

# INVESTIGATION OF THE TWO-CUT PHASE REGION IN THE COMPLEX CUBIC ENSEMBLE OF RANDOM MATRICES

AHMAD BARHOUMI, PAVEL BLEHER, ALFREDO DEAÑO, AND MAXIM YATTSELEV

*Dedicated to the memory of Freeman Dyson*

**ABSTRACT.** We investigate the phase diagram of the complex cubic unitary ensemble of random matrices with the potential  $V(M) = -\frac{1}{3}M^3 + tM$  where  $t$  is a complex parameter. As proven in our previous paper [7], the whole phase space of the model,  $t \in \mathbb{C}$ , is partitioned into two phase regions,  $O_{\text{one-cut}}$  and  $O_{\text{two-cut}}$ , such that in  $O_{\text{one-cut}}$  the equilibrium measure is supported by one Jordan arc (cut) and in  $O_{\text{two-cut}}$  by two cuts. The regions  $O_{\text{one-cut}}$  and  $O_{\text{two-cut}}$  are separated by critical curves, which can be calculated in terms of critical trajectories of an auxiliary quadratic differential. In [7] the one-cut phase region was investigated in detail. In the present paper we investigate the two-cut region. We prove that in the two-cut region the endpoints of the cuts are analytic functions of the real and imaginary parts of the parameter  $t$ , but not of the parameter  $t$  itself (so that the Cauchy–Riemann equations are violated for the endpoints). We also obtain the semiclassical asymptotics of the orthogonal polynomials associated with the ensemble of random matrices and their recurrence coefficients. The proofs are based on the Riemann–Hilbert approach to semiclassical asymptotics of the orthogonal polynomials and the theory of  $S$ -curves and quadratic differentials.

## 1. INTRODUCTION

This work is a continuation of the works of Bleher and Deaño [5, 6] and Bleher, Deaño, and Yattselev [7]. The main goal of the whole project is to analyze the large  $N$  asymptotics of the partition function and correlation functions in the unitary ensemble of random matrices with the cubic potential  $V(M) = -\frac{1}{3}M^3 + tM$ , where  $t$  is a complex parameter. The partition function of this ensemble is formally defined by the matrix integral over the space of  $N \times N$  Hermitian matrices,

$$Z_N^{\text{RM}}(t) = \int_{\mathcal{H}_N} e^{-N\text{Tr}(-\frac{1}{3}M^3 + tM)} dM.$$

The integral is formal because it is divergent, and it needs a regularization. To define such a regularization consider the (formal) identity

$$Z_N^{\text{RM}}(t) = \frac{\pi^{N(N-1)/2}}{\prod_{k=1}^N k!} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \prod_{1 \leq j < k \leq N} (z_j - z_k)^2 \prod_{j=1}^N e^{-N(-\frac{1}{3}z_j^3 + tz_j)} dz_1 \dots dz_N$$

see, [8, formula (1.2.18)]. The random matrix partition function  $Z_N^{\text{RM}}(t)$  can now be regularized by changing the contour of integration from the real line  $\mathbb{R}$  to the contour  $\Gamma$  depicted in Figure 1,

---

*Date:* January 30, 2022.

*2010 Mathematics Subject Classification.* 33C47, 30E15, 31A25, 15B52.

*Key words and phrases.* cubic random matrix model, partition function, equilibrium measure,  $S$ -curve, quadratic differential, orthogonal polynomials, non-Hermitian orthogonality, Riemann–Hilbert problem, nonlinear steepest descent method.

The first author (A.B.) was partially supported by the National Science Foundation (NSF) under DMS-1812625. The work of the second author (P.B.) is supported in part by the NSF Grants DMS-1265172 and DMS-1565602. The second author (P.B.) acknowledges hospitality and support from MSRI (Mathematical Sciences Research Institute) during the research program “Universality and Integrability in Random Matrix Theory and Interacting Particle Systems” in September–December 2021. The third author (A.D.) acknowledges financial support from the EPSRC grant “Painlevé equations: analytical properties and numerical computation”, reference EP/P026532/1, as well as support by the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement with UC3M in the line of Excellence of University Professors (EPUC3M23), and in the context of the V PRICIT (Regional Programme of Research and Technological Innovation). The research of the fourth author (M.Y.) is supported by a grant from the Simons Foundation, CGM-706591.

which goes from  $(-\infty)$  to  $(e^{\pi i/3}\infty)$ . Hence, we are interested in studying the regularized eigenvalue partition function defined as

$$(1.1) \quad Z_N(t) := \int_{\Gamma} \dots \int_{\Gamma} \prod_{1 \leq j < k \leq N} (z_j - z_k)^2 \prod_{j=1}^N e^{-N(-\frac{1}{3}z_j^3 + tz_j)} dz_1 \dots dz_N,$$

which is well defined (the integrals are convergent) for all values  $t \in \mathbb{C}$ .

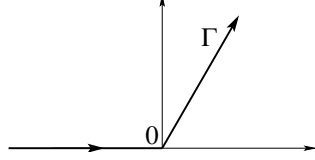


FIGURE 1. The contour  $\Gamma$  of integration.

As shown in [5], the partition function  $Z_N(t)$  can be used to study the topological expansion in the cubic ensemble of random matrices and enumeration of regular graphs of degree 3 on Riemannian surfaces. To be more precise, let us consider the cubic potential in the form  $M^2/2 - uM^3$ , where  $u > 0$ , and the corresponding eigenvalue matrix integral is given by

$$(1.2) \quad \Xi_N(u) = \int_{\Gamma} \dots \int_{\Gamma} \prod_{1 \leq j < k \leq N} (\zeta_j - \zeta_k)^2 \prod_{j=1}^N e^{-N(\frac{1}{2}\zeta_j^2 - u\zeta_j^3)} d\zeta_1 \dots d\zeta_N.$$

We would like to consider all the possible values, including the complex ones, of the parameter  $u$ . Formula (1.2) is not very convenient for this purpose because the contour of integration  $\Gamma$  should be rotated to secure the convergence of the integral depending on the argument of  $u$ . Instead, let us make the change of variables

$$\zeta_j = (3u)^{-1/3} z_j + \frac{1}{6u},$$

which yields that

$$\frac{\zeta_j^2}{2} - u\zeta_j^3 - \frac{1}{108u^2} = -\frac{z_j^3}{3} + tz_j,$$

where

$$(1.3) \quad t = \frac{1}{4(3u)^{4/3}}.$$

This implies the relation between the partition functions,

$$\Xi_N(u) = (4t)^{N^2/4} e^{-2t^{3/2}N/3} Z_N(t).$$

As proven in [5], the free energy of the cubic model  $\mathcal{F}_N(u)$  admits an asymptotic expansion as  $N \rightarrow \infty$  in powers of  $N^{-2}$ , that is,

$$(1.4) \quad \mathcal{F}_N(u) := \frac{1}{N^2} \ln \frac{\Xi_N(u)}{\Xi_N(0)} \sim \sum_{g=0}^{\infty} \frac{F^{(2g)}(u)}{N^{2g}}$$

for any  $u$  in the interval  $0 \leq u < u_c$ , where  $u_c := 3^{1/4}18^{-1}$  is a critical point. In addition, the functions  $F^{(2g)}(u)$  admit an analytic continuation to the disk  $|u| < u_c$  on the complex plane, and if we expand them in powers of  $u$ ,

$$(1.5) \quad F^{(2g)}(u) = \sum_{j=1}^{\infty} \frac{f_{2j}^{(2g)} u^{2j}}{(2j)!},$$

then the coefficient  $f_{2j}^{(2g)}$  is a positive integer that counts the number of 3-valent connected graphs with  $2j$  vertices on a Riemann surface of genus  $g$ . Asymptotic expansion (1.4) is called a *topological*

*expansion*. For more details on this aspect of the theory, we refer the reader to the classical papers of Bessis, Itzykson and Zuber [4], Brézin, Itzykson, Parisi and Zuber [9], the monograph of Forrester [22, Section 1.6], the works of Mulase [31], Di Francesco [16], Ercolani and McLaughlin [18, 19], Eynard [20], and references therein, or a very readable introduction by Zvonkin [40]. It is noteworthy that the general idea of a topological expansion goes back to the classical work of 't Hooft [39].

As shown in [5], the coefficients  $f_{2j}^{(2g)}/(2j)!$  of power series (1.5) behave, when  $j \rightarrow \infty$ , as

$$\frac{f_{2j}^{(2g)}}{(2j)!} = \frac{K_{2g} j^{\frac{5g-7}{2}}}{u_c^{2j}} \left(1 + O\left(j^{-1/2}\right)\right), \quad K_{2g} > 0.$$

This implies that  $u_c$  is the radius of convergence of each power series (1.5). In fact,  $u = u_c$  is a singular point of each of the functions (1.5). The topological expansion in a neighborhood of the critical point  $u_c$  has been obtained in the work of Bleher and Deaño [6] and it is closely related to the Painlevé I equation as follows: consider a formal series

$$(1.6) \quad \sum_{g=0}^{\infty} \Gamma\left(\frac{5g-1}{2}\right) \frac{u_c^g K_{2g}}{6 \cdot 3^{1/4}} \lambda^{\frac{1-5g}{2}},$$

it can be shown that there exists a one-parameter family of solutions  $\{y_\alpha(\lambda)\}$ ,  $\alpha \in \mathbb{R}$ , of the Painlevé I differential equation

$$y''(\lambda) = a_0 y^2(\lambda) - a_1 \lambda,$$

with  $a_0 = 2^{\frac{3}{2}} 3^{\frac{9}{4}}$  and  $a_1 = 2^{\frac{3}{2}} 3^{-\frac{5}{4}}$ , that admits an asymptotic expansion (1.6) as  $\lambda \rightarrow -\infty$ . This family consists of the Boutroux *tronquée* solutions, which means that they are asymptotically free of poles in one of the five canonical sectors of angle  $2\pi/5$  in the complex plane that appear naturally when considering rotational symmetries of the Painlevé I equation. The parameter  $\alpha$  appears in the Painlevé I Riemann–Hilbert problem, as described by Kapaev in [28], and in the context of the cubic model, it is related to the choice of the original contour of integration  $\Gamma$ , see [5, 6]. In the present case, we have  $\alpha = 1$ , which makes the solution of Painlevé I *tritonquée*, that is, asymptotically free of poles in four of the five canonical sectors of angle  $2\pi/5$  in the complex plane. We refer the reader to [6, Section 2.1] and also the work of Joshi and Kitaev [27] for more details.

It is noteworthy that the key ingredient in the proof of topological expansion (1.4) in [5] is the derivation of semiclassical asymptotic formulae for the recurrence coefficients  $\gamma_n^2, \beta_n$ , see (1.3) and (1.13) further below, of the corresponding monic orthogonal polynomials  $P_n(z) = z^n + \dots$  defined via orthogonality relations

$$\int_{\Gamma} P_n(z) z^k e^{-N(\frac{1}{2}z^2 - uz^3)} dz = 0, \quad k \in \{0, 1, \dots, n-1\}.$$

The idea of using the orthogonal polynomials in ensembles of random matrices goes back to the classical works of Dyson and Mehta, see the works [17] and [30] and references therein. Namely, as proven in [5], for any  $u \in [0, u_c)$ , there exists  $\varepsilon > 0$  such that as  $N, n \rightarrow \infty$  with  $1 - \varepsilon \leq \frac{n}{N} \leq 1 + \varepsilon$ , the recurrence coefficients  $\gamma_n^2$  and  $\beta_n$  admit asymptotic expansions in powers of  $\frac{1}{N^2}$ :

$$(1.7) \quad \begin{cases} \gamma_n^2 \sim \sum_{k=0}^{\infty} \frac{1}{N^{2k}} g_{2k}\left(\frac{n}{N}, u\right), \\ \beta_n \sim \sum_{k=0}^{\infty} \frac{1}{N^{2k}} b_{2k}\left(\frac{n}{N} + \frac{1}{2N}, u\right), \end{cases}$$

where the functions  $g_{2k}(s, u)$ ,  $b_{2k}(s, u)$  do not depend on  $n$  and  $N$  and are analytic in  $s$  at  $s = 1$ .

In [6] this asymptotic expansion is extended to the double scaling asymptotic expansion of the recurrence coefficients at the critical point  $u_c$ . In the double scaling regime we set

$$\frac{n}{N} = 1 + vN^{-4/5}.$$

where  $v \in \mathbb{R}$  is a scaling variable. Then as proven in [6], at  $u = u_c$  the recurrence coefficients  $\gamma_n^2$  and  $\beta_n$  admit asymptotic expansions in powers of  $N^{-2/5}$  as  $N \rightarrow \infty$ :

$$(1.8) \quad \begin{cases} \gamma_n^2 \sim \gamma_c^2 + \sum_{k=1}^{\infty} \frac{1}{N^{2k/5}} p_{2k}(v), \\ \beta_n \sim \beta_c + \sum_{k=1}^{\infty} \frac{1}{N^{2k/5}} q_{2k}(\tilde{v}), \end{cases}$$

where the functions  $p_{2k}(v)$ ,  $q_{2k}(\tilde{v})$  are expressed in terms of the Boutroux tritronquée solution to Painlevé I equation,  $\gamma_c^2, \beta_c$  are known explicitly, and  $\tilde{v} = v + N^{-1/5}/2$ . As shown in [6], expansions (1.7) and (1.8) can be extended for large  $N$  to  $u$  in overlapping intervals:  $[0, u_c - N^{-0.79}]$  for (1.7) and  $[u_c - N^{-0.65}, u_c]$  for (1.8), and this can be used to obtain the double scaling asymptotic formula for the partition function at  $u_c$ . Namely, let

$$u - u_c = 2^{-\frac{12}{5}} 3^{-\frac{7}{4}} \lambda N^{-\frac{4}{5}},$$

where  $\lambda$  is a complex scaling variable in the double scaling regime. Then for  $\lambda$  outside of a neighborhood of the poles of  $y(\lambda)$ , the Boutroux tritronquée solution to Painlevé I equation, the partition function  $\Xi_N(u)$  can be written as

$$(1.9) \quad \Xi_N(u) = \Xi_N^{\text{reg}}(u) \Xi_N^{\text{sing}}(\lambda) (1 + \mathcal{O}(N^{-\varepsilon})), \quad \varepsilon > 0,$$

where the regular and singular factors are given by

$$\Xi_N^{\text{reg}}(u) = e^{N^2[a+b(u-u_c)+c(u-u_c)^2]+d} \quad \text{and} \quad \Xi_N^{\text{sing}}(\lambda) = e^{-Y(\lambda)},$$

where  $a, b, c, d$  are some explicit constants and  $Y(\lambda)$  is a solution of the differential equation  $Y''(\lambda) = y(\lambda)$  with the boundary condition

$$Y(\lambda) = \frac{2\sqrt{6}}{45}(-\lambda)^{5/2} - \frac{1}{48} \log(-\lambda) + \mathcal{O}((-\lambda)^{-5/2}), \quad \lambda \rightarrow -\infty.$$

Asymptotic formula (1.9) is used in [6] to prove the conjecture of David [10, 11] that the poles of the tritronquée solution  $y(\lambda)$  give rise to zeros of  $\Xi_N(u)$ .

As we have stressed before, the topological expansions from [5, 6, 7] were obtained by first analyzing the associated orthogonal polynomials defined by

$$(1.10) \quad \int_{\Gamma} z^k P_n(z; t, N) e^{-NV(z;t)} dz = 0, \quad k \in \{0, \dots, n-1\},$$

where

$$(1.11) \quad V(z; t) := -\frac{z^3}{3} + tz, \quad t \in \mathbb{C}.$$

Due to the non-Hermitian character of the above relations, it might happen that polynomial satisfying (1.10) is non-unique. In this case we understand by  $P_n(z; t, N)$  the monic non-identically zero polynomial of the smallest degree (such a polynomial is always unique). One way of connecting the partition function to orthogonal polynomials is via *three term recurrence relations*. More precisely, let

$$(1.12) \quad h_n(t, N) := \int_{\Gamma} P_n^2(z; t, N) e^{-NV(z;t)} dz = \frac{D_n(t, N)}{D_{n-1}(t, N)},$$

where  $D_{-1}(t, N) \equiv 1$  and  $D_n(t, N) := \det \left[ \int_{\Gamma} z^{i+j} e^{-NV(z;t)} dz \right]_{i,j=0}^n$  is the Hankel determinant of the moments of the measure of integration in (1.10). It is easy to see that each  $D_n(t, N)$  is an entire function of  $t$  and therefore each  $h_n(t, N)$  is meromorphic in  $\mathbb{C}$ . Hence, given  $n$ , the set of the values  $t$  for which there exists  $k \in \{0, \dots, n\}$  such that  $h_k(t, N) = 0$  is countable with no limit points in the finite plane. Outside of this set, the standard argument using orthogonality (1.10) shows that

$$(1.13) \quad zP_n(z; t, N) = P_{n+1}(z; t, N) + \beta_n(t, N)P_n(z; t, N) + \gamma_n^2(t, N)P_{n-1}(z; t, N),$$

and by analytic continuation (1.13) extends to those values of  $t$  for which  $D_{n-1}(t, N)D_n(t, N) \neq 0$  (that is,  $n + 1$ -st and  $n$ -th polynomials appearing (1.13) have the prescribed degrees), where

$$(1.14) \quad \gamma_n^2(t, N) = h_n(t, N)/h_{n-1}(t, N).$$

Notice that (1.13) remains meaningful even if  $D_{n-2}(t, N) = 0$ . In this case  $h_{n-2}(t, N) = 0$ ,  $h_{n-1}(t, N) = \infty$ , and  $h_n(t, N)$  is finite, which means that  $P_{n-2}(z; t, N)$  is orthogonal to itself and therefore is equal to  $P_{n-1}(z; t, N)$  but not to  $P_n(z; t, N)$ . Thus,  $\gamma_n^2(t, N) = 0$  and the last term in (1.13) is zero. Notice also that if  $h_{n-1}(t, N) \neq 0$  while  $h_n(t, N) = 0$  (that is,  $D_{n-1}(t, N) \neq 0$  and  $D_n(t, N) = 0$ ), then we simply have that  $\gamma_N^2(t, N) = 0$  by analytic continuation. However, this means that the polynomial  $P_n(z; t, N)$  is orthogonal to itself and therefore  $P_{n+1}(z; t, N) = P_n(z; t, N)$ . Then the analytic continuation argument necessitates that  $\beta_n(t, N) = \infty$ , which also can be seen from the determinantal representation of  $\beta_n(t, N)$  as it is the difference of the subleading coefficients of  $P_n(z; t, N)$  and  $P_{n+1}(z; t, N)$ . It is further known that the recurrence coefficients  $\gamma_N^2(t, N)$  satisfy the *Toda equation*:

$$\frac{\partial^2 F_N(t)}{\partial t^2} = \gamma_N^2(t, N), \quad F_N(t) = \frac{1}{N^2} \log Z_N(t).$$

Another way of connecting partition function to orthogonal polynomials is through the formula

$$Z_N(t) = N! \prod_{n=0}^{N-1} h_n(t, N) (= N! D_{N-1}(t, N)).$$

As it turns out, there are two different regions in the complex  $t$ -plane for which the behavior of the polynomials  $P_n(z; t, N)$ , and therefore of the partition function, is different. Colloquially, we call them one-cut and two-cut regions, a distinction that will become clear later. In [5, 6, 7] the partition function has been analyzed in and on the boundary of the one-cut region. The goal of this work is to start the analysis in the two-cut region. More precisely, here we only consider the asymptotics of the orthogonal polynomials and their recurrence coefficients and postpone the analysis of the partition function for the future project. The structure of the paper is as follows:

- In Sections 2 and 3 we describe equilibrium measures and corresponding  $S$ -curves for the cubic model under consideration (those describe asymptotic behavior of the normalized counting measures of zeros of the orthogonal polynomials in the weak\* sense). This leads us to a precise description of the phase diagram (i.e., of the one- and two-cut regions) of the cubic model on the complex  $t$ -plane.
- In Section 4 we present the main results of the paper: asymptotic formulae for the orthogonal polynomials and their recurrence coefficients in the two-cut phase region.
- Section 5–9 are devoted to the proof of our main results, in them we obtain various results about the detailed structure of the  $S$ -curves, derive main properties of the dominant terms of the expansions, apply the Riemann–Hilbert approach to asymptotics of the orthogonal polynomials, and finally prove the expansions themselves.

## 2. EQUILIBRIUM MEASURES AND S-PROPERTY

It is well understood that the zeros of polynomials satisfying (1.10) asymptotically distribute as a certain *weighted equilibrium measure* on an  $S$ -contour corresponding to the weight function (1.11). In this section we discuss these notions in greater detail. Our consideration will use the works of Huybrechs, Kuijlaars, and Lejon [25] and Kuijlaars and Silva [29]. Let us start with some definitions.

**Definition 2.1.** *Let  $V$  be an entire function. The logarithmic energy in the external field  $\operatorname{Re}V$  of a measure  $\nu$  in the complex plane is defined as*

$$E_V(\nu) = \iint \log \frac{1}{|s-t|} d\nu(s) d\nu(t) + \int \operatorname{Re}V(s) d\nu(s).$$

*The equilibrium energy of a contour  $\Gamma$  in the external field  $\operatorname{Re}V$  is equal to*

$$(2.1) \quad \mathcal{E}_V(\Gamma) = \inf_{\nu \in \mathcal{M}(\Gamma)} E_V(\nu),$$

where  $\mathcal{M}(\Gamma)$  denotes the space of Borel probability measures on  $\Gamma$ .

When  $\operatorname{Re}V(s) - \log|s| \rightarrow +\infty$  as  $\Gamma \ni s \rightarrow \infty$ , there exists a unique minimizing measure for (2.1), which is called the *weighted equilibrium measure* of  $\Gamma$  in the external field  $\operatorname{Re}V$ , say  $\mu_\Gamma$ , see [33, Theorem I.1.3] or [25]. The support of  $\mu_\Gamma$ , say  $J_\Gamma$ , is a compact subset of  $\Gamma$ . The equilibrium measure  $\mu = \mu_\Gamma$  is characterized by the Euler–Lagrange variational conditions:

$$(2.2) \quad 2U^\mu(z) + \operatorname{Re}V(z) \begin{cases} = \ell, & z \in J_\Gamma, \\ \geq \ell, & z \in \Gamma \setminus J_\Gamma, \end{cases}$$

where  $\ell = \ell_\Gamma$  is a constant, the Lagrange multiplier, and  $U^\mu(z) := -\int \log|z-s|d\mu(s)$  is the logarithmic potential of  $\mu$ , see [33, Theorem I.3.3].

Observe that due to the analyticity of the integrand in (1.1),  $\Gamma$  can be varied without changing the value of  $Z_N(t)$ . Henceforth, we suppose that partition function (1.1) and orthogonal polynomials (1.10) are defined with  $\Gamma \in \mathcal{T}$ , where  $\mathcal{T}$  is the following class of contours.

**Definition 2.2.** We shall denote by  $\mathcal{T}$  the collection of all piecewise smooth contours that extend to infinity in both directions, each admitting a parametrization  $z(s)$ ,  $s \in \mathbb{R}$ , for which there exists  $\epsilon \in (0, \pi/6)$  and  $s_0 > 0$  such that

$$(2.3) \quad \begin{cases} |\arg(z(s)) - \pi/3| \leq \epsilon, & s \geq s_0, \\ |\arg(z(s)) - \pi| \leq \epsilon, & s \leq -s_0, \end{cases}$$

where  $\arg(z(s)) \in [0, 2\pi)$ .

Despite the above flexibility, it is well understood in the theory of non-Hermitian orthogonal polynomials, starting with the works of Stahl [34, 35, 36] and Gonchar and Rakhmanov [24], that one should use the contour whose equilibrium measure has support *symmetric* (with the  $S$ -property) in the external field  $\operatorname{Re}V$ . We make this idea precise in the following definition.

**Definition 2.3.** The support  $J_\Gamma$  has the  $S$ -property in the external field  $\operatorname{Re}V$ , if it consists of a finite number of open analytic arcs and their endpoints, and on each arc it holds that

$$(2.4) \quad \frac{\partial}{\partial n_+} (2U^{\mu_\Gamma} + \operatorname{Re}V) = \frac{\partial}{\partial n_-} (2U^{\mu_\Gamma} + \operatorname{Re}V),$$

where  $\frac{\partial}{\partial n_+}$  and  $\frac{\partial}{\partial n_-}$  are the normal derivatives from the (+)- and (-)-side of  $\Gamma$ . We shall say that a curve  $\Gamma \in \mathcal{T}$  is an  $S$ -curve in the field  $\operatorname{Re}V$ , if  $J_\Gamma$  has the  $S$ -property in this field.

It is also understood that geometrically  $J_\Gamma$  is comprised of *critical trajectories* of a certain quadratic differential. Recall that if  $Q$  is a meromorphic function, a *trajectory* (resp. *orthogonal trajectory*) of a quadratic differential  $-Q(z)dz^2$  is a maximal regular arc on which

$$-Q(z(s))(z'(s))^2 > 0 \quad (\text{resp.} \quad -Q(z(s))(z'(s))^2 < 0)$$

for any local uniformizing parameter. A trajectory is called *critical* if it is incident with a *finite critical point* (a zero or a simple pole of  $-Q(z)dz^2$ ) and it is called *short* if it is incident only with finite critical points. We designate the expression *critical (orthogonal) graph* of  $-Q(z)dz^2$  for the totality of the critical (orthogonal) trajectories  $-Q(z)dz^2$ .

The following theorem is a specialization to  $V(z; t)$  of [29, Theorems 2.3 and 2.4].

**Theorem 2.1.** Let  $V(z; t)$  be given by (1.11).

(1) There exists a contour  $\Gamma_t \in \mathcal{T}$  such that

$$(2.5) \quad \mathcal{E}_V(\Gamma_t) = \sup_{\Gamma \in \mathcal{T}} \mathcal{E}_V(\Gamma).$$

(2) The support  $J_t$  of the equilibrium measure  $\mu_t := \mu_{\Gamma_t}$  has the  $S$ -property in the external field  $\operatorname{Re}V(z; t)$  and the measure  $\mu_t$  is uniquely determined by (2.4). Therefore,  $\mu_t$  and its support  $J_t$  are the same for every  $\Gamma_t$  satisfying (2.5).

(3) *The function*

$$(2.6) \quad Q(z; t) = \left( \frac{V'(z; t)}{2} - \int \frac{d\mu_t(s)}{z-s} \right)^2, \quad z \in \mathbb{C} \setminus J_t,$$

is a polynomial of degree 4.

(4) *The support  $J_t$  consists of some short critical trajectories of the quadratic differential  $-Q(z; t)dz^2$  and the equation*

$$(2.7) \quad d\mu_t(z) = -\frac{1}{\pi i} Q_+^{1/2}(z; t) dz, \quad z \in J_t,$$

holds on each such critical trajectory, where  $Q^{1/2}(z; t) = \frac{1}{2}z^2 + \mathcal{O}(z)$  as  $z \rightarrow \infty$  (in what follows,  $Q^{1/2}(z; t)$  will always stand for such a branch).

Much information on the structure of the critical graphs of a quadratic differential can be found in the excellent monographs [26, 32, 37]. Since  $\deg Q = 4$ ,  $J_t$  consists of one or two arcs, corresponding (respectively) to the cases where  $Q(z; t)$  has two simple zeros and one double zero, and the case where it has four simple zeros. Away from  $J_t$ , one has freedom in choosing  $\Gamma_t$ . In particular, let

$$(2.8) \quad \mathcal{U}(z; t) := \operatorname{Re} \left( 2 \int_e^z Q^{1/2}(s; t) ds \right) = \ell_{\Gamma_t} - \operatorname{Re}(V(z; t)) - 2U^{\mu_t}(z),$$

where  $e \in J_t$  is any and the second equality follows from (2.6) (since the constant  $\ell_{\Gamma_t}$  in (2.2) is the same for both connected components of  $J_t$  and the integrand is purely imaginary on  $J_t$ , the choice of  $e$  is indeed not important). Clearly,  $\mathcal{U}(z; t)$  is a subharmonic function (harmonic away from  $J_t$ ) which is equal to zero  $J_t$  by (2.2). The trajectories of  $-Q(z; t)dz^2$  emanating out of the endpoints of  $J_t$  belong to the set  $\{z : \mathcal{U}(z; t) = 0\}$  and it follows from the variational condition (2.2) that  $\Gamma_t \setminus J_t \subset \{z : \mathcal{U}(z; t) < 0\}$ . However, within the region  $\{z : \mathcal{U}(z; t) < 0\}$  the set  $\Gamma_t \setminus J_t$  can be varied freely. The geometry of the set  $\{z : \mathcal{U}(z; t) < 0\}$  is described further below in Theorems 3.1 and 3.2.

