Global action-angle variables for the periodic Toda lattice

Kappeler, T; Henrici, A
Global action-angle variables for the periodic Toda lattice

Andreas Henrici    Thomas Kappeler

February 27, 2008

Abstract
In this paper we construct global action-angle variables for the periodic Toda lattice.

1 Introduction
Consider the Toda lattice with period $N$ ($N \geq 2$),

$$\dot{q}_n = \partial_{p_n} H, \quad \dot{p}_n = -\partial_{q_n} H$$

for $n \in \mathbb{Z}$, where the (real) coordinates $(q_n, p_n)_{n \in \mathbb{Z}}$ satisfy $(q_{n+N}, p_{n+N}) = (q_n, p_n)$ for any $n \in \mathbb{Z}$ and the Hamiltonian $H_{Toda}$ is given by

$$H_{Toda} = \frac{1}{2} \sum_{n=1}^{N} p_n^2 + \alpha^2 \sum_{n=1}^{N} e^{q_n - q_{n+1}}$$

(1)

where $\alpha$ is a positive parameter, $\alpha > 0$. For the standard Toda lattice, $\alpha = 1$. The Toda lattice was introduced by Toda [19] and studied extensively in the sequel. It is an important model for an integrable system of $N$ particles in one space dimension with nearest neighbor interaction and belongs to the family of lattices introduced and numerically investigated by Fermi, Pasta, and Ulam in their seminal paper [5]. To prove the integrability of the Toda lattice, Flaschka introduced in [3] the (noncanonical) coordinates

$$b_n := -p_n \in \mathbb{R}, \quad a_n := \alpha e^{\frac{1}{2}(q_n - q_{n+1})} \in \mathbb{R}_{>0} \quad (n \in \mathbb{Z}).$$

These coordinates describe the motion of the Toda lattice relative to the center of mass. Note that the total momentum is conserved by the Toda flow, hence any trajectory of the center of mass is a straight line.

*Supported in part by the Swiss National Science Foundation, the programme SPECT, and the European Community through the FP6 Marie Curie RTN ENIGMA (MRTN-CT-2004-5652)

12000 Mathematics Subject Classification: 37J35, 39A12, 39A70, 70H06
In these coordinates the Hamiltonian $H_{Toda}$ takes the simple form

$$H = \frac{1}{2} \sum_{n=1}^{N} b_n^2 + \sum_{n=1}^{N} a_n^2,$$

and the equations of motion are

$$\begin{aligned}
\dot{b}_n &= a_n^2 - a_{n-1}^2 \\
\dot{a}_n &= \frac{1}{4} a_n (b_{n+1} - b_n)
\end{aligned} \quad (n \in \mathbb{Z}).$$

(2)

Note that $(b_{n+N}, a_{n+N}) = (b_n, a_n)$ for any $n \in \mathbb{Z}$, and $\prod_{n=1}^{N} a_n = a^N$. Hence we can identify the sequences $(b_n)_{n \in \mathbb{Z}}$ and $(a_n)_{n \in \mathbb{Z}}$ with the vectors $(b_n)_{1 \leq n \leq N} \in \mathbb{R}^N$ and $(a_n)_{1 \leq n \leq N} \in \mathbb{R}_{>0}^N$. Our aim is to study the normal form of the system of equations (2) on the phase space

$$\mathcal{M} := \mathbb{R}^N \times \mathbb{R}_{>0}^N.$$

This system is Hamiltonian with respect to the nonstandard Poisson structure $J \equiv J_{b,a}$, defined at a point $(b, a) = ((b_n, a_n)_{1 \leq n \leq N}$ by

$$J = \begin{pmatrix}
0 & A \\
-tA & 0
\end{pmatrix},$$

(3)

where $A$ is the $b$-independent $N \times N$-matrix

$$A = \frac{1}{2} \begin{pmatrix}
a_1 & 0 & \ldots & 0 & -a_N \\
-a_1 & a_2 & 0 & \ddots & 0 \\
0 & -a_2 & a_3 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & \ldots & 0 & -a_{N-1} & a_N
\end{pmatrix},$$

(4)

The Poisson bracket corresponding to (3) is then given by

$$\{F, G\} J(b, a) = \langle (\nabla_b F, \nabla_a F), J (\nabla_b G, \nabla_a G) \rangle_{\mathbb{R}^{2N}}$$

$$= \langle \nabla_b F, A \nabla_a G \rangle_{\mathbb{R}^N} - \langle \nabla_a F, A^T \nabla_b G \rangle_{\mathbb{R}^N}. $$

(5)

where $F, G \in C^1(\mathcal{M})$ and where $\nabla_b$ and $\nabla_a$ denote the gradients with respect to the $N$-vectors $b = (b_1, \ldots, b_N)$ and $a = (a_1, \ldots, a_N)$, respectively. Therefore, equations (2) can alternatively be written as $\dot{b}_n = \{b_n, H\} J_1, \dot{a}_n = \{a_n, H\} J_1$ ($1 \leq n \leq N$). Further note that

$$\{b_n, a_n\} J = \frac{a_n}{2}, \quad \{b_{n+1}, a_n\} J = -\frac{a_n}{2},$$

(6)

while $\{b_n, a_k\} J = 0$ for any $n, k$ with $n \notin \{k, k + 1\}$. 

Since the matrix $A$ defined by (4) has rank $N - 1$, the Poisson structure $J$ is degenerate. It admits the two Casimir functions:

$$C_1 := -\frac{1}{N} \sum_{n=1}^{N} b_n \quad \text{and} \quad C_2 := \left(\prod_{n=1}^{N} a_n\right)^{\frac{1}{N}}$$

(7)

whose gradients $\nabla_{b,a} C_i = (\nabla_b C_i, \nabla_a C_i)$ ($i = 1, 2$), given by

$$\nabla_b C_1 = -\frac{1}{N} (1, \ldots, 1), \quad \nabla_a C_1 = 0,$$

(8)

$$\nabla_b C_2 = 0, \quad \nabla_a C_2 = \frac{C_2}{N} \left(\frac{1}{a_1}, \ldots, \frac{1}{a_N}\right),$$

(9)

are linearly independent at each point $(b,a)$ of $M$.

The main result of this paper is the following one:

**Theorem 1.1.** The periodic Toda lattice admits globally defined action-angle variables. More precisely:

(i) There exist real analytic functions $(I_n)_{1 \leq n \leq N-1}$ on $M$ which are pairwise in involution and which Poisson commute with the Toda Hamiltonian $H$ and the two Casimir functions $C_1, C_2$, i.e. for any $1 \leq m, n \leq N - 1$, $i = 1, 2$,

$$\{I_m, I_n\}_J = 0 \quad \text{on } M$$

and

$$\{H, I_n\}_J = 0 \quad \text{and} \quad \{C_i, I_n\}_J = 0 \quad \text{on } M.$$

(ii) For any $1 \leq n \leq N - 1$ there exist a real analytic submanifold $D_n$ of codimension 2 and a function $\theta_n : M \setminus D_n \to \mathbb{R}$, defined mod $2\pi$ and real analytic when considered mod $\pi$, so that on $M \setminus \bigcup_{n=1}^{N-1} D_n$, $(\theta_n)_{1 \leq n \leq N-1}$ and $(I_n)_{1 \leq n \leq N-1}$ are conjugate variables. More precisely, for any $1 \leq m, n \leq N - 1$, $i = 1, 2$,

$$\{I_m, \theta_n\}_J = \delta_{mn} \quad \text{and} \quad \{C_i, \theta_n\}_J = 0 \quad \text{on } M \setminus D_n$$

and

$$\{\theta_m, \theta_n\}_J = 0 \quad \text{on } M \setminus (D_m \cup D_n).$$

Let $M_{\beta, \alpha} := \{(b,a) \in \mathbb{R}^{2N} : (C_1, C_2) = (\beta, \alpha)\}$ denote the level set of $(C_1, C_2)$ for $(\beta, \alpha) \in \mathbb{R} \times \mathbb{R}_{>0}$. Note that $(-\beta 1_N, \alpha 1_N) \in M_{\beta, \alpha}$ where $1_N = (1, \ldots, 1) \in \mathbb{R}^N$. As the gradients $\nabla_{b,a} C_1$ and $\nabla_{b,a} C_2$ are linearly independent everywhere on $M$, the sets $M_{\beta, \alpha}$ are (real analytic) submanifolds of $M$ of codimension two. Furthermore the Poisson structure $J$, restricted to $M_{\beta, \alpha}$, becomes nondegenerate everywhere on $M_{\beta, \alpha}$ and therefore induces a symplectic structure $\nu_{\beta, \alpha}$ on $M_{\beta, \alpha}$. In this way, we obtain a symplectic foliation of $M$ with $M_{\beta, \alpha}$ being the symplectic leaves.

---

2A smooth function $C : M \to \mathbb{R}$ is a Casimir function for $J$ if $\{C, \cdot\}_J \equiv 0$. 
Corollary 1.2. On each symplectic leaf $M_{\beta,\alpha}$, the action variables $(I_n)_{1 \leq n \leq N-1}$ are a maximal set of functionally independent integrals in involution of the periodic Toda lattice.

In subsequent work [9], we will use Theorem 1.1 to construct global Birkhoff coordinates for the periodic Toda lattice. More precisely, we introduce the model space $\mathcal{P} := \mathbb{R}^{2(N-1)} \times \mathbb{R} \times \mathbb{R}_{>0}$ endowed with the degenerate Poisson structure $J_0$ whose symplectic leaves are $\mathbb{R}^{2(N-1)} \times \{\beta\} \times \{\alpha\}$ endowed with the standard Poisson structure, and prove the following theorem:

Theorem 1.3. There exists a real analytic, canonical diffeomorphism

$$
\Omega : (\mathcal{M}, J) \rightarrow (\mathcal{P}, J_0)
(b, a) \mapsto (x_n, y_n)_{1 \leq n \leq N-1}, C_1, C_2)
$$

such that the coordinates $(x_n, y_n)_{1 \leq n \leq N-1}, C_1, C_2$ are global Birkhoff coordinates for the periodic Toda lattice, i.e. $(x_n, y_n)_{1 \leq n \leq N-1}$ are canonical coordinates, $C_1, C_2$ are the Casimirs and the transformed Toda Hamiltonian $\hat{H} = H \circ \Omega^{-1}$ is a function of the actions $(I_n)_{1 \leq n \leq N-1}$ and $C_1, C_2$ alone.

In [10] we used Theorem 1.3 to obtain a KAM theorem for Hamiltonian perturbations of the periodic Toda lattice.

