STABILITY OF THE PERIODIC TODA LATTICE UNDER SHORT RANGE PERTURBATIONS

SPYRIDON KAMVISSIS AND GERALD TESCHL

Abstract. We consider the stability of the periodic Toda lattice (and slightly more generally of the algebro-geometric finite-gap 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 the solitons travelling on the quasi-periodic background, the n/t-pane contains $g+2$ areas where the perturbed solution is close to a finite-gap solution in 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 matrix 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 [24] states that a short range perturbation of the constant solution of a soliton equation eventually splits into a number of stable solitons. The solitons constitute the stable part of arbitrary short range initial conditions. This is the motivation for the result presented here. Our aim is to

---

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

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

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

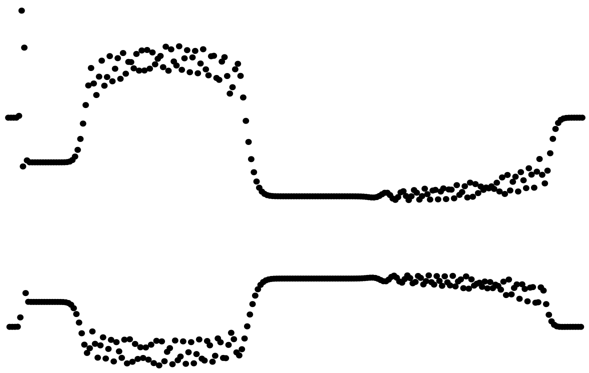

FIGURE 1. Numerically computed solution of the Toda lattice, with initial condition a period two solution perturbed at one point in the middle.

investigate the case where the constant background solution is replaced by a periodic one.

Of course periodic solutions, as well as solitons travelling on a periodic background, are well understood and one might expect the natural generalization in this situation. However, as pointed out in [13] this naive expectation is wrong: In Figure 1 the two observed lines express the variables $a(n,t)$ 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. We provide the details in the case of the Toda lattice though it is clear that our methods apply to other soliton equations as well.

Consider the doubly infinite Toda lattice in Flaschka's variables (see e.g. [20], [21], or [23])

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

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

$$\sum_n |n|(|a(n,t) - a_q(n,t)| + |b(n,t) - b_q(n,t)|) < \infty \tag{1.2}$$

for one (and hence for all, see [7]) $t \in \mathbb{R}$. The perturbed solution can be computed via the inverse scattering transform. The case where (a_q, b_q) is constant is classical (see again [20] or [23]), the more general case we want here was solved only recently in [7] (see also [15]). We will also assume that the reflection coefficient is continuous. This is for example the case if we replace $|n|$ by $|n|^2$ in (1.2).

Assume for simplicity that the Jacobi operator H_q 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

$$\begin{aligned} \prod_{j=n}^{\infty}(\frac{a_l(j,t)}{a_q(j,t)})^2 =& \frac{\theta(\underline{z}(n,t))}{\theta(\underline{z}(n-1,t))} \frac{\theta(\underline{z}(n-1,t) + \underline{\delta}(n,t))}{\theta(\underline{z}(n,t) + \underline{\delta}(n,t))} \times \\ & \times \exp\left(\frac{1}{2\pi \mathrm{i}} \int_{C(n/t)} \log(1 - |R|^2)\omega_{\infty_+\infty_-}\right), \\ \delta_\ell(n,t) =& \frac{1}{2\pi\mathrm{i}} \int_{C(n/t)} \log(1-|R|^2)\zeta_\ell, \end{aligned} \tag{1.3}$$

where R is the associated reflection coefficient, $\omega_{\infty_+\infty_-}$ is an Abelian differential of the third kind defined in (2.14), $C(n/t) = \pi^{-1}(\sigma(H_q) \cap (-\infty, z_j(n/t)))$, and $z_j(n/t)$ is a special stationary phase point for the phase defined in the beginning of Section 4. 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*

$$\prod_{j=n}^{\infty} \frac{a_l(j,t)}{a(j,t)} \to 1 \tag{1.4}$$

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

The proof of this theorem will be given in Section 5 of this paper.

Remark 1.2. *(i) If eigenvalues are present we can apply appropriate Darboux transformations to add the effect of such eigenvalues (*[9]*). What we then see asymptotically is travelling solitons in a periodic background. Note that this will change the asymptotics on one side. However, our method works for such situations (cf.* [8]*) unaltered.*
(ii) Employing the very same methods of the paper it is very easy to show that in any region $|\frac{n}{t}| > C$*, one has*

$$\prod_{j=n}^{\infty} \frac{a_l(j,t)}{a(j,t)} \to 1 \tag{1.5}$$

uniformly in t*, as* $n \to \infty$*.*
(iii) 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* $t \to \infty$*. We expect a 'collisionless shock' phenomenon to appear (*[5], [12]*). We will study this in a future paper.*

By dividing in (1.3) one recovers the $a(n,t)$. It follows from the main Theorem and the last remark above that

$$|a(n,t) - a_l(n,t)| \to 0 \tag{1.6}$$

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*

$$\sum_{j=n}^{\infty} \big(b_l(j,t) - b_q(j,t)\big) \to 0 \tag{1.7}$$

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

$$\begin{aligned} \sum_{j=n}^{\infty} \big(b_l(j,t) - b_q(j,t)\big) =& \frac{1}{2\pi \mathrm{i}} \int_{C(n/t)} \log(1 - |R|^2)\Omega_0 \\ &+ \frac{1}{2} \frac{d}{ds} \log\left(\frac{\theta(\underline{z}(n,s) + \underline{\delta}(n,t))}{\theta(\underline{z}(n,s))} \right) \Big|_{s=t}, \end{aligned} \tag{1.8}$$

where Ω_0 *is an Abelian differential of the second kind defined in (2.15).*

The proof of this theorem will also be given in Section 5 of this paper.

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, [2]. In particular, this class contains all periodic solutions. We will use the same notation as in [20], where we also refer to for proofs. As a reference for Riemann surfaces in this context we recommend [10].

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

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

$g \in \mathbb{N}$. $\mathbb{M}$ 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} counter-clockwise 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

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

The differentials fulfill

$$\int_{a_j} \zeta_k = \delta_{j,k}, \qquad \int_{b_j} \zeta_k = \tau_{j,k}, \qquad \tau_{j,k} = \tau_{k,j}, \qquad 1 \le j,k \le g. \tag{2.6}$$

Now pick g numbers (the Dirichlet eigenvalues)

$$(\hat{\mu}_j)_{j=1}^g = (\mu_j, \sigma_j)_{j=1}^g \tag{2.7}$$

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

$$\begin{aligned} \underline{z}(p,n,t) &= \underline{\hat{A}}_{p_0}(p) - \underline{\hat{\alpha}}_{p_0}(\mathcal{D}_{\underline{\hat{\mu}}}) - n\underline{\hat{A}}_{\infty_-}(\infty_+) + t\underline{U}_0 - \underline{\hat{\Xi}}_{p_0} \in \mathbb{C}^g, \\ \underline{z}(n,t) &= \underline{z}(\infty_+, n, t), \end{aligned} \tag{2.8}$$

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

$$\hat{\Xi}_{p_0,j} = \frac{1 - \sum_{k=1}^g \tau_{j,k}}{2}, \qquad p_0 = (E_0, 0), \tag{2.9}$$

$\underline{U}_0$ 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 $\theta(\underline{z}(p,n))$ has precisely g zeros $\hat{\mu}_j(n)$ (with $\hat{\mu}_j(0) = \hat{\mu}_j$), where $\theta(\underline{z})$ is the Riemann theta function of $\mathbb{M}$.

Then our background solution is given by

$$\begin{aligned} a_q(n,t)^2 &= \tilde{a}^2 \frac{\theta(\underline{z}(n+1,t))\theta(\underline{z}(n-1,t))}{\theta(\underline{z}(n,t))^2}, \\ b_q(n,t) &= \tilde{b} + \frac{1}{2}\frac{d}{dt} \log\Big(\frac{\theta(\underline{z}(n,t))}{\theta(\underline{z}(n-1,t))}\Big). \end{aligned} \tag{2.10}$$

The constants $\tilde{a}$, $\tilde{b}$ depend only on the Riemann surface (see [20, Section 9.2]).

