potential of a measure μ .

PROPOSITION 2.2 *The measures μ_1 and μ_2 satisfy for some constant $\ell \in \mathbb{R}$,*

$$(2.13) \quad 2U^{\mu_1}(x) = U^{\mu_2}(x) - V(x) + a|x| - \ell, \quad x \in \text{supp}(\mu_1),$$

$$(2.14) \quad 2U^{\mu_1}(x) \geq U^{\mu_2}(x) - V(x) + a|x| - \ell, \quad x \in \mathbb{R} \setminus \text{supp}(\mu_1),$$

$$(2.15) \quad 2U^{\mu_2}(x) = U^{\mu_1}(x), \quad x \in \text{supp}(\sigma - \mu_2),$$

$$(2.16) \quad 2U^{\mu_2}(x) < U^{\mu_1}(x), \quad x \in i\mathbb{R} \setminus \text{supp}(\sigma - \mu_2).$$

These relations follow directly from the variational conditions of the equilibrium problem.

2.3 Regular and Singular Cases

We say that μ_1 is *regular* if in (2.5) we have $h(x) > 0$ on

$$\text{supp}(\mu_1) = \bigcup_{j=1}^N [a_j, b_j],$$

and if the variational inequality (2.14) is strict for every $x \in \mathbb{R} \setminus \text{supp}(\mu_1)$. Otherwise μ_1 is called *singular* [17, 18].

The measure μ_2 has a density on $i\mathbb{R}$ that is bounded by $\frac{a}{\pi}$. We say that μ_2 is *singular* if equality in this restriction is attained at $z = 0$ and at no other point of $i\mathbb{R}$. In all other cases (in particular, if $c > 0$), the measure μ_2 is called *regular*.

For the analysis in this paper, we will assume that both μ_1 and μ_2 are regular. The following lemma follows immediately from this assumption and is stated only for further reference.

LEMMA 2.3 *The measures μ_1 and $\sigma - \mu_2$ satisfy the following square root behavior near their endpoints a_j , b_j , and $\pm ic$:*

- (i) *If the measure μ_1 is regular, then it has a density of the form (2.5) with h strictly positive on $\bigcup_{j=1}^N [a_j, b_j]$.*
- (ii) *If $c > 0$ then the measure $\sigma - \mu_2$ has a density of the form*

$$\frac{a}{\pi} - \frac{d\mu_2(z)}{|dz|} = k(z) \sqrt{|z|^2 - c^2},$$

where k is an analytic, strictly positive function on $(-i\infty, -ic] \cup [ic, i\infty)$.

Lemma 2.3(i) follows immediately from the definition of μ_1 being regular. Lemma 2.3(ii) follows from (2.11).

2.4 Limiting Eigenvalue Distribution

Our main result deals with the global distribution of eigenvalues as $n \rightarrow \infty$.

THEOREM 2.4 *Let V be an even polynomial, and let A be a diagonal matrix with two eigenvalues $\pm a$ of equal multiplicity. Let (μ_1, μ_2) be the solution of the equilibrium problem in Section 2.1, and assume that both μ_1 and μ_2 are regular in the sense explained above. Then the mean eigenvalue distribution of a matrix M from the random matrix model*

$$\frac{1}{Z_n} \exp(-n \operatorname{Tr}(V(M) - AM)) dM$$

has the limit μ_1 as $n \rightarrow \infty$.

We strongly expect that the conclusion of Theorem 2.4 remains valid in the case where μ_1 and/or μ_2 is singular.

Theorem 2.4 will be proved in Section 5.7.

2.5 About the Proof of Theorem 2.4

The proof of Theorem 2.4 is based on the Riemann-Hilbert problem for multiple orthogonal polynomials and its connection with the external source model (1.1). The multiple orthogonal polynomials $P_n(x)$ in question are orthogonal with respect to the weights

$$(2.17) \quad w_1(x) = e^{-n(V(x)-ax)}, \quad w_2(x) = e^{-n(V(x)+ax)}.$$

More precisely, $P_n(x)$ is a monic polynomial of degree n that is characterized by the multiple orthogonality conditions (we assume n is even)

$$(2.18) \quad \int_{-\infty}^{\infty} P_n(x) w_j(x) x^k dx = 0, \quad j = 1, 2, \quad k = 0, \dots, \frac{n}{2} - 1.$$

The polynomial $P_n(x)$ is also the average characteristic polynomial

$$(2.19) \quad P_n(z) = \mathbb{E}[\det(zI_n - M)],$$

where the expectation \mathbb{E} is taken with respect to the model (1.1); see [9].

The Riemann-Hilbert problem (RH problem) for $P_n(x)$ is as follows. We look for an analytic matrix-valued function $Y : \mathbb{C} \setminus \mathbb{R} \rightarrow \mathbb{C}^{3 \times 3}$ with jump properties

$$(2.20) \quad Y_+(x) = Y_-(x) \begin{pmatrix} 1 & w_1(x) & w_2(x) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in \mathbb{R},$$

and asymptotic condition

$$(2.21) \quad Y(z) = (I + O(1/z)) \begin{pmatrix} z^n & 0 & 0 \\ 0 & z^{-n/2} & 0 \\ 0 & 0 & z^{-n/2} \end{pmatrix} \quad \text{as } z \rightarrow \infty.$$

Here and below we use the following standard notation. If Γ is an oriented contour in the complex plane, then the side that is on the left (right) when traversing Γ according to its orientation is called the $+$ side ($-$ side), and for any $z \in \Gamma$ we use $Y_+(z)$ ($Y_-(z)$) to denote the limiting values of $Y(z)$ along the $+$ side ($-$ side) of Γ . In (2.20) we have the contour $\Gamma = \mathbb{R}$ oriented from left to right.

The RH problem (2.20)–(2.21) has a unique solution. The $(1, 1)$ -entry of $Y(z)$ is the multiple orthogonal polynomial $P_n(z)$ characterized by (2.18). Note that the RH problem (2.20)–(2.21) is a generalization of the RH problem for orthogonal polynomials [26] to multiple orthogonal polynomials [38].

It is known that the eigenvalues of the random matrix model with external source (1.3) form a determinantal point process with correlation kernel [9]

$$(2.22) \quad K_n(x, y) = \frac{1}{2\pi i(x-y)} \begin{pmatrix} 0 & w_1(y) & w_2(y) \end{pmatrix} Y_+(y)^{-1} Y_+(x) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Theorem 2.4 then comes down to the following statement about the limiting behavior of the kernels K_n :

$$(2.23) \quad \lim_{n \rightarrow \infty} \frac{1}{n} K_n(x, x) = \frac{d\mu_1(x)}{dx}, \quad x \in \mathbb{R}.$$

We will establish (2.23) in Section 5.7, thereby proving Theorem 2.4.

From the RH analysis it is possible to obtain universality results for the local eigenvalue correlations as well. In the regular cases, this leads to the usual sine kernel in the bulk and Airy kernel at the edge points of the spectrum. We will not discuss this any further and refer to the papers [2, 8, 10, 17, 18, 21], among others, for a detailed analysis in a similar context.

2.6 Organization of the Paper

The rest of the paper is organized as follows. In Section 3 we discuss the structure of the equilibrium measures and we prove Theorem 2.1. In Section 4 we introduce the Riemann surface built from the solution of the equilibrium problem. Section 5 contains the steepest-descent analysis of the RH problem for $Y(z)$, leading to the proof of Theorem 2.4. In Section 6 we make some general remarks on the expected phase transitions of our model, and in Section 7 we study this in detail for the case of a quartic potential.

3 The Equilibrium Problem

3.1 Existence of the Minimizer

In this section we prove the existence of the minimizer (μ_1, μ_2) of the equilibrium problem in Section 2.1. To this end we follow [21, sec. 4].

PROOF: The energy functional (2.1) can be written as

$$(3.1) \quad E(\mu_1, \mu_2) = \frac{3}{4} I(\mu_1) + \frac{1}{4} I(\mu_1 - 2\mu_2) + \int (V(x) - a|x|) d\mu_1(x)$$

where

$$I(\nu) = \iint \log \frac{1}{|x-y|} d\nu(x) d\nu(y)$$

denotes the logarithmic energy of a signed measure ν . Occasionally we will also write

$$I(\nu_1, \nu_2) = \iint \log \frac{1}{|x-y|} d\nu_1(x) d\nu_2(y)$$

to denote the mixed energy of a pair of measures ν_1 and ν_2 .

Since $I(\nu) \geq 0$ if ν is a signed measure with $\int d\nu = 0$, we find from (3.1) that

$$\begin{aligned} E(\mu_1, \mu_2) &\geq \frac{3}{4} I(\mu_1) + \int (V(x) - a|x|) d\mu_1(x) \\ &\geq \min_{\mu_1} \left[\frac{3}{4} I(\mu_1) + \int (V(x) - a|x|) d\mu_1(x) \right] > -\infty, \end{aligned}$$

where the last inequality follows from standard logarithmic potential theory with external fields; see, e.g., [36]. Thus the energy functional is bounded from below.

If we fix μ_2 on $i\mathbb{R}$ and we minimize (2.1) with respect to μ_1 only, then the problem for μ_1 is to minimize

$$I(\mu_1) + \int (V(x) - a|x| - U^{\mu_2}(x)) d\mu_1(x).$$

The extra term $U^{\mu_2}(x)$ comes from the interaction between μ_1 and μ_2 . It is a term that attracts the μ_1 mass towards the origin. It can indeed be proved (as in [21]) that if the minimizer in external field $V(x) - a|x|$ is contained in $[-X, X]$, then the minimizer in external field $V(x) - a|x| - U^{\mu_2}(x)$ is also contained in $[-X, X]$ (and so X is independent of μ_2).

If we fix μ_1 on $[-X, X]$, then the problem for μ_2 is to minimize

$$I(\mu_2) - \int U^{\mu_1}(x) d\mu_2(x)$$

among all $\mu_2 \leq \sigma$ with total mass $\frac{1}{2}$. As in [21], equality in the constraint is attained precisely on an interval of the form $[-ic, ic]$ for certain $c \geq 0$. We will show further that the minimizer μ_2 satisfying this constraint is given explicitly by (2.8)–(2.12). From these explicit formulas it follows immediately that for a measure μ_1 on $[-X, X]$, the corresponding minimizer μ_2 satisfies

$$\frac{d\mu_2(z)}{|dz|} \leq \frac{K}{|z|^2}, \quad z \in i\mathbb{R},$$

with a constant K that only depends on X .

Let $(\mu_{1,n}, \mu_{2,n})_{n=1}^{\infty}$ be a sequence of vectors of measures satisfying

$$\text{supp}(\mu_{1,n}) \subset \mathbb{R}, \quad \text{supp}(\mu_{2,n}) \subset i\mathbb{R}, \quad \int d\mu_{1,n} = 1, \quad \int d\mu_{2,n} = \frac{1}{2},$$

and $\mu_{2,n} \leq \sigma$, such that

$$E(\mu_{1,n}, \mu_{2,n}) \leq \frac{1}{n} + \inf E(\mu_1, \mu_2).$$

As shown above, we may assume in addition that

$$\text{supp}(\mu_{1,n}) \subset [-X, X], \quad \frac{d\mu_{2,n}(z)}{|dz|} \leq \frac{K}{|z|^2}, \quad z \in i\mathbb{R},$$

with X and K independent of n .

Then it follows as in [21] that the sequences $(\mu_{1,n})_{n=1}^{\infty}$ and $(\mu_{2,n})_{n=1}^{\infty}$ are tight. There is a convergent subsequence of $(\mu_{1,n}, \mu_{2,n})_{n=1}^{\infty}$, and the limit is the vector of minimizing measures; see also [28].

Summarizing, we have now proved the existence of the solution (μ_1, μ_2) to the equilibrium problem. The uniqueness of the solution follows in a standard way from the convexity of the energy functional; see, e.g., (3.1) and [21]. \square

3.2 Proof of Theorem 2.1

The proof of Theorem 2.1(i) follows as in [16], while Theorem 2.1(iii) is evident from the symmetry of the problem.

It remains to prove (2.7)–(2.12) in Theorem 2.1(ii). For $k = 1, 2$, define the Cauchy transforms

$$(3.2) \quad F_k(z) = \int \frac{1}{z-s} d\mu_k(s), \quad z \in \mathbb{C} \setminus \text{supp}(\mu_k).$$

By differentiating the variational condition (2.15), we find that

$$(3.3) \quad F_{2+}(z) + F_{2-}(z) - F_1(z) = 0, \quad z \in (-i\infty, -ic) \cup (ic, i\infty).$$

Here we assume that the imaginary axis is oriented from bottom to top, so that the $+$ side is on the left and the $-$ side is on the right, as usual. If $c > 0$, then $\frac{d\mu_2}{|dz|} = \frac{d\sigma(z)}{|dz|} = \frac{a}{\pi}$ on $(-ic, ic)$, from which it follows that

$$(3.4) \quad F_{2+}(z) - F_{2-}(z) = -2a, \quad z \in (-ic, ic).$$

We can solve equations (3.3)–(3.4) for F_2 . We consider the two cases $c = 0$ and $c > 0$.

Case 1. $c = 0$. In this case we have equation (3.3) on $i\mathbb{R}$,

$$(3.5) \quad F_{2+}(z) + F_{2-}(z) - F_1(z) = 0, \quad z \in (-i\infty, i\infty).$$

Define the Cauchy transforms of the restrictions of the measure μ_1 to the positive and negative half-axes,

$$F_1^+(z) = \int_0^\infty \frac{d\mu_1(s)}{z-s}, \quad z \in \mathbb{C} \setminus (\mathbb{R}_+ \cap \text{supp}(\mu_1)),$$

$$F_1^-(z) = \int_{-\infty}^0 \frac{d\mu_1(s)}{z-s}, \quad z \in \mathbb{C} \setminus (\mathbb{R}_- \cap \text{supp}(\mu_1)).$$

Then

$$F_1(z) = F_1^+(z) + F_1^-(z),$$

and, due to the uniqueness of the solution of the scalar Riemann-Hilbert problem (3.5), we have

$$F_2(z) = \begin{cases} F_1^+(z), & \text{Re } z \leq 0, \\ F_1^-(z), & \text{Re } z \geq 0. \end{cases}$$

Therefore, the Sokhotski-Plemelj formula implies that for $z \in i\mathbb{R}$,

$$\begin{aligned} \frac{d\mu_2(z)}{|dz|} &= -\frac{1}{2\pi}[F_{2+}(z) - F_{2-}(z)] = -\frac{1}{2\pi}[F_1^+(z) - F_1^-(z)] \\ &= -\frac{1}{2\pi} \left[\int_0^\infty \frac{d\mu_1(s)}{z-s} - \int_0^\infty \frac{d\mu_1(s)}{z+s} \right] = \frac{1}{\pi} \int_0^\infty \frac{s d\mu_1(s)}{|z|^2 + s^2}, \end{aligned}$$

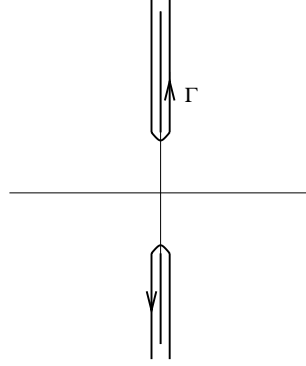
which is equivalent to (2.8). The case $c = 0$ is valid if and only if the density (2.8) is bounded by $\frac{a}{\pi}$. Since (2.8) assumes its maximum for $|z| = 0$, this happens if and only if (2.7) holds.