Related work: Theorem 1.1 and Theorem 1.3 improve on earlier work on the normal form of the periodic Toda lattice in [1, 2]. In particular, we construct global Birkhoff coordinates on all of $\mathcal{M}$ instead of a single symplectic leaf and show that techniques recently developed for treating the KdV equation (cf. [11, 12]) and the defocusing NLS equation (cf. [8, 16]) can also be applied for the Toda lattice.

Outline of the paper: In section 2 we review the Lax pair of the periodic Toda lattice and collect some auxiliary results on the spectrum of the Jacobi matrix $L(b, a)$ associated to an element $(b, a) \in \mathcal{M}$. In section 3 we study the action variables $(I_n)_{1 \leq n \leq N-1}$, and in section 4 we define the angle variables $(\theta_n)_{1 \leq n \leq N-1}$ on $\mathcal{M} \setminus \cup_{n=1}^{N-1} D_n$ using holomorphic differentials defined on the hyperelliptic Riemann surface associated to the spectrum of $L(b, a)$. In sections 5 and 6 we establish formulas of the gradients of the actions and angles in terms of products of fundamental solutions and prove orthogonality relations between such products which are then used in section 7 to show that $(I_n)_{1 \leq n \leq N-1}$ and $(\theta_n)_{1 \leq n \leq N-1}$ are canonical variables and to prove Theorem 1.1 and Corollary 1.2.
2 Preliminaries

It is well known (cf. e.g. [19]) that the system (2) admits a Lax pair formulation

\[ \dot{L} = \partial_t L = [B, L], \]

where

\[
L = L + (b, a) \equiv \begin{pmatrix}
  b_1 & a_1 & 0 & \ldots & \pm a_N \\
  a_1 & b_2 & a_2 & \ddots & \\
  0 & a_2 & b_3 & \ddots & 0 \\
  \vdots & \ddots & \ddots & \ddots & a_N \\
  \pm a_N & \ldots & 0 & a_{N-1} & b_N
\end{pmatrix},
\]

and \(B\) the skew-symmetric matrix

\[
B = \begin{pmatrix}
  0 & a_1 & 0 & \ldots & -a_N \\
  -a_1 & 0 & a_2 & \ddots & \\
  0 & -a_2 & \ddots & \ddots & 0 \\
  \vdots & \ddots & \ddots & \ddots & a_N \\
  a_N & \ldots & 0 & -a_{N-1} & 0
\end{pmatrix}.
\]

Hence the flow of \(\dot{L} = [B, L]\) is isospectral.

**Proposition 2.1.** For a solution \((b(t), a(t))\) of the periodic Toda lattice (2), the eigenvalues \((\lambda_j \pm j)_{1 \leq j \leq N}\) of \(L(b(t), a(t))\) are conserved quantities.

Let us now collect a few results from [17] and [19] of the spectral theory of Jacobi matrices needed in the sequel. Denote by \(\mathcal{M}^C\) the complexification of the phase space \(M\),

\[ \mathcal{M}^C = \{ (b, a) \in \mathbb{C}^{2N} : \text{Re } a_j > 0 \quad \forall 1 \leq j \leq N \}. \]

For \((b, a) \in \mathcal{M}^C\) we consider for any complex number \(\lambda\) the difference equation

\[ (R_{b,a}y)(k) = \lambda y(k) \quad (k \in \mathbb{Z}) \]

where \(y(\cdot) = y(k)_{k \in \mathbb{Z}} \in \mathbb{C}^\mathbb{Z}\) and \(R_{b,a}\) is the difference operator

\[ R_{b,a} = a_{k-1}S^{-1} + b_k S^0 + a_k S^1 \]

with \(S^m\) denoting the shift operator of order \(m \in \mathbb{Z}\), i.e.

\[ (S^m y)(k) = y(k + m) \text{ for } k \in \mathbb{Z}. \]

**Fundamental solutions:** The two fundamental solutions \(y_1(\cdot, \lambda)\) and \(y_2(\cdot, \lambda)\) of (11) are defined by the standard initial conditions \(y_1(0, \lambda) = 1, y_1(1, \lambda) = 0\) and \(y_2(0, \lambda) = 0, y_2(1, \lambda) = 1\). They satisfy the Wronskian identity

\[ W(n) := y_1(n, \lambda)y_2(n + 1, \lambda) - y_1(n + 1, \lambda)y_2(n, \lambda) = \frac{a_N}{a_n}. \]
Note that for \( n = N \) one gets
\[ W(N) = 1. \tag{14} \]

For each \( k \in \mathbb{N}, y_i(k, \lambda, b, a), i = 1, 2, \) is a polynomial in \( \lambda \) of degree at most \( k - 1 \) and depends real analytically on \((b, a)\) (see [17]). In particular, one easily verifies that \( y_2(N + 1, \lambda, b, a) \) is a polynomial in \( \lambda \) of degree \( N \) with leading coefficient \( \alpha^{-N} \).

**Wronskian:** More generally, one defines for any two sequences \((v(n))_{n \in \mathbb{Z}}\) and \((w(n))_{n \in \mathbb{Z}}\) the Wronskian sequence \((W(n))_{n \in \mathbb{Z}} = (W(v, w)(n))_{n \in \mathbb{Z}}\) by
\[ W(n) := v(n)w(n + 1) - v(n + 1)w(n). \]

Let us recall the following properties of the Wronskian, which can be easily verified.

**Lemma 2.2.**
(i) If \( y \) and \( z \) are solutions of (11) for \( \lambda = \lambda_1 \) and \( \lambda = \lambda_2 \), respectively, then \( W = W(y, z) \) satisfies for any \( k \in \mathbb{Z} \)
\[ a_k W(k) = a_{k-1} W(k-1) + (\lambda_2 - \lambda_1) y(k)z(k). \tag{15} \]

(ii) If \( y(\cdot, \lambda) \) is a 1-parameter-family of solutions of (11) which is continuously differentiable with respect to the parameter \( \lambda \) and \( \dot{y}(k, \lambda) := \frac{\partial}{\partial \lambda} y(k, \lambda) \), then \( W = W(y, \dot{y}) \) satisfies for any \( k \in \mathbb{Z} \)
\[ a_k W(k) = a_{k-1} W(k-1) + y(k, \lambda)^2. \tag{16} \]

**Discriminant:** We denote by \( \Delta(\lambda) \equiv \Delta(\lambda, b, a) \) the discriminant of (11), defined by
\[ \Delta(\lambda) := y_1(N, \lambda) + y_2(N + 1, \lambda). \tag{17} \]

In the sequel, we will often write \( \Delta_\lambda \) for \( \Delta(\lambda) \). Note that \( y_2(N + 1, \lambda) \) is a polynomial in \( \lambda \) of degree \( N \) with leading term \( \alpha^{-N} \lambda^N \), whereas \( y_1(N, \lambda) \) is a polynomial in \( \lambda \) of degree less than \( N \), hence \( \Delta(\lambda, b, a) \) is a polynomial in \( \lambda \) of degree \( N \) with leading term \( \alpha^{-N} \lambda^N \), and it depends real analytically on \((b, a)\) (see e.g. [19]). According to Floquet’s Theorem (see e.g. [18]), for \( \lambda \in \mathbb{C} \) given, (11) admits a periodic or antiperiodic solution of period \( N \) if the discriminant \( \Delta(\lambda) \) satisfies \( \Delta(\lambda) = +2 \) or \( \Delta(\lambda) = -2 \), respectively. (These solutions correspond to eigenvectors of \( L^+ \) or \( L^- \), respectively, with \( L^\pm \) defined by (10).) It turns out to be more convenient to combine these two cases by
Figure 1: Examples of the discriminant $\Delta(\lambda)$

considering the periodic Jacobi matrix $Q \equiv Q(b,a)$ of size $2N$ defined by

$$Q = \begin{pmatrix}
    b_1 & a_1 & \ldots & 0 & 0 & \ldots & 0 & a_N \\
    a_1 & b_2 & \ddots & \vdots & \vdots & & \vdots \\
    \vdots & \ddots & \ddots & a_{N-1} & \vdots & & \vdots \\
    0 & \ddots & a_{N-1} & b_N & a_N & \ldots & 0 & 0 \\
    0 & \ldots & 0 & a_N & b_1 & a_1 & \ldots & 0 \\
    0 & \ldots & 0 & a_1 & b_2 & \ddots & \vdots \\
    \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
    a_N & \ldots & 0 & 0 & 0 & \ddots & a_{N-1} & b_N
\end{pmatrix}.$$ 

Then the spectrum of the matrix $Q$ is the union of the spectra of the matrices $L^+$ and $L^-$ and therefore the zero set of the polynomial $\Delta^2 - 4$. The function $\Delta^2 - 4$ is a polynomial in $\lambda$ of degree $2N$ and admits a product representation

$$\Delta^2 - 4 = \alpha^{-2N} \prod_{j=1}^{2N} (\lambda - \lambda_j). \quad (18)$$

The factor $\alpha^{-2N}$ in (18) comes from the above mentioned fact that the leading term of $\Delta(\lambda)$ is $\alpha^{-N} \lambda^N$.

For any $(b,a) \in \mathcal{M}$, the matrix $Q$ is symmetric and hence the eigenvalues $(\lambda_j)_{1 \leq j \leq 2N}$ of $Q$ are real. When listed in increasing order and with their algebraic multiplicities, they satisfy the following relations (cf. [17])

$$\lambda_1 < \lambda_2 \leq \lambda_3 < \lambda_4 \leq \ldots \lambda_{2N-2} \leq \lambda_{2N-1} < \lambda_{2N}.$$ 

As explained above, the $\lambda_j$ are periodic or antiperiodic eigenvalues of $L$ and thus eigenvalues of $L^+$ or $L^-$ according to whether $\Delta(\lambda) = 2$ or $\Delta(\lambda) = -2$. One has (cf. [17])

$$\Delta(\lambda_1) = (-1)^N \cdot 2, \quad \Delta(\lambda_{2n}) = \Delta(\lambda_{2n+1}) = (-1)^{n+N} \cdot 2, \quad \Delta(\lambda_{2N}) = 2. \quad (19)$$
Since $\Delta_\lambda$ is a polynomial of degree $N$, $\hat{\Delta}_\lambda = \hat{\Delta}(\lambda) = \frac{d^N}{d\lambda^N}\Delta(\lambda)$ is a polynomial of degree $N - 1$, and it admits a product representation of the form

$$\hat{\Delta}_\lambda = N\alpha^{-N} \prod_{k=1}^{N-1} (\lambda - \hat{\lambda}_k).$$  

(20)

The zeroes $(\hat{\lambda}_n)_{1 \leq n \leq N-1}$ of $\hat{\Delta}_\lambda$ satisfy $\lambda_{2n} \leq \hat{\lambda}_n \leq \lambda_{2n+1}$ for any $1 \leq n \leq N - 1$. The intervals $(\lambda_{2n}, \lambda_{2n+1})$ are referred to as the $n$-th spectral gap and $\gamma_n := \lambda_{2n+1} - \lambda_{2n}$ as the $n$-th gap length. Note that $|\Delta(\lambda)| > 2$ on the spectral gaps. We say that the $n$-th gap is open if $\gamma_n > 0$ and collapsed otherwise. The set of elements $(b, a) \in \mathcal{M}$ for which the $n$-th gap is collapsed is denoted by $D_n,$

$$D_n := \{(b, a) \in \mathcal{M} : \gamma_n = 0\}. \quad (21)$$

By writing the condition $\gamma_n = 0$ as $\gamma_n^2 = 0$ and exploiting the fact that $\gamma_n^2$ (unlike $\gamma_n$) is a real analytic function on $\mathcal{M}$, it can be shown as in [12] that $D_n$ is a real analytic submanifold of $\mathcal{M}$ of codimension 2.