Introduce the time dependent Baker-Akhiezer function

$$\psi_q(p,n,t) = C(n,0,t) \frac{\theta(\underline{z}(p,n,t))}{\theta(\underline{z}(p,0,0))} \exp\Big(n \int_{E_0}^p \omega_{\infty_+,\infty_-} + t \int_{E_0}^p \Omega_0\Big), \tag{2.11}$$

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

$$C(n,0,t)^2 = \frac{\theta(\underline{z}(0,0))\theta(\underline{z}(-1,0))}{\theta(\underline{z}(n,t))\theta(\underline{z}(n-1,t))}, \tag{2.12}$$

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

$$\theta(\underline{z}) = \sum_{\underline{m}\in\mathbb{Z}^g} \exp 2\pi\mathrm{i}\left(\langle\underline{m},\underline{z}\rangle + \frac{\langle\underline{m},\underline{\tau}\,\underline{m}\rangle}{2}\right), \qquad \underline{z}\in\mathbb{C}^g, \tag{2.13}$$

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

$$\omega_{\infty_+\infty_-} = \frac{\prod_{j=1}^g(\pi-\lambda_j)}{R_{2g+2}^{1/2}}d\pi \tag{2.14}$$

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

$$\Omega_0 = \frac{\prod_{j=0}^g(\pi-\tilde{\lambda}_j)}{R_{2g+2}^{1/2}}d\pi, \qquad \sum_{j=0}^g \tilde{\lambda}_j = \frac{1}{2}\sum_{j=0}^{2g+1} E_j, \tag{2.15}$$

is the Abelian differential of the second kind with second order poles at ∞_+ respectively ∞_- (see [20, 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 $\mathbb{M}\backslash\{\infty_\pm\}$ with an essential singularity at $\infty_\pm$. The two branches are denoted by

$$\psi_{q,\pm}(z,n,t) = \psi_q(p,n,t), \qquad p=(z,\pm) \tag{2.16}$$

and it satisfies

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

where H_q, $P_{q,2}$ 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

$$\sigma(H_q) = \bigcup_{j=0}^g [E_{2j}, E_{2j+1}]. \tag{2.17}$$

For further information and proofs we refer to [20, 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 [6] and [7]. 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 defined by

$$\lim_{n\to\pm\infty} w(z)^{\mp n}(\psi_\pm(z,n,t) - \psi_{q,\pm}(z,n,t)) = 0, \tag{3.1}$$

where $w(z)$ is the quasimomentum map

$$w(z) = \exp(\int_{E_0}^{p} \omega_{\infty_+,\infty_-}), \quad p=(z,+). \tag{3.2}$$

The asymptotics of the two projections of the Jost function are

$$\begin{aligned}\psi_\pm(z,n,t) = & \frac{z^{\mp n}\Big(\prod_{j=0}^{n-1} a_q(j,t)\Big)^{\pm 1}}{A_\pm(n,t)} \times \\ & \times \Big(1 + \Big(B_\pm(n,t) \pm \sum_{j=1}^{n} b_q(j - {\scriptstyle 0 \atop 1}, t)\Big)\frac{1}{z} + O(\frac{1}{z^2})\Big),\end{aligned} \tag{3.3}$$

as $z\to\infty$, where

$$\begin{aligned} A_+(n,t) &= \prod_{j=n}^{\infty} \frac{a(j,t)}{a_q(j,t)}, \quad B_+(n,t) = \sum_{j=n+1}^{\infty} (b_q(j,t) - b(j,t)),\\ A_-(n,t) &= \prod_{j=-\infty}^{n-1} \frac{a(j,t)}{a_q(j,t)}, \quad B_-(n,t) = \sum_{j=-\infty}^{n-1} (b_q(j,t) - b(j,t)).\end{aligned} \tag{3.4}$$

One has the scattering relations

$$T(z)\psi_\mp(z,n,t) = \overline{\psi_\pm(z,n,t)} + R_\pm(z)\psi_\pm(z,n,t), \qquad z\in\sigma(H_q). \tag{3.5}$$

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

$$T(z)\overline{R_+(z)} + \overline{T(z)}R_-(z) = 0, \qquad |T(z)|^2 + |R_\pm(z)|^2 = 1. \tag{3.6}$$

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:

$$m(p,n,t) = \left\{ \begin{array}{ll} \begin{pmatrix} T(z)\psi_-(z,n,t) & \psi_+(z,n,t)\end{pmatrix}, & p=(z,+)\\ \begin{pmatrix} \psi_+(z,n,t) & T(z)\psi_-(z,n,t)\end{pmatrix}, & p=(z,-)\end{array}\right. . \tag{3.7}$$

Note that $m(p,n,t)$ inherits the poles at $\hat{\mu}_j(n,t)$ 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,

$$m_\pm(p_0)=\lim_{\Pi_\pm\ni p\to p_0} m(p), \qquad p_0\in\Sigma. \tag{3.8}$$

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

$$p^* = \begin{cases} (z,\mp) & \text{for } p=(z,\pm),\\ \infty_\mp & \text{for } p=\infty_\pm, \end{cases} \tag{3.9}$$

and complex conjugation

$$\overline{p} = \begin{cases} (\overline{z},\pm) & \text{for } p=(z,\pm)\not\in\Sigma,\\ (z,\mp) & \text{for } p=(z,\pm)\in\Sigma,\\ \infty_\pm & \text{for } p=\infty_\pm. \end{cases} \tag{3.10}$$

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 (3.5) and (3.6), we obtain

$$m_+(p,n,t)=m_-(p,n,t)\begin{pmatrix} |T(p)|^2 & -\overline{R(p)}\\ R(p) & 1\end{pmatrix}, \tag{3.11}$$

where we have extend our definition of T to Σ such that it is equal to $T(z)$ on Σ_+ and equal to $\overline{T(z)}$ on Σ_-. Similarly for $R(z)$. 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

$$m^2(p,n,t)=m(p,n,t)\begin{pmatrix} \psi_q(p^*,n,t)^{-1} & 0\\ 0 & \psi_q(p,n,t)^{-1}\end{pmatrix}. \tag{3.12}$$

Its divisor satisfies

$$(m_1^2) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)^*}, \qquad (m_2^2) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)}, \tag{3.13}$$

and the jump conditions become

$$\begin{aligned} m_+^2(p,n,t) &= m_-^2(p,n,t) J^2(p,n,t) \\ J^2(p,n,t) &= \begin{pmatrix} 1-|R(p)|^2 & -\overline{R(p)\Theta(p,n,t)}\mathrm{e}^{-t\phi(p)} \\ R(p)\Theta(p,n,t)\mathrm{e}^{t\phi(p)} & 1 \end{pmatrix}, \end{aligned} \tag{3.14}$$

where

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

and

$$\phi(p,\frac{n}{t}) = 2\int_{E_0}^p \Omega_0 + 2\frac{n}{t}\int_{E_0}^p \omega_{\infty_+,\infty_-} \in \mathrm{i}\mathbb{R} \tag{3.15}$$

for $p \in \Sigma$. Furthermore,

$$m^2(\infty_+,n,t) = \begin{pmatrix} A_+(n,t) & \frac{1}{A_+(n,t)} \end{pmatrix}. \tag{3.16}$$

and

$$m^2(\infty_-,n,t) = \begin{pmatrix} \frac{1}{A_+(n,t)} & A_+(n,t) \end{pmatrix}. \tag{3.17}$$

Here we have used

$$T(\infty) = A_-(n,t)A_+(n,t)\big(1 - \frac{B_+(n,t)+B_-(n,t)}{z} + O(\frac{1}{z^2}) \tag{3.18}$$

where $A_\pm(n,t)$ and $B_\pm(n,t)$ are defined in (3.4).

Next we show how to normalize the problem at infinity.