### 3. STRUCTURE OF $\Gamma_t$

The partition of the phase space into one- and two-cut regions as well as the structure of  $\Gamma_t$  and its dependence on  $t$  has been heuristically described in [1, 2]. Even before these works, the phase diagram in Figure 3 was heuristically described in [10]. Mathematically rigorous description was provided in [7], but only in the one-cut region. Let us quickly recall the important notions from [7].

Denote by  $C$  the critical graph of an auxiliary quadratic differential

$$(3.1) \quad -(1 + 1/s)^3 ds^2,$$

see Figure 2(a). It was shown in [7, Section 5] that  $C$  consists of 5 critical trajectories emanating from  $-1$  at the angles  $2\pi k/5$ ,  $k \in \{0, 1, 2, 3, 4\}$ , one of them being  $(-1, 0)$ , other two forming a loop crossing the real line approximately at 0.635, and the last two approaching infinity along the imaginary axis without changing the half-plane (upper or lower). Given  $C$ , define

$$\Delta := \{x : 2x^3 \in C\}.$$

Further, put  $\Omega_{\text{one-cut}}$  to be the shaded region on Figure 2(b) and set

$$\partial\Omega_{\text{one-cut}} = \Delta_{\text{birth}}^b \cup \{-2^{-1/3}\} \cup \Delta_{\text{split}} \cup \{e^{\pi i/3} 2^{-1/3}\} \cup \Delta_{\text{birth}}^a,$$

where  $\Delta_{\text{split}}$  connects  $-2^{-1/3}$  and  $e^{\pi i/3} 2^{-1/3}$ ,  $\Delta_{\text{birth}}^b$  extends to infinity in the direction of the angle  $7\pi/6$  while  $\Delta_{\text{birth}}^a$  extends to infinity in the direction of the angle  $\pi/6$ . Let

$$t(x) := (x^3 - 1)/x$$

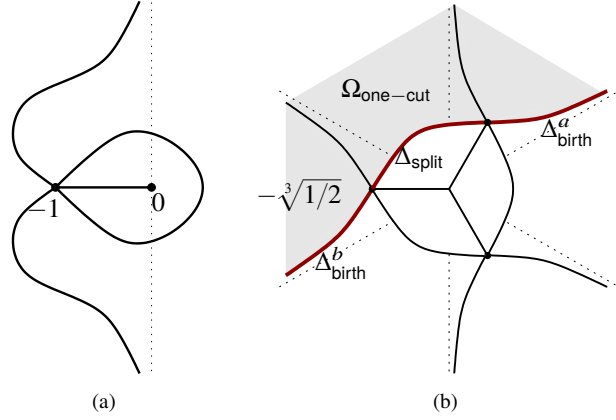


FIGURE 2. Schematic representation of (a) the critical graph  $C$ ; (b) the set  $\Delta$  (solid lines) and the domain  $\Omega_{\text{one-cut}}$  (shaded region).

and set

$$(3.2) \quad \begin{cases} t_{\text{cr}} := 3 \cdot 2^{-2/3} = t(-2^{-1/3}), \\ O_{\text{one-cut}} := t(\Omega_{\text{one-cut}}), \\ C_{\text{split}} := t(\Delta_{\text{split}}), \quad C_{\text{birth}}^b := t(\Delta_{\text{birth}}^b), \quad C_{\text{birth}}^a := t(\Delta_{\text{birth}}^a), \\ S := (t_{\text{cr}}, \infty), \quad e^{2\pi i/3} S := \{z : e^{-2\pi i/3} z \in S\}, \end{cases}$$

see Figure 3. The function  $t(x)$  is holomorphic in  $\Omega_{\text{one-cut}}$  with non-vanishing derivative there.

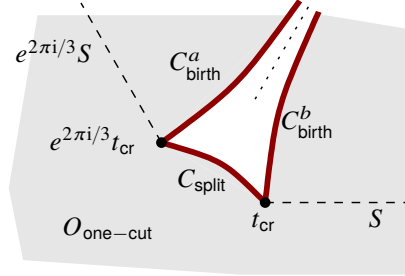


FIGURE 3. Domain  $O_{\text{one-cut}}$  (shaded region);  $\partial O_{\text{one-cut}}$  consisting of the open bounded arc  $C_{\text{split}}$ , two open semi-unbounded arcs  $C_{\text{birth}}^a$  and  $C_{\text{birth}}^b$ , and two points  $t_{\text{cr}}$  and  $e^{2\pi i/3} t_{\text{cr}}$ ; the semi-unbounded open horizontal rays  $S$  and  $e^{2\pi i/3} S$  (dashed lines).

It maps  $\Omega_{\text{one-cut}}$  onto  $O_{\text{one-cut}}$  in a one-to-one fashion. Hence, the inverse map  $x(t)$  exists and is holomorphic.

**Convention 1.** Below, we adopt the following convention:  $\Gamma(z_1, z_2)$  (resp.  $\Gamma[z_1, z_2]$ ) stands for the trajectory or orthogonal trajectory (resp. the closure of) of the differential  $-Q(z; t)dz^2$  connecting  $z_1$  and  $z_2$ , oriented from  $z_1$  to  $z_2$ , and  $\Gamma(z, e^{i\theta}\infty)$  (resp.  $\Gamma(e^{i\theta}\infty, z)$ ) stands for the orthogonal trajectory ending at  $z$ , approaching infinity at the angle  $\theta$ , and oriented away from  $z$  (resp. oriented towards  $z$ ).<sup>1</sup>

The following theorem has been proven in [7, Theorem 3.2] and it describes the geometry of  $\Gamma_t$  when  $t \in \overline{O_{\text{one-cut}}}$ .

<sup>1</sup>This notation is unambiguous as the corresponding trajectories are unique for polynomial differentials as follows from Teichmüller's lemma.



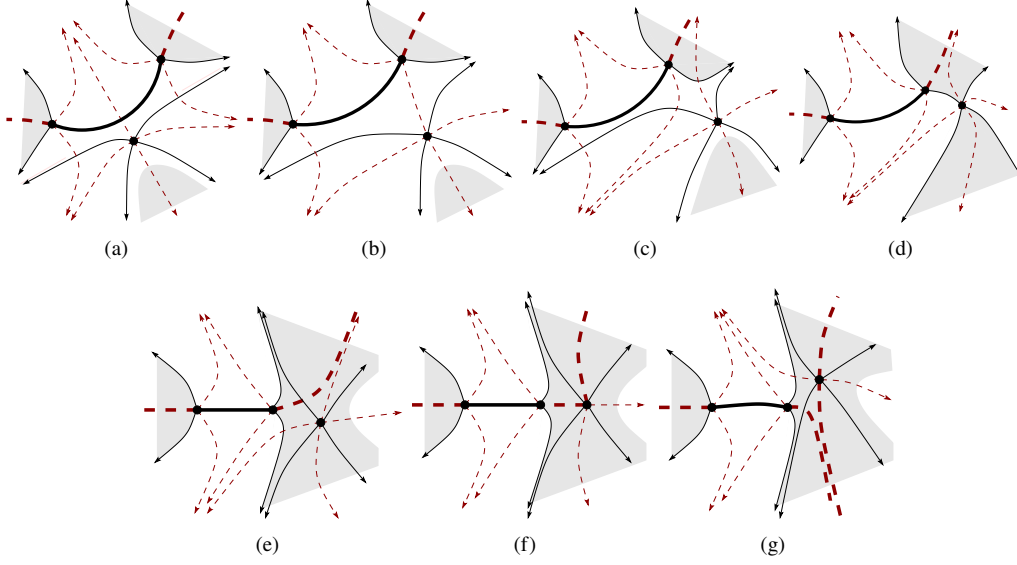


FIGURE 4. Schematic representation of the critical (solid) and critical orthogonal (dashed) graphs of  $-Q(z;t)dz^2$  when  $t \in O_{\text{one-cut}}$ . The bold curves represent the preferred S-curve  $\Gamma_t$ . Shaded region is the set  $\{\mathcal{U}(z;t) < 0\}$ .

**Theorem 3.1.** Let  $\mu_t$  and  $Q(z;t)$  be as in Theorem 2.1, and let  $J_t = \text{supp}(\mu_t)$ . When  $t \in \overline{O_{\text{one-cut}}}$ , the polynomial  $Q(z;t)$  is of the form

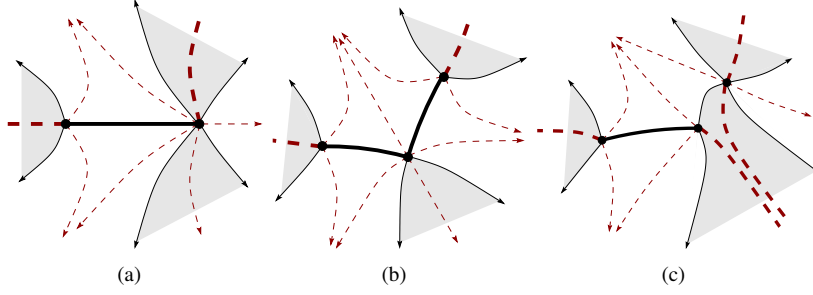
$$(3.3) \quad Q(z;t) = \frac{1}{4}(z - a(t))(z - b(t))(z - c(t))^2$$

with  $a(t)$ ,  $b(t)$ , and  $c(t)$  given by

$$(3.4) \quad \begin{cases} a(t) & := x(t) - i\sqrt{2}/\sqrt{x}(t), \\ b(t) & := x(t) + i\sqrt{2}/\sqrt{x}(t), \\ c(t) & := -x(t), \end{cases}$$

where  $\sqrt{x}(t)$  is the branch holomorphic in  $O_{\text{one-cut}}$  satisfying  $\sqrt{x}(0) = e^{\pi i/3}$ . The set  $J_t$  consists of a single arc and

- (I) if  $t \in O_{\text{one-cut}}$ , then  $J_t = \Gamma[a, b]$  and an S-curve  $\Gamma_t \in \mathcal{T}$  can be chosen as
  - (a)  $\Gamma(e^{\pi i \infty}, a) \cup J_t \cup \Gamma(b, e^{\pi i/3} \infty)$  when  $t$  belongs to the connected component bounded by  $S \cup C_{\text{split}} \cup e^{2\pi i/3} S$ , see Figure 4(a–e);
  - (b)  $\Gamma(e^{\pi i \infty}, a) \cup J_t \cup \Gamma(b, c) \cup \Gamma(c, e^{\pi i/3} \infty)$  when  $t \in S$ , see Figure 4(f);
  - (c)  $\Gamma(e^{\pi i \infty}, c) \cup \Gamma(c, a) \cup J_t \cup \Gamma(b, e^{\pi i/3} \infty)$  when  $t \in e^{2\pi i/3} S$ ;
  - (d)  $\Gamma(e^{\pi i \infty}, a) \cup J_t \cup \Gamma(b, e^{-\pi i/3} \infty) \cup \Gamma(e^{-\pi i/3} \infty, c) \cup \Gamma(c, e^{\pi i/3} \infty)$  when  $t$  belongs to the connected component bounded by  $S \cup C_{\text{birth}}^b$ , see Figure 4(g);
  - (e)  $\Gamma(e^{\pi i \infty}, c) \cup \Gamma(c, e^{-\pi i/3}) \cup \Gamma(e^{-\pi i/3} \infty, a) \cup J_t \cup \Gamma(b, e^{\pi i/3} \infty)$  when  $t$  belongs to the connected component bounded by  $e^{2\pi i/3} S \cup C_{\text{birth}}^a$ .
- (II) if  $t = t_{\text{cr}}$  (resp.  $t = e^{2\pi i/3} t_{\text{cr}}$ ), then  $J_t = \Gamma[a, b]$ ,  $c$  coincides with  $b$  (resp.  $a$ ), and an S-curve  $\Gamma_t \in \mathcal{T}$  can be chosen as in Case I(a), see Figure 5(a).
- (III) if  $t \in C_{\text{split}}$ , then  $J_t = \Gamma[a, c] \cup \Gamma[c, b]$  and an S-curve  $\Gamma_t \in \mathcal{T}$  can be chosen as in Case I(a), see Figure 5(b).
- (IV) if  $t \in C_{\text{birth}}^b$  (resp.  $t \in C_{\text{birth}}^a$ ), then  $J_t = \Gamma[a, b]$  and an S-curve  $\Gamma_t \in \mathcal{T}$  can be chosen as in Case I(d) (resp. Case I(e)), see Figure 5(c).

FIGURE 5. This is a continuation of Figure 4 for the case  $t \in \partial O_{\text{one-cut}}$ .

Now, let  $O_{\text{two-cut}} := \mathbb{C} \setminus \overline{O_{\text{one-cut}}}$ . Then the following theorem holds.

**Theorem 3.2.** *Let  $\mu_t$  and  $Q(z; t)$  be as in Theorem 2.1,  $J_t = \text{supp}(\mu_t)$ . When  $t \in O_{\text{two-cut}}$ , the polynomial  $Q(z; t)$  is of the form*

$$(3.5) \quad Q(z; t) = \frac{1}{4}(z - a_1(t))(z - b_1(t))(z - a_2(t))(z - b_2(t))$$

with  $a_1(t)$ ,  $b_1(t)$ ,  $a_2(t)$ , and  $b_2(t)$  all distinct. The real and imaginary parts of  $a_i(t)$ ,  $b_i(t)$  are real analytic functions of  $\text{Re}(t)$  and  $\text{Im}(t)$  when  $t \in O_{\text{two-cut}}$ ; however, at no point of  $O_{\text{two-cut}}$  any of the functions  $a_i(t)$ ,  $b_i(t)$  is analytic. The S-curve  $\Gamma_t$  can be chosen as

$$\Gamma(e^{\pi i} \infty, a_1(t)) \cup J_{t,1} \cup \Gamma(b_1(t), e^{-\pi i/3} \infty) \cup \Gamma(e^{-\pi i/3} \infty, a_2(t)) \cup J_{t,2} \cup \Gamma(b_2(t), e^{\pi i/3} \infty),$$

where  $J_t = J_{t,1} \cup J_{t,2}$  and  $J_{t,i} := \Gamma[a_i(t), b_i(t)]$ ,  $i \in \{1, 2\}$ , see Figure 6 (this also explains how we choose the labeling of the zeros of  $Q(z; t)$  in the considered case). Moreover, it holds that

$$(3.6) \quad a_1(t), b_1(t) \rightarrow a(t^*), \quad b_1(t), a_2(t) \rightarrow c(t^*), \quad \text{and} \quad a_2(t), b_2(t) \rightarrow b(t^*)$$

as  $t \rightarrow t^*$  with  $t^* \in C_{\text{birth}}^a \cup \{e^{2\pi i/3} t_{\text{cr}}\}$ ,  $t^* \in C_{\text{split}}$ , and  $t^* \in C_{\text{birth}}^b \cup \{t_{\text{cr}}\}$ , respectively.

We prove Theorem 3.2 in Section 5.

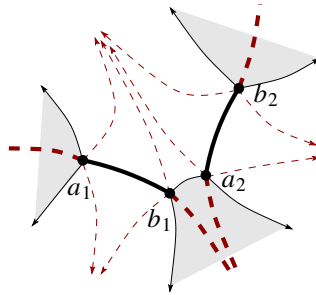


FIGURE 6. The schematic representation of the critical and critical orthogonal graphs of  $-Q(z; t) dz^2$  when  $t \in O_{\text{two-cut}}$ . The bold curves represent the preferred S-curve  $\Gamma_t$ . Shaded region is the set  $\{\mathcal{U}(z; t) < 0\}$ .

#### 4. MAIN RESULTS

In this section we assume that  $t \in O_{\text{two-cut}}$  and that  $Q(z; t)$  and  $J_t = J_{t,1} \cup J_{t,2}$  are as in Theorem 3.2. We also put

$$E_t := \{a_1(t), b_1(t), a_2(t), b_2(t)\}, \quad J_{t,i}^\circ := J_{t,i} \setminus E_t, \quad \text{and} \quad J_t^\circ := J_{t,1}^\circ \cup J_{t,2}^\circ.$$

**Convention 2.** When it comes to the definition of the S-contour  $\Gamma_t$ , it will be more practical for us to change the choice of  $\Gamma_t$  made in Theorem 3.2. Rather than including unbounded trajectories  $\Gamma(b_1(t), e^{-\pi i/3}\infty)$  and  $\Gamma(e^{-\pi i/3}\infty, a_2(t))$ , we shall replace them by a smooth Jordan arc, say  $I_t$ , connecting  $b_1(t)$  and  $a_2(t)$ , such that  $I_t^\circ := I_t \setminus E_t$  lies entirely within the region  $\{\mathcal{U}(z; t) < 0\}$  while there exist  $s_1(t) \in \Gamma(b_1(t), e^{-\pi i/3}\infty)$  and  $s_2(t) \in \Gamma(e^{-\pi i/3}\infty, a_2(t))$  for which  $\Gamma(b_1(t), s_1(t)) \cup \Gamma(s_2(t), a_2(t)) \subset I_t^\circ$ , see Figure 14 further below.

To describe the asymptotics of the orthogonal polynomials themselves, we need to construct the Szegő function of  $e^{V(z;t)}$ .

**Proposition 4.1.** *Let constants  $\varsigma(t)$  and  $d_i(t)$  be given by*

$$(4.1) \quad \varsigma(t) := \frac{2t}{3d_0(t)} \quad \text{and} \quad d_i(t) := \int_{I_t} \frac{s^i ds}{Q^{1/2}(s;t)},$$

where, as usual, we use the branch  $Q^{1/2}(z; t) = \frac{1}{2}z^2 + \mathcal{O}(z)$  as  $z \rightarrow \infty$ . Then the Szegő function

$$(4.2) \quad \mathcal{D}(z; t) := \exp \left\{ \frac{1}{2}V(z; t) + \frac{1}{3} \left( z + \int_{I_t} \frac{3\varsigma(t)}{s-z} \frac{ds}{Q^{1/2}(s;t)} \right) Q^{1/2}(z; t) \right\}$$

is holomorphic and non-vanishing in  $\mathbb{C} \setminus \Gamma[a_1(t), b_2(t)]$  with continuous traces on  $J_t^\circ \cup I_t^\circ$  that satisfy

$$(4.3) \quad \begin{cases} \mathcal{D}_+(s; t)\mathcal{D}_-(s; t) = e^{V(s;t)}, & s \in J_t^\circ, \\ \mathcal{D}_+(s; t) = \mathcal{D}_-(s; t)e^{2\pi i\varsigma(t)}, & s \in I_t^\circ. \end{cases}$$

Denote by  $\mu_1(t) := \int s d\mu_t(s)$  the first moment of the equilibrium measure  $\mu_t$ . Then

$$(4.4) \quad \mathcal{D}(z; t) = \exp \left\{ \frac{1}{3} - \frac{(\varsigma d_1)(t)}{2} \right\} \left( 1 + \frac{t^2 + \mu_1(t) - 3(\varsigma d_2)(t)/2}{3z} + \mathcal{O}\left(\frac{1}{z^2}\right) \right)$$

as  $z \rightarrow \infty$ . In particular, it holds that  $\mathcal{D}(\infty; t) = \exp \left\{ \frac{1-t(d_1/d_0)(t)}{3} \right\}$ .

We shall also denote by  $D(z; t) := \mathcal{D}(z; t)/\mathcal{D}(\infty; t)$  the normalized Szegő function. We prove Proposition 4.1 in Section 6.1.

To describe the geometric growth of the orthogonal polynomials, let us define

$$(4.5) \quad \mathcal{Q}(z; t) := \int_{b_2(t)}^z Q^{1/2}(s; t) ds, \quad z \in \mathbb{C} \setminus \Gamma(e^{\pi i}\infty, b_2(t)],$$

(where we deviate slightly from Convention 1 and denote by  $\Gamma[s_1, s_2]$  a subarc of  $\Gamma_t$  connecting  $s_1$  to  $s_2$ ; for example,  $\Gamma(e^{\pi i}\infty, a_2]$  includes  $I_t$  rather than the short trajectories connecting  $b_1$  to  $a_2$ ). Observe that  $\mathcal{U}(z; t) = 2\text{Re}(\mathcal{Q}(z; t))$  as defined in (2.8). This function has the following properties.

**Proposition 4.2.** *Let constants  $\tau(t), \omega(t)$  be given by*

$$(4.6) \quad \tau(t) := -\frac{1}{\pi i} \int_{I_t} Q^{1/2}(s; t) ds \quad \text{and} \quad \omega(t) := -\frac{1}{\pi i} \int_{J_{t,1}} Q_+^{1/2}(s; t) ds.$$

These constants are necessarily real (in fact,  $\omega(t) = \mu_t(J_{t,1})$ , see (2.7)). The function  $e^{\mathcal{Q}(z;t)}$  is holomorphic in  $\mathbb{C} \setminus \Gamma[a_1(t), b_2(t)]$  and there exists a constant  $\ell_*(t)$  such that

$$(4.7) \quad \exp \left\{ \frac{V(z; t) - \ell_*(t)}{2} + \mathcal{Q}(z; t) \right\} = z + \mathcal{O}(1) \quad \text{as} \quad z \rightarrow \infty.$$

Moreover,  $\mathcal{Q}(z; t)$  possesses continuous traces on  $J_t^\circ \cup I_t^\circ$  that are purely imaginary on  $J_t^\circ$  and satisfy

$$(4.8) \quad \begin{cases} e^{\mathcal{Q}_+(s;t) + \mathcal{Q}_-(s;t)} = 1, & s \in J_{t,2}^\circ, \\ e^{\mathcal{Q}_+(s;t) + \mathcal{Q}_-(s;t)} = e^{2\pi i\tau(t)}, & s \in J_{t,1}^\circ, \\ e^{\mathcal{Q}_+(s;t)} = e^{\mathcal{Q}_-(s;t) - 2\pi i\omega(t)}, & s \in I_t^\circ. \end{cases}$$

We prove Proposition 4.2 in Section 6.2. Observe that it follows from Theorem 3.2 that  $|e^{\mathcal{Q}(z;t)}|$  is less than 1 when  $\mathcal{U}(z;t) < 0$  (the shaded areas of Figure 6), is equal to 1 on the critical trajectories (black curves), and otherwise is greater than 1. The constant  $\ell_*(t)$  is such that  $\operatorname{Re}(\ell_*(t)) = \ell_{\Gamma_t}$ , see (2.2). Moreover, the following identity holds.

**Proposition 4.3.** *Set*

$$(4.9) \quad \mathbf{B}(t) := - \left( \int_{J_{t,1}} \frac{ds}{\mathcal{Q}_+^{1/2}(s;t)} \right) / \left( \int_{I_t} \frac{ds}{\mathcal{Q}^{1/2}(s;t)} \right).$$

Let  $\theta(u; \mathbf{B}(t))$  be the Riemann theta function associated with  $\mathbf{B}(t)$ , i.e.,

$$(4.10) \quad \theta(u; \mathbf{B}(t)) = \sum_{k \in \mathbb{Z}} \exp \{ \pi i \mathbf{B}(t) k^2 + 2\pi i u k \}, \quad u \in \mathbb{C}.$$

Then it holds that

$$(4.11) \quad e^{\ell_*(t)} = \left( 4\mathcal{D}(\infty; t) \frac{e^{\pi i \tau(t)(\varsigma(t) + \omega(t) + \mathbf{B}(t)\tau(t))}}{b_1(t) - a_1(t) + b_2(t) - a_2(t)} \frac{\theta(\varsigma(t) + \omega(t) + \mathbf{B}(t)\tau(t))}{\theta(0)} \right)^2.$$

We prove this proposition in Section 7.6. We also explain right after the statement of Proposition 7.1 further below that  $b_1(t) - a_1(t) + b_2(t) - a_2(t) \neq 0$ .

Another auxiliary function we need is given by

$$(4.12) \quad A(z; t) := \frac{1}{2} \left( \left( \frac{z - b_2(t)}{z - a_2(t)} \frac{z - b_1(t)}{z - a_1(t)} \right)^{1/4} + \left( \frac{z - b_2(t)}{z - a_2(t)} \frac{z - b_1(t)}{z - a_1(t)} \right)^{-1/4} \right)$$

for  $z \in \overline{\mathbb{C}} \setminus J_t$ , where the branches are chosen so that the summands are holomorphic in  $\overline{\mathbb{C}} \setminus J_t$  and have value 1 at infinity. As explained in Section 7.2, this function can be analytically continued through each side of  $J_t^\circ$  and is non-vanishing in the domain of the definition.

Given a sequence  $\{N_n\}_{n \in \mathbb{N}}$ , we define further below in (7.23) functions  $\Theta_n(z; t)$ , which are ratios of Riemann theta functions (4.10) with various arguments. To shorten the presentation of the main results, we only discuss main features of the functions  $\Theta_n(z; t)$  and defer the detailed construction and description of further properties to Section 7.

**Proposition 4.4.** *Functions  $\Theta_n(z; t)$  are holomorphic in  $\overline{\mathbb{C}} \setminus \Gamma[a_1(t), b_2(t)]$  with at most one zero there. They have continuous traces on  $I_t^\circ$  that satisfy*

$$\Theta_{n+}(s; t) = \Theta_{n-}(s; t) e^{2\pi i (n\omega(t) + (n - N_n)\varsigma(t))}.$$

Assume that there exists a constant  $N_*$  such that  $|n - N_n| \leq N_*$  for all  $n \in \mathbb{N}$ . Then for any  $\delta > 0$  there exists a constant  $C(t, \delta, N_*)$  such that

$$|A(z; t) \Theta_n(z; t)| \leq C(t, \delta, N_*), \quad z \in \overline{\mathbb{C}} \setminus \bigcup_{e \in E_t} \{|z - e| < \delta\},$$

that is, including the traces on  $\Gamma[a_1(t), b_2(t)]$ . Given  $\varepsilon > 0$ , let  $\mathbb{N}(t, \varepsilon)$  be a subsequence of indices  $n$  such that  $\Theta_n(z; t)$  is non-vanishing in  $\{|z| \geq 1/\varepsilon\}$ . Then there exists a constant  $c(t, \varepsilon) > 0$  such that

$$|\Theta_n(\infty; t)| \geq c(t, \varepsilon), \quad n \in \mathbb{N}(t, \varepsilon).$$

As in the case of the Szegő functions  $\mathcal{D}(z; t)$ , it would be convenient for us to renormalize  $\Theta_n(z; t)$  at infinity. Thus, we set  $\vartheta_n(z; t) := \Theta_n(z; t) / \Theta_n(\infty; t)$ . Observe that  $\vartheta_n(z; t) D^{N_n - n}(z; t) e^{n\mathcal{Q}(z;t)}$  are functions holomorphic in  $\mathbb{C} \setminus J_t$ .

Proposition 4.4 has substance only if the sets  $\mathbb{N}(t, \varepsilon)$  have infinite cardinality. Recall  $\mathbf{B}(t)$  from (4.9) above. It follows from the general theory of Riemann surfaces, see Section 7.1, that  $\operatorname{Im}(\mathbf{B}(t)) > 0$ . In particular, any  $s \in \mathbb{C}$  can be uniquely written as  $x + \mathbf{B}(t)y$  for some  $x, y \in \mathbb{R}$ .

**Proposition 4.5.** *Write  $\varsigma(t) =: x(t) + \mathbf{B}(t)y(t)$ ,  $x(t), y(t) \in \mathbb{R}$ . It holds that*

- (i) if  $x(t), y(t) \in \mathbb{Z}$ , then the functions  $\Theta_n(z; t)$  do not depend on the choice of  $N_n$  and at least one of the integers  $n, n + 1$  belongs to  $\mathbb{N}(t, \varepsilon)$ <sup>2</sup>;
- (ii) otherwise, for any natural number  $N_* > 0$  there exists a choice of  $N_n$  such that  $|n - N_n| \leq N_*$  and  $\mathbb{N}(t, \varepsilon) = \mathbb{N}$  for all  $\varepsilon$  small enough.

On the other hand if the sequence  $\{N_n\}_{n \in \mathbb{N}}$  is fixed, we can claim the following.

**Proposition 4.6.** *Let  $\{N_n\}_{n \in \mathbb{N}}$  be a sequence such that  $|n - N_n| \leq N_*$  for some  $N_* \geq 0$ . The sequence  $\mathbb{N}(t, \varepsilon)$  is infinite for all  $\varepsilon$  small enough unless there exist integers  $d > 0, k, i_1, i_2, m_1, m_2$  such that*

$$(4.13) \quad \varsigma(t) = (i_1 + \mathbf{B}(t)i_2)/d, \quad \omega(t)d = (k - 1)i_1 + m_1d, \quad \text{and} \quad \tau(t)d = (k - 1)i_2 + m_2d,$$

where at least one of the fractions  $i_1/d, i_2/d$  is irreducible,  $d$  is even,  $i_1$  and  $i_2$  are odd, and the sequence  $\{N_n\}$  is such that all but finitely many numbers  $nk - N_n$  are divisible by  $d/2$  but not by  $d$ . Moreover, the following special cases take place:

- (i) if there exists an infinite subsequence  $\{n_l\}$  such that  $N_{n_l+1} - N_{n_l} \in \{0, 1\}$  (in particular, this happens when  $N_n = n$ , in which case the differences are always equal to 1), then at least one of the integers  $n_l, n_l + 1$  belongs to  $\mathbb{N}(t, \varepsilon)$ .
- (ii) if one of the triples  $\omega(t), x(t), 1$  or  $\tau(t), y(t), 1$  is rationally independent, then at least one of the integers  $n, n + 1$  belongs to  $\mathbb{N}(t, \varepsilon)$ .