Case 2. $c > 0$. Consider the function

$$\tilde{F}_2(z) = \frac{F_2(z)}{\sqrt{R(z)}}, \quad R(z) = z^2 + c^2,$$

where $\sqrt{R(z)} = \sqrt{z^2 + c^2}$ is defined with a cut $(-i\infty, -ic] \cup [ic, i\infty)$, and $\sqrt{R(0)} = c$. From equation (3.3) we obtain that

$$(3.6) \quad \tilde{F}_{2+}(z) - \tilde{F}_{2-}(z) = \frac{F_1(z)}{\sqrt{R(z)}_+}, \quad z \in (-i\infty, -ic] \cup [ic, i\infty),$$

FIGURE 3.1. The contour Γ .

where $+$ again denotes the limiting value from the left half-plane. Observe that for $y > c$,

$$\sqrt{R(iy)}_+ = -i\sqrt{y^2 - c^2}.$$

In addition, from equation (3.4) we obtain that

$$(3.7) \quad \tilde{F}_{2+}(z) - \tilde{F}_{2-}(z) = -\frac{2a}{\sqrt{R(z)}}, \quad z \in (-ic, ic).$$

By (3.6) and (3.7) the Sokhotski-Plemelj formula implies that

$$(3.8) \quad \begin{aligned} \tilde{F}_2(z) = & -\frac{1}{2\pi i} \left(\int_{-i\infty}^{-ic} + \int_{ic}^{i\infty} \right) \frac{F_1(s) ds}{(z-s)\sqrt{R(s)}_+} \\ & + \frac{1}{2\pi i} \int_{-ic}^{ic} \frac{2a ds}{(z-s)\sqrt{R(s)}}. \end{aligned}$$

The first term in the right-hand side of (3.8) can be written as

$$(3.9) \quad -\frac{1}{2\pi i} \left(\int_{-i\infty}^{-ic} + \int_{ic}^{i\infty} \right) \frac{F_1(s) ds}{(z-s)\sqrt{R(s)}_+} = \frac{1}{4\pi i} \oint_{\Gamma} \frac{F_1(s) ds}{(z-s)\sqrt{R(s)}},$$

where the contour Γ is depicted in Figure 3.1. From (3.2) and Fubini's theorem,

$$(3.10) \quad \begin{aligned} & \frac{1}{4\pi i} \oint_{\Gamma} \frac{F_1(s) ds}{(z-s)\sqrt{R(s)}} \\ & = \frac{1}{4\pi i} \int_{-\infty}^{\infty} \oint_{\Gamma} \frac{ds}{(z-s)(s-\zeta)\sqrt{R(s)}} d\mu_1(\zeta) \\ & = \frac{1}{4\pi i} \int_{-\infty}^{\infty} \frac{1}{z-\zeta} \left(\oint_{\Gamma} \frac{ds}{(z-s)\sqrt{R(s)}} + \oint_{\Gamma} \frac{ds}{(s-\zeta)\sqrt{R(s)}} \right) d\mu_1(\zeta). \end{aligned}$$

By contour deformation and Cauchy's theorem we have that

$$\frac{1}{4\pi i} \oint_{\Gamma} \frac{ds}{(z-s)\sqrt{R(s)}} = \frac{1}{2\sqrt{R(z)}}, \quad z \in \mathbb{C} \setminus ((-i\infty, -ic] \cup [ic, i\infty));$$

hence (3.10) reduces to

$$\begin{aligned} & \frac{1}{4\pi i} \oint_{\Gamma} \frac{F_1(s)ds}{(z-s)\sqrt{R(s)}} \\ (3.11) \quad &= \int_{-\infty}^{\infty} \frac{1}{2(z-\zeta)} \left(\frac{1}{\sqrt{R(z)}} - \frac{1}{\sqrt{R(\zeta)}} \right) d\mu_1(\zeta) \\ &= \frac{1}{\sqrt{R(z)}} \int_{-\infty}^{\infty} \frac{d\mu_1(\zeta)}{2(z-\zeta)} - \int_{-\infty}^{\infty} \frac{d\mu_1(\zeta)}{2(z-\zeta)\sqrt{R(\zeta)}}. \end{aligned}$$

For the second term in the right-hand side of (3.8) we have

$$(3.12) \quad \frac{1}{2\pi i} \int_{-ic}^{ic} \frac{2a ds}{(z-s)\sqrt{R(s)}} = \frac{a \operatorname{sgn} \operatorname{Re} z}{\sqrt{R(z)}}.$$

By inserting (3.11) and (3.12) in (3.8), we obtain that

$$\tilde{F}_2(z) = \frac{1}{\sqrt{R(z)}} \int_{-\infty}^{\infty} \frac{d\mu_1(s)}{2(z-s)} - \int_{-\infty}^{\infty} \frac{d\mu_1(s)}{2(z-s)\sqrt{R(s)}} + \frac{a \operatorname{sgn} \operatorname{Re} z}{\sqrt{R(z)}};$$

hence

$$(3.13) \quad F_2(z) = \frac{F_1(z)}{2} - \sqrt{z^2 + c^2} \int_{-\infty}^{\infty} \frac{d\mu_1(s)}{2(z-s)\sqrt{s^2 + c^2}} + a \operatorname{sgn} \operatorname{Re} z.$$

Note that by taking $z \rightarrow +\infty$ in (3.13), we find the relation (2.9) between c and a . Now the density of μ_2 is equal to

$$\begin{aligned} \frac{d\mu_2(z)}{|dz|} &= -\frac{1}{2\pi} [F_{2+}(z) - F_{2-}(z)] \\ &= \frac{a}{\pi} - \frac{i\sqrt{|z|^2 - c^2}}{\pi} \int_{-\infty}^{\infty} \frac{d\mu_1(s)}{2(z-s)\sqrt{s^2 + c^2}} \\ &= \frac{a}{\pi} - \frac{i\sqrt{|z|^2 - c^2}}{\pi} \int_0^{\infty} \left(\frac{1}{z-s} + \frac{1}{z+s} \right) \frac{d\mu_1(s)}{2\sqrt{s^2 + c^2}} \\ &= \frac{a}{\pi} - \frac{1}{\pi} \int_0^{\infty} \frac{|z|\sqrt{|z|^2 - c^2}}{|z|^2 + s^2} \frac{d\mu_1(s)}{\sqrt{s^2 + c^2}}, \end{aligned}$$

which is equivalent to (2.11). Then (2.12) follows from this and (2.9). \square

3.3 Structure of the Equilibrium Measures in the Regular Case

Theorem 2.1 implies that in the regular case, the structure of the equilibrium measures near the origin is described by one of the following three cases. This distinction will be important at several places in our RH steepest-descent analysis.

Case I. $c = 0$ and N is even.

Case II. $c > 0$ and N is even.

Case III. $c > 0$ and N is odd.

Here we recall the definitions of c and N in (2.4)–(2.6). Thus Case I could be formulated equivalently as $\text{supp}(\sigma - \mu_2) = i\mathbb{R}$ and $0 \notin \text{supp}(\mu_1)$. Similar formulations can be given for Case II and Case III. Note that the situation where $c = 0$ and N is odd is not in the above list since it would imply that μ_1 has singular behavior at the origin.

For the quadratic potential $V(x) = x^2/2$, it turns out that we are in Case I (with $N = 2$) for large values of a and in Case III (with $N = 1$) for small values of a . Case II does not occur.

In general, one expects Case I to happen for large values of a . Note that the upper constraint σ is not active in this case and therefore it could be removed from the equilibrium problem. A consequence of this is that $2\mu_2$ is equal to the balayage of μ_1 onto the imaginary axis; see also (2.8). This means that $2U^{\mu_2}(z) = U^{\mu_1}(z)$ for $z \in i\mathbb{R}$.

Let μ_1^L and μ_1^R denote the restrictions of μ_1 to the negative and positive real axis, respectively. By symmetry, we then have in Case I that μ_2 is the balayage of either μ_1^L or μ_1^R onto the imaginary axis, and moreover

$$(3.14) \quad U^{\mu_2}(z) = U^{\mu_1^L}(z) \quad \text{for } \text{Re } z \geq 0,$$

$$(3.15) \quad U^{\mu_2}(z) = U^{\mu_1^R}(z) \quad \text{for } \text{Re } z \leq 0.$$

Note that the equality is valid not only on the imaginary axis but also in a full half-plane. This follows from an easy application of the minimum and maximum principles for harmonic functions [36, chap. 0].

Then the following chain of equations is easy to verify:

$$(3.16) \quad \begin{aligned} I(\mu_2) &= \int U^{\mu_2} d\mu_2 = \int U^{\mu_1^L} d\mu_2 \\ &= \int U^{\mu_2} d\mu_1^L = \int U^{\mu_1^R} d\mu_1^L = I(\mu_1^L, \mu_1^R), \end{aligned}$$

and in a similar way

$$(3.17) \quad I(\mu_1, \mu_2) = \int U^{\mu_2} d\mu_1^L + \int U^{\mu_2} d\mu_1^R = 2I(\mu_1^L, \mu_1^R).$$

Since also

$$(3.18) \quad I(\mu_1) = I(\mu_1^L + \mu_1^R) = I(\mu_1^L) + I(\mu_1^R) + 2I(\mu_1^L, \mu_1^R),$$

it then follows from (3.16)–(3.18) that the energy functional (2.1) can be rewritten in Case I as

$$E(\mu_1, \mu_2) = I(\mu_1^L) + I(\mu_1^R) + I(\mu_1^L, \mu_1^R) \\ + \int (V(x) + ax) d\mu_1^L(x) + \int (V(x) - ax) d\mu_1^R(x).$$

Therefore the equilibrium problem is equivalent to the following equilibrium problem of Angelesco type for μ_1^L and μ_1^R : Minimize

$$(3.19) \quad E(\mu_1^L, \mu_1^R) = \iint \log \frac{1}{|x-y|} d\mu_1^L(x) d\mu_1^L(y) \\ + \iint \log \frac{1}{|x-y|} d\mu_1^R(x) d\mu_1^R(y) \\ + \iint \log \frac{1}{|x-y|} d\mu_1^L(x) d\mu_1^R(y) \\ + \int (V(x) + ax) d\mu_1^L(x) + \int (V(x) - ax) d\mu_1^R(x),$$

with respect to all pairs of measures (μ_1^L, μ_1^R) satisfying

- μ_1^L is a measure on \mathbb{R}_- with total mass $\frac{1}{2}$,
- μ_1^R is a measure on \mathbb{R}_+ with total mass $\frac{1}{2}$.

For the quadratic potential $V(x) = \frac{1}{2}x^2$, this equilibrium problem was described in [8].

4 Riemann Surface

From the minimizer (μ_1, μ_2) of the vector equilibrium problem we construct a three-sheeted Riemann surface \mathcal{R} whose three sheets are given as follows:

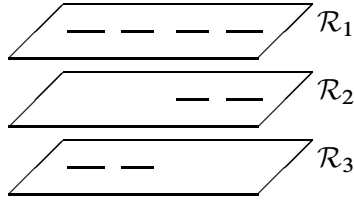
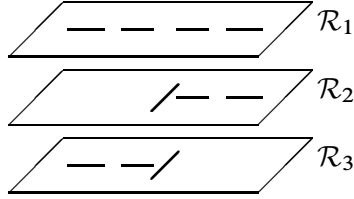
$$(4.1) \quad \mathcal{R}_1 = \overline{\mathbb{C}} \setminus \bigcup_{j=1}^N [a_j, b_j],$$

$$(4.2) \quad \mathcal{R}_2 = \overline{\mathbb{C}} \setminus \left(\left(\mathbb{R}_+ \cap \bigcup_{j=1}^N [a_j, b_j] \right) \cup \overline{(-ic, ic)} \right),$$

$$(4.3) \quad \mathcal{R}_3 = \overline{\mathbb{C}} \setminus \left(\left(\mathbb{R}_- \cap \bigcup_{j=1}^N [a_j, b_j] \right) \cup \overline{(-ic, ic)} \right).$$

Here $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ denotes the Riemann sphere.

The sheet \mathcal{R}_1 is connected with \mathcal{R}_2 via the intervals $[a_j, b_j]$ on the positive real line, \mathcal{R}_1 is connected with \mathcal{R}_3 via the intervals $[a_j, b_j]$ on the negative real line, and (in Case II and Case III) \mathcal{R}_2 is connected to \mathcal{R}_3 via the interval $[-ic, ic]$ on the

FIGURE 4.1. Riemann surface for Case I (with $N = 4$).FIGURE 4.2. Riemann surface for Case II (with $N = 4$).

imaginary axis. The connections are in the usual crosswise manner. The Riemann surface is compact and has genus

$$(4.4) \quad g = \begin{cases} N - 2 & \text{in Case I,} \\ N - 1 & \text{in Case II and Case III.} \end{cases}$$

Here Cases I, II, and III were defined in Section 3.3. An illustration of the Riemann surface for each of these three cases is shown in Figures 4.1, 4.2, and 4.3.

Recall functions F_1 and F_2 in (3.2). These functions are used to define a meromorphic function on the Riemann surface; compare with [21, lemma 5.1]:

PROPOSITION 4.1 *For $k = 1, 2, 3$, let $\xi_k(z)$ on the sheet \mathcal{R}_k be defined by*

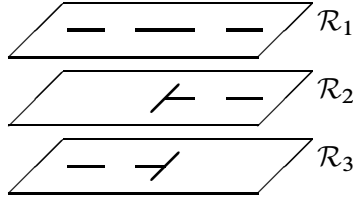
$$(4.5) \quad \begin{aligned} \xi_1(z) &= V'(z) - F_1(z), & z \in \mathcal{R}_1, \\ \xi_2(z) &= a + F_1(z) - F_2(z), & z \in \mathcal{R}_2, \operatorname{Re} z > 0, \\ \xi_2(z) &= a + F_2(z), & z \in \mathcal{R}_2, \operatorname{Re} z < 0, \\ \xi_3(z) &= -a + F_2(z), & z \in \mathcal{R}_3, \operatorname{Re} z > 0, \\ \xi_3(z) &= -a + F_1(z) - F_2(z), & z \in \mathcal{R}_3, \operatorname{Re} z < 0. \end{aligned}$$

Then these functions have an analytic continuation to a meromorphic function (denoted by $\xi(z)$) on the Riemann surface whose only pole is at the point at infinity on the first sheet.