Theorem 3.1. *Let ℓ be a fixed positive integer $\leq 2g+1$. Consider the following Riemann–Hilbert problem. Find a 2×2 matrix $m^3(p)$ that is meromorphic off Σ,*

$$(m_{j1}^3) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)^*}, \quad (m_{j2}^3) \geq -\mathcal{D}_{\underline{\hat{\mu}}(n,t)}, \qquad j=1,2; \tag{3.19}$$

with the following behavior at the band edges E_k, $1 \leq k \leq \ell$:

$$(m_{j1}^3)(p) = O(T(z)), \quad (m_{j2}^3)(p) = O(T(z)^{-1}), \qquad j=1,2, \tag{3.20}$$

for $p=(z,+)$ near E_k on the top sheet and

$$(m_{j1}^3)(p) = O(T(z)^{-1}), \quad (m_{j2}^3)(p) = O(T(z)), \qquad j=1,2, \tag{3.21}$$

for $p=(z,-)$ near E_k on the bottom sheet; that is bounded near $\Sigma \setminus (\cup_{j>l} E_j \cup \cup_j [\hat{\mu}_j(n,t) \cup \hat{\mu}_j(n,t)^*])$ and that has same jumps as $m^2(p)$, *that is*

$$m^3_+(p) = m^3_-(p) J^3(p),$$
$$J^3(p) = \begin{pmatrix} 1 - |R(p)|^2 & -\overline{R(p)\Theta(p,n,t)}\mathrm{e}^{-t\phi(p)} \\ R(p)\Theta(p,n,t)\mathrm{e}^{t\phi(p)} & 1 \end{pmatrix}, \tag{3.22}$$

$z \in \Sigma \setminus \cup_j\{E_j\}$, *but such that*

$$m^3(\infty_+) = \mathbb{I}. \tag{3.23}$$

A unique such $m^3(p)$ exists and furthermore the Jacobi operator can be recovered from $m^3(\infty_-)$ via the formulae

$$A_+(n,t) = \sqrt{\frac{1 + (m^3_{12}(\infty_-, n, t))}{(m^3_{11}(\infty_-, n, t))}}. \tag{3.24}$$

The integer ℓ will be chosen later according to the location of the stationary phase points.

Remark 3.2. *It is not a priori obvious that a solution of the above Riemann–Hilbert problem exists. However, existence and uniqueness follow from a general theorem in Section A. On the other hand existence will also be justified a posteriori, since existence of the solution Q of problem (4.27) guaranteed by Theorem 5.4 will imply existence of the solution of the problem centered on the small crosses of Section 4, which will then imply existence of the solution m^5 of the Riemann–Hilbert problem (4.24), which is equivalent to existence of the problem above, using definitions (4.23) and (4.10).*

Uniqueness is actually not necessary for our purposes. All we need is uniqueness at ∞_- and that follows easily, without the use of the vanishing theorem of Section A. Indeed, note that the jump matrix J^3 has determinant one and hence $\det(m^3)$ *has no jump. Thus it extends to a meromorphic function by the Schwarz reflection principle. The singularities at the band edges are removable by (3.20)–(3.21) (unless one of the $\hat{\mu}_j$ is sitting there). So its divisor satisfies* $(\det(m^3)) \geq -\mathcal{D}_{\underline{\hat{\mu}}} - \mathcal{D}_{\underline{\hat{\mu}}^*}$ *and the Riemann–Roch theorem implies*

$$\det(m^3(p)) = \prod_{j=1}^{g} \frac{z - \nu_j}{z - \mu_j}, \qquad z = \pi(p), \tag{3.25}$$

for some complex numbers ν_j. Hence, if $\tilde{m}^3$ is a second solution, then $\tilde{m}^3(m^3)^{-1}$ has no jump and by a similar argument

$$\tilde{m}^3(p) = \frac{h(z)}{\prod_{j=1}^{g}(z - \nu_j)} m^3(p), \tag{3.26}$$

where $h(z)$ is a polynomial matrix with h_{11}, h_{22} monic polynomials of degree g and h_{12}, h_{21} polynomials of degree $g-1$. Hence m^3 is uniquely defined at ∞_-.

We also note that the behavior at the band edges E_j is imposed by the zeros and poles introduced (later in the text) by the sequence of transformations from m^6 to m^3.

Proof. As noted above, existence and uniqueness follows from a general theorem in Section A. Next, note the symmetry

$$J^3(p) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} J^3(p^*)^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.27}$$

of the jump matrix. It follows that

$$m^3(p) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} (m^3(\infty_-))^{-1} m^3(p^*) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.28}$$

Indeed, it is simple to check that both sides satisfy the same Riemann–Hilbert problem (3.22)–(3.23). Hence, by uniqueness, they are equal.

Now suppose

$$m^3(\infty_-) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

Noting that $\det(m^3(\infty_-)) = 1$ and setting $p = \infty_-$ in (3.28) we get $\beta = -\gamma$. So

$$m^3(\infty_-) = \begin{pmatrix} \alpha & \beta \\ -\beta & \frac{1-\beta^2}{\alpha} \end{pmatrix}. \tag{3.29}$$

Next note that both the matrix $m^3(p)$ and the row vector $m^2(p)$ satisfy the same jump conditions (3.22). It follows that the row vector $m^2(p)(m^3(p))^{-1}$ has no jumps and is bounded near the band edges. Arguing as in the remark above we see

$$m^2(p) = \frac{k(z)}{\prod_{j=1}^g (z-\nu_j)} m^3(p), \tag{3.30}$$

where, by (3.16), $k(z)$ is of the form

$$k(z) = \begin{pmatrix} A_+ \prod_{j=1}^g (z-\nu_{1j}) & \frac{1}{A_+} \prod_{j=1}^g (z-\nu_{2j}) \end{pmatrix}. \tag{3.31}$$

Setting $p = \infty_-$ in (3.30) we get, in view of (3.17) and (3.29),

$$\begin{pmatrix} \frac{1}{A_+} & A_+ \end{pmatrix} = \begin{pmatrix} A_+ & \frac{1}{A_+} \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ -\beta & \frac{1-\beta^2}{\alpha} \end{pmatrix}.$$

We thus arrive at

$$A_+ = \sqrt{\frac{1+\beta}{\alpha}}$$

and the theorem follows by recalling (3.29). □

Note that setting $R(z) \equiv 0$ we recover the purely periodic solution, as we should. Moreover, by $J(p) = \overline{J(\overline{p})}$, $p \in \Sigma$, we also have $m^3(p) = \overline{m^3(\overline{p})}$ (since both sides satisfy the same Riemann–Hilbert problem).

Remark 3.3. *It is important to note here that the Riemann–Hilbert problem arising from the analysis of the perturbed periodic problem (for any discrete in space integrable system) has to be set on a Riemann surface. There is no way one can use a symmetry like (3.27) to normalize the Riemann–Hilbert problem at infinity (i.e. get rid of the inappropriate condition (3.16) at ∞_+) if one stays in the complex plane.*

Remark 3.4. *In the proof of Theorem 1.3 we will show that in (3.30) we actually have*

$$m^2(p) = \begin{pmatrix} A_+ & \frac{1}{A_+} \end{pmatrix} m^3(p). \tag{3.32}$$

4. The stationary phase points

The phase in the factorization problem (3.14) is $t\,\phi$ where ϕ was defined in (3.15). Invoking (2.14) and (2.15), 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, $0 \le j \le g$, are real and different with precisely one lying in each spectral gap, say λ_j in the j'th gap. Similarly, $\tilde{\lambda}_j$, $0 \le j \le g$, are real and different and $\tilde{\lambda}_j$, $1 \le j \le g$, sits in the j'th gap. However $\tilde{\lambda}_0$ can be anywhere (see [20, 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 $z_j(\eta)$, $0 \le j \le g$, 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,

$$z_j' = -\frac{q(z_j)}{\tilde{q}'(z_j) + \eta q'(z_j)},$$

where

$$\tilde{q}(z) = \prod_{j=0}^{g} (z - \tilde{\lambda}_j), q(z) = \prod_{j=1}^{g} (z - \lambda_j),$$

shows that $z_j(\eta)$ is monotone decreasing (note that the denominator cannot vanish since the z_j's are always different) and must stay between λ_j and λ_{j+1}. □