We prove Propositions 4.4, 4.5 and 4.6 in Sections 7.5, 7.3, and 7.4, respectively.

**Theorem 4.1.** *Let  $t \in O_{\text{two-cut}}$  and  $\{N_n\}_{n=1}^\infty$  be a sequence such that  $|n - N_n| \leq N_*$  for some  $N_*$  fixed. Let  $P_n(z; t, N)$  be defined by (1.11)–(1.10) and*

$$\psi_n(z; t) := P_n(z; t, N_n) e^{-n(V(z;t) - \ell_*(t))/2}.$$

Given  $\varepsilon > 0$ , let  $\mathbb{N}(t, \varepsilon)$  be as in Proposition 4.4. Then for all  $n \in \mathbb{N}(t, \varepsilon)$  large enough it holds that<sup>3</sup>

$$(4.14) \quad \psi_n(z; t) = \left( (A\vartheta_n D^{N_n - n})(z; t) + O_\varepsilon(n^{-1}) \right) e^{nQ(z;t)}$$

locally uniformly in  $\mathbb{C} \setminus J_t$ ; moreover,

$$(4.15) \quad \psi_n(s; t) = (A\vartheta_n D^{N_n - n})_+(s; t) e^{nQ_+(s;t)} + (A\vartheta_n D^{N_n - n})_-(s; t) e^{nQ_-(s;t)} + O_\varepsilon(n^{-1})$$

locally uniformly on  $J_t^\circ$ .

Recall that each  $\vartheta_n(z; t)$  might have a single zero in  $\mathbb{C} \setminus J_t$ . If these zeros accumulate at a point  $z_*$  along a subsequence of  $\mathbb{N}(t, \varepsilon)$ , then the polynomials  $P_n(z; t, N_n)$  will have a single zero approaching  $z_*$  along this subsequence by (4.14) and Rouché's theorem. With this exception, it also follows from (4.14) that  $P_n(z; t, N_n)$  are eventually zero free on compact subsets  $\mathbb{C} \setminus J_t$ . The main part of the proof of Theorem 4.1 is carried out in Section 8 with auxiliary details relegated to Sections 6 and 7.

The functions  $\vartheta_n(z; t)$  were defined as pull-backs from one of the sheets (a copy of  $\overline{\mathbb{C}} \setminus J_t$ ) of certain ratios of Riemann theta functions defined on the (two-sheeted) Riemann surface of  $Q^{1/2}(z; t)$ . To describe asymptotics of the recurrence coefficients we shall need the pull-backs from the other sheet as well, which, for the purpose of the next theorem, we denote by  $\vartheta_n^*(z; t)$ . To complicate matters further, our analysis requires another related family of ratios of theta functions, which, in the next theorem, we will denote by  $\tilde{\vartheta}_n(z; t)$ , see (7.24) for the rigorous definition of these functions. The proof of Proposition 4.4 in fact shows that  $\vartheta_n^*(\infty; t) = O_\varepsilon(1)$  and  $\tilde{\vartheta}_n(\infty; t) = O_\varepsilon(1)$  for  $n \in \mathbb{N}(t, \varepsilon)$ .

**Theorem 4.2.** *Let  $S_k(t) := \sum_{i=1}^2 (b_i^k(t) - a_i^k(t))$  for  $k \in \{1, 2\}$ . In the setting of Theorem 4.1, it holds for all  $n \in \mathbb{N}(t, \varepsilon)$  large enough that*

$$(4.16) \quad \begin{cases} h_n(t, N_n) &= \frac{\pi}{2} e^{-n\ell_*(t)} S_1(t) \mathcal{D}^{2(n-N_n)}(\infty; t) \vartheta_n^*(\infty; t) + O_\varepsilon(n^{-1}), \\ \gamma_n^2(t, N_n) &= \frac{1}{16} S_1^2(t) \vartheta_n^*(\infty; t) \tilde{\vartheta}_n(\infty; t) + O_\varepsilon(n^{-1}), \end{cases}$$

<sup>2</sup>In fact, this claim can be improved depending on the arithmetic properties of the numbers  $\omega(t)$  and  $\tau(t)$ .

<sup>3</sup>By writing  $f(n) = O_\varepsilon(n^{-1})$  we mean that there exists a constant  $C_\varepsilon$  depending on  $\varepsilon$  for which  $|f(n)| \leq C_\varepsilon n^{-1}$ .

where  $\mathcal{D}(\infty; t)$  is as in Proposition 4.1,  $\ell_*(t)$  as in (4.7), and  $\mathcal{O}_\varepsilon(\cdot)$  is interpreted in the same way as in Theorem 4.1. Furthermore, it holds that

$$(4.17) \quad \beta_n(t, N_n) = \frac{\vartheta_n^*(\infty; t) \left( \frac{S_2(t)}{2S_1(t)} + \frac{d}{dz} (\log \vartheta_n(1/z; t) + \log \vartheta_n^*(1/z; t)) \Big|_{z=0} \right) + \mathcal{O}_\varepsilon(n^{-1})}{\vartheta_n^*(\infty; t) + \mathcal{O}_\varepsilon(n^{-1})},$$

where we agree that  $\vartheta_n^*(\infty; t) \frac{d}{dz} (\log \vartheta_n^*(1/z; t)) \Big|_{z=0} = \frac{d}{dz} \vartheta_n^*(1/z; t) \Big|_{z=0}$  even when  $\vartheta_n^*(\infty; t) = 0$ .

Of course, if there exists a subsequence of  $\mathbb{N}(t, \varepsilon)$  along which the values  $\vartheta_n^*(\infty; t)$  are separated away from zero, then formula (4.17) can be further simplified. However, such a subsequence might not exist (in particular, one can deduce from (7.7) and (7.20) further below that it does not exist when  $N_n = n$  and  $\omega(t) = \tau(t) = 1/2$ ). Notice that formulae (4.16) are not immediately reflective of the equality in (1.14). However, as we point out further below in (9.12), the leading term of the asymptotics of  $\gamma_n^2(t, N_n)$  can be rewritten to make this connection more transparent. Theorem 4.2 is proven in Section 9, but its proof heavily relies on the material of Section 8.

## 5. S-CURVES

In this section we prove Theorem 3.2. We do it in several steps. In Section 5.1 we gather results about quadratic differentials that will be important to us throughout the proof. In Section 5.2 we show the validity of formula (3.5); that is, we prove that we are indeed in the two-cut case when  $t \in \mathcal{O}_{\text{two-cut}}$ . In Section 5.3 we show that the critical and critical orthogonal graphs of

$$\varpi_t(z) := -Q(z; t) dz^2$$

do look like as depicted on Figure 6. In Section 5.4 we describe the dependence of the zeros of  $Q(z; t)$  on  $t$  by showing that the variational condition (2.2) and the S-property (2.6) yield that the zeros satisfy a certain system of real equations with non-zero Jacobian, see (5.17) and (5.18), and that this system is, in fact, uniquely solved by them. Finally, in Section 5.5 we establish the limits in (3.6).

**5.1. On Quadratic Differentials.** To start, let us also recall the following important result, known as Teichmüller's lemma, see [37, Theorem 14.1]. Let  $P$  be a geodesic polygon of a quadratic differential, that is, a Jordan curve in  $\overline{\mathbb{C}}$  that consists of a finite number of trajectories and orthogonal trajectories of this differential. Then it holds that

$$(5.1) \quad \sum_{z \in P} \left( 1 - \theta(z) \frac{2 + \text{ord}(z)}{2\pi} \right) = 2 + \sum_{z \in \text{int}(P)} \text{ord}(z),$$

where  $\text{ord}(z)$  is the order of  $z$  with respect to the considered differential and  $\theta(z) \in [0, 2\pi]$ ,  $z \in P$ , is the interior angle of  $P$  at  $z$ . Both sums in (5.1) are finite since only critical points of the differential have a non-zero contribution.

Let us briefly recall the main properties of the differential  $\varpi_t(z)$ . The only critical points of  $\varpi_t(z)$  are the zeros of  $Q(z; t)$  and the point at infinity. Regular points have order 0, the order of a zero of  $Q(z; t)$  is equal to its multiplicity, and infinity is a critical point of order  $-8$ . Through each regular point passes exactly one trajectory and one orthogonal trajectory of  $\varpi_t(z)$ , which are orthogonal to each other at the point. Two distinct (orthogonal) trajectories meet only at critical points [37, Theorem 5.5]. As  $Q(z; t)$  is a polynomial, no finite union of (orthogonal) trajectories can form a closed Jordan curve, while a trajectory and an orthogonal trajectory can intersect at most once [32, Lemma 8.3]. Furthermore, (orthogonal) trajectories of  $\varpi_t(z)$  cannot be recurrent (dense in two-dimensional regions) [26, Theorem 3.6]. From each critical point of order  $m > 0$  there emanate  $m + 2$  critical trajectories whose consecutive tangent lines at the critical point form an angle  $2\pi/(m + 2)$ . Furthermore, since infinity is a pole of order 8, the critical trajectories can approach infinity only in six distinguished directions, namely, asymptotically to the lines  $L_{-\pi/6}$ ,  $L_{\pi/6}$ , and  $L_{\pi/2}$ , where  $L_\theta = \{z : z = re^{i\theta}, r \in (-\infty, \infty)\}$ . In fact, there exists a neighborhood of infinity such that any trajectory entering it necessarily tends to infinity [37, Theorem 7.4]. This discussion

also applies to orthogonal trajectories. In particular, they can approach infinity asymptotically to the lines  $L_0$ ,  $L_{\pi/3}$ , and  $L_{2\pi/3}$  only.

Denote by  $\mathcal{G}$  the critical graph of  $\varpi_t(z)$ , that is, the totality of all the critical trajectories of  $\varpi_t(z)$ . Then, see [26, Theorem 3.5], the complement of  $\mathcal{G}$  can be written as a disjoint union of either half-plane or strip domains. Recall that a *half-plane (or end) domain* is swept by trajectories unbounded in both directions that approach infinity along consecutive critical directions. Its boundary is connected and consists of a union of two unbounded critical trajectories and a finite number (possibly zero) of short trajectories of  $\varpi_t(z)$ . The map  $z \mapsto \int^z \sqrt{-\varpi_t}$  maps end domains conformally onto half planes  $\{z \in \mathbb{C} \mid \operatorname{Re}(z) > c\}$  for some  $c \in \mathbb{R}$  that depends on the domain, and extends continuously to the boundary. Similarly, a *strip domain* is again swept by trajectories unbounded in both directions, but its boundary consists of two disjoint  $\varpi_t(z)$ -paths, each of which is comprised of two unbounded critical trajectories and a finite number (possibly zero) of short trajectories. The map  $z \mapsto \int^z \sqrt{-\varpi_t}$  maps strip domains conformally onto vertical strips  $\{w \in \mathbb{C} \mid c_1 < \operatorname{Re}(w) < c_2\}$  for some  $c_1, c_2 \in \mathbb{R}$  depending on the domain, and extends continuously to their boundaries. The number  $c_2 - c_1$  is known as the *width of a strip domain* and can be calculated in terms of  $\varpi_t(z)$  as

$$(5.2) \quad \left| \operatorname{Re} \int_p^q \sqrt{-\varpi_t} \right|$$

where  $p, q$  belong to different components of the boundary of the domain.

**5.2. Proof of (3.5).** Assume to the contrary that for a given  $t \in O_{\text{two-cut}}$  we are in one-cut case. That is, there exists a choice of  $a, b$ , and  $c$  such that the polynomial  $Q(z; t)$  from (2.6) has the form (3.3). It follows from (2.6) in conjunction with (1.11) that

$$(5.3) \quad Q(z; t) = \left( \frac{-z^2 + t}{2} - \frac{1}{z} + \mathcal{O}(z^{-2}) \right)^2 = \frac{(z^2 - t)^2}{4} + z + C$$

for some constant  $C$ . Then, by equating the coefficients in (3.3) and (5.3), we obtain a system of equations

$$(5.4) \quad \begin{cases} a + b + 2c = 0, \\ ab + c^2 + 2(a + b)c = -2t, \\ 2abc + (a + b)c^2 = -4. \end{cases}$$

Setting  $x := (a + b)/2$  and eliminating the product  $ab$  from the second and third relations in (5.4) yields

$$(5.5) \quad x^3 - tx - 1 = 0,$$

which is exactly the equation appearing before (3.2). Given any solution of (5.5), say  $x(t)$ , then  $a(t)$ ,  $b(t)$ , and  $c(t)$  are necessarily expressed via (3.4). Theorem 2.1 and the variational condition (2.2) imply that there must exist a contour  $\Gamma_t \in \mathcal{T}$  (this class of contours was defined right after (2.3)) such that

$$(5.6) \quad \mathcal{U}(z; t) \leq 0 \quad \text{for all } z \in \Gamma_t,$$

see (2.8). In what follows, we shall show that no such contour exists in  $\mathcal{T}$  for any of the three possible choices of  $x(t)$  solving (5.5) when  $t \in O_{\text{two-cut}}$  and  $Q(z; t)$  is given by (3.3) and (3.4).

In accordance with the above strategy, observe that the solutions of (5.5) can be written as

$$(5.7) \quad x_k(t) = u_k(t) + \frac{t}{3u_k(t)}, \quad u_k(t) := \left( \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{t^3}{27}} \right)^{1/3} e^{2k\pi i/3},$$

$k \in \{0, 1, 2\}$ , with all branches being principal. It can be readily verified that  $x_1(t)$  is analytic in  $\mathbb{C} \setminus (\overline{e^{2\pi i/3}S}, \overline{S})$  (here,  $\overline{\cdot}$  means topological closure), see (3.2) for the definition of the ray  $S$ , and

$$(5.8) \quad \begin{cases} x_0(t) &= e^{4\pi i/3} x_1(t e^{-2\pi i/3}), \\ x_1(t) &= e^{4\pi i/3} x_1(\overline{t} e^{2\pi i/3}), \\ x_2(t) &= x_1(\overline{t}). \end{cases}$$

Furthermore, noting that the function  $x(t)$ , defined after (3.2), maps  $\Omega_{\text{one-cut}}$  onto  $O_{\text{one-cut}}$ , it can be easily checked that  $x(t)$  is evaluated as shown on Figure 7, where the dashed lines are the chosen branch cuts of  $x_1(t)$ . In particular,  $x(t)$  can be analytically continued across  $C_{\text{split}}$ ,  $C_{\text{birth}}^a$ , and  $C_{\text{birth}}^b$ , see (3.2) and Figure 7. In what follows, we consider what happens in the case of each of these continuations.

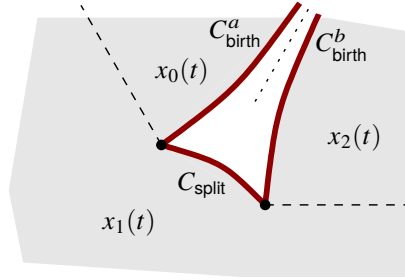


FIGURE 7. Determination of  $x(t)$

Continue  $x(t)$  into  $O_{\text{two-cut}}$  by either  $x_2(t)$  or  $x_0(t)$ , that is, analytically across either  $C_{\text{birth}}^b$  or  $C_{\text{birth}}^a$ , see Figure 7. The first and the last symmetries in (5.8) then yield that  $\varpi_t(z)$  is either equal to

$$\overline{\varpi_{\overline{t}}(\overline{z})} \quad \text{or} \quad \varpi_{t e^{-2\pi i/3}}(z e^{-4\pi i/3}).$$

Since the set  $O_{\text{two-cut}}$  is symmetric with respect to the line  $L_{\pi/3}$ , its rotation by  $-2\pi/3$  is equal to its reflection across the real axis. Thus, the critical graph  $\varpi_t(z)$  when  $t \in O_{\text{two-cut}}$  and  $x(t)$  is continued by either  $x_2(t)$  or  $x_0(t)$  is equal to the reflection across the real axis or the rotation by  $4\pi/3$  of the critical graph of  $\varpi_{t_*}(z)$  for some  $t_*$  such that  $\overline{t_*} \in O_{\text{two-cut}}$ . These graphs were studied in [7, Theorems 3.2 and 3.4] and determined to have the structure as depicted on Figure 4(a–c) (or the reflection of these three panels across the line  $L_{2\pi/3}$ ). A direct examination shows that none of these critical graphs form a curve in  $\mathcal{T}$  for which (5.6) holds (such a curve must belong to the closure of the gray regions on Figure 4).

Suppose now that we continue  $x(t)$  by  $x_1(t)$ , that is, analytically across  $C_{\text{split}}$ . For such a choice of  $x(t)$ , the critical graph of  $\varpi_t(z)$  was studied in [25] when  $t \in L_{\pi/3}$  (in which case the critical graph is symmetric with respect to  $L_{2\pi/3}$ ). In particular, it was shown that  $x(t) \in L_{2\pi/3}$ , no union of critical trajectories join  $a$  and  $b$ , and no critical trajectory of  $\varpi_t(z)$  crosses the line  $L_{2\pi/3}$  when  $t \in L_{\pi/3} \cap O_{\text{two-cut}}$ , see [25, Lemma 3.2]. Since critical trajectories cannot intersect, can approach infinity asymptotically to the lines  $L_{\pi/6}$ ,  $L_{\pi/2}$ , and  $L_{-\pi/6}$  only and must obey Teichmüller's lemma (5.1), the critical graph of  $\varpi_t(z)$  must be as on Figure 8. Clearly, this critical graph does not yield a curve in  $\mathcal{T}$  for which (5.6) holds. Thus, to complete the proof, we need to argue that the structure of the critical trajectories of  $\varpi_t(z)$  remains the same for all  $t \in O_{\text{two-cut}}$  in the considered case. Observe that it is enough to show that all the trajectories out of  $b$  approach infinity.

Recall that the trajectories emanating out of  $b$  are part of the level set  $\{\mathcal{U}(z; t) = 0\}$ , see (2.8). When  $t \in L_{\pi/3} \cap O_{\text{two-cut}}$ , these trajectories approach infinity at the angles  $-\pi/6$ ,  $\pi/6$ , and  $\pi/2$ , see Figure 8. Since the values of  $\mathcal{U}(z; t)$  analytically depend on  $t$ , the same must be true in a neighborhood of each such  $t$ . This will remain so until one of the trajectories hits a critical point different from the one at infinity. As can be seen from Figure 8, this critical point must necessarily be  $c = -x$ , see (3.4). That is, as long as  $\mathcal{U}(-x; t) \neq 0$ , the trajectories out of  $b$  will asymptotically



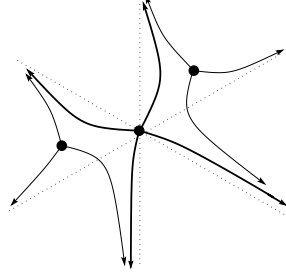


FIGURE 8. The critical graph of  $\varpi_t(z)$  when  $x(t)$  is analytically continued across  $C_{\text{split}}$ .

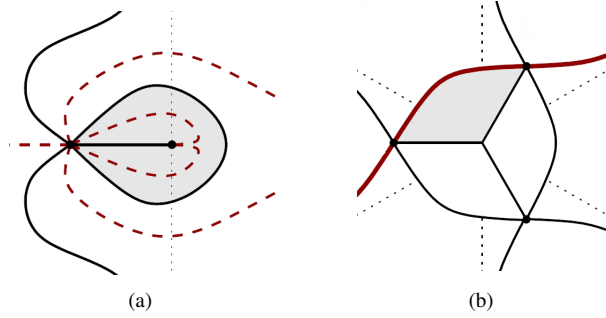


FIGURE 9. Shaded regions represent the domains within which  $2x_1^3(t)$  (panel a) and  $x_1(t)$  (panel b) change when  $t \in O_{\text{two-cut}}$ .

behave as on Figure 8. When  $x(t)$  is continued into  $O_{\text{two-cut}}$  by  $x_1(t)$ , its values lie within the gray region on Figure 9(b), see also Figure 2(b). Respectively, the values  $2x^3(t)$  lie within the gray region on Figure 9(a), see also Figure 2(a). It was verified in [7, Section 5.3] that

$$\mathcal{U}(-x; t) = \operatorname{Re} \left( \frac{2}{3} \int_{-1}^{2x^3} \left( 1 + \frac{1}{s} \right)^{3/2} ds \right),$$

where the path of integration lies within the shaded domain on Figure 9(a). Hence,  $\mathcal{U}(-x; t) = 0$  if and only if  $2x^3$  belongs to a trajectory of  $-(1 + 1/s)^3 ds^2$  emanating from  $-1$ . These trajectories are drawn on Figures 2 and 9(a) (black lines). Thus,  $\mathcal{U}(-x; t) \neq 0$  in the considered case as claimed.

**5.3. Critical Graph of  $\varpi_t(z)$ .** Let, as usual,  $Q(z; t)$  be the polynomial guaranteed by Theorem 2.1. According to what precedes, it has the form (3.5) when  $t \in O_{\text{two-cut}}$ . Recall the properties of the differential  $\varpi_t(z) = -Q(z; t)dz^2$  described at the beginning of Section 5.1. In particular, it has four critical points of order 1, which, for a moment, we label as  $z_1(t), z_2(t), z_3(t), z_4(t)$  (these are the zeros of  $Q(z; t)$ ), a critical point of order  $-8$  at infinity, and no other critical points. It follows from Theorem 2.1(4) and (2.2) with (2.8) that

$$(5.9) \quad \operatorname{Re} \left( \int_{\Gamma_t[z_i(t), z_j(t)]} Q_+^{1/2}(z; t) dz \right) = 0,$$

where  $\Gamma_t[z_i(t), z_j(t)]$  is the subarc of  $\Gamma_t$  with endpoints  $z_i(t), z_j(t)$  and  $Q_+^{1/2}(z; t)$  is the trace of  $Q^{1/2}(z; t)$  on the positive side of  $\Gamma_t$ . Equations (5.9) imply existence of three short critical trajectories of  $\varpi_t(z)$ . Indeed, if all three critical trajectories out of a zero  $z_i(t)$  approach infinity, then  $z_i(t)$  must belong to a boundary of at least one strip domain. Let  $z_j(t)$  be a different zero of  $Q(z; t)$  belonging to the other component of the boundary of this strip domain. Then it follows from (5.2) and (5.9) that the width of this strip domain is 0, which is impossible. Thus, each zero of  $Q(z; t)$  must be

coincident with at least one short trajectory. Therefore, either there is a zero, say  $z_4(t)$ , connected by short trajectories to the remaining three zeros or there are at least two short trajectories connecting two pairs of zeros. In the latter case, label these zeros by  $a_1(t), b_1(t)$  and  $a_2(t), b_2(t)$ . If the other two trajectories out of both  $a_1(t)$  and  $b_1(t)$  approach infinity, one of these zeros again must belong to the boundary of a strip domain with either  $a_2(t)$  or  $b_2(t)$  belonging to the other component of the boundary. As before, (5.9) yields that the width of this strip domain is 0, which, again, is impossible. Thus, in this case there also exists a third short critical trajectory. Then we choose a labeling of the zeros so that  $b_1(t)$  and  $a_2(t)$  are connected by this trajectory.

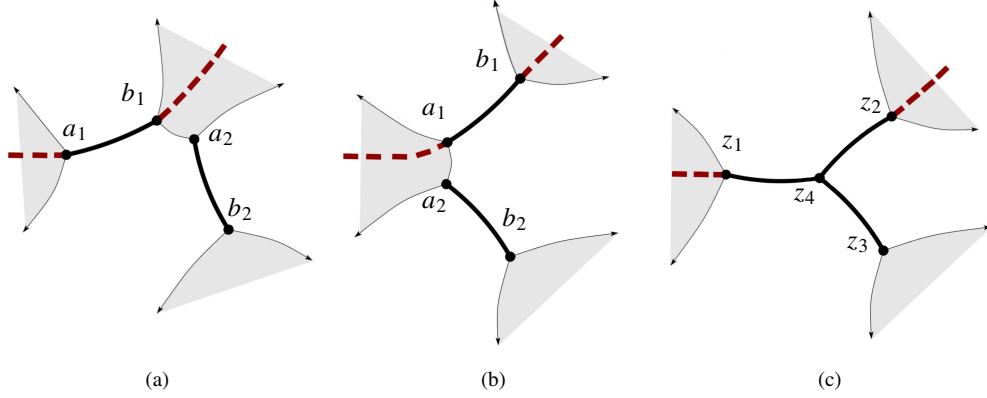


FIGURE 10. Geometries of the critical graph of  $\varpi_t(z)$ . Shaded regions represent the open set  $\{\mathcal{U}(z; t) < 0\}$ , the white regions represent the open  $\{\mathcal{U}(z; t) > 0\}$ , and the red dashed arcs form  $\Gamma_t \setminus J_t$ .

Since short critical trajectories cannot form closed curves, there cannot be any more of them. That is, the remaining critical trajectories are unbounded. Consider the two unbounded critical trajectories out of  $z_1(t)$  in the case where short ones form a threefold, see Figure 10(c). Since critical trajectories cannot intersect and the remaining zeros are connected to  $z_1(t)$  by short critical trajectories, the unbounded critical trajectories out of  $z_1(t)$  delimit a half-plane domain and, in particular, must approach infinity along consecutive critical directions (those are given by the angles  $(2k + 1)\pi/6$ ,  $k \in \{0, \dots, 5\}$ , see Section 5.1). Clearly, the same is true for the unbounded critical trajectories out of  $z_2(t)$  and  $z_3(t)$  as well as for the unbounded critical trajectories out of  $a_1(t)$ ,  $b_2(t)$ , and the union of the unbounded critical trajectory out of  $b_1(t)$ , the short critical trajectory connecting  $b_1(t)$  to  $a_2(t)$ , and the unbounded critical trajectory out of  $a_2(t)$  in the case where short critical trajectories form a Jordan arc, see Figure 10.

Now, let  $\mathcal{U}(z; t)$  be given by (2.8). Clearly,  $\mathcal{U}(z; t)$  is a subharmonic function which is equal to zero on  $J_t$ , see (2.2). Since  $\mathcal{U}(z; t)$  must have the same sign on both sides of each subarc of  $J_t$  by (2.4), it follows from the maximum principle for subharmonic functions that it is positive there. Further, since trajectories of  $\varpi_t(z)$  cannot form a closed Jordan curve, all the connected components of the open set  $\{\mathcal{U}(z; t) < 0\}$  must necessarily extend to infinity. Since  $\operatorname{Re}(V(z; t))$  is the dominant term of  $\mathcal{U}(z; t)$  around infinity, see (2.8), for any  $\delta > 0$  there exists  $R > 0$  sufficiently large so that

$$(5.10) \quad \begin{cases} (S_{\pi/3, \delta} \cup S_{\pi, \delta} \cup S_{-\pi/3, \delta}) \cap \{|z| > R\} & \subset \{\mathcal{U}(z; t) < 0\}, \\ (S_{0, \delta} \cup S_{2\pi/3, \delta} \cup S_{-2\pi/3, \delta}) \cap \{|z| > R\} & \subset \{\mathcal{U}(z; t) > 0\}, \end{cases}$$

where  $S_{\theta, \delta} := \{|\arg(z) - \theta| < \pi/6 - \delta\}$ . Altogether, the critical graph of  $\varpi_t(z)$  must look like either on Figure 6 or on Figure 10.

It remains to show that  $\varpi_t(z)$  cannot have the critical graph as on any of the panels of Figure 10. To this end, recall that the contour  $\Gamma_t$  must contain  $J_t$  and two unbounded arcs extending to infinity in the directions  $\pi/3$  and  $\pi$  (red dashed unbounded arcs on Figure 10). Let  $\Gamma_*$  be obtained from

$\Gamma_t$  by dropping the short trajectory that is a part of  $J_t$  and whose removal keeps  $\Gamma_*$  connected (this can be done for any of the panels on Figure 10). Observe that  $\Gamma_*$  also belongs to  $\mathcal{T}$ . Let  $\mu_*$  be the weighted equilibrium distribution on  $\Gamma_*$  as defined in Definition 2.1. Since  $\Gamma_* \subset \Gamma_t$ , it holds that  $\mu_* \in \mathcal{M}(\Gamma_t)$ . Moreover, since  $\mu_* \neq \mu_t$ ,  $\mathcal{E}_V(\Gamma_*) = E_V(\mu_*) > E_V(\mu_t) = \mathcal{E}_V(\Gamma_t)$ , see (2.1). However, the last inequality clearly contradicts (2.5).