PROOF: Let us first check that ξ_k is analytic on the sheet \mathcal{R}_k , $k = 2, 3$. This reduces to showing the equality

$$F_{2,+}(z) + F_{2,-}(z) - F_1(z) = 0, \quad z \in i\mathbb{R} \setminus [-ic, ic],$$

which is a direct consequence of the variational condition (2.15); see also (3.3).

FIGURE 4.3. Riemann surface for Case III (with $N = 3$).

Next, we must show that the ξ -functions are each other's analytic continuation when passing through one of the cuts of the Riemann surface. For example, the fact that $\xi_{1,+} = \xi_{2,-}$ along the cut $\mathbb{R}_+ \cap \bigcup_{j=1}^N [a_j, b_j]$ reduces to showing the equality

$$F_{1,+}(x) + F_{1,-}(x) - F_2(x) - V'(x) + a = 0, \quad x \in \mathbb{R}_+ \cap \bigcup_{j=1}^N (a_j, b_j),$$

which is a direct consequence of the variational condition (2.13). The other equalities are checked similarly; see also (3.4). \square

It follows from Proposition 4.1 that the function $\xi(z)$ is an algebraic function satisfying an equation of the third degree in ξ , known as the spectral curve:

$$(4.6) \quad \xi^3 + p_2(z)\xi^2 + p_1(z)\xi + p_0(z) = 0$$

where p_0 , p_1 , and p_2 are polynomials. Here

$$(4.7) \quad p_2(z) = -\xi_1(z) - \xi_2(z) - \xi_3(z) = -V'(z)$$

is known, but the determination of the polynomials

$$p_1(z) = \xi_1(z)\xi_2(z) + \xi_1(z)\xi_3(z) + \xi_2(z)\xi_3(z),$$

$$p_0(z) = -\xi_1(z)\xi_2(z)\xi_3(z),$$

cannot be done explicitly in general.

However, we can find expressions for p_1 and p_0 in terms of μ_1 and μ_2 . See [4, theorem 5.1] for a similar result for the spectral curve in the case of the Cauchy two-matrix model. The method of proof of Lemma 4.2 is also similar to the one in [4].

LEMMA 4.2 *We have*

$$(4.8) \quad p_1(z) = \int \frac{V'(z) - V'(t)}{z - t} d\mu_1(t) - a^2$$

and

$$(4.9) \quad p_0(z) = a^2 V'(z) + a \int \operatorname{sgn}(t) \frac{V'(z) - V'(t)}{z - t} d\mu_1(t) - \iint \frac{V'(z) - V'(t)}{z - t} \frac{d\mu_1(t) d\mu_2(s)}{t - s}.$$

PROOF: From equation (4.5) and the fact that $F_j(z) = O(1/z)$ as $z \rightarrow \infty$ for $j = 1, 2$, we find that

$$(4.10) \quad \begin{aligned} p_1(z) &= \xi_1(z)\xi_2(z) + \xi_1(z)\xi_3(z) + \xi_2(z)\xi_3(z) \\ &= V'(z)F_1(z) - a^2 + O(1/z) \end{aligned}$$

as $z \rightarrow \infty$. We write

$$(4.11) \quad \begin{aligned} V'(z)F_1(z) &= \int \frac{V'(z) - V'(t)}{z - t} d\mu_1(t) + \int \frac{V'(t)}{z - t} d\mu_1(t) \\ &= \int \frac{V'(t) - V'(z)}{z - t} d\mu_1(t) + O(1/z), \end{aligned}$$

which we insert into (4.10). Since $p_1(z)$ is a polynomial, we can drop all $O(1/z)$ terms and we obtain (4.8).

From equation (4.5) we also find, for $\operatorname{Re} z > 0$,

$$(4.12) \quad \begin{aligned} p_0(z) &= -\xi_1(z)\xi_2(z)\xi_3(z) \\ &= a^2V'(z) + aV'(z)F_1(z) \\ &\quad + V'(z)[F_2^2(z) - 2aF_2(z) - F_1(z)F_2(z)] + O(1/z) \end{aligned}$$

as $z \rightarrow \infty$. The term $V'(z)F_1(z)$ we rewrite as in (4.11). To evaluate the next term in the right-hand side of (4.12), we note that by (3.3)

$$(4.13) \quad F_{2+}(z) + F_{2-}(z) = F_1(z), \quad z \in \operatorname{supp}(\sigma - \mu_2).$$

Also, by the Sokhotski-Plemelj formula,

$$(4.14) \quad F_{2+}(z) - F_{2-}(z) = -2\pi i \frac{d\mu_2(z)}{dz}, \quad z \in \operatorname{supp}(\mu_2),$$

which in particular for $z \in i\mathbb{R} \setminus \operatorname{supp}(\sigma - \mu_2) = (-ic, ic)$ implies that

$$(4.15) \quad F_{2+}(z) - F_{2-}(z) = -2\pi i \frac{d\sigma(z)}{dz} = -2a, \quad z \in i\mathbb{R} \setminus \operatorname{supp}(\sigma - \mu_2).$$

Combining (4.13)–(4.15) we find

$$(4.16) \quad \begin{aligned} (F_2(z) + a)_+ + (F_2(z) - a)_- &= F_1(z), & z \in \operatorname{supp}(\sigma - \mu_2), \\ (F_2(z) + a)_+ - (F_2(z) - a)_- &= 0, & z \in i\mathbb{R} \setminus \operatorname{supp}(\sigma - \mu_2), \\ (F_2(z) + a)_+ - (F_2(z) - a)_- &= 2a - 2\pi i \frac{d\mu_2}{dz}, & z \in \operatorname{supp}(\sigma - \mu_2), \end{aligned}$$

Multiplying the equations in (4.16), we find

$$\begin{aligned} (F_2(z) + a)_+^2 - (F_2(z) - a)_-^2 &= 0, & z \in i\mathbb{R} \setminus \operatorname{supp}(\sigma - \mu_2), \\ (F_2(z) + a)_+^2 - (F_2(z) - a)_-^2 &= \left(2a - 2\pi i \frac{d\mu_2}{dz}\right) F_1(z), & z \in \operatorname{supp}(\sigma - \mu_2), \end{aligned}$$

which implies that

$$(4.17) \quad (F_2(z) \mp a)^2 = a^2 - \frac{1}{2\pi i} \int_{i\mathbb{R}} \frac{(2a - 2\pi i \frac{d\mu_2(s)}{ds})}{z-s} F_1(s) ds, \quad \pm \operatorname{Re} z > 0.$$

For $\operatorname{Re} z > 0$ we rewrite (4.17) as

$$(4.18) \quad \begin{aligned} F_2^2(z) - 2aF_2(z) &= -\frac{a}{\pi i} \int_{i\mathbb{R}} \frac{F_1(s)}{z-s} ds + \int \frac{F_1(s)}{z-s} d\mu_2(s) \\ &= -\frac{a}{\pi i} \int \int_{i\mathbb{R}} \frac{1}{(z-s)(s-t)} ds d\mu_1(t) \\ &\quad + \iint \frac{d\mu_1(t)d\mu_2(s)}{(z-s)(s-t)}. \end{aligned}$$

By a residue calculation we have for $\operatorname{Re} z > 0$ and $t \in \mathbb{R}$,

$$\frac{1}{2\pi i} \int_{i\mathbb{R}} \frac{1}{(z-s)(s-t)} ds = \begin{cases} 0, & \text{if } t > 0, \\ \frac{1}{z-t}, & \text{if } t < 0, \end{cases}$$

so that we obtain from (4.18)

$$F_2^2(z) - 2aF_2(z) = -2a \int_{-\infty}^0 \frac{1}{z-t} d\mu_1(t) + \iint \frac{d\mu_1(t)d\mu_2(s)}{(z-s)(s-t)}$$

for $\operatorname{Re} z > 0$, and

$$\begin{aligned} &F_2^2(z) - 2aF_2(z) - F_1(z)F_2(z) \\ &= -2a \int_{-\infty}^0 \frac{1}{z-t} d\mu_1(t) + \iint \frac{d\mu_1(t)d\mu_2(s)}{(s-t)(z-s)} - \iint \frac{d\mu_1(t)d\mu_2(s)}{(z-t)(z-s)} \\ &= -2a \int_{-\infty}^0 \frac{1}{z-t} d\mu_1(t) - \iint \frac{d\mu_1(t)d\mu_2(s)}{(z-t)(t-s)}, \quad \operatorname{Re} z > 0. \end{aligned}$$

Thus

$$(4.19) \quad \begin{aligned} &V'(z)[F_2^2(z) - 2aF_2(z) - F_1(z)F_2(z)] \\ &= -2a \int_{-\infty}^0 \frac{V'(z) - V'(t)}{z-t} d\mu_1(t) \\ &\quad - \iint \frac{V'(z) - V'(t)}{z-t} \frac{d\mu_1(t)d\mu_2(s)}{t-s} + O(1/z) \end{aligned}$$

as $z \rightarrow \infty$ with $\operatorname{Re} z > 0$. By substituting (4.11) and (4.19) into (4.12) and by dropping the $O(1/z)$ terms, we obtain (4.9) for $\operatorname{Re} z > 0$. Since both sides of (4.9) are polynomials in z , the identity (4.9) is valid everywhere. \square

Remark 4.3. Another expression for p_0 can be found from (4.12) as follows:

The term $V'(z)F_1(z)$ we can rewrite as in (4.11). The similar term $V'(z)F_2(z)$ in (4.12) cannot be rewritten as

$$(4.20) \quad \int \frac{V'(z) - V'(s)}{z - s} d\mu_2(s) + O(z^{-1})$$

since the integral in (4.20) is divergent (unless V is quadratic). However, by (2.8) and (2.12) we have that

$$(4.21) \quad \frac{d\mu_2(z)}{dz} = \rho_2(z)$$

where ρ_2 is analytic in a neighborhood of infinity, and

$$(4.22) \quad \rho_2(z) = O(z^{-2}) \quad \text{as } z \rightarrow \infty.$$

Indeed, we have

$$\rho_2(z) = \begin{cases} \frac{1}{2\pi i} \int \frac{|s|}{-z^2 + s^2} d\mu_1(s), & c = 0, \\ \frac{1}{2\pi i} \int \frac{-z^2 + s^2 + z^2 \sqrt{1 + c^2/z^2}}{(-z^2 + s^2)\sqrt{s^2 + c^2}} d\mu_1(s), & c > 0. \end{cases}$$

Then, by (4.21) and (4.22) we have that $F_2(z)$ for $\operatorname{Re} z > 0$ can be written as

$$(4.23) \quad F_2(z) = \int_{-i\infty}^{i\infty} \frac{\rho_2(s) ds}{z - s} = \oint_{\Gamma} \frac{\rho_2(s) ds}{z - s},$$

where Γ is the contour that goes from $(-iR)$ to (iR) along the imaginary axis and then from (iR) to $(-iR)$ along the semicircle centered at the origin in the left half-plane, with $R > 0$ chosen big enough so that the function $\rho_2(s)$ is analytic for $|s| \geq R$. Then it follows from (4.23) that for $\operatorname{Re} z > 0$,

$$(4.24) \quad \begin{aligned} V'(z)F_2(z) &= \int_{\Gamma} \frac{V'(z)}{z - s} \rho_2(s) ds \\ &= \int_{\Gamma} \frac{V'(z) - V'(s)}{z - s} \rho_2(s) ds + O(z^{-1}) \end{aligned}$$

as $z \rightarrow \infty$.

To handle the other terms in (4.12) we introduce

$$(4.25) \quad \begin{aligned} V_1(z, t) &= \frac{V'(z) - V'(t)}{z - t}, \\ V_2(z, t, s) &= \frac{\frac{V'(z) - V'(t)}{z - t} - \frac{V'(s) - V'(t)}{s - t}}{z - s} \\ &= \frac{(s - t)V'(z) + (t - z)V'(s) + (z - s)V'(t)}{(z - t)(s - t)(z - s)}, \end{aligned}$$

and observe that $V_1(z, t)$ is a symmetric polynomial in z and t , and $V_2(z, t, s)$ is a symmetric polynomial in z, t , and s . Note that

$$(4.26) \quad \begin{aligned} \frac{V'(z)}{(z-t)(z-s)} &= V_2(z, t, s) + \frac{V_1(s, t)}{z-s} + \frac{V'(t)}{(z-t)(z-s)} \\ &= V_2(z, t, s) + O(z^{-1}) \end{aligned}$$

as $z \rightarrow \infty$. Then for $\operatorname{Re} z > 0$, we have by (4.23) and (4.26)

$$(4.27) \quad \begin{aligned} V'(z)F_1(z)F_2(z) &= \int_{\Gamma} \frac{V'(z)}{(z-t)(z-s)} d\mu_1(t)\rho_2(s)ds \\ &= \int_{\Gamma} \int_{\Gamma} V_2(z, t, s)d\mu_1(t)\rho_2(s)ds + O(z^{-1}), \end{aligned}$$

and similarly

$$(4.28) \quad V'(z)F_2^2(z) = \int_{\Gamma} \int_{\Gamma} V_2(z, t, s)\rho_2(s)ds\rho_2(t)dt + O(z^{-1}).$$

By substituting (4.11), (4.24), (4.27), and (4.28) into (4.12) and by dropping $O(z^{-1})$ terms, we obtain for $\operatorname{Re} z > 0$,

$$(4.29) \quad \begin{aligned} p_0(z) &= a^2V'(z) + a \int V_1(z, s) d\mu_1(s) \\ &\quad - 2a \int_{\Gamma} V_1(z, s) \rho_2(s)ds - \int_{\Gamma} \int_{\Gamma} V_2(z, t, s)d\mu_1(t)\rho_2(s)ds \\ &\quad + \int_{\Gamma} \int_{\Gamma} V_2(z, t, s)\rho_2(t)\rho_2(s)dt ds, \end{aligned}$$

where the notation in (4.25) is used. Both sides of (4.29) are polynomials in z , and therefore the equality holds in the full complex plane.