**Isolating neighborhoods:** Let $(b, a) \in \mathcal{M}$ be given. The strict inequality $\lambda_{2n-1} < \lambda_{2n}$ guarantees the existence of a family of mutually disjoint open subsets $(U_n)_{1 \leq n \leq N-1}$ of $\mathbb{C}$ so that for any $1 \leq n \leq N - 1$, $U_n$ is a neighborhood of the closed interval $[\lambda_{2n}, \lambda_{2n+1}]$. Such a family of neighborhoods is referred to as a family of isolating neighborhoods (for $(b, a)$).

In the case where $(b, a) \in \mathcal{M}^C$, we list the eigenvalues $(\lambda_j)_{1 \leq j \leq 2N}$ in lexicographic order:

$$\lambda_1 < \lambda_2 < \lambda_3 < \ldots < \lambda_{2N}.$$  

We then extend the gap lengths $\gamma_n$ to all of $\mathcal{M}^C$ by

$$\gamma_n := \lambda_{2n+1} - \lambda_{2n} \quad (1 \leq n \leq N - 1)$$

and define

$$D_n^C := \{(b, a) \in \mathcal{W} : \gamma_n = 0\}. \quad (23)$$

In the sequel, we will omit the superscript and always write $D_n$ for $D_n^C$.

Similarly, we do this for the zeroes $(\hat{\lambda}_n)_{1 \leq n \leq N-1}$ of $\hat{\Delta}_\lambda$. The $\hat{\lambda}_i$'s and $\lambda_i$'s no longer depend continuously on $(b, a) \in \mathcal{M}^C$. However, if we choose a small enough complex neighborhood $\mathcal{W}$ of $\mathcal{M}$ in $\mathcal{M}^C$, then for any $(b, a) \in \mathcal{W}$ the closed intervals $G_n \subseteq \mathbb{C} (1 \leq n \leq N - 1)$ defined by

$$G_n := \{(1 - t)\lambda_{2n} + t\lambda_{2n+1} : 0 \leq t \leq 1\}$$  

(24)

are pairwise disjoint, and hence, as in the real case, there exists a family of isolating neighborhoods $(U_n)_{1 \leq n \leq N-1}$.

---

3The lexicographic ordering $a < b$ for complex numbers $a$ and $b$ is defined by

$$a < b \iff \begin{cases} \text{Re } a < \text{Re } b \\ \text{or} \\ \text{Re } a = \text{Re } b \text{ and } \text{Im } a \leq \text{Im } b. \end{cases}$$  

(22)
Lemma 2.3. There exists a neighborhood $W$ of $M$ in $M^C$ such that for any $(b,a) \in W$, there are neighborhoods $U_n$ of $G_n$ in $\mathbb{C}$ ($1 \leq n \leq N-1$) which are pairwise disjoint.

Remark 2.4. In the sequel, we will have to shrink the complex neighborhood $W$ several times, but continue to denote it by the same letter.

- **Contours $\Gamma_n$:** For any $(b,a) \in W$ and any $1 \leq n \leq N-1$, we denote by $\Gamma_n$ a circuit in $U_n$ around $G_n$ with counterclockwise orientation.

- **Isospectral set:** For $(b,a) \in M$, the set $\text{Iso}(b,a)$ of all elements $(b',a') \in M$ so that $Q(b',a')$ has the same spectrum as $Q(b,a)$ is described with the help of the Dirichlet eigenvalues $\mu_1 < \mu_2 < \ldots < \mu_{N-1}$ of (11) defined by

$$y_1(N + 1, \mu_n) = 0.$$  \hspace{1cm} (25)

They coincide with the eigenvalues of the $(N-1) \times (N-1)$-matrix $L_2 = L_2(b,a)$ given by

$$
\begin{pmatrix}
    b_2 & a_2 & 0 & \ldots & 0 \\
    a_2 & \ddots & \ddots & \ddots & \vdots \\
    0 & \ddots & \ddots & \ddots & 0 \\
    \vdots & \ddots & \ddots & \ddots & a_{N-1} \\
    0 & \ldots & 0 & a_{N-1} & b_N
\end{pmatrix}.
$$

In the sequel, we will also refer to $\mu_1, \ldots, \mu_{N-1}$ as the Dirichlet eigenvalues of $L(b,a)$. Evaluating the Wronskian identity (13) at $\lambda = \mu_n$ one sees that $\mu_n$ lies in the closure of the $n$-th spectral gap. More precisely, substituting $y_1(N + 1, \mu_n) = 0$ in the identity (13) with $\lambda = \mu_n$ yields

$$y_1(N, \mu_n)y_2(N + 1, \mu_n) = 1.$$  \hspace{1cm} (26)

Hence the value of the discriminant at $\mu_n$ is given by

$$\Delta(\mu_n) = y_2(N + 1, \mu_n) + \frac{1}{y_2(N + 1, \mu_n)}$$  \hspace{1cm} (27)

and $|\Delta(\mu_n)| \geq 2$. By Lemma 2.6 below, given the point $(b,a)$ with $b_1 = \ldots = b_N = \beta$ and $a_1 = \ldots = a_N = \alpha$, one has $\lambda_{2n} = \lambda_{2n+1}$ and hence $\mu_n = \lambda_{2n}$ for any $1 \leq n \leq N - 1$. It then follows from a straightforward deformation argument that $\lambda_{2n} \leq \mu_n \leq \lambda_{2n+1}$ everywhere in the real space $M$.

Conversely, according to van Moerbeke [17], given any (real) Jacobi matrix $Q$ with spectrum $\lambda_1 < \lambda_2 \leq \lambda_3 < \lambda_4 \leq \lambda_5 < \ldots \lambda_{2N-2} \leq \lambda_{2N-1} < \lambda_{2N}$ and any sequence $(\mu_n)_{1 \leq n \leq N-1}$ with $\lambda_{2n} \leq \mu_n \leq \lambda_{2n+1}$ for $n = 1, \ldots, N-1$, there are exactly $2^r$ $N$-periodic Jacobi matrices $Q$ with spectrum $(\lambda_n)_{1 \leq n \leq 2N}$ and Dirichlet spectrum $(\mu_n)_{1 \leq n \leq N-1}$, where $r$ is the number of $n$'s with $\lambda_{2n} < \mu_n < \lambda_{2n+1}$.

In the case where $(b,a) \in M^C$, we continue to define the Dirichlet eigenvalues $(\mu_n)_{1 \leq n \leq N-1}$ by (25), and we list them in lexicographic ordering $\mu_1 < \mu_2 <$...
... ≺ μ_{N-1}. Then the μᵢ’s no longer depend continuously on (b, a) ∈ M^C. However, if we choose the complex neighborhood W of M in M^C of Lemma 2.3 small enough, then for any (b, a) ∈ W and 1 ≤ n ≤ N − 1, μᵢ is contained in the neighborhood Uᵢ of Gᵢ (but not necessarily in Gᵢ itself).

**Riemann surface Σ_{b,a}:** Denote by Σ_{b,a} the Riemann surface obtained as the compactification of the affine curve C_{b,a} defined by

\[ \{ (λ, z) ∈ C^2 : z^2 = Δ^2(λ, b, a) - 4 \}. \] (28)

Note that C_{b,a} and Σ_{b,a} are spectral invariants. (Strictly speaking, Σ_{b,a} is a Riemann surface only if the spectrum of Q(b, a) is simple - see e.g. Appendix A in [13] for details in this case. If the spectrum of Q(b, a) is not simple, Σ(b, a) becomes a Riemann surface after doubling the multiple eigenvalues - see e.g. section 2 of [13].)

**Dirichlet divisors:** To the Dirichlet eigenvalue μᵢ (1 ≤ n ≤ N − 1) we associate the point μᵢ on the surface Σ_{b,a},

\[ μᵢ := (μᵢ, √{Δ^2ᵢ - 4}) \text{ with } √{Δ^2ᵢ - 4} := y₁(N, μᵢ) - y₂(N + 1, μᵢ), \] (29)

where we used that, in view of (26) and the Wronskian identity (14),

\[ Δ^2ᵢ - 4 = (y₁(N, μᵢ) - y₂(N + 1, μᵢ))^2. \]

**Standard root:** The standard root or s-root for short, √1 − λ^2, is defined for λ ∈ C \ [−1, 1] by

\[ √1 - λ^2 := iλ \sqrt{1 - λ^{-2}}. \] (30)

More generally, we define for λ ∈ C \ {ta + (1 − t)b | 0 ≤ t ≤ 1} the s-root of a radicand of the form (b − λ)(λ − a) with a < b, a ≠ b by

\[ √(b - λ)(λ - a) := γ \sqrt{1 - w^2}, \] (31)

where γ := b − a, τ := b + a 2 and w := a − τ 2.