We have shown that the critical graph of  $\varpi_t(z)$  must look like on Figure 6. As the critical orthogonal and critical trajectories cannot intersect, the structure of the critical orthogonal graph is uniquely determined by structure of the critical graph. Now, we can completely fix the labeling of the zeros of  $Q(z; t)$  by given the label  $a_1(t)$  to one that is incident with the orthogonal critical trajectory extending to infinity asymptotically to the ray  $\arg(z) = \pi$ .

**5.4. Dependence on  $t$ .** We start with some general considerations. Let  $f(z)$  and  $g(z)$  be analytic functions of  $z = x + iy$ . Consider a determinant of the form

$$D = \begin{vmatrix} \partial_x \operatorname{Re}(f) & \partial_y \operatorname{Re}(f) & * \\ \partial_x \operatorname{Im}(f) & \partial_y \operatorname{Im}(f) & * \\ \partial_x \operatorname{Re}(g) & \partial_y \operatorname{Re}(g) & * \end{vmatrix},$$

where the entries of the third column are not important for the forthcoming computation. Due to Cauchy-Riemann relations it holds that  $f' = \partial_x \operatorname{Re}(f) + i\partial_x \operatorname{Im}(f) = \partial_y \operatorname{Im}(f) - i\partial_y \operatorname{Re}(f)$ . Therefore,

$$D = \begin{vmatrix} \operatorname{Re}(f') & -\operatorname{Im}(f') & * \\ \operatorname{Im}(f') & \operatorname{Re}(f') & * \\ \operatorname{Re}(g') & -\operatorname{Im}(g') & * \end{vmatrix} = \begin{vmatrix} f' & if' & * \\ \operatorname{Im}(f') & \operatorname{Re}(f') & * \\ \operatorname{Re}(g') & -\operatorname{Im}(g') & * \end{vmatrix} = \frac{i}{2} \begin{vmatrix} f' & if' & * \\ \overline{f'} & -i\overline{f'} & * \\ \operatorname{Re}(g') & -\operatorname{Im}(g') & * \end{vmatrix}$$

by adding the second row times  $i$  to the first one and then multiplying the second row by  $-2i$  and adding the first row to it. It further holds that

$$D = \frac{i}{2} \begin{vmatrix} 2f' & if' & * \\ 0 & -i\overline{f'} & * \\ g' & -\operatorname{Im}(g') & * \end{vmatrix} = \frac{i}{2} \begin{vmatrix} 2f' & 0 & * \\ 0 & -i\overline{f'} & * \\ g' & -ig'/2 & * \end{vmatrix} = \frac{1}{2} \begin{vmatrix} f' & 0 & * \\ 0 & \overline{f'} & * \\ g' & g' & * \end{vmatrix},$$

where we added the second column times  $-i$  to the first one, then added the first column times  $-i/2$  to the second one, and then factored 2 from the first column,  $-i$  from the second one, and  $1/2$  from the third row.

Now, let  $f_j(z_1, z_2, z_3, z_4)$ ,  $j \in \{1, 2, 3, 4, 5\}$ , be analytic functions in each variable  $z_i = x_i + iy_i$ . We would like to compute the Jacobian of the following system of real-valued functions of  $x_1, y_1, \dots, x_4, y_4$ :

$$(5.11) \quad \operatorname{Re}(f_1), \operatorname{Im}(f_1), \operatorname{Re}(f_2), \operatorname{Im}(f_2), \operatorname{Re}(f_3), \operatorname{Im}(f_3), \operatorname{Re}(f_4), \operatorname{Re}(f_5).$$

That is, we are interested in

$$\operatorname{Jac} = \begin{vmatrix} \operatorname{Re}(f_{11}) & -\operatorname{Im}(f_{11}) & \operatorname{Re}(f_{12}) & -\operatorname{Im}(f_{12}) & \operatorname{Re}(f_{13}) & -\operatorname{Im}(f_{13}) & \operatorname{Re}(f_{14}) & -\operatorname{Im}(f_{14}) \\ \operatorname{Im}(f_{11}) & \operatorname{Re}(f_{11}) & \operatorname{Im}(f_{12}) & \operatorname{Re}(f_{12}) & \operatorname{Im}(f_{13}) & \operatorname{Re}(f_{13}) & \operatorname{Im}(f_{14}) & \operatorname{Re}(f_{14}) \\ \operatorname{Re}(f_{21}) & -\operatorname{Im}(f_{21}) & \operatorname{Re}(f_{22}) & -\operatorname{Im}(f_{22}) & \operatorname{Re}(f_{23}) & -\operatorname{Im}(f_{23}) & \operatorname{Re}(f_{24}) & -\operatorname{Im}(f_{24}) \\ \operatorname{Im}(f_{21}) & \operatorname{Re}(f_{21}) & \operatorname{Im}(f_{22}) & \operatorname{Re}(f_{22}) & \operatorname{Im}(f_{23}) & \operatorname{Re}(f_{23}) & \operatorname{Im}(f_{24}) & \operatorname{Re}(f_{24}) \\ \operatorname{Re}(f_{31}) & -\operatorname{Im}(f_{31}) & \operatorname{Re}(f_{32}) & -\operatorname{Im}(f_{32}) & \operatorname{Re}(f_{33}) & -\operatorname{Im}(f_{33}) & \operatorname{Re}(f_{34}) & -\operatorname{Im}(f_{34}) \\ \operatorname{Im}(f_{31}) & \operatorname{Re}(f_{31}) & \operatorname{Im}(f_{32}) & \operatorname{Re}(f_{32}) & \operatorname{Im}(f_{33}) & \operatorname{Re}(f_{33}) & \operatorname{Im}(f_{34}) & \operatorname{Re}(f_{34}) \\ \operatorname{Re}(f_{41}) & -\operatorname{Im}(f_{41}) & \operatorname{Re}(f_{42}) & -\operatorname{Im}(f_{42}) & \operatorname{Re}(f_{43}) & -\operatorname{Im}(f_{43}) & \operatorname{Re}(f_{44}) & -\operatorname{Im}(f_{44}) \\ \operatorname{Re}(f_{51}) & -\operatorname{Im}(f_{51}) & \operatorname{Re}(f_{52}) & -\operatorname{Im}(f_{52}) & \operatorname{Re}(f_{53}) & -\operatorname{Im}(f_{53}) & \operatorname{Re}(f_{54}) & -\operatorname{Im}(f_{54}) \end{vmatrix},$$

where  $f_{ji} := \partial_{z_i} f_j$ . By performing the same row and column operations as for the determinant  $D$  above, we get that

$$\text{Jac} = \frac{1}{2i} \begin{vmatrix} f_{11} & 0 & f_{12} & 0 & f_{13} & 0 & f_{14} & 0 \\ 0 & \bar{f}_{11} & 0 & \bar{f}_{12} & 0 & \bar{f}_{13} & 0 & \bar{f}_{14} \\ f_{21} & 0 & f_{22} & 0 & f_{23} & 0 & f_{24} & 0 \\ 0 & \bar{f}_{21} & 0 & \bar{f}_{22} & 0 & \bar{f}_{23} & 0 & \bar{f}_{24} \\ f_{31} & 0 & f_{32} & 0 & f_{33} & 0 & f_{34} & 0 \\ 0 & \bar{f}_{31} & 0 & \bar{f}_{32} & 0 & \bar{f}_{33} & 0 & \bar{f}_{34} \\ f_{41} & \bar{f}_{41} & f_{42} & \bar{f}_{42} & f_{43} & \bar{f}_{43} & f_{44} & \bar{f}_{44} \\ f_{51} & \bar{f}_{51} & f_{52} & \bar{f}_{52} & f_{53} & \bar{f}_{53} & f_{54} & \bar{f}_{54} \end{vmatrix}$$

with the constant in front of the determinant coming from  $(i/2)^3 \times 2^4 \times (-i)^4 \times (1/2)^2$ , where the first factor is due to multiplications of the second, fourth and sixth rows by  $-2i$ , the second is due to factoring 2 from the odd columns, the third one comes from factoring  $-i$  from the even columns, and the fourth factor is due to factoring  $1/2$  from the last two rows. Assume further that

$$(5.12) \quad \begin{cases} f_1(z_1, z_2, z_3, z_4) &= z_1 + z_2 + z_3 + z_4, \\ f_2(z_1, z_2, z_3, z_4) &= z_1 z_2 + z_1 z_3 + z_1 z_4 + z_2 z_3 + z_2 z_4 + z_3 z_4, \\ f_3(z_1, z_2, z_3, z_4) &= z_2 z_3 z_4 + z_1 z_3 z_4 + z_1 z_2 z_4 + z_1 z_2 z_3. \end{cases}$$

Then, by using the above explicit expressions and subtracting the first (resp. second) column from the third, fifth, and seventh (resp. fourth, sixth, and eighth), we get that

$$\text{Jac} = \frac{1}{2i} \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & z_1 - z_2 & 0 & z_1 - z_3 & 0 & z_1 - z_4 & 0 & 0 \\ * & * & 0 & \bar{z}_1 - \bar{z}_2 & 0 & \bar{z}_1 - \bar{z}_3 & 0 & 0 & \bar{z}_1 - \bar{z}_4 \\ * & * & \frac{z_1 - z_2}{(z_3 + z_4)^{-1}} & 0 & \frac{z_1 - z_3}{(z_2 + z_4)^{-1}} & 0 & \frac{z_1 - z_4}{(z_2 + z_3)^{-1}} & 0 & 0 \\ * & * & 0 & \frac{\bar{z}_1 - \bar{z}_2}{(\bar{z}_3 + \bar{z}_4)^{-1}} & 0 & \frac{\bar{z}_1 - \bar{z}_3}{(\bar{z}_2 + \bar{z}_4)^{-1}} & 0 & 0 & \frac{\bar{z}_1 - \bar{z}_4}{(\bar{z}_2 + \bar{z}_3)^{-1}} \\ * & * & (z_1 - z_2)g_{42} & (\bar{z}_1 - \bar{z}_2)\bar{g}_{42} & (z_1 - z_3)g_{43} & (\bar{z}_1 - \bar{z}_3)\bar{g}_{43} & (z_1 - z_4)g_{44} & (\bar{z}_1 - \bar{z}_4)\bar{g}_{44} & 0 \\ * & * & (z_1 - z_2)g_{52} & (\bar{z}_1 - \bar{z}_2)\bar{g}_{52} & (z_1 - z_3)g_{53} & (\bar{z}_1 - \bar{z}_3)\bar{g}_{53} & (z_1 - z_4)g_{54} & (\bar{z}_1 - \bar{z}_4)\bar{g}_{54} & 0 \end{vmatrix},$$

where  $g_{ji}(z_1, z_2, z_3, z_4) := (z_1 - z_i)^{-1}(f_{ji} - f_{j1})(z_1, z_2, z_3, z_4)$ ,  $j \in \{4, 5\}$  and  $i \in \{2, 3, 4\}$ . Hence,

$$\text{Jac} = \frac{|z_1 - z_2|^2 |z_1 - z_3|^2 |z_1 - z_4|^2}{2i} \begin{vmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ z_3 + z_4 & 0 & z_2 + z_4 & 0 & z_2 + z_3 & 0 \\ 0 & \bar{z}_3 + \bar{z}_4 & 0 & \bar{z}_2 + \bar{z}_4 & 0 & \bar{z}_2 + \bar{z}_3 \\ g_{42} & \bar{g}_{42} & g_{43} & \bar{g}_{43} & g_{44} & \bar{g}_{44} \\ g_{52} & \bar{g}_{52} & g_{53} & \bar{g}_{53} & g_{54} & \bar{g}_{54} \end{vmatrix}.$$

Absolutely analogous computation now implies that

$$\text{Jac} = \frac{|z_1 - z_2|^2 |z_1 - z_3|^2 |z_1 - z_4|^2 |z_2 - z_3|^2 |z_2 - z_4|^2}{2i} \begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ h_{43} & \bar{h}_{43} & h_{44} & \bar{h}_{44} \\ h_{53} & \bar{h}_{53} & h_{54} & \bar{h}_{54} \end{vmatrix},$$

where  $h_{ji}(z_1, z_2, z_3, z_4) := (z_2 - z_i)^{-1}(g_{ji} - g_{j2})(z_1, z_2, z_3, z_4)$ ,  $j \in \{4, 5\}$  and  $i \in \{3, 4\}$ . The above expression immediately yields that

$$(5.13) \quad \text{Jac} = \frac{\prod_{i < j} |z_i - z_j|^2}{2i} \begin{vmatrix} k_4 & \bar{k}_4 \\ k_5 & \bar{k}_5 \end{vmatrix} = \prod_{i < j} |z_i - z_j|^2 \text{Im}(k_4 \bar{k}_5),$$

where  $k_j(z_1, z_2, z_3, z_4) := (z_3 - z_4)^{-1}(g_{j4} - g_{j3})(z_1, z_2, z_3, z_4)$ ,  $j \in \{4, 5\}$ . Finally, let

$$w(z) := \sqrt{(z - z_1)(z - z_2)(z - z_3)(z - z_4)}$$

be a branch such that  $w(z) = z^2 + \mathcal{O}(z)$  as  $z \rightarrow \infty$  with branch cuts  $\gamma_{12}$  and  $\gamma_{34}$  that are bounded, disjoint, and smooth, and where  $\gamma_{ij}$  connects  $z_i$  to  $z_j$ . Further, select a smooth arc  $\gamma_{32}$  disjoint (except for the endpoints) from the previous two. Set

$$(5.14) \quad f_4(z_1, z_2, z_3, z_4) := 4 \int_{\gamma_{12}} w(z) dz \quad \text{and} \quad f_5(z_1, z_2, z_3, z_4) := 4 \int_{\gamma_{32}} w(z) dz,$$

where we integrate  $w(z)$  on the positive side of  $\gamma_{12}$ . Let  $O \subset \{z_i \neq z_j, i \neq j, i, j \in \{1, 2, 3, 4\}\}$  be a domain such that there exist arcs  $\gamma_{ij}(z_1, z_2, z_3, z_4)$  with the above properties for each  $(z_1, z_2, z_3, z_4) \in O$ , which, in addition, possess parameterizations that depend continuously on each variable  $z_1, z_2, z_3, z_4$ . Then the functions  $f_j(z_1, z_2, z_3, z_4)$ ,  $j \in \{4, 5\}$ , are analytic in each variable  $z_i$  for  $(z_1, z_2, z_3, z_4) \in O$ . Furthermore, it can be readily computed that

$$g_{ji} = 2 \int \frac{w(z) dz}{(z - z_1)(z - z_i)}, \quad h_{ji} = -2 \int \frac{w(z) dz}{(z - z_1)(z - z_2)(z - z_i)}, \quad \text{and} \quad k_j = 2 \int \frac{dz}{w(z)},$$

where the integrals are taken over  $\gamma_{12}$  when  $j = 4$  and  $\gamma_{32}$  when  $j = 5$ . Trivially, (5.13) can be rewritten as

$$(5.15) \quad \text{Jac} = 4 \prod_{i < j} |z_i - z_j|^2 \text{Im} \left( \int_{\gamma_{12}} \frac{dz}{w(z)} \int_{\gamma_{32}} \overline{\frac{dz}{w(z)}} \right).$$

Now, consider the Riemann surface  $\mathfrak{S} := \{z := (z, w) : w^2 = (z - z_1)(z - z_2)(z - z_3)(z - z_4)\}$ . Denote by  $\pi : \mathfrak{S} \rightarrow \overline{\mathbb{C}}$  the natural projection  $\pi(z) = z$  and write  $w(z)$  for a rational function on  $\mathfrak{S}$  such that  $z = (z, w(z))$ . Let  $\beta := \pi^{-1}(\gamma_{12})$  and  $\alpha := \pi^{-1}(\gamma_{32})$ . Orient these cycles so that

$$2 \int_{\gamma_{12}} \frac{dz}{w(z)} = \oint_{\beta} \frac{dz}{w(z)} \quad \text{and} \quad 2 \int_{\gamma_{32}} \frac{dz}{w(z)} = \oint_{\alpha} \frac{dz}{w(z)}.$$

Observe that the cycles  $\alpha, \beta$  form the right pair at the point of their intersection and that  $\mathfrak{S} \setminus \{\alpha \cup \beta\}$  is simply connected. Hence, the cycles  $\alpha, \beta$  form a homology basis on  $\mathfrak{S}$ . Since the genus of  $\mathfrak{S}$  is 1, it has a unique (up to multiplication by a constant) holomorphic differential. It is quite easy to check that this differential is  $dz/w(z)$ . Hence, we get from (5.15) that

$$(5.16) \quad \text{Jac} = \prod_{i < j} |z_i - z_j|^2 \text{Im} \left( \oint_{\beta} \frac{dz}{w(z)} \oint_{\alpha} \overline{\frac{dz}{w(z)}} \right) > 0$$

when  $(z_1, z_2, z_3, z_4) \in O$ , where the last inequality was shown by Riemann.

Now, let  $Q(z; t) = \frac{1}{4}(z - a_1(t))(z - b_1(t))(z - a_2(t))(z - b_2(t))$  be the polynomial from Theorem 2.1. It can be easily deduced from (2.6) that

$$(5.17) \quad \begin{cases} f_1(a_1, b_1, a_2, b_2) &= 0, \\ f_2(a_1, b_1, a_2, b_2) &= -2t, \\ f_3(a_1, b_1, a_2, b_2) &= -4, \end{cases}$$

where the functions  $f_i(z_1, z_2, z_3, z_4)$ ,  $i \in \{1, 2, 3\}$ , are given by (5.12).

Fix  $t^* \in O_{\text{two-cut}}$  and let  $\delta^* > 0$  be small enough so that all four disks  $\{|z - z_i^*| \leq \delta^*\}$  are disjoint, where  $z_1^* = a_1(t^*)$ ,  $z_2^* = b_1(t^*)$ ,  $z_3^* = a_2(t^*)$ , and  $z_4^* = b_2(t^*)$ . Let  $O := \{(z_1, z_2, z_3, z_4) : |z_i - z_i^*| < \delta^*\}$ ,  $s_i^*$  be the point of intersection of  $\{|z - z_i^*| = \delta^*\}$  and  $J_{t^*}$ ,  $i \in \{1, 2, 3, 4\}$ , and  $u_i^*$  be the point of intersection of  $\{|z - z_i^*| = \delta^*\}$  and  $\Gamma(b_1(t^*), a_2(t^*))$ ,  $i \in \{2, 3\}$ , where, as usual,  $\Gamma(a, b)$  is the subarc of the trajectories of  $\varpi_t(z)$  connecting  $a$  and  $b$ . Then we can choose  $\gamma_{12} = [z_1, s_1^*] \cup \Gamma(s_1^*, s_2^*) \cup [s_2^*, z_2]$ ,  $\gamma_{32} = [z_3, u_3^*] \cup \Gamma(u_3^*, u_2^*) \cup [u_2^*, z_2]$ , and  $\gamma_{34} = [z_3, s_3^*] \cup \Gamma(s_3^*, s_4^*) \cup [s_4^*, z_4]$ , where  $[a, b]$  is the line segment connecting  $a$  and  $b$  in  $\mathbb{C}$ , see Figure

11. Clearly, the arcs  $\gamma_{ij}$  continuously depend on  $(z_1, z_2, z_3, z_4) \in O$ . Now, the relations (5.9) can be rewritten as

$$(5.18) \quad \operatorname{Re}(f_4(a_1, b_1, a_2, b_2)) = 0 \quad \text{and} \quad \operatorname{Re}(f_5(a_1, b_1, a_2, b_2)) = 0$$

with  $f_j(z_1, z_2, z_3, z_4)$ ,  $j \in \{4, 5\}$ , given by (5.14), where we set  $w(z) := Q^{1/2}(z; t)$ . It follows

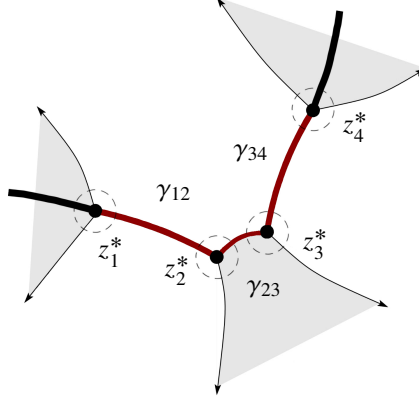


FIGURE 11. A schematic illustration of  $\gamma_{ij}$  and  $z_i^*$ .

from (5.16) and the implicit function theorem that there exists a neighborhood of  $t^*$  in which system (5.17) and (5.18) is uniquely solvable and the solution, say  $(a_1^*(t), b_1^*(t), a_2^*(t), b_2^*(t))$ , is such that the real and imaginary parts of  $a_i^*(t), b_i^*(t)$  are real analytic functions of  $\operatorname{Re}(t)$  and  $\operatorname{Im}(t)$  for  $t$  in this neighborhood. These local solutions are unique only locally around the point  $(a_1(t^*), b_1(t^*), a_2(t^*), b_2(t^*))$  and we still need to argue that they do coincide with the zeros  $a_i(t), b_i(t)$  of  $Q(z; t)$  (of course, it holds that  $a_i^*(t^*) = a_i(t^*)$  and  $b_i^*(t^*) = b_i(t^*)$ ).

In what follows, we always assume that  $t$  belongs to a disk centered at  $t^*$  of small enough radius so that the functions  $a_i^*(t), b_i^*(t)$  are defined and continuous in this disk. Let

$$(5.19) \quad Q^*(z; t) := \frac{1}{4}(z - a_1^*(t))(z - b_1^*(t))(z - a_2^*(t))(z - b_2^*(t)) \quad \text{and} \quad \varpi_t^*(z) := -Q^*(z; t)dz^2.$$

Further, let  $\mathcal{U}^*(z; t)$  be defined as in (2.8) with  $Q(z; t)$  replaced by  $Q^*(z; t)$ . For the moment, choose the branch cut for  $Q^*(z; t)^{1/2}$  as in the paragraph between (5.17) and (5.18). Each function  $\mathcal{U}^*(z; t)$  is harmonic off the chosen branch cut and can be continued harmonically across it by  $-\mathcal{U}^*(z; t)$ . Moreover, it follows immediately from their definition that the functions  $\mathcal{U}^*(z; t)$  are uniformly bounded above and below on any compact set for all considered values of the parameter  $t$ . Thus, they converge to  $\mathcal{U}(z; t^*)$  locally uniformly in  $\mathbb{C} \setminus \{a_1(t^*), b_1(t^*), a_2(t^*), b_2(t^*)\}$  as  $t \rightarrow t^*$ . Since the critical graph of  $\varpi_t^*(z)$  is the zero-level set of  $\mathcal{U}^*(z; t)$ , it converges to the critical graph of  $\varpi_{t^*}(z)$  in any disk  $\{|z| < R\}$ . Due to relations (5.18), the argument at the beginning of Section 5.3 also shows that the critical graph of  $\varpi_t^*(z)$  has three short critical trajectories, which, due to uniform convergence, necessarily connect  $a_1^*(t)$  to  $b_1^*(t)$ ,  $b_1^*(t)$  to  $a_2^*(t)$ , and  $a_2^*(t)$  to  $b_2^*(t)$  (the disk around  $t^*$  can be decreased if necessary). Thus, arguing as in Section 5.3 and using uniform convergence, we can show that Figure 6 also schematically represents the critical and critical orthogonal graphs of  $\varpi_t^*(z)$ . Moreover, let us now take the branch cut for  $Q^*(z; t)^{1/2}$ , say  $J_t^*$ , along the short critical trajectories of  $\varpi_t^*(z)$  connecting  $a_i^*(t)$  to  $b_i^*(t)$ . Then the shading on Figure 6 corresponds to regions where  $\mathcal{U}^*(z; t)$  is positive (white) and negative (gray).

Define  $\Gamma_t^*$  to be the union of the critical orthogonal trajectory of  $\varpi_t^*(z)$  that connects infinity to  $a_1^*(t)$ , its short critical trajectories, and the critical orthogonal trajectory that connects  $b_2^*(t)$  to infinity. Orient it so that the positive direction proceeds from  $a_1^*(t)$  to  $b_2^*(t)$ . Let the measures  $\mu_t^*$  be given by (2.7) with  $Q(z; t)$  replaced by  $Q^*(z; t)$  and  $J_t$  replaced by  $J_t^*$ . Clearly, each  $\mu_t^*$  is a positive measure.

Moreover, it has a unit mass by the Cauchy theorem and since  $Q^*(z; t)^{1/2} = (z^2 - t)/2 + 1/z + \mathcal{O}(1/z^2)$  due to (5.17), see also (5.3). Thus, it holds that

$$F^*(z; t) := Q^*(z; t)^{1/2} + \frac{V'(z; t)}{2} - \int \frac{d\mu_t^*(s)}{z - s} = \mathcal{O}(z^{-2})$$

as  $z \rightarrow \infty$  and  $F^*(z; t)$  is holomorphic in  $\overline{\mathbb{C}} \setminus J_t^*$ . It follows from the well known behavior of Cauchy integrals of smooth densities, see [23, Section I.8], that the traces of  $F^*(z; t)$  on  $J_t^*$  are bounded. It further follows from the Sokhotski-Plemelj formulae, see [23, Section I.4], that

$$\begin{aligned} F_+^*(s; t) - F_-^*(s; t) &= Q_+^*(s; t)^{1/2} - Q_-^*(s; t)^{1/2} - \left( \int \frac{Q_+^*(w; t)^{1/2} dw}{w - z} \frac{d\pi i}{\pi i} \right)_+ + \left( \int \frac{Q_+^*(w; t)^{1/2} dw}{w - z} \frac{d\pi i}{\pi i} \right)_- \\ &= Q_+^*(s; t)^{1/2} - Q_-^*(s; t)^{1/2} - 2Q_+^*(s; t)^{1/2} \equiv 0 \end{aligned}$$

for  $s \in J_t^*$ . Hence,  $F^*(z; t)$  is an entire function and therefore is identically zero. This observation, in particular, yields that

$$\mathcal{U}^*(z; t) := \operatorname{Re} \left( 2 \int_{b_2^*(t)}^z Q^*(s; t)^{1/2} ds \right) = \ell_t^* - \operatorname{Re}(V(z; t)) - 2U\mu_t^*(z)$$

for some constant  $\ell_t^*$ , see also (2.8). Since  $\mathcal{U}^*(z; t)$  can be harmonically continued across  $J_t^*$  by  $-\mathcal{U}^*(z; t)$ , we get that  $\mu_t^*$  satisfies (2.4); that is  $J_t^*$  has the S-property in the field  $\operatorname{Re}(V(z; t))$ . Since  $\Gamma_t^* \in \mathcal{T}$ , it follows from the uniqueness part of Theorem 2.1(2) that  $\mu_t^* = \mu_t$ . In particular,  $a_i^*(t) = a_i(t)$  and  $b_i^*(t) = b_i(t)$ ,  $i \in \{1, 2\}$ .

Since any compact subset of  $\mathcal{O}_{\text{two-cut}}$  can be covered by finitely many disks where the above considerations hold, the functions  $a_i(t)$ ,  $b_i(t)$  continuously depend on  $t \in \mathcal{O}_{\text{two-cut}}$  and, moreover, their real and imaginary parts are real analytic functions of  $\operatorname{Re}(t)$  and  $\operatorname{Im}(t)$ .