Note that if we replace integration over Γ with respect to $\rho_2(s)ds$ by integration over $i\mathbb{R}$ with respect to $d\mu_2(s)$, then the integrals in (4.29) become divergent.

Example 4.4. We can calculate p_0 , p_1 , and p_2 explicitly in the quadratic case. If

$$(4.30) \quad V(z) = \frac{1}{2}z^2,$$

then $p_2(z) = -z$ by (4.7) and $p_1(z) = 1 - a^2$ by (4.8), since μ_1 has total mass 1. From (4.8) it follows that $p_0(z) = a^2z$ since the integrals in (4.8) vanish for quadratic V , as μ_1 and μ_2 are symmetric measures. So the spectral curve for (4.30) is

$$(4.31) \quad \xi^3 - z\xi^2 + (1 - a^2)\xi + a^2z = 0.$$

This is known as Pastur's equation [35]. It plays an important role in [2, 8, 10].

Example 4.5. Let us now consider the quartic case

$$(4.32) \quad V(z) = \frac{1}{4}z^4 - \frac{t}{2}z^2.$$

Then $p_2(z) = -(z^3 - tz)$ by (4.7). From (4.8) and (4.9) we now conclude that $p_1(z) = z^2 + \mathcal{O}(1)$ and $p_0(z) = a^2z^3 + \mathcal{O}(z)$ as $z \rightarrow \infty$. By symmetry we have that p_1 is an even polynomial and p_0 is an odd polynomial, so that p_1 and p_0 are determined up to two constants. Thus for some $\alpha, \beta \in \mathbb{R}$, we have

$$p_1(z) = z^2 + \alpha, \quad p_0(z) = a^2z^3 + \beta z,$$

and the spectral curve is

$$(4.33) \quad \xi^3 - (z^3 - tz)\xi^2 + (z^2 + \alpha)\xi + a^2z^3 + \beta z = 0.$$

This is McLaughlin's equation, named after K. T.-R. McLaughlin [31], who derived it first for the case $t = 0$; see also [3]. We will analyze the quartic potential in more detail in Section 7 below where (4.33) will play a crucial role.

5 Proof of Theorem 2.4

Recall the RH problem for $Y(z)$ in (2.20)–(2.21). In Sections 5.1–5.6 we will perform a Deift-Zhou steepest-descent analysis of this RH problem. This will then lead to the proof of Theorem 2.4 in Section 5.7.

5.1 First Transformation $Y \mapsto X$

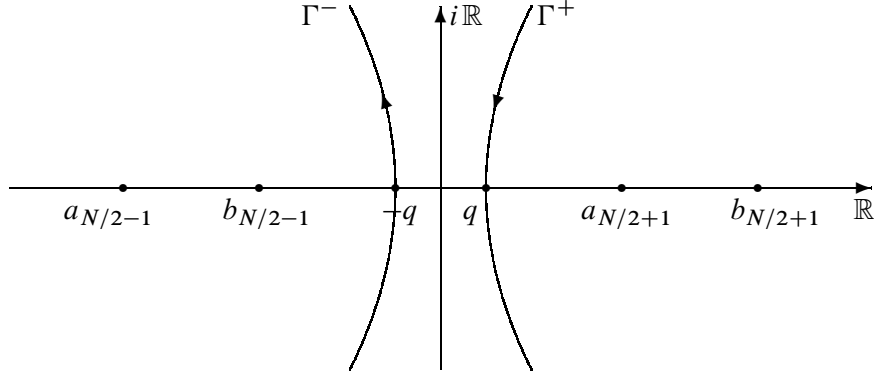
In the first transformation we open up an unbounded lens around $\text{supp}(\sigma - \mu_2)$, which is bounded by a contour $\Gamma = \Gamma^+ \cup \Gamma^-$. We choose the lens so that it is symmetric under reflection with respect to both the real and the imaginary axes. The construction of the lens depends on whether we are in Case I or in Cases II/III.

First assume that we are in Case I so that $\text{supp}(\sigma - \mu_2) = i\mathbb{R}$. Then we take Γ as in Figure 5.1. That is, we take $q \in (0, a_{N/2} + 1)$ and let Γ^+ be an unbounded Jordan curve in the right half-plane, symmetric in the real axis and crossing the real axis in q . Also, Γ^+ is asymptotic to the half-ray $\arg z = \pm\theta$ for some $\theta \in (0, \frac{\pi}{2})$. Then Γ^- is the reflection of Γ^+ in the imaginary axis. We orient Γ as shown in Figure 5.1.

Next, consider Cases II and III. Then we take $\Gamma = \Gamma^+ \cup \Gamma^-$ as in Figure 5.2. The part of Γ^+ in the upper half-plane is a Jordan curve going from ∞ at an angle $\theta \in (0, \frac{\pi}{2})$ to the point ic . The other part of Γ^+ is its reflection in the real axis, and Γ^- is obtained from Γ^+ by reflection in the imaginary axis. We orient Γ as shown in Figure 5.2.

The precise way to choose the contour Γ will be described further on.

The contour Γ divides the complex plane into an inner and an outer part. By definition, we say that $\text{supp}(\sigma - \mu_2)$ is inside the lens, while $\text{supp}(\mu_1)$ is outside the lens. Note that our definitions are such that the outside of the lens is always on the left when traversing Γ according to its orientation.

FIGURE 5.1. Lens around $\text{supp}(\sigma - \mu_2) = i\mathbb{R}$ in Case I.

We define a new 3×3 matrix-valued function X by

$$(5.1) \quad X(z) = \begin{cases} Y(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -e^{-2naz} \\ 0 & 0 & 1 \end{pmatrix} & \text{for } \text{Re } z > 0 \text{ outside the lens,} \\ Y(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -e^{2naz} & 1 \end{pmatrix} & \text{for } \text{Re } z < 0 \text{ outside the lens,} \\ Y(z) & \text{for } z \text{ inside the lens.} \end{cases}$$

Recall that q is used to denote the intersection of Γ^+ with the real axis (in Case I). In Cases II and III we put $q = 0$.

Then X satisfies the following RH problem:

- (1) X is analytic in $\mathbb{C} \setminus (\mathbb{R} \cup [-ic, ic] \cup \Gamma)$.
- (2) X satisfies the jump properties

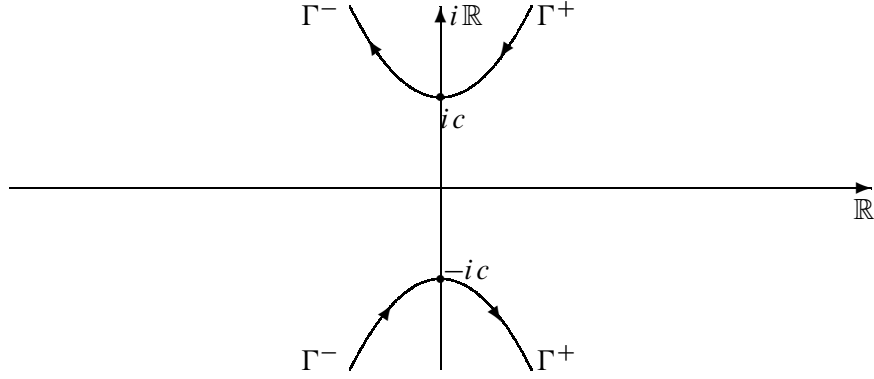
$$(5.2) \quad 2X_+(x) = X_-(x) \begin{pmatrix} 1 & e^{-n(V(x)-ax)} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } x > q,$$

$$(5.3) \quad X_+(x) = X_-(x) \begin{pmatrix} 1 & 0 & e^{-n(V(x)+ax)} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } x < -q,$$

$$(5.4) \quad X_+(x) = X_-(x) \begin{pmatrix} 1 & e^{-n(V(x)-ax)} & e^{-n(V(x)+ax)} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for $-q < x < q$ (in Case I)

on the real line,

FIGURE 5.2. Lens around $\text{supp}(\sigma - \mu_2)$ in Cases II and III.

$$(5.5) \quad X_+(z) = X_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & e^{-2naz} \\ 0 & -e^{2naz} & 1 \end{pmatrix}$$

for $z \in (-ic, ic)$ (in Cases II, III)

on the interval $(-ic, ic)$ oriented upwards, and

$$(5.6) \quad X_+(z) = X_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -e^{-2naz} \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } z \in \Gamma^+,$$

$$(5.7) \quad X_+(z) = X_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -e^{2naz} & 1 \end{pmatrix} \quad \text{for } z \in \Gamma^-,$$

on the contour Γ .

(3) As $z \rightarrow \infty$ we have

$$(5.8) \quad X(z) = (I + O(1/z)) \begin{pmatrix} z^n & 0 & 0 \\ 0 & z^{-n/2} & 0 \\ 0 & 0 & z^{-n/2} \end{pmatrix}.$$

5.2 Second Transformation $X \mapsto T$

In the second transformation $X \mapsto T$ we use the minimizers μ_1 and μ_2 of the vector equilibrium problem and the associated g -functions

$$(5.9) \quad g_k(z) = \int \log(z-s) d\mu_k(s).$$

For $k = 1$ we choose the branch of the logarithm in (5.9) in the standard way, with a branch cut $(-\infty, s]$ along the negative real axis. Thus $g_1(z)$ is defined and analytic in $\mathbb{C} \setminus (-\infty, b_N]$ and

$$g_1(z) = \log z + O(1/z) \quad \text{as } z \rightarrow \infty.$$

For $k = 2$ the logarithm in (5.9) is taken with branch cut $(-i\infty, s]$ along the imaginary axis. Thus $\log(z-s) = \log|z-s| + i \arg(z-s)$ with $-\frac{\pi}{2} < \arg(z-s) < \frac{3\pi}{2}$. Then $g_2(z)$ is defined and analytic in $\mathbb{C} \setminus i\mathbb{R}$, and

$$g_2(z) = \frac{1}{2} \log z + O(1) \quad \text{as } z \rightarrow \infty,$$

with the above branch of the logarithm.

The behavior of the real and imaginary parts of the g -functions is described in the following lemma:

LEMMA 5.1 *For $x \in \mathbb{R} \setminus \{0\}$ we have*

$$(5.10) \quad \begin{aligned} g_{1,+}(x) + g_{1,-}(x) &= -2U^{\mu_1}(x), \\ g_{1,+}(x) - g_{1,-}(x) &= 2\pi i \mu_1([x, +\infty)), \\ g_2(x) &= -U^{\mu_2}(x) + \begin{cases} 0, & x > 0, \\ \frac{\pi i}{2}, & x < 0. \end{cases} \end{aligned}$$

For $z \in i\mathbb{R} \setminus \{0\}$ we have

$$(5.11) \quad \begin{aligned} g_{2,+}(z) + g_{2,-}(z) &= -2U^{\mu_2}(z) + \frac{\pi i}{2}, \\ g_{2,+}(z) - g_{2,-}(z) &= \begin{cases} 2\pi i \mu_2([z, i\infty)), & z \in i\mathbb{R}, \\ \frac{\pi i}{2} - 2az, & z \in [-ic, ic], \end{cases} \\ g_1(z) &= -U^{\mu_1}(z) + \begin{cases} \frac{\pi i}{2}, & z \in i\mathbb{R}_+, \\ -\frac{\pi i}{2}, & z \in i\mathbb{R}_-. \end{cases} \end{aligned}$$

PROOF: The equalities follow immediately from the definitions, where we have to be careful with the choice of branches of the logarithm as discussed above.

Only the equality in (5.11) for $z \in [-ic, ic]$ needs some extra comment. We have that $\mu_2 = \sigma$ on $[-ic, ic]$ and so if $z \in [0, ic]$,

$$\mu_2([0, z]) = \sigma([0, z]) = \frac{a}{\pi} |z| = -\frac{a}{\pi} iz,$$

and therefore

$$(5.12) \quad \mu_2([z, i\infty)) = \mu_2([0, i\infty)) - \mu_2([0, z]) = \frac{1}{4} + \frac{a}{\pi} iz,$$

since μ_2 is symmetric in the origin and $\mu_2(i\mathbb{R}) = \frac{1}{2}$. Equality (5.12) is proved for $z \in [-ic, 0]$ in a similar way, and thus

$$g_{2,+}(z) - g_{2,-}(z) = 2\pi i \mu_2([z, i\infty)) = \frac{\pi i}{2} - 2az, \quad z \in [-ic, ic].$$

□

Now we introduce the λ -functions, which are defined as the following antiderivatives of the ξ -functions (4.5). Recall that ℓ is the constant in the variational conditions (2.13)–(2.14).

$$(5.13) \quad \begin{aligned} \lambda_1(z) &= V(z) + \ell - g_1(z), & z \in \mathcal{R}_1, \\ \lambda_2(z) &= az + g_1(z) - g_2(z), & z \in \mathcal{R}_2, \operatorname{Re} z > 0, \\ \lambda_2(z) &= az + g_2(z), & z \in \mathcal{R}_2, \operatorname{Re} z < 0, \\ \lambda_3(z) &= -az + \frac{\pi i}{2} + g_2(z), & z \in \mathcal{R}_3, \operatorname{Re} z > 0, \\ \lambda_3(z) &= -az + \frac{\pi i}{2} + g_1(z) - g_2(z), & z \in \mathcal{R}_3, \operatorname{Re} z < 0. \end{aligned}$$

Note that $\lambda_2(z)$ and $\lambda_3(z)$ are defined with a cut along the entire imaginary axis. In fact, from (5.11)–(5.13) it follows that

$$(5.14) \quad \lambda_{3,+}(z) - \lambda_{3,-}(z) = \begin{cases} 0, & z \in [ic, i\infty), \\ -\pi i, & z \in (-i\infty, -ic], \end{cases}$$

and a similar formula holds for λ_2 .