**Canonical sheet and canonical root:** For (b, a) ∈ M the canonical sheet of Σ_{b,a} is given by the set of points (λ, √{Δ^2 − 4}) in C_{b,a}, where the c-root √Δ^2 − 4 is defined on C \ ∪₀≤n≤N λ₂n, λ₂n₊₁ (where λ₀ := −∞ and λ₂N₊₁ := ∞) and determined by the sign condition

\[ −i \sqrt{Δ^2 - 4} > 0 \text{ for } λ₂N₋₁ < λ < λ₂N. \] (32)

As a consequence one has for any 1 ≤ n ≤ N

\[ \text{sign } √{Δ^2 − 4} = (-1)^{N+n-1} \text{ for } λ₁n < λ < λ₂n₊₁. \] (33)

The definition of the canonical sheet and the c-root can be extended to the neighborhood W of M in M^C of Lemma 2.3.
The $s$-root and the $e$-root will be used together in the following way: By the product representations (20) and (18) of $\Delta_\lambda$ and $\Delta_\lambda^2 - 4$, respectively, one sees that for any $(b, a)$ in $W \setminus D_n$ with $1 \leq n \leq N - 1$, 
\[
\frac{\Delta_\lambda}{\sqrt{\Delta_\lambda^2 - 4}} = \frac{N(\lambda - \hat{\lambda}_n)}{\sqrt{(\lambda_{2n+1} - \lambda)(\lambda - \lambda_{2n})}} \chi_n(\lambda) \quad \forall \lambda \in \Gamma_n
\] 
where 
\[
\chi_n(\lambda) = \frac{(-1)^{N+n-1}}{\sqrt{(\lambda - \lambda_1)(\lambda_{2N} - \lambda)}} \prod_{m \neq n} \frac{\lambda - \hat{\lambda}_m}{\sqrt{(\lambda - \lambda_{2m+1})(\lambda - \lambda_{2m})}}.
\] 
Note that the principal branches of the square roots in (35) are well defined for $\lambda$ near $G_n$ and that the function $\chi_n$ is analytic and nonvanishing on $U_n$. In addition, for $(b, a)$ real, $\chi_n$ is nonnegative on the interval $(\lambda_{2n}, \lambda_{2n+1})$.

**Abelian differentials:** Let $(b, a) \in M$ and $1 \leq n \leq N - 1$. Then there exists a unique polynomial $\psi_n(\lambda)$ of degree at most $N - 2$ such that for any $1 \leq k \leq N - 1$
\[
\frac{1}{2\pi} \int_{c_k} \frac{\psi_n(\lambda)}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda = \delta_{kn}.
\] 
Here, for any $1 \leq k \leq N - 1$, $c_k$ denotes the lift of the contour $\Gamma_k$ to the canonical sheet of $\Sigma_{b,a}$. For any $k \neq n$ with $\lambda_{2k} \neq \lambda_{2k+1}$, it follows from (36) that
\[
\frac{1}{\pi} \int_{\lambda_{2k}}^{\lambda_{2k+1}} \frac{\psi_n(\lambda)}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda = 0.
\] 
Hence in every gap $(\lambda_{2k}, \lambda_{2k+1})$ with $k \neq n$ the polynomial $\psi_n$ has a zero which we denote by $\sigma^n_k$. If $\lambda_{2k} = \lambda_{2k+1}$ then it follows from (36) and Cauchy’s theorem that $\sigma^n_k = \lambda_{2k} = \lambda_{2k+1}$. As $\psi_n(\lambda)$ is a polynomial of degree at most $N - 2$, one has
\[
\psi_n(\lambda) = M_n \prod_{\substack{1 \leq k \leq N - 1 \\k \neq n}} (\lambda - \sigma^n_k),
\] 
where $M_n \equiv M_n(b, a) \neq 0$.

In a straightforward way one can prove that there exists a neighborhood $W$ of $M$ in $M_C$, so that for any $(b, a) \in W$ and any $1 \leq n \leq N - 1$, there is a unique polynomial $\psi_n(\lambda)$ of degree at most $N - 2$ satisfying (36) for any $1 \leq k \leq N - 1$ as well as the product representation (38), and so that the zeroes are analytic functions on $W$.

**Lemma 2.5.** Let $1 \leq n \leq N - 1$ be fixed. Then the zeroes $(\hat{\lambda}_k)_{1 \leq k \leq N - 1}$ of $\hat{\Delta}(\lambda)$ and $(\sigma^n_k)_{1 \leq k \leq N - 1, k \neq n}$ of $\psi_n(\lambda)$ satisfy the estimates 
\[
\hat{\lambda}_k - \tau_k = O(\gamma_k^2),
\]
\[
\sigma^n_k - \tau_k = O(\gamma_k^2).
\] 
\(\text{near any given point } (b, a) \in W, \text{ where } \tau_k = \frac{1}{2}(\lambda_{2k+1} + \lambda_{2k}).\)
Proof. To verify (39), write $\Delta_n^2 - 4$ in the form

$$\Delta_n^2 - 4 = (\lambda - \lambda_{2n})(\lambda_{2n+1} - \lambda)p_n(\lambda)$$

(41)

where $p_n$ is a polynomial which does not vanish for $\lambda \in U_n$. Then (39) follows by differentiating (41) with respect to $\lambda$ at $\lambda_n$.

Fix $1 \leq k, n \leq N - 1$ with $k \neq n$. In a first step we prove that $\sigma_n^k - \tau_k = O(\gamma_k)$ near any given point $(b, a) \in W$. If $\gamma_k = 0$, then $\sigma_n^k = \tau_k$, and (40) is clearly fulfilled. Hence we assume in the sequel that $\gamma_k \neq 0$. By the product formulas (38) and (18) for $\psi_n(\lambda)$ and $\Delta_n^2 - 4$, respectively, we obtain, for $\lambda$ near $G_k$,

$$\frac{\psi_n(\lambda)}{\sqrt{\Delta_n^2 - 4}} = \frac{\lambda - \sigma_n^k}{\sqrt{(\lambda_{2k+1} - \lambda)(\lambda - \lambda_{2k})}} \zeta^n_k(\lambda)$$

(42)

where

$$\zeta^n_k(\lambda) = \frac{M_n'(b, a)}{(\lambda - \sigma_n^k) \sqrt{(\lambda - \lambda_1)(\lambda_{2N} - \lambda)}} \prod_{m \neq k} \frac{\lambda - \sigma_m^n}{\sqrt{(\lambda_{2m+1} - \lambda)(\lambda_{2m} - \lambda)}},$$

(43)

with $\sigma_n^k := \tau_n$ and $M_n'(b, a) \neq 0$. The function $\zeta^n_k$ is analytic and nonvanishing in $U_k$. Substituting (42) into (39) one gets

$$\frac{1}{2\pi} \int_{\Gamma_k} \frac{\lambda - \sigma_n^k}{\sqrt{(\lambda_{2k+1} - \lambda)(\lambda - \lambda_{2k})}} \zeta^n_k(\lambda) d\lambda = 0.$$  

(44)

We now drop the superscript $n$ for the remainder of this proof and write $\zeta_k$ as $\zeta_k(\lambda) = \xi_k + (\zeta_k(\lambda) - \xi_k)$ with $\xi_k := \zeta_k(\tau_k) \neq 0$. Note that

$$\frac{1}{2\pi} \int_{\Gamma_k} \frac{\lambda - \sigma_k}{\sqrt{(\lambda_{2k+1} - \lambda)(\lambda - \lambda_{2k})}} d\lambda = \tau_k - \sigma_k$$

and hence (44) becomes

$$(\sigma_k - \tau_k)\xi_k = \frac{1}{2\pi} \int_{\Gamma_k} \frac{(\lambda - \sigma_k)(\xi_k(\lambda) - \xi_k)}{\sqrt{(\lambda_{2k+1} - \lambda)(\lambda - \lambda_{2k})}} d\lambda.$$  

(45)

To estimate the integral on the right hand side of (45), note that

$$\left|\frac{1}{2\pi} \int_{\Gamma_k} \frac{f(\lambda)}{\sqrt{(\lambda_{2k+1} - \lambda)(\lambda - \lambda_{2k})}} d\lambda\right| \leq \max_{\lambda \in G_k} |f(\lambda)|$$

(46)

for an arbitrary function $f$ analytic on $U_k$. We want to apply (46) for $f(\lambda) = (\lambda - \sigma_k)(\zeta_k(\lambda) - \xi_k)$. Note that for $\lambda \in G_k$,

$$|\zeta_k(\lambda) - \xi_k| = |\zeta_k(\lambda) - \zeta_k(\tau_k)| \leq M|\gamma_k|,$$

where $M = \sup_{1 \leq k \leq N - 1} \{|\zeta_k(\lambda)| : \lambda \in G_k\}$. Hence (46) leads to

$$|\sigma_k - \tau_k||\xi_k| = \sup_{\lambda \in G_k} |\lambda - \sigma_k| O(\gamma_k).$$
Dividing by $|\xi_k| \neq 0$, we get
\[ |\sigma_k - \tau_k| = \sup_{\lambda \in G_k} |\lambda - \sigma_k| O(\gamma_k) \] (47)
and in particular $|\sigma_k - \tau_k| = O(\gamma_k)$.

In a second step, we now improve the estimate (47). Note that
\[ \sup_{\lambda \in G_k} |\lambda - \sigma_k| \leq |\sigma_k - \tau_k| + \sup_{\lambda \in G_k} |\lambda - \tau_k| = O(\gamma_k) \] (48).
Combining (47) and (48), we obtain the claimed estimate (40).

For later use, we compute the spectra of $Q(b,a)$ and $L_2(b,a)$ in the special case $(b,a) = (\beta_1 N, \alpha_1 N)$ with $\beta \in \mathbb{R}$ and $\alpha > 0$. Here $1_N$ denotes the vector $(1, \ldots, 1) \in \mathbb{R}^N$. These points are the equilibrium points (of the restrictions) of the Toda Hamiltonian vector field (to the symplectic leaves $M_{\beta,\alpha}$). We compute the spectrum $(\lambda_j)_{1 \leq j \leq 2N}$ of the matrix $Q(\beta_1 N, \alpha_1 N)$ and the Dirichlet eigenvalues $(\mu_l)_{1 \leq l \leq N-1}$ of $L = L(\beta_1 N, \alpha_1 N)$ together with a normalized eigenvector $g_l = (g_l(j))_{1 \leq j \leq N}$ of $\mu_l$, i.e. $Lg_l = \mu_l g_l$, $g_l(1) = 0$, and a vector $h_l = (h_l(j))_{1 \leq j \leq N}$, which is the normalized solution of $L = \mu_l y$ orthogonal to $g_l$ satisfying $W(h_l, g_l)(N) > 0$.