Let us now show that at no point in  $\mathcal{O}_{\text{two-cut}}$  are the functions  $a_i(t)$ ,  $b_i(t)$  analytic in  $t$ . Suppose, for the sake of contradiction, that one of these functions, say  $b_2(t)$ , is analytic in  $t$  at some point  $t_0 \in \mathcal{O}_{\text{two-cut}}$ . Then all the other endpoint functions,  $a_1(t)$ ,  $b_1(t)$ , and  $a_2(t)$ , are analytic in  $t$  at  $t_0$  as well. Indeed, from (5.12) and (5.17) we have that

$$\begin{cases} a_1 + b_1 + a_2 = -b_2, \\ a_1 b_1 + a_1 a_2 + b_1 a_2 = -2t - b_2(a_1 + b_1 + a_2) = -2t + b_2^2, \\ a_1 b_1 a_2 = -4 - b_2(a_1 b_1 + a_1 a_2 + b_1 a_2) = -4 - b_2(-2t + b_2^2) = -4 + 2tb_2 - b_2^3. \end{cases}$$

Hence  $a_1(t)$ ,  $b_1(t)$ ,  $a_2(t)$  are roots of the cubic equation

$$z^3 + b_2 z^2 + (-2t + b_2^2)z + (4 - 2tb_2 + b_2^3) = 0.$$

Since  $a_1(t)$ ,  $b_1(t)$ ,  $a_2(t)$  are pairwise different, they analytically depend on the coefficients of the cubic equation. That is, they are analytic functions of  $t$  and  $b_2$ . This implies that  $a_1(t)$ ,  $b_1(t)$ , and  $a_2(t)$  are analytic in  $t$  at  $t_0$ . The analyticity of  $a_1(t)$ ,  $b_1(t)$ ,  $a_2(t)$ , and  $b_2(t)$  yields that the integral

$$\frac{1}{\pi i} \int_{\Gamma_t} \int_{[a_1(t), b_1(t)]} Q_+^{1/2}(s; t) ds$$

is also an analytic function of  $t$  at  $t_0$ . However, the integral above is equal to  $-\omega(t)$ , see (4.6), which is a real number, see (2.7). Hence,  $\omega(t)$  must be constant in  $U$ . On the other hand, it follows from (3.6), the proof of which in the next subsection is independent of the current considerations, that  $\omega(t)$  is not constant in  $\mathcal{O}_{\text{two-cut}}$ . We also have already shown that  $\omega(t)$  is real analytic in  $\operatorname{Re}(t)$  and  $\operatorname{Im}(t)$  and therefore cannot be locally constant. This contradiction proves that the endpoints  $a_1(t)$ ,  $b_1(t)$ ,  $a_2(t)$ , and  $b_2(t)$  are not analytic functions of  $t$  at any point in  $\mathcal{O}_{\text{two-cut}}$ .

**5.5. Proof of (3.6).** It follows from [3, Theorem 5.11] that if  $t$  remains in a bounded set, so do the zeros of  $Q(z; t)$ . Fix  $t^* \in \partial O_{\text{two-cut}}$  and let  $\{t_m\}$  be a sequence such that  $t_m \rightarrow t^*$  as  $m \rightarrow \infty$ . Restricting to a subsequence if necessary, we see that there exist  $a_i^*, b_i^*$  such that  $e(t_m) \rightarrow e^*$  as  $m \rightarrow \infty$ , where  $e \in \{a_1, b_1, a_2, b_2\}$ . Clearly, polynomials  $Q(z; t_m)$  converge uniformly on compact subsets of  $\mathbb{C}$  to  $Q^*(z)$ , the polynomial with zeros  $a_i^*, b_i^*$  and leading coefficient  $1/4$ . Let  $\varpi^*(z) := -Q^*(z)dz^2$ . Then repeating the argument after (5.19), we get that the critical graphs of  $\varpi_{t_m}(z)$  converge to the critical graph of  $\varpi^*(z)$  in any disk  $|z| < R$ . Furthermore, as stated in [26, Theorem 3.3], there exists  $R_t > 0$  such that every trajectory of  $\varpi_t(z)$  entering  $\{|z| > R_t\}$  necessarily remains in  $\{|z| > R_t\}$  and tends to infinity. Moreover, examination of the proof of [26, Theorem 3.3] also shows that one can take  $R_t = 2^{-1/2} \max_i \{|a_i(t)|, |b_i(t)|\}$  and that if a trajectory enters  $S_{(2k-1)\pi/3, \delta} \cap \{|z| > R_t\}$ ,  $k \in \{0, 1, 2\}$ , it stays in this sector, see (5.10). Put  $R^* := \sup_m R_{t_m}$ . As all the differentials  $\varpi_{t_m}(z)$  have structurally “the same” critical graph, see Figure 6, and all the trajectories of  $\varpi_{t_m}(z)$  and  $\varpi^*(z)$  entering  $S_{(2k-1)\pi/3, \delta} \cap \{|z| > R^*\}$  must remain there, tend to infinity, and not intersect, the unbounded critical trajectories of  $\varpi^*(z)$  behave like on Figure 6.

Observe that the first and the third equations of (5.17), see also (5.12), must remain true for  $a_1^*, b_1^*, a_2^*, b_2^*$  as well. Hence, we cannot simultaneously have that  $a_1^* = b_1^*$  and  $a_2^* = b_2^*$ . As short trajectories cannot form loops and their tangent vectors cannot become parallel at each critical point, the above considerations yield that the critical and therefore critical orthogonal graphs of  $\varpi^*(t)$  look like either on Figure 6, or on Figure 5, or as the graphs obtained by reflection across the line  $L_{2\pi/3}$  of the graphs on Figure 5. Repeating the arguments at the end of Section 5.4, we see that  $\varpi^*(t)$  gives rise to an S-contour in  $\mathcal{T}$ . The limits in (3.6) now follows from the uniqueness of such a S-contour and Theorem 3.1 (in particular, the critical graph  $\varpi_*(t)$  cannot look like as on Figure 6).

## 6. FUNCTIONS $\mathcal{D}(z; t)$ AND $Q(z; t)$

In this section we prove Propositions 4.1 and 4.2 as well as discuss other properties of  $Q(z; t)$ . We consider the parameter  $t \in O_{\text{two-cut}}$  to be fixed and stop indicating the dependence on  $t$  of the various quantities appearing below whenever this does not introduce ambiguity and is convenient.

**6.1. Proof of Proposition 4.1.** It follows from (1.11), (2.6), and the choice of the branch of  $Q^{1/2}(z)$  that

$$(6.1) \quad Q^{1/2}(z) = \frac{z^2 - t}{2} + \frac{1}{z} + \frac{\mu_1}{z^2} + O\left(\frac{1}{z^3}\right).$$

It further follows from the choice of the constant  $\varsigma, d_1$  in (4.1) that

$$z + \int_{I_t} \frac{3\varsigma}{s - z} \frac{ds}{Q^{1/2}(s)} = z - \frac{2t}{z} - \frac{3\varsigma d_1}{z^2} - \frac{3\varsigma d_2}{z^3} + O\left(\frac{1}{z^4}\right)$$

as  $z \rightarrow \infty$ . Thus, the analyticity properties of  $\mathcal{D}(z)$  as well as (4.4) now follow from the fact that the product of the above functions behaves like

$$-\frac{3V(z)}{2} + 1 - \frac{3\varsigma d_1}{2} + \frac{t^2 + \mu_1 - 3\varsigma d_2/2}{z} + O\left(\frac{1}{z^2}\right)$$

as  $z \rightarrow \infty$ . Moreover, since  $Q_+^{1/2}(s) = -Q_-^{1/2}(s)$  for  $s \in J_t$ , we get the first relation in (4.3). The second relation in (4.3) follows from Plemelj-Sokhotski formula

$$\left( \int_{I_t} \frac{\varsigma}{s - z} \frac{ds}{Q^{1/2}(s)} \right)_+ - \left( \int_{I_t} \frac{\varsigma}{s - z} \frac{ds}{Q^{1/2}(s)} \right)_- = \frac{2\pi i \varsigma}{Q^{1/2}(z)}, \quad z \in I_t^\circ.$$

**6.2. Proof of Proposition 4.2.** Since the arc  $I_t$  is homologous to the short critical trajectory of  $-Q(z)dz^2$  connecting  $b_1$  and  $a_2$  and  $J_{t,1} = \Gamma[a_1, b_1]$  is such a trajectory, see Figure 6 and Conventions 1 and 2, the constants  $\tau, \omega$  are indeed real. Let

$$(6.2) \quad g(z) := \int \log(z - s) d\mu_t(s), \quad z \in \mathbb{C} \setminus \Gamma(e^{\pi i} \infty, b_2],$$



where we take the principal branch of  $\log(\cdot - s)$  holomorphic outside of  $\Gamma(e^{\pi i \infty}, s]$  and  $\mu_t$  is the equilibrium measure defined in (2.7). It follows directly from definition (6.2) that

$$\partial_z g(z) = \int \frac{d\mu_t(s)}{z - s},$$

where, as usual,  $\partial_z := (\partial_x - i\partial_y)/2$ . Therefore, it can be deduced from (2.2) and (2.6) that

$$(6.3) \quad g(z) = \frac{V(z) - \ell_*}{2} + \int_{b_2}^z Q^{1/2}(s) ds = \frac{V(z) - \ell_*}{2} + \mathcal{Q}(z),$$

where, as usual, we take the branch  $Q^{1/2}(z) = \frac{1}{2}z^2 + \mathcal{O}(z)$ ,  $\ell_*$  is a constant such that the equality holds at  $b_2$  (notice that  $\text{Re}(\ell_*) = \ell$ , see (2.2)), and  $\mathcal{Q}(z)$  is given by (4.5). Property (4.7) clearly follows from (6.2) and (6.3). In the view of (6.3), let us define

$$(6.4) \quad \phi_e(z) := 2 \int_e^z Q^{1/2}(s) ds, \quad e \in \{a_1, b_1, a_2, b_2\},$$

holomorphically in  $\mathbb{C} \setminus \Gamma(e^{\pi i \infty}, b_2]$  when  $e = b_2$ , in  $\mathbb{C} \setminus \Gamma[a_1, e^{\pi i/3 \infty})$  when  $e = a_1$ , and in  $\mathbb{C} \setminus (\Gamma(e^{\pi i \infty}, b_1) \cup \Gamma(a_2, e^{\pi i/3 \infty}))$  when  $e \in \{b_1, a_2\}$ . Clearly,  $\phi_{b_2}(z) = 2\mathcal{Q}(z)$ . One can readily check that

$$(6.5) \quad \phi_{b_2}(z) = \begin{cases} \phi_{a_2}(z) \pm 2\pi i(1 - \omega), \\ \phi_{b_1}(z) \pm 2\pi i(1 - \omega) + 2\pi i\tau, & z \in \mathbb{C} \setminus \Gamma, \\ \phi_{a_1}(z) \pm 2\pi i + 2\pi i\tau, \end{cases}$$

where the plus sign is used if  $z$  lies to the left of  $\Gamma$  and the minus sign if  $z$  lies to the right of it, and

$$(6.6) \quad \phi_{b_2\pm}(s) = \begin{cases} \pm 2\pi i \mu_t(\Gamma[s, b_2]), & s \in \Gamma(a_2, b_2), \\ \pm 2\pi i \mu_t(\Gamma[s, b_2]) + 2\pi i\tau, & s \in \Gamma(a_1, b_1). \end{cases}$$

The jump relations in (4.8) now easily follow from (6.5) and (6.6).

For future use let us record that (6.3), (6.5), and (6.6) imply that

$$(6.7) \quad g_+(s) - g_-(s) = \begin{cases} 0, & s \in \Gamma(b_2, e^{\pi i/3 \infty}), \\ \pm \phi_{b_2\pm}(s), & s \in \Gamma(a_2, b_2), \\ 2\pi i(1 - \omega), & s \in \Gamma(b_1, a_2), \\ \pm(\phi_{b_2\pm}(s) - 2\pi i\tau), & s \in \Gamma(a_1, b_1), \\ 2\pi i, & s \in \Gamma(e^{\pi i \infty}, a_1), \end{cases}$$

and that

$$(6.8) \quad g_+(s) + g_-(s) - V(s) + \ell_* = \begin{cases} \phi_{b_2}(s), & s \in \Gamma(b_2, e^{\pi i/3 \infty}), \\ 0, & s \in \Gamma(a_2, b_2), \\ \phi_{a_2}(s), & s \in \Gamma(b_1, a_2), \\ 2\pi i\tau, & s \in \Gamma(a_1, b_1), \\ \phi_{a_1}(s) + 2\pi i\tau, & s \in \Gamma(e^{\pi i \infty}, a_1). \end{cases}$$

**6.3. Local Analysis at  $e \in \{a_1, b_1, a_2, b_2\}$ .** Given  $e \in \{a_1, b_1, a_2, b_2\}$ , let

$$(6.9) \quad U_e := \{z : |z - e| < \delta_e \rho(t)/3\},$$

where  $\delta_e \in (0, 1]$  to be adjusted later and we shall specify the function  $\rho(t)$  at the end of this subsection. Set

$$(6.10) \quad J_e := U_e \cap J_t \quad \text{and} \quad I_e := U_e \cap (\Gamma_t \setminus J_t),$$

where the arcs  $J_e$  and  $I_e$  inherit their orientation from  $\Gamma_t$  and we assume that the value of  $\rho(t)$  is small enough so that these arcs are connected and that  $I_e$  is a subarc of the orthogonal critical trajectory

of  $-Q(z)dz^2$  emanating from  $e$  (see the remarks about  $\Gamma_t$  at the very beginning of Section 4). The latter fact and Theorem 3.2 yield that

$$(6.11) \quad \phi_e(s) < 0, \quad s \in I_e,$$

see Figure 6. In fact, the same reasoning shows that (6.11) holds not only on  $I_e$ , but on  $\Gamma(e^{\pi i \infty}, a_1)$  when  $e = a_1$ , on  $\Gamma(b_2, e^{\pi i/3 \infty})$  when  $e = b_2$ , and, for  $\operatorname{Re}(\phi_e(z))$  on  $\Gamma(b_1, a_2)$  when  $e \in \{b_1, a_2\}$  (observe that these functions are also monotone on the respective arcs). Furthermore, each function  $\phi_e(z)$  is analytic in  $U_e \setminus J_e$  and its traces on  $J_e$  satisfy

$$(6.12) \quad \phi_{e\pm}(s) = \pm 2\pi i v_e \mu(J_{s,e}) = 2\pi e^{\pm 3\pi i v_e/2} \mu(J_{s,e}),$$

where  $J_{s,e}$  is the subarc of  $J_e$  with endpoints  $e$  and  $s$ ,

$$(6.13) \quad v_e := \begin{cases} 1, & e \in \{b_1, b_2\}, \\ -1, & e \in \{a_1, a_2\}, \end{cases}$$

and the second equality follows from (2.7) and (6.4). Since  $|\phi_e(z)| \sim |z - e|^{3/2}$  as  $z \rightarrow e$ , it follows from (6.11) and (6.12) that we can define an analytic branch of  $(-\phi_e)^{2/3}(z)$  in  $U_e$  that is positive on  $I_e$  and satisfies  $(-\phi_e)^{2/3}(s) = -(2\pi\mu_t(J_{s,e}))^{2/3}$ ,  $s \in J_e$ . Since  $(-\phi_e)^{2/3}(z)$  has a simple zero at  $e$ , it is conformal in  $U_e$  for all radii small enough. Altogether,  $(-\phi_e)^{2/3}(z)$  maps  $e$  into the origin, is conformal in  $U_e$ , and satisfies

$$(6.14) \quad \begin{cases} (-\phi_e)^{2/3}(J_e) \subset (-\infty, 0), \\ (-\phi_e)^{2/3}(I_e) \subset (0, \infty). \end{cases}$$

Furthermore, if we define  $(-\phi_e)^{1/6}(z)$  to be holomorphic in  $U_e \setminus J_e$  and positive on  $I_e$ , then

$$(6.15) \quad (-\phi_e)_+^{1/6}(s) = v_e i (-\phi_e)_-^{1/6}(s), \quad s \in J_e.$$

To specify  $\rho(t)$ , let  $\rho_e(t)$  be the radius of the largest disk around  $e$  for which  $J_e, I_e$  are connected and in which  $(-\phi_e)^{2/3}(z)$  is conformal. Observe that the disk around  $e$  of radius  $\rho_e(t)$  cannot contain other endpoints of  $J_t$  besides  $e$ . We set  $\rho(t) := \min_e \{\rho_e(t)\}$ . Then the disks  $U_e$  in (6.9) are necessarily disjoint. Observe also that  $\rho(t)$  is non-zero for all  $t \in \mathcal{O}_{\text{two-cut}}$  and continuously depends on  $t$  due to continuous dependence on  $t$  of  $\phi_e(z)$ , which in itself follows from Theorem 3.2 and (6.4).

## 7. FUNCTIONS $\Theta_n(z; t)$

In this section we prove Proposition 4.3–4.6 as well as discuss some related results. As in the previous section, we omit indicating the explicit dependence on  $t$  whenever convenient.

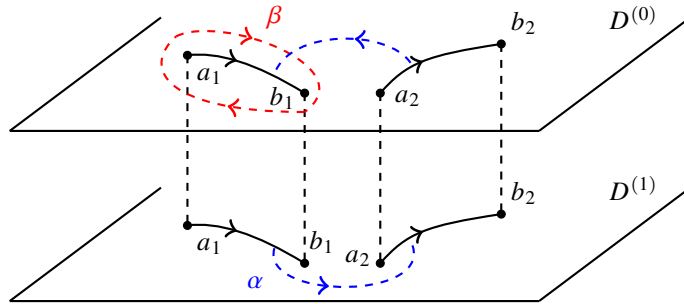


FIGURE 12. Schematic plot of the Riemann surface  $\mathfrak{S}$  and the cycles  $\alpha$  and  $\beta$ .

**7.1. Riemann Surface.** In this subsection we discuss properties of the Riemann surface  $\mathfrak{S}$  that has already appeared between (5.15) and (5.16). Once again, set

$$(7.1) \quad \mathfrak{S} := \{z := (z, w) : w^2 = Q(z)\}.$$

We denote by  $\pi : \mathfrak{S} \rightarrow \overline{\mathbb{C}}$  the natural projection  $\pi(z) = z$  and by  $\cdot^*$  a holomorphic involution on  $\mathfrak{S}$  acting according to the rule  $z^* = (z, -w)$ . We use notation  $z, s, a$  for points on  $\mathfrak{S}$  with natural projections  $z, s, a$ .

The function  $w(z)$ , defined by  $w^2(z) = Q(z)$ , is meromorphic on  $\mathfrak{S}$  with simple zeros at the ramification points  $E_t = \{a_1, b_1, a_2, b_2\} = \pi^{-1}(E_t)$ , double poles at the points on top of infinity, and is otherwise non-vanishing and finite. Put

$$\Delta := \pi^{-1}(J_t) \quad \text{and} \quad \mathfrak{S} = D^{(0)} \cup \Delta \cup D^{(1)},$$

where the domains  $D^{(k)}$  project onto  $\overline{\mathbb{C}} \setminus J_t$  with labels chosen so that  $2w(z) = (-1)^k z^2 + \mathcal{O}(z)$  as  $z$  approaches the point on top of infinity within  $D^{(k)}$ . For  $z \in \overline{\mathbb{C}} \setminus J_t$  we let  $z^{(k)}$  stand for  $z \in D^{(k)}$  with, as agreed,  $\pi(z) = z$ . We define a homology basis on  $\mathfrak{S}$  in the following way: we let

$$\alpha := \pi^{-1}(I_t) \quad \text{and} \quad \beta := \pi^{-1}(J_{t,1}),$$

where  $\alpha$  is oriented towards  $b_1$  within  $D^{(0)}$  and  $\beta$  is oriented so that  $\alpha, \beta$  form the right pair at  $b_1$ , see Figures 12 and 13. It also will be convenient to put  $\mathfrak{S}_{\alpha, \beta} := \mathfrak{S} \setminus \{\alpha \cup \beta\}$ , which is simply connected.

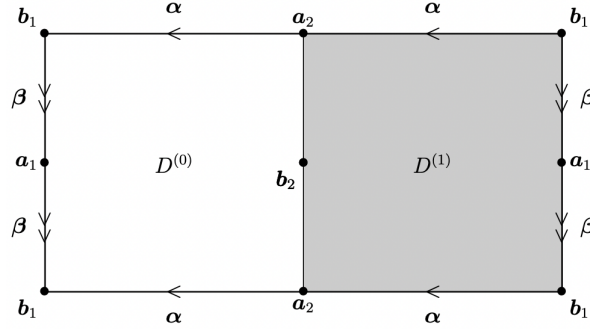


FIGURE 13. Planar model of the Riemann surface  $\mathfrak{S}$  and the cycles  $\alpha$  and  $\beta$ .

The surface  $\mathfrak{S}$  has genus one. Thus, there exists a unique holomorphic differential on  $\mathfrak{S}$  normalized to have a unit period on  $\alpha$ , say  $\mathcal{H}$ . In fact, it can be explicitly expressed as

$$(7.2) \quad \mathcal{H}(z) = \left( \oint_{\alpha} \frac{ds}{w(s)} \right)^{-1} \frac{dz}{w(z)}.$$

With this notation, the choice of the homology basis, and the definition of  $w(z)$ , (4.9) becomes

$$(7.3) \quad \mathbf{B} = \oint_{\beta} \mathcal{H}, \quad \text{Im}(\mathbf{B}) > 0,$$

where the last inequality is a classical result of Riemann, see [41].

Given the normalized holomorphic differential, we define

$$(7.4) \quad \mathbf{a}(z) := \int_{b_2}^z \mathcal{H},$$

where we restrict  $z$  as well as the path of integration to  $\mathfrak{S}_{\alpha, \beta}$ . The function  $\mathbf{a}(z)$  is a holomorphic function in  $\mathfrak{S}_{\alpha, \beta}$  with continuous traces on  $\alpha, \beta$  (away from the point of their intersection) that satisfy

$$(7.5) \quad \mathbf{a}_+(s) - \mathbf{a}_-(s) = \begin{cases} -\mathbf{B}, & s \in \alpha \setminus \{b_1\}, \\ 1, & s \in \beta \setminus \{b_1\}, \end{cases}$$

by the normalization of  $\mathcal{H}$  and the definition of  $\mathbf{B}$ . Moreover, it continuously extends to  $\partial\mathfrak{S}_{\alpha,\beta}$ , the topological boundary of  $\mathfrak{S}_{\alpha,\beta}$  (the rectangle on Figure 13). Notice that

$$(7.6) \quad \tau = \frac{1}{2\pi i} \oint_{\alpha} w(s) ds \quad \text{and} \quad \omega = -\frac{1}{2\pi i} \oint_{\beta} w(s) ds,$$

where  $\tau, \omega$  are the constants from (4.6). It readily follows from (7.5) and (7.6) that

$$\oint_{\partial\mathfrak{S}_{\alpha,\beta}} (w\mathbf{a})(s) ds = \oint_{\beta} w(s)(\mathbf{a}_+ - \mathbf{a}_-)(s) ds + \oint_{\alpha} w(s)(\mathbf{a}_+ - \mathbf{a}_-)(s) ds = -2\pi i(\omega + \mathbf{B}\tau),$$

where  $\partial\mathfrak{S}_{\alpha,\beta}$  is oriented so that  $\mathfrak{S}_{\alpha,\beta}$  remains on the left when  $\partial\mathfrak{S}_{\alpha,\beta}$  is traversed in the positive direction. On the other hand, the function  $(w\mathbf{a})(z)$  is meromorphic in  $\mathfrak{S}_{\alpha,\beta}$  with only two singularities, both polar, at  $\infty^{(0)}$  and  $\infty^{(1)}$ . Moreover, since  $w(z^*) = -w(z)$  and  $\mathbf{a}(z^*) = -\mathbf{a}(z)$ , the residues at those poles coincide. Therefore, it holds that

$$\omega + \mathbf{B}\tau = -\frac{1}{2\pi i} \oint_{\partial\mathfrak{S}_{\alpha,\beta}} (w\mathbf{a})(s) ds = -2\text{res}_{z=\infty^{(0)}}(w\mathbf{a})(z).$$

Recall that  $w(z^{(0)}) = \mathcal{Q}^{1/2}(z)$  and therefore it has expansion (6.1) as  $z^{(0)} \rightarrow \infty^{(0)}$ . This and (7.4) also allow us to deduce that

$$\mathbf{a}(z^{(0)}) = \mathbf{a}(\infty^{(0)}) + \int_{\infty^{(0)}}^{z^{(0)}} \mathcal{H} = \mathbf{a}(\infty^{(0)}) - \left( \oint_{\alpha} \frac{ds}{w(s)} \right)^{-1} \left( \frac{2}{z} + \frac{2t}{3z^3} + \mathcal{O}\left(\frac{1}{z^4}\right) \right)$$

as  $z^{(0)} \rightarrow \infty^{(0)}$ . Therefore, taking into account (4.1) with the definition of  $\alpha$  and the symmetry  $\mathbf{a}(z^*) = -\mathbf{a}(z)$ , we get that

$$(7.7) \quad \zeta + \omega + \mathbf{B}\tau = \zeta + 2\mathbf{a}(\infty^{(0)}) + \frac{4t}{3} \left( \oint_{\alpha} \frac{ds}{w(s)} \right)^{-1} = 2 \int_{b_2}^{\infty^{(0)}} \mathcal{H} = \int_{\infty^{(1)}}^{\infty^{(0)}} \mathcal{H},$$

where the path of integration lies entirely in  $\mathfrak{S}_{\alpha,\beta}$  and is involution-symmetric.

**7.2. Jacobi Inversion Problem.** Denote by  $\text{Jac}(\mathfrak{S}) := \mathbb{C}/\{\mathbb{Z} + \mathbf{B}\mathbb{Z}\}$  the Jacobi variety of  $\mathfrak{S}$ . We shall represent elements of  $\text{Jac}(\mathfrak{S})$  as equivalence classes  $[s] = \{s + j + m\mathbf{B} : j, m \in \mathbb{Z}\}$ , where  $s \in \mathbb{C}$ . One can readily deduce from (7.5) that  $\mathbf{a}(z)$ , defined in (7.4), is essentially a holomorphic bijection from  $\mathfrak{S}$  onto  $\text{Jac}(\mathfrak{S})$ . Formally, we define this bijection, known as Abel's map, by

$$(7.8) \quad z \in \mathfrak{S} \mapsto \left[ \int_{b_2}^z \mathcal{H} \right] \in \text{Jac}(\mathfrak{S}).$$

Given  $s \in \mathbb{C}$ , we denote by  $z_{[s]}$  the unique point in  $\mathfrak{S}$  such that  $\left[ \int_{b_2}^{z_{[s]}} \mathcal{H} \right] = [s]$ .

**Proposition 7.1.** *Let  $\tau, \omega$  be given by (4.6) and  $\zeta$  by (4.1). Further, let  $\{N_n\}$  be a sequence as in Theorem 4.1. Denote by  $z_{n,k} = z_{n,k}(t)$  the unique solution  $z_{[s_{n,k}(t)]}$  of the Jacobi inversion problem with*

$$(7.9) \quad s_{n,k}(t) := \int_{b_2}^{p^{(k)}} \mathcal{H} + (n - N_n)\zeta + (\omega + \mathbf{B}\tau)n, \quad p = p(t) := \frac{b_1 b_2 - a_1 a_2}{(b_2 - a_2) + (b_1 - a_1)},$$

$k \in \{0, 1\}$ . Then for any subsequence  $\mathbb{N}_*$  the point  $\infty^{(0)}$  is a topological limit point of  $\{z_{n,1}\}_{n \in \mathbb{N}_*}$  if and only if  $\infty^{(1)}$  is a topological limit point of  $\{z_{n,0}\}_{n \in \mathbb{N}_*}$ . Moreover, it holds that  $[s_{n,0}(t; N_n)] = [s_{n-1,1}(t; N_n)]$ ; that is,  $z_{n,0}(N_n) = z_{n-1,1}(N_n)$ .

The behavior of the points  $z_{n,k}$  with respect to  $n$  can be extremely chaotic. Assuming that  $n - N_n$  is constant, it is known that if the numbers  $\omega$  and  $\tau$  are rational, then there exist only finitely many distinct points  $z_{n,k}$ ; when  $\omega$  and  $\tau$  are irrational, all the points  $z_{n,k}$  are distinct, lie on a Jordan curve if 1,  $\omega$  and  $\tau$  are rationally dependent, and are dense on the whole surface  $\mathfrak{S}$  otherwise [38].

Let us also point out that the point  $p$  is always finite. Indeed, it follows from (1.11) and (2.6) that  $b_1 + b_2 + a_1 + a_2 = 0$ , see also (5.12) and (5.17). Thus, if  $b_2 + b_1 - a_1 - a_2 = 0$  were true, then we would have  $b_1 = -b_2$  and  $a_1 = -a_2$ . In this case  $a_1 b_1 a_2 + a_1 b_1 b_2 + a_1 a_2 b_2 + b_1 a_2 b_2 = 0$ . However, this expression must be equal to  $-4$  according to the above references.

