LONG-TIME ASYMPTOTICS OF THE PERIODIC TODA LATTICE UNDER SHORT-RANGE PERTURBATIONS

SPYRIDON KAMVISSIS AND GERALD TESCHL

Abstract. We consider the long-time asymptotics of periodic (and slightly more generally of algebro-geometric finite-gap) solutions of the doubly infinite Toda lattice under a short-range perturbation. We prove that the perturbed lattice asymptotically approaches a modulated lattice.

More precisely, let g be the genus of the hyperelliptic curve associated with the unperturbed solution. We show that, apart from the phenomenon of solitons travelling on the quasi-periodic background, the n/t-plane contains $g+2$ areas where the perturbed solution is close to a finite-gap solution on the same isospectral torus. In between there are $g+1$ regions where the perturbed solution is asymptotically close to a modulated lattice which undergoes a continuous phase transition (in the Jacobian variety) and which interpolates between these isospectral solutions. In the special case of the free lattice ($g=0$) the isospectral torus consists of just one point and we recover the known result.

Both the solutions in the isospectral torus and the phase transition are explicitly characterized in terms of Abelian integrals on the underlying hyperelliptic curve.

Our method relies on the equivalence of the inverse spectral problem to a vector Riemann–Hilbert problem defined on the hyperelliptic curve and generalizes the so-called nonlinear stationary phase/steepest descent method for Riemann–Hilbert problem deformations to Riemann surfaces.

1. Introduction

A classical result going back to Zabusky and Kruskal [46] states that a decaying (fast enough) perturbation of the constant solution of a soliton equation eventually splits into a number of "solitons": localized travelling waves that preserve their shape and velocity after interaction, plus a decaying radiation part. This is the motivation for the result presented here. Our aim is to investigate the case where the constant background solution is replaced by a periodic one. We provide the detailed analysis in the case of the Toda lattice though it is clear that our methods apply to other soliton equations as well.

In the case of the Korteweg–de Vries equation the asymptotic result was first shown by Šabat [37] and by Tanaka [40]. Precise asymptotics for the radiation part were first formally derived by Zakharov and Manakov [45] and by Ablowitz and Segur [1], [38] with further extensions by Buslaev and Sukhanov [5]. A detailed rigorous justification not requiring any a priori information on the asymptotic form of the solution was first given by Deift and Zhou [6] for the case of the modified

2000 *Mathematics Subject Classification.* Primary 37K40, 37K45; Secondary 35Q15, 37K10.

Key words and phrases. Riemann–Hilbert problem, Toda lattice.

J. Math. Phys. **53**, 073706 (2012).

Research supported in part by the ESF programme MISGAM, and the Austrian Science Fund (FWF) under Grant No. P17762 and Y330.

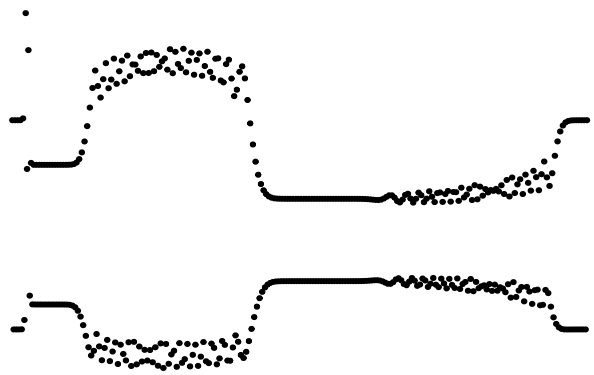

FIGURE 1. Numerically computed solution of the Toda lattice, with initial condition a period two solution perturbed at one point in the middle.

Korteweg–de Vries equation, inspired by earlier work of Manakov [31] and Its [19] (see also [20], [21], [22]). For further information on the history of this problem we refer to the survey by Deift, Its, and Zhou [8].

A naive guess would be that the perturbed periodic lattice approaches the unperturbed one in the uniform norm. However, as pointed out in [25] this is wrong: In Figure 1 the two observed lines express the variables $a(n,t)$ of the Toda lattice (see (1.1) below) at a frozen time t. In areas where the lines seem to be continuous this is due to the fact that we have plotted a huge number of particles and also due to the 2-periodicity in space. So one can think of the two lines as the even- and odd-numbered particles of the lattice. We first note the single soliton which separates two regions of apparent periodicity on the left. Also, after the soliton, we observe three different areas with apparently periodic solutions of period two. Finally there are some transitional regions in between which interpolate between the different period two regions. It is the purpose of this paper to give a rigorous and complete mathematical explanation of this picture. This will be done by formulating the inverse spectral problem as a vector Riemann–Hilbert problem on the underlying hyperelliptic curve and extending the nonlinear steepest descent method to this new setting. While Riemann–Hilbert problem on Riemann surfaces have been considered in detail before, see for example the monograph by Rodin [36], we extend this theory as well (see e.g. our novel solution formula for scalar Riemann–Hilbert problems in Theorem 4.3).

Consider the doubly infinite Toda lattice in Flaschka's variables (see e.g. [15], [41], [42], or [44])

$$\begin{aligned} \dot b(n,t) &= 2(a(n,t)^2 - a(n-1,t)^2), \\ \dot a(n,t) &= a(n,t)(b(n+1,t) - b(n,t)), \end{aligned} \tag{1.1}$$

$(n,t) \in \mathbb{Z} \times \mathbb{R}$, where the dot denotes differentiation with respect to time.

In case of a constant background the long-time asymptotics were first computed by Novokshenov and Habibullin [34] and later made rigorous by Kamvissis [23] under the additional assumption that no solitons are present. The full case (with solitons) was only recently presented by Krüger and Teschl in [28] (for a review see also [29]).

Here we will consider a quasi-periodic algebro-geometric background solution (a_q, b_q), to be described in the next section, plus a short-range perturbation (a, b) satisfying

$$\sum_{n\in\mathbb{Z}} n^6(|a(n,t) - a_q(n,t)| + |b(n,t) - b_q(n,t)|) < \infty \tag{1.2}$$

for $t = 0$ and hence for all (see e.g. [11]) $t \in \mathbb{R}$. The perturbed solution can be computed via the inverse scattering transform. The case where (a_q, b_q) is constant is classical (see again [15], [41] or [44]), while the more general case we want here was solved only recently in [11] (see also [32]).

To fix our background solution, consider a hyperelliptic Riemann surface of genus g with real moduli $E_0, E_1, ..., E_{2g+1}$. Choose a Dirichlet divisor $\mathcal{D}_{\underline{\hat{\mu}}}$ and introduce

$$\underline{z}(n,t) = \underline{\hat{A}}_{p_0}(\infty_+) - \underline{\hat{\alpha}}_{p_0}(\mathcal{D}_{\underline{\hat{\mu}}}) - n\underline{\hat{A}}_{\infty_-}(\infty_+) + t\underline{U}_0 - \underline{\hat{\Xi}}_{p_0} \in \mathbb{C}^g, \tag{1.3}$$

where $\underline{A}_{p_0}$ ($\underline{\alpha}_{p_0}$) is Abel's map (for divisors) and $\underline{\Xi}_{p_0}$, $\underline{U}_0$ are some constants defined in Section 2. Then our background solution is given in terms of Riemann theta functions (defined in (2.14)) by

$$\begin{aligned} a_q(n,t)^2 &= \tilde{a}^2 \frac{\theta(\underline{z}(n+1,t))\theta(\underline{z}(n-1,t))}{\theta(\underline{z}(n,t))^2}, \\ b_q(n,t) &= \tilde{b} + \frac{1}{2}\frac{d}{dt}\log\Big(\frac{\theta(\underline{z}(n,t))}{\theta(\underline{z}(n-1,t))}\Big), \end{aligned} \tag{1.4}$$

where $\tilde{a}$, $\tilde{b} \in \mathbb{R}$ are again some constants.

We can of course view this hyperelliptic Riemann surface as formed by cutting and pasting two copies of the complex plane along bands. Having this picture in mind, we denote the standard projection to the complex plane by π.

Assume for simplicity that the Jacobi operator

$$H(t)f(n) = a(n,t)f(n+1) + a(n-1,t)f(n-1) + b(n,t)f(n), \quad f \in \ell^2(\mathbb{Z}), \tag{1.5}$$

corresponding to the perturbed problem (1.1) has no eigenvalues. In this paper we prove that for long times the perturbed Toda lattice is asymptotically close to the following limiting lattice defined by

$$\begin{aligned} \prod_{j=n}^{\infty}\Big(\frac{a_l(j,t)}{a_q(j,t)}\Big)^2 =& \frac{\theta(\underline{z}(n,t))}{\theta(\underline{z}(n-1,t))}\frac{\theta(\underline{z}(n-1,t)+\underline{\delta}(n,t))}{\theta(\underline{z}(n,t)+\underline{\delta}(n,t))}\times \\ & \times\exp\left(\frac{1}{2\pi\mathrm{i}}\int_{C(n/t)}\log(1-|R|^2)\omega_{\infty_+\,\infty_-}\right), \\ \delta_\ell(n,t) =& \frac{1}{2\pi\mathrm{i}}\int_{C(n/t)}\log(1-|R|^2)\zeta_\ell, \end{aligned} \tag{1.6}$$

where R is the associated reflection coefficient, ζ_ℓ is a canonical basis of holomorphic differentials, $\omega_{\infty_+\,\infty_-}$ is an Abelian differential of the third kind defined in (2.15), and $C(n/t)$ is a contour on the Riemann surface. More specific, $C(n/t)$ is obtained by taking the spectrum of the unperturbed Jacobi operator H_q between $-\infty$ and a special stationary phase point $z_j(n/t)$, for the phase of the underlying Riemann–Hilbert problem defined in the beginning of Section 4, and lifting it to the Riemann surface (oriented such that the upper sheet lies to its left). The point $z_j(n/t)$ will move from $-\infty$ to $+\infty$ as n/t varies from $-\infty$ to $+\infty$. From the products above, one easily recovers $a_l(n,t)$. More precisely, we have the following.

Theorem 1.1. *Let C be any (large) positive number and δ be any (small) positive number. Let $E_s \in S$ be the 'resonance points' defined by $S = \{E_s : |R(E_s)| = 1\}$. (There are at most $2g+2$ such points, since they are always endpoints E_j of the bands that constitute the spectrum of the Jacobi operator.) Consider the region $D = \{(n,t) : |\frac{n}{t}| < C\} \cap \{(n,t) : |z_j(\frac{n}{t}) - E_s| > \delta\}$, where $z_j(\frac{n}{t})$ is the special stationary phase point for the phase defined in the beginning of Section 4. Then one has*

$$\prod_{j=n}^{\infty} \frac{a_l(j,t)}{a(j,t)} \to 1 \tag{1.7}$$

uniformly in D, as $t \to \infty$.

The proof of this theorem will be given in Section 4 of this paper.

Remark 1.2. *(i) It is easy to see how the asymptotic formula above describes the picture given by the numerics. Recall that the spectrum $\sigma(H_q)$ of H_q consists of $g+1$ bands whose band edges are the branch points of the underlying hyperelliptic Riemann surface. If $\frac{n}{t}$ is small enough, $z_j(n/t)$ is to the left of all bands implying that $C(n/t)$ is empty and thus $\delta_\ell(n,t) = 0$; so we recover the purely periodic lattice. At some value of $\frac{n}{t}$ a stationary phase point first appears in the first band of $\sigma(H_q)$ and begins to move form the left endpoint of the band towards the right endpoint of the band. (More precisely we have a pair of stationary phase points z_j and z_j^*, one in each sheet of the hyperelliptic curve, with common projection $\pi(z_j)$ on the complex plane.) So $\delta_\ell(n,t)$ is now a non-zero quantity changing with $\frac{n}{t}$ and the asymptotic lattice has a slowly modulated non-zero phase. Also the factor given by the exponential of the integral is non-trivially changing with $\frac{n}{t}$ and contributes to a slowly modulated amplitude. Then, after the stationary phase point leaves the first band there is a range of $\frac{n}{t}$ for which no stationary phase point appears in the spectrum $\sigma(H_q)$, hence the phase shift $\delta_\ell(n,t)$ and the integral remain constant, so the asymptotic lattice is periodic (but with a non-zero phase shift). Eventually a stationary phase point appears in the second band, so a new modulation appears and so on. Finally, when $\frac{n}{t}$ is large enough, so that all bands have been traversed by the stationary phase point(s), the asymptotic lattice is again periodic. Periodicity properties of theta functions easily show that phase shift is actually cancelled by the exponential of the integral and we recover the original periodic lattice with no phase shift at all.*

(ii) If eigenvalues are present we can apply appropriate Darboux transformations to add the effect of such eigenvalues ([13]*). What we then see asymptotically is travelling solitons in a periodic background. Note that this will change the asymptotics on one side. In any case, our method works unaltered for such situations (cf.* [12]*) as well.*

(iii) Employing the very same methods of the paper it is very easy to show that in any region $|\frac{n}{t}| > C$, one has

$$\prod_{j=n}^{\infty} \frac{a_l(j,t)}{a(j,t)} \to 1 \tag{1.8}$$

uniformly in t, as $n \to \infty$.

(iv) The effect of the resonances E_s is only felt locally (and to higher order in $1/t$) in some small (decaying as $t \to \infty$) region, where in fact $|z_j(\frac{n}{t}) - E_s| \to 0$ as $t \to \infty$. So the above theorem is actually true in $\{(n,t) : |\frac{n}{t}| < C\}$. Near the resonances we expect both a "collisionless shock" phenomenon and a Painlevé region to appear ([9], [6], [23], [24]). A proof of this can be given using the results of [9] and [6].

(v) For the proof of Theorem 1.1 and Theorem 1.3 it would suffice to assume (1.2) with n^6 replaced by $|n|^3$ (or even $|n|$ plus the requirement that the associated reflection coefficient is Hölder continuous). Our stronger assumption is only required for the detailed decay estimates in Theorem 1.4 below.

By dividing in (1.6) one recovers the $a(n,t)$. It follows from the main Theorem and the last remark above that

$$|a(n,t) - a_l(n,t)| \to 0 \tag{1.9}$$

uniformly in D, as $t \to \infty$. In other words, the perturbed Toda lattice is asymptotically close to the limiting lattice above.

A similar theorem can be proved for the velocities $b(n,t)$.

Theorem 1.3. *In the region $D = \{(n,t) : |\frac{n}{t}| < C\} \cap \{(n,t) : |z_j(\frac{n}{t}) - E_s| > \delta\}$, of Theorem 1.1 we also have*

$$\sum_{j=n}^{\infty} \big(b_l(j,t) - b_q(j,t)\big) \to 0 \tag{1.10}$$

uniformly in D, as $t \to \infty$, where b_l is given by

$$\begin{aligned}\sum_{j=n}^{\infty} \big(b_l(j,t) - b_q(j,t)\big) =& \frac{1}{2\pi \mathrm{i}} \int_{C(n/t)} \log(1 - |R|^2)\Omega_0 \\ &+ \frac{1}{2}\frac{d}{ds} \log\left(\frac{\theta(\underline{z}(n,s) + \underline{\delta}(n,t))}{\theta(\underline{z}(n,s))}\right)\Big|_{s=t}\end{aligned} \tag{1.11}$$

and Ω_0 is an Abelian differential of the second kind defined in (2.16).

The proof of this theorem will also be given in Section 4 of this paper.

The next question we address here concerns the higher order asymptotics. Namely, what is the rate at which the perturbed lattice approaches the limiting lattice? Even more, what is the exact asymptotic formula?

Theorem 1.4. *Let D_j be the sector $D_j = \{(n,t), : z_j(n/t) \in [E_{2j} + \varepsilon, E_{2j+1} - \varepsilon]$ for some $\varepsilon > 0$. Then one has*

$$\prod_{j=n}^{\infty} \left(\frac{a(j,t)}{a_l(j,t)}\right)^2 = 1 + \sqrt{\frac{\mathrm{i}}{\phi''(z_j(n/t))t}} 2\mathrm{Re}\left(\overline{\beta(n,t)}\mathrm{i}\Lambda_0(n,t)\right) + O(t^{-\alpha}) \tag{1.12}$$

and

$$\sum_{j=n+1}^{\infty} \big(b(j,t) - b_l(j,t)\big) = \sqrt{\frac{\mathrm{i}}{\phi''(z_j(n/t))t}} 2\mathrm{Re}\left(\overline{\beta(n,t)}\mathrm{i}\Lambda_1(n,t)\right) + O(t^{-\alpha}) \tag{1.13}$$

for any $\alpha < 1$ uniformly in D_j, as $t \to \infty$. Here

$$\phi''(z_j)/\mathrm{i} = \frac{\prod_{k=0,k\neq j}^{g}(z_j - z_k)}{\mathrm{i}R_{2g+2}^{1/2}(z_j)} > 0, \tag{1.14}$$

(where $\phi(p,n/t)$ is the phase function defined in (3.17) *and $R_{2g+2}^{1/2}(z)$ the square root of the underlying Riemann surface),*

$$\Lambda_0(n,t) = \omega_{\infty_-\,\infty_+}(z_j) + \sum_{k,\ell} c_{k\ell}(\underline{\hat{\nu}}(n,t)) \int_{\infty_+}^{\infty_-} \omega_{\hat{\nu}_\ell(n,t),0}\zeta_k(z_j),$$

$$\Lambda_1(n,t) = \omega_{\infty_-,0}(z_j) - \sum_{k,\ell} c_{k\ell}(\underline{\hat{\nu}}(n,t))\omega_{\hat{\nu}_\ell(n,t),0}(\infty_+)\zeta_k(z_j), \tag{1.15}$$

with $c_{k\ell}(\underline{\hat{\nu}}(n,t))$ some constants defined in (5.14)*, $\omega_{q,0}$ an Abelian differential of the second kind with a second order pole at q (cf. Remark 5.2),*

$$\begin{aligned}\beta =& \sqrt{\nu}\mathrm{e}^{\mathrm{i}(\pi/4-\arg(R(z_j)))+\arg(\Gamma(\mathrm{i}\nu))-2\nu\alpha(z_j))}\left(\frac{\phi''(z_j)}{\mathrm{i}}\right)^{\mathrm{i}\nu}\mathrm{e}^{-t\phi(z_j)}t^{-\mathrm{i}\nu}\times\\ &\times\frac{\theta(\underline{z}(z_j,n,t)+\underline{\delta}(n,t))}{\theta(\underline{z}(z_j,0,0))}\frac{\theta(\underline{z}(z_j^*,0,0))}{\theta(\underline{z}(z_j^*,n,t)+\underline{\delta}(n,t))}\times\\ &\times\exp\left(\frac{1}{2\pi\mathrm{i}}\int_{C(n/t)}\log\left(\frac{1-|R|^2}{1-|R(z_j)|^2}\right)\omega_{p\,p^*}\right),\end{aligned} \tag{1.16}$$

where $\Gamma(z)$ is the gamma function,

$$\nu = -\frac{1}{2\pi}\log(1-|R(z_j)|^2) > 0, \tag{1.17}$$

and $\alpha(z_j)$ is a constant defined in (4.24)*.*

The proof of this theorem will be given in Section 5 of this paper. The idea of the proof is that even when a Riemann-Hilbert problem needs to be considered on an algebraic variety, a localized parametrix Riemann-Hilbert problem need only be solved in the complex plane and the local solution can then be glued to the global Riemann-Hilbert solution on the variety.