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). Let us fix $\ell = 2j$ in Theorem 3.1 and 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

$$\mathrm{Im}(\phi') < 0, \quad \text{in } [E_{2j}, z_j], \qquad \mathrm{Im}(\phi') > 0, \quad \text{in } [z_j, E_{2j+1}] \tag{4.4}$$

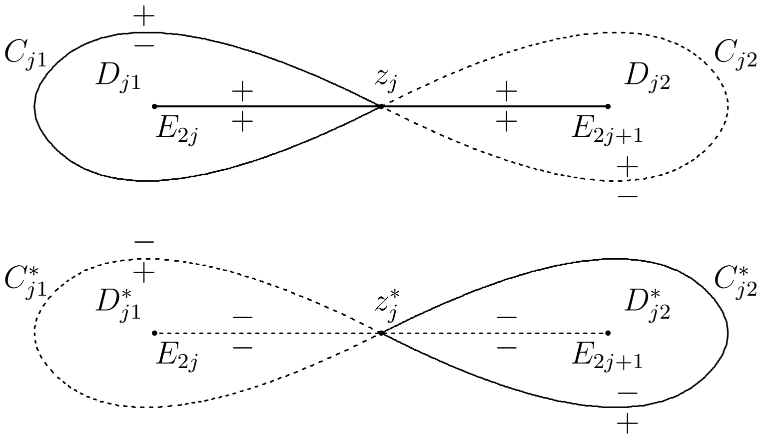

FIGURE 2. The lens contour near a band containing a stationary phase point z_j and its flipping image containing z_j^*. Views from the top and bottom sheet. Dotted curves lie in the bottom sheet.

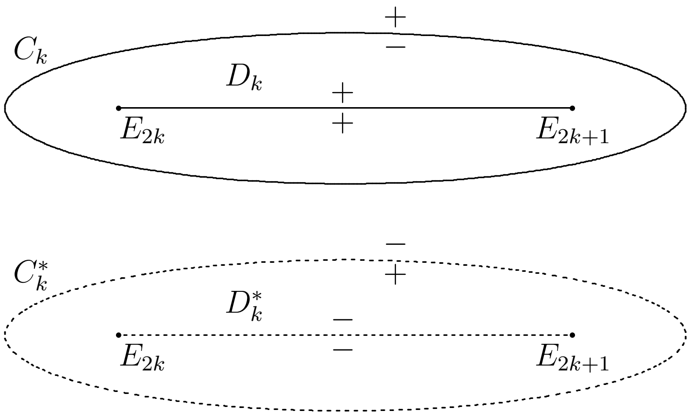

FIGURE 3. The lens contour near a band not including any stationary phase point. Views from the top and bottom sheet.

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

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

Now observe that our jump condition (3.22) has the following important factorization

$$J^3 = (b_-)^{-1} b_+, \tag{4.5}$$

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

$$J^3 = (B_-)^{-1} \begin{pmatrix} 1 - |R|^2 & 0 \\ 0 & \frac{1}{1-|R|^2} \end{pmatrix} B_+, \tag{4.6}$$

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 (1.3)).

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

$$\underline{\alpha}_{p_0}(\mathcal{D}_{\underline{\hat{\nu}}(n,t)}) = \underline{\alpha}_{p_0}(\mathcal{D}_{\underline{\hat{\mu}}(n,t)}) + \underline{\delta}(n,t), \tag{4.7}$$

where

$$\delta_\ell(n,t) = \frac{1}{2\pi\mathrm{i}} \int_{C(n/t)} \log(1 - |R|^2) \zeta_\ell. \tag{4.8}$$

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 (2.14) one checks that δ_ℓ is real. Hence it follows from [20, 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 [20, Lem. A.20]. □

Now we can formulate the scalar Riemann–Hilbert problem required to eliminate the diagonal part in the factorization (4.6):

$$\begin{aligned} &d_+(p,n,t) = d_-(p,n,t)(1-|R(p)|^2), \quad p \in C(n/t),\\ &(d) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)}, \\ &d(\infty_+, n, t) = 1, \end{aligned} \tag{4.9}$$

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. [18, Thm. 5.2]).

Theorem 4.3. *The unique solution of (4.9) is given by*

$$\begin{aligned} d(p,n,t) =& \frac{\theta(\underline{z}(n,t)+\underline{\delta}(n,t))}{\theta(\underline{z}(n,t))} \frac{\theta(\underline{z}(p,n,t))}{\theta(\underline{z}(p,n,t)+\underline{\delta}(n,t))} \times \\ & \times \exp\left(\frac{1}{2\pi \mathrm{i}} \int_{C(n/t)} \log(1-|R|^2)\omega_{p\,\infty_+}\right), \end{aligned} \tag{4.10}$$

where $\underline{\delta}(n,t)$ is defined in (4.8) and $\omega_{p\,q}$ is the Abelian differential of the third kind with poles at p and q.

The function $d(p)$ is meromorphic in $\mathbb{M}\setminus\Sigma$ with first order poles at $\hat{\nu}_j(n,t)$ and first oder 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 [18, Sec. 1.8]. It is the core of our explicit formula (1.3) 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. [18, Sec 1.4] respectively Section 5). 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

$$\theta(\underline{z}+\underline{m}+\underline{\tau}\,\underline{n}) = \exp[2\pi\mathrm{i}\left(-\langle\underline{n},\underline{z}\rangle - \frac{\langle\underline{n},\underline{\tau}\,\underline{n}\rangle}{2}\right)]\theta(\underline{z}), \quad \underline{n},\underline{m}\in\mathbb{Z}^g, \tag{4.11}$$

where $\underline{\tau}$ is the matrix of b-periods defined in (2.6) and $\langle ., .. \rangle$ denotes the scalar product in $\mathbb{R}^g$ (cf., e.g. [10] or [20, 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 (4.10) 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 (4.11) 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 (4.11) 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 (4.8), the jumps of the ration of theta functions and the exponential term compensate each other which shows that (4.10) is single-valued.

To prove 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 $(\tilde{d}/d) \geq -\mathcal{D}_{\underline{\hat{\mu}}}$. But $\mathcal{D}_{\underline{\hat{\mu}}}$ 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 (4.9). $\square$

Remark 4.4. *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* [22]*). Here $A = A_+(n,t)A_-(n,t) = T(\infty)$.*

In particular,

$$\begin{aligned} d(\infty_-, n, t) =& \frac{\theta(\underline{z}(n-1,t))}{\theta(\underline{z}(n,t))} \frac{\theta(\underline{z}(n,t) + \underline{\delta}(n,t))}{\theta(\underline{z}(n-1,t) + \underline{\delta}(n,t))} \times \\ & \times \exp\left(\frac{1}{2\pi\mathrm{i}} \int_{C(n/t)} \log(1 - |R|^2) \omega_{\infty_-\infty_+} \right), \end{aligned} \tag{4.12}$$

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 (2.14) 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 4.5. *For p near a stationary phase point z_j or z_j^* (not equal to a band edge) we have*

$$d(p) = \left(\frac{z - z_j}{z - E_{2j}} \right)^{\pm \mathrm{i}\nu} e_{\pm}(z), \quad p = (z, \pm), \tag{4.13}$$

where $e_{\pm}(z)$ is holomorphic near z_j and

$$\nu = -\frac{1}{2\pi} \log(1 - |R(z_j)|^2) > 0. \tag{4.14}$$

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

$$d(p) = T^{\pm 1}(z) \tilde{e}_{\pm}(z), \quad p = (z, \pm), \tag{4.15}$$

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 is a local property of the Cauchy integral and hence follows as in the case of Cauchy integrals in the plane (cf. [16]) by factorizing the jump according to $1 - |R(p)|^2 = (1 - |R(z_j)|^2) \frac{1 - |R(p)|^2}{1 - |R(z_j)|^2}$.