**Lemma 2.6.** The spectrum $(\lambda_j)_{1 \leq j \leq 2N}$ of $Q(\beta_1 N, \alpha_1 N)$ and the Dirichlet eigenvalues $(\mu_l)_{1 \leq l \leq N-1}$ of $L(\beta_1 N, \alpha_1 N)$ are given by
\[
\lambda_1 = \beta - 2\alpha, \quad \lambda_{2l} = \lambda_{2l+1} = \mu_l = \beta - 2\alpha \cos \frac{l\pi}{N} \quad (1 \leq l \leq N-1),
\]
\[
\lambda_{2N} = \beta + 2\alpha.
\]

In particular, all spectral gaps of $Q(\beta_1 N, \alpha_1 N)$ are collapsed. For any $1 \leq l \leq N - 1$, the vectors $g_l$ and $h_l$ defined by
\[
g_l(j) = (-1)^{j+1} \sqrt{\frac{2}{N}} \sin \left(\frac{(j-1)l\pi}{N}\right) \quad (1 \leq j \leq N),
\]
\[
h_l(j) = (-1)^j \sqrt{\frac{2}{N}} \cos \left(\frac{(j-1)l\pi}{N}\right) \quad (1 \leq j \leq N)
\]
satisfy $L = \mu_l y$ and the normalization conditions
\[
\sum_{j=1}^{N} g_l(j)^2 = \sum_{j=1}^{N} h_l(j)^2 = 1, \quad g_l(0) > 0, \quad g_l(1) = 0;
\]
\[
W(h_l, g_l)(N) > 0, \quad \langle h_l, g_l \rangle_{\mathbb{R}^N} = 0.
\]
Remark 2.7. For \((b,a) = (\beta_1 N, \alpha_1 N)\) the fundamental solutions \(y_1\) and \(y_2\) are given by
\[
y_1(j, \lambda) = -\frac{\sin(\rho(j-1))}{\sin \rho} \quad (j \in \mathbb{Z}) \tag{51}
\]
\[
y_2(j, \lambda) = \frac{\sin(\rho j)}{\sin \rho} \quad (j \in \mathbb{Z}) \tag{52}
\]
where \(\pi < \rho < 2\pi\) is determined by \(\cos \rho = \frac{\lambda - \beta}{2\alpha}\).

Proof. For any \(\lambda \in \mathbb{R}\), the difference equation (11) for \((\beta_1 N, \alpha_1 N)\) reads
\[
(R_{\beta, \alpha} y)(k) := \beta y(k) + \alpha y(k-1) + \alpha y(k+1) = \lambda y(k) \tag{53}
\]
and can be written as
\[
y(k-1) + y(k+1) = \frac{\lambda - \beta}{\alpha} y(k). \tag{54}
\]
Since we are looking for periodic solutions of (54), we make the ansatz \(y(k) = e^{\pm i \rho k}\). This leads to the characteristic equation
\[
2 \cos \rho \equiv e^{i \rho} + e^{-i \rho} = \frac{\lambda - \beta}{\alpha}. \tag{55}
\]
For the solution to be \(2N\)-periodic, it is required that \(\rho \in \frac{\pi}{N} \mathbb{Z}\). To put the eigenvalues in ascending order, introduce \(\rho_l = (1 + \frac{l}{N}) \pi\) with \(0 \leq l \leq N\). Then for any \(1 \leq j \leq 2N\), there exists \(0 \leq l \leq N\) such that
\[
\lambda_j = \beta - 2\alpha \cos \rho_l = \beta - 2\alpha \cos \frac{l \pi}{N}.
\]
Note that for \(l = 0\), \(\lambda_1 = \beta - 2\alpha\) is an eigenvalue of \(Q(\beta_1 N, \alpha_1 N)\) with eigenvector \(y(k) = e^{i \pi k} = (-1)^k\). Similarly, for \(l = N\), \(\lambda_{2N} = \beta + 2\alpha\) is an eigenvalue with eigenvector \(y(k) \equiv 1\). For the eigenvalue \(\lambda_{2l} = \beta - 2\alpha \cos \frac{l \pi}{N}\) \((1 \leq l \leq N - 1)\),
\[
y_{\pm}(k) = e^{\pm i \rho_l k}
\]
are two linearly independent eigenvectors. As there are \(2N\) eigenvalues altogether, \(\lambda_{2l}\) is double for any \(1 \leq l \leq N - 1\), and \(\lambda_1\) and \(\lambda_{2N}\) are both simple. It follows that all \(N-1\) gaps are collapsed and hence \(\mu_l = \lambda_{2l}\) for all \(1 \leq l \leq N-1\).

Turning to the computation of \(g_k\) and \(h_k\), one easily verifies that for any real number \(\lambda \neq \pm 2\alpha + \beta\), the fundamental solution \(y_1(\cdot, \lambda)\) of (54) with \(y_1(0, \lambda) = 1\) and \(y_1(1, \lambda) = 0\) is given by
\[
y_1(j, \lambda) = -\sin \left(\frac{\rho(j-1)}{\sin \rho}\right) \quad (j \in \mathbb{Z})
\]
where \(\pi < \rho < 2\pi\) is determined by \(\cos \rho = \frac{\lambda - \beta}{2\alpha}\), thus proving (51). In the same way, one verifies (52). For \(\lambda = \mu_1 = \beta - 2\alpha \cos \frac{l \pi}{N}\) we then get
\[
\sin \left(\rho_l(j-1)\right) = \sin \left((1 + \frac{l}{N})\pi(j-1)\right) = (-1)^{j+1} \sin \left(\frac{(j-1)l \pi}{N}\right).
In particular, \( \sin(\rho_l(j-1)) = 0 \) for \( j = 1 \) and \( j = N+1 \). As

\[
\sum_{j=1}^{N} \sin^2 \left( \frac{(j-1)l\pi}{N} \right) = \sum_{j=1}^{N} \cos^2 \left( \frac{(j-1)l\pi}{N} \right)
\]

and these two sums add up to \( N \), one sees that

\[
\sum_{j=1}^{N} \sin^2 \left( \frac{(j-1)l\pi}{N} \right) = \frac{N}{2},
\]

yielding the claimed formula \((55)\) for \( g_l \).

By the same argument one shows that \( \tilde{h}_l \) given by \((-1)^j \sqrt{2} \cos \frac{(j-1)l\pi}{N} \) (i.e. the right side of \((50)\)) satisfies \( R_{\beta,\alpha} \tilde{h}_l = \mu_l \tilde{h}_l \) and the normalization condition \( \sum_{j=1}^{N} \tilde{h}_l(j)^2 = 1 \). Using standard trigonometric identities one verifies that

\[
\langle g_l, \tilde{h}_l \rangle = \sum_{j=1}^{N} g_l(j) \tilde{h}_l(j) = 0
\]

and \( W(\tilde{h}_l, g_l)(N) \) can be computed to be

\[
\tilde{h}_l(N)g_l(N+1) - \tilde{h}_l(N+1)g_l(N) = -\tilde{h}_l(N+1)g_l(N) = -\tilde{h}_l(1)g_l(0) > 0.
\]

Hence \( \tilde{h}_l \) is indeed the eigenvector with the required normalization, i.e. \( h_l = \tilde{h}_l \), thus proving \((50)\).

3 Action Variables

In the next two sections, we define the candidates for action-angle variables on the phase space \( M \) and investigate some of their properties. In this section we introduce globally defined action variables \( (I_n)_{1 \leq n \leq N-1} \) as proposed by Flaschka-McLaughlin \([4]\).

**Definition 3.1.** Let \( (b, a) \in M \). For \( 1 \leq n \leq N-1 \),

\[
I_n := \frac{1}{2\pi} \int_{\Gamma_n} \lambda \frac{\Delta_{\lambda}}{\sqrt{\Delta_{\lambda}^2 - 4}} \, d\lambda
\]

where \( \Delta_{\lambda} = \frac{d}{d\lambda} \Delta_{\lambda} \) is the \( \lambda \)-derivative of the discriminant \( \Delta_{\lambda} = \Delta(\lambda, b, a) \) and the contour \( \Gamma_n \) and the canonical root \( \sqrt{\cdot} \) are given as in section \([3]\).

**Remark 3.2.** The contours \( \Gamma_n \) can be chosen locally independently of \( (b, a) \). In view of the fact that \( \Delta_{\lambda} \) is a spectral invariant of \( L(b, a) \) the actions \( I_n \) are entirely determined by the spectrum of \( L(b, a) \). In particular, \( (I_n)_{1 \leq n \leq N-1} \) are constants of motion, since by Proposition \([2,7]\) the Toda flow is isospectral.
Remark 3.3. The variables \((I_n)_{1 \leq n \leq N-1}\) can also be represented as integrals on the Riemann surface \(\Sigma_{b,a}\). If \(c_n\) denotes the lift of \(\Gamma_n\) to the canonical sheet of \(\Sigma_{b,a}\), (57) becomes

\[
I_n = \frac{1}{2\pi} \int_{c_n} \lambda \frac{\dot{\Delta}_\lambda}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda \quad (1 \leq n \leq N-1).
\]

From the definition (56), the following result can be deduced:

**Proposition 3.4.** On the real space \(\mathcal{M}\), each function \(I_n\) is real, nonnegative, and it vanishes if \(\gamma_n = 0\).

**Proof.** Since

\[
\int_{\Gamma_n} \frac{\dot{\Delta}_\lambda}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda = 0,
\]

it follows that

\[
I_n = \frac{1}{2\pi} \int_{\Gamma_n} (\lambda - \lambda_n) \frac{\dot{\Delta}_\lambda}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda.
\]

By shrinking the contour of integration to the real interval, we get

\[
I_n = \frac{1}{\pi} \int_{\lambda_2n}^{\lambda_{2n+1}} (-1)^{N+n-1}(\lambda - \lambda_n) \frac{\dot{\Delta}_\lambda}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda,
\]

by taking into account the definition (52) of the \(c\)-root. Since \(\text{sign}(\lambda - \lambda_n)\dot{\Delta}_\lambda = (-1)^{N+n-1}\) on \([\lambda_{2n}, \lambda_{2n+1}] \setminus \{\lambda_n\}\), the integrand is real and nonnegative, hence \(I_n\) is real and nonnegative on \(\mathcal{M}\), as claimed.

If \(\gamma_n = 0\), then \(\lambda_{2n} = \lambda_{2n+1}\). Hence \(\dot{\lambda}_n = \lambda_{2n+1} = \tau_n\) and

\[
\lambda - \dot{\lambda}_n = i \sqrt{(\lambda_{2n+1} - \lambda)(\lambda - \lambda_{2n})}.
\]

Therefore the integrand in (58) is holomorphic in the interior of the contour \(\Gamma_n\), and by Cauchy’s theorem the integral in (58) vanishes.

The action variables \((I_n)_{1 \leq n \leq N-1}\) can be extended in a straightforward way to a complex neighborhood \(\mathcal{W}\) of \(\mathcal{M}\) in \(\mathcal{M}^C\).

**Theorem 3.5.** There exists a complex neighborhood \(\mathcal{W}\) of \(\mathcal{M}\) in \(\mathcal{M}^C\) such that for all \(1 \leq n \leq N-1\), the functions \(I_n\) defined by (56) extend analytically to \(\mathcal{W}\), \(I_n : \mathcal{W} \to \mathbb{C}\).