The same idea can produce the asymptotics in the two resonance regions mentioned above: a "collisionless shock" phenomenon and a Painlevé region, for every resonance pint E_s, by simply using the results of ([9], [6]). We leave the details to the reader.

Remark 1.5. *(i) The current work combines two articles that have appeared previously in the arXiv as arXiv:0705.0346 and arXiv:0805.3847 but have not been published otherwise. The necessary changes needed to include solitons are given in* [30] *which was based on arXiv:0705.0346 (see also* [13], [28]*, and* [43]*).*

(ii) Combining our technique with the one from [7] *can lead to a complete asymptotic expansion.*

(iii) Finally, we note that the same proof works even if there are different spatial asymptotics as $n \to \pm\infty$ as long as they lie in the same isospectral class (cf. [12]*).*

2. Algebro-geometric quasi-periodic finite-gap solutions

As a preparation we need some facts on our background solution (a_q, b_q) which we want to choose from the class of algebro-geometric quasi-periodic finite-gap solutions, that is the class of stationary solutions of the Toda hierarchy, [3], [17], [41]. In particular, this class contains all periodic solutions. We will use the same notation as in [41], where we also refer to for proofs. As a reference for Riemann surfaces in this context we recommend [16].

To set the stage let $\mathbb{M}$ be the Riemann surface associated with the following function

$$R_{2g+2}^{1/2}(z), \qquad R_{2g+2}(z) = \prod_{j=0}^{2g+1}(z - E_j), \qquad E_0 < E_1 < \cdots < E_{2g+1}, \tag{2.1}$$

$g \in \mathbb{N}$. $\mathbb{M}$ is a compact, hyperelliptic Riemann surface of genus g. We will choose $R_{2g+2}^{1/2}(z)$ as the fixed branch

$$R_{2g+2}^{1/2}(z) = -\prod_{j=0}^{2g+1}\sqrt{z - E_j}, \tag{2.2}$$

where $\sqrt{.}$ is the standard root with branch cut along $(-\infty, 0)$.

A point on $\mathbb{M}$ is denoted by $p = (z, \pm R_{2g+2}^{1/2}(z)) = (z, \pm)$, $z \in \mathbb{C}$, or $p = (\infty, \pm) = \infty_\pm$, and the projection onto $\mathbb{C} \cup \{\infty\}$ by $\pi(p) = z$. The points $\{(E_j, 0), 0 \le j \le 2g+1\} \subseteq \mathbb{M}$ are called branch points and the sets

$$\Pi_\pm = \{(z, \pm R_{2g+2}^{1/2}(z)) \mid z \in \mathbb{C} \setminus \bigcup_{j=0}^{g}[E_{2j}, E_{2j+1}]\} \subset \mathbb{M} \tag{2.3}$$

are called upper, lower sheet, respectively.

Let $\{a_j, b_j\}_{j=1}^g$ be loops on the surface $\mathbb{M}$ representing the canonical generators of the fundamental group $\pi_1(\mathbb{M})$. We require a_j to surround the points E_{2j-1}, E_{2j} (thereby changing sheets twice) and b_j to surround E_0, E_{2j-1} counterclockwise on the upper sheet, with pairwise intersection indices given by

$$a_i \circ a_j = b_i \circ b_j = 0, \qquad a_i \circ b_j = \delta_{i,j}, \qquad 1 \le i, j \le g. \tag{2.4}$$

The corresponding canonical basis $\{\zeta_j\}_{j=1}^g$ for the space of holomorphic differentials can be constructed by

$$\underline{\zeta} = \sum_{j=1}^{g} \underline{c}(j) \frac{\pi^{j-1} d\pi}{R_{2g+2}^{1/2}}, \tag{2.5}$$

where the constants $\underline{c}(.)$ are given by

$$c_j(k) = C_{jk}^{-1}, \qquad C_{jk} = \int_{a_k} \frac{\pi^{j-1} d\pi}{R_{2g+2}^{1/2}} = 2\int_{E_{2k-1}}^{E_{2k}} \frac{z^{j-1} dz}{R_{2g+2}^{1/2}(z)} \in \mathbb{R}. \tag{2.6}$$

The differentials fulfill

$$\int_{a_j} \zeta_k = \delta_{j,k}, \qquad \int_{b_j} \zeta_k = \tau_{j,k}, \qquad \tau_{j,k} = \tau_{k,j}, \qquad 1 \le j, k \le g. \tag{2.7}$$

Now pick g numbers (the Dirichlet eigenvalues)

$$(\hat{\mu}_j)_{j=1}^g = (\mu_j, \sigma_j)_{j=1}^g \tag{2.8}$$

whose projections lie in the spectral gaps, that is, $\mu_j \in [E_{2j-1}, E_{2j}]$. Associated with these numbers is the divisor $\mathcal{D}_{\underline{\hat{\mu}}}$ which is one at the points $\hat{\mu}_j$ and zero else. Using this divisor we introduce

$$\begin{aligned} &\underline{z}(p, n, t) = \underline{\hat{A}}_{p_0}(p) - \underline{\hat{\alpha}}_{p_0}(\mathcal{D}_{\underline{\hat{\mu}}}) - n\underline{\hat{A}}_{\infty_-}(\infty_+) + t\underline{U}_0 - \underline{\hat{\Xi}}_{p_0} \in \mathbb{C}^g, \\ &\underline{z}(n, t) = \underline{z}(\infty_+, n, t), \end{aligned} \tag{2.9}$$

where $\underline{\Xi}_{p_0}$ is the vector of Riemann constants

$$\hat{\Xi}_{p_0,j} = \frac{j + \sum_{k=1}^{g} \tau_{j,k}}{2}, \qquad p_0 = (E_0, 0), \tag{2.10}$$

$\underline{U}_0$ are the b-periods of the Abelian differential Ω_0 defined below, and $\underline{A}_{p_0}$ ($\underline{\alpha}_{p_0}$) is Abel's map (for divisors). The hat indicates that we regard it as a (single-valued) map from $\hat{\mathbb{M}}$ (the fundamental polygon associated with $\mathbb{M}$ by cutting along the a and b cycles) to $\mathbb{C}^g$. We recall that the function $\theta(\underline{z}(p,n,t))$ has precisely g zeros $\hat{\mu}_j(n,t)$ (with $\hat{\mu}_j(0,0) = \hat{\mu}_j$), where $\theta(\underline{z})$ is the Riemann theta function of $\mathbb{M}$.

Then our background solution is given by

$$\begin{aligned} a_q(n,t)^2 &= \tilde{a}^2 \frac{\theta(\underline{z}(n+1,t))\theta(\underline{z}(n-1,t))}{\theta(\underline{z}(n,t))^2}, \\ b_q(n,t) &= \tilde{b} + \frac{1}{2}\frac{d}{dt} \log\Big(\frac{\theta(\underline{z}(n,t))}{\theta(\underline{z}(n-1,t))}\Big). \end{aligned} \tag{2.11}$$

The constants $\tilde{a}$, $\tilde{b}$ depend only on the Riemann surface (see [41, Section 9.2]).

Introduce the time dependent Baker-Akhiezer function

$$\psi_q(p,n,t) = C(n,0,t)\frac{\theta(\underline{z}(p,n,t))}{\theta(\underline{z}(p,0,0))} \exp\Big(n\int_{E_0}^{p} \omega_{\infty_+\,\infty_-} + t\int_{E_0}^{p} \Omega_0\Big), \tag{2.12}$$

where $C(n,0,t)$ is real-valued,

$$C(n,0,t)^2 = \frac{\theta(\underline{z}(0,0))\theta(\underline{z}(-1,0))}{\theta(\underline{z}(n,t))\theta(\underline{z}(n-1,t))}, \tag{2.13}$$

and the sign has to be chosen in accordance with $a_q(n,t)$. Here

$$\theta(\underline{z}) = \sum_{\underline{m}\in\mathbb{Z}^g} \exp 2\pi\mathrm{i}\left(\langle\underline{m},\underline{z}\rangle + \frac{\langle\underline{m},\underline{\tau}\,\underline{m}\rangle}{2}\right), \qquad \underline{z}\in\mathbb{C}^g, \tag{2.14}$$

is the Riemann theta function associated with $\mathbb{M}$,

$$\omega_{\infty_+\,\infty_-} = \frac{\prod_{j=1}^{g}(\pi-\lambda_j)}{R_{2g+2}^{1/2}}d\pi \tag{2.15}$$

is the Abelian differential of the third kind with poles at ∞_+ and ∞_- and

$$\Omega_0 = \frac{\prod_{j=0}^{g}(\pi-\tilde{\lambda}_j)}{R_{2g+2}^{1/2}}d\pi, \qquad \sum_{j=0}^{g}\tilde{\lambda}_j = \frac{1}{2}\sum_{j=0}^{2g+1}E_j, \tag{2.16}$$

is the Abelian differential of the second kind with second order poles at ∞_+ respectively ∞_- (see [41, Sects. 13.1, 13.2]). All Abelian differentials are normalized to have vanishing a_j periods.

The Baker-Akhiezer function is a meromorphic function on $\mathbb{M}\setminus\{\infty_\pm\}$ with an essential singularity at $\infty_\pm$. The two branches are denoted by

$$\psi_{q,\pm}(z,n,t) = \psi_q(p,n,t), \qquad p = (z,\pm) \tag{2.17}$$

and it satisfies

$$\begin{aligned} H_q(t)\psi_q(p,n,t) &= \pi(p)\psi_q(p,n,t), \\ \frac{d}{dt}\psi_q(p,n,t) &= P_{q,2}(t)\psi_q(p,n,t), \end{aligned} \tag{2.18}$$

where

$$H_q(t)f(n) = a_q(n,t)f(n+1) + a_q(n-1,t)f(n-1) + b_q(n,t)f(n), \tag{2.19}$$

$$P_{q,2}(t)f(n) = a_q(n,t)f(n+1) - a_q(n-1,t)f(n-1) \tag{2.20}$$

are the operators from the Lax pair for the Toda lattice.

It is well known that the spectrum of $H_q(t)$ is time independent and consists of $g+1$ bands

$$\sigma(H_q) = \bigcup_{j=0}^{g} [E_{2j}, E_{2j+1}]. \tag{2.21}$$

For further information and proofs we refer to [41, Chap. 9 and Sect. 13.2].

3. The Inverse scattering transform and the Riemann–Hilbert problem

In this section our notation and results are taken from [10] and [11]. Let $\psi_{q,\pm}(z,n,t)$ be the branches of the Baker-Akhiezer function defined in the previous section. Let $\psi_\pm(z,n,t)$ be the Jost functions for the perturbed problem

$$a(n,t)\psi_\pm(z,n+1,t) + a(n-1,t)\psi_\pm(z,n-1,t) + b(n,t)\psi_\pm(z,n,t) = z\psi_\pm(z,n,t) \tag{3.1}$$

defined by the asymptotic normalization

$$\lim_{n\to\pm\infty} w(z)^{\mp n}(\psi_\pm(z,n,t) - \psi_{q,\pm}(z,n,t)) = 0, \tag{3.2}$$

where $w(z)$ is the quasimomentum map

$$w(z) = \exp(\int_{E_0}^{p} \omega_{\infty_+\,\infty_-}), \quad p = (z,+). \tag{3.3}$$

The asymptotics of the two projections of the Jost function are

$$\begin{aligned} \psi_\pm(z,n,t) = \psi_{q,\pm}(z,0,t) &\frac{z^{\mp n}\Big(\prod_{j=0}^{n-1} a_q(j,t)\Big)^{\pm 1}}{A_\pm(n,t)} \times \\ &\times \Big(1 + \Big(B_\pm(n,t) \pm \sum_{j=1}^{n} b_q(j - {\textstyle{0 \atop 1}}, t)\Big)\frac{1}{z} + O(\frac{1}{z^2})\Big), \end{aligned} \tag{3.4}$$

as $z \to \infty$, where

$$\begin{aligned} A_+(n,t) &= \prod_{j=n}^{\infty} \frac{a(j,t)}{a_q(j,t)}, \quad B_+(n,t) = \sum_{j=n+1}^{\infty} (b_q(j,t) - b(j,t)), \\ A_-(n,t) &= \prod_{j=-\infty}^{n-1} \frac{a(j,t)}{a_q(j,t)}, \quad B_-(n,t) = \sum_{j=-\infty}^{n-1} (b_q(j,t) - b(j,t)). \end{aligned} \tag{3.5}$$

One has the scattering relations

$$T(z)\psi_\mp(z,n,t) = \overline{\psi_\pm(z,n,t)} + R_\pm(z)\psi_\pm(z,n,t), \qquad z \in \sigma(H_q), \tag{3.6}$$

where $T(z)$, $R_\pm(z)$ are the transmission respectively reflection coefficients. Here $\psi_\pm(z,n,t)$ is defined such that $\psi_\pm(z,n,t) = \lim_{\varepsilon\downarrow 0}\psi_\pm(z+\mathrm{i}\varepsilon,n,t)$, $z \in \sigma(H_q)$. If we take the limit from the other side we have $\overline{\psi_\pm(z,n,t)} = \lim_{\varepsilon\downarrow 0}\psi_\pm(z-\mathrm{i}\varepsilon,n,t)$.

The transmission $T(z)$ and reflection $R_\pm(z)$ coefficients satisfy

$$T(z)\overline{R_+(z)} + \overline{T(z)}R_-(z) = 0, \qquad |T(z)|^2 + |R_\pm(z)|^2 = 1. \tag{3.7}$$

In particular one reflection coefficient, say $R(z) = R_+(z)$, suffices.

We will define a Riemann–Hilbert problem on the Riemann surface $\mathbb{M}$ as follows:

$$m(p,n,t) = \begin{cases} \begin{pmatrix} T(z)\psi_-(z,n,t) & \psi_+(z,n,t) \end{pmatrix}, & p = (z,+) \\ \begin{pmatrix} \psi_+(z,n,t) & T(z)\psi_-(z,n,t) \end{pmatrix}, & p = (z,-) \end{cases}. \tag{3.8}$$

Note that $m(p,n,t)$ inherits the poles at $\hat{\mu}_j(0,0)$ and the essential singularity at $\infty_\pm$ from the Baker–Akhiezer function.

We are interested in the jump condition of $m(p,n,t)$ on Σ, the boundary of $\Pi_\pm$ (oriented counterclockwise when viewed from top sheet Π_+). It consists of two copies $\Sigma_\pm$ of $\sigma(H_q)$ which correspond to non-tangential limits from $p = (z,+)$ with $\pm\mathrm{Im}(z) > 0$, respectively to non-tangential limits from $p = (z,-)$ with $\mp\mathrm{Im}(z) > 0$.

To formulate our jump condition we use the following convention: When representing functions on Σ, the lower subscript denotes the non-tangential limit from Π_+ or Π_-, respectively,

$$m_\pm(p_0) = \lim_{\Pi_\pm \ni p \to p_0} m(p), \qquad p_0 \in \Sigma. \tag{3.9}$$

Using the notation above implicitly assumes that these limits exist in the sense that $m(p)$ extends to a continuous function on the boundary away from the band edges.

Moreover, we will also use symmetries with respect to the the sheet exchange map

$$p^* = \begin{cases} (z,\mp) & \text{for } p = (z,\pm), \\ \infty_\mp & \text{for } p = \infty_\pm, \end{cases} \tag{3.10}$$

and complex conjugation

$$\overline{p} = \begin{cases} (\overline{z},\pm) & \text{for } p = (z,\pm) \notin \Sigma, \\ (z,\mp) & \text{for } p = (z,\pm) \in \Sigma, \\ \infty_\pm & \text{for } p = \infty_\pm. \end{cases} \tag{3.11}$$

In particular, we have $\overline{p} = p^*$ for $p \in \Sigma$.

Note that we have $\tilde{m}_\pm(p) = m_\mp(p^*)$ for $\tilde{m}(p) = m(p^*)$ (since $*$ reverses the orientation of Σ) and $\tilde{m}_\pm(p) = \overline{m_\pm(p^*)}$ for $\tilde{m}(p) = \overline{m(\overline{p})}$.

With this notation, using (3.6) and (3.7), we obtain

$$m_+(p,n,t) = m_-(p,n,t) \begin{pmatrix} 1 - |R(p)|^2 & -\overline{R(p)} \\ R(p) & 1 \end{pmatrix}, \tag{3.12}$$

where we have extended our definition of R to Σ such that it is equal to $R(z)$ on Σ_+ and equal to $\overline{R(z)}$ on Σ_-. In particular, the condition on Σ_+ is just the complex conjugate of the one on Σ_- since we have $R(p^*) = \overline{R(p)}$ and $m_\pm(p^*,n,t) = \overline{m_\pm(p,n,t)}$ for $p \in \Sigma$.