For the second claim note that (compare the Remark 4.4)

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

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. □

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

$$m^4(p) = \begin{pmatrix} \frac{1}{d(\infty_-)} & 0 \\ 0 & 1 \end{pmatrix} m^3(p) D(p), \quad D(p) = \begin{pmatrix} d(p^*) & 0 \\ 0 & d(p) \end{pmatrix}, \tag{4.16}$$

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

$$\begin{aligned} & m^4_+(p) = m^4_-(p) J^4(p), \quad p \in \Sigma, \\ & (m^4_{j1}) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)^*}, \quad (m^3_{j2}) \geq -\mathcal{D}_{\underline{\hat{\nu}}(n,t)}, \quad j=1,2, \\ & m^4(\infty_+) = \mathbb{I}, \end{aligned} \tag{4.17}$$

where the jump is given by

$$J^4(p) = D_-(p)^{-1} J^3(p) D_+(p), \quad p \in \Sigma. \tag{4.18}$$

In particular, m^4 has its poles shifted from $\hat{\mu}_j(n,t)$ to $\hat{\nu}_j(n,t)$ and the behavior (3.20), (3.21) at the band edges is canceled by $d(p)$ (cf. Lemma 4.5).

Furthermore, J^4 can be factorized as

$$J^4 = \begin{pmatrix} 1-|R|^2 & -\frac{d}{d^*}\overline{R\Theta}\mathrm{e}^{-t\phi} \\ \frac{d^*}{d}\Theta R \mathrm{e}^{t\phi} & 1 \end{pmatrix} = (\tilde{b}_-)^{-1} \tilde{b}_+, \quad p \in \Sigma \setminus C(n/t), \tag{4.19}$$

where

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

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

$$J^4 = \begin{pmatrix} 1 & -\frac{d_+}{d^*_-}\overline{R\Theta}\mathrm{e}^{-t\phi} \\ \frac{d^*_-}{d_+}\Theta R \mathrm{e}^{t\phi} & 1-|R|^2 \end{pmatrix} = (\tilde{B}_-)^{-1} \tilde{B}_+, \quad p \in C(n/t), \tag{4.20}$$

where

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

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

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

$$\frac{d_-^*(p)}{d_+(p)} = \frac{d_-^*(p)}{d_-(p)} \frac{1}{1-|R(p)|^2} = \frac{\overline{d_+(p)}}{d_+(p)}, \qquad p \in C(n/t), \tag{4.21}$$

respectively

$$\frac{d_+(p)}{d_-^*(p)} = \frac{d_+(p)}{d_+^*(p)} \frac{1}{1-|R(p)|^2} = \frac{\overline{d_-^*(p)}}{d_-^*(p)}, \qquad p \in C(n/t). \tag{4.22}$$

We finally define m^5 by

$$\begin{aligned}
m^5 &= m^4 \tilde{B}_+^{-1}, \quad p \in D_k,\ k < j,\\
m^5 &= m^4 \tilde{B}_-^{-1}, \quad p \in D_k^*,\ k < j,\\
m^5 &= m^4 \tilde{B}_+^{-1}, \quad p \in D_{j1},\\
m^5 &= m^4 \tilde{B}_-^{-1}, \quad p \in D_{j1}^*,\\
m^5 &= m^4 \tilde{b}_+^{-1}, \quad p \in D_{j2},\\
m^5 &= m^4 \tilde{b}_-^{-1}, \quad p \in D_{j2}^*,\\
m^5 &= m^4 \tilde{b}_+^{-1}, \quad p \in D_k,\ k > j,\\
m^5 &= m^4 \tilde{b}_-^{-1}, \quad p \in D_k^*,\ k > j,\\
m^5 &= m^4, \quad \text{otherwise},
\end{aligned} \tag{4.23}$$

where we assume that the deformed contour is sufficiently close to the original one (in particular so close such that we don't cross any poles).

The new jump matrix is given by

$$
\begin{aligned}
m^5_+(p,n,t) &= m^5_-(p,n,t) J^5(p,n,t),\\
J^5 &= \tilde{B}_+, \quad p \in C_k,\ k<j,\\
J^5 &= \tilde{B}_-^{-1}, \quad p \in C_k^*,\ k<j,\\
J^5 &= \tilde{B}_+, \quad p \in C_{j1},\\
J^5 &= \tilde{B}_-^{-1}, \quad p \in C_{j1}^*,\\
J^5 &= \tilde{b}_+, \quad p \in C_{j2},\\
J^5 &= \tilde{b}_-^{-1}, \quad p \in C_{j2}^*,\\
J^5 &= \tilde{b}_+, \quad p \in C_k,\ k>j,\\
J^5 &= \tilde{b}_-^{-1}, \quad p \in C_k^*,\ k>j.
\end{aligned} \tag{4.24}
$$

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. We refer to the discussion of [4] for the details.

Eventually we will see that, asymptotically, $m^5(p) \to \mathbb{I}$ uniformly, as $p \to \infty_\pm$, hence

$$
\begin{aligned}
m^3(p) &= \begin{pmatrix} d(\infty_-) & 0 \\ 0 & 1 \end{pmatrix} m^4(p) \begin{pmatrix} d(p^*)^{-1} & 0 \\ 0 & d(p)^{-1} \end{pmatrix}\\
&= \begin{pmatrix} d(\infty_-) & 0 \\ 0 & 1 \end{pmatrix} m^5(p) \begin{pmatrix} d(p^*)^{-1} & 0 \\ 0 & d(p)^{-1} \end{pmatrix}\\
&\sim \begin{pmatrix} d(\infty_-) d(p^*)^{-1} & 0 \\ 0 & d(p)^{-1} \end{pmatrix}
\end{aligned} \tag{4.25}
$$

for p near $\infty_\pm$. Finally, by Theorem 3.1 we will recover

$$
A_+(n,t)^2 \sim d(\infty_-, n, t)^{-1}. \tag{4.26}
$$

Remark 4.6. *$|R|$ is generically 1 at the endpoints E_j and hence $\frac{R}{1-|R|^2}$ is generically singular there, so if m^5 is to be regular (which is still to be proved) then m^3 must be considered (generically) as singular at the endpoints. This observation is the origin of the imposed behavior of m^3 at E_j in the definition of m^3 in Theorem 3.1. If we are in the non-generic case where $|R| < 1$ at some of the band edge E_j the imposed conditions on m^3 are of course void.*

The crucial observation now is that the jumps J^5 on the oriented paths $C_{j1}, C_{j2}, C_{j1}^*, C_{j2}^*$ 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 $t \to \infty$). Then, we can delete the contours $C_{j1}, C_{j2}, C_{j1}^*, C_{j2}^*$ from the picture. In any case, the jumps across C_k, C_k^*, $k \neq j$, are uniformly exponentially small. So we hope we can also delete the contours C_k, C_k^*, $k \neq j$.

Consider then the following "model" Riemann–Hilbert problem which is meant to "approximate" (4.24). Let m^6 be the function meromorphic of index 0 in $X \setminus L$ where L is the set of two crosses centered at z_j, z_j^* each branch of which has length $= t^{-1/2+\varepsilon}$, for some small $\varepsilon > 0$ which we do not need to specify. (We also do not specify the exact angles between the branches of the crosses as long as the situation is as shown in the two figures.) Assume the jump along the cross is given by

$$m^6_+(p,n,t) = m^6_-(p,n,t) J^6(p,n,t), \quad p \neq z_j, z_j^*, \tag{4.27}$$

where $J^6(p,n,t) = J^5(p,n,t)$, $p \in L$.

Note that the Riemann–Hilbert problems for m^3, m^4, m^5 are exactly, not approximately, equivalent. The step from m^5 to m^6 is the first where an approximation is involved.

Once we show that a solution of (4.27) exists, it is easy to see by simple rescaling that the solution is decaying as $t \to \infty$. Thus the remaining cross can also be neglected, up to leading order. To be more precise, once we have chosen the crosses so that their length is of order $t^{-1/2+\varepsilon}$, $\varepsilon > 0$, then the complement of the crosses will be deleted since the exponential terms are uniformly small there, while the small crosses will eventually be ignored simply because their length is small. All this will be done rigorously in Sections 5 and A.

Existence of solutions m^3, m^4, m^5 and the approximate problem (4.27) will follow directly as a corollary of Theorem A.1.

We expect (4.27) to be explicitly solvable since it is easily reducible to a problem with constant jump matrix, even though the underlying space is a Riemann surface. In any case, finding an exact solution is not necessary so long as we are only interested in leading asymptotics for the perturbed lattice.