We can now reformulate Lemma 5.1 in terms of the λ -functions. This leads to the following two lemmas:

LEMMA 5.2 *We have*

$$(5.15) \quad \begin{aligned} \lambda_{2,\pm}(x) - \lambda_{1,\mp}(x) &\begin{cases} = 0, & x \in \mathbb{R}_+ \cap \operatorname{supp}(\mu_1), \\ < 0, & x \in \mathbb{R}_+ \setminus \operatorname{supp}(\mu_1), \end{cases} \\ \lambda_{3,\pm}(x) - \lambda_{1,\mp}(x) &\begin{cases} = 0, & x \in \mathbb{R}_- \cap \operatorname{supp}(\mu_1), \\ < 0, & x \in \mathbb{R}_- \setminus \operatorname{supp}(\mu_1), \end{cases} \\ \operatorname{Re}(\lambda_{3,+}(z) - \lambda_{3,-}(z)) &\begin{cases} = 0, & z \in i\mathbb{R} \setminus (-ic, ic), \\ < 0, & z \in (-ic, ic). \end{cases} \end{aligned}$$

PROOF: These are reformulations of the variational conditions (2.13)–(2.16) associated with the equilibrium problem, taking into account (5.10)–(5.13). The fact that we have strict inequalities follows from our assumption that the measure μ_1 is regular. \square

LEMMA 5.3 *We have*

$$(5.16) \quad \begin{aligned} \lambda_{1,+}(x) - \lambda_{1,-}(x) &= -2\pi i \mu_1([x, +\infty)), & x \in \mathbb{R}, \\ \lambda_{2,+}(x) - \lambda_{2,-}(x) &= 2\pi i \mu_1([x, +\infty)), & x \in \mathbb{R}_+, \\ \lambda_{3,+}(x) - \lambda_{3,-}(x) &= 2\pi i \mu_1([x, +\infty)), & x \in \mathbb{R}_-, \\ \lambda_{2,+}(z) - \lambda_{3,-}(z) &= \lambda_{2,-}(z) - \lambda_{3,+}(z) \\ &= \begin{cases} 0, & z \in (-ic, ic), \\ 2\pi i \mu_2([z, i\infty)) - \frac{\pi i}{2} + 2az, & z \in i\mathbb{R}. \end{cases} \end{aligned}$$

The last expression is purely imaginary and its imaginary part is strictly increasing in terms of $\text{Im}(z)$ as $z \in (-i\infty, -ic) \cup (ic, i\infty)$.

PROOF: This is a straightforward calculation using (5.10)–(5.13). \square

We define the new matrix-valued function T as

$$(5.17) \quad \begin{aligned} T(z) &= \text{diag}(e^{-n\ell}, 1, e^{-n\frac{1}{2}\pi i})X(z) \\ &\quad \times \text{diag}(e^{n(\lambda_1(z)-V(z))}, e^{n(\lambda_2(z)-az)}, e^{n(\lambda_3(z)+az)}). \end{aligned}$$

Then T satisfies the following RH problem.

- (1) T is analytic on $\mathbb{C} \setminus (\mathbb{R} \cup [-ic, ic] \cup \Gamma)$.
- (2) The jumps for T are

$$(5.18) \quad T_+(x) = T_-(x) \begin{pmatrix} e^{n(\lambda_{1,+}-\lambda_{1,-})} & e^{n(\lambda_{2,+}-\lambda_{1,-})} & 0 \\ 0 & e^{n(\lambda_{2,+}-\lambda_{2,-})} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } x > q,$$

$$(5.19) \quad T_+(x) = T_-(x) \begin{pmatrix} e^{n(\lambda_{1,+}-\lambda_{1,-})} & 0 & e^{n(\lambda_{3,+}-\lambda_{1,-})} \\ 0 & 1 & 0 \\ 0 & 0 & e^{n(\lambda_{3,+}-\lambda_{3,-})} \end{pmatrix} \quad \text{for } x < -q,$$

$$(5.20) \quad T_+(x) = T_-(x) \begin{pmatrix} 1 & e^{n(\lambda_2-\lambda_1)} & e^{n(\lambda_3-\lambda_1)} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } -q < x < q$$

on the real axis,

$$(5.21) \quad T_+(z) = T_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & e^{n(\lambda_{3,+}-\lambda_{3,-})} \end{pmatrix} \quad \text{for } z \in (-ic, ic)$$

on the imaginary interval, and

$$(5.22) \quad T_+(z) = T_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -e^{n(\lambda_3-\lambda_2)} \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } z \in \Gamma^+,$$

$$(5.23) \quad T_+(z) = T_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -e^{n(\lambda_2 - \lambda_3)} & 1 \end{pmatrix} \quad \text{for } z \in \Gamma^-$$

on the curve Γ .

(3) As $z \rightarrow \infty$ we have that

$$T(z) = I + O(1/z).$$

Note that for the (2, 3) and (3, 2) entries of the jump matrix in (5.21), we used the last formula in (5.16). Also observe that $T(z)$ does not have jumps on $i\mathbb{R} \setminus [-ic, ic]$. Indeed, in principle one should consider the jump relation

$$(5.24) \quad T_+(z) = T_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{n(\lambda_{2,+} - \lambda_{2,-})} & 0 \\ 0 & 0 & e^{n(\lambda_{3,+} - \lambda_{3,-})} \end{pmatrix},$$

$$z \in (-i\infty, -ic) \cup (ic, i\infty).$$

But from the jump relations in (5.14) we see that

$$e^{n(\lambda_{3,+}(z) - \lambda_{3,-}(z))} = e^0 = 1, \quad z \in (ic, i\infty),$$

and also

$$e^{n(\lambda_{3,+}(z) - \lambda_{3,-}(z))} = e^{-n\pi i} = 1, \quad z \in (-i\infty, -ic),$$

by the fact that n is even. In a similar way one shows that $e^{n(\lambda_{2,+}(z) - \lambda_{2,-}(z))} = 1$, and so the jump matrix in (5.24) is just the identity matrix.

By virtue of (5.15)–(5.16), we may rewrite the jumps on \mathbb{R} as follows:

$$(5.25) \quad T_+ = T_- \begin{pmatrix} e^{n(\lambda_{1,+} - \lambda_{1,-})} & 1 & 0 \\ 0 & e^{-n(\lambda_{1,+} - \lambda_{1,-})} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in \text{supp}(\mu_1), \quad x > q,$$

$$(5.26) \quad T_+ = T_- \begin{pmatrix} e^{n(\lambda_{1,+} - \lambda_{1,-})} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-n(\lambda_{1,+} - \lambda_{1,-})} \end{pmatrix}, \quad x \in \text{supp}(\mu_1), \quad x < -q,$$

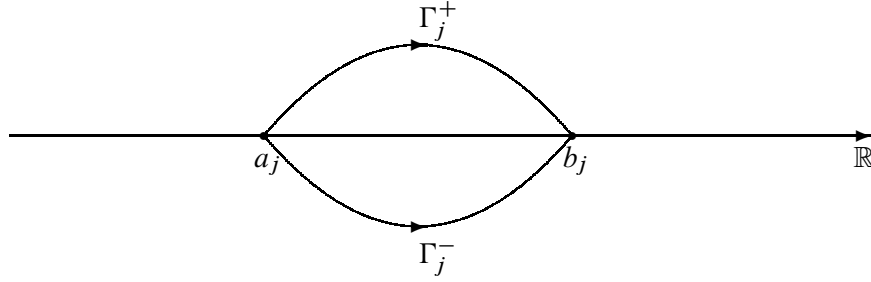
and

$$(5.27) \quad T_+ = T_- \begin{pmatrix} e^{-2\pi i n \alpha_j} & e^{n(\lambda_{2,+} - \lambda_{1,-})} & 0 \\ 0 & e^{2\pi i n \alpha_j} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in (b_j, a_{j+1}), \quad x > q,$$

$$(5.28) \quad T_+ = T_- \begin{pmatrix} e^{-2\pi i n \alpha_j} & 0 & e^{n(\lambda_{3,+} - \lambda_{1,-})} \\ 0 & 1 & 0 \\ 0 & 0 & e^{2\pi i n \alpha_j} \end{pmatrix}, \quad x \in (b_j, a_{j+1}), \quad x < -q,$$

where $b_0 = -\infty$, $a_{N+1} = +\infty$, and

$$(5.29) \quad \alpha_j = \mu_1([a_{j+1}, +\infty)), \quad j = 0, \dots, N.$$

FIGURE 5.3. Lens around the interval $[a_j, b_j]$.

We would like the jump matrices on Γ to be exponentially close to the identity matrix as $n \rightarrow \infty$. From the above RH problem, we see that this is achieved provided Γ lies in the region where $\operatorname{Re}(\lambda_3 - \lambda_2) < 0$ if $\operatorname{Re} z > 0$ and $\operatorname{Re}(\lambda_3 - \lambda_2) > 0$ if $\operatorname{Re} z < 0$. The fact that Γ can indeed be chosen in this way follows by applying the Cauchy-Riemann equations to the last equality in Lemma 5.3 and using the last line in the statement of that lemma.

5.3 Lenses around the Intervals $[a_j, b_j]$

Around each of the intervals $[a_j, b_j]$ we open up a small lens to transform the oscillatory entries of the jump matrix into exponentially decaying entries. Since the nontrivial part of the jump matrix is locally of size 2×2 only, this can be done in the standard way [15].

More precisely, we take Jordan curves Γ_j^+ and Γ_j^- surrounding the interval $[a_j, b_j]$ as in Figure 5.3. The region between these curves is called the lens, and Γ_j^+ and Γ_j^- are the upper and lower lips of the lens, respectively. We choose them sufficiently close to the real axis so that

$$(5.30) \quad \operatorname{Re}(\lambda_1(z) - \lambda_2(z)) < 0 \quad \text{for } z \text{ inside the lens, } z \notin [a_j, b_j], \operatorname{Re}(z) > 0,$$

$$(5.31) \quad \operatorname{Re}(\lambda_1(z) - \lambda_3(z)) < 0 \quad \text{for } z \text{ inside the lens, } z \notin [a_j, b_j], \operatorname{Re}(z) < 0.$$

The fact that this is possible follows from applying the Cauchy-Riemann equations to the first equation of (5.16); cf. [15].

The curves Γ_j^\pm are chosen disjoint from each other, disjoint from the lens Γ in Section 5.1, and also disjoint from the imaginary axis. The only case where an intersection with the imaginary axis occurs is in Case III with $j = (N + 1)/2$, and we assume that in this case, Γ_j^\pm intersects $i\mathbb{R}$ at a point $\pm iz_0$ strictly inside $(-ic, ic)$.

We define

$$(5.32) \quad S(z) = \begin{cases} T(z) \begin{pmatrix} 1 & 0 & 0 \\ -e^{n(\lambda_1(z)-\lambda_2(z))} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & z \text{ in upper part lens, } \operatorname{Re} z > 0, \\ T(z) \begin{pmatrix} 1 & 0 & 0 \\ e^{n(\lambda_1(z)-\lambda_2(z))} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & z \text{ in lower part lens, } \operatorname{Re} z > 0, \\ T(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -e^{n(\lambda_1(z)-\lambda_3(z))} & 0 & 1 \end{pmatrix}, & z \text{ in upper part lens, } \operatorname{Re} z < 0, \\ T(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ e^{n(\lambda_1(z)-\lambda_3(z))} & 0 & 1 \end{pmatrix}, & z \text{ in lower part lens, } \operatorname{Re} z < 0, \\ T(z), & \text{elsewhere.} \end{cases}$$

Then S satisfies the following RH problem:

- (1) S is analytic in $\mathbb{C} \setminus (\mathbb{R} \cup i\mathbb{R} \cup \Gamma \cup \bigcup_j \Gamma_j)$.
- (2) For $x \in \operatorname{supp}(\mu_1)$ we have that

$$S_+(x) = S_-(x) \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in \operatorname{supp}(\mu_1) \cap \mathbb{R}_+,$$

$$S_+(x) = S_-(x) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad x \in \operatorname{supp}(\mu_1) \cap \mathbb{R}_-.$$

On the lips Γ_j of the lenses we have

$$S_+(z) = S_-(z) \begin{pmatrix} 1 & 0 & 0 \\ e^{n(\lambda_1-\lambda_2)} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for } z \in \Gamma_j^\pm, \operatorname{Re} z > 0,$$

$$S_+(z) = S_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ e^{n(\lambda_1-\lambda_3)} & 0 & 1 \end{pmatrix} \quad \text{for } z \in \Gamma_j^\pm, \operatorname{Re} z < 0.$$

In Case III we also have

$$(5.33) \quad S_+(z) = S_-(z) \quad \text{for } z \in (-iz_0, iz_0).$$

The jumps of $S(z)$ on the other contours are the same as those for $T(z)$.

- (3) As $z \rightarrow \infty$, we have that

$$S(z) = I + O(1/z).$$

From (5.15) and (5.30)–(5.31), it can be checked that all the nonconstant entries in the jump matrices for $S(z)$ tend to 0 as $n \rightarrow \infty$, uniformly for z bounded away from the branch points a_j , b_j , and $\pm ic$. The only case that requires more explanation is the (3, 1) entry in the jump matrix in (5.33). In that case, one can factorize

$$e^{n(\lambda_1(z)-\lambda_{3,-}(z))} = e^{n(\lambda_1(z)-\lambda_{3,+}(z))} e^{n(\lambda_{3,+}(z)-\lambda_{3,-}(z))},$$

and observe from (5.15) and (5.31) that for $z \in [-iz_0, iz_0]$, the leftmost factor is uniformly bounded by 1 while the rightmost factor is uniformly exponentially decaying as $n \rightarrow \infty$.

5.4 Global Parametrix

The global parametrix we look for is a 3×3 matrix-valued function M with jumps (obtained from the jumps of S by ignoring all entries that are exponentially small for $n \rightarrow \infty$).

- (1) M is analytic in $\mathbb{C} \setminus ([a_1, b_N] \cup [-ic, ic])$.
- (2) The jumps for M are

$$(5.34) \quad M_+(x) = M_-(x) \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in \text{supp}(\mu_1) \cap \mathbb{R}_+,$$

$$(5.35) \quad M_+(x) = M_-(x) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad x \in \text{supp}(\mu_1) \cap \mathbb{R}_-,$$

$$(5.36) \quad M_+(x) = M_-(x) \begin{pmatrix} e^{-2\pi i n \alpha_j} & 0 & 0 \\ 0 & e^{2\pi i n \alpha_j} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in (b_j, a_{j+1}) \cap \mathbb{R}_+,$$

$$(5.37) \quad M_+(x) = M_-(x) \begin{pmatrix} e^{-2\pi i n \alpha_j} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{2\pi i n \alpha_j} \end{pmatrix}, \quad x \in (b_j, a_{j+1}) \cap \mathbb{R}_-$$

on the real line, and

$$(5.38) \quad M_+(z) = M_-(z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad z \in (-ic, ic).$$

- (3) At infinity we have

$$M(z) = I + O(1/z) \quad \text{as } z \rightarrow \infty,$$

- (4) M has at most fourth-root singularities at the branch points $a_1, b_1, \dots, a_N, b_N, \pm ic$.