To remove the essential singularity at $\infty_\pm$ and to get a meromorphic Riemann–Hilbert problem we set

$$m^2(p,n,t) = m(p,n,t) \begin{pmatrix} \psi_q(p^*,n,t)^{-1} & 0 \\ 0 & \psi_q(p,n,t)^{-1} \end{pmatrix}. \tag{3.13}$$

Its divisor satisfies

$$(m_1^2) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)^*}, \qquad (m_2^2) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)}, \tag{3.14}$$

and the jump conditions become

$$\begin{aligned} m_+^2(p,n,t) &= m_-^2(p,n,t) J^2(p,n,t) \\ J^2(p,n,t) &= \begin{pmatrix} 1-|R(p)|^2 & -\overline{R(p)\Theta(p,n,t)}\mathrm{e}^{-t\phi(p)} \\ R(p)\Theta(p,n,t)\mathrm{e}^{t\phi(p)} & 1 \end{pmatrix}, \end{aligned} \tag{3.15}$$

where

$$\Theta(p,n,t) = \frac{\theta(\underline{z}(p,n,t))}{\theta(\underline{z}(p,0,0))} \frac{\theta(\underline{z}(p^*,0,0))}{\theta(\underline{z}(p^*,n,t))} \tag{3.16}$$

and

$$\phi(p,\frac{n}{t}) = 2\int_{E_0}^p \Omega_0 + 2\frac{n}{t}\int_{E_0}^p \omega_{\infty_+\,\infty_-} \in \mathrm{i}\mathbb{R} \tag{3.17}$$

for $p \in \Sigma$. Note

$$\frac{\psi_q(p,n,t)}{\psi_q(p^*,n,t)} = \Theta(p,n,t)\mathrm{e}^{t\phi(p)}.$$

Observe that

$$m^2(p) = \overline{m^2(\overline{p})}$$

and

$$m^2(p^*) = m^2(p)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which follow directly from the definition (3.13). They are related to the symmetries

$$J^2(p) = \overline{J^2(\overline{p})} \quad \text{and} \quad J^2(p) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} J^2(p^*)^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now we come to the normalization condition at ∞_+. To this end note

$$m(p,n,t) = \Big(A_+(n,t)(1 - B_+(n-1,t)\tfrac{1}{z}) \quad \tfrac{1}{A_+(n,t)}(1 + B_+(n,t)\tfrac{1}{z})\Big) + O(\frac{1}{z^2}), \tag{3.18}$$

for $p = (z,+) \to \infty_+$, with $A_\pm(n,t)$ and $B_\pm(n,t)$ are defined in (3.5). The formula near ∞_- follows by flipping the columns. Here we have used

$$T(z) = A_-(n,t)A_+(n,t)\Big(1 - \frac{B_+(n,t) + b_q(n,t) - b(n,t) + B_-(n,t)}{z} + O(\frac{1}{z^2})\Big). \tag{3.19}$$

Using the properties of $\psi(p,n,t)$ and $\psi_q(p,n,t)$ one checks that its divisor satisfies

$$(m_1) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)^*}, \qquad (m_2) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)}. \tag{3.20}$$

Next we show how to normalize the problem at infinity. The use of the above symmetries is necessary and it makes essential use of the second sheet of the Riemann surface (see also the Conclusion of this paper).

Theorem 3.1. *The function*

$$m^3(p) = \frac{1}{A_+(n,t)} m^2(p,n,t) \tag{3.21}$$

with $m^2(p,n,t)$ *defined in* (3.13) *is meromorphic away from* Σ *and satisfies:*

$$m^3_+(p) = m^3_-(p) J^3(p), \quad p \in \Sigma,$$
$$(m^3_1) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)^*}, \quad (m^3_2) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)}, \tag{3.22}$$
$$m^3(p^*) = m^3(p) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$m^3(\infty_+) = \begin{pmatrix} 1 & * \end{pmatrix}, \tag{3.23}$$

where the jump is given by

$$J^3(p,n,t) = \begin{pmatrix} 1 - |R(p)|^2 & -\overline{R(p)\Theta(p,n,t)} \mathrm{e}^{-t\phi(p)} \\ R(p)\Theta(p,n,t) \mathrm{e}^{t\phi(p)} & 1 \end{pmatrix}. \tag{3.24}$$

Setting $R(z) \equiv 0$ we clearly recover the purely periodic solution, as we should. Moreover, note

$$m^3(p) = \begin{pmatrix} \frac{1}{A_+(n,t)^2} & 1 \end{pmatrix} + \begin{pmatrix} \frac{B_+(n,t)}{A_+(n,t)^2} & -B_+(n-1,t) \end{pmatrix} \frac{1}{z} + O(\frac{1}{z^2}). \tag{3.25}$$

for $p = (z,-)$ near ∞_-.

While existence of a solution follows by construction, uniqueness follows from Theorem B.1 and Remark B.2.

Theorem 3.2. *The solution of the Riemann–Hilbert problem of Theorem 3.1 is unique.*

4. The stationary phase points and corresponding contour deformations

The phase in the factorization problem (3.15) is $t\,\phi$ where ϕ was defined in (3.17). Invoking (2.15) and (2.16), we see that the stationary phase points are given by

$$\prod_{j=0}^{g} (z - \tilde{\lambda}_j) + \frac{n}{t} \prod_{j=1}^{g} (z - \lambda_j) = 0. \tag{4.1}$$

Due to the normalization of our Abelian differentials, the numbers λ_j, $1 \leq j \leq g$, are real and different with precisely one lying in each spectral gap, say λ_j in the j'th gap. Similarly, $\tilde{\lambda}_j$, $0 \leq j \leq g$, are real and different and $\tilde{\lambda}_j$, $1 \leq j \leq g$, sits in the j'th gap. However $\tilde{\lambda}_0$ can be anywhere (see [41, Sect. 13.5]).

As a first step let us clarify the dependence of the stationary phase points on $\frac{n}{t}$.

Lemma 4.1. *Denote by* $z_j(\eta)$, $0 \leq j \leq g$, *the stationary phase points, where* $\eta = \frac{n}{t}$. *Set* $\lambda_0 = -\infty$ *and* $\lambda_{g+1} = \infty$, *then*

$$\lambda_j < z_j(\eta) < \lambda_{j+1} \tag{4.2}$$

and there is always at least one stationary phase point in the j*'th spectral gap. Moreover,* $z_j(\eta)$ *is monotone decreasing with*

$$\lim_{\eta \to -\infty} z_j(\eta) = \lambda_{j+1} \quad \text{and} \quad \lim_{\eta \to \infty} z_j(\eta) = \lambda_j. \tag{4.3}$$

Proof. Due to the normalization of the Abelian differential $\Omega_0 + \eta\omega_{\infty_+\,\infty_-}$ there is at least one stationary phase point in each gap and they are all different. Furthermore,

$$z_j' = -\frac{q(z_j)}{\tilde{q}'(z_j) + \eta q'(z_j)} = -\frac{\prod_{k=1}^{g}(z_j - \lambda_k)}{\prod_{k=0, k\neq j}^{g} z_j - z_k},$$

where

$$\tilde{q}(z) = \prod_{k=0}^{g} (z - \tilde{\lambda}_k), \quad q(z) = \prod_{k=1}^{g} (z - \lambda_k).$$

Since the points λ_k are fixed points of this ordinary first order differential equation (note that the denominator cannot vanish since the z_j's are always different), the numbers z_j cannot cross these points. Combining the behavior as $\eta \to \pm\infty$ with the fact that there must always be at least one of them in each gap, we conclude that z_j must stay between λ_j and λ_{j+1}. This also shows $z_j' < 0$ and thus $z_j(\eta)$ is monotone decreasing. □

In summary, the lemma tells us that we have the following picture: As $\frac{n}{t}$ runs from $-\infty$ to $+\infty$ we start with $z_g(\eta)$ moving from ∞ towards E_{2g+1} while the others stay in their spectral gaps until $z_g(\eta)$ has passed the first spectral band. After this has happened, $z_{g-1}(\eta)$ can leave its gap, while $z_g(\eta)$ remains there, traverses the next spectral band and so on. Until finally $z_0(\eta)$ traverses the last spectral band and escapes to $-\infty$.

So, depending on n/t there is at most one single stationary phase point belonging to the union of the bands $\sigma(H_q)$, say $z_j(n/t)$. On the Riemann surface, there are two such points z_j and its flipping image z_j^* which may (depending on n/t) lie in Σ.

There are three possible cases.

(i) One stationary phase point, say z_j, belongs to the interior of a band $[E_{2j}, E_{2j+1}]$ and all other stationary phase points lie in open gaps.
(ii) $z_j = z_j^* = E_j$ for some j and all other stationary phase points lie in open gaps.
(iii) No stationary phase point belongs to $\sigma(H_q)$.

Case (i). Note that in this case

$$\phi''(z_j)/\mathrm{i} = \frac{\prod_{k=0,k\neq j}^{g}(z_j - z_k)}{\mathrm{i}R_{2g+2}^{1/2}(z_j)} > 0. \tag{4.4}$$

Let us introduce the following "lens" contour near the band $[E_{2j}, E_{2j+1}]$ as shown in Figure 2. The oriented paths $C_j = C_{j1} \cup C_{j2}$, $C_j^* = C_{j1}^* \cup C_{j2}^*$ are meant to be close to the band $[E_{2j}, E_{2j+1}]$.

We have

$$\mathrm{Re}(\phi) > 0, \quad \text{in } D_{j1}, \qquad \mathrm{Re}(\phi) < 0, \quad \text{in } D_{j2}.$$

Indeed

$$\mathrm{Im}(\phi') < 0, \quad \text{in } [E_{2j}, z_j], \qquad \mathrm{Im}(\phi') > 0, \quad \text{in } [z_j, E_{2j+1}] \tag{4.5}$$

noting that ϕ is imaginary in $[E_{2j}, E_{2j+1}]$ and writing $\phi' = d\phi/dz$. Using the Cauchy-Riemann equations we find that the above inequalities are true, as long as C_{j1}, C_{j2} are close enough to the band $[E_{2j}, E_{2j+1}]$. A similar picture appears in the lower sheet.

Concerning the other bands, one simply constructs a "lens" contour near each of the other bands $[E_{2k}, E_{2k+1}]$ and $[E_{2k}^*, E_{2k+1}^*]$ as shown in Figure 3. The oriented paths C_k, C_k^* are meant to be close to the band $[E_{2k}, E_{2k+1}]$. The appropriate transformation is now obvious. Arguing as before, for all bands $[E_{2k}, E_{2k+1}]$ we will have

$$\mathrm{Re}(\phi) < (>) 0, \quad \text{in } D_k, \quad k > (<) j.$$

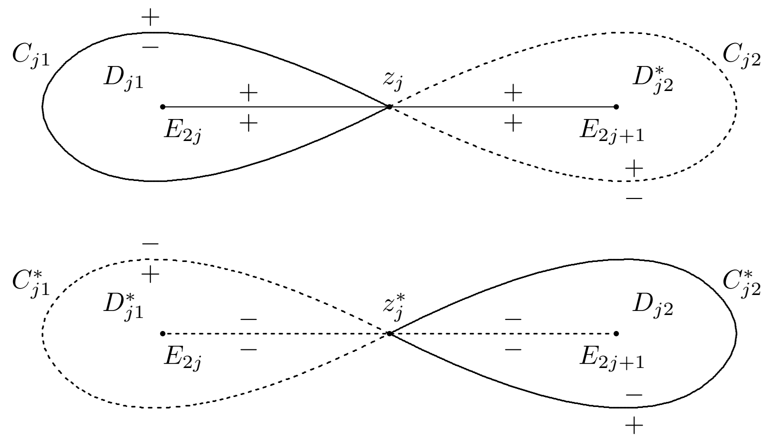

FIGURE 2. The lens contour near a band containing a stationary phase point z_j and its flipping image containing z_j^*. Views from the top and bottom sheet. Dotted curves lie in the bottom sheet.

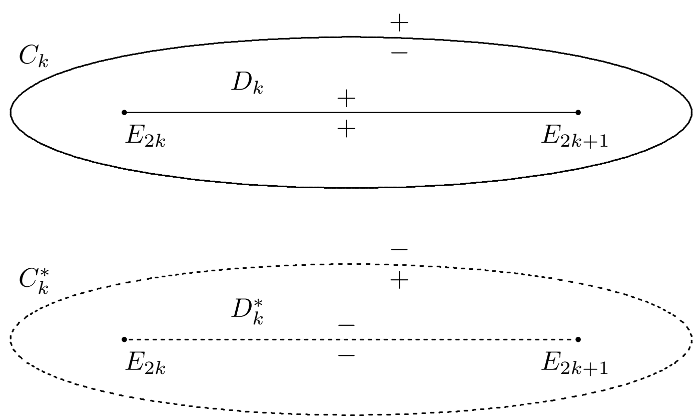

FIGURE 3. The lens contour near a band not including any stationary phase point. Views from the top and bottom sheet.

Now observe that our jump condition (3.24) has the following important factorization

$$J^3 = (b_-)^{-1} b_+, \tag{4.6}$$

where

$$b_- = \begin{pmatrix} 1 & \overline{R\Theta} \mathrm{e}^{-t\,\phi} \\ 0 & 1 \end{pmatrix}, \qquad b_+ = \begin{pmatrix} 1 & 0 \\ R\Theta \mathrm{e}^{t\,\phi} & 1 \end{pmatrix}.$$

This is the right factorization for $z > z_j(n/t)$. Similarly, we have

$$J^3 = (B_-)^{-1} \begin{pmatrix} 1-|R|^2 & 0 \\ 0 & \frac{1}{1-|R|^2} \end{pmatrix} B_+, \tag{4.7}$$

where

$$B_- = \begin{pmatrix} 1 & 0 \\ -\frac{R\Theta \mathrm{e}^{t\,\phi}}{1-|R|^2} & 1 \end{pmatrix}, \qquad B_+ = \begin{pmatrix} 1 & -\frac{\overline{R\Theta}\mathrm{e}^{-t\,\phi}}{1-|R|^2} \\ 0 & 1 \end{pmatrix}.$$

This is the right factorization for $z < z_j(n/t)$. To get rid of the diagonal part we need to solve the corresponding scalar Riemann–Hilbert problem. Again we have to search for a meromorphic solution. This means that the poles of the scalar Riemann–Hilbert problem will be added to the resulting Riemann–Hilbert problem. On the other hand, a pole structure similar to the one of m^3 is crucial for uniqueness. We will address this problem by choosing the poles of the scalar problem in such a way that its zeros cancel the poles of m^3. The right choice will turn out to be $\mathcal{D}_{\underline{\hat{\nu}}}$ (that is, the Dirichlet divisor corresponding to the limiting lattice defined in (1.6)).

Lemma 4.2. *Define a divisor $\mathcal{D}_{\underline{\hat{\nu}}(n,t)}$ of degree g via*

$$\underline{\alpha}_{p_0}(\mathcal{D}_{\underline{\hat{\nu}}(n,t)}) = \underline{\alpha}_{p_0}(\mathcal{D}_{\underline{\hat{\mu}}(n,t)}) + \underline{\delta}(n,t), \tag{4.8}$$

where

$$\delta_\ell(n,t) = \frac{1}{2\pi \mathrm{i}} \int_{C(n/t)} \log(1-|R|^2)\zeta_\ell. \tag{4.9}$$

Then $\mathcal{D}_{\underline{\hat{\nu}}(n,t)}$ is nonspecial and $\pi(\hat{\nu}_j(n,t)) = \nu_j(n,t) \in \mathbb{R}$ with precisely one in each spectral gap.

Proof. Using (2.15) one checks that δ_ℓ is real. Hence it follows from [41, Lem. 9.1] that the ν_j are real and that there is one in each gap. In particular, the divisor $\mathcal{D}_{\underline{\hat{\nu}}}$ is nonspecial by [41, Lem. A.20]. □

Now we can formulate the scalar Riemann–Hilbert problem required to eliminate the diagonal part in the factorization (4.7):

$$\begin{aligned} & d_+(p,n,t) = d_-(p,n,t)(1-|R(p)|^2), \quad p \in C(n/t), \\ & (d) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)}, \\ & d(\infty_+,n,t) = 1, \end{aligned} \tag{4.10}$$

where $C(n/t) = \Sigma \cap \pi^{-1}((-\infty, z_j(n/t))$. Since the index of the (regularized) jump is zero (see remark below), there will be no solution in general unless we admit g additional poles (see e.g. [36, Thm. 5.2]).

Theorem 4.3. *The unique solution of* (4.10) *is given by*

$$\begin{aligned} d(p,n,t) = & \frac{\theta(\underline{z}(n,t) + \underline{\delta}(n,t))}{\theta(\underline{z}(n,t))} \frac{\theta(\underline{z}(p,n,t))}{\theta(\underline{z}(p,n,t) + \underline{\delta}(n,t))} \times \\ & \times \exp\left(\frac{1}{2\pi \mathrm{i}} \int_{C(n/t)} \log(1-|R|^2) \omega_{p\,\infty_+} \right), \end{aligned} \tag{4.11}$$

where $\underline{\delta}(n,t)$ is defined in (4.9) *and $\omega_{p\,q}$ is the Abelian differential of the third kind with poles at p and q (cf. Remark 4.4 below).*

The function $d(p)$ is meromorphic in $\mathbb{M} \setminus \Sigma$ with first order poles at $\hat{\nu}_j(n,t)$ and first order zeros at $\hat{\mu}_j(n,t)$. Also $d(p)$ is uniformly bounded in n, t away from the poles.

In addition, we have $d(p) = \overline{d(\overline{p})}$.

Note that this formula is different (in fact much simpler) from the explicit solution formula from Rodin [36, Sec. 1.8]. It is the core of our explicit formula (1.6) for the limiting lattice.

Proof. On the Riemann sphere, a scalar Riemann–Hilbert problem is solved by the Plemelj–Sokhotsky formula. On our Riemann surface we need to replace the Cauchy kernel $\frac{d\lambda}{\lambda - z}$ by the Abelian differential of the third kind $\omega_{p\,\infty_+}$. But now it is important to observe that this differential is not single-valued with respect to p. In fact, if we move p across the a_ℓ cycle, the normalization $\int_{a_\ell} \omega_{p\,\infty_+} = 0$ enforces a jump by $2\pi\mathrm{i}\zeta_\ell$. One way of compensating for these jumps is by adding to $\omega_{p\,\infty_+}$ suitable integrals of Abelian differentials of the second kind (cf. [36, Sec 1.4] or Appendix A). Since this will produce essential singularities after taking exponentials we prefer to rather leave $\omega_{p\,\infty_+}$ as it is and compensate for the jumps (after taking exponentials) by proper use of Riemann theta functions.

To this end recall that the Riemann theta function satisfies

$$\theta(\underline{z} + \underline{m} + \underline{\tau}\,\underline{n}) = \exp[2\pi\mathrm{i}\left(-\langle \underline{n}, \underline{z}\rangle - \frac{\langle \underline{n}, \underline{\tau}\,\underline{n}\rangle}{2}\right)]\theta(\underline{z}), \quad \underline{n}, \underline{m} \in \mathbb{Z}^g, \tag{4.12}$$

where $\underline{\tau}$ is the matrix of b-periods defined in (2.7) and $\langle ., ..\rangle$ denotes the scalar product in $\mathbb{R}^g$ (cf., e.g. [16] or [41, App. A]). By definition both the theta functions (as functions on $\mathbb{M}$) and the exponential term are only defined on the "fundamental polygon" $\hat{\mathbb{M}}$ of $\mathbb{M}$ and do not extend to single-valued functions on $\mathbb{M}$ in general. However, multi-valuedness apart, d is a (locally) holomorphic solution of our Riemann–Hilbert problem which is one at ∞_+ by our choice of the second pole of the Cauchy kernel $\omega_{p\,\infty_+}$. The ratio of theta functions is, again apart from multi-valuedness, meromorphic with simple zeros at $\hat{\mu}_j$ and simple poles at $\hat{\nu}_j$ by Riemann's vanishing theorem. Moreover, the normalization is chosen again such that the ratio of theta functions is one at ∞_+. Hence it remains to verify that (4.11) gives rise to a single-valued function on $\mathbb{M}$.