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 might expect different local asymptotics expressed in terms of (generalized) Painlevé functions. In the case $|R(E_k)| < 1$ the two crosses coalesce and the discussion of Section A goes through virtually unaltered. If $|R(E_k)| = 1$ the problem is singular in an essential way and we expect a "collisionless shock" phenomenon in the region where $z_j(n/t) \sim E_k$, similar to the one studied in [1], [5], [12]. The main difficulty arises from the singularity of $\frac{R}{1-|R|^2}$. We expect that an appropriate "local" Riemann–Hilbert problem however is still explicitly solvable. It is definitely interesting to understand the actual contribution of the band edges and the way it is different from the free case due to the fact that the underlying Riemann surface has nonzero genus. We plan to investigate this issue in a future publication. But in the present work, we will assume that the stationary phase points stay away from the E_k (see Theorems 1.1, 5.4).

Case (iii). In the case where no stationary phase points lie in the spectrum the situation is similar to the case (i). 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 (4.23) we set

$$\begin{aligned} m^5 &= m^4 \tilde{B}_+^{-1}, \quad p \in D_k,\ k < j,\\ m^5 &= m^4 \tilde{B}_-^{-1}, \quad p \in D_k^*,\ k < j,\\ m^5 &= m^4 \tilde{b}_+^{-1}, \quad p \in D_k,\ k > j,\\ m^5 &= m^4 \tilde{b}_-^{-1}, \quad p \in D_k^*,\ k > j,\\ m^5 &= m^4, \quad \text{otherwise}. \end{aligned} \tag{4.28}$$

It is now easy to check that in both cases (i) and (iii) formula (4.12) is still true.

5. 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 [3]). In our case the underlying space is a Riemann surface. Hence we have 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 [18, Sec. 4]. Allowing poles at the nonspecial divisor $\mathcal{D}_{\underline{\hat{\mu}}}$ the corresponding Cauchy kernel is given by

$$\Omega_p^{\underline{\hat{\mu}}} = \omega_{p \infty_+} + \sum_{j=1}^{g} I_j^{\underline{\hat{\mu}}}(p)\zeta_j, \tag{5.1}$$

where

$$I_j^{\underline{\hat{\mu}}}(p) = \sum_{\ell=1}^{g} c_{j\ell} \int_{\infty_+}^{p} \omega_{\hat{\mu}_\ell,0}. \tag{5.2}$$

Note that $I_j^{\underline{\hat{\mu}}}(p)$ has first order poles at the points $\underline{\hat{\mu}}$.

The constants $c_{j\ell}$ 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 [10, Sect. III.3] or [20, 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 $1 \leq \ell \leq g$, then the divisor of $\zeta = \sum_{k=1}^{g} c_k\zeta_k$ satisfies $(\zeta) \geq \mathcal{D}_{\underline{\hat{\mu}}}$. 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 $c_k = 0$.

We will always assume that none of the poles lie on our contour Σ. This can be done without loss of generality since we can move the contour a little without changing the value at ∞_- (which is the only value we are eventually interested in). Furthermore, we will, for simplicity, also assume that Σ is sufficiently smooth, compact, and does not contain $\infty_\pm$. We will abbreviate $L^p(\Sigma) = L^p(\Sigma, \mathbb{C}^{2\times 2})$.

Theorem 5.1. *Set*

$$\underline{\Omega}_p^{\underline{\hat{\mu}}} = \begin{pmatrix} \Omega_p^{\underline{\hat{\mu}}^*} & 0 \\ 0 & \Omega_p^{\underline{\hat{\mu}}} \end{pmatrix} \tag{5.3}$$

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

$$(Cf)(p) = \frac{1}{2\pi \mathrm{i}} \int_\Sigma f\, \underline{\Omega}_p^{\underline{\hat{\mu}}}, \quad \textit{for} \quad p \notin \Sigma, \tag{5.4}$$

and

$$(C_\pm f)(q) = \lim_{p\to q\in\Sigma} (Cf)(p) \tag{5.5}$$

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*

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

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 $\Omega_q^{\hat{\mu}} = (\frac{1}{z-z(q)} + O(1))dz$ and hence the first part follows as in the Cauchy case on the complex plane (cf. [16]) 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 $p \to \infty_+$. Moreover, $I_j^{\underline{\hat{\mu}}^*}(\infty_+) = 0$ and hence (iii) holds. □

Now, given any $b_-, b_+ \in L^\infty(\Sigma)$ with determinant equal to 1, let the operator $C_w : L^2(\Sigma) \to L^2(\Sigma)$ be defined by

$$C_w f = C_+(fw_-) + C_-(fw_+) \tag{5.6}$$

for a 2×2 matrix valued f, where

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

Theorem 5.2. *Assume that* μ *solves the singular integral equation*

$$\mu = \mathbb{I} + C_w\mu \quad \textit{in} \quad L^2(\Sigma). \tag{5.7}$$

Let Q *be defined by the integral formulae*

$$Q = \mathbb{I} + C(\mu w) \quad \textit{on } \mathbb{M}\setminus\Sigma, \tag{5.8}$$

where $w = w_+ + w_-$. *Then* Q *is a solution of the following meromorphic Riemann–Hilbert problem.*

$$\begin{aligned} Q_+(p) &= Q_-(p) b_-^{-1}(p) b_+(p), \quad p \in \Sigma,\\ Q(\infty_+) &= \mathbb{I},\\ (Q_{j1}) \ge -\mathcal{D}_{\underline{\hat{\mu}}^*}, &\quad (Q_{j2}) \ge -\mathcal{D}_{\underline{\hat{\mu}}}. \end{aligned} \tag{5.9}$$

Proof. To show that Q defined above solves (5.9) note that

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

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

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

and similarly $Q_- = \mu(\mathbb{I} - w_-)$. Hence $Q_+ b_+^{-1} = \mu = Q_- b_-^{-1}$ and thus thus $Q_+ = Q_-(b_-)^{-1} b_+$. This proves the jump condition. That Q has the right devisor and the correct normalization at ∞_+ follows from Theorem 5.1 (ii) and (iii), respectively. □

Remark 5.3. *A few remarks are in order:*

(i). The theorem stated above does not address uniqueness. However, we do have uniqueness of $Q(\infty_-)$. For the proof of this, just follow Remark 3.2.

(ii). If we only assume $\mu = \mathbb{I} + C_w \mu$ in $L^2(\Sigma)$, we are only guaranteed an L^2-solution of the Riemann–Hilbert problem, that is, the limits are taken in the L^2 sense and the values of Q_+, Q_- on Σ are L^2. This is all we need here. On the other hand, we could have alternatively used the Hölder theory described in [14]. *Of course, in our current concrete situation the analyticity of the jumps for m^4, m^5 implies the continuity of the limiting values $m^4_\pm$, $m^5_\pm$ (except at the cross center) and thus also of $m^3_\pm$ away from the band edges E_j.*

(iii). 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 (5.8) evaluated at ∞_-. We write it as

$$\begin{aligned} Q(\infty_-) &= (\mathbb{I} + C(\mu w))(\infty_-) \\ &= \mathbb{I} + \int_\Sigma (\mathbb{I} - C_w)^{-1}(\mathbb{I})\, w\, \underline{\Omega}^{\hat{\underline{\mu}}}_{\infty_-} \end{aligned} \tag{5.10}$$

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

Theorem 5.4. *A unique solution Q of the Riemann–Hilbert problems (4.24) and (4.27) exists such that Q is uniformly bounded in (n, t) away from the poles.*

Proof. Existence and uniqueness follows from Theorem A.1. Moreover, note that the jump for both m^5 and m^6 is of the form $\mathbb{I} +$ *off-diagonal* where the off-diagonal part is bounded in norm by $|R|$.