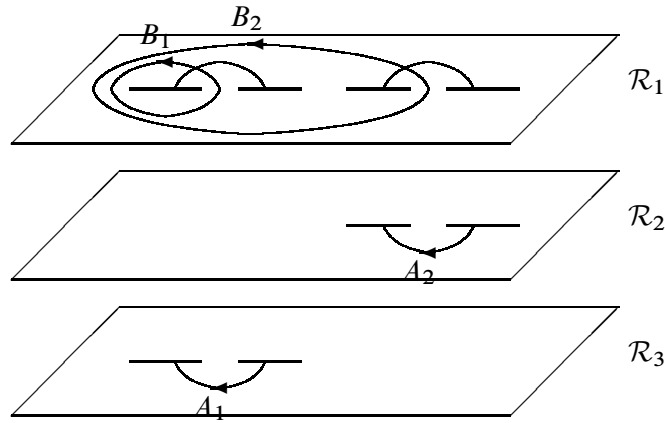


FIGURE 5.4. Canonical homology basis in Case I (with $N = 4$).

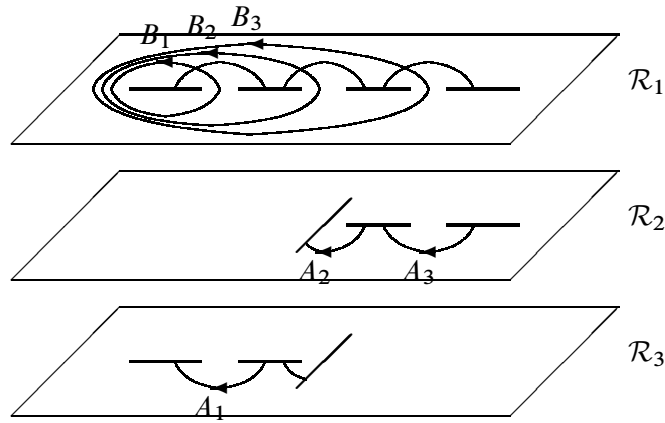


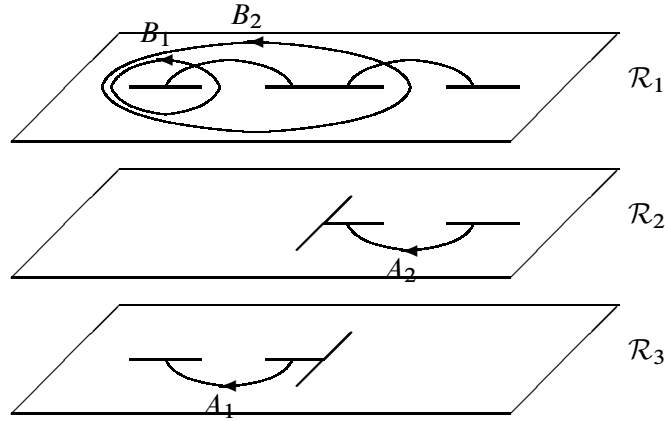
FIGURE 5.5. Canonical homology basis in Case II (with $N = 4$).

We can solve this problem with the help of meromorphic differentials on the Riemann surface. Such a construction was first used in [32] and later developed further in [22, 30].

To the Riemann surface we associate a canonical homology basis $\{A_1, \dots, A_g, B_1, \dots, B_g\}$ where g is the genus. The details of the construction depend on whether we are in Case I, II, or III; see Figures 5.4–5.6.

For brevity, we give a detailed description only for Cases II and III. Then the genus is $g = N - 1$. The cycles B_j are on the first sheet and B_j encircles $[a_j, b_j]$ once in the counterclockwise direction. The cycles A_j are partly in the upper half-plane on the first sheet and partly in the lower half-plane on the second or third sheet. A_j passes through $[a_j, b_j]$ and $[a_{j+1}, b_{j+1}]$.

The antiholomorphic involution ϕ is defined by mapping z to \bar{z} on the same sheet. The fixed point set of ϕ is the disjoint union of $N = g + 1$ closed curves

FIGURE 5.6. Canonical homology basis in Case III (with $N = 3$).

$\bigcup_{j=0}^g \Sigma_j$ on the Riemann surface. Here Σ_j is homotopic to A_j as a closed curve, $j = 1, \dots, g$, while Σ_0 is the unbounded component.

LEMMA 5.4 *If $P_j \in \Sigma_j$ for $j = 1, \dots, g$, then the divisor*

$$D = \sum_{j=1}^g P_j$$

is nonspecial.

PROOF: A detailed proof of this theorem in a similar setting will be given in [22]. Here we only outline some of the key steps. We use $u : \mathcal{R} \rightarrow \mathbb{C}^g/L$ to denote the Abel map, mapping the Riemann surface \mathcal{R} to \mathbb{C}^g/L , where L is the lattice defined from the columns of the Riemann period matrix. We also use $\theta : \mathbb{C}^g \rightarrow \mathbb{C}^g$ to denote the corresponding Riemann theta function on \mathbb{C}^g , and \vec{K} denotes the vector of Riemann constants; see, e.g., [24].

The proof is based on the following result, which can be found in [20, theorem 2.4.2]. The divisor D is nonspecial if and only if $\theta(u(P) - u(D) - \vec{K})$ does not vanish identically for P on the Riemann surface.

The next step is to show that for $\vec{s} \in \mathbb{R}^g$ we have $\theta(\vec{s}) > 0$. This relies on the fact that the Riemann surface \mathcal{R} is the Schottky double of a bordered Riemann surface; see [25, cor. 6.13] and [32].

Using the antiholomorphic involution ϕ , it can be shown that the Riemann period matrix is purely imaginary. Finally, one then shows that $u(D) + \vec{K}$ has a real representative modulo the lattice L [22]. By taking into account the results in the last two paragraphs, the desired result then follows. \square

We now basically follow [30]. Given (P_1, \dots, P_g) with $P_j \in \Sigma_j$, we define a meromorphic differential ω_P so that

- ω_P has simple poles in $a_1, b_1, \dots, a_N, b_N, \pm ic, P_1, \dots, P_g, \infty_2$, and ∞_3 with residues

$$\begin{aligned} \operatorname{Res}(\omega_P, a_j) &= \operatorname{Res}(\omega_P, b_j) = -\frac{1}{2}, \quad j = 1, \dots, N, \\ \operatorname{Res}(\omega_P, \pm ic) &= -\frac{1}{2}, \quad \operatorname{Res}(\omega_P, P_j) = 1, \quad j = 1, \dots, g, \\ \operatorname{Res}(\omega_P, \infty_2) &= \operatorname{Res}(\omega_P, \infty_3) = 1, \end{aligned}$$

- ω_P is holomorphic elsewhere,
- ω_P has vanishing A -periods:

$$\int_{A_j} \omega_P = 0, \quad j = 1, \dots, g.$$

The total sum of the residues is 0. These conditions determine ω_P uniquely. The B -periods are purely imaginary and are well-defined modulo $2\pi i\mathbb{Z}$, and the mapping

$$\Psi : \Sigma_1 \times \dots \times \Sigma_g \rightarrow (\mathbb{R}/\mathbb{Z})^g : (P_1, \dots, P_g) \mapsto \frac{1}{2\pi i} \left(\int_{B_1} \omega_P, \dots, \int_{B_g} \omega_P \right)$$

is a bijection. These claims follow in the same way as in [30].

Thus there exist $P_j^{(1)} \in \Sigma_j$ so that

$$\Psi(P_1^{(1)}, \dots, P_g^{(1)}) = (n\alpha_1, \dots, n\alpha_g) \pmod{\mathbb{Z}}.$$

Let $\omega_P^{(1)}$ be the corresponding meromorphic differential.

We take the base point $P_0 = \infty_1$ and define three functions $v_1(z)$, $v_2(z)$, and $v_3(z)$ of a complex variable z as follows: We have

$$v_k(z) = \exp \left(\int_{P_0}^z \omega_P^{(1)} \right)$$

where z is considered as a point on the k^{th} sheet of \mathcal{R} , and where the path of integration is as follows:

- For $k = 1$, the path of integration is on the first sheet and does not intersect the real line.
- For $k = 2$, the path of integration is on the first and second sheets; for $\operatorname{Im} z > 0$ ($\operatorname{Im} z < 0$), the path starts in the lower (upper) half-plane of the first sheet, crosses one of the cuts $[a_j, b_j]$ in \mathbb{R}_+ to the upper (lower) half-plane of the second sheet, and stays there until it ends at z .
- For $k = 3$, the path of integration is on the first and third sheets; for $\operatorname{Im} z > 0$ ($\operatorname{Im} z < 0$), the path starts in the lower (upper) half-plane of the first sheet, crosses one of the cuts $[a_j, b_j]$ in \mathbb{R}_- to the upper (lower) half-plane of the third sheet, and stays there until it ends at z .

Then the vector (v_1, v_2, v_3) is well-defined and analytic in $\mathbb{C} \setminus (\mathbb{R} \cup i\mathbb{R})$ with jumps

$$(v_1, v_2, v_3)_+ = (v_1, v_2, v_3)_- J_v$$

where the jump matrices on the real line are

$$\begin{aligned} J_v(x) &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} && \text{for } x \in \text{supp}(\mu_1) \cap \mathbb{R}_+, \\ J_v(x) &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} && \text{for } x \in \text{supp}(\mu_1) \cap \mathbb{R}_-, \\ J_v(x) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} && \text{for } x < a_1 \text{ and } x > b_N, \\ J_v(x) &= \begin{pmatrix} e^{-2\pi i n \alpha_j} & 0 & 0 \\ 0 & -e^{2\pi i n \alpha_j} & 0 \\ 0 & 0 & -1 \end{pmatrix} && \text{for } b_j < x < a_{j+1}, x > 0, \\ J_v(x) &= \begin{pmatrix} e^{-2\pi i n \alpha_j} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -e^{2\pi i n \alpha_j} \end{pmatrix} && \text{for } b_j < x < a_{j+1}, x < 0, \end{aligned}$$

while on the imaginary axis they are

$$\begin{aligned} J_v(z) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{for } z \in (-i\infty, -ic) \cup (ic, i\infty), \\ J_v(z) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} && \text{for } z \in (-ic, ic). \end{aligned}$$

Also, as $z \rightarrow \infty$, we have

$$v_1(z) = 1 + O(1/z), \quad v_2(z) = O(1/z), \quad v_3(z) = O(1/z),$$

and all v_j -functions have fourth-root singular behavior at the branch points $a_1, b_1, \dots, a_N, b_N, \pm ic$.

Now define the first row (M_{11}, M_{12}, M_{13}) as follows:

$$(M_{11}, M_{12}, M_{13}) = \begin{cases} (v_1, v_2, v_3) & \text{if } \text{Im } z > 0, \\ (v_1, -v_2, -v_3) & \text{if } \text{Im } z < 0. \end{cases}$$

This gives the correct jumps for M . The other rows of M can be constructed in a similar way or by a simple transformation of the first row [30].

5.5 Local Parametrices

In the regular case (the one we are considering) we construct local parametrices out of Airy functions near each of the branch points. We denote the local parametrices by P . The local parametrices match with the global parametrix on the boundary of a small circle around the branch points. Since the nontrivial part of the jump matrix is locally of size 2×2 only, and since in the regular case we have square root behavior near all the branch points (Lemma 2.3), this construction can be done in the usual way [15]. We omit the details.

5.6 Final RH problem

We define the final RH matrix

$$R(z) = \begin{cases} T(z)M(z)^{-1} & \text{away from the branch points,} \\ T(z)P(z)^{-1} & \text{near the branch points.} \end{cases}$$

Then $R(z)$ has jump matrices tending to the identity matrix as $n \rightarrow \infty$, and it is normalized in the sense that $R(z) = I + O(1/z)$ as $z \rightarrow \infty$. Standard arguments then show that $R(z)$ itself also tends to the identity matrix, uniformly for z on compact subsets of \mathbb{C} [15].

5.7 Proof of Theorem 2.4

Having performed the steepest-descent analysis of the RH problem, we can now prove Theorem 2.4. The proof follows the same pattern as in the papers [2, 8, 10, 19].

First assume that $x, y \in (a_j, b_j)$ with $x, y > 0$. We will transform (2.22) under the series of transformations $Y \mapsto X \mapsto T \mapsto S$. By virtue of (5.1) we have

$$K_n(x, y) = \frac{1}{2\pi i(x-y)} \begin{pmatrix} 0 & w_1(y) & 0 \end{pmatrix} X_+(y)^{-1} X_+(x) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

By using (5.17) this becomes

$$K_n(x, y) = \frac{e^{n(V(x)-V(y))}}{2\pi i(x-y)} \begin{pmatrix} 0 & e^{n\lambda_{2,+}(y)} & 0 \end{pmatrix} T_+(y)^{-1} T_+(x) \begin{pmatrix} e^{-n\lambda_{1,+}(x)} \\ 0 \\ 0 \end{pmatrix}.$$

From (5.32) we get

$$(5.39) \quad K_n(x, y) = \frac{e^{n(V(x)-V(y))}}{2\pi i(x-y)} \begin{pmatrix} -e^{n\lambda_{1,+}(y)} & e^{n\lambda_{2,+}(y)} & 0 \end{pmatrix} \\ \times S_+(y)^{-1} S_+(x) \begin{pmatrix} e^{-n\lambda_{1,+}(x)} \\ e^{-n\lambda_{2,+}(x)} \\ 0 \end{pmatrix}.$$

Now it follows by standard arguments (e.g., [8, sec. 9]) that

$$S^{-1}(y)S(x) = I + O(x-y) \quad \text{as } y \rightarrow x$$

uniformly in n . Inserting this into (5.39) and setting $h(x) = V(x) - \operatorname{Re}(\lambda_1(x))$ yields

$$K_n(x, y) = e^{n(h(x)-h(y))} \frac{\sin(n \operatorname{Im}(\lambda_{1,+}(x) - \lambda_{1,+}(y)))}{\pi(x-y)} + O(1), \quad y \rightarrow x,$$

uniformly in n . By letting $y \rightarrow x$ and using l'Hôpital's rule we find

$$K_n(x, x) = \frac{n \operatorname{Im}(\xi_{1,+})}{\pi} + O(1)$$

as $n \rightarrow \infty$. From (4.5) and the Stieltjes inversion principle we conclude that

$$\lim_{n \rightarrow \infty} \frac{1}{n} K_n(x, x) = \frac{d\mu_1(x)}{dx}.$$

This proves (2.23) for $x \in (a_j, b_j)$, $x > 0$. The proof for the other values of $x \in \mathbb{R}$ is similar or can be obtained from symmetry considerations.

6 Phase Transitions: General Discussion

Recall Cases I, II, and III describing the structure of the equilibrium measures in Section 3.3. Intuitively one expects the following possible behavior in terms of the parameter a . For large a we are in Case I. The measure μ_1 is supported on two (or more) disjoint intervals with a gap around 0. The constraint σ is not active.