Let us start by looking at the values from the left/right on the cycle b_ℓ. Since our path of integration in $\underline{z}(p)$ is forced to stay in $\hat{\mathbb{M}}$, the difference between the limits from the right and left is the value of the integral along a_ℓ. So by (4.12) the limits of the theta functions match. Similarly, since $\omega_{p\,\infty_+}$ is normalized along a_ℓ cycles, the limits from the left/right of $\omega_{p\,\infty_+}$ coincide. So the limits of the exponential terms from different sides of b_ℓ match as well.

Next, let us compare the values from the left/right on the cycle a_ℓ. Since our path of integration in $\underline{z}(p)$ is forced to stay in $\hat{\mathbb{M}}$, the difference between the limits from the right and left is the value of the integral along b_ℓ. So by (4.12) the limits of the theta functions will differ by a multiplicative factor $\exp(2\pi\mathrm{i}\delta_\ell)$. On the other hand, since $\omega_{p\,\infty_+}$ is normalized along a_ℓ cycles, the values from the right and left will differ by $-2\pi\mathrm{i}\zeta_\ell$. By our definition of $\underline{\delta}$ in (4.9), the jumps of the ration of theta functions and the exponential term compensate each other which shows that (4.11) is single-valued.

To see uniqueness let $\tilde{d}$ be a second solution and consider $\tilde{d}/d$. Then $\tilde{d}/d$ has no jump and the Schwarz reflection principle implies that it extends to a meromorphic function on $\mathbb{M}$. Since the poles of d cancel the poles of $\tilde{d}$, its divisor satisfies $(\tilde{d}/d) \geq -\mathcal{D}_{\hat{\underline{\mu}}}$. But $\mathcal{D}_{\hat{\underline{\mu}}}$ is nonspecial and thus $\tilde{d}/d$ must be constant by the Riemann–Roch theorem. Setting $p = \infty_+$ we see that this constant is one, that is, $\tilde{d} = d$ as claimed.

Finally, $d(p) = \overline{d(\overline{p})}$ follows from uniqueness since both functions solve (4.10). $\square$

Remark 4.4. *The Abelian differential $\omega_{p\,q}$ used in the previous theorem is explicitly given by*

$$\omega_{p\,q} = \left(\frac{R_{2g+2}^{1/2} + R_{2g+2}^{1/2}(p)}{2(\pi - \pi(p))} - \frac{R_{2g+2}^{1/2} + R_{2g+2}^{1/2}(q)}{2(\pi - \pi(q))} + P_{pq}(\pi) \right) \frac{d\pi}{R_{2g+2}^{1/2}}, \tag{4.13}$$

where $P_{pq}(z)$ is a polynomial of degree $g-1$ which has to be determined from the normalization $\int_{a_\ell} \omega_{p\,p^*} = 0$. For $q = \infty_\pm$ we have

$$\omega_{p\,\infty_\pm} = \left(\frac{R_{2g+2}^{1/2} + R_{2g+2}^{1/2}(p)}{2(\pi - \pi(p))} \mp \frac{1}{2}\pi^g + P_{p\infty_\pm}(\pi) \right) \frac{d\pi}{R_{2g+2}^{1/2}}. \tag{4.14}$$

Remark 4.5. *Once the last stationary phase point has left the spectrum, that is, once $C(n/t) = \Sigma$, we have $d(p) = A^{-1} T(z)^{\pm 1}$, $p = (z, \pm)$ (compare* [43]*). Here $A = A_+(n,t) A_-(n,t) = T(\infty)$.*

In particular,

$$\begin{aligned} d(\infty_-, n, t) =& \frac{\theta(\underline{z}(n-1,t))}{\theta(\underline{z}(n,t))} \frac{\theta(\underline{z}(n,t) + \underline{\delta}(n,t))}{\theta(\underline{z}(n-1,t) + \underline{\delta}(n,t))} \times \\ & \times \exp\left(\frac{1}{2\pi\mathrm{i}} \int_{C(n/t)} \log(1 - |R|^2) \omega_{\infty_-\,\infty_+} \right), \end{aligned} \tag{4.15}$$

since $\underline{z}(\infty_-, n, t) = \underline{z}(\infty_+, n-1, t) = \underline{z}(n-1, t)$. Note that $\overline{d(\infty_-, n, t)} = d(\overline{\infty_-}, n, t) = d(\infty_-, n, t)$ shows that $d(\infty_-, n, t)$ is real-valued. Using (2.15) one can even show that it is positive.

The next lemma characterizes the singularities of $d(p)$ near the stationary phase points and the band edges.

Lemma 4.6. *For p near a stationary phase point z_j or z_j^* (not equal to a band edge) we have*

$$d(p) = (z - z_j)^{\pm \mathrm{i}\nu} e^{\pm}(z), \quad p = (z, \pm), \tag{4.16}$$

where $e^{\pm}(z)$ is Hölder continuous of any exponent less than 1 near z_j and

$$\nu = -\frac{1}{2\pi} \log(1 - |R(z_j)|^2) > 0. \tag{4.17}$$

Here $(z - z_j)^{\pm \mathrm{i}\nu} = \exp(\pm \mathrm{i}\nu \log(z - z_j))$, where the branch cut of the logarithm is along the negative real axis.

For p near a band edge $E_k \in C(n/t)$ we have

$$d(p) = T^{\pm 1}(z) \tilde{e}^{\pm}(z), \quad p = (z, \pm), \tag{4.18}$$

where $\tilde{e}^{\pm}(z)$ is holomorphic near E_k if none of the ν_j is equal to E_k and $\tilde{e}_\pm(z)$ has a first order pole at $E_k = \nu_j$ else.

Proof. The first claim we first rewrite (4.11) as

$$\begin{aligned} d(p, n, t) =& \exp\left(\mathrm{i}\nu \int_{C(n/t)} \omega_{p\,\infty_+} \right) \frac{\theta(\underline{z}(n,t) + \underline{\delta}(n,t))}{\theta(\underline{z}(n,t))} \frac{\theta(\underline{z}(p,n,t))}{\theta(\underline{z}(p,n,t) + \underline{\delta}(n,t))} \times \\ & \times \exp\left(\frac{1}{2\pi\mathrm{i}} \int_{C(n/t)} \log\left(\frac{1 - |R|^2}{1 - |R(z_j)|^2} \right) \omega_{p\,\infty_+} \right). \end{aligned} \tag{4.19}$$

Next observe

$$\frac{1}{2}\int_{C(n/t)} \omega_{p\,p^*} = \pm\log(z-z_j) \pm \alpha(z_j) + O(z-z_j), \quad p=(z,\pm), \tag{4.20}$$

where $\alpha(z_j)\in\mathbb{R}$, and hence

$$\int_{C(n/t)} \omega_{p\,\infty_+} = \pm\log(z-z_j) \pm \alpha(z_j) + \frac{1}{2}\int_{C(n/t)} \omega_{\infty_-\,\infty_+} + O(z-z_j), \quad p=(z,\pm), \tag{4.21}$$

from which the first claim follows.

For the second claim note that

$$t(p) = \frac{1}{T(\infty)} \begin{cases} T(z), & p=(z,+)\in\Pi_+, \\ T(z)^{-1}, & p=(z,-)\in\Pi_-, \end{cases}$$

satisfies the (holomorphic) Riemann–Hilbert problem

$$\begin{aligned} & t_+(p) = t_-(p)(1-|R(p)|^2), \quad p\in\Sigma, \\ & t(\infty_+) = 1. \end{aligned}$$

Hence $d(p)/t(p)$ has no jump along $C(n,t)$ and is thus holomorphic near $C(n/t)$ away from band edges $E_k=\nu_j$ (where there is a simple pole) by the Schwarz reflection principle. □

Furthermore,

Lemma 4.7. *We have*

$$e^{\pm}(z) = \overline{e^{\mp}(z)}, \qquad p=(z,\pm)\in\Sigma\backslash C(n/t), \tag{4.22}$$

and

$$\begin{aligned} e^+(z_j) = & \exp\left(\mathrm{i}\nu\alpha(z_j) + \frac{\mathrm{i}\nu}{2}\int_{C(n/t)} \omega_{\infty_-\,\infty_+}\right)\times \\ & \times \frac{\theta(\underline{z}(n,t)+\underline{\delta}(n,t))}{\theta(\underline{z}(n,t))} \frac{\theta(\underline{z}(z_j,n,t))}{\theta(\underline{z}(z_j,n,t)+\underline{\delta}(n,t))}\times \\ & \times \exp\left(\frac{1}{2\pi\mathrm{i}}\int_{C(n/t)} \log\left(\frac{1-|R|^2}{1-|R(z_j)|^2}\right)\left(\omega_{z_j\,z_j^*} + \omega_{\infty_-\,\infty_+}\right)\right), \end{aligned} \tag{4.23}$$

where

$$\alpha(z_j) = \lim_{p\to z_j} \frac{1}{2}\int_{C(n/t)} \omega_{p\,p^*} - \log(\pi(p)-z_j). \tag{4.24}$$

Here $\alpha(z_j)\in\mathbb{R}$ and $\omega_{p\,p^*}$ is real whereas $\omega_{\infty_-\,\infty_+}$ is purely imaginary on $C(n/t)$.

Proof. The first claim follows since $d(p^*)=d(\overline{p})=\overline{d(p)}$ for $p\in\Sigma\backslash C(n/t)$. The second claim follows from (4.19) using $\int_{C(n/t)} f\,\omega_{p\,\infty_+} = \frac{1}{2}\int_{C(n/t)} f\,(\omega_{p\,p}+\omega_{\infty_-\,\infty_+})$ for symmetric functions $f(q)=f(q^*)$. □

Having solved the scalar problem above for d we can introduce the new Riemann–Hilbert problem

$$m^4(p) = d(\infty_-)^{-1} m^3(p) D(p), \quad D(p) = \begin{pmatrix} d(p^*) & 0 \\ 0 & d(p) \end{pmatrix}. \tag{4.25}$$

where $d^*(p) = d(p^*)$ is the unique solution of

$$\begin{aligned} & d_+^*(p) = d_-^*(p)(1-|R(p)|^2)^{-1}, \quad p\in C(n/t),\\ & (d^*) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)^*},\\ & d^*(\infty_-) = 1. \end{aligned}$$

Note that

$$\det(D(p)) = d(p) d(p^*) = d(\infty_-)\prod_{j=1}^{g}\frac{z-\mu_j}{z-\nu_j}.$$

Then a straightforward calculation shows that m^4 satisfies

$$\begin{aligned} & m_+^4(p) = m_-^4(p) J^4(p), \quad p\in\Sigma,\\ & (m_1^4) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)^*}, \quad (m_2^4) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)},\\ & m^4(p^*) = m^4(p)\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}\\ & m^4(\infty_+) = \begin{pmatrix} 1 & * \end{pmatrix}, \end{aligned} \tag{4.26}$$

where the jump is given by

$$J^4(p) = D_-(p)^{-1} J^3(p) D_+(p), \quad p\in\Sigma. \tag{4.27}$$

In particular, m^4 has its poles shifted from $\hat{\mu}_j(n,t)$ to $\hat{\nu}_j(n,t)$.

Furthermore, J^4 can be factorized as

$$J^4 = \begin{pmatrix} 1-|R|^2 & -\frac{d}{d^*}\overline{R\Theta}\mathrm{e}^{-t\,\phi}\\ \frac{d^*}{d}R\Theta\mathrm{e}^{t\,\phi} & 1\end{pmatrix} = (\tilde{b}_-)^{-1}\tilde{b}_+, \quad p\in\Sigma\setminus C(n/t), \tag{4.28}$$

where $\tilde{b}_\pm = D^{-1} b_\pm D$, that is,

$$\tilde{b}_- = \begin{pmatrix} 1 & \frac{d}{d^*}\overline{R\Theta}\mathrm{e}^{-t\,\phi}\\ 0 & 1\end{pmatrix}, \qquad \tilde{b}_+ = \begin{pmatrix} 1 & 0\\ \frac{d^*}{d}R\Theta\mathrm{e}^{t\,\phi} & 1\end{pmatrix}, \tag{4.29}$$

for $\pi(p) > z_j(n/t)$ and

$$J^4 = \begin{pmatrix} 1 & -\frac{d_+}{d_-^*}\overline{R\Theta}\mathrm{e}^{-t\,\phi}\\ \frac{d_-^*}{d_+}R\Theta\mathrm{e}^{t\,\phi} & 1-|R|^2\end{pmatrix} = (\tilde{B}_-)^{-1}\tilde{B}_+, \quad p\in C(n/t), \tag{4.30}$$

where $\tilde{B}_\pm = D_\pm^{-1} B_\pm D_\pm$, that is,

$$\tilde{B}_- = \begin{pmatrix} 1 & 0\\ -\frac{d_-^*}{d_-}\frac{R\Theta}{1-|R|^2}\mathrm{e}^{t\,\phi} & 1\end{pmatrix}, \qquad \tilde{B}_+ = \begin{pmatrix} 1 & -\frac{d_+}{d_+^*}\frac{\overline{R\Theta}}{1-|R|^2}\mathrm{e}^{-t\,\phi}\\ 0 & 1\end{pmatrix}, \tag{4.31}$$

for $\pi(p) < z_j(n/t)$.

Note that by $\overline{d(p)} = d(\overline{p})$ we have

$$\frac{d_-^*(p)}{d_+(p)} = \frac{d_-^*(p)}{d_-(p)}\frac{1}{1-|R(p)|^2} = \frac{\overline{d_+(p)}}{d_+(p)}, \qquad p\in C(n/t), \tag{4.32}$$

respectively

$$\frac{d_+(p)}{d_-^*(p)} = \frac{d_+(p)}{d_+^*(p)}\frac{1}{1-|R(p)|^2} = \frac{\overline{d_-^*(p)}}{d_-^*(p)}, \qquad p\in C(n/t). \tag{4.33}$$

We finally define m^5 by

$$
\begin{aligned}
m^5 &= m^4 \tilde{B}_+^{-1}, \quad p \in D_k,\ k < j,\\
m^5 &= m^4 \tilde{B}_-^{-1}, \quad p \in D_k^*,\ k < j,\\
m^5 &= m^4 \tilde{B}_+^{-1}, \quad p \in D_{j1},\\
m^5 &= m^4 \tilde{B}_-^{-1}, \quad p \in D_{j1}^*,\\
m^5 &= m^4 \tilde{b}_+^{-1}, \quad p \in D_{j2},\\
m^5 &= m^4 \tilde{b}_-^{-1}, \quad p \in D_{j2}^*,\\
m^5 &= m^4 \tilde{b}_+^{-1}, \quad p \in D_k,\ k > j,\\
m^5 &= m^4 \tilde{b}_-^{-1}, \quad p \in D_k^*,\ k > j,\\
m^5 &= m^4, \quad \text{otherwise},
\end{aligned} \tag{4.34}
$$

where we assume that the deformed contour is sufficiently close to the original one. The new jump matrix is given by

$$
\begin{aligned}
m_+^5(p,n,t) &= m_-^5(p,n,t) J^5(p,n,t),\\
J^5 &= \tilde{B}_+, \quad p \in C_k,\ k < j,\\
J^5 &= \tilde{B}_-^{-1}, \quad p \in C_k^*,\ k < j,\\
J^5 &= \tilde{B}_+, \quad p \in C_{j1},\\
J^5 &= \tilde{B}_-^{-1}, \quad p \in C_{j1}^*,\\
J^5 &= \tilde{b}_+, \quad p \in C_{j2},\\
J^5 &= \tilde{b}_-^{-1}, \quad p \in C_{j2}^*,\\
J^5 &= \tilde{b}_+, \quad p \in C_k,\ k > j,\\
J^5 &= \tilde{b}_-^{-1}, \quad p \in C_k^*,\ k > j.
\end{aligned} \tag{4.35}
$$

Here we have assumed that the function $R(p)$ admits an analytic extension in the corresponding regions. Of course this is not true in general, but we can always evade this obstacle by approximating $R(p)$ by analytic functions in the spirit of [6]. We will provide the details in Section 6.

The crucial observation now is that the jumps J^5 on the oriented paths C_k, C_k^* are of the form $\mathbb{I} +$ *exponentially small* asymptotically as $t \to \infty$, at least away from the stationary phase points z_j, z_j^*. We thus hope we can simply replace these jumps by the identity matrix (asymptotically as $t \to \infty$) implying that the solution should asymptotically be given by the constant vector $\begin{pmatrix} 1 & 1 \end{pmatrix}$. That this can in fact be done will be shown in the next section by explicitly computing the contribution of the stationary phase points thereby showing that they are of the order $O(t^{-1/2})$, that is,

$$
m^5(p) = \begin{pmatrix} 1 & 1 \end{pmatrix} + O(t^{-1/2})
$$

uniformly for p a way from the jump contour. Hence all which remains to be done to prove Theorem 1.1 and Theorem 1.3 is to trace back the definitions of m^4 and m^3 and comparing with (3.25). First of all, since m^5 and m^4 coincide near ∞_- we have

$$
m^4(p) = \begin{pmatrix} 1 & 1 \end{pmatrix} + O(t^{-1/2})
$$

uniformly for p in a neighborhood of ∞_-. Consequently, by the definition of m^4 from (4.25), we have

$$m^3(p) = d(\infty_-)\left(d(p^*)^{-1} \quad d(p)^{-1}\right) + O(t^{-1/2})$$

again uniformly for p in a neighborhood of ∞_-. Finally, comparing this last identity with (3.25) shows

$$A_+(n,t)^2 = d(\infty_-, n, t) + O(t^{-1/2}), \quad B_+(n,t) = -d_1(n,t) + O(t^{-1/2}), \tag{4.36}$$

where d_1 is defined via

$$d(p) = 1 + \frac{d_1}{z} + O(\frac{1}{z^2}), \qquad p = (z, +) \text{ near } \infty_+$$

Hence it remains to compute d_1. Proceeding as in [41, Thm. 9.4] respectively [43, Sec. 4] one obtains

$$\begin{aligned} d_1 = & -\frac{1}{2\pi\mathrm{i}} \int_{C(n/t)} \log(1-|R|^2)\Omega_0 \\ & -\frac{1}{2}\frac{d}{ds}\log\left(\frac{\theta(\underline{z}(n,s) + \underline{\delta}(n,t))}{\theta(\underline{z}(n,s))}\right)\Big|_{s=t}, \end{aligned}$$

where Ω_0 is the Abelian differential of the second kind defined in (2.16).