Furthermore the operator $(\mathbb{I}-C_w)^{-1}$ corresponding to both problems (with trivial factorization of the jump) is clearly uniformly bounded in norm. This in turn follows from the Neumann series

$$(\mathbb{I}-C_w)^{-1} = \mathbb{I} + C_w + C_w^2 + C_w^3 + ... \tag{5.11}$$

and the fact that $\|C_{w^6}\|$ is bounded by a constant that is strictly less than 1 and independent of n, t. The bound for $\|C_{w^6}\|$ follows from the following facts:

The singular part of the Cauchy operators $C_\pm$ as defined by (5.5) has norm equal to one (see e.g. [19], p.58). Hence the singular part of C_{w^6} has norm bounded by the supremum of $|R|$, which is less than one in the domain we are interested.

The nonsingular part of the Cauchy operators is small (for large times) because the contour where J^6 lies is of length $= t^{-1/2+\varepsilon}$.

So, after all, $(\mathbb{I}-C_{w^6})^{-1}$ is uniformly bounded in norm, and by, say, the second resolvent identity $(\mathbb{I}-C_{w^5})^{-1}$ is uniformly bounded in norm, hence m^5, m^6 are uniformly bounded in (n,t) away from the poles. □

The next theorem will enable us to

(i) first remove the bulk of the contour in problem (4.24), away from the stationary phase points,
(ii) also remove the rest of the contour.

Theorem 5.5. *Suppose we have a Riemann–Hilbert problem as in (5.9) above with w depending on parameters n, t such that the solution given by (5.8), or equivalently $(\mathbb{I}-C_w)^{-1}(\mathbb{I})(s)$, is uniformly bounded in n, t on Σ. Suppose first that the jump matrix is such that $\|w_\pm\|_\infty < \varepsilon$, where ε is a given small number (independent of n, t).*

Then the solution of the Riemann–Hilbert problem given by (5.8) satisfies $\|Q(\infty_-)-\mathbb{I}\| < C\varepsilon$, where C is some constant independent of n, t.

Suppose then that the length of the contour is $o(1)$ as $t \to \infty$. Then the solution of the Riemann–Hilbert problem given by (5.8) satisfies $\|Q(\infty_-)-\mathbb{I}\| \to 0$, as $t \to \infty$ uniformly in n, t.

Proof. Both assertions follow from the integral formula (5.10) above. □

Theorem 1.1 follows in view of Theorem 3.1.

Proof of Theorem 1.1. In order to compare the solution m^6 of (4.27) and the solution m^5 of (4.24) observe that the difference of the jumps lives in the complement of the two small crosses L with respect to the jump for m^5, that is $\cup_j (C_j \cup C_j^*) \setminus L$. Hence the difference of the jumps is a uniformly exponentially small quantity, so both solutions are uniformly close at least near ∞_- (which is all we care here) by the second resolvent identity.

From Theorem 5.5 above we see that $m^6(\infty_-)$ is uniformly close to the identity. Then $m^5(\infty_-)$ is uniformly close to the identity and by tracing back to m^4 and m^3 (according to the discussion in Section 4) and using (3.24) we have our result. □

Proof of Theorem 1.3. We begin by showing

$$m^2(p) = \begin{pmatrix} A_+ & \frac{1}{A_+} \end{pmatrix} m^3(p).$$

In order to avoid the possible singularities of m^3 at the band edges, we will show the equivalent statement

$$\tilde{m}^2(p) = \begin{pmatrix} d(\infty_-)A_+ & \frac{1}{A_+} \end{pmatrix} m^4(p),$$

where

$$\tilde{m}^2(p) = m^2(p)D(p)$$

(compare the definition (4.16) of m^4). By the factor $T(z)$ in the definition (3.12) of m^2, the new quantity $\tilde{m}^2$ is bounded near the band edges and thus both sides solve the vector Riemann-Hilbert problem

$$\begin{aligned} &\tilde{m}^2_+(p) = \tilde{m}^2_-(p) J^4(p), \quad p \in \Sigma, \\ &(\tilde{m}^2_1) \geq -\mathcal{D}_{\underline{\hat{\nu}}^*}, \quad (\tilde{m}^2_2) \geq -\mathcal{D}_{\underline{\hat{\nu}}}, \\ &\tilde{m}^2(\infty_+) = \begin{pmatrix} d(\infty_-)A_+ & \frac{1}{A_+} \end{pmatrix}, \end{aligned} \tag{5.12}$$

whose solution is unique by Theorem A.1 (cf. Remark A.2).

Combining this with (4.25) we obtain

$$m^2(p,n,t) \sim \begin{pmatrix} \frac{A_+(n,t)d(\infty_-,n,t)}{d(p^*,n,t)} & \frac{1}{A_+(n,t)d(p,n,t)} \end{pmatrix} \tag{5.13}$$

for p near $\infty_\pm$. Moreover, for $p=(z,+)$ we have

$$m^2(p) = \begin{pmatrix} A_+ & \frac{1}{A_+} \end{pmatrix} + \begin{pmatrix} -A_+B_+ & \frac{B_+}{A_+} \end{pmatrix} \frac{1}{z} + O(\frac{1}{z^2})$$

and thus

$$B_+(n,t) \sim -d_1(n,t),$$

where d_1 is given by

$$d(p) = 1 + \frac{d_1}{z} + O(\frac{1}{z^2}).$$

Hence it remains to compute d_1. Proceeding as in [20, Thm. 9.4] respectively [22, 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 (2.14). □

6. 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. [4], [11]) 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).

Although the most celebrated applications of the deformation method initiated by [4] for the asymptotic evaluation of solutions of Riemann-Hilbert factorization problems have been in statistical mechanics (via orthogonal polynomials, random matrices and combinatorial probability), most mathematical innovations have appeared in the study of nonlinear dispersive PDEs or systems of ODEs (cf. [4], [5], [14]). It is thus interesting that another mathematical extension of the theory (setting the factorization problem on a Riemann surface rather than the complex plane) arises in the study of an innocent looking stability problem for the periodic Toda lattice.

Appendix A. An existence 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 [25]. In this section we state and prove an analogous theorem where the underlying spectral curve is our hyperelliptic curve with real branch cuts.

We state and prove the following theorem assuming local analyticity of the jump and thus proving continuity of the limiting values. But (cf. Remark 5.3 (ii)) we could alternatively drop the analyticity condition and then only guarantee an L^2-solution of the Riemann–Hilbert problem, which is all we need here. On the other hand, we could have alternatively used the Hölder theory described in [14].

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

Theorem A.1. *The following matrix Riemann–Hilbert problem on* $\mathbb{M}$ *has always a unique solution and the limits* Q_+, Q_- *on* Σ *are continuous.*

$$\begin{aligned} Q_+(p) = Q_-(p)J(p), \quad p \in \Sigma, \\ Q(\infty_+) = \mathbb{I}, \quad (Q_{j1}) \geq -\mathcal{D}_{\underline{\hat{\mu}}^*}, \quad (Q_{j2}) \geq -\mathcal{D}_{\underline{\hat{\mu}}} \end{aligned} \tag{A.1}$$

where Σ *is an oriented contour, consisting of a union of finitely many smooth arcs, which 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.

The divisor $\mathcal{D}_{\underline{\hat{\mu}}}$ *is such that* $\pi(\mu_j) \in [E_{2j-1}, E_{2j}]$.

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. Continuity of the limits Q_+, Q_- follows immediately from the analyticity of J.

Let $J(z) = b_-^{-1}(z)b_+(z) = \mathbb{I}J(z)$ be the trivial factorization $b_- = \mathbb{I}$, $b_+ = J$. The operator $\mathbb{I} - C_w$ is defined as in Section 5 with $w_+ = J - \mathbb{I}$ and $w_- = 0$. By Theorem 5.2, we only need to show that the operator $\mathbb{I} - C_w$ is invertible. Note first that $\mathbb{I} - C_w$ is a Fredholm operator with index, $\mathrm{ind}(\mathbb{I} - C_w)$, zero. Indeed one can easily check that

$$(\mathbb{I} - C_w)(\mathbb{I} - C_{\tilde{w}}) = \mathbb{I} + T_w, \tag{A.2}$$

where $T_w(f) = C_-[C_-(f\tilde{w}_+)w_+]$ and

$$\tilde{w}_+ = -b_+ + \mathbb{I} = -J + \mathbb{I}, \qquad \tilde{w}_- = \mathbb{I} - \mathbb{I} = 0. \tag{A.3}$$

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 $w = w_-$ 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,

$$T_w f_n(p) = \frac{1}{2\pi\mathrm{i}} \int_{\Sigma'} (C_-(f_n w)\tilde{w}) \underline{\Omega}_p^{\underline{\hat{\mu}}}. \tag{A.4}$$

Now clearly $(C_-(f_n w)\tilde{w})(p) \to 0$ as $n \to \infty$. Also $|(C_-(f_n w)\tilde{w})(p)| < const \, \|f_n\|_{L^2} \|\tilde{w}\|_{L^2} < const$. By the dominated convergence theorem, $\|T_w f_n\|_{L^2} < const \, \|(C_-(f_n w)\tilde{w})(p)\| \to 0$.