*Proof of Proposition 7.1.* Define

$$(7.10) \quad \gamma(z) := \left( \frac{z - b_2}{z - a_2} \frac{z - b_1}{z - a_1} \right)^{1/4}, \quad z \in \overline{\mathbb{C}} \setminus J_t,$$

where  $\gamma(z)$  is holomorphic off  $J_t = \text{supp}(\mu_t)$  and the branch is chosen so that  $\gamma(\infty) = 1$ . Further, set

$$(7.11) \quad A(z) = \frac{\gamma(z) + \gamma^{-1}(z)}{2} \quad \text{and} \quad B(z) := \frac{\gamma(z) - \gamma^{-1}(z)}{-2i}.$$

Observe that the function  $A(z)$  was already defined in (4.12). The functions  $A(z)$  and  $B(z)$  are holomorphic in  $\overline{\mathbb{C}} \setminus J_t$  and satisfy

$$(7.12) \quad A(\infty) = 1, \quad B(\infty) = 0, \quad \text{and} \quad A_{\pm}(s) = \pm B_{\mp}(s), \quad s \in J_t^{\circ}.$$

Notice that the equation  $(AB)(z) = 0$  can be rewritten as  $\gamma^4(z) = 1$  and has two solutions, namely,  $\infty$  and the point  $p$  from (7.9). In fact, unless  $p \in J_t^{\circ}$ , it is a zero of  $B(z)$ . Indeed, it is enough to show that  $\gamma(p) = 1$  in the latter case. Let  $L_i := \gamma^4(J_{t,i})$ ,  $i \in \{1, 2\}$ , which are unbounded arcs connecting the origin to the point at infinity. Let  $L \subset \overline{\mathbb{C}} \setminus J_t$  be an arc connecting the point at infinity and  $p$ . Then  $\gamma^4(L)$  is a closed curve that contains 1 and does not intersect the arcs  $L_i$  and therefore does not wind around the origin. Thus, analytic continuation of the principal branch of the 1/4-root from 1 along  $\gamma^4(L)$  leads back to the value 1 at the point 1. However, this continuation is exactly the continuation of  $\gamma(z)$  from the point at infinity to  $p$  along  $L$ , which does imply that  $\gamma(p) = 1$  as claimed.

It follows from (7.12) that

$$(7.13) \quad \begin{cases} (B/A)(z), & z \in D^{(0)}, \\ -(A/B)(z), & z \in D^{(1)}, \end{cases}$$

is a rational function on  $\mathfrak{S}$  with two simple zeros  $\infty^{(0)}$  and  $p^{(0)}$  and two simple poles  $\infty^{(1)}$  and  $p^{(1)}$  (if it happens that  $p \in J_t^{\circ}$ , then we choose  $p^{(0)} \in \mathfrak{S}$  precisely in such a way that it is a zero of (7.13) and so  $p^{(1)}$  is a pole of (7.13); it is, of course, still true that these points are distinct and  $\pi(p^{(k)}) = p$ ). Therefore, Abel's theorem yields that

$$(7.14) \quad \left[ \int_{p^{(0)}}^{\infty^{(1)}} \mathcal{H} \right] = \left[ \int_{p^{(1)}}^{\infty^{(0)}} \mathcal{H} \right]$$

while the relations (7.9), in particular, imply that

$$(7.15) \quad \left[ \int_{p^{(0)}}^{z_{n,0}} \mathcal{H} \right] = \left[ \int_{p^{(1)}}^{z_{n,1}} \mathcal{H} \right].$$

Let  $z_k$  be a topological limit of a subsequence  $\{z_{n_i,k}\}$ . Holomorphy of the differential  $\mathcal{H}$  implies that

$$\int_{p^{(k)}}^{z_{n_i,k}} \mathcal{H} = \int_{p^{(k)}}^{z_k} \mathcal{H} + \int_{z_k}^{z_{n_i,k}} \mathcal{H} \rightarrow \int_{p^{(k)}}^{z_k} \mathcal{H}$$

as  $i \rightarrow \infty$ , where the integral from  $z_k$  to  $z_{n_i,k}$  is taken along the path that projects into a segment joining  $z_k$  and  $z_{n_i,k}$ . The first claim of the proposition now follows from (7.14), (7.15), and the unique solvability of the Jacobi inversion problem on  $\mathfrak{S}$ . Observe that

$$(7.16) \quad \left[ \int_{p^{(0)}}^{p^{(1)}} \mathcal{H} \right] = \left[ \int_{\infty^{(1)}}^{\infty^{(0)}} \mathcal{H} \right] = [\mathcal{S} + \omega + \mathbf{B}\tau].$$

by (7.14) and (7.7). Since

$$s_{n,0}(N_n) = s_{n-1,1}(N_n) + \int_{p^{(1)}}^{p^{(0)}} \mathcal{H} + \zeta + \omega + \mathbf{B}\tau$$

the second conclusion of the proposition easily follows.  $\square$

**7.3. Proof of Proposition 4.5.** As we shall show further below, the functions  $\Theta_n(z; t)$  from Proposition 4.4 vanish at  $z_{n,1}$  when it belongs to  $D^{(0)}$  and do not vanish at all when  $z_{n,1}$  does not belong to  $D^{(0)}$ . Hence, the subsequences  $\mathbb{N}(\varepsilon) = \mathbb{N}(t, \varepsilon)$  from Proposition 4.4 can be equivalently defined as

$$\mathbb{N}(\varepsilon) := \left\{ n \in \mathbb{N} : z_{n,1} \notin D^{(0)} \cap \pi^{-1}(\{|z| \geq 1/\varepsilon\}) \right\},$$

Denote by  $r[s]$  the distance from 0 to  $[s]$ , viewed as a lattice in  $\mathbb{C}$ . That is,

$$r[s] := \min_{j, m \in \mathbb{Z}} |s + j + \mathbf{B}m|,$$

where  $s$  is any representative of  $[s]$ . Further, for each  $r[s] > 0$ , choose  $\varepsilon_s > 0$  small enough so that

$$|\mathbf{a}(z) - \mathbf{a}(\infty^{(0)})| < r[s]/3, \quad z \in U_{\varepsilon_s}, \quad U_{\varepsilon_s} := D^{(0)} \cap \pi^{-1}(\{|z| > 1/\varepsilon_s\}),$$

where we also assume that  $\varepsilon_s$  is small enough so that always  $U_{\varepsilon_s} \subset \mathfrak{S}_{\alpha, \beta}$ . Then

$$(7.17) \quad [\mathbf{a}(z_1) - \mathbf{a}(z_2)] = [s] \neq [0] \quad \Rightarrow \quad \text{either } z_1 \notin U_{\varepsilon_s} \text{ or } z_2 \notin U_{\varepsilon_s}$$

as otherwise it must hold  $0 < r[s] \leq |\mathbf{a}(z_1) - \mathbf{a}(z_2)| < (2/3)r[s]$ .

Assume that either  $x$  or  $y$  is not an integer. Let  $z_{n,1}^-, z_{n,1}, z_{n,1}^+$  be solutions of Jacobi inversion problem (7.9) with  $N$  equal to  $n-1, n, n+1$ , respectively. Then it follows from (7.9) that

$$[\mathbf{a}(z_{n,1}^\pm) - \mathbf{a}(z_{n,1})] = [\mp \zeta].$$

As  $[\pm \zeta] \neq [0]$  by assumption, the second assertion of the proposition follows from (7.17).

Assume now that both  $x$  and  $y$  are integers. It readily follows from (7.9) that the definition of  $z_{n,k}$  does not depend on the choice of  $N_n$  in this case. The first assertion of the proposition holds since

$$[\mathbf{a}(z_{n+1,1}) - \mathbf{a}(z_{n,1})] = [\omega + \mathbf{B}\tau] \neq [0],$$

where the last conclusion follows from the fact  $\omega = \mu_t(J_{t,1}) \in (0, 1)$  by the very definition in (4.6).

**7.4. Proof of Proposition 4.6.** Let  $k = N_{n+1} - N_n$ . It follows from (7.9) that

$$(7.18) \quad [\mathbf{a}(z_{n+1,1}) - \mathbf{a}(z_{n,1})] = [(1-k)\zeta + \omega + \mathbf{B}\tau].$$

Set  $\mathbb{K} := \{k \in \mathbb{Z} : r[(1-k)\zeta + \omega + \mathbf{B}\tau] = 0\}$ . We have already shown that  $r[\omega + \mathbf{B}\tau] > 0$ . It also holds that  $r[\zeta + \omega + \mathbf{B}\tau] > 0$  since  $[\zeta + \omega + \mathbf{B}\tau] \neq [0]$  by (7.7) and the unique solvability of the Jacobi inversion problem. The following lemma easily follows from (7.17).

**Lemma 7.2.** *Assume that  $|n - N_n| \leq N_*$  and let  $\varepsilon_k := \varepsilon_{(1-k)\zeta + \omega + \mathbf{B}\tau}$ .*

- (1) *If  $\mathbb{K} = \emptyset$ , then at least one of the integers  $n, n+1$  belongs to  $\mathbb{N}(\varepsilon)$  for all  $\varepsilon \leq \min_{|k| \leq 2N_*+1} \varepsilon_k$ .*
- (2) *If  $\mathbb{K} = \{k'\}$ , there exists an infinite subsequence  $\{n_l\}$  such that  $N_{n_l+1} - N_{n_l} \neq k'$  and therefore at least one of the integers  $n_l, n_l+1$  belongs to  $\mathbb{N}(\varepsilon)$  for all  $\varepsilon \leq \min_{|k| \leq 2N_*+1, k \neq k'} \varepsilon_k$ .*
- (3) *If there exists an infinite subsequence  $\{n_l\}$  such that  $N_{n_l+1} - N_{n_l} \in \{0, 1\}$ , then at least one of the integers  $n_l, n_l+1$  belongs to  $\mathbb{N}(\varepsilon)$  for all  $\varepsilon \leq \min\{\varepsilon_0, \varepsilon_1\}$ .*

Inclusion  $k \in \mathbb{K}$  means that  $[(1-k)\zeta + \omega + \mathbf{B}\tau] = [0]$ . Hence, both triples  $\omega, x, 1$  and  $\tau, y, 1$  are rationally dependent,  $\zeta = x + \mathbf{B}y$ . Assume that  $k', k'' \in \mathbb{K}, k' \neq k''$ . Then it follows from (7.18) that  $[\omega + \mathbf{B}\tau] = [(k' - 1)\zeta]$  and  $[(k'' - k')\zeta] = [0]$ . The latter relation implies the first representation in (4.13) while the former gives the other two. It is easy to see in this case that  $\mathbb{K} = k' + d\mathbb{Z}$ . That is, if  $\mathbb{K}$  has at least two elements, then it is an arithmetic progression,  $\omega, \tau$  are rational numbers,  $\zeta$  has rational coordinates in the basis  $1, \mathbf{B}$ , and the second and third relations of (4.13) must be satisfied. Thus, we can claim the following.

**Lemma 7.3.** *If one of the triples  $\omega, x, 1$  or  $\tau, y, 1$  is rationally independent, then  $\mathbb{K} = \emptyset$ . If not all numbers  $\omega, \tau, x, y$  are rational or they all rational but the second and third relations of (4.13) do not hold, then either  $\mathbb{K} = \emptyset$  or  $\mathbb{K} = \{k'\}$ .*

Assume now that all three relations of (4.13) take place. That is,  $[\omega + \mathbf{B}\tau] = [(k-1)\zeta]$  and  $[d\zeta] = [0]$  for some integers  $k, d$ . It follows from (7.14) that

$$\int_{\mathbf{b}_2}^{p^{(1)}} \mathcal{H} = \frac{1}{2} \int_{p^{(0)}}^{p^{(1)}} \mathcal{H} = \frac{1}{2} \left( \int_{\infty^{(1)}}^{\infty^{(0)}} \mathcal{H} + j + \mathbf{B}m \right) = \int_{\mathbf{b}_2}^{\infty^{(0)}} \mathcal{H} + \frac{j + \mathbf{B}m}{2}$$

for some  $j, m \in \mathbb{Z}$ , where we use involution-symmetric paths of integration. Notice that  $j, m$  cannot be simultaneously even as this would contradict unique solvability of the Jacobi inversion problem. In fact, it holds that

$$(7.19) \quad \left[ \int_{p^{(1)}}^{\infty^{(0)}} \mathcal{H} \right] = \left[ \int_{\mathbf{b}_2}^{\infty^{(0)}} \mathcal{H} + \int_{\mathbf{b}_2}^{p^{(0)}} \mathcal{H} \right] = \left[ \int_{\mathbf{b}_2}^{\mathbf{b}_1} \mathcal{H} \right] = \left[ \frac{1 + \mathbf{B}}{2} \right],$$

i.e., both  $j, m$  must be odd. Indeed, the first equality follows from the symmetry  $\alpha(z^*) = -\alpha(z)$ . Next, observe that the function  $(-1)^k \gamma^{-2}(z^{(k)}) - 1$  is meromorphic on  $\mathfrak{S}$  with zeros at  $\infty^{(0)}$  and  $p^{(0)}$  and poles at  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , all simple. Hence, the second equality is a consequence of the same symmetry and Abel's theorem. Further, consider an involution-symmetric cycle represented by the anti-diagonal on Figure 13, oriented so that it proceeds towards  $\mathbf{b}_1$  in  $D^{(0)}$ . It is clearly homologous to  $\alpha + \beta$ . Then the symmetry of  $\alpha(z)$  and the normalization in (7.2) and (7.3) yield the final equality in (7.19). Notice that the conclusion of (7.19) a posteriori holds with the bounds of integration in left most integral of (7.19) in flipped. Therefore, adding  $\int_{\infty^{(0)}}^{\mathbf{b}_2} \mathcal{H}$  to both sides of (7.9) gives us

$$(7.20) \quad \left[ \int_{\infty^{(0)}}^{z_{n,1}} \mathcal{H} \right] = \left[ \frac{1 + \mathbf{B}}{2} + (n - N_n)\zeta + (\omega + \mathbf{B}\tau)n \right] = \left[ \frac{1 + \mathbf{B}}{2} + (nk - N_n)\zeta \right].$$

Since  $\zeta$  has rational coordinates in the basis  $1, \mathbf{B}$  with denominator  $d$ , the right-hand side of (7.20) has at most  $d$  distinct values that depend only on  $\varrho \in \{0, \dots, d-1\}$ , the remainder of the division of  $nk - N_n$  by  $d$ . Let  $z_\varrho, \varrho \in \{0, \dots, d-1\}$ , be such that

$$\left[ \int_{\infty^{(0)}}^{z_\varrho} \mathcal{H} \right] = \left[ \frac{1 + \mathbf{B}}{2} + \varrho\zeta \right].$$

Clearly,  $\{z_{n,1}\}_{n \in \mathbb{N}} \subseteq \{z_\varrho\}_{\varrho=0}^{d-1}$ . Thus, it only remains to investigate when  $z_\varrho = \infty^{(0)}$ , or equivalently, when  $[(1 + \mathbf{B})/2 + \varrho\zeta] = [0]$ . From the representation of  $\zeta$  in (4.13), it must simultaneously hold that  $\varrho = d(2l_j - 1)/(2i_j), j \in \{1, 2\}$ , for some  $l_1, l_2 \in \mathbb{Z}$ . Since one of the pairs  $(i_j, d)$  is co-prime and  $\varrho \in \{0, \dots, d-1\}$ , this is possible only if  $\varrho = 0$  or  $\varrho = d/2$  (in the second case, of course,  $d$  must be even and  $i_j = 2l_j - 1$  must be odd), and the former is ruled out since  $[(1 + \mathbf{B})/2] \neq [0]$  (or equivalently, since  $0 \neq 2l_j - 1$ ).

**Lemma 7.4.** *If all three relations of (4.13) take place, then Jacobi inversion problem (7.9) for  $z_{n,1}$  has only finitely many distinct solutions and  $\infty^{(0)}$  is one of them if and only if  $d$  is even,  $i_1, i_2$  are odd, and  $nk - N_n$  is divisible by  $d/2$  but not by  $d$ .*

**7.5. Proof of Proposition 4.4.** Recall that Abel's map (7.8) is essentially carried out by the function  $\alpha(z)$  defined in (7.4). We shall consider the extension  $\tilde{\alpha}(z)$  of  $\alpha(z)$  to the whole surface  $\mathfrak{S}$  defined by setting  $\tilde{\alpha}(s) := \alpha_+(s)$  for  $s \in \alpha$  and  $s \in \beta \setminus \{\mathbf{b}_1\}$ . Given such an extension and (7.9), there exist unique integers  $j_{n,k}, m_{n,k}$  such that

$$(7.21) \quad \tilde{\alpha}(z_{n,k}) = \tilde{\alpha}(p^{(k)}) + (n - N_n)\zeta + (\omega + \mathbf{B}\tau)n + j_{n,k} + \mathbf{B}m_{n,k}, \quad k \in \{0, 1\}.$$

Recall the definition of  $\theta(u)$  in (4.10). The function  $\theta(u)$  is holomorphic in  $\mathbb{C}$  and enjoys the following periodicity properties:

$$(7.22) \quad \theta(u + j + \mathbf{B}m) = \exp\{-\pi i \mathbf{B}m^2 - 2\pi i u m\} \theta(u), \quad j, m \in \mathbb{Z}.$$

It is also known that  $\theta(u)$  vanishes only at the points of the lattice  $[\frac{B+1}{2}]$ . Let

$$(7.23) \quad \Theta_{n,k}(z) = \exp \left\{ -2\pi i (m_{n,k} + \tau n) \alpha(z) \right\} \frac{\theta(\alpha(z) - \tilde{\alpha}(z_{n,k}) - \frac{B+1}{2})}{\theta(\alpha(z) - \tilde{\alpha}(p^{(k)}) - \frac{B+1}{2})}.$$

Then we define  $\Theta_n(z; t)$  from Proposition 4.4 as well as the other functions appearing in the statement of Theorem 4.2 by

$$(7.24) \quad \Theta_n(z; t) := \Theta_{n,1}^{(0)}(z), \quad \vartheta_n(z) := \frac{\Theta_{n,1}^{(0)}(z)}{\Theta_{n,1}^{(0)}(\infty)}, \quad \vartheta_n^*(z) := \frac{\Theta_{n,1}^{(1)}(z)}{\Theta_{n,1}^{(0)}(\infty)}, \quad \text{and} \quad \tilde{\vartheta}_n(z) := \frac{\Theta_{n,0}^{(0)}(z)}{\Theta_{n,0}^{(1)}(\infty)},$$

where  $F^{(i)}(z)$ ,  $i \in \{0, 1\}$ , stands for the pull-back under  $\pi(z)$  of a function  $F(z)$  from  $D^{(i)}$  into  $\overline{\mathbb{C}} \setminus J_t$ .

The functions  $\Theta_{n,k}(z)$  are meromorphic on  $\mathfrak{S}_{\alpha,\beta}$  with exactly one pole, which is simple and located at  $p^{(k)}$ , and exactly one zero, which is also simple and located at  $z_{n,k}$  (observe that these functions can be analytically continued as multiplicatively multivalued functions on the whole surface  $\mathfrak{S}$ ; thus, we can talk about simplicity of a pole or zero regardless whether it belongs to the cycles of a homology basis or not). Moreover, according to (7.5), (7.21), and (7.22), they possess continuous traces on  $\alpha, \beta$  away from  $\mathbf{b}_1$  that satisfy

$$(7.25) \quad \Theta_{n,k+}(s) = \Theta_{n,k-}(s) \begin{cases} \exp \{ -2\pi i (\omega n + (n - N_n) \varsigma) \}, & s \in \alpha \setminus \{\mathbf{b}_1\}, \\ \exp \{ -2\pi i \tau n \}, & s \in \beta \setminus \{\mathbf{b}_1\}. \end{cases}$$

Recall functions  $A(z), B(z)$  from (7.11) and (7.12). To discuss boundedness properties of  $\Theta_{n,k}(z)$  and for the asymptotic analysis in the following section it will be convenient to define

$$(7.26) \quad M_{n,0}(z) = \Theta_{n,0}(z) \begin{cases} B(z), & z \in D^{(0)}, \\ A(z), & z \in D^{(1)}, \end{cases} \quad \text{and} \quad M_{n,1}(z) = \Theta_{n,1}(z) \begin{cases} A(z), & z \in D^{(0)}, \\ -B(z), & z \in D^{(1)}. \end{cases}$$

These functions are holomorphic on  $\mathfrak{S} \setminus \{\alpha \cup \beta \cup \Delta\}$  since the pole of  $\Theta_{n,k}(z)$  is canceled by the zero of  $B(z)$ . Each function  $M_{n,k}(z)$  has exactly two zeros, namely,  $z_{n,k}$  and  $\infty^{(k)}$ . It follows from (7.12) and (7.25) that

$$(7.27) \quad \begin{cases} M_{n,k\pm}^{(0)}(s) = \mp M_{n,k\mp}^{(1)}(s), & s \in J_{t,2}^\circ, \\ M_{n,k\pm}^{(0)}(s) = \mp e^{-2\pi i \tau n} M_{n,k\mp}^{(1)}(s), & s \in J_{t,1}^\circ, \\ M_{n,k\pm}^{(i)}(s) = e^{(-1)^i 2\pi i (n\omega + (n - N_n) \varsigma)} M_{n,k\mp}^{(i)}(s), & s \in I_t^\circ. \end{cases}$$

It further follows from (7.10) and (7.11) that  $|M_{n,k}(z)| \sim |z - e|^{-1/4}$  as  $z \rightarrow e \in \mathbf{E}$  unless  $z_{n,k}$  coincides with  $e$  in which case the exponent becomes  $1/4$ . Assume now that there exists  $N_* \geq 0$  such that  $|n - N_n| \leq N_*$  for all  $n \in \mathbb{N}$ . Then for each  $\delta > 0$  there exists  $C(\delta, N_*)$  independent of  $n$  such that

$$(7.28) \quad |M_{n,k}(z)| \leq C(\delta, N_*), \quad z \in O_\delta := \mathfrak{S} \setminus \cup_{e \in \mathbf{E}} \pi^{-1}\{|z - e| < \delta\}.$$

Indeed, observe that  $\{\log |M_{n,k}(z)|\}$  is a family of subharmonic functions in  $O_\delta \setminus \alpha$  (the jump of  $M_{n,k}(z)$  is unimodular on  $\beta$ ). By the maximum principle for subharmonic functions,  $\log |M_{n,k}(z)|$  reaches its maximum on  $\partial(O_\delta \setminus \alpha)$ , where the maximum is clearly finite. Since the sequence  $\{n - N_n\}$  is bounded by assumption and the range of  $\tilde{\alpha}(z)$  is bounded by construction, so are the sequences  $\{m_{n,k} + \tau n\}$  and  $\{j_{n,k} + \omega n\}$ , see (7.21) and recall that  $j_{n,k} + \omega n$  are real and  $\text{Im}(B) > 0$ . Thus, any limit point of  $\{\log |M_{n,k}(z)|\}$  is obtained by taking simultaneous limit points of  $\{n - N_n\}$ ,  $\{m_{n,k} + \tau n\}$ , and  $\{j_{n,k} + \omega n\}$ , computing the corresponding solution  $z_k$  of the Jacobi inversion problem (7.21) and plugging all of these quantities into the right-hand side of (7.23). Hence, all these limit functions are also bounded above on the closure of  $O_\delta \setminus \alpha$ , which proves (7.28). Finally, it holds that

$$(7.29) \quad |M_{n,k}(\infty^{(1-k)})| \geq c_\varepsilon, \quad n \in \mathbb{N}(\varepsilon),$$



for some constant  $c_\varepsilon > 0$  by a similar compactness argument combined with the definition of  $\mathbb{N}(\varepsilon)$  in Proposition 4.4, the observation that  $M_{n,k}(\mathbf{z})$  is non-zero at  $\infty^{(1-k)}$  when  $\infty^{(1-k)} \neq \mathbf{z}_{n,k}$ , and the last conclusion of Proposition 7.1.

**7.6. Proof of Proposition 4.3.** Define a function  $F(z)$  on  $\mathfrak{S} \setminus \{\alpha \cup \beta\}$  by setting  $F(z^{(1)}) := 1/F(z^{(0)})$  and

$$F(z^{(0)}) := \mathcal{D}^{-1}(z) \exp \{V(z)/2 + Q(z)\},$$

see (4.2) and (4.5). It follows from (4.3) and (4.8) that

$$(7.30) \quad F_+(s) = F_-(s) \begin{cases} \exp \{2\pi i(\omega + \varsigma)\}, & s \in \alpha \setminus \{\mathbf{b}_1\}, \\ \exp \{2\pi i\tau\}, & s \in \beta \setminus \{\mathbf{b}_1\}, \end{cases}$$

where one needs to recall that  $I_t$  is oriented towards  $a_2$  while  $\alpha \cap \mathfrak{S}^{(0)}$  is oriented towards  $\mathbf{b}_1$ . Moreover, it follows from (4.7) that

$$(7.31) \quad F(z^{(0)}) = e^{\ell^*/2} \mathcal{D}^{-1}(\infty)z + \mathcal{O}(1)$$

as  $z \rightarrow \infty$ . On the other hand, we get from (7.16) that there exist integers  $j, m$  such that

$$(7.32) \quad \tilde{\mathbf{a}}(p^{(1)}) - \tilde{\mathbf{a}}(p^{(0)}) = \varsigma + \omega + \mathbf{B}\tau + j + \mathbf{B}m.$$

Thus, we get similarly to (7.23) and (7.25) that the function

$$(7.33) \quad \Theta(z) := \exp \left\{ -2\pi i(m + \tau)\mathbf{a}(z) \right\} \frac{\theta(\mathbf{a}(z) - \tilde{\mathbf{a}}(p^{(1)}) - \frac{\mathbf{B}+1}{2})}{\theta(\mathbf{a}(z) - \tilde{\mathbf{a}}(p^{(0)}) - \frac{\mathbf{B}+1}{2})}$$

is meromorphic in  $\mathfrak{S} \setminus \{\alpha \cup \beta\}$ , with a simple zero at  $p^{(1)}$ , a simple pole at  $p^{(0)}$ , and otherwise non-vanishing and finite, whose traces on the cycles of the homology basis satisfy

$$(7.34) \quad \Theta_+(s) = \Theta_-(s) \begin{cases} \exp \{ -2\pi i(\omega + \varsigma) \}, & s \in \alpha \setminus \{\mathbf{b}_1\}, \\ \exp \{ -2\pi i\tau \}, & s \in \beta \setminus \{\mathbf{b}_1\}. \end{cases}$$

Combining (7.30), (7.31), and (7.34) yields that the function  $(F\Theta)(z)$  is rational on  $\mathfrak{S}$  with the divisor  $p^{(1)} + \infty^{(1)} - p^{(0)} - \infty^{(0)}$ . That is, this function is a constant multiple of the reciprocal of (7.13). Observe that

$$A(\infty) = 1 \quad \text{and} \quad B(z) = \frac{S_1}{4i} \frac{1}{z} + \mathcal{O}(z^{-2})$$

as  $z \rightarrow \infty$ . Then we get from (7.13), (7.31), and the symmetry  $F(z)F(z^*) \equiv 1$  that

$$(7.35) \quad e^{\ell^*} = \mathcal{D}^2(\infty) \frac{16 \Theta(\infty^{(1)})}{S_1^2 \Theta(\infty^{(0)})}.$$

Thus, it only remains to obtain the expression for the last fraction. We get from (7.19) that

$$\mathbf{a}(\infty^{(0)}) - \tilde{\mathbf{a}}(p^{(1)}) = \frac{1 + \mathbf{B}}{2} + u + \mathbf{B}v$$

for some integers  $u, v$ . Hence, we get from (7.7), (7.32), and (7.22) that

$$\Theta(\infty^{(0)}) = \frac{e^{-\pi i(m+\tau)(\varsigma+\omega+\mathbf{B}\tau)} \theta(\mathbf{B}v)}{\theta(\varsigma + \omega + \mathbf{B}\tau + \mathbf{B}(m+v))} = \frac{e^{\pi i(2v+m-\tau)(\varsigma+\omega+\mathbf{B}\tau) + \pi i\mathbf{B}(2v+m)} \theta(0)}{\theta(\varsigma + \omega + \mathbf{B}\tau)}$$

and similarly

$$\frac{1}{\Theta(\infty^{(1)})} = \frac{e^{-\pi i(m+\tau)(\varsigma+\omega+\mathbf{B}\tau)} \theta(\mathbf{B}(m+v))}{\theta(\mathbf{B}v - \varsigma - \omega - \mathbf{B}\tau)} = \frac{e^{-\pi i(2v+m+\tau)(\varsigma+\omega+\mathbf{B}\tau) - \pi i\mathbf{B}(2v+m)} \theta(0)}{\theta(-\varsigma - \omega - \mathbf{B}\tau)}.$$

Since  $\theta(\cdot)$  is an even function, this finishes the proof of the proposition.

## 8. ASYMPTOTIC ANALYSIS

**8.1. Initial Riemann-Hilbert Problem.** As agreed before, we omit the dependence on  $t$  whenever it does not cause ambiguity. In what follows, it will be convenient to set

$$\mathbf{I} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We are seeking solutions of the following sequence of Riemann-Hilbert problems for  $2 \times 2$  matrix functions (RHP- $\mathbf{Y}$ ):

- (a)  $\mathbf{Y}(z)$  is analytic in  $\mathbb{C} \setminus \Gamma$  and  $\lim_{\mathbb{C} \setminus \Gamma \ni z \rightarrow \infty} \mathbf{Y}(z)z^{-n\sigma_3} = \mathbf{I}$ ;
- (b)  $\mathbf{Y}(z)$  has continuous traces on  $\Gamma \setminus \{a_1, b_1, a_2, b_2\}$  that satisfy

$$\mathbf{Y}_+(s) = \mathbf{Y}_-(s) \begin{pmatrix} 1 & e^{-N_n V(s)} \\ 0 & 1 \end{pmatrix},$$

where, as before,  $V(z)$  is given by (1.11) and the sequence  $\{N_n\}$  is such that  $|n - N_n| \leq N_*$  for some  $N_* > 0$ .