Case (ii). In the special case where the two stationary phase points coincide (so $z_j = z_j^* = E_k$ for some k) the Riemann–Hilbert problem arising above is of a different nature, even in the simpler non-generic case $|R(E_k)| < 1$. In analogy to the case of the free lattice one expects different local asymptotics expressed in terms of Painlevé functions. In the case $|R(E_k)| < 1$ the two crosses coalesce and the discussion of Section B goes through virtually unaltered. If $|R(E_k)| = 1$ the problem is singular in an essential way and we expect an extra "collisionless shock" phenomenon (on top of the Painlevé phenomenon) in the region where $z_j(n/t) \sim E_k$, similar to the one studied in [1], [9], [24]. The main difficulty arises from the singularity of $\frac{R}{1-|R|^2}$. An appropriate "local" Riemann–Hilbert problem however is still explicitly solvable and the actual contribution of the band edges is similar to the free case. All this can be studied as in Section 5 (see also our discussion of this in the Introduction). But in the present work, we will assume that the stationary phase points stay away from the E_k.

Case (iii). In the case where no stationary phase points lie in the spectrum the situation is similar to the case (i). In fact, it is much simpler since there is no contribution from the stationary phase points: There is a gap (the j-th gap, say) in which two stationary phase points exist. We construct "lens-type" contours C_k around every single band lying to the left of the j-th gap and make use of the factorization $J^3 = (\tilde{b}_-)^{-1}\tilde{b}_+$. We also construct "lens-type" contours C_k around every single band lying to the right of the j-th gap and make use of the factorization

$J^3 = (\tilde{B}_-)^{-1}\tilde{B}_+$. Indeed, in place of (4.34) we set

$$\begin{aligned}
m^5 &= m^4 \tilde{B}_+^{-1}, \quad p \in D_k,\ k < j,\\
m^5 &= m^4 \tilde{B}_-^{-1}, \quad p \in D_k^*,\ k < j,\\
m^5 &= m^4 \tilde{b}_+^{-1}, \quad p \in D_k,\ k > j,\\
m^5 &= m^4 \tilde{b}_-^{-1}, \quad p \in D_k^*,\ k > j,\\
m^5 &= m^4, \quad \text{otherwise}.
\end{aligned} \tag{4.37}$$

It is now easy to check that in both cases (i) and (iii) formula (4.15) is still true.

Remark 4.8. *We have asymptotically reduced our Riemann–Hilbert problem to one defined on two small crosses. If we are only interested in showing that the contribution of these crosses is small (i.e that the solution of the Riemann–Hilbert problem is uniformly small for large times) we can evoke the existence theorem in the second appendix as well as some rescaling argument.*

Since we are interested in actually computing the higher order asymptotic term, a more detailed analysis of the local parametrix Riemann–Hilbert problem is required.

5. The "local" Riemann–Hilbert problems on the small crosses

In the previous section we have shown how the long-time asymptotics can be read off from the Riemann–Hilbert problem

$$\begin{aligned}
& m_+^5(p,n,t) = m_-^5(p,n,t) J^5(p,n,t), \quad p \in \Sigma^5,\\
& (m_1^5) \ge -\mathcal{D}_{\underline{\hat{\nu}}(n,t)^*}, \quad (m_2^5) \ge -\mathcal{D}_{\underline{\hat{\nu}}(n,t)},\\
& m^5(p^*,n,t) = m^5(p,n,t) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\\
& m^5(\infty_+,n,t) = \begin{pmatrix} 1 & * \end{pmatrix}.
\end{aligned} \tag{5.1}$$

In this section we are interested in the actual asymptotic rate at which $m^5(p) \to \begin{pmatrix} 1 & 1 \end{pmatrix}$. We have already seen in the previous section that the jumps J^5 on the oriented paths C_k, C_k^* for $k \ne j$ are of the form $\mathbb{I} +$ *exponentially small* asymptotically as $t \to \infty$. The same is true for the oriented paths $C_{j1}, C_{j2}, C_{j1}^*, C_{j2}^*$ at least away from the stationary phase points z_j, z_j^*. On these paths, and in particular near the stationary phase points (see Figure 4), the jumps read

$$\begin{aligned}
J^5 &= \tilde{B}_+ = \begin{pmatrix} 1 & -\frac{d}{d^*}\frac{R^*\Theta^*}{1-R^*R}\mathrm{e}^{-t\,\phi} \\ 0 & 1 \end{pmatrix}, \quad p \in C_{j1},\\
J^5 &= \tilde{B}_-^{-1} = \begin{pmatrix} 1 & 0 \\ \frac{d^*}{d}\frac{R\Theta}{1-R^*R}\mathrm{e}^{t\,\phi} & 1 \end{pmatrix}, \quad p \in C_{j1}^*,\\
J^5 &= \tilde{b}_+ = \begin{pmatrix} 1 & 0 \\ \frac{d^*}{d}R\Theta\mathrm{e}^{t\,\phi} & 1 \end{pmatrix}, \quad p \in C_{j2},\\
J^5 &= \tilde{b}_-^{-1} = \begin{pmatrix} 1 & -\frac{d}{d^*}R^*\Theta^*\mathrm{e}^{-t\,\phi} \\ 0 & 1 \end{pmatrix}, \quad p \in C_{j2}^*.
\end{aligned} \tag{5.2}$$

Note that near the stationary phase points the jumps are given by (cf. Lemma 4.6)

$$\hat{B}_+ = \begin{pmatrix} 1 & -\left(\sqrt{\frac{\phi''(z_j)}{\mathrm{i}}}(z-z_j)\right)^{2\mathrm{i}\nu} \frac{\overline{r}}{1-|r|^2}\mathrm{e}^{-t\,\phi} \\ 0 & 1 \end{pmatrix}, \quad p \in L_{j1},$$

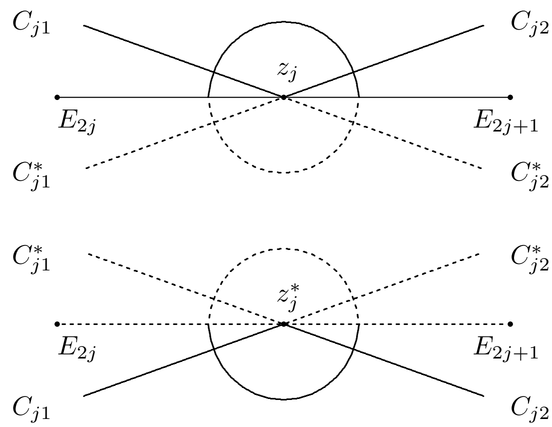

FIGURE 4. The small cross containing the stationary phase point z_j and its flipping image containing z_j^*. Views from the top and bottom sheet. Dotted curves lie in the bottom sheet.

$$\hat{B}_-^{-1} = \begin{pmatrix} 1 & 0 \\ \left(\sqrt{\frac{\phi''(z_j)}{\mathrm{i}}}(z - z_j)\right)^{-2\mathrm{i}\nu} \frac{r}{1-|r|^2}\mathrm{e}^{t\,\phi} & 1 \end{pmatrix}, \quad p \in L_{j1}^*,$$

$$\hat{b}_+ = \begin{pmatrix} 1 & 0 \\ \left(\sqrt{\frac{\phi''(z_j)}{\mathrm{i}}}(z - z_j)\right)^{-2\mathrm{i}\nu} r\mathrm{e}^{t\,\phi} & 1 \end{pmatrix}, \quad p \in L_{j2},$$

$$\hat{b}_-^{-1} = \begin{pmatrix} 1 & -\left(\sqrt{\frac{\phi''(z_j)}{\mathrm{i}}}(z - z_j)\right)^{2\mathrm{i}\nu} \overline{r}\mathrm{e}^{-t\,\phi} \\ 0 & 1 \end{pmatrix}, \quad p \in L_{j2}^*, \tag{5.3}$$

where (cf. (3.16) and (4.16))

$$r = R(z_j)\Theta(z_j, n, t)\frac{\overline{e^+(z_j)}}{e^+(z_j)}\left(\frac{\phi''(z_j)}{\mathrm{i}}\right)^{\mathrm{i}\nu}. \tag{5.4}$$

Since the reflection coefficients are continuously differentiable by our decay assumption (1.2) and by Lemma 4.6 the error terms will satisfy appropriate Hölder estimates, that is

$$\|\tilde{B}_+(p) - \hat{B}_+(p)\| \le C|z - z_j|^\alpha, \qquad p = (z, +) \in C_{j1}, \tag{5.5}$$

for any $\alpha < 1$ and similarly for the other matrices.

To reduce our Riemann–Hilbert problem to the one corresponding to the two crosses we proceed as follows: We take a small disc D around $z_j(n/t)$ and project it to the complex plane using the canonical projection π. Now consider the (holomorphic) matrix Riemann–Hilbert problem *in the complex plane* with the very jump obtained by projection and normalize it to be $\mathbb{I}$ near ∞. Denote this solution by $M(z)$. Then, as is shown in [6] (see also [29, Thm. A.1]), the solution of this matrix Riemann–Hilbert problem on a small cross in the complex plane is asymptotically of the form

$$M(z) = \mathbb{I} + \frac{M_0}{z - z_j} \frac{1}{t^{1/2}} + O(t^{-\alpha}), \tag{5.6}$$

for any $\alpha < 1$ and z outside a neighborhood of z_j, where

$$M_0 = \mathrm{i}\sqrt{\mathrm{i}/\phi''(z_j)} \begin{pmatrix} 0 & -\beta(t) \\ \overline{\beta(t)} & 0 \end{pmatrix},$$

$$\beta(t) = \sqrt{\nu}\mathrm{e}^{\mathrm{i}(\pi/4 - \arg(r) + \arg(\Gamma(\mathrm{i}\nu)))} \mathrm{e}^{-\mathrm{i}t\phi(z_j)} t^{-\mathrm{i}\nu}. \tag{5.7}$$

Now we lift this solution back to the small disc on our Riemann-surface by setting $M(p) = M(z)$ for $p \in D$ and $M(p) = \overline{M(\overline{z})}$ for $p \in D^*$. We define

$$m^6(p) = \begin{cases} m^5(p) M^{-1}(p), & p \in D \cup D^* \\ m^5(p), & \text{else.} \end{cases} \tag{5.8}$$

Note that m^6 has no jump inside $D \cup D^*$. Its jumps on the boundary are given by

$$m^6_+(p) = m^6_-(p) M^{-1}(p), \qquad p \in \partial D \cup \partial D^* \tag{5.9}$$

and the remaining jumps are unchanged. In summary, all jumps outside $D \cup D^*$ are of the form $\mathbb{I} +$ *exponentially small* and the jump on $\partial D \cup \partial D^*$ is of the form $\mathbb{I} + O(t^{-1/2})$.

In order to identify the leading behaviour it remains to rewrite the Riemann–Hilbert problem for m^6 as a singular integral equation following Appendix A. Let the operator $C_{w^6} : L^2(\Sigma^6) \to L^2(\Sigma^6)$ be defined by

$$C_{w^6} f = C_-(f w^6) \tag{5.10}$$

for a vector valued f, where $w^6 = J^6 - \mathbb{I}$ and

$$(C_\pm f)(q) = \lim_{p \to q \in \Sigma^6} \frac{1}{2\pi\mathrm{i}} \int_{\Sigma^6} f \, \underline{\Omega}_p^{\underline{\hat{\nu}}}, \qquad \underline{\Omega}_p^{\underline{\hat{\nu}}} = \begin{pmatrix} \Omega_p^{\underline{\hat{\nu}}^*, \infty_+} & 0 \\ 0 & \Omega_p^{\underline{\hat{\nu}}, \infty_-} \end{pmatrix}, \tag{5.11}$$

are the Cauchy operators for our Riemann surface. In particular, $\Omega_p^{\underline{\hat{\nu}}, q}$ is the Cauchy kernel given by

$$\Omega_p^{\underline{\hat{\nu}}, q} = \omega_{p\, q} + \sum_{j=1}^{g} I_j^{\underline{\hat{\nu}}, q}(p) \zeta_j, \tag{5.12}$$

where

$$I_j^{\underline{\hat{\nu}}, q}(p) = \sum_{\ell=1}^{g} c_{j\ell}(\underline{\hat{\nu}}) \int_q^p \omega_{\hat{\nu}_\ell, 0}. \tag{5.13}$$

Here $\omega_{q,0}$ is the (normalized) Abelian differential of the second kind with a second order pole at q (cf. Remark 5.2 below). Note that $I_j^{\underline{\hat{\nu}}, q}(p)$ has first order poles at the points $\underline{\hat{\nu}}$.

The constants $c_{j\ell}(\underline{\hat{\nu}})$ are chosen such that $\Omega_p^{\underline{\hat{\nu}}, q}$ is single valued, that is,

$$(c_{\ell k}(\underline{\hat{\nu}}))_{1 \le \ell, k \le g} = \left(\sum_{j=1}^{g} c_k(j) \frac{\mu_\ell^{j-1} d\pi}{R_{2g+2}^{1/2}(\hat{\mu}_\ell)} \right)_{1 \le \ell, k \le g}^{-1} \tag{5.14}$$

where $c_k(j)$ are defined in (2.6) (cf. Lemma A.3).

Next, consider the solution μ^6 of the singular integral equation

$$\mu = \begin{pmatrix} 1 & 1 \end{pmatrix} + C_{w^6}\mu \quad \text{in} \quad L^2(\Sigma^6). \tag{5.15}$$

Then the solution of our Riemann–Hilbert problem is given by

$$m^6(p) = \begin{pmatrix} 1 & 1 \end{pmatrix} + \frac{1}{2\pi\mathrm{i}} \int_{\Sigma^6} \mu^6\, w^6\, \underline{\Omega}_p^{\hat{\underline{\nu}}}. \tag{5.16}$$

Since $\|w^6\|_\infty = O(t^{-1/2})$ Neumann's formula implies

$$\mu^6(q) = (\mathbb{I} - C_{w^6})^{-1} \begin{pmatrix} 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} + O(t^{-1/2}). \tag{5.17}$$

Moreover,

$$w^6(p) = \begin{cases} -\frac{M_0}{z-z_j}\frac{1}{t^{1/2}} + O(t^{-\alpha}), & p \in \partial D, \\ -\frac{\overline{M_0}}{z-z_j}\frac{1}{t^{1/2}} + O(t^{-\alpha}), & p \in \partial D^*. \end{cases} \tag{5.18}$$

Hence we obtain

$$\begin{aligned} m^6(p) &= \begin{pmatrix} 1 & 1 \end{pmatrix} - \frac{\begin{pmatrix} 1 & 1 \end{pmatrix} M_0}{t^{1/2}} \frac{1}{2\pi\mathrm{i}} \int_{\partial D} \frac{1}{\pi - z_j}\, \underline{\Omega}_p^{\hat{\underline{\nu}}} \\ &\quad - \frac{\begin{pmatrix} 1 & 1 \end{pmatrix} \overline{M_0}}{t^{1/2}} \frac{1}{2\pi\mathrm{i}} \int_{\partial D^*} \frac{1}{\pi - z_j}\, \underline{\Omega}_p^{\hat{\underline{\nu}}} + O(t^{-\alpha}) \\ &= \begin{pmatrix} 1 & 1 \end{pmatrix} - \frac{\begin{pmatrix} 1 & 1 \end{pmatrix} M_0}{t^{1/2}} \underline{\Omega}_p^{\hat{\underline{\nu}}}(z_j) - \frac{\begin{pmatrix} 1 & 1 \end{pmatrix} \overline{M_0}}{t^{1/2}} \underline{\Omega}_p^{\hat{\underline{\nu}}}(z_j^*) + O(t^{-\alpha}) \\ &= \begin{pmatrix} 1 & 1 \end{pmatrix} \\ &\quad - \sqrt{\frac{\mathrm{i}}{\phi''(z_j) t}} \begin{pmatrix} \mathrm{i}\overline{\beta}\Omega_p^{\hat{\underline{\nu}}^*,\infty_+}(z_j) - \mathrm{i}\beta\Omega_p^{\hat{\underline{\nu}}^*,\infty_+}(z_j^*) & -\mathrm{i}\beta\Omega_p^{\hat{\underline{\nu}},\infty_-}(z_j) + \mathrm{i}\overline{\beta}\Omega_p^{\hat{\underline{\nu}},\infty_-}(z_j^*) \end{pmatrix} \\ &\quad + O(t^{-\alpha}). \end{aligned} \tag{5.19}$$

Note that the right-hand side is real-valued for $p \in \pi^{-1}(\mathbb{R})\backslash\Sigma$ since $\overline{\Omega_p^{\hat{\underline{\nu}},\infty_\pm}(\overline{q})} = \Omega_{\overline{p}}^{\hat{\underline{\nu}},\infty_\pm}(q)$ implies

$$\Omega_p^{\hat{\underline{\nu}},\infty_\pm}(z_j^*) = \overline{\Omega_p^{\hat{\underline{\nu}},\infty_\pm}(z_j)}, \qquad p \in \pi^{-1}(\mathbb{R})\backslash\Sigma. \tag{5.20}$$

Since we need the asymptotic expansions around ∞_- we note

Lemma 5.1. *We have*

$$\Omega_p^{\hat{\underline{\nu}},\infty_+}(z_j) = \Lambda_0^{\hat{\underline{\nu}}} + \Lambda_1^{\hat{\underline{\nu}}}\frac{1}{z} + O(\frac{1}{z^2}) \tag{5.21}$$

for $p = (z, -)$ *near* ∞_-, *where*

$$\Lambda_0^{\hat{\underline{\nu}}} = \Omega_{\infty_-}^{\hat{\underline{\nu}},\infty_+}(z_j) = \Omega_{\infty_-}^{\hat{\underline{\nu}}^*,\infty_+}(z_j) = \omega_{\infty_-\,\infty_+}(z_j) + \sum_{k,\ell} c_{k\ell}(\hat{\underline{\nu}}) \int_{\infty_+}^{\infty_-} \omega_{\hat{\nu}_\ell,0}\zeta_k(z_j) \tag{5.22}$$

and

$$\begin{aligned} \Lambda_1^{\hat{\underline{\nu}}} &= \omega_{\infty_-,0}(z_j) + \sum_{k,\ell} c_{k\ell}(\hat{\underline{\nu}})\omega_{\hat{\nu}_\ell,0}(\infty_-)\zeta_k(z_j) \\ &= \omega_{\infty_-,0}(z_j) - \sum_{k,\ell} c_{k\ell}(\hat{\underline{\nu}}^*)\omega_{\hat{\nu}_\ell^*,0}(\infty_+)\zeta_k(z_j). \end{aligned} \tag{5.23}$$

Proof. To see $\Omega^{\hat{\underline{\nu}}}_{\infty_-}(z_j) = \Omega^{\hat{\underline{\nu}}^*}_{\infty_-}(z_j)$ note $c_{k\ell}(\hat{\underline{\nu}}^*) = -c_{k\ell}(\hat{\underline{\nu}})$ and $\int_{\infty_+}^{\infty_-} \omega_{\hat{\nu}_\ell^*,0} = \int_{\infty_-}^{\infty_+} \omega_{\hat{\nu}_\ell,0}$. □

Observe that since $c_{k\ell}(\hat{\underline{\nu}}) \in \mathbb{R}$ and $\int_{\infty_+}^{\infty_-} \omega_{\hat{\nu}_\ell,0} \in \mathbb{R}$ we have $\Lambda_0^{\hat{\underline{\nu}}} \in \mathrm{i}\mathbb{R}$.