When a decreases, the gap around 0 shrinks. Then one of two things could happen. It could happen that for a certain value of a the constraint becomes active, while the gap in the support of μ_1 around 0 is still there. Then we are in Case II. Then if a further decreases the gap may close or not. The latter depends on whether in the unitary matrix model with potential V (without external source) 0 is in the support or not. If the support closes then we are in Case III.

The other situation that could happen is that the constraint σ remains inactive all the way until for a certain value of a the gap in the support of μ_1 is closed. Then if a further decreases the constraint becomes active. The transition is then from Case I to Case III without passing through Case II. This is precisely what happens in the quadratic case $V(x) = \frac{1}{2}x^2$. More generally, one expects this kind of behavior when the potential $V(x)$ is convex or “nearly” convex.

For those values of a for which a transition between one of Cases I, II, or III to another takes place, one expects that the local eigenvalue correlations near the origin are described by special functions related to ODEs. For typical cases one expects such special functions as Pearcey integrals [1, 10, 37] and the Hastings-McLeod solution to the Painlevé II equation [7, 13]. However, our model allows for new kinds of critical and multicritical behavior as well, but it remains an open problem to describe these new critical phenomena.

In Section 7 we will illustrate the above considerations in detail for the case of a quartic potential. See in particular Figure 7.1.

7 A Case Study: The Quartic Potential

Let us investigate the case of a quartic potential

$$V(z) = \frac{1}{4}z^4 - \frac{t}{2}z^2$$

and the associated McLaughlin equation (4.33):

$$(7.1) \quad \xi^3 - (z^3 - tz)\xi^2 + (z^2 + \alpha)\xi + a^2z^3 + \beta z = 0.$$

The discriminant of the McLaughlin equation (with respect to ξ) is a polynomial $D_{12}(z)$ of degree 12 in z . We calculated it with Maple, but it is too long and not too interesting to reproduce here in full. The first terms are

$$(7.2) \quad D_{12}(z) = -4\alpha^3 + (\alpha^2t^2 + 18\alpha\beta t - 12\alpha^2 - 27\beta^2)z^2 + O(z^4) \quad \text{as } z \rightarrow 0.$$

The branch points of the Riemann surface are among the zeros of $D_{12}(z)$. There are other zeros, and they should come with higher multiplicities.

For general α and β the McLaughlin equation has genus 4 (according to Maple). The special choices for α and β that are relevant to us will lead to a reduction of the genus. The genus can be at most 1, as the following lemma shows.

A similar result occurs in [21, prop. 5.2.5], but the proof given there is incorrect. Here we give a self-contained proof that may be used for the situation in [21] as well. The proof uses an idea due to Lun Zhang (personal communication).

LEMMA 7.1 *Assume that $x \mapsto V(\sqrt{x})$ is convex for $x > 0$. Then the support of μ_1 is either one interval (in case $0 \in \text{supp}(\mu_1)$) or a disjoint union of two intervals (in case $0 \notin \text{supp}(\mu_1)$), and the measure μ_1 can have singular behavior only at 0.*

PROOF: By fixing $\mu_2 \leq \sigma$ in the energy functional (2.1), we see that μ_1 is the unique minimizer for the energy functional

$$I(\mu) + \int_{-\infty}^{\infty} (V(x) - a|x| - U^{\mu_2}(x))d\mu(x)$$

among probability measures μ on \mathbb{R} . Since the external field is symmetric, it follows from [36, theorem IV.1.10(f)] that $d\mu_1(t) = d\tilde{\mu}_1(t^2)/2$ where $\tilde{\mu}_1$ is the unique minimizer of the energy functional

$$(7.3) \quad I(\mu) + 2 \int_0^{\infty} (V(\sqrt{x}) - a\sqrt{x} - U^{\mu_2}(\sqrt{x})) d\mu(x)$$

over all probability measures μ on $[0, \infty)$.

We are going to show that the external field $V(\sqrt{x}) - a\sqrt{x} - U^{\mu_2}(\sqrt{x})$ in (7.3) is convex for $x > 0$. Since

$$U^{\mu_2}(\sqrt{x}) = -\frac{1}{2} \int \log(x + |z|^2) d\mu_2(z),$$

we have

$$\frac{d^2}{dx^2} (U^{\mu_2}(\sqrt{x})) = \frac{1}{2} \int \frac{1}{(x + |z|^2)^2} d\mu_2(z),$$

which due to the constraint $\mu_2 \leq \sigma$ can be bounded by

$$\frac{d^2}{dx^2} (U^{\mu_2}(\sqrt{x})) < \frac{a}{2\pi} \int_{i\mathbb{R}} \frac{1}{(x + |z|^2)^2} |dz| = \frac{a}{4x^{3/2}} = \frac{d^2}{dx^2} (-a\sqrt{x}).$$

It follows that $-a\sqrt{x} - U^{\mu_2}(\sqrt{x})$ is convex. Due to the assumption in the lemma on the convexity of $V(\sqrt{x})$, it then follows that $(V(\sqrt{x}) - a\sqrt{x} - U^{\mu_2}(\sqrt{x}))$ is indeed convex for $x > 0$.

From the convexity it follows that $\tilde{\mu}_1$ is supported on one interval. Then μ_1 is supported on either one or two intervals, depending on whether 0 belongs to the support of $\tilde{\mu}_1$ or not. The convexity also implies that the density of $\tilde{\mu}_1$ does not vanish in the interior of its support and has square root behavior at its nonzero endpoint(s); see, e.g., [13, lemma 3.5]. The same properties then apply to μ_1 . \square

The three cases in Section 3.3 now specialize as follows:

Case I. In the first case there are four real branch points $\pm b_1$ and $\pm b_2$ (with $b_1 > b_2 > 0$) and no other branch points. The remaining zeros of the discriminant (7.2) come as four double zeros. The Riemann surface has genus 0.

Case II. In the second case there are four real branch points $\pm b_1$ and $\pm b_2$ (with $b_1 > b_2 > 0$) and two purely imaginary branch points $\pm ic$ (with $c > 0$). The remaining zeros of the discriminant come in the form of a sixfold zero at $z = 0$, see Lemma 7.2 below. The Riemann surface has genus 1.

Case III. In the third case there are two real branch points at $\pm b_1$ (with $b_1 > 0$) and two purely imaginary branch points at $\pm ic$ (with $c > 0$). There are four double zeros of the discriminant. The Riemann surface has genus 0.

It remains to investigate in more detail Cases I, II, and III. More precisely, for any $t \in \mathbb{R}$ and $a > 0$ we want to determine which of the three cases applies. We also want to find the curves in the ta -plane where a transition from one case to another takes place.

The Genus 1 Region

Let us first investigate Case II. This is the case of genus 1. It turns out that in this case, the parameters α and β in the McLaughlin equation (7.1) take on a particularly simple form: they are both equal to 0. This is the content of the next lemma.

LEMMA 7.2 (Case II) *Let $t \in \mathbb{R}$ and $a > 0$ be such that the McLaughlin equation (7.1) is of genus 1. Then one has that $\alpha = \beta = 0$. Moreover, the discriminant of the McLaughlin equation can then be factored as*

$$(7.4) \quad D_{12}(z) = z^6 D_6(z)$$

where D_6 is the degree 6 polynomial

$$(7.5) \quad D_6(z) = 4a^2 z^6 + (1 - 12ta^2)z^4 + ((12t^2 - 18)a^2 - 2t)z^2 + (-27a^4 + (18t - 4t^3)a^2 - 4 + t^2).$$

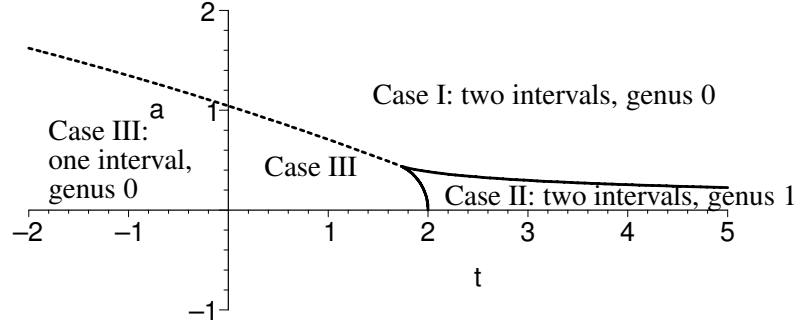


FIGURE 7.1. Phase diagram for the quartic potential $\frac{1}{4}x^4 - \frac{t}{2}x^2$. The dark curve represents the Painlevé II transition. The dotted curve represents the Pearcey transition.

PROOF: From the general descriptions above we know that in the genus 1 case, the McLaughlin equation has six simple branch points $\pm b_1$, $\pm b_2$, and $\pm ic$ ($b_1, b_2, c > 0$), which are simple zeros of the discriminant. The remaining zeros of the discriminant should come as three double zeros, possibly coalescing. By symmetry 0 is a double zero, and this forces $\alpha = 0$; cf. (7.2).

If we substitute $\alpha = 0$ in (7.1) and calculate the discriminant with respect to ξ , we obtain the 12th degree polynomial

$$(7.6) \quad \begin{aligned} & 4a^2z^{12} - (-4\beta + 12a^2t - 1)z^{10} - (18a^2 - 12t^2a^2 + 2t + 12t\beta)z^8 \\ & - (4 - 12t^2\beta - t^2 + 18\beta - 18a^2t + 4t^3a^2 + 27a^4)z^6 \\ & - (54a^2\beta - 18t\beta + 4t^3\beta)z^4 - 27\beta^2z^2. \end{aligned}$$

We can take out the factor z^2 in (7.6), corresponding to the double zero at $z = 0$. The remaining factor is of degree 10 and we consider it as a fifth-degree polynomial $D_5(y)$ in $y = z^2$. We already know that this polynomial has three simple roots b_1^2 , b_2^2 , and $-c^2$ and one double root, call it d . Since by (7.6) the sum of the roots equals $(-4\beta + 12a^2t - 1)/(4a^2) \in \mathbb{R}$, it follows that d is real. Moreover, from (7.6) we see that the product of the roots equals $27\beta^2/(4a^2) \geq 0$. Since this product can also be written as $-b_1^2b_2^2c^2d^2 \leq 0$, it then follows that $\beta = d = 0$.

In conclusion, we have shown now that $\alpha = \beta = 0$. Inserting this in the McLaughlin equation and computing its discriminant by a direct calculation then leads to (7.4)–(7.5). \square

Note that the factor z^6 in (7.4) corresponds to the sixfold zero at $z = 0$, while the zeros of $D_6(z)$ should yield the branch points $\pm b_1$, $\pm b_2$, and $\pm ic$. In particular, four of these zeros should be real and the other two purely imaginary. The next lemma describes when this happens.

LEMMA 7.3 *The polynomial (7.5) has four real and two purely imaginary zeros precisely for those $(t, a) \in \mathbb{R} \times \mathbb{R}_+$ lying in the open region \mathcal{D} bounded by the*

points $(t_1, a_1) = (2, 0)$, $(t_2, a_2) = (\sqrt{3}, 1/\sqrt[4]{27})$, and $(t_3, a_3) = (+\infty, 0)$, the straight-line segment between (t_1, a_1) and (t_3, a_3) , and the two branches of the curve

$$(7.7) \quad D_6(0) = -27a^4 + (18t - 4t^3)a^2 - 4 + t^2 = 0$$

that connect the point (t_2, a_2) with (t_1, a_1) and (t_3, a_3) , respectively. The region \mathcal{D} is shown in the bottom right part of Figure 7.1.

PROOF: Rewrite $D_6(z)$ as a cubic polynomial in the variable $y = z^2$:

$$4a^2y^3 + (1 - 12ta^2)y^2 + ((12t^2 - 18)a^2 - 2t)y + (-27a^4 + (18t - 4t^3)a^2 - 4 + t^2).$$

We are looking for the values of $(t, a) \in \mathbb{R} \times \mathbb{R}_+$ for which this cubic polynomial has two strictly positive and one strictly negative zero. This is a standard routine whose description we omit. \square

The two previous lemmas show that the genus of the McLaughlin equation can only be 1 if (t, a) lies in the region \mathcal{D} . Outside \mathcal{D} the genus must necessarily be 0.

It remains to show that inside \mathcal{D} the genus is exactly 1 (and not 0). This is taken care of by the next lemma.

LEMMA 7.4

- (i) Inside \mathcal{D} the genus is either identically 1 or identically 0.
- (ii) There exists at least one point (t, a) in \mathcal{D} for which the genus is 1.

PROOF: For part (i) we invoke the following continuity result, which follows from general potential theory: The parameters α and β in the McLaughlin equation depend continuously on $t \in \mathbb{R}$ and $a > 0$. Using this, we see that the discriminant of the McLaughlin equation also depends continuously on t and a and therefore also the branch points since these are simple roots of the discriminant.

Now let $\mathcal{D}_1 \subset \mathcal{D}$ be the region formed by those $(t, a) \in \mathcal{D}$ for which the genus is 1. We show that \mathcal{D}_1 is both open and closed in \mathcal{D} . To show that it is open, let $(t, a) \in \mathcal{D}_1$. Then the discriminant has six simple zeros and by continuity the same holds in an open neighborhood of (t, a) . To show that \mathcal{D}_1 is closed in \mathcal{D} , we take a sequence of points $(t_k, a_k) \in \mathcal{D}_1$, $k = 1, 2, \dots$, which converges to a limit point $(t, a) \in \mathcal{D}$. By Lemma 7.2 we have $\alpha = \beta = 0$ for each (t_k, a_k) so by continuity the same must hold for the limit point (t, a) . But then Lemma 7.3 shows that the discriminant has six distinct simple zeros, which implies that the genus is 1. Hence $(t, a) \in \mathcal{D}_1$.

For part (ii), we only outline a proof. The idea is to show that for any fixed $t > 2$, we have $(t, a) \in \mathcal{D}_1$ for all a small enough. This relies on the fact that for $t > 2$ the eigenvalues in the unitary matrix model with potential $\frac{1}{4}x^4 - \frac{t}{2}x^2$ (without external source) are supported on two intervals [5, 7]. The claim then follows from a continuity argument for $a \rightarrow 0$; we do not go into the details.