**Proof.** Let \(\mathcal{W}\) denote a neighborhood of \(\mathcal{M}\) in \(\mathcal{M}^C\) of Lemma 2.3 and define for any \(1 \leq n \leq N-1\) the functions \(I_n\) on \(\mathcal{W}\) by the formula (57). Let \((b,a) \in \mathcal{W}\) be given. Then there exists a neighborhood \(\mathcal{W}_{b,a}\) of \((b,a)\) in \(\mathcal{W}\) so that the integration contours \(\Gamma_n\) in (57) can be chosen to be the same for any element in \(\mathcal{W}_{b,a}\) and \(\Delta_\lambda/\sqrt{\Delta_\lambda^2 - 4}\) is analytic on \(B_r(\Gamma_n) \times \mathcal{W}_{b,a}\), where \(B_r(\Gamma_n) := \{\lambda \in \mathbb{C} | \text{dist}(\lambda, \Gamma_n) < \epsilon\}\) is the \(\epsilon\)-neighborhood of \(\Gamma_n\) with \(\epsilon\) sufficiently small. This shows that \(I_n\) is analytic on \(\mathcal{W}\).
Proposition 3.6. There exists a complex neighborhood $W$ of $M$ in $\mathcal{M}^C$ such that for any $1 \leq n \leq N - 1$, the quotient $I_n/\gamma_n^2$ extends analytically from $\mathcal{M} \setminus D_n$ to all of $W$ and has strictly positive real part on $W$. As a consequence, $\xi_n = \sqrt{2I_n/\gamma_n^2}$ is an analytic and nonvanishing function on $W$, where $\sqrt{2}$ is the principal branch of the square root on $\mathbb{C} \setminus (-\infty, 0]$.

Proof. Let $W$ be the complex neighborhood of Theorem 3.5. Substituting (34) into (58) leads to the following identity on $W \setminus D_n$:

$$I_n = \frac{N}{2\pi} \int_{\Gamma_n} \frac{(\lambda - \lambda_n)^2}{\sqrt{(\lambda_{2n+1} - \lambda)(\lambda - \lambda_{2n})}} \chi_n(\lambda) d\lambda,$$

where $\chi_n$ is given by (35). Upon the substitution $\lambda(\zeta) = \tau_n + \frac{2\pi}{\gamma_n} \zeta$, with $\tau_n = \frac{1}{2}(\lambda_{2n} + \lambda_{2n+1})$ and $\delta_n = \frac{2(\lambda_n - \tau_n)}{\gamma_n}$, one then obtains

$$\frac{2I_n}{\gamma_n^2} = \frac{N}{4\pi} \int_{\Gamma_n'} \frac{(\zeta - \delta_n)^2}{\sqrt{1 - \zeta^2}} \chi_n(\tau_n + \frac{\gamma_n}{2}\zeta) d\zeta,$$

where $\Gamma_n'$ is the pullback of $\Gamma_n$ under the substitution $\lambda = \lambda(\zeta)$, i.e. a circuit in $\mathbb{C}$ around $[-1, 1]$. By (39), $\lambda_n - \tau_n = O(\gamma_n^2)$, and hence $\delta_n \to 0$ as $\gamma_n \to 0$. We conclude that

$$\lim_{\gamma_n \to 0} \frac{2I_n}{\gamma_n^2} = \frac{N}{4\pi} \int_{\Gamma_n} \chi_n(\tau_n) \frac{\zeta^2}{\sqrt{1 - \zeta^2}} d\zeta = \chi_n(\tau_n) \frac{N}{2\pi} \int_{-1}^{1} \frac{t^2}{\sqrt{1 - t^2}} dt = \frac{N}{4} \chi_n(\tau_n).$$

By defining $\frac{2I_n}{\gamma_n^2}$ by $\frac{N}{4\pi} \chi_n(\tau_n)$ on $W \cap D_n$, it follows that $\frac{2I_n}{\gamma_n^2}$ is a continuous function on all of $W$. This extended function is analytic on $W \setminus D_n$ as is its restriction to $W \cap D_n$. By Theorem A.6 in [12] it then follows that $\frac{2I_n}{\gamma_n^2}$ is analytic on all of $W$.

By Lemma 3.7 below, the quotient $I_n/\gamma_n^2$ can be bounded away from zero on $M$. $\frac{2I_n}{\gamma_n^2} \geq \frac{1}{\pi\gamma_n(\lambda_{2n} - \lambda_1)}$. By shrinking $W$, if necessary, it then follows that for any $1 \leq n \leq N - 1$, the real part of $I_n/\gamma_n^2$ is positive and never vanishes on $W$. Hence the principal branch of the square root of $2I_n/\gamma_n^2$ is well defined on $W$ and $\xi_n$ has the claimed properties.

To show that $\sqrt{2I_n/\gamma_n^2}$ is well defined on $W$, we used in the proof of Proposition 3.6 the following auxiliary result, which we prove in Appendix A.

Lemma 3.7. For any $(b, a) \in M$ and any $1 \leq n \leq N - 1$,

$$\gamma_n^2 \leq 3\pi(\lambda_{2n} - \lambda_1)I_n.$$

(60)

From the definition (56), Proposition 3.4 and the estimate (60) one obtains

Corollary 3.8. For any $(b, a) \in M$ and any $1 \leq n \leq N - 1$,

$$I_n = 0 \quad \text{if and only if} \quad \gamma_n = 0.$$
Actually, Lemma 3.7 can be improved. We finish this section with an a priori estimate of the gap lengths $\gamma_n$ in terms of the action variables and the value of the Casimir $C_2$ alone, which will be shown in Appendix B.

**Theorem 3.9.** For any $(b, a) \in M_{\beta, \alpha}$ with $\beta \in \mathbb{R}$, $\alpha > 0$ arbitrary,

$$\sum_{n=1}^{N-1} \gamma_n^2 \leq 12\pi^2 \alpha \left( \sum_{n=1}^{N-1} I_n \right) + 9\pi^2 (N-1) \left( \sum_{n=1}^{N-1} I_n \right)^2. \quad (61)$$

### 4 Angle Variables

In this section, we define and study the angle coordinates $(\theta_n)$ for any $1 \leq n \leq N - 1$. Each $\theta_n$ is defined mod $2\pi$ on $W \setminus D_n$, where $W$ is a complex neighborhood of $M$ in $\mathcal{M}^C$ as in Lemma 2.3 and $D_n$ is given by (23).

**Definition 4.1.** For any $1 \leq n \leq N - 1$, the function $\theta_n$ is defined for $(b, a) \in M \setminus D_n$ by

$$\theta_n := \eta_n + \sum_{n \neq k=1}^{N-1} \beta_k^n \text{ mod } 2\pi, \quad (62)$$

where for $k \neq n$,

$$\beta_k^n = \int_{\mu_k^{-}}^{\mu_k^{+}} \frac{\psi_n(\lambda)}{\sqrt{\Delta^2 - 4}} d\lambda, \quad \eta_n = \int_{\lambda_{2n}}^{\lambda_{2n}^+} \frac{\psi_n(\lambda)}{\sqrt{\Delta^2 - 4}} d\lambda \text{ (mod } 2\pi), \quad (63)$$

and where for $1 \leq k \leq N - 1$, $\mu_k^-$ is the Dirichlet divisor defined in (29), and $\lambda_{2n}$ is identified with the ramification point $(\lambda_{2k}, 0)$ on the Riemann surface $\Sigma_{b, a}$. The integration paths on $\Sigma_{b, a}$ in (63) are required to be admissible in the sense that their image under the projection $\pi : \Sigma_{b, a} \rightarrow \mathbb{C}$ on the first component stays inside the isolating neighborhoods $U_k$.

Note that, in view of the normalization conditions (30) of $\psi_n$, the above restriction of the paths of integration in (63) implies that $\eta_n$ and hence $\theta_n$ are well-defined mod $2\pi$.

**Theorem 4.2.** Let $W$ be the complex neighborhood of $M$ in $\mathcal{M}^C$ introduced in Lemma 2.3. Then for any $1 \leq n \leq N - 1$, the function $\theta_n : W \setminus D_n \rightarrow \mathbb{C} \text{ (mod } \pi)$ is analytic.

**Remark 4.3.** As the lexicographic ordering of the eigenvalues of $Q(b, a)$ is not continuous on $W$, it follows that $\eta_n$ and hence $\theta_n$ are only continuous mod $\pi$ on $W$.

**Proof of Theorem 4.2.** To see that $\theta_n : W \setminus D_n \rightarrow \mathbb{C} \text{ (mod } \pi)$ is analytic, define for any $1 \leq k \leq N - 1$ the set

$$E_k := \{(b, a) \in \mathcal{M}^C : \mu_k(b, a) \in \{\lambda_{2k}(b, a), \lambda_{2k+1}(b, a)\}\}. \quad (64)$$
Below, we show that for any $1 \leq k \leq N - 1$ with $k \neq n$, $\beta^n_k$ is analytic on $W \setminus (D_k \cup E_k)$, that its restrictions to $D_k \cap W$ and $E_k \cap W$ are weakly analytic and that it is continuous on $W$. Together with the fact that $E_k \cap W$ and $D_k \cap W$ are analytic subvarieties of $W$ it then follows that $\beta^n_k$ is analytic on $W$ - see Theorem A.6 in [12]. Similar results can be shown for $\beta^n_n = \eta_n \pmod{\pi}$ on $W \setminus D_n$, and one concludes that $\theta_n \pmod{\pi}$ is analytic on $W \setminus D_n$.

To prove that $\beta^n_k$, $k \neq n$, is analytic on $W \setminus (D_k \cup E_k)$, note that since $\lambda_{2k}$ is a simple eigenvalue on $W \setminus D_k$, it is analytic there. Furthermore, $\beta^n_k$ is an analytic function on the (sufficiently small) neighborhood $W$ of $M$ in $\mathcal{M}$. On $W \setminus (D_k \cup E_k)$ we can use the substitution $\lambda = \lambda_{2k} + z$ to get

$$\beta^n_k = \int_{\lambda_{2k}}^{\lambda_{2k+1}} \frac{\psi_n(\lambda)}{\sqrt{\Delta_2^{(2k+2)} - 4}} d\lambda = \int_0^{\mu_k^*} \frac{\psi_n(\lambda_{2k} + z)}{\sqrt{\Delta_2^2 / D(z)}} d\lambda,$$

where $D(z) = \Delta_2^{(2k+2)} - 4$ is analytic near $z = 0$ and $D(0) \neq 0$. Note that $D(z)$ does not vanish for $z$ on an admissible integration path not going through $\lambda_{2k+1}$. Such a path exists since $(b, a)$ is in the complement of $E_k$. Furthermore $\psi_n(\lambda_{2k} + z)$ and $D(z)$ are analytic in $z$ near such a path and depend analytically on $(b, a) \in W \setminus (D_k \cup E_k)$. Combining these arguments shows that $\beta^n_k$ is analytic on $W \setminus (D_k \cup E_k)$.