Remark 5.2. *Note that the Abelian integral appearing in the previous lemma is explicitly given by*

$$\omega_{\infty_-,0} = \frac{-\pi^{g+1} + \frac{1}{2}\sum_{j=0}^{2g+1} E_j \pi^g + P_{\infty_-,0}(\pi) + R_{2g+2}^{1/2}}{R_{2g+2}^{1/2}} d\pi, \tag{5.24}$$

with $P_{\infty_-,0}$ a polynomial of degree $g-1$ which has to be determined from the normalization.

Similarly,

$$\omega_{\hat{\nu},0} = \frac{R_{2g+2}^{1/2} + R_{2g+2}^{1/2}(\hat{\nu}) + \frac{R'_{2g+2}(\hat{\nu})}{2R_{2g+2}^{1/2}(\hat{\nu})}(\pi-\nu) + P_{\hat{\nu},0}(\pi)\cdot(\pi-\nu)^2}{2(\pi-\nu)^2 R_{2g+2}^{1/2}} d\pi, \tag{5.25}$$

with $P_{\hat{\nu},0}$ a polynomial of degree $g-1$ which has to be determined from the normalization.

As in the previous section, the asymptotics can be read off by using

$$m^3(p) = d(\infty_-) m^6(p) \begin{pmatrix} \frac{1}{d(p^*)} & 0 \\ 0 & \frac{1}{d(p)} \end{pmatrix} \tag{5.26}$$

for p near ∞_- and comparing with (3.25). We obtain

$$A_+(n,t)^2 = \frac{1}{d(\infty_-)} \left(1 + \sqrt{\frac{\mathrm{i}}{\phi''(z_j)t}} \left(\mathrm{i}\overline{\beta}\Lambda_0^{\hat{\underline{\nu}}} - \mathrm{i}\beta\overline{\Lambda_0^{\hat{\underline{\nu}}}} \right) \right) + O(t^{-\alpha}) \tag{5.27}$$

and

$$B_+(n,t) = -d_1 - \sqrt{\frac{\mathrm{i}}{\phi''(z_j)t}} \left(\mathrm{i}\overline{\beta}\Lambda_1^{\hat{\underline{\nu}}^*} - \mathrm{i}\beta\overline{\Lambda_1^{\hat{\underline{\nu}}^*}} \right) + O(t^{-\alpha}), \tag{5.28}$$

for any $\alpha < 1$. Theorem 1.4 and hence also Theorems 1.1 and 1.3 are now proved under the assumption that $R(p)$ admits an analytic extension (which will be true if in our decay assumption (1.2) the weight n^6 is replaced by $\exp(-\varepsilon|n|)$ for some $\varepsilon > 0$) to be able to make our contour deformations. We will show how to get rid of this assumption by analytic approximation in the next section.

Summarizing, let us emphasize that the general significance of the method developed in this section is this: even when a Riemann-Hilbert problem needs to be considered on an algebraic variety, a localized parametrix Riemann-Hilbert problem need only be solved in the complex plane and the local solution can then be glued to the global Riemann-Hilbert solution on the variety. After this gluing procedure the resulting Riemann-Hilbert problem on the variety is asymptotically small and can be solved asymptotically (on the variety) by virtue of the associated singular integral equations.

The method described in this section can thus provide the higher order asymptotics *also* in the collisonless shock and Painlevé regions mentioned in the Introduction, by using existing results in ([9], [6]).

6. Analytic Approximation

In this section we want to show how to get rid of the analyticity assumption on the reflection coefficient $R(p)$. To this end we will split $R(p)$ into an analytic part $R_{a,t}$ plus a *small* residual term $R_{r,t}$ following the ideas of [6] (see also [29, Sect. 6]). The analytic part will be moved to regions of the Riemann surface while the residual term remains on $\Sigma = \pi^{-1}\big(\sigma(H_q)\big)$. This needs to be done in such a way that the residual term is of $O(t^{-1})$ and the growth of the analytic part can be controlled by the decay of the phase.

In order to avoid problems when one of the poles ν_j hits Σ, we have to make the approximation in such a way that the nonanalytic residual term vanishes at the band edges. That is, split R according to

$$\begin{aligned} R(p) =& R(E_{2j})\frac{z-E_{2j}}{E_{2j+1}-E_{2j}} + R(E_{2j+1})\frac{z-E_{2j+1}}{E_{2j}-E_{2j+1}} \\ & \pm \sqrt{z-E_{2j}}\sqrt{z-E_{2j+1}}\tilde{R}(p), \qquad p=(z,\pm), \end{aligned} \tag{6.1}$$

and approximate $\tilde{R}$. Note that if $R \in C^l(\Sigma)$, then $\tilde{R} \in C^{l-1}(\Sigma)$.

We will use different splittings for different bands depending on whether the band contains our stationary phase point $z_j(n/t)$ or not. We will begin with some preparatory lemmas.

For the bands containing no stationary phase points we will use a splitting based on the following Fourier transform associated with the background operator H_q. Given $R \in C^l(\Sigma)$ we can write

$$R(p) = \sum_{n\in\mathbb{Z}} \hat{R}(n)\psi_q(p,n,0), \tag{6.2}$$

where $\psi_q(p,x,t)$ denotes the time-dependent Baker–Akhiezer function and (cf. [10], [11])

$$\hat{R}(n) = \frac{1}{2\pi\mathrm{i}}\oint_\Sigma R(p)\psi_q(p^*,n,0)\frac{\mathrm{i}\prod_{j=1}^g(\pi(p)-\mu_j)}{R_{2g+2}^{1/2}(p)}d\pi(p). \tag{6.3}$$

If we make use of (2.12), the above expression for $R(p)$ is of the form

$$R(p) = \sum_{n\in\mathbb{Z}} \hat{R}(n)\theta_q(p,n,0)\exp\big(\mathrm{i}nk(p)\big). \tag{6.4}$$

where $k(p) = -\mathrm{i}\int_{E_0}^p \omega_{\infty_+\,\infty_-}$ and $\theta_q(p,n,t)$ collects the remaining parts in (2.12).

Using $k(p)$ as a new coordinate and performing l integration by parts one obtains

$$|\hat{R}(n)| \le \frac{const}{1+|n|^l} \tag{6.5}$$

provided $R \in C^l(\Sigma)$.

Lemma 6.1. *Suppose $\hat{R} \in \ell^1(\mathbb{Z})$, $n^l\hat{R}(n) \in \ell^1(\mathbb{Z})$ and let $\beta > 0$ be given. Then we can split $R(p)$ according to*

$$R(p) = R_{a,t}(p) + R_{r,t}(p),$$

such that $R_{a,t}(p)$ is analytic for in the region $0 < \mathrm{Im}(k(p)) < \varepsilon$ and

$$|R_{a,t}(p)\mathrm{e}^{-\beta t}| = O(t^{-l}), \quad 0 < \mathrm{Im}(k(p)) < \varepsilon, \tag{6.6}$$

$$|R_{r,t}(p)| = O(t^{-l}), \quad p \in \Sigma. \tag{6.7}$$

Proof. We choose

$$R_{a,t}(p) = \sum_{n=-N(t)}^{\infty} \hat{R}(n)\theta_q(p,n,0)\exp\big(\mathrm{i}nk(p)\big)$$

with $N(t) = \lfloor \frac{\beta_0}{\varepsilon} t \rfloor$ for some positive $\beta_0 < \beta$. Then, for $0 < \mathrm{Im}(k(p)) < \varepsilon$,

$$\begin{aligned} \big|R_{a,t}(k)\mathrm{e}^{-\beta t}\big| &\le C\mathrm{e}^{-\beta t} \sum_{n=-N(t)}^{\infty} |\hat{R}(n)|\mathrm{e}^{-\mathrm{Im}(k(p))n} \\ &\le C\mathrm{e}^{-\beta t}\mathrm{e}^{N(t)\varepsilon}\|F\|_1 = \|\hat{R}\|_1 \mathrm{e}^{-(\beta-\beta_0)t}, \end{aligned}$$

which proves the first claim. Similarly, for $p \in \Sigma$,

$$|R_{r,t}(k)| \le C \sum_{n=N(t)+1}^{\infty} \frac{n^l|\hat{R}(-n)|}{n^l} \le C\frac{\|n^l\hat{R}(-n)\|_{\ell^1(\mathbb{N})}}{N(t)^l} \le \frac{\tilde{C}}{t^l}$$

□

For the band which contains $z_j(n/t)$ we need to take the small vicinities of the stationary phase points into account. Since the phase is cubic near these points, we cannot use it to dominate the exponential growth of the analytic part away from Σ. Hence we will take the phase as a new variable and use the Fourier transform with respect to this new variable. Since this change of coordinates is singular near the stationary phase points, there is a price we have to pay, namely, requiring additional smoothness for $R(p)$.

Without loss of generality we will choose the path of integration in our phase $\phi(p)$, defined in (3.17), such that $\phi(p)$ is continuous (and thus analytic) in $D_{j,1}$ with continuous limits on the boundary (cf. Figure 2). We begin with

Lemma 6.2. *Suppose $R(p) \in C^5(\Sigma)$. Then we can split $R(p)$ according to*

$$R(p) = R_0(p) + (\pi(p) - \pi(z_j))H(p), \qquad p \in \Sigma \cap D_{j,1}, \tag{6.8}$$

where $R_0(p)$ is a real rational function on $\mathbb{M}$ such that $H(p)$ vanishes at z_j, z_j^* of order three and has a Fourier series

$$H(p) = \sum_{n\in\mathbb{Z}} \hat{H}(n)\mathrm{e}^{n\omega_0\phi(p)}, \qquad \omega_0 = \frac{2\pi\mathrm{i}}{\phi(z_j) - \phi(z_j^*)} > 0, \tag{6.9}$$

with $n\hat{H}(n)$ summable. Here ϕ denotes the phase defined in (3.17).

Proof. We begin by choosing a rational function $R_0(p) = a(z) + b(z)R_{2g+2}^{1/2}(p)$ with $p = (z, \pm)$ such that $a(z)$, $b(z)$ are real-valued polynomials which are chosen such that $a(z)$ matches the values of $\mathrm{Re}(R(p))$ and its first four derivatives at z_j and $\mathrm{i}^{-1}b(z)R_{2g+2}^{1/2}(p)$ matches the values of $\mathrm{Im}(R(p))$ and its first four derivatives at z_j. Since $R(p)$ is C^5 we infer that $H(p) \in C^4(\Sigma)$ and it vanishes together with its first three derivatives at z_j, z_j^*.

Note that $\phi(p)/\mathrm{i}$, where ϕ is defined in (3.17) has a maximum at z_j^* and a minimum at z_j. Thus the phase $\phi(p)/\mathrm{i}$ restricted to $\Sigma \cap D_{j,1}$ gives a one to one coordinate transform $\Sigma \cap D_{j,1} \to [\phi(z_j^*)/\mathrm{i}, \phi(z_j)/\mathrm{i}]$ and we can hence express $H(p)$ in this new coordinate. The coordinate transform locally looks like a cube root near z_j and z_j^*, however, due to our assumption that H vanishes there, H is still C^2 in this new coordinate and the Fourier transform with respect to this new coordinates exists and has the required properties. □

Moreover, as in Lemma 6.1 we obtain:

Lemma 6.3. *Let $H(p)$ be as in the previous lemma. Then we can split $H(p)$ according to $H(p) = H_{a,t}(p) + H_{r,t}(p)$ such that $H_{a,t}(p)$ is analytic in the region* $\mathrm{Re}(\phi(p)) < 0$ *and*

$$|H_{a,t}(p)\mathrm{e}^{\phi(p)t/2}| = O(1),\ p \in \overline{D_{j,1}}, \quad |H_{r,t}(p)| = O(t^{-1}),\ p \in \Sigma. \tag{6.10}$$

Proof. We choose $H_{a,t}(p) = \sum_{n=-K(t)}^{\infty} \hat{H}(n)\mathrm{e}^{n\omega_0\phi(p)}$ with $K(t) = \lfloor t/(2\omega_0)\rfloor$. Then we can proceed as in Lemma 6.1:

$$|H_{a,t}(p)\mathrm{e}^{\phi(p)t/2}| \le \|\hat{H}\|_1 |\mathrm{e}^{-K(t)\omega_0\phi(p)+\phi(p)t/2}| \le \|\hat{H}\|_1$$

and

$$|H_{r,t}(p)| \le \frac{1}{K(t)} \sum_{n=K(t)+1}^{\infty} n|\hat{H}(-n)| \le \frac{C}{t}.$$

□

Clearly an analogous splitting exists for $p \in \Sigma \cap D_{j2}$.

Now we are ready for our analytic approximation step. First of all recall that our jump is given in terms $\tilde{b}_\pm$ and $\tilde{B}_\pm$ defined in (4.29) and (4.31), respectively. While $\tilde{b}_\pm$ are already in the correct form for our purpose, this is not true for $\tilde{B}_\pm$ since they contain the non-analytic expression $|T(p)|^2$. To remedy this we will rewrite $\tilde{B}_\pm$ in terms of the left rather than the right scattering data. For this purpose let us use the notation $R_r(p) \equiv R_+(p)$ for the right and $R_l(p) \equiv R_-(p)$ for the left reflection coefficient. Moreover, let $d_r(p,x,t) = d(p,x,t)$ and $d_l(p,x,t) \equiv T(p)/d(p,x,t)$.

With this notation we have

$$J^4(p) = \begin{cases} \tilde{b}_-(p)^{-1}\tilde{b}_+(p), & \pi(p) > z_j(n/t), \\ \tilde{B}_-(p)^{-1}\tilde{B}_+(p), & \pi(p) < z_j(n/t), \end{cases} \tag{6.11}$$

where

$$\tilde{b}_- = \begin{pmatrix} 1 & \frac{d_r(p,x,t)}{d_r(p^*,x,t)} R_r(p^*)\Theta(p^*)\mathrm{e}^{-t\phi(p)} \\ 0 & 1 \end{pmatrix},$$

$$\tilde{b}_+ = \begin{pmatrix} 1 & 0 \\ \frac{d_r(p^*,x,t)}{d_r(p,x,t)} R_r(p)\Theta(p)\mathrm{e}^{-t\phi(p)} & 1 \end{pmatrix},$$

and

$$\tilde{B}_- = \begin{pmatrix} 1 & 0 \\ -\frac{d_{r,-}(p^*,x,t)}{d_{r,-}(p,x,t)} \frac{R_r(p)\Theta(p)}{|T(p)|^2}\mathrm{e}^{t\,\phi(p)} & 1 \end{pmatrix},$$

$$\tilde{B}_+ = \begin{pmatrix} 1 & -\frac{d_{r,+}(p,x,t)}{d_{r,+}(p^*,x,t)} \frac{R_r(p^*)\Theta(p^*)}{|T(p)|^2}\mathrm{e}^{-t\,\phi(p)} \\ 0 & 1 \end{pmatrix}.$$

Using (3.7) we can write

$$\tilde{B}_- = \begin{pmatrix} 1 & 0 \\ \frac{d_l(p^*,x,t)}{d_l(p,x,t)} R_l(p)\Theta(p)\mathrm{e}^{-t\phi(p)} & 1 \end{pmatrix},$$

$$\tilde{B}_+ = \begin{pmatrix} 1 & \frac{d_l(p,x,t)}{d_l(p^*,x,t)} R_l(p^*)\Theta(p^*)\mathrm{e}^{-t\phi(p)} \\ 0 & 1 \end{pmatrix}.$$

Now we split $R_r(p) = R_{a,t}(p) + R_{r,t}(p)$ by splitting $\tilde{R}_r(p)$ defined via (6.1) according to Lemma 6.1 for $\pi(p) \in [E_{2k}, E_{2k+1}]$ with $k < j$ (i.e., not containing $z_j(n/t)$) and according to Lemma 6.3 for $\pi(p) \in [E_{2j}, z_j(n/t)]$. In the same way we split $R_l(p) = R_{a,t}(p) + R_{r,t}(p)$ for $\pi(p) \in [z_j(n/t), E_{2j+1}]$ and $\pi(p) \in [E_{2k}, E_{2k+1}]$ with $k > j$. For β in Lemma 6.1 we can choose

$$\beta = \begin{cases} \min_{p\in C_k} -\mathrm{Re}(\phi(p)) > 0, & \pi(p) > z_j(n/t), \\ \min_{p\in C_k} \mathrm{Re}(\phi(p)) > 0, & \pi(p) < z_j(n/t). \end{cases} \tag{6.12}$$

In this way we obtain

$$\tilde{b}_\pm(p) = \tilde{b}_{a,t,\pm}(p)\tilde{b}_{r,t,\pm}(p) = \tilde{b}_{r,t,\pm}(p)\tilde{b}_{a,t,\pm}(p),$$
$$\tilde{B}_\pm(p) = \tilde{B}_{a,t,\pm}(p)\tilde{B}_{r,t,\pm}(p) = \tilde{B}_{r,t,\pm}(p)\tilde{B}_{a,t,\pm}(p).$$

Here $\tilde{b}_{a,t,\pm}(p)$, $\tilde{b}_{r,t,\pm}(p)$ (resp. $\tilde{B}_{a,t,\pm}(p)$, $\tilde{B}_{r,t,\pm}(p)$) denote the matrices obtained from $\tilde{b}_\pm(p)$ (resp. $\tilde{B}_\pm(p)$) by replacing $R_r(p)$ (resp. $R_l(p)$) with $R_{a,t}(p)$, $R_{r,t}(p)$, respectively. Now we can move the analytic parts into regions of the Riemann surface as in Section 4 while leaving the rest on Σ. Hence, rather than (4.35), the jump now reads

$$J^5(p) = \begin{cases} \tilde{b}_{a,t,+}(p), & p \in C_k, \quad \pi(p) > z_j(n/t), \\ \tilde{b}_{a,t,-}(p)^{-1}, & p \in C_k^*, \quad \pi(p) > z_j(n/t), \\ \tilde{b}_{r,t,-}(p)^{-1}\tilde{b}_{r,t,+}(p), & p \in \Sigma, \quad \pi(p) > z_j(n/t), \\ \tilde{B}_{a,t,+}(p), & p \in C_k, \quad \pi(p) < z_j(n/t), \\ \tilde{B}_{a,t,-}(p)^{-1}, & p \in C_k^*, \quad \pi(p) < z_j(n/t), \\ \tilde{B}_{r,t,-}(p)^{-1}\tilde{B}_{r,t,+}(p), & p \in \Sigma, \quad \pi(p) < z_j(n/t). \end{cases} \tag{6.13}$$

By construction $R_{a,t}(p) = R_0(p) + (\pi(p) - \pi(z_j))H_{a,t}(p)$ will satisfy the required Lipschitz estimate in a vicinity of the stationary phase points (uniformly in t) and the jump will be $J^5(p) = \mathbb{I} + O(t^{-1})$. The remaining parts of Σ can be handled analogously and hence we can proceed as in Section 5.