Similarly, one can show that there exists a compact operator T_v such that

$$(\mathbb{I} - C_{\tilde{w}})(\mathbb{I} - C_w) = \mathbb{I} + T_v. \tag{A.5}$$

Hence by [17, Thm. 1.4.3] $\mathbb{I} - C_w$ is Fredholm.

Now considering $\mathrm{ind}(\mathbb{I} - \varepsilon C_w)$ for $0 \le \varepsilon \le 1$ and recall that C_w is bounded. It is a fact [17, Thm. 1.3.8] that $\mathrm{ind}(\mathbb{I} - \varepsilon C_w)$ is continuous with ε. Since it is an integer, it has to be constant. So in particular $\mathrm{ind}(\mathbb{I} - C_w) = \mathrm{ind}(\mathbb{I}) = 0$.

By the Fredholm alternative, it follows that to show the bounded invertibility of $\mathbb{I} - C_w$ we only need to show that $\ker(\mathbb{I} - C_w) = 0$.

So suppose there is a ν such that $(\mathbb{I} - C_w)\nu = 0$. We will prove that $\nu = 0$.

Define

$$M = C(\nu w) = C(\nu w_+). \tag{A.6}$$

Then, as in the proof of Theorem 5.2 one verifies $M_- = \nu$ and $M_+ = \nu J$. Hence M solves (A.1) except for the normalization at ∞_+, which now reads $M(\infty_+) = 0$. In other words M solves the associated vanishing Riemann–Hilbert problem. Also, as above, M_+, M_- are continuous.

Next we want to apply Cauchy's integral theorem to $M(p)M^*(\overline{p}^*)$. 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 a holomorphic away from the contour.

Indeed let

$$d\Omega = -\mathrm{i} \frac{\prod_{j=1}^{g} (\pi - \mu_j)}{R_{2g+2}^{1/2}} d\pi \tag{A.7}$$

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

$$\int_D M(p)M^*(\overline{p}^*)d\Omega(p), \tag{A.8}$$

where D is a $\overline{*}$-invariant contour consisting of one small loop in every connected component of $\mathbb{M} \setminus \Sigma$. 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_ν of $\Sigma\backslash\pi^{-1}(\sigma(H_q))$ there are two contributions to the integral on the deformed contour:

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

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 $M_- = \nu = 0$ and hence $M = 0$.

To prove uniqueness, suppose there were two solutions. Their difference would then satisfy the same jump and it would be zero at ∞_+. Using the above argument for M we see that the difference would have to vanish everywhere. □

Remark A.2. *The proof of the previous theorem also shows uniqueness for the following vector Riemann–Hilbert problem on* $\mathbb{M}$

$$\begin{aligned} Q_+(p) &= Q_-(p)J(p), \quad p \in \Sigma, \\ Q(\infty_+) = \begin{pmatrix}\alpha & \beta\end{pmatrix}, &\quad (Q_1) \geq -\mathcal{D}_{\underline{\hat{\mu}}^*}, \quad (Q_2) \geq -\mathcal{D}_{\underline{\hat{\mu}}} \end{aligned} \tag{A.9}$$

where $J(z)$, Σ, *and* $\mathcal{D}_{\underline{\hat{\mu}}}$ *satisfy the same assumptions as in the previous theorem and* $\begin{pmatrix}\alpha & \beta\end{pmatrix}$ *is some constant vector.*

The above Theorem does not apply directly to the Riemann–Hilbert problem for m^3 of Theorem 3.1, because of the (possible) singularities at the band edges (see (3.20), (3.21)). However, since condition (ii) holds for the Riemann–Hilbert problem (4.17) for m^4 (while condition (i) is irrelevant), we get existence and uniqueness for m^4 and thus also for m^3 by (4.16).

Theorem A.3. *The Riemann–Hilbert problem of Theorem 3.1 admits a unique solution.*

Again, the analyticity condition (iii) is not always satisfied for the jump matrix of Theorem 3.1. But, as noted by [4] (see also the discussion after (4.22)), one can always approximate the reflection coefficient by a holomorphic function near the jump contour. Since the nearby holomorphic Riemann–Hilbert problem has a solution, so does the possibly non-holomorphic one (see [4] for the details).

Alternatively, we can use the L^2 version of Theorem 6.1, without the analyticity condition.

Acknowledgments.
Spyridon Kamvissis gratefully acknowledges the support of the European Science Foundation (MISGAM program) during two visits to the University of Vienna in 2005 and 2006. He also acknowledges the support of the Austrian Science Fund (FWF) during a second visit in 2006.

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] 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).
[3] R. Beals, R. Coifman, *Scattering and inverse scattering for first order systems*, Comm. in Pure and Applied Math. **37**, 39–90 (1984).
[4] P. Deift, X. Zhou, *A steepest descent method for oscillatory Riemann–Hilbert problems*, Ann. of Math. (2) **137**, 295–368 (1993).
[5] 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).
[6] I. Egorova, J. Michor, and G. Teschl, *Scattering theory for Jacobi operators with quasi-periodic background*, Comm. Math. Phys. **264-3**, 811–842 (2006).
[7] 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).
[8] I. Egorova, J. Michor, and G. Teschl, *Scattering theory for Jacobi operators with steplike quasi-periodic background*, Inverse Problems **23**, 905–918 (2007).
[9] I. Egorova, J. Michor, and G. Teschl, *Soliton solutions of the Toda hierarchy on quasi-periodic background revisited*, Math. Nach. (to appear).
[10] H. Farkas and I. Kra, *Riemann Surfaces*, 2$^{\text{nd}}$ edition, GTM 71, Springer, New York, 1992.
[11] 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).

[12] S. Kamvissis, *Long time behavior for the focusing nonlinear Schrödinger equation with real spectral singularities,* Comm. Math. Phys., **180-2**, 325–343 (1996).

[13] S. Kamvissis and G. Teschl, *Stability of periodic soliton equations under short range perturbations*, Phys. Lett. A, **364-6**, 480–483 (2007).

[14] S. Kamvissis, K. McLaughlin, and P. Miller, *Semiclassical Soliton Ensembles for the Focusing Nonlinear Schrödinger Equation,* Annals of Mathematics, Study 154, Princeton University Press, Princeton, 2003.

[15] 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.

[16] N.I. Muskhelishvili, *Singular Integral Equations*, P. Noordhoff Ltd., Groningen, 1953.

[17] S. Prössdorf, *Some Classes of Singular Equations*, North-Holland, Amsterdam, 1978.

[18] Yu. Rodin, *The Riemann Boundary Problem on Riemann Surfaces*, Mathematics and its Applications (Soviet Series) **16**, D. Reidel Publishing Co., Dordrecht, 1988.

[19] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

[20] G. Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*, Math. Surv. and Mon. **72**, Amer. Math. Soc., Rhode Island, 2000.

[21] G. Teschl, *Almost everything you always wanted to know about the Toda equation*, Jahresber. Deutsch. Math.-Verein. **103**, no. 4, 149–162 (2001).

[22] G. Teschl, *Algebro-geometric constraints on solitons with respect to quasi-periodic backgrounds*, Bull. London Math. Soc. (to appear).

[23] M. Toda, *Theory of Nonlinear Lattices*, 2nd enl. ed., Springer, Berlin, 1989.

[24] 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 (1963).

[25] 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`

Faculty of Mathematics, 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/`