For $k \neq n$ with $\lambda_{2k} \neq \lambda_{2k+1}$ one has

$$\int_{\lambda_{2k}}^{\lambda_{2k+1}} \frac{\psi_n(\lambda)}{\sqrt{\Delta_2^{(2k+2)} - 4}} d\lambda = 0. \quad (64)$$

As $\sigma^n_k = \lambda_{2k}$ if $\lambda_{2k} = \lambda_{2k+1}$ one sees that (64) continues to hold for $(b, a) \in E_k \cap W$ with $\lambda_{2k} = \lambda_{2k+1}$ and we have $\beta^n_k |_{E_k \cap W} \equiv 0$. To prove the analyticity of $\beta^n_k |_{D_k \cap W}$ consider the representation $[\psi_n(\lambda)] / \sqrt{\Delta_2^{(2k+2)} - 4}$ of $\frac{\psi_n(\lambda)}{\sqrt{\Delta_2^{(2k+2)} - 4}}$. For $(b, a) \in D_k \cap W$, one has

$$\lambda_{2k} = \lambda_{2k+1} = \tau_k = \sigma^n_k,$$

which implies that the factor $\sqrt{\lambda - \sigma^n_k} / \sqrt{(\lambda_{2k+1} - \lambda)(\lambda - \lambda_{2k})}$ in (64) equals $\pm i$. Hence we can write

$$\beta^n_k = \int_{\lambda_{2k}}^{\mu_k^*} \frac{\psi_n(\lambda)}{\sqrt{\Delta_2^{(2k+2)} - 4}} d\lambda = \pm i \int_{\tau_k}^{\mu_k^*} \zeta^n_k(\lambda) d\lambda.$$

As $\mu_k$ is analytic on $W$, it then follows that $\beta^n_k |_{D_k \cap W}$ is analytic. To see that $\beta^n_k$ is continuous on $W$, one separately shows that $\lambda_{2k}$ is continuous at points in $W \setminus (D_k \cup E_k)$, $E_k \cap W \setminus D_k$, $D_k \cap W \setminus E_k$ and $D_k \cap E_k \cap W$, where for the proof of the continuity of $\beta^n_k$ at points in $D_k \cap E_k \cap W$ we use [12] and the estimate $\sigma^n_k - \tau_k = O(\gamma_k^2)$ of Lemma 2.5.

By (64), $\eta_n$ vanishes mod $\pi$ on $E_n \cap W \setminus D_n$. Arguing in a similar way as for $\beta^n_k$ one then concludes that $\eta_n \pmod{\pi}$ is analytic on $W \setminus D_n$.

\footnote{Let $E$ and $F$ be complex Banach spaces, and let $U \subset E$ be open. The map $f : U \to F$ is weakly analytic on $U$, if for each $u \in U$, $h \in E$ and $L \in F^*$, the function $z \mapsto Lf(u + zh)$ is analytic in some neighborhood of the origin in $C$.}
5 Gradients

In this section we establish formulas of the gradients of $I_n, \theta_n$ ($1 \leq n \leq N - 1$) on $M$ in terms of products of the fundamental solutions $y_1$ and $y_2$.

Consider the discriminant for a fixed value of $\lambda$ as a function on $M$,

$$\Delta_\lambda(b, a) = y_1(N) + y_2(N + 1).$$

Then $\Delta_\lambda$ is a real analytic function on $M$. To obtain a formula for the gradients of $y_1(N)$ and $y_2(N + 1)$ with respect to $b$, differentiate $R_{b,a}y_i = \lambda y_i$ with respect to $b$ in the direction $v \in \mathbb{R}^N$ to get

$$(R_{b,a} - \lambda) \langle \nabla_b y_i, v \rangle(k) = -v_k y_i(k). \quad (65)$$

Differentiating $R_{b,a}y_i = \lambda y_i$ with respect to $a$ in the direction $u \in \mathbb{R}^N$ leads to

$$(R_{b,a} - \lambda) \langle \nabla_a y_i, u \rangle(k) = -u_{k-1} y_i(k-1) - u_k y_i(k+1). \quad (66)$$

Taking the sum of (65) and (66) yields

$$(R_{b,a} - \lambda) \left( \langle \nabla_b y_i, v \rangle + \langle \nabla_a y_i, u \rangle \right)(k) = -(R_{v,u}y_i)(k) \quad (67)$$

which we can rewrite as

$$(R_{b,a} - \lambda) \langle \nabla_{b,a} y_i, (v,u) \rangle(k) = -(R_{v,u}y_i)(k), \quad (68)$$

where $\langle \cdot, \cdot \rangle$ in (68) now denotes the standard scalar product in $\mathbb{R}^{2N}$, whereas in (65), (66), and (67) it is the one in $\mathbb{R}^N$. The inhomogeneous Jacobi difference equation (68) for the sequence $\langle \nabla_{b,a} y_i, (v,u) \rangle(k)$ can be integrated using the discrete analogue of the method of the variation of constants used for inhomogeneous differential equations. As $\langle \nabla_{b,a} y_i, (v,u) \rangle(0) = \langle \nabla_{b,a} y_i, (v,u) \rangle(1) = 0$, one obtains in this way for $m \geq 1$

$$\langle \nabla_{b,a} y_i, (v,u) \rangle(m) = -\left( \frac{y_2(m)}{a_N} \sum_{k=1}^{m} y_1(k)(R_{v,u}y_i)(k) \right) - \left( \frac{-y_1(m)}{a_N} \sum_{k=1}^{m} y_2(k)(R_{v,u}y_i)(k) \right). \quad (69)$$

In the sequel, we will use (69) to derive various formulas for the gradients. The common feature among these formulas is that they involve products between the fundamental solutions $y_1$ and $y_2$ of (11). Whereas the gradients with respect to $b = (b_1, \ldots, b_N)$ involve products computed by componentwise multiplication, the gradients with respect to $a = (a_1, \ldots, a_N)$ involve products obtained by multiplying shifted components, reflecting the fact that the $b_j$ are the diagonal elements of the symmetric matrix $L(b,a)$, whereas the $a_j$ are the off-diagonal elements of $L(b,a)$. 
To simplify notation for the formulas in this section, we define for sequences \((v(j))_{j \in \mathbb{Z}}, (w(j))_{j \in \mathbb{Z}} \subseteq \mathbb{C}\) the \(N\)-vectors
\[
\begin{align*}
v \cdot w & := (v(j)w(j))_{1 \leq j \leq N}, \\
v \cdot Sw & := (v(j)w(j+1))_{1 \leq j \leq N},
\end{align*}
\] (70)
(71)
where \(S\) denotes the shift operator of order 1. Combining (70) and (71), we define the \(2N\)-vector
\[
v \cdot S w := (v \cdot w, v \cdot Sw + w \cdot Sv).
\] (72)
In case \(v = w\) we also use the shorter notation
\[
v^2 := v \cdot S v.
\] (73)
Written componentwise, \(v \cdot S w\) is the \(2N\)-vector
\[
(v \cdot S w)(j) = \begin{cases} v(j)w(j) & (1 \leq j \leq N) \\
v(j-N)w(j-N+1) + v(j-N+1)w(j-N) & (N < j \leq 2N) \end{cases}
\]
Proposition 5.1. For any \((b,a) \in \mathcal{M}\), the gradient \(\nabla_{b,a} \Delta_\lambda = (\nabla_b \Delta_\lambda, \nabla_a \Delta_\lambda)\) is given by
\[
\begin{align*}
-a_N \nabla_b \Delta_\lambda & = y_2(N) y_1 \cdot y_1 - y_1(N+1) y_2 \cdot y_2 + (y_2(N+1) - y_1(N)) y_1 \cdot y_2, \\
-a_N \nabla_a \Delta_\lambda & = 2 y_2(N) y_1 \cdot S y_1 - 2 y_1(N+1) y_2 \cdot S y_2 \\
& \quad + (y_2(N+1) - y_1(N))(y_1 \cdot S y_2 + y_2 \cdot S y_1)
\end{align*}
\] (74)
(75)
or in the notation introduced above
\[
\nabla_{b,a} \Delta_\lambda = - \frac{1}{a_N} \left( y_2(N) y_1^2 - y_1(N+1) y_2^2 + (y_2(N+1) - y_1(N)) y_1 \cdot S y_2 \right).
\] (76)
The gradients \(\nabla_b \Delta_\lambda\) and \(\nabla_a \Delta_\lambda\) admit the representations \(1 \leq m \leq N\)
\[
\begin{align*}
\frac{\partial \Delta_\lambda}{\partial b_m} & = - \frac{1}{a_m} y_2(N, \lambda, S^m b, S^m a), \\
\frac{\partial \Delta_\lambda}{\partial a_m} & = - \left( \frac{1}{a_m} y_2(N+1, \lambda, S^m b, S^m a) + \frac{1}{a_{m+1}} y_2(N, \lambda, S^{m+1} b, S^{m+1} a) \right).
\end{align*}
\] (77)
(78)
Proof. The claimed formula (76) follows from the definition of \(\Delta_\lambda\) and formula (61). Indeed, evaluate (61) for \(i = 1\) and \(m = N\) to get
\[
\begin{align*}
\langle \nabla_{b,a} y_1, (v,u) \rangle & \langle (v,u) \rangle = - \frac{y_2(N)}{a_N} \sum_{k=1}^N y_1(k) (u_{k-1}y_1(k-1) + v_k y_1(k) + u_k y_1(k+1)) \\
& \quad + \frac{y_1(N)}{a_N} \sum_{k=1}^N y_2(k) (u_{k-1}y_1(k-1) + v_k y_1(k) + u_k y_1(k+1)).
\end{align*}
\] (79)
In order to identify these two sums with $\langle y_1^2, (v, u) \rangle$ and $\langle y_1 \cdot y_2, (v, u) \rangle$, respectively, note that

$$\sum_{k=1}^{N} u_{k-1} y_1(k)y_1(k - 1) = \sum_{k=1}^{N} u_k y_1(k)y_1(k + 1) + u_N T_1$$

where

$$T_1 := y_1(0)y_1(1) - y_1(N)y_1(N + 1).$$

For the second sum in (79), we get an expression of the same type with a similar correction term

$$T_2 := y_1(0)y_2(1) - y_1(N)y_2(N + 1).$$

Taking into account the initial conditions of the fundamental solutions and the Wronskian identity (13), one sees that $y_2(N)T_1 - y_1(N)T_2$ vanishes. Hence we have the formula

$$\langle \nabla_{b,a} y_1, (v, u) \rangle(N) = -\frac{1}{a_N} \left( y_2(N)\langle y_1^2, (v, u) \rangle - y_1(N)\langle y_1 \cdot y_2, (v, u) \rangle \right). \quad (80)$$

Similarly, evaluating formula (59) for $i = 2$ and $m = N + 1$ leads to

$$\langle \nabla_{b,a} y_2, (v, u) \rangle(N+1) = -\frac{1}{a_N} \left( y_2(N+1)\langle y_1 \cdot y_2, (v, u) \rangle - y_1(N+1)\langle y_2^2, (v, u) \rangle \right). \quad (81)$$

Here we used that the value of the right side of (69) does not change when we omit the term for $k = m = N + 1$ in both sums.