7. Conclusion

We have considered here the stability problem for the periodic Toda lattice under a short-range perturbation. We have discovered that a nonlinear stationary phase method (cf. [6], [23]) is applicable and as a result we have shown that the long-time behavior of the perturbed lattice is described by a modulated lattice which undergoes a continuous phase transition (in the Jacobian variety).

We have extended the well-known nonlinear stationary phase method of Deift and Zhou to Riemann–Hilbert problems living in an algebraic variety. Even though the studied example involves a hyperelliptic Riemann surface the method is easily extended to surfaces with several sheets. We were forced to tackle such Riemann–Hilbert problems by the very problem, since there is no way we could use the symmetries needed to normalize the Riemann–Hilbert problem of Section 3 without including a second sheet. We believe that this is one significant novelty of our contribution.

Although the most celebrated applications of the deformation method initiated by [6] for the asymptotic evaluation of solutions of Riemann–Hilbert factorization problems have been in the areas orthogonal polynomials, random matrices and combinatorial probability, most mathematical innovations have appeared in the study of nonlinear dispersive PDEs or systems of ODEs (e.g. [6], [9], [26]). It is thus interesting that another mathematical extension of the theory arises in the study of an innocent looking stability problem for the periodic Toda lattice.

On the other hand, we see the current work as part of a more general program. The next step is to consider initial data that are a short pertrubation of a finite gap solution at $\pm\infty$ but with different genus at each infinity, a generalized "Toda shock" problem. Then a similar picture arises (modulation regions separated by "periodic" regions) but now the genus of the modulated solution can also jump between different regions of the (n, t)-plane. The understanding of the more general picture is crucial for the understanding of the following very interesting problem.

Consider the Toda lattice on the quarter plane $n, t \geq 0$ with initial data that are asymptotically periodic (or constant) as $n \to \infty$ and periodic data $a_0(t)$ and $b_0(t)$. What is the long time behavior of the system?

Special cases of this problem correspond to the generalized Toda shock described above. A full understanding of the periodic forcing problem thus requires an understanding of the setting described in this paper.

A related publication is for example [4] where the authors study such a periodic forcing problem (for NLS rather than Toda) by extending the inverse scattering method of Fokas (e.g. [14]) for integrable systems in the quarter plane and actually arrive at a Riemann–Hilbert problem living in a Riemann surface. We thus expect our methods to have a wide applicability.

Appendix A. A singular integral equation

In the complex plane, the solution of a Riemann–Hilbert problem can be reduced to the solution of a singular integral equation (see [2]) via a Cauchy-type formula. In our case the underlying space is a Riemann surface $\mathbb{M}$. The purpose of this appendix is to produce a more general Cauchy-type formula to Riemann–Hilbert problems of the type

$$\begin{aligned} & m_+(p) = m_-(p) J(p), \quad p \in \Sigma, \\ & (m_1) \geq -\mathcal{D}_{\underline{\hat{\mu}}^*}, \quad (m_2) \geq -\mathcal{D}_{\underline{\hat{\mu}}}, \\ & m(\infty_+) = m_0 \in \mathbb{C}^2. \end{aligned} \tag{A.1}$$

Once one has such an integral formula, it is easy to "perturb" it and prove that small changes in the data produce small changes in the solution of the Riemann-Hilbert problem.

Concerning the jump contour Σ and the jump matrix J we will make the following assumptions:

Hypothesis H. A.1. *Let Σ consist of a finite number of smooth oriented finite curves in $\mathbb{M}$ which intersect at most finitely many times with all intersections being transversal. The divisor $\mathcal{D}_{\underline{\hat{\mu}}}$ is nonspecial. The contour Σ does neither contain $\infty_\pm$ nor any of the points $\underline{\hat{\mu}}$ and that the jump matrix J is nonsingular and can be* *factorized according to* $J = b_-^{-1} b_+ = (\mathbb{I} - w_-)^{-1}(\mathbb{I} + w_+)$, *where* $w_\pm = \pm(b_\pm - \mathbb{I})$ *are continuous.*

Remark A.2. *(i). We dropped our symmetry requirement*

$$m(p^*) = m(p) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{A.2}$$

here since it only is important in the presence of solitons. However, if both Σ *and* $w_\pm$ *are compatible with this symmetry, then one can restrict all operators below to the corresponding symmetric subspaces implying a symmetric solution. Details will be given in* [29].

(ii). The assumption that none of the poles $\underline{\hat{\mu}}$ *lie on our contour* Σ *can be made without loss of generality if the jump is analytic since we can move the contour a little without changing the value at* ∞_- *(which is the only value we are eventually interested in). Alternatively, the case where one (or more) of the poles* $\hat{\mu}_j$ *lies on* Σ *can be included if one assumes that* $w_\pm$ *has a first order zero at* $\hat{\mu}_j$. *In fact, in this case one can replace* $\mu(s)$ *by* $\tilde{\mu}(s) = (\pi(s) - \mu_j)\mu(s)$ *and* $w_\pm(s)$ *by* $\tilde{w}_\pm(s) = (\pi(s) - \mu_j)^{-1} w_\pm(s)$.

Otherwise one could also assume that the matrices $w_\pm$ *are Hölder continuous and vanish at such points. Then one can work with the weighted measure* $-\mathrm{i} R_{2g+2}^{1/2}(p) d\pi$ *on* Σ. *In fact, one can show that the Cauchy operators are still bounded in this weighted Hilbert space (cf.* [18, Thm. 4.1]*).*

Our first step is to replace the classical Cauchy kernel by a "generalized" Cauchy kernel appropriate to our Riemann surface. In order to get a single valued kernel we need again to admit g poles. We follow the construction from [36, Sec. 4].

Lemma A.3. *Let* $\mathcal{D}_{\underline{\hat{\mu}}}$ *be nonspecial and introduce the differential*

$$\Omega_p^{\underline{\hat{\mu}}} = \omega_{p\,\infty_+} + \sum_{j=1}^{g} I_j^{\underline{\hat{\mu}}}(p) \zeta_j, \tag{A.3}$$

where

$$I_j^{\underline{\hat{\mu}}}(p) = \sum_{\ell=1}^{g} c_{j\ell}(\underline{\hat{\mu}}) \int_{\infty_+}^{p} \omega_{\hat{\mu}_\ell,0}. \tag{A.4}$$

Here $\omega_{p\,,q}$ *is the (normalized) Abelian differential of the third kind with poles at* p, q *(cf. Remark 4.4) and* $\omega_{q,0}$ *is the (normalized) Abelian differential of the second kind with a second order pole at* q *(cf. Remark 5.2) and the matrix* $c_{j\ell}$ *is defined as the inverse matrix of* $\eta_\ell(\hat{\mu}_j)$, *where* $\zeta_\ell = \eta_\ell(z)dz$ *is the chart expression in a local chart near* $\hat{\mu}_j$ *(the same chart used to define* $\omega_{\hat{\mu}_j,0}$*).*

Then $\Omega_p^{\underline{\hat{\mu}}}$ *is single valued as a function of* p *with first order poles at the points* $\underline{\hat{\mu}}$.

Proof. Note that $I_j^{\underline{\hat{\mu}}}(p)$ has first order poles at the points $\underline{\hat{\mu}}$ hence it remains to show that the constants $c_{j\ell}(\underline{\hat{\mu}})$ are chosen such that $\Omega_p^{\underline{\hat{\mu}}}$ is single valued (cf. the discussion in the proof of Theorem 4.3). That is,

$$\int_{b_k} dI_j^{\underline{\hat{\mu}}} = \sum_{\ell=1}^{g} c_{j\ell} \int_{b_k} \omega_{\hat{\mu}_\ell,0} = \sum_{\ell=1}^{g} c_{j\ell} \eta_k(\hat{\mu}_\ell) = \delta_{jk},$$

where $\zeta_k = \eta_k(z)dz$ is the chart expression in a local chart near $\hat{\mu}_\ell$ (here the b_k periods are evaluated using the usual bilinear relations, see [16, Sect. III.3] or [41, Sect. A.2]). That the matrix $\eta_k(\hat{\mu}_\ell)$ is indeed invertible can be seen as follows: If $\sum_{k=1}^g \eta_k(\hat{\mu}_\ell)c_k = 0$ for $1 \le \ell \le g$, then the divisor of $\zeta = \sum_{k=1}^g c_k\zeta_k$ satisfies $(\zeta) \ge \mathcal{D}_{\underline{\hat{\mu}}}$. But since we assumed the divisor $\mathcal{D}_{\underline{\hat{\mu}}}$ to be nonspecial, $i(\mathcal{D}_{\underline{\hat{\mu}}}) = 0$, we have $\zeta = 0$ implying $c_k = 0$. □

Next we show that the Cauchy kernel introduced in (A.3) has indeed the correct properties. We will abbreviate $L^p(\Sigma) = L^p(\Sigma, \mathbb{C}^2)$.

Theorem A.4. *Set*

$$\underline{\Omega}_p^{\underline{\hat{\mu}}} = \begin{pmatrix} \Omega_p^{\underline{\hat{\mu}}^*} & 0 \\ 0 & \Omega_p^{\underline{\hat{\mu}}} \end{pmatrix} \tag{A.5}$$

and define the matrix operators as follows. Given a 2×2 matrix f defined on Σ with Hölder continuous entries, let

$$(Cf)(p) = \frac{1}{2\pi \mathrm{i}} \int_\Sigma f\, \underline{\Omega}_p^{\underline{\hat{\mu}}}, \quad \text{for} \quad p \notin \Sigma, \tag{A.6}$$

and

$$(C_\pm f)(q) = \lim_{p \to q \in \Sigma} (Cf)(p) \tag{A.7}$$

from the left and right of Σ respectively (with respect to its orientation). Then

(i) *The operators $C_\pm$ are given by the Plemelj formulas*

$$(C_+ f)(q) - (C_- f)(q) = f(q),$$
$$(C_+ f)(q) + (C_- f)(q) = \frac{1}{\pi \mathrm{i}} \fint_\Sigma f\, \underline{\Omega}_q^{\underline{\hat{\mu}}},$$

and extend to bounded operators on $L^2(\Sigma)$. Here $\fint$ denotes the principal value integral, as usual.

(ii) *Cf is a meromorphic function off Σ, with divisor given by $((Cf)_{j1}) \ge -\mathcal{D}_{\underline{\hat{\mu}}^*}$ and $((Cf)_{j2}) \ge -\mathcal{D}_{\underline{\hat{\mu}}}$.*

(iii) *$(Cf)(\infty_+) = 0$.*

Proof. In a chart $z = z(p)$ near $q_0 \in \Sigma$, the differential $\Omega_q^{\underline{\hat{\mu}}} = (\frac{1}{z - z(q)} + O(1))dz$ and hence the first part follows as in the Cauchy case on the complex plane (cf. [33] or [39]) using a partition of unity. To see (ii) note that the integral over $\omega_{p\,\infty_+}$ is a (multivalued) holomorphic function, while the integral over the rest is a linear combination of the (multivalued) meromorphic functions $I_j^{\underline{\hat{\mu}}}$ respectively $I_j^{\underline{\hat{\mu}}^*}$. By construction, $I_j^{\underline{\hat{\mu}}}$ has at most simple poles at the points $\underline{\hat{\mu}}$ and thus (ii) follows. Finally, to see (iii) observe that $\omega_{p\,\infty_+}$ restricted to Σ converges uniformly to zero as $p \to \infty_+$ (cf. (4.14)). Moreover, $I_j^{\underline{\hat{\mu}}^*}(\infty_+) = 0$ and hence (iii) holds. □

Now, let the operator $C_w : L^2(\Sigma) \to L^2(\Sigma)$ be defined by

$$C_w f = C_+(f w_-) + C_-(f w_+) \tag{A.8}$$

for a 2×2 matrix valued f, where

$$w_+ = b_+ - \mathbb{I} \quad \text{and} \quad w_- = \mathbb{I} - b_-.$$

Theorem A.5. *Assume Hypothesis A.1 and let $m_0 \in \mathbb{C}^2$ be given.*

Assume that μ solves the singular integral equation

$$\mu = m_0 + C_w \mu \quad \text{in} \quad L^2(\Sigma). \tag{A.9}$$

Then m be defined by the integral formula

$$m = m_0 + C(\mu w) \quad \text{on } \mathbb{M} \setminus \Sigma, \tag{A.10}$$

where $w = w_+ + w_-$, is a solution of the meromorphic Riemann–Hilbert problem (A.1).

Conversely, if m is a solution of (A.1)*, then μ defined via $\mu = m_\pm b_\pm^{-1}$ solves* (A.9).

Proof. Suppose μ solves (A.9). To show that m defined above solves (A.1) note that

$$m_\pm = \mathbb{I} + C_\pm(\mu w).$$

Thus, using $C_+ - C_- = \mathbb{I}$ and the definition of C_w we obtain

$$\begin{aligned} m_+ &= (m_0 + C_+(\mu w)) = (m_0 + C_+(\mu w_+) + C_+(\mu w_-)) \\ &= (m_0 + \mu w_+ + C_-(\mu w_+) + C_+(\mu w_-)) = (m_0 + \mu w_+ + C_w \mu) \\ &= \mu(\mathbb{I} + w_+) \end{aligned}$$

and similarly $m_- = \mu(\mathbb{I} - w_-)$. Hence $m_+ b_+^{-1} = \mu = m_- b_-^{-1}$ and thus $m_+ = m_-(b_-)^{-1} b_+$. This proves the jump condition. That m has the right devisor and the correct normalization at ∞_+ follows from Theorem A.4 (ii) and (iii), respectively.

Conversely, if m is a solution of the Riemann–Hilbert problem (A.1), then we can set $\mu = m_+ b_+^{-1} = m_- b_-^{-1}$ and define $\tilde{m}$ by (A.10). To see that in fact $m = \tilde{m}$ holds, observe that both satisfy the same additive jump condition $m_+ - m_- = \tilde{m}_+ - \tilde{m}_- = \mu w$. Hence the difference $m - \tilde{m}$ has no jump and thus must be meromorphic. Moreover, by the divisor conditions $(m_1 - \tilde{m}_1) \geq -\mathcal{D}_{\underline{\hat{\mu}}^*}$ and $(m_2 - \tilde{m}_2) \geq -\mathcal{D}_{\underline{\hat{\mu}}}$, the Riemann–Roch theorem implies that $m - \tilde{m}$ is constant. By our normalization at ∞_+ this constant must be the zero vector. Thus $m = \tilde{m}$ and as before one computes

$$m_+ = \mu b_+ - \mu + m_0 + C_w \mu,$$

showing that (A.9) holds. □

Remark A.6. *(i). The theorem stated above does not address uniqueness. This will be done in Theorem B.1 under an additional symmetry assumption.*

(ii). The notation b_+, b_- is meant to make one think of the example $J^3 = (b_-)^{-1} b_+$ in Section 4, but the theorem above is fairly general. In particular it also applies to the trivial factorizations $J^3 = \mathbb{I} J^3 = J^3 \mathbb{I}$.

We are interested in the formula (A.10) evaluated at ∞_-. We write it as

$$\begin{aligned} m(\infty_-) &= (m_0 + C(\mu w))(\infty_-) \\ &= m_0 + \int_\Sigma (\mathbb{I} - C_w)^{-1}(m_0)\, w\, \Omega_{\infty_-}^{\underline{\hat{\mu}}} \end{aligned} \tag{A.11}$$

and we perturb it with respect to w while keeping the contour Σ fixed.

Hence we have a formula for the solution of our Riemann–Hilbert problem $m(z)$ in terms of $(\mathbb{I} - C_w)^{-1} m_0$ and this clearly raises the question of bounded invertibility of $\mathbb{I} - C_w$. This follows from Fredholm theory (cf. e.g. [47]):

Lemma A.7. *Assume Hypothesis A.1. Then the operator* $\mathbb{I} - C_w$ *is Fredholm of index zero,*

$$\operatorname{ind}(\mathbb{I} - C_w) = 0. \tag{A.12}$$

Proof. Using the Bishop–Kodama theorem [27] we can approximate $w_\pm$ by functions which are analytic in a neighborhood of Σ and hence, since the norm limits of compact operators are compact, we can assume that $w_\pm$ are analytic in a neighborhood of Σ without loss of generality.

First of all one can easily check that

$$(\mathbb{I} - C_w)(\mathbb{I} - C_{-w}) = (\mathbb{I} - C_{-w})(\mathbb{I} - C_w) = \mathbb{I} - T_w, \tag{A.13}$$

where $T_w(f) = C_-[C_-(fw_+)w_+]$. But $T_w(f)$ is a compact operator. Indeed, suppose $f_n \in L^2(\Sigma)$ converges weakly to zero. We will show that $\|T_w f_n\|_{L^2} \to 0$.

Using the analyticity of w_+ in a neighborhood of Σ and the definition of C_-, we can slightly deform the contour Σ to some contour Σ' close to Σ, on the right, and have, by Cauchy's theorem,

$$T_w f_n(p) = \frac{1}{2\pi \mathrm{i}} \int_{\Sigma'} (C(f_n w_+)w_+) \underline{\Omega}_{\underline{p}}^{\hat{\underline{\mu}}}. \tag{A.14}$$

Now clearly $(C(f_n w_+)w_+)(p) \to 0$ as $n \to \infty$. and since also $|(C(f_n w_+)w_+)(p)| < \mathit{const}\, \|f_n\|_{L^2} \|w_+\|_{L^\infty} < \mathit{const}$ we infer $\|T_w f_n\|_{L^2} \to 0$ by virtue of the dominated convergence theorem.