The connection of RHP- $\mathbf{Y}$  to orthogonal polynomials was first demonstrated by Fokas, Its, and Kitaev in [21] and lies in the following. If the solution of RHP- $\mathbf{Y}$  exists, then it is necessarily of the form

$$(8.1) \quad \mathbf{Y}(z) = \begin{pmatrix} P_n(z) & (CP_n e^{-N_n V})(z) \\ -\frac{2\pi i}{h_{n-1}} P_{n-1}(z) & -\frac{2\pi i}{h_{n-1}} (CP_{n-1} e^{-N_n V})(z) \end{pmatrix},$$

where  $P_n(z) = P_n(z; t, N_n)$  are the polynomial satisfying orthogonality relations (1.10),  $h_n = h_n(t, N_n)$  are the constants defined in (1.12), and  $Cf(z)$  is the Cauchy transform of a function  $f$  given on  $\Gamma$ , i.e.,

$$(Cf)(z) := \frac{1}{2\pi i} \int_{\Gamma} \frac{f(s)}{s - z} ds.$$

Below, we show the solvability of RHP- $\mathbf{Y}$  for all  $n \in \mathbb{N}(t, \varepsilon)$  large enough following the framework of the steepest descent analysis introduced by Deift and Zhou [15]. The latter lies in a series of transformations which reduce RHP- $\mathbf{Y}$  to a problem with jumps asymptotically close to identity.

**8.2. Renormalized Riemann-Hilbert Problem.** Suppose that  $\mathbf{Y}(z)$  is a solution of RHP- $\mathbf{Y}$ . Put

$$(8.2) \quad \mathbf{T}(z) := e^{n\ell_* \sigma_3 / 2} \mathbf{Y}(z) e^{-n(g(z) + \ell_*/2)\sigma_3},$$

where the function  $g(z)$  is defined by (6.2) and  $\ell_*$  was introduced in (6.3). Then

$$\mathbf{T}_+(s) = \mathbf{T}_-(s) \begin{pmatrix} e^{-n(g_+(s) - g_-(s))} & e^{n(g_+(s) + g_-(s) - V(s) + \ell_*) + (n - N_n)V(s)} \\ 0 & e^{-n(g_-(s) - g_+(s))} \end{pmatrix},$$

$s \in \Gamma$ , and therefore we deduce from (4.7) and (6.6)–(6.8) that  $\mathbf{T}(z)$  solves RHP- $\mathbf{T}$ :

- (a)  $\mathbf{T}(z)$  is analytic in  $\mathbb{C} \setminus \Gamma$  and  $\lim_{\mathbb{C} \setminus \Gamma \ni z \rightarrow \infty} \mathbf{T}(z) = \mathbf{I}$ ;
- (b)  $\mathbf{T}(z)$  has continuous traces on  $\Gamma \setminus \{a_1, b_1, a_2, b_2\}$  that satisfy

$$\mathbf{T}_+(s) = \mathbf{T}_-(s) \begin{cases} \begin{pmatrix} 1 & e^{n(2\pi i \tau + \phi_{a_1}(s)) + (n - N_n)V(s)} \\ 0 & 1 \end{pmatrix}, & s \in \Gamma(e^{i\pi} \infty, a_1), \\ \begin{pmatrix} 1 & e^{n\phi_{b_2}(s) + (n - N_n)V(s)} \\ 0 & 1 \end{pmatrix}, & s \in \Gamma(b_2, e^{\pi i/3} \infty), \\ \begin{pmatrix} e^{2\pi i \omega n} & e^{n\phi_{a_2}(s) + (n - N_n)V(s)} \\ 0 & e^{-2\pi i \omega n} \end{pmatrix}, & s \in \Gamma(b_1, a_2), \end{cases}$$

and

$$\mathbf{T}_+(s) = \mathbf{T}_-(s) \begin{cases} \begin{pmatrix} e^{-n\phi_{b_2^+}(s)} & e^{(n-N_n)V(s)} \\ 0 & e^{-n\phi_{b_2^-}(s)} \end{pmatrix}, & s \in \Gamma(a_2, b_2), \\ \begin{pmatrix} e^{-n(\phi_{b_2^+}(s)-2\pi i\tau)} & e^{2\pi i\tau n+(n-N_n)V(s)} \\ 0 & e^{-n(\phi_{b_2^-}(s)-2\pi i\tau)} \end{pmatrix}, & s \in \Gamma(a_1, b_1). \end{cases}$$

Clearly, if **RHP-T** is solvable and  $\mathbf{T}(z)$  is the solution, then by inverting (8.2) one obtains a matrix  $\mathbf{Y}(z)$  that solves **RHP-Y**.

**8.3. Lens Opening.** As usual in the steepest descent analysis of matrix Riemann-Hilbert problems for orthogonal polynomials, the next step is based on the identity

$$\begin{pmatrix} e^{-n(\phi_{b_2^+}(s)-C)} & e^{nC+(n-N_n)V(s)} \\ 0 & e^{-n(\phi_{b_2^-}(s)-C)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ e^{-n\phi_{b_2^-}(s)-(n-N_n)V(s)} & 1 \end{pmatrix} \times \\ \begin{pmatrix} 0 & e^{nC+(n-N_n)V(s)} \\ -e^{-nC-(n-N_n)V(s)} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ e^{-n\phi_{b_2^+}(s)-(n-N_n)V(s)} & 1 \end{pmatrix}$$

that follows from (6.6), where  $C = 2\pi i\tau$  when  $s \in \Gamma(a_1, b_1)$  and  $C = 0$  when  $s \in \Gamma(a_2, b_2)$ . To carry it out, we shall introduce two additional system of arcs.

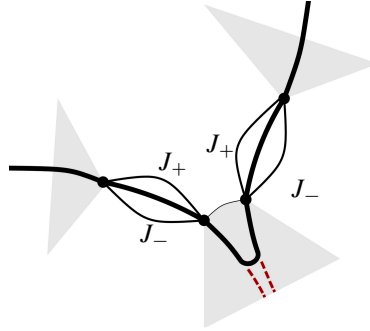


FIGURE 14. The thick curves represent  $\Gamma$  and thinner black curves represent  $J_{\pm}$ . The shaded part represents regions where  $\operatorname{Re}(\phi_e(z)) < 0$  (they are the same for each  $e \in \{a_1, b_1, a_2, b_2\}$ ).

Denote by  $J_{\pm}$  smooth homotopic deformations of  $J_t$  within the region  $\operatorname{Re}(\phi_{b_2}(z)) > 0$  such that  $J_+$  lies to the left and  $J_-$  to the right of  $J_t$ , see Figure 14 (both connected subarcs of  $J_+$ , resp.  $J_-$ , are oriented from  $a_i$  to  $b_i$ ,  $i \in \{1, 2\}$ ). We shall fix the way these arcs emanate from  $e \in \{a_1, b_1, a_2, b_2\}$ . Namely, let  $U_e$  be given by (6.9) and  $(-\phi_e)^{2/3}(z)$  be as in (6.14). Then we require that

$$(8.3) \quad \arg((-\phi_e)^{2/3}(z)) = \pm \nu_e(2\pi/3), \quad z \in U_e \cap J_{\pm},$$

where  $\nu_e$  is defined by (6.13). This requirement always can be fulfilled due to conformality  $(-\phi_e)^{2/3}(z)$  in  $U_e$  and the choice of the branch in (6.14).

Denote by  $O_{\pm}$  the open sets delimited by  $J_{\pm}$  and  $J_t$ . Set

$$(8.4) \quad \mathbf{S}(z) := \mathbf{T}(z) \begin{cases} \begin{pmatrix} 1 & 0 \\ \mp e^{-n\phi_{b_2}(z)-(n-N_n)V(z)} & 1 \end{pmatrix}, & z \in O_{\pm}, \\ \mathbf{I}, & \text{otherwise.} \end{cases}$$

Then, if  $\mathbf{T}(z)$  solves **RHP-T**,  $\mathbf{S}(z)$  solves **RHP-S**:

- (a)  $\mathbf{S}(z)$  is analytic in  $\mathbb{C} \setminus (\Gamma \cup J_+ \cup J_-)$  and  $\lim_{\mathbb{C} \setminus \Gamma \ni z \rightarrow \infty} \mathbf{S}(z) = \mathbf{I}$ ;

- (b)  $\mathbf{S}(z)$  has continuous traces on  $\Gamma \setminus \{a_1, b_1, a_2, b_2\}$  that satisfy **RHP-T**(b) on  $\Gamma(e^{\pi i \infty}, a_1)$ ,  $\Gamma(b_1, a_2)$ , and  $\Gamma(b_2, e^{\pi i/3 \infty})$ , as well as

$$\mathbf{S}_+(s) = \mathbf{S}_-(s) \begin{cases} \begin{pmatrix} 0 & e^{(n-N_n)V(s)} \\ -e^{-(n-N_n)V(s)} & 0 \end{pmatrix}, & s \in \Gamma(a_2, b_2), \\ \begin{pmatrix} 0 & e^{2\pi i \tau n + (n-N_n)V(s)} \\ -e^{-2\pi i \tau n - (n-N_n)V(s)} & 0 \end{pmatrix}, & s \in \Gamma(a_1, b_1), \\ \begin{pmatrix} 1 & 0 \\ e^{-n\phi_{b_2}(s) - (n-N_n)V(s)} & 1 \end{pmatrix}, & s \in J_{\pm}. \end{cases}$$

As before, since transformation (8.4) is invertible, a solution of **RHP-S** yields a solution of **RHP-T**.

**8.4. Global Parametrix.** The Riemann-Hilbert problem for the global parametrix is obtained from **RHP-S** by removing the quantities that are asymptotically zero from the jump matrices in **RHP-S**(b). The latter can be easily identified with the help of (6.5) and by recalling that the constant  $\tau$  is real. Thus, we are seeking the solution of **RHP-N**:

- (a)  $N(z)$  is analytic in  $\overline{\mathbb{C}} \setminus \Gamma[a_1, b_2]$  and  $N(\infty) = \mathbf{I}$ ;  
(b)  $N(z)$  has continuous traces on  $\Gamma(a_1, b_2) \setminus \{b_1, a_2\}$  that satisfy

$$N_+(s) = N_-(s) \begin{cases} \begin{pmatrix} 0 & e^{(n-N_n)V(s)} \\ -e^{-(n-N_n)V(s)} & 0 \end{pmatrix}, & s \in \Gamma(a_2, b_2), \\ \begin{pmatrix} 0 & e^{2\pi i \tau n + (n-N_n)V(s)} \\ -e^{-2\pi i \tau n - (n-N_n)V(s)} & 0 \end{pmatrix}, & s \in \Gamma(a_1, b_1), \\ \begin{pmatrix} e^{2\pi i \omega n} & 0 \\ 0 & e^{-2\pi i \omega n} \end{pmatrix}, & s \in \Gamma(b_1, a_2). \end{cases}$$

We shall solve this problem only for  $n \in \mathbb{N}(\varepsilon) = \mathbb{N}(t, \varepsilon)$  from Proposition 4.4. In fact, to solve **RHP-N** we only need to exclude indices  $n$  for which  $z_{n,1} = \infty^{(0)}$  (in this case  $z_{n,0} = \infty^{(1)}$ , see Proposition 7.1) as should be clear from (8.5). However, for further analysis we shall need estimate (7.29), where it is crucial that  $n \in \mathbb{N}(\varepsilon)$  for some  $\varepsilon > 0$ .

Let the functions  $M_{n,k}(z)$  be given by (7.26) and  $\mathcal{D}(z) = \mathcal{D}(z; t)$  be defined by (4.2). With the notation introduced right after (7.23), a solution of **RHP-N** is given by

$$(8.5) \quad N(z) = \mathbf{M}^{-1}(\infty) \mathbf{M}(z), \quad \mathbf{M}(z) := \begin{pmatrix} M_{n,1}^{(0)}(z) & M_{n,1}^{(1)}(z) \\ M_{n,0}^{(0)}(z) & M_{n,0}^{(1)}(z) \end{pmatrix} \mathcal{D}^{(N_n - n)\sigma_3}(z).$$

Indeed, **RHP-N**(a) follows from holomorphy of  $\mathcal{D}(z)$  and  $M_{n,k}(z)$  discussed in Proposition 4.1 and right after (7.26). Fulfillment of **RHP-N**(b) can be checked by using (4.3) and (7.27). Observe also that  $\det(N(z)) \equiv 1$ . Indeed, as the jump matrices in **RHP-N**(b) have unit determinants,  $\det(N(z))$  is holomorphic through  $\Gamma(a_1, b_1)$ ,  $\Gamma(b_1, a_2)$ , and  $\Gamma(a_2, b_2)$ . It also has at most square root singularities at  $\{a_1, b_1, a_2, b_2\}$  as explained right after (7.27). Thus, it is holomorphic throughout  $\overline{\mathbb{C}}$  and therefore is a constant. The normalization at infinity implies that this constant is 1.

**8.5. Local Parametrix.** The jumps discarded in **RHP-N** are not uniformly close to the identity around each point  $e \in \{a_1, b_1, a_2, b_2\}$ . The goal of this section is to solve **RHP-S** in the disks  $U_e$ , see (6.9), with a certain matching condition on the boundary of the disks. More precisely, we are looking for a matrix functions  $\mathbf{P}_e(z)$  that solves **RHP-P**:

- (a)  $\mathbf{P}_e(z)$  has the same analyticity properties as  $\mathbf{S}(z)$  restricted to  $U_e$ , see **RHP-S**(a);  
(b)  $\mathbf{P}_e(z)$  satisfies the same jump relations as  $\mathbf{S}(z)$  restricted to  $U_e$ , see **RHP-S**(b);  
(c)  $\mathbf{P}_e(z) = N(z)(\mathbf{I} + \mathcal{O}(n^{-1}))$  holds uniformly on  $\partial U_e$  as  $n \rightarrow \infty$ .

Again, we shall solve **RHP-P** only for  $n \in \mathbb{N}(\varepsilon)$  with  $\mathcal{O}(\cdot) = \mathcal{O}_{\varepsilon}(\cdot)$  in **RHP-P**(c).

Let  $U_e, J_e$ , and  $I_e, e \in \{a_1, b_1, a_2, b_2\}$ , be as in (6.9) and (6.10). Further, let  $A(\zeta)$  be the Airy matrix [13, 14]. That is, it is analytic in  $\mathbb{C} \setminus ((-\infty, \infty) \cup L_- \cup L_+)$ ,  $L_\pm := \{\zeta : \arg(\zeta) = \pm 2\pi/3\}$ , and satisfies

$$\mathbf{A}_+(s) = \mathbf{A}_-(s) \begin{cases} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, & s \in (-\infty, 0), \\ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, & s \in L_\pm, \\ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, & s \in (0, \infty), \end{cases}$$

where the real line is oriented from  $-\infty$  to  $\infty$  and the rays  $L_\pm$  are oriented towards the origin. It is known that  $A(\zeta)$  has the following asymptotic expansion<sup>4</sup> at infinity:

$$(8.6) \quad A(\zeta) e^{\frac{2}{3}\zeta^{3/2}\sigma_3} \sim \frac{\zeta^{-\sigma_3/4}}{\sqrt{2}} \sum_{k=0}^{\infty} \begin{pmatrix} s_k & 0 \\ 0 & t_k \end{pmatrix} \begin{pmatrix} (-1)^k & i \\ (-1)^k i & 1 \end{pmatrix} \left(\frac{2}{3}\zeta^{3/2}\right)^{-k},$$

where the expansion holds uniformly in  $\mathbb{C} \setminus ((-\infty, \infty) \cup L_- \cup L_+)$ , and

$$s_0 = t_0 = 1, \quad s_k = \frac{\Gamma(3k + 1/2)}{54^k k! \Gamma(k + 1/2)}, \quad t_k = -\frac{6k + 1}{6k - 1} s_k, \quad k \geq 1.$$

Let us write  $A_e := A$  if  $e \in \{b_1, b_2\}$  and  $A_e := \sigma_3 A \sigma_3$  if  $e \in \{a_1, a_2\}$ . It can be easily checked that  $\sigma_3 A \sigma_3$  has the same jumps as  $A$  only with the reversed orientation of the rays. Moreover, one needs to replace  $i$  by  $-i$  in (8.6) when describing the behavior of  $\sigma_3 A \sigma_3$  at infinity. Let  $\zeta_e(z) := [-n(3/4)\phi_e(z)]^{2/3}$ , which is conformal in  $U_e$ , see (6.14). Further, put

$$\mathbf{J}_e(z) := e^{(N_n - n)V(z)\sigma_3/2} \begin{cases} \mathbf{I}, & e = b_2, \\ e^{\pi i(\pm\omega)n\sigma_3}, & e = a_2, \\ e^{\pi i(\pm\omega - \tau)n\sigma_3}, & e = b_1, \\ e^{-\pi i\tau n\sigma_3}, & e = a_1, \end{cases}$$

where we use  $\omega$  if  $z$  lies to the left of  $\Gamma$  and use  $-\omega$  if  $z$  lies to the right of  $\Gamma$ . Then it can be readily verified by using (6.5) that

$$(8.7) \quad \mathbf{P}_e(z) := \mathbf{E}_e(z) \mathbf{A}_e(\zeta_e(z)) e^{(2/3)\zeta_e^{3/2}(z)\sigma_3} \mathbf{J}_e(z),$$

satisfies **RHP-P**(a,b) for any matrix  $\mathbf{E}_e(z)$  holomorphic in  $U_e$ . It follows immediately from (8.6) that **RHP-P**(c) will be satisfied if

$$(8.8) \quad \mathbf{E}_e(z) := (\mathbf{N} \mathbf{J}_e^{-1})(z) \begin{pmatrix} 1 & -v_e i \\ -v_e i & 1 \end{pmatrix} \frac{\zeta_e^{\sigma_3/4}(z)}{\sqrt{2}},$$

provided this matrix function is holomorphic in  $U_e$ , where  $v_e$  was defined in (6.13). By using **RHP-N**(b) and (6.15) one can readily check that  $\mathbf{E}_e(z)$  is holomorphic in  $U_e \setminus \{e\}$ . Since  $\zeta_e(z)$  has a simple zero at  $e$ , it also follows from (8.5) and the claim after (7.27) that  $\mathbf{E}_e(z)$  can have at most square root singularity at  $e$  and therefore is in fact holomorphic in the entire disk  $U_e$  as needed.

In fact, it follows from (8.6)–(8.8) that

$$(8.9) \quad \mathbf{P}_e(z) \sim \mathbf{N}(z) \left( \mathbf{I} + \frac{1}{n} \sum_{k=0}^{\infty} \frac{\mathbf{P}_{e,k}(z)}{n^k} \right),$$

<sup>4</sup>As usual, we say that a function  $F(p)$ , which might depend on other variables as well, admits an asymptotic expansion  $F(p) \sim \sum_{k=1}^{\infty} c_k \psi_k(p)$ , where the functions  $\psi_k(p)$  depend only on  $p$  while the coefficients  $c_k$  might depend on other variables but not  $p$ , if for any natural number  $K$  it holds that  $F(p) - \sum_{k=0}^{K-1} c_k \psi_k(p) = O_K(\psi_K(p))$ .

where the expansion inside the parentheses holds uniformly on  $\partial U_e$  and locally uniformly for  $t \in \mathcal{O}_{\text{two-cut}}$ , and

$$(8.10) \quad \mathbf{P}_{e,k-1}(z) = \mathbf{J}_e^{-1}(z) \begin{pmatrix} 1 & -v_e i \\ -v_e i & 1 \end{pmatrix} \begin{pmatrix} s_k & 0 \\ 0 & t_k \end{pmatrix} \begin{pmatrix} (-1)^k & v_e i \\ v_e (-1)^k i & 1 \end{pmatrix} \mathbf{J}_e(z) \left( -\frac{\phi_e(z)}{2} \right)^{-k}, \quad k \geq 1.$$

**8.6. RH Problem with Small Jumps.** Set  $\Sigma := ([(\Gamma \setminus J_t) \cup J_+ \cup J_-] \cap D) \cup (\cup_e \partial U_e)$ ,  $D := \mathbb{C} \setminus \cup_e \bar{U}_e$ . We shall show that for all  $n \in \mathbb{N}(\varepsilon)$  large enough there exists a matrix function  $\mathbf{R}(z)$  that solves the following Riemann-Hilbert problem (RHP- $\mathbf{R}$ ):

- (a)  $\mathbf{R}(z)$  is holomorphic in  $\mathbb{C} \setminus \Sigma$  and  $\lim_{\mathbb{C} \setminus \Gamma \ni z \rightarrow \infty} \mathbf{R}(z) = \mathbf{I}$ ;
- (b)  $\mathbf{R}(z)$  has continuous traces on  $\Sigma^\circ$  (points with the well defined tangent) that satisfy

$$\mathbf{R}_+(s) = \mathbf{R}_-(s) \begin{cases} \mathbf{P}_e(s) \mathbf{N}^{-1}(s), & s \in \partial U_e, \\ \mathbf{N}(s) \begin{pmatrix} 1 & \\ e^{-n\phi_{b_2}(s) - (n-N_n)V(s)} & 0 \end{pmatrix} \mathbf{N}^{-1}(s), & s \in J_\pm \cap D, \end{cases}$$

where  $\partial U_e$  is oriented clockwise, and

$$\mathbf{R}_+(s) = \mathbf{R}_-(s) \begin{cases} \mathbf{N}(s) \begin{pmatrix} 1 & e^{n(2\pi i \tau + \phi_{a_1}(s)) + (n-N_n)V(s)} \\ 0 & 1 \end{pmatrix} \mathbf{N}^{-1}(s), & s \in \Gamma(e^{i\pi} \infty, a_1) \cap D, \\ \mathbf{N}_-(s) \begin{pmatrix} e^{2\pi i \omega n} & e^{n\phi_{a_2}(s) + (n-N_n)V(s)} \\ 0 & e^{-2\pi i \omega n} \end{pmatrix} \mathbf{N}_+^{-1}(s), & s \in \Gamma(b_1, a_2) \cap D, \\ \mathbf{N}(s) \begin{pmatrix} 1 & e^{n\phi_{b_2}(s) + (n-N_n)V(s)} \\ 0 & 1 \end{pmatrix} \mathbf{N}^{-1}(s), & s \in \Gamma(b_2, e^{\pi i/3} \infty) \cap D. \end{cases}$$

Observe that RHP- $\mathbf{R}$  is a well posed problem as  $\det(\mathbf{N}(z)) \equiv 1$ , as explained after (8.5), and therefore the matrix is invertible. Recall also that the entries of  $\mathbf{N}(z)$  and  $\mathbf{N}^{-1}(z)$  are uniformly bounded on  $\Sigma$  for  $n \in \mathbb{N}(\varepsilon)$  according to (7.28) and (7.29).

To prove solvability of RHP- $\mathbf{R}$ , let us show that the jump matrices in RHP- $\mathbf{R}$ (b) are close to the identity. To this end, set

$$(8.11) \quad \Delta(s) := (\mathbf{R}_-^{-1} \mathbf{R}_+)(s) - \mathbf{I}, \quad s \in \Sigma.$$

Since the entries of  $\mathbf{N}(z)$  are uniformly bounded on each  $\partial U_e$  with respect to  $n \in \mathbb{N}(\varepsilon)$ , it holds by RHP- $\mathbf{P}$ (c) and (8.9) that

$$(8.12) \quad \Delta(s) \sim \frac{1}{n} \sum_{k=0}^{\infty} \frac{(\mathbf{N} \mathbf{P}_{e,k} \mathbf{N}^{-1})(s)}{n^k},$$

where this pseudo expansion<sup>5</sup> is valid uniformly on  $\partial U_e$ . Thus, it holds that

$$(8.13) \quad \|\Delta\|_{L^\infty(\cup_e \partial U_e)} = \mathcal{O}_\varepsilon(n^{-1}).$$

Moreover, it follows from (6.11) and the sentence right after that there exists a constant  $C_D < 1$ , depending on the radii of the disks  $U_e$ , such that  $|e^{\phi_{b_2}(s)}| < C_D$  for  $s \in \Gamma(b_2, e^{\pi i/3} \infty) \cap D$ ,  $|e^{\phi_{a_1}(s)}| < C_D$  for  $s \in \Gamma(e^{\pi i} \infty, a_1) \cap D$ , and  $|e^{\phi_{a_2}(s)}| < C_D$  for  $s \in \Gamma(b_1, a_2) \cap D$ . Therefore,

$$(8.14) \quad \Delta(s) = \mathbf{N}_-(s) \begin{pmatrix} 0 & e^{n(\phi_e(s)+C) + (n-N_n)V(s)} \\ 0 & 0 \end{pmatrix} \mathbf{N}_+^{-1}(s) = \mathcal{O}(C_D^n)$$

on  $(\Sigma \cap D) \setminus (J_+ \cup J_-)$  since  $C$  is either zero or purely imaginary, the entries of  $\mathbf{N}(z)$  are bounded and so is the sequence  $\{n - N_n\}$ , the subscripts  $\pm$  are needed only on  $\Gamma(b_1, a_2) \cap D$ , and we used the fact

$$\mathbf{N}_-(s) e^{2\pi i \omega n \sigma_3} \mathbf{N}_+^{-1}(s) = \mathbf{I}, \quad s \in \Gamma(a_1, b_2),$$

<sup>5</sup>We say that a function  $F(p)$  admits a pseudo expansion  $F(p) \sim \sum_{k=1}^{\infty} c_k \psi_k(p)$  if the coefficients  $c_k$  do depend on  $p$ , but the estimates  $F(p) - \sum_{k=0}^{K-1} c_k \psi_k(p) = \mathcal{O}_K(\psi_K(p))$  remain valid for any  $K$ .

see **RHP-N**(b). Similarly, we get that

$$(8.15) \quad \Delta(s) = \mathbf{N}(s) \begin{pmatrix} 0 & 0 \\ e^{-n\phi_{b_2}(s)-(n-N_n)V(s)} & 0 \end{pmatrix} \mathbf{N}^{-1}(s) = \mathcal{O}(C_D^n)$$

on  $J_{\pm} \cap D$  for a possibly adjusted constant  $C_D$ , where we used the fact that  $\operatorname{Re}(\phi_{b_2}(s)) > 0$  for  $s \in J_{\pm} \setminus E$ , see Figure 14.

Equations (8.13), (8.14), and (8.15) show that  $\Delta(s)$  is uniformly close to zero. Since the entries of  $\mathbf{N}(z)$  are holomorphic at infinity and  $e^{n\phi_e(s)}$  is geometrically small as  $\Gamma \ni s \rightarrow \infty$ ,  $\Delta(s)$  is close to zero in  $L^2$ -norm as well. Then it follows from the same analysis as in [12, Corollary 7.108] that  $\mathbf{R}(z)$  exists for all  $n \in \mathbb{N}(\varepsilon)$  large enough and it holds uniformly in  $\mathbb{C}$  that

$$(8.16) \quad \mathbf{R}(z) = \mathbf{I} + \mathcal{O}_{\varepsilon}(n^{-1}).$$

**8.7. Solution of RHP-Y.** Given  $\mathbf{R}(z)$ ,  $\mathbf{N}(z)$ , and  $\mathbf{P}_e(z)$ , solutions of **RHP-R**, **RHP-N**, and **RHP-P**, respectively, it is a trivial verification to check that **RHP-S** is solved by

$$(8.17) \quad \mathbf{S}(z) = \begin{cases} (\mathbf{RN})(z) & \text{in } D \setminus [(\Gamma \setminus J_t) \cup J_+ \cup J_-], \\ (\mathbf{RP}_e)(z) & \text{in } U_e, e \in \{a_1, b_1, a_2, b_2\}. \end{cases}$$

Let  $K$  be a compact set in  $\mathbb{C} \setminus \Gamma$ . We can always arrange so that the set  $K$  lies entirely within the unbounded component of the contour  $\Sigma$ . Then it follows from (8.2), (8.4), and (8.17) that

$$(8.18) \quad \mathbf{Y}(z) = e^{-n\ell^* \sigma_3/2} (\mathbf{RN})(z) e^{n(g(z) + \ell^*/2)\sigma_3}, \quad z \in K.$$

Subsequently, by using (8.1) and (6.3), we see that

$$P_n(z) = [\mathbf{Y}(z)]_{11} = ([\mathbf{R}(z)]_{11}[\mathbf{N}(z)]_{11} + [\mathbf{R}(z)]_{12}[\mathbf{N}(z)]_{21}) e^{nQ(z) + \frac{n}{2}(V(z) - \ell^*)}.$$

Therefore, it follows from (8.5) and (8.16) that

$$(8.19) \quad \psi_n(z) e^{-nQ(z)} = (1 + \mathcal{O}_{\varepsilon}(n^{-1})) D^{N_n - n}(z) \frac{M_{n,1}^{(0)}(z)}{M_{n,1}^{(0)}(\infty)} + \mathcal{O}_{\varepsilon}(n^{-1}) \frac{D^{N_n - n}(z)}{\mathcal{D}^{2(n-N_n)}(\infty)} \frac{M_{n,0}^{(0)}(z)}{M_{n,0}^{(1)}(\infty)}.$$

Asymptotic formula (4.14) now follows from (7.28) and (7.29), boundedness of  $\{N_n - n\}$ , and definitions of  $M_{n,1}(z)$  in (7.26), of  $\Theta_n(z)$  right before (7.23), and of  $\vartheta_n(z)$  right after Proposition 4.4.