An alternative approach to prove part (ii) would be to pick a numerical point $(t, a) \in \mathcal{D}$ and show by direct means (using the McLaughlin equation with $\alpha = \beta = 0$) that this algebraic curve makes the RH steepest-descent analysis work, in a similar vein as in [2, 8, 10]. \square

We summarize our findings with the following:

PROPOSITION 7.5 *Let \mathcal{D} be the region defined in Lemma 7.3. For $(t, a) \in \mathcal{D}$ the McLaughlin equation has genus 1 and we have $\alpha = \beta = 0$. For $(t, a) \in (\mathbb{R} \times \mathbb{R}_+) \setminus \mathcal{D}$ the genus is 0.*

The Genus 0 Region

Now we focus on the genus 0 region $(t, a) \in (\mathbb{R} \times \mathbb{R}_+) \setminus \mathcal{D}$. A useful representation for the parameters α and β in this region is given in [3]. It is shown there that the following parametrization holds:

$$(7.8) \quad \alpha = \frac{(1-u)(c^4 + 2c^4u - u^2)^2}{c^2u^3},$$

$$(7.9) \quad \beta = -\frac{(3c^4 - u)(c^4 + 2c^4u - u^2)^2}{c^4u^2},$$

where

$$(7.10) \quad u := 2c^4 + ac$$

and where c is a root of the equation

$$(7.11) \quad 2c^6 - 2tc^4 + 3c^2 - ac^3 - tac - a^2 = 0.$$

According to [3], c is actually the *largest positive* root of equation (7.11), but we will not need this in what follows.

Let us seek the values of $(t, a) \in (\mathbb{R} \times \mathbb{R}_+) \setminus \mathcal{D}$ for which a phase transition occurs between Cases I, II, and III. For such (t, a) the discriminant of the McLaughlin equation should have a zero at $z = 0$. By virtue of (7.2) this implies $\alpha = 0$. From (7.8) this implies that either $u = 1$ or $c^4 + 2c^4u - u^2 = 0$. Let us first consider the case $u = 1$. Then we have from (7.10) that

$$(7.12) \quad 2c^4 + ac = 1.$$

Thus the equations (7.11)–(7.12) have a common root c . In other words, the resultant of these two equations with respect to the variable c should be 0. Computing this resultant with Maple yields the following condition on (t, a) :

$$(7.13) \quad 54a^4 + (72t - t^3)a^2 - (t^4 - 16t^2 + 64) = 0.$$

Next we consider the case $c^4 + 2c^4u - u^2 = 0$. From (7.8)–(7.9) this implies that $\alpha = \beta = 0$ and we know from earlier considerations (or from a similar resultant calculation as above) that this is only possible if (t, a) is such that (7.7) holds. The phase transition on this curve will be discussed in the next section. The phase transition on the curve (7.13) will be discussed in the section thereafter.

Painlevé II Transition

At the two curved boundaries of the region \mathcal{D} in Figure 7.1, we have a transition from genus 0 to genus 1. Recall that these boundaries are described by the relevant branches of the equation

$$D_6(0) = -27a^4 + (18t - 4t^3)a^2 - 4 + t^2 = 0.$$

More precisely, these branches are given by

$$A_1 = a^2 = \frac{t}{3} - \frac{2}{27}[t^3 - (t^2 - 3)^{3/2}], \quad t \geq \sqrt{3},$$

$$A_2 = a^2 = \frac{t}{3} - \frac{2}{27}[t^3 + (t^2 - 3)^{3/2}], \quad \sqrt{3} \leq t \leq 2.$$

We have $A_1 > 0$ for every $t \geq \sqrt{3}$, whereas $A_2 > 0$ only for $\sqrt{3} \leq t < 2$ and $0 < A_2 \leq A_1$ for these t -values. See Figure 7.1.

On the above curves we have a transition from genus 0 to genus 1, and we expect that the phase transition is of Painlevé II type [7, 13, 19]. More precisely, we expect that the following happens. If one lets a decrease towards the curve $a^2 = A_1$ (with $t > \sqrt{3}$), then the constraint σ on the imaginary axis becomes active, and we have a transition from Case I to Case II. If one further decreases a towards the curve $a^2 = A_2$ ($\sqrt{3} < t < 2$), then the gap in the support of μ_1 closes and hence we have a transition from Case II to Case III.

On the curve $a^2 = A_2$ the phase transition involves the eigenvalue measure μ_1 , and therefore we expect Painlevé II behavior in the local eigenvalue correlations at the origin [7, 13]. On the curve $a^2 = A_1$, however, the phase transition takes place on the “nonphysical” sheets of the Riemann surface and therefore it is not felt in the eigenvalue statistics. But then we expect Painlevé II behavior in the recurrence coefficients for the associated multiple orthogonal polynomials, as in [19].

Note that for $t > 2$ the transition at $a^2 = A_2$ does not occur. This is consistent with the fact that for $t > 2$ the eigenvalues in the unitary matrix model with potential $\frac{1}{4}x^4 - \frac{t}{2}x^2$ (without external source) are supported on two intervals [5, 7], as mentioned before.

Pearcey Transition

Now we consider the equation (7.13),

$$54a^4 + (72t - t^3)a^2 - (t^4 - 16t^2 + 64) = 0.$$

This equation has the two solutions

$$A_3 = a^2 = -\frac{2t}{3} + \frac{1}{108}[t^3 + (t^2 + 24)^{3/2}]$$

$$A_4 = a^2 = -\frac{2t}{3} + \frac{1}{108}[t^3 - (t^2 + 24)^{3/2}].$$

The branch A_4 is negative, and so is irrelevant for us. The other branch A_3 is positive, and we expect that for $t < \sqrt{3}$ a phase transition of the Pearcey type [1, 10, 34, 37] takes place for $a^2 = A_3$. See Figure 7.1.

Acknowledgment. We are grateful to the referee for pointing out the expressions in Lemma 4.2 to us.

The first author is supported in part by National Science Foundation (NSF) Grant DMS-0652005. The second author is a Postdoctoral Fellow of the Fund for Scientific Research—Flanders (Belgium). The third author is supported in part by FWO Flanders Project G.0427.09, by K.U.Leuven Research Grant OT/08/33, by the Belgian Interuniversity Attraction Pole P06/02, by the European Science Foundation Program MISGAM, and by grant MTM2008-06689-C02-01 of the Spanish Ministry of Science and Innovation.

Bibliography

- [1] Adler, M.; van Moerbeke, P. PDEs for the Gaussian ensemble with external source and the Pearcey distribution. *Comm. Pure Appl. Math.* **60** (2007), no. 9, 1261–1292.
- [2] Aptekarev, A. I.; Bleher, P. M.; Kuijlaars, A. B. J. Large n limit of Gaussian random matrices with external source. II. *Comm. Math. Phys.* **259** (2005), no. 2, 367–389.
- [3] Aptekarev, A. I.; Lysov, V. G.; Tulyakov, D. N. The global eigenvalue distribution regime of random matrices with an anharmonic potential and an external source. *Teoret. Mat. Fiz.* **159** (2009), no. 1, 34–57.
- [4] Bertola, M.; Gekhtman, M.; Szmigielski, J. The Cauchy two-matrix model. *Comm. Math. Phys.* **287** (2009), no. 3, 983–1014.
- [5] Bleher, P. M.; Its, A. Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. *Ann. of Math. (2)* **150** (1999), no. 1, 185–266.
- [6] Bleher, P. M.; Its, A. R. Asymptotics of the partition function of a random matrix model. *Ann. Inst. Fourier (Grenoble)* **55** (2003), no. 6, 1943–2000.
- [7] Bleher, P. M.; Its, A. Double scaling limit in the random matrix model: the Riemann-Hilbert approach. *Comm. Pure Appl. Math.* **56** (2003), no. 4, 433–516.
- [8] Bleher, P. M.; Kuijlaars, A. B. J. Large n limit of Gaussian random matrices with external source. I. *Comm. Math. Phys.* **252** (2004), no. 1-3, 43–76.
- [9] Bleher, P. M.; Kuijlaars, A. B. J. Random matrices with external source and multiple orthogonal polynomials. *Int. Math. Res. Not.* **2004** (2004) no. 3, 109–129.
- [10] Bleher, P. M.; Kuijlaars, A. B. J. Large n limit of Gaussian random matrices with external source. III. Double scaling limit. *Comm. Math. Phys.* **270** (2007), no. 2, 481–517.
- [11] Brézin, E.; Hikami, S. Level spacing of random matrices in an external source. *Phys. Rev. E (3)* **58** (1998), no. 6, part A, 7176–7185.
- [12] Brézin, E.; Hikami, S. Universal singularity at the closure of the gap in a random matrix theory. *Phys. Rev. E* **57** (1998), 4140–4149.
- [13] Claeys, T.; Kuijlaars, A. B. J. Universality of the double scaling limit in random matrix models. *Comm. Pure Appl. Math.* **59** (2006), no. 11, 1573–1603.
- [14] Claeys, T.; Vanlessen, M. Universality of a double scaling limit near singular edge points in random matrix models. *Comm. Math. Phys.* **273** (2007), no. 2, 499–532.
- [15] Deift, P. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*. Courant Lecture Notes in Mathematics, 3. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, R.I., 1999.

- [16] Deift, P.; Kriecherbauer, T.; McLaughlin K. T.-R. New results on the equilibrium measure for logarithmic potentials in the presence of an external field. *J. Approx. Theory* **95** (1998), no. 3, 388–475.
- [17] Deift, P.; Kriecherbauer, T.; McLaughlin K. T.-R.; Venakides, S.; Zhou, X. Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52** (1999), no. 12, 1491–1552.
- [18] Deift, P.; Kriecherbauer, T.; McLaughlin K. T.-R.; Venakides, S.; Zhou, X. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* **52** (1999), no. 11, 1335–1425.
- [19] Delvaux, S.; Kuijlaars, A. B. J. A phase transition for nonintersecting Brownian motions, and the Painlevé equation. *Int. Math. Res. Not. IMRN* **2009** (2009), no. 19, 3639–3725.
- [20] Dubrovin, B. A. Theta-functions and nonlinear equations. *Uspekhi Mat. Nauk* **36** (1981), no. 2(218), 11–80; English translation: *Russian Math. Surveys* **36** (1981), 11–92.
- [21] Duits, M.; Kuijlaars, A. B. J. Universality in the two-matrix model: a Riemann-Hilbert steepest-descent analysis. *Comm. Pure Appl. Math.* **62** (2009), no. 8, 1076–1153.
- [22] Duits, M.; Kuijlaars, A. B. J.; Mo, M. Y. The Hermitian two matrix model with an even quartic potential. In preparation.
- [23] Dyson, F. J. A Brownian-motion model for the eigenvalues of a random matrix. *J. Mathematical Phys.* **3** (1962), 1191–1198.
- [24] Farkas, H. M.; Kra, I. *Riemann surfaces*. Graduate Texts in Mathematics, 71. Springer, New York–Berlin, 1980.
- [25] Fay, J. D. *Theta functions on Riemann surfaces*. Lecture Notes in Mathematics, 352. Springer, Berlin–New York, 1973.
- [26] Fokas, A. S.; Its, A. R.; Kitaev, A. V. The isomonodromy approach to matrix models in 2D quantum gravity. *Comm. Math. Phys.* **147** (1992), no. 2, 395–430.
- [27] Kuijlaars, A. B. J. Multiple orthogonal polynomial ensembles. *Recent trends in orthogonal polynomials and approximation theory*, 155–176. Contemporary Mathematics, 507. American Mathematical Society, Providence, R.I., 2010.
- [28] Kuijlaars, A. B. J.; Dragnev, P. D. Equilibrium problems associated with fast decreasing polynomials. *Proc. Amer. Math. Soc.* **127** (1999), no. 4, 1065–1074.
- [29] Kuijlaars, A. B. J.; Martínez-Finkelshtein, A.; Wielonsky, F. Non-intersecting squared Bessel paths and multiple orthogonal polynomials for modified Bessel weights. *Comm. Math. Phys.* **286** (2009), no. 1, 217–275.
- [30] Kuijlaars, A. B. J.; Mo, M. Y. The global parametrix in the Riemann-Hilbert steepest descent analysis for orthogonal polynomials. arXiv:0909.5626v1.
- [31] McLaughlin, K. T.-R. Asymptotic analysis of random matrices with external source and a family of algebraic curves. *Nonlinearity* **20** (2007), 1547–1571.
- [32] Mo, M. Y. Universality in the two matrix model with a monomial quartic and a general even polynomial potential. *Comm. Math. Phys.* **291** (2009), no. 3, 863–894.
- [33] Nikishin, E. M.; Sorokin, V. N. *Rational approximations and orthogonality*. Translations of Mathematical Monographs, 92. American Mathematical Society, Providence, R.I., 1991.
- [34] Okounkov, A.; Reshetikhin, N. Random skew plane partitions and the Pearcey process. *Comm. Math. Phys.* **269** (2007), no. 3, 571–609.
- [35] Pastur, L. The spectrum of random matrices. (Russian). *Teoret. Mat. Fiz.* **10** (1972), no. 1, 102–112; English translation: *Theoret. and Math. Phys.* **10** (1972), no. 1, 67–74.
- [36] Saff, E. B.; Totik, V. *Logarithmic potentials with external field*. Grundlehren der mathematischen Wissenschaften, 316. Springer, Berlin, 1997.
- [37] Tracy, C. A.; Widom, H. The Pearcey process. *Comm. Math. Phys.* **263** (2006), no. 2, 381–400.
- [38] Van Assche, W.; Geronimo, J. S.; Kuijlaars, A. B. J. Riemann-Hilbert problems for multiple orthogonal polynomials. *Special functions 2000: current perspective and future directions*

(*Tempe, AZ, 2000*), 23–59. NATO Science Series II: Mathematics, Physics and Chemistry, 30. Kluwer, Dordrecht, The Netherlands, 2001.

- [39] Zinn-Justin, P. Random Hermitian matrices in an external field. *Nuclear Phys. B* **497** (1997), no. 3, 725–732.
- [40] Zinn-Justin, P. Universality of correlation functions of Hermitian random matrices in an external field. *Comm. Math. Phys.* **194** (1998), no. 3, 631–650.

P. BLEHER
Indiana University–Purdue University
Indianapolis
Department of Mathematical Sciences
402 N. Blackford Street
Indianapolis, IN 46202
E-mail: bleher@math.iupui.edu

S. DELVAUX
Katholieke Universiteit Leuven
Department of Mathematics
Celestijnenlaan 200B
B-3001 Leuven
BELGIUM
E-mail: steven.delvaux@
wis.kuleuven.be

A. B. J. KUIJLAARS
Katholieke Universiteit Leuven
Department of Mathematics
Celestijnenlaan 200B
B-3001 Leuven
BELGIUM
E-mail: arno.kuijlaars@
wis.kuleuven.be

Received January 2010.