Hence by [35, Thm. 1.4.3] $\mathbb{I} - C_w$ is Fredholm. Moreover, consider $\operatorname{ind}(\mathbb{I} - \varepsilon C_w)$ for $0 \le \varepsilon \le 1$ and recall that $\operatorname{ind}(\mathbb{I} - \varepsilon C_w)$ is continuous with respect to ε ([35, Thm. 1.3.8]). Since it is an integer, it has to be constant, that is, $\operatorname{ind}(\mathbb{I} - C_w) = \operatorname{ind}(\mathbb{I}) = 0$. □

By the Fredholm alternative, it follows that to show the bounded invertibility of $\mathbb{I} - C_w$ we only need to show that $\ker(\mathbb{I} - C_w) = 0$. The latter being equivalent to unique solvability of the corresponding vanishing Riemann–Hilbert problem.

Corollary A.8. *Assume Hypothesis A.1.*

A unique solution of the Riemann–Hilbert problem (A.1) *exists if and only if the corresponding vanishing Riemann–Hilbert problem, where the normalization condition is given by* $m(\infty_+) = \begin{pmatrix} 0 & 0 \end{pmatrix}$, *has at most one solution.*

We are interested in comparing two Riemann–Hilbert problems associated with respective jumps w_0 and w with $\|w - w_0\|_\infty$ small, where

$$\|w\|_\infty = \|w_+\|_{L^\infty(\Sigma)} + \|w_-\|_{L^\infty(\Sigma)}. \tag{A.15}$$

For such a situation we have the following result:

Theorem A.9. *Assume that for some data* w_0^t *the operator*

$$\mathbb{I} - C_{w_0^t} : L^2(\Sigma) \to L^2(\Sigma) \tag{A.16}$$

has a bounded inverse, where the bound is independent of t.

Furthermore, assume w^t *satisfies*

$$\|w^t - w_0^t\|_\infty \le \alpha(t) \tag{A.17}$$

for some function $\alpha(t) \to 0$ *as* $t \to \infty$. *Then* $(\mathbb{I} - C_{w^t})^{-1} : L^2(\Sigma) \to L^2(\Sigma)$ *also exists for sufficiently large* t *and the associated solutions of the Riemann–Hilbert problems* (A.1) *only differ by* $O(\alpha(t))$.

Proof. Follows easily by the Cauchy-type integral formula proved above, the boundedness of the Cauchy transform and the second resolvent identity.

More precisely, by the boundedness of the Cauchy transform, one has

$$\|(C_{w^t} - C_{w_0^t})\| \le const \|w\|_\infty.$$

Thus, by the second resolvent identity, we infer that $(\mathbb{I} - C_{w^t})^{-1}$ exists for large t and

$$\|(\mathbb{I} - C_{w^t})^{-1} - (\mathbb{I} - C_{w_0^t})^{-1}\| = O(\alpha(t)).$$

The claim now follows, since this implies $\|\mu^t - \mu_0^t\|_{L^2} = O(\alpha(t))$ where μ_0^t is defined in the obvious way as in (A.9) and thus $m^t(z) - m_0^t(z) = O(\alpha(t))$ uniformly in z away from Σ. □

Appendix B. A uniqueness theorem for factorization problems on a Riemann surface

In the case where the underlying spectral curve is the complex plane it is often useful to have a theorem guaranteeing existence of a solution of a Riemann–Hilbert problem under some symmetry conditions. One such is, for example, the Schwarz reflection theorem provided in [47]. In this section we state and prove an analogous theorem where the underlying spectral curve is our hyperelliptic curve with real branch cuts.

For any matrix (or vector) M we denote its adjoint (transpose of complex conjugate) as M^*. Then we have

Theorem B.1. *Assume in addition to Hypothesis A.1 assume that $\mu_j \in [E_{2j-1}, E_{2j}]$ and that Σ is symmetric under sheet exchange plus conjugation ($\Sigma = \overline{\Sigma}^*$) such that*

(i) $J(p^*) = J(\overline{p})^*$, for $p \in \Sigma \setminus \pi^{-1}(\sigma(H_q))$,

(ii) $\mathrm{Re}(J(p)) = \frac{1}{2}(J(p) + J(p)^*)$ is positive definite for $p \in \pi^{-1}(\sigma(H_q))$,

(iii) J is analytic in a neighborhood of Σ.

Then the vector Riemann–Hilbert problem (A.1) *on $\mathbb{M}$ has always a unique solution.*

Note here that the $+$-side of the contour is mapped to the $-$-side under sheet exchange. In particular, the theorem holds if $J = \mathbb{I}$, that is there is no jump, on $\pi^{-1}(\sigma(H_q))$.

Proof. By Corollary A.8 it suffices to show that the corresponding vanishing problem has only the trivial solution.

Our strategy is to apply Cauchy's integral theorem to

$$m(p) m^*(\overline{p}^*) = m_1(p)\overline{m_1(\overline{p}^*)} + m_2(p)\overline{m_2(\overline{p}^*)}.$$

To this end we will multiply it by a meromorphic differential $d\Omega$ which has zeros at $\underline{\mu}$ and $\underline{\mu}^*$ and simple poles at $\infty_\pm$ such that the differential $m(p)m^*(\overline{p}^*)d\Omega(p)$ is holomorphic away from the contour.

Indeed let

$$d\Omega = -\mathrm{i} \frac{\prod_{j=1}^g (\pi - \mu_j)}{R_{2g+2}^{1/2}} d\pi \tag{B.1}$$

and note that $\frac{\prod_j (z-\mu_j)}{R_{2g+2}^{1/2}(z)}$ is a Herglotz–Nevanlinna function. That is, it has positive imaginary part in the upper half-plane (and it is purely imaginary on $\sigma(H_q)$). Hence $m(p)\overline{m^T(p)}d\Omega(p)$ will be positive on $\pi^{-1}(\sigma(H_q))$.

Consider then the integral

$$\int_D m(p)m^*(\overline{p}^*)d\Omega(p), \tag{B.2}$$

where D is a $\overline{*}$-invariant contour consisting of one small loop in every connected component of $\mathbb{M}\backslash\Sigma$. Clearly the above integral is zero by Cauchy's residue theorem. We will deform D to a $\overline{*}$-invariant contour consisting of two parts, one, say D_+, wrapping around the part of Σ lying on Π_+ and the $+$ side of $\pi^{-1}(\sigma(H_q))$ and the other being $D_- = \overline{D_+}^*$.

For each component Σ_j of $\Sigma \setminus \pi^{-1}(\sigma(H_q))$ there are two contributions to the integral on the deformed contour:

$$\int_{\Sigma_j} m_+(p)m_-^*(\overline{p}^*)d\Omega = \int_{\Sigma_j} m_-(p)J(p)m_-^*(\overline{p}^*)d\Omega \quad \text{and}$$
$$\int_{-\Sigma_j} m_-(p)m_+^*(\overline{p}^*)d\Omega = \int_{-\Sigma_j} m_-(p)J^*(\overline{p}^*)m_-^*(\overline{p}^*)d\Omega.$$

Because of condition (i) the two integrals cancel each other.

In view of the above and using Cauchy's theorem, one gets

$$\begin{aligned}
0 &= \int_D m(p)m^*(\overline{p}^*)d\Omega \\
&= \int_{\pi^{-1}(\sigma(H_q))} [m_+(p)m_-^*(\overline{p}^*) + m_-(p)m_+^*(\overline{p}^*)]d\Omega \\
&= \int_{\pi^{-1}(\sigma(H_q))} m_-(p)(J(p) + J^*(\overline{p}^*))m_-^*(\overline{p}^*)d\Omega.
\end{aligned}$$

By condition (ii) it now follows that $m_- = 0$ and hence $m = C(\mu w)$ with $\mu = m_- = 0$ by Theorem A.5 (where we used the trivial factorization $b_- = \mathbb{I}$ and $b_+ = J$). $\square$

Remark B.2. *The same proof also shows uniqueness for the following symmetric vector Riemann–Hilbert problem on* $\mathbb{M}$

$$\begin{aligned}
m_+(p) &= m_-(p)J(p), \quad p \in \Sigma, \\
m(p^*) &= m(p)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
m(\infty_+) = \begin{pmatrix} 1 & * \end{pmatrix}, &\quad (m_1) \ge -\mathcal{D}_{\underline{\hat{\mu}}^*}, \quad (m_2) \ge -\mathcal{D}_{\underline{\hat{\mu}}}
\end{aligned} \tag{B.3}$$

where $J(z)$, Σ, *and* $\mathcal{D}_{\underline{\hat{\mu}}}$ *satisfy the same assumptions as in the previous theorem. Just note that in this case the symmetry assumption implies* $m(p)m^*(\overline{p}^*) = m_1(p)\overline{m_2(\overline{p})} + m_2(p)\overline{m_1(\overline{p})}$.

Acknowledgments. We thank I. Egorova, H. Krüger and A. Mikikits-Leitner for pointing out errors in a previous version of this article. G.T. would like to thank P. Deift for discussions on this topic.

S.K. gratefully acknowledges the support of the European Science Foundation (MISGAM program) and the Austrian Science Fund (FWF) during several visits to the University of Vienna in 2005–2007. G.T. gratefully acknowledges the extraordinary hospitality of the Courant Institute of Mathematical Sciences, where part of this research was done.

References

[1] M.J. Ablowitz, H. Segur, *Asymptotic solutions of the Korteweg–de Vries equation*, Studies in Appl. Math. **57-1**, 13–44, (1976/77).

[2] R. Beals, R. Coifman, *Scattering and inverse scattering for first order systems*, Comm. in Pure and Applied Math. **37**, 39–90 (1984).

[3] W. Bulla, F. Gesztesy, H. Holden, and G. Teschl, *Algebro-geometric quasi-periodic finite-gap solutions of the Toda and Kac-van Moerbeke hierarchies*, Memoirs Amer. Math. Soc. **135**, no. 641, 1–79 (1998).

[4] A. Boutet de Monvel and V. Kotlyarov, *The focusing nonlinear Schrödinger equation on the quarter plane with time-periodic boundary condition: a Riemann–Hilbert approach*, J. Inst. Math. Jussieu **6:4**, 579–611 (2007).

[5] V. S. Buslaev and V. V. Sukhanov, *Asymptotic behavior of solutions of the Korteweg-de Vries equation*, Jour. Sov. Math. **34**, 1905–1920 (1986).

[6] P. Deift, X. Zhou, *A steepest descent method for oscillatory Riemann–Hilbert problems*, Ann. of Math. (2) **137**, 295–368 (1993).

[7] P. Deift, X. Zhou, *Long-time asymptotics for integrable systems. higher order theory* , Commun. Math. Phys. **165**, 175–191 (1994).

[8] P. A. Deift, A. R. Its, A. R., and X. Zhou, *Long-time asymptotics for integrable nonlinear wave equations*, in "Important developments in soliton theory", 181–204, Springer Ser. Nonlinear Dynam., Springer, Berlin, 1993.

[9] P. Deift, S. Venakides, and X. Zhou, *The collisionless shock region for the long time behavior of solutions of the KdV equation*, Comm. in Pure and Applied Math. **47**, 199–206 (1994).

[10] I. Egorova, J. Michor, and G. Teschl, *Scattering theory for Jacobi operators with quasi-periodic background*, Comm. Math. Phys. **264-3**, 811–842 (2006).

[11] I. Egorova, J. Michor, and G. Teschl, *Inverse scattering transform for the Toda hierarchy with quasi-periodic background*, Proc. Amer. Math. Soc. **135**, 1817–1827 (2007).

[12] I. Egorova, J. Michor, and G. Teschl, *Scattering theory for Jacobi operators with steplike quasi-periodic background*, Inverse Problems **23**, 905–918 (2007).

[13] I. Egorova, J. Michor, and G. Teschl, *Soliton solutions of the Toda hierarchy on quasi-periodic background revisited*, Math. Nach. **282:4**, 526–539 (2009).

[14] A.S.Fokas, A.R.Its, L.Y.Sung, *The Nonlinear Schrödinger Equation on the Half-Line,* Nonlinearity, **18**, 1771-1822, (1995).

[15] L. D. Faddeev, L. A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons*, Springer, Berlin, 1987.

[16] H. Farkas and I. Kra, *Riemann Surfaces*, 2$^{\text{nd}}$ edition, GTM 71, Springer, New York, 1992.

[17] F. Gesztesy, H. Holden, J. Michor, and G. Teschl, *Soliton Equations and Their Algebro-Geometric Solutions. Volume II:* $(1+1)$*-Dimensional Discrete Models*, Cambridge Stud. in Adv. Math. **114**, Cambridge Univ. Press, Cambridge, 2008.

[18] I. Gohberg and N. Krupnik, *One-Dimensional Linear Singular Integral Equations*, Birkhäuser, Basel, 1992.

[19] A. R. Its, *Asymptotic behavior of the solutions to the nonlinear Schrödinger equation, and isomonodromic deformations of systems of linear differential equations*, Soviet. Math. Dokl. **24:3**, 452–456 (1981).

[20] A. R. Its, *"Isomonodromy" solutions of equations of zero curvature*, Math. USSR-Izv. **26:3**, 497–529 (1986).

[21] A. R. Its, *Asymptotic behavior of the solution of the Cauchy problem for the modified Korteweg-de Vries equation (Russian)*, in Wave propagation. Scattering theory, 214–224, 259, Probl. Mat. Fiz. **12**, Leningrad. Univ., Leningrad, 1987.

[22] A. R. Its and V. È. Petrov, *"Isomonodromic" solutions of the sine-Gordon equation and the time asymptotics of its rapidly decreasing solutions*, Soviet Math. Dokl. **26:1**, 244–247 (1982).

[23] S. Kamvissis, *On the long time behavior of the doubly infinite Toda lattice under initial data decaying at infinity,* Comm. Math. Phys., **153-3**, 479–519 (1993).
[24] S. Kamvissis, *Long time behavior for the focusing nonlinear Schrödinger equation with real spectral singularities,* Comm. Math. Phys., **180-2**, 325–343 (1996).
[25] S. Kamvissis and G. Teschl, *Stability of periodic soliton equations under short range perturbations*, Phys. Lett. A, **364-6**, 480–483 (2007).
[26] S. Kamvissis, K. McLaughlin, and P. Miller, *Semiclassical Soliton Ensembles for the Focusing Nonlinear Schrödinger Equation,* Annals of Mathematics, Study 154, Princeton Univ. Press, Princeton, 2003.
[27] L. K. Kodama, *Boundary measures of analytic differentials and uniform approximation on a Riemann surface*, Pacific J. Math. **15**, 1261–1277 (1965).
[28] H. Krüger and G. Teschl, *Long-time asymptotics for the Toda lattice in the soliton region*, Math. Z. **262**, 585–602 (2009).
[29] H. Krüger and G. Teschl, *Long-time asymptotics of the Toda lattice for decaying initial data revisited*, Rev. Math. Phys. **21:1**, 61–109 (2009).
[30] H. Krüger and G. Teschl, *Stability of the periodic Toda lattice in the soliton region*, Int. Math. Res. Not. **2009**, Art. ID rnp077, 36pp (2009).
[31] S. V. Manakov, *Nonlinear Frauenhofer diffraction*, Sov. Phys. JETP **38:4**, 693–696 (1974).
[32] J. Michor and G. Teschl, *Trace formulas for Jacobi operators in connection with scattering theory for quasi-periodic background*, in Operator Theory, Analysis and Mathematical Physics, J. Janas, et al. (eds.), 51–57, Oper. Theory Adv. Appl. **174**, Birkhäuser, Basel, 2007.
[33] N.I. Muskhelishvili, *Singular Integral Equations*, P. Noordhoff Ltd., Groningen, 1953.
[34] V.Yu. Novokshenov and I.T. Habibullin, *Nonlinear differential-difference schemes integrable by the method of the inverse scattering problem. Asymptotics of the solution for* $t \to \infty$, Sov. Math. Doklady **23/2**, 304–307 (1981).
[35] S. Prössdorf, *Some Classes of Singular Equations*, North-Holland, Amsterdam, 1978.
[36] Yu. Rodin, *The Riemann Boundary Problem on Riemann Surfaces*, Mathematics and its Applications (Soviet Series) **16**, D. Reidel Publishing Co., Dordrecht, 1988.
[37] A. B. Šabat, *On the Korteweg-de Vries equation*, Soviet Math. Dokl. **14**, 1266–1270 (1973).
[38] H. Segur and M. J. Ablowitz, *Asymptotic solutions of nonlinear evolution equations and a Painléve transcendent*, Phys. D **3**, 165–184 (1981).
[39] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.
[40] S. Tanaka, *Korteweg–de Vries equation; Asymptotic behavior of solutions*, Publ. Res. Inst. Math. Sci. **10**, 367–379 (1975).
[41] G. Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*, Math. Surv. and Mon. **72**, Amer. Math. Soc., Rhode Island, 2000.
[42] G. Teschl, *Almost everything you always wanted to know about the Toda equation*, Jahresber. Deutsch. Math.-Verein. **103**, no. 4, 149–162 (2001).
[43] G. Teschl, *Algebro-geometric constraints on solitons with respect to quasi-periodic backgrounds*, Bull. London Math. Soc. **39-4**, 677–684 (2007).
[44] M. Toda, *Theory of Nonlinear Lattices*, 2nd enl. ed., Springer, Berlin, 1989.
[45] V. E. Zakharov and S. V. Manakov, *Asymptotic behavior of nonlinear wave systems integrated by the inverse method*, Sov. Phys. JETP **44**, 106–112 (1976).
[46] N. J. Zabusky and M. D. Kruskal, *Interaction of solitons in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett. **15**, 240–243 (1965).
[47] X. Zhou, *The Riemann–Hilbert problem and inverse scattering*, SIAM J. Math. Anal. **20-4**, 966–986 (1989).

Department of Applied Mathematics, University of Crete, 714 09 Knossos, Greece
E-mail address: spyros@tem.uoc.gr
URL: http://www.tem.uoc.gr/~spyros/

Faculty of Mathematics, University of Vienna, Nordbergstrasse 15, 1090 Wien, Austria, and International Erwin Schrödinger Institute for Mathematical Physics, Boltzmanngasse 9, 1090 Wien, Austria
E-mail address: Gerald.Teschl@univie.ac.at
URL: http://www.mat.univie.ac.at/~gerald/