Now, let  $K$  be any compact set in  $\mathbb{C} \setminus J_t$ . Write  $K = K_1 \cup K_2$ , where  $K_1, K_2$  are compact,  $K_1$  does not intersect  $\Gamma$  and  $K_2$  lies entirely within the region  $\{\operatorname{Re}(\phi_{b_2}(z)) < 0\}$ , see Figures 6 and 14. Again, the lens  $\Sigma$  can be adjusted so that  $K_1$  lies in the unbounded component of the complement of  $\Sigma$ . Hence, the estimate (4.14) on  $K_1$  follows as before. To obtain it on  $K_2$ , recall that we had a lot of freedom in choosing  $\Gamma$  away from  $J_t$ . That is,  $\Gamma$  can be deformed into  $\Gamma'$  that avoids  $K_2$  and belongs to  $\{\operatorname{Re}(\phi_{b_2}(z)) < 0\}$  away from  $J_t$ . Then **RHP-Y**, formulated on  $\Gamma'$ , can be solved exactly as before since estimates (8.14) and (8.15) remain the same (with a possibly modified constant  $C_D$ ), and therefore (4.14) can be shown via (8.18)–(8.19).

Finally, take  $K \subset J_t^{\circ}$ . It again follows from (8.1), (8.2), (8.4), and (8.17) that

$$P_n(s) = [\mathbf{Y}(s)]_{11} = \left( [\mathbf{R}(s)]_{11} \left( [\mathbf{N}(s)]_{11+} + [\mathbf{N}(s)]_{12+} e^{-n\phi_{b_2}(s)-(n-N_n)V(s)} \right) + [\mathbf{R}(s)]_{12} \left( [\mathbf{N}(s)]_{21+} + [\mathbf{N}(s)]_{22+} e^{-n\phi_{b_2}(s)-(n-N_n)V(s)} \right) \right) e^{ng+(s)}, \quad s \in K.$$

Now, (6.7) and **RHP-N**(b) yield that

$$[\mathbf{N}(s)]_{i2+} e^{-n\phi_{b_2}(s)-(n-N_n)V(s) + ng+(s)} = [\mathbf{N}(s)]_{i1-} e^{ng-(s)}$$

for  $s \in J_t^{\circ}$  and  $i \in \{1, 2\}$ . Hence, we get from (6.3) and (8.16) that

$$\begin{aligned} \psi_n(s) &= (1 + \mathcal{O}_{\varepsilon}(n^{-1})) \left( [\mathbf{N}(s)]_{11+} e^{nQ+(s)} + [\mathbf{N}(s)]_{11-} e^{nQ-(s)} \right) + \\ &\quad \mathcal{O}_{\varepsilon}(n^{-1}) \left( [\mathbf{N}(s)]_{21+} e^{nQ+(s)} + [\mathbf{N}(s)]_{21-} e^{nQ-(s)} \right) \end{aligned}$$

for  $s \in K$ . Since the traces  $2Q_{\pm}(s) = \phi_{b_2\pm}(s)$ ,  $s \in J_t$ , are purely imaginary by (6.6), the above asymptotic formula yields (4.15) in the same way (8.19) yielded (4.14).

### 9. RECURRENCE COEFFICIENTS

In this section, we will use the notation  $P_n(z; t, N) = z^n + \sum_{k=0}^{n-1} (P_{n,t,N})_k z^k$  and

$$\mathbf{K}(z) = \mathbf{I} + z^{-1}\mathbf{K}_1(n, t, N) + z^{-2}\mathbf{K}_2(n, t, N) + \mathcal{O}(z^{-3}) \quad \text{as } z \rightarrow \infty,$$

where  $\mathbf{K} \in \{\mathbf{M}, \mathbf{N}, \mathbf{R}, \mathbf{T}, \mathbf{Y}z^{-n\sigma_3}\}$ . To begin, recall (1.13) and (1.14). That is,

$$(9.1) \quad \begin{cases} \gamma_n^2(t, N) &= h_n(t, N)/h_{n-1}(t, N), \\ \beta_n(t, N) &= (P_{n,t,N})_{n-1} - (P_{n+1,t,N})_n, \end{cases}$$

where for the second formula in (9.1) to be valid it must hold that  $\deg(P_{n+1}(\cdot; t, N)) = n + 1$ , see the discussion after (1.14). It follows from the orthogonality relations (1.10) and (1.12) that

$$\begin{aligned} z^n (CP_n e^{-NV})(z) &= -\frac{1}{z} \frac{1}{2\pi i} \int_{\Gamma} s^n P_n e^{-NV} ds - \frac{1}{z^2} \frac{1}{2\pi i} \int_{\Gamma} s^{n+1} P_n e^{-NV} ds + \mathcal{O}(z^{-3}) \\ &= \frac{h_n(t, N)}{2\pi i} \left( -\frac{1}{z} + \frac{(P_{n+1,t,N})_n}{z^2} \right) + \mathcal{O}(z^{-3}), \end{aligned}$$

where the second expression for the coefficient next to  $z^{-2}$  is obtained under the condition that  $\deg(P_{n+1}(\cdot; t, N)) = n + 1$  (the first expression, of course, is valid regardless this condition) and is based on the observation

$$\int_{\Gamma} s^{n+1} P_n e^{-NV} ds = \int_{\Gamma} (s^{n+1} - P_{n+1}) P_n e^{-NV} ds = (P_{n+1,N})_n \int_{\Gamma} s^n P_n e^{-NV} ds,$$

where all polynomials correspond to the same parameter  $N$ . Hence, in the considered case, it follows from (8.1) that

$$(9.2) \quad \mathbf{Y}_1(n, t, N) = \begin{pmatrix} (P_{n,t,N})_{n-1} & -\frac{h_n(t, N)}{2\pi i} \\ -\frac{2\pi i}{h_{n-1}(t, N)} & -(P_{n,t,N})_{n-1} \end{pmatrix} \quad \text{and} \quad \mathbf{Y}_2(n, N) = \begin{pmatrix} * & \frac{h_n(t, N)}{2\pi i} (P_{n+1,t,N})_n \\ * & * \end{pmatrix}.$$

This, in turn, yields the formulae

$$(9.3) \quad \begin{cases} h_n(t, N) &= -2\pi i [\mathbf{Y}_1(n, t, N)]_{12}, \\ \gamma_n^2(t, N) &= [\mathbf{Y}_1(n, t, N)]_{12} [\mathbf{Y}_1(n, t, N)]_{21}, \\ \beta_n(t, N) &= [\mathbf{Y}_1(n, t, N)]_{11} + [\mathbf{Y}_2(n, t, N)]_{12} [\mathbf{Y}_1(n, t, N)]_{12}^{-1}. \end{cases}$$

Assume now that  $\deg(P_{n+1}(\cdot; t^*, N)) = n$  for some value  $t = t^*$ . As explained after (1.14),  $\beta_n(t^*, N) = \infty$  in this case. Moreover, we also pointed out after (1.13) that if  $\deg(P_k(\cdot; t, N)) = k$  for  $k \in \{n, n+1\}$ , then  $\beta_n(t, N)$  is finite. Hence, for  $n \in \mathbb{N}(t, \varepsilon)$ ,  $P_{n+1}(z; t, N)$  degenerates if and only if  $\beta_n(t, N)$  is infinite. Thus, due to meromorphic dependence of  $\beta_n(t, N)$  on the parameter  $t$ , if  $P_{n+1}(z; t^*, N)$  degenerates, then  $\deg(P_{n+1}(\cdot; t, N)) = n + 1$  in some small punctured neighborhood of  $t^*$ . Moreover, it follows from Theorem 3.2 and (7.9) that  $z_{n,1}(t)$  is real analytic in real and imaginary parts of  $t$ . Hence, if  $n \in \mathbb{N}(t^*, \varepsilon)$ , then  $n \in \mathbb{N}(t, \varepsilon)$  for some sequence of values  $t$  converging to  $t^*$ . Clearly, along this sequence the third relation in (9.3) remains valid. Since the entries of  $\mathbf{Y}(z)$  meromorphically depend on  $t$ , then the third relation in (9.3) must remain valid at  $t^*$  as well (this simply means that  $\beta_n(t, N) = \infty$  corresponds to  $[\mathbf{Y}_2(n, t, N)]_{12} \neq 0$  and  $[\mathbf{Y}_1(n, t, N)]_{12} = 0$ ). Hence, (9.3) holds regardless the degree of  $P_{n+1}(z; t, N)$  and we drop indicating the dependence on  $t$  altogether.

As before, let  $\mu_1 = \mu_1(t)$  be the first moment of the equilibrium measure  $\mu_t$  (hopefully this slight abuse of notation will not cause a confusion). Then (6.2) implies that

$$(9.4) \quad e^{-ng(z)\sigma_3} = z^{-n\sigma_3} \left[ \mathbf{I} + \frac{n\mu_1}{z} \sigma_3 + \mathcal{O}\left(\frac{1}{z^2}\right) \right],$$



where  $\mathcal{O}(\cdot)$  is a diagonal matrix. Combining (9.4) with (8.2) yields

$$(9.5) \quad \begin{cases} [\mathbf{T}_1(n, N)]_{12} &= e^{n\ell_*} [\mathbf{Y}_1(n, N)]_{12}, & \left\{ \begin{aligned} [\mathbf{T}_1(n, N)]_{11} &= [\mathbf{Y}_1(n, N)]_{11} + n\mu_1, \\ [\mathbf{T}_1(n, N)]_{21} &= e^{-n\ell_*} [\mathbf{Y}_1(n, N)]_{21}, \end{aligned} \right. \\ [\mathbf{T}_2(n, N)]_{12} &= e^{n\ell_*} ([\mathbf{Y}_2(n, N)]_{12} - n\mu_1 [\mathbf{Y}_1(n, N)]_{12}). \end{cases}$$

This, in turn, leads to

$$(9.6) \quad \begin{cases} h_n(N) &= -2\pi i e^{-n\ell_*} [\mathbf{T}_1(n, N)]_{12}, \\ \gamma_n^2(N) &= [\mathbf{T}_1(n, N)]_{12} [\mathbf{T}_1(n, N)]_{21}, \\ \beta_n(N) &= [\mathbf{T}_1(n, N)]_{11} + [\mathbf{T}_2(n, N)]_{12} [\mathbf{T}_1(n, N)]_{12}^{-1}. \end{cases}$$

It was shown in [14, Theorem 7.10] that  $\mathbf{R}(z)$  admits an asymptotic expansion

$$(9.7) \quad \mathbf{R}(z) = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{R}^{(k)}(z)}{n^k}, \quad n \rightarrow \infty \quad \text{and} \quad z \in \mathbb{C} \setminus \cup_e \partial U_e.$$

The matrices  $\mathbf{R}^{(k)}(z)$  satisfy an additive Riemann-Hilbert Problem:

- (a)  $\mathbf{R}^{(k)}(z)$  is analytic in  $\mathbb{C} \setminus \cup_e \partial U_e$ ;
- (b)  $\mathbf{R}_+^{(k)}(z) = \mathbf{R}_-^{(k)}(z) + \sum_{j=1}^k \mathbf{R}_-^{(k-j)}(z) \Delta_j(z)$  for  $z \in \cup_e \partial U_e$ ;
- (c) as  $z \rightarrow \infty$ ,  $\mathbf{R}^{(k)}(z)$  admits an expansion of the form

$$(9.8) \quad \mathbf{R}^{(k)}(z) = \frac{\mathbf{R}_1^{(k)}}{z} + \frac{\mathbf{R}_2^{(k)}}{z^2} + \mathcal{O}\left(\frac{1}{z^3}\right).$$

Recall that  $\mathbf{T}(z) = (\mathbf{R}\mathbf{N})(z)$  for  $z$  in the vicinity of infinity by (8.2) and (8.18). It also follows from (8.5), (7.28) and (7.29) that matrices  $\mathbf{N}(z)$  form a normal family in  $n \in \mathbb{N}(\varepsilon)$  in the vicinity of infinity. Plugging expansions (9.7) and (9.8) into the product  $\mathbf{T}(z) = (\mathbf{R}\mathbf{N})(z)$  yields

$$(9.9) \quad \begin{cases} \mathbf{T}_1(n, N) &= \mathbf{N}_1(n, N) + \frac{\mathbf{R}_1^{(1)}}{n} + \frac{\mathbf{R}_1^{(2)}}{n^2} + \mathcal{O}_\varepsilon\left(\frac{1}{n^3}\right), \\ \mathbf{T}_2(n, N) &= \mathbf{N}_2(n, N) + \frac{\mathbf{R}_1^{(1)} \mathbf{N}_1(n, N) + \mathbf{R}_2^{(1)}}{n} + \frac{\mathbf{R}_1^{(2)} \mathbf{N}_1(n, N) + \mathbf{R}_2^{(2)}}{n^2} + \mathcal{O}_\varepsilon\left(\frac{1}{n^3}\right), \end{cases}$$

where we used the normality of  $\mathbf{N}(z)$  to deduce uniform boundedness of the entries of  $\mathbf{N}_k(n, N)$ ,  $k \in \{1, 2\}$ , to maintain  $\mathcal{O}_\varepsilon(n^{-3})$  as the order of the error terms (it is, of course, now depends on  $\varepsilon$  as do the bounds for the entries  $\mathbf{N}_k(n, N)$ ).

To use (9.9), we move on to computing  $\mathbf{N}_k(n, N)$ . To that end, it follows from (4.4) that

$$\mathcal{D}^{(N-n)\sigma_3}(z) = \mathcal{D}^{(N-n)\sigma_3}(\infty) \left( \mathbf{I} + \frac{(N-n)D_1}{3z} \sigma_3 + \mathcal{O}\left(\frac{1}{z^2}\right) \right)$$

as  $z \rightarrow \infty$ , where  $\mathcal{O}(z^{-2})$  is a diagonal matrix and  $D_1 := t^2 + \mu_1 - 3\zeta d_2/2$ . Further, recall the function  $\gamma(z)$  from (7.10). Then

$$\gamma(z) = 1 - \frac{S_1}{4z} + \frac{C - 3S_1^2/8}{4z^2} + \mathcal{O}(z^{-3}) \quad \text{and} \quad \gamma^{-1}(z) = 1 + \frac{S_1}{4z} - \frac{C - 5S_1^2/8}{4z^2} + \mathcal{O}(z^{-3})$$

as  $z \rightarrow \infty$ , where  $S_k = \sum_{i=1}^2 (b_i^k - a_i^k)$  and  $C = a_1^2 + a_2^2 + b_1 b_2 + a_1 a_2 - \sum_{i,j=1}^2 b_i a_j$ . Then it follows from (7.11) that

$$A(z) = 1 + \mathcal{O}(z^{-2}) \quad \text{and} \quad B(z) = \frac{S_1}{4i} \frac{1}{z} + \frac{S_2}{8i} \frac{1}{z^2} + \mathcal{O}(z^{-3})$$

as  $z \rightarrow \infty$  since  $S_1^2 - 2C = S_2$ . The above power series expansions, along with (7.26) and (8.5), yield that

$$(9.10) \quad N_1(n, N) = \frac{(N-n)D_1}{3}\sigma_3 + \begin{pmatrix} \left. \frac{d}{dz} \left( \log \Theta_{n,1}^{(0)}(1/z) \right) \right|_{z=0} & -\mathcal{D}^{2(n-N)}(\infty) \frac{S_1}{4i} \frac{\Theta_{n,1}^{(1)}(\infty)}{\Theta_{n,1}^{(0)}(\infty)} \\ \mathcal{D}^{2(N-n)}(\infty) \frac{S_1}{4i} \frac{\Theta_{n,0}^{(0)}(\infty)}{\Theta_{n,0}^{(1)}(\infty)} & \left. \frac{d}{dz} \left( \log \Theta_{n,0}^{(1)}(1/z) \right) \right|_{z=0} \end{pmatrix}$$

and

$$(9.11) \quad [N_2(n, N)]_{12} = \frac{\mathcal{D}^{2(n-N)}(\infty) S_1}{\Theta_{n,1}^{(0)}(\infty)} \frac{1}{4i} \left[ \left( \frac{(N-n)D_1}{3} - \frac{S_2}{2S_1} \right) \Theta_{n,1}^{(1)}(\infty) - \left. \frac{d}{dz} \left( \Theta_{n,1}^{(1)}(1/z) \right) \right|_{z=0} \right].$$

Using (9.10) and (9.9) in (9.6) gives

$$h_n(N_n) = \frac{\pi S_1}{2} e^{-n\ell^*} \mathcal{D}^{2(n-N)}(\infty) \frac{\Theta_{n,1}^{(1)}(\infty)}{\Theta_{n,1}^{(0)}(\infty)} + \mathcal{O}_\varepsilon \left( \frac{1}{n} \right)$$

as  $\mathbb{N}(\varepsilon) \ni n \rightarrow \infty$ , which proves the first formula in (4.16) (recall (7.24)). Similarly,

$$\gamma_n^2(t, N) = \frac{S_1^2 \Theta_{n,1}^{(1)}(\infty) \Theta_{n,0}^{(0)}(\infty)}{16 \Theta_{n,1}^{(0)}(\infty) \Theta_{n,0}^{(1)}(\infty)} + \mathcal{O}_\varepsilon \left( \frac{1}{n} \right) \quad \text{as } \mathbb{N}(\varepsilon) \ni n \rightarrow \infty,$$

which finishes the proof of (4.16). Define  $\Theta_{n-1,1}(z; N_n)$  as in (7.23) but with  $\mathbf{z}_{n-1,1}(N_{n-1})$  replaced by  $\mathbf{z}_{n-1,1}(N_n)$ . It follows from (7.25), the last claim of Proposition 7.1, and (7.34) that  $\Theta_{n,0}(z)$  is a constant multiple of  $\Theta(z)\Theta_{n-1,1}(z; N_n)$ . Then we get from (7.35) that

$$(9.12) \quad \gamma_n^2(t, N) = e^{-\ell^*} \mathcal{D}^2(\infty) \frac{\Theta_{n,1}^{(1)}(\infty) \Theta_{n-1,1}^{(0)}(\infty; N_n)}{\Theta_{n,1}^{(0)}(\infty) \Theta_{n-1,1}^{(1)}(\infty; N_n)} + \mathcal{O}_\varepsilon \left( \frac{1}{n} \right) \quad \text{as } \mathbb{N}(\varepsilon) \ni n \rightarrow \infty.$$

Finally, (9.6) and (9.9) yield

$$\beta_n(N_n) = \frac{[N_1]_{11}[N_1]_{12} + [N_2]_{12} + \mathcal{O}_\varepsilon(n^{-1})}{[N_1]_{12} + \mathcal{O}_\varepsilon(n^{-1})},$$

which, in view of (9.10) and (9.11) gives

$$\beta_n(N_n) = \frac{\Theta_{n,1}^{(1)}(\infty) \left( \frac{S_2}{2S_1} + \left. \frac{d}{dz} \left( \log \Theta_{n,1}^{(0)}(1/z) + \log \Theta_{n,1}^{(1)}(1/z) \right) \right|_{z=0} \right) + \mathcal{O}_\varepsilon(n^{-1})}{\Theta_{n,1}^{(1)}(\infty) + \mathcal{O}_\varepsilon(n^{-1})},$$

as claimed in (4.17) (we need to divide both numerator and denominator by  $\Theta_{n,1}^{(0)}(\infty)$ ).

## REFERENCES

- [1] G. Álvarez, L. Martínez Alonso, and E. Medina. Determination of  $S$ -curves with applications to the theory of non-hermitian orthogonal polynomials. *J. Stat. Mech.* (2013) P06006. 7
- [2] G. Álvarez, L. Martínez Alonso, and E. Medina. Phase structure and asymptotic zero densities of orthogonal polynomials in the cubic model. *J. Comput. Appl. Math.*, **284**, (2015), 10–25. 7
- [3] M. Bertola and A. Tovbis. On asymptotic regimes of orthogonal polynomials with complex varying quartic exponential weight. *SIGMA*, **12**, (2016), 118, 50 pages. 24
- [4] D. Bessis, X. Itzykson, and J. B. Zuber. Quantum Field Theory Techniques in Graphical Enumeration. *Adv. Appl. Math.* **1** (1980), 109–157. 3
- [5] P. M. Bleher and A. Deaño. Topological expansion in the cubic random matrix model. *Int. Math. Res. Not.* **12** (2013), 2699–2755. 1, 2, 3, 4, 5
- [6] P. M. Bleher and A. Deaño. Painlevé I double scaling limit in the cubic matrix model. *Random Matrices Theory Appl.* **5**, 2 (2016), 1650004, 58. 1, 3, 4, 5

- [7] P. Bleher, A. Deaño, and M. Yattselev. Topological expansion in the complex cubic log-gas model. One-cut case. *J. Statist. Phys.*, **166**(3–4), (2017), 784–827. [1](#), [4](#), [5](#), [7](#), [8](#), [16](#), [17](#)
- [8] P. Bleher and K. Liechty. *Random Matrices and the Six-Vertex Model (CRM Monograph)*. American Mathematical Society, 2015, 224 pages, ISBN-13: 978-1470409616 [1](#)
- [9] E. Brézin, C. Itzykson, G. Parisi, and J.-B. Zuber. Planar diagrams. *Commun. Math. Phys.* **59** (1978), 35–51. [3](#)
- [10] F. David. Phases of the large  $N$  matrix model and non-perturbative effects in 2D gravity. *Nucl. Phys. B* **348** (1991), 507–524. [4](#), [7](#)
- [11] F. David. Non-perturbative effects in matrix models and vacua of two dimensional gravity *Phys. Lett. B* **302** (1993), 403–410. [4](#)
- [12] P. Deift. *Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach*. American Mathematical Society, 2000. [39](#)
- [13] P. Deift, T. Kriecherbauer, K. T.–R. McLaughlin, S. Venakides, and X. Zhou. 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), 1335–1425. [37](#)
- [14] P. Deift, T. Kriecherbauer, K. T.–R. McLaughlin, S. Venakides, and X. Zhou. Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52** (1999), 1491–1552. [37](#), [41](#)
- [15] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation. *Ann. Math.* **137** (1993), 295–368. [34](#)
- [16] P. Di Francesco. Matrix model combinatorics: Applications to folding and coloring. *Random Matrices and Their Applications (P. M. Bleher and A. R. Its, eds.)*. Mathematical Sciences Research Institute Publications, Vol. 40, 2001. [3](#)
- [17] F.J. Dyson. Correlations between the eigenvalues of a random matrix, *Commun. Math. Phys.* **19** (1970) 235–250. [3](#)
- [18] N.M. Ercolani and K.T-R. McLaughlin. Asymptotics of the partition function for random matrices via Riemann–Hilbert techniques, and applications to graphical enumeration. *Int. Math. Res. Not.* **14** (2003), 755–820. [3](#)
- [19] N. M. Ercolani, K. T.–R. McLaughlin, and V. U. Pierce. Random matrices, graphical enumeration and the continuum limit of the Toda lattices, *Comm. Math. Phys.* **278** (2008) 31–81. [3](#)
- [20] B. Eynard. Counting Surfaces, CRM Aisenstadt Chair lectures, Progress in Mathematical Physics **70**, Birkhäuser, Basel, 2016. [3](#)
- [21] A.S. Fokas, A.R. Its, and A.V. Kitaev. The isomonodromy approach to matrix models in 2D quantum gravity. *Comm. Math. Phys.* **147** (1992), 395–430. [34](#)
- [22] P. Forrester. *Log-Gases and Random Matrices*, volume 34 of *The London Mathematical Society Monographs Series*. Princeton University Press, 2010. [3](#)
- [23] F.D. Gakhov. *Boundary Value Problems*. Dover Publications, Inc., New York, 1990. [23](#)
- [24] A.A. Gonchar and E.A. Rakhmanov. Equilibrium distributions and the degree of rational approximation of analytic functions. *Mat. Sb.* **134(176)** (1987), no. 3, 306–352, 447 English transl. in *Math. USSR-Sb.* **62** (1989), no 2, 305–348. [6](#)
- [25] D. Huybrechs, A. Kuijlaars, and N. Lejon. Zero distribution of complex orthogonal polynomials with respect to exponential weights. *J. Approx. Theory*, **184** (2014), 28–54. [5](#), [6](#), [16](#)
- [26] J. Jenkins. *Univalent functions and conformal maps*. Springer, 1965. [7](#), [14](#), [15](#), [24](#)
- [27] N. Joshi, A. V. Kitaev. On Boutroux’s tritronquée solutions of the first Painlevé equation. *Stud. Appl. Math.* **107**, **3** (2001), 253–291. [3](#)
- [28] A. A. Kapaev. Quasi-linear Stokes phenomenon for the Painlevé first equation. *J. Phys. A* **37**, **46** (2004), 11149–11167. [3](#)
- [29] A. Kuijlaars and G. Silva. S-curves in Polynomial External Fields. *J. Approx. Theory*, **191** (2015), 1–37. [5](#), [6](#)
- [30] M.L. Mehta. *Random Matrices*. Second Edition, Academic Press, San Diego e.a., 1991. [3](#)
- [31] M. Mulase. Lectures on the asymptotic expansion of a Hermitian matrix integral. *Lecture Notes in Physics* 502, 91–134 (H. Aratyn et al., eds). Springer Verlag, 1998. [3](#)
- [32] Ch. Pommerenke. *Univalent Functions*, Vandenhoeck & Ruprecht, Göttingen, 1975. [7](#), [14](#)
- [33] E.B. Saff and V. Totik. *Logarithmic Potentials with External Fields*, volume 316 of *Grundlehren der Math. Wissenschaften*. Springer-Verlag, Berlin, 1997. [6](#)
- [34] H. Stahl. Extremal domains associated with an analytic function. I, II. *Complex Variables Theory Appl.* **4** (1985), 311–324, 325–338. [6](#)
- [35] H. Stahl. Structure of extremal domains associated with an analytic function. *Complex Variables Theory Appl.* **4** (1985), 339–356. [6](#)
- [36] H. Stahl. Orthogonal polynomials with complex valued weight function. I, II. *Constr. Approx.* **2** (1986), no. 3, 225–240, 241–251. [6](#)
- [37] K. Strebel. *Quadratic differentials*. Springer, 1984. [7](#), [14](#)
- [38] S.P. Suetin. Convergence of Chebyshev continued fractions for elliptic functions. *Mat. Sb.*, 194(12):63–92, 2003. English transl. in *Math. Sb.* 194(12):1807–1835, 2003. [28](#)
- [39] G. ’t Hooft. A Planar Diagram Theory for Strong Interactions. *Nucl. Phys. B* **72** (1974), 461–473. [3](#)
- [40] A. Zvonkin. Matrix Integrals and Map Enumeration: An Accessible Introduction. *Mathl. Comput. Modelling* **26**, **8-10** (1997), 281–304. [3](#)
- [41] D. Mumford and C. Musili. *Tata Lectures on Theta. I (Modern Birkhäuser classics)*. Birkhäuser Boston Incorporated, 2007. [27](#)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, EAST HALL, 530 CHURCH STREET, ANN ARBOR, MI 48109, USA

*Email address:* `barhouri@umich.edu`

DEPARTMENT OF MATHEMATICAL SCIENCES, INDIANA UNIVERSITY-PURDUE UNIVERSITY INDIANAPOLIS, 402 NORTH BLACKFORD STREET, INDIANAPOLIS, IN 46202, USA

*Email address:* `pbleher@iupui.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSIDAD CARLOS III DE MADRID, AVDA. DE LA UNIVERSIDAD 30, 28911 LEGANÉS, MADRID, SPAIN

*Email address:* `alfredo.deanho@uc3m.es`

DEPARTMENT OF MATHEMATICAL SCIENCES, INDIANA UNIVERSITY-PURDUE UNIVERSITY INDIANAPOLIS, 402 NORTH BLACKFORD STREET, INDIANAPOLIS, IN 46202, USA

*Email address:* `maxyatts@iupui.edu`