It remains to prove the two formulas (77) and (78). We first note that

$$y_2(n, \lambda, S^m b, S^m a) = \frac{a_m}{a_N} \left( y_2(n+m, \lambda, b, a)y_1(m, \lambda, b, a) - y_1(n+m, \lambda, b, a)y_2(m, \lambda, b, a) \right), \quad (82)$$

since both sides of (82) are solutions of $R_{S^m b, S^m a} y = \lambda y$ (for fixed $m \in \mathbb{Z}$) with the same initial conditions at $n = 0, 1$. For $n = 1$ this follows from the Wronskian identity (13). Similarly, one shows that

$$y_1(N+m, \lambda) = y_1(N, \lambda)y_1(m, \lambda) + y_1(N+1, \lambda)y_2(m, \lambda), \quad (83)$$

$$y_2(N+m, \lambda) = y_2(N, \lambda)y_1(m, \lambda) + y_2(N+1, \lambda)y_2(m, \lambda). \quad (84)$$

for any $(b, a) \in \mathcal{M}$. Hence, suppressing the variable $\lambda$, we get

$$y_2(N, S^m b, S^m a) = \frac{a_m}{a_N} \left( (y_2(N)y_1(m) + y_2(N+1)y_2(m))y_1(m) - (y_1(N)y_1(m) + y_1(N+1)y_2(m))y_2(m) \right)$$

$$= \frac{a_m}{a_N} \left( y_2(N)y_1(m)^2 + (y_2(N+1) - y_1(N))y_1(m)y_2(m) - y_1(N+1)y_2(m)^2 \right).$$
By (74) this leads to

\[ y_2(N, S^m b, S^m a) = -a_m \frac{\partial \Delta \lambda}{\partial b_{m}} \]

and formula (77) is established. To prove (78), we first conclude from (82) that

\[ \frac{a_N}{a_{m+1}} y_2(N-1, S^{m+1} b, S^{m+1} a) = y_2(N + m, b, a)y_1(m + 1, b, a) 
- y_1(N + m, b, a)y_2(m + 1, b, a) \]  
(85)

and

\[ \frac{a_N}{a_m} y_2(N + 1, S^m b, S^m a) = y_2(N + m + 1, b, a)y_1(m, b, a) 
- y_1(N + m + 1, b, a)y_2(m, b, a). \]  
(86)

Now expand the right hand sides of (85) and (86) according to (83) and (84). By (74), this leads to

\[ \sum y \sum \lambda \text{ is established. To prove (78), we first conclude from (82) that} \]

By (74), the sum of (85) and (86) is

\[ -a_N \frac{\partial \Delta \lambda}{\partial y_{m}} \text{, thus proving (78).} \]

As a next step, we compute the gradients of the Dirichlet and periodic eigenvalues. In the following lemma, we consider the fundamental solution \( y_1(\cdot, \mu) \) as an \( N \)-vector \( y_1(j, \mu) \leq j \leq N \). Let \( \| y_1(\mu) \|^2 = \sum_{j=1}^{N} y_1(j, \mu)^2 \), and denote by \( \cdot \) the derivative with respect to \( \lambda \).

**Lemma 5.2.** If \( \mu \) is a Dirichlet eigenvalue of \( L(b, a) \), then

\[ a_N y_1(N, \mu) \hat{y}_1(N + 1, \mu) = \| y_1(\mu) \|^2 > 0. \]  
(87)

In particular, \( \hat{y}_1(N + 1, \mu) \neq 0 \), which implies that all Dirichlet eigenvalues are simple.

**Proof.** This follows from adding up the relations (16). \( \square \)

As the Dirichlet eigenvalues \( (\mu_n)_{1 \leq n \leq N-1} \) of \( L(b, a) \) coincide with the roots of \( y_1(N+1, \mu) \) and these roots are simple, they are real analytic on \( M \). Similarly, the eigenvalues \( \lambda_1 \) and \( \lambda_{2N} \) are real analytic on \( M \), whereas for any \( 1 \leq n \leq N - 1 \), \( \lambda_{2n} \) and \( \lambda_{2n+1} \) are real analytic on \( M \setminus D_n \). Note that for \( (b, a) \in M \setminus D_n \) and \( i \in \{2n, 2n + 1\} \), we have \( \Delta \lambda_i \neq 0 \) as \( \lambda_i \) is a simple eigenvalue.

**Proposition 5.3.** For any \( 1 \leq n \leq N - 1 \), the gradients of the periodic eigenvalues \( \lambda_i \) \( (i = 2n, 2n + 1) \) on \( M \setminus D_n \) and of the Dirichlet eigenvalues \( \mu_n \) on \( M \) are given by

\[ \nabla_{b,a} \lambda_i = -\frac{\nabla_{b,a} \Delta \lambda |_{\lambda=\lambda_i}}{\Delta \lambda_i} = f_i^2 \quad \text{and} \quad \nabla_{b,a} \mu_n = g_n^2, \]  
(88)

where we denote by \( f_i \) the eigenvector of \( L(b, a) \) associated to \( \lambda_i \), normalized by

\[ \sum_{j=1}^{N} f_i(j)^2 = 1 \quad \text{and} \quad (f_i(1), f_i(2)) \in (\mathbb{R}_{>0} \times \mathbb{R}) \cup (\{0\} \times \mathbb{R}_{>0}), \]

and where \( g_n = (g_n(j))_{1 \leq j \leq N} \) is the fundamental solution \( y_1(\cdot, \mu_n) \) normalized so that \( \sum_{j=1}^{N} g_n(j)^2 = 1 \).
**Proof.** We first show the second formula in (88). Differentiating \( y_1(N + 1, \mu_n) = 0 \) with respect to \((b, a)\), one obtains

\[
\nabla_{b,a} \mu_n = -\frac{\nabla_{b,a} y_1(N + 1, \lambda)|_{\lambda = \mu_n}}{y_1(N + 1, \mu_n)}.
\] (89)

Here we used that \( y_1(N + 1, \mu_n) \neq 0 \) by Lemma 5.2. To compute the gradient \( \nabla_{b,a} y_1(N + 1, \lambda)|_{\lambda = \mu_n} \), we evaluate (90) for \( i = 1 \) and \( m = N + 1 \). In view of \( y_1(N + 1, \mu_n) = 0 \) and taking into account (26), one then gets

\[
\nabla_{b,a} \mu_n = \frac{y_1^2(\mu_n)}{a_N y_1(N, \mu_n)y_1(N + 1, \mu_n)}.
\] (90)

The claimed formula \( \nabla_{b,a} \mu_n = y_1^2 \) now follows from Lemma 5.2. By differentiating \( \Delta_{\lambda_i} = \pm 2 \) with respect to \((b, a)\), one obtains \( \nabla_{b,a} \lambda_i = -\nabla_{b,a} \Delta_{\lambda_i}|_{\lambda = \lambda_i}/\Delta_{\lambda_i} \) in a similar fashion. To see that \( \nabla_{b,a} \lambda_i = f_i^2 \), differentiate \( R_{b,a} f_i = \lambda_i f_i \) with respect to \((b, a)\) in the direction \((v, u) \in \mathbb{R}^{2N}\),

\[
R_{b,a}(\nabla_{b,a} f_i, (v, u))(k) + (R_{v,u} f_i)(k) = \langle \nabla_{b,a} \lambda_i, (v, u) \rangle f_i(k) + \lambda_i \langle \nabla_{b,a} f_i, (v, u) \rangle(k),
\]

where \( \langle \cdot, \cdot \rangle \) denotes the standard scalar product in \( \mathbb{R}^{2N} \). Take the scalar product (in \( \mathbb{R}^N \)) of the above equation with \( f_i \). Now use that

\[
\langle \nabla_{b,a} f_i(v, u), R_{b,a} f_i \rangle_{\mathbb{R}^N} = \lambda_i \langle \nabla_{b,a} f_i(v, u), f_i \rangle_{\mathbb{R}^N},
\]

\[
\langle f_i, f_i \rangle_{\mathbb{R}^N} = 1,
\]

and

\[
\langle R_{v,u} f_i, f_i \rangle_{\mathbb{R}^N} = \langle f_i^2, (v, u) \rangle_{\mathbb{R}^{2N}},
\]

to conclude that \( \nabla_{b,a} \lambda_i = f_i^2 \) holds.

To compute the Poisson brackets involving angle variables we need to establish some additional auxiliary results. Recall from section 3 that for \( 1 \leq k, n \leq N - 1 \) with \( k \neq n \) and \((b, a) \in \mathcal{M}\), \( \beta_k^n \) is given by

\[
\beta_k^n = \int_{\lambda_{2k}}^{\lambda_{2k}^*} \frac{\psi_n(\lambda)}{\sqrt{\Delta_{\lambda} - 4}} \, d\lambda,
\] (91)

whereas

\[
\beta_k^n := \eta_n = \int_{\lambda_{2n}}^{\lambda_{2n}^*} \frac{\psi_n(\lambda)}{\sqrt{\Delta_{\lambda} - 4}} \, d\lambda \pmod{2\pi}.
\] (92)

By Theorem 4.2, the functions \( \beta_k^n \) with \( k \neq n \) are real analytic on \( \mathcal{M}\), whereas \( \beta_k^n \), when considered mod \( \pi \), is real analytic on \( \mathcal{M} \setminus D_n \).

**Proposition 5.4.** Let \( 1 \leq k \leq N - 1 \) and \((b, a) \in \mathcal{M}\). If \( \gamma_k > 0 \) and \( \lambda_{2k} = \mu_k \), then for any \( 1 \leq n \leq N - 1 \),

\[
\nabla_{b,a} \beta_k^n = \frac{\psi_n(\mu_k)}{a_N \Delta_{\mu_k}} g_k \cdot s^k h_k,
\]
where \( h_k \) denotes the solution of \( R_{b,a}y = \mu_k y \) orthogonal to \( g_k \), i.e.
\[
\sum_{j=1}^{N} g_k(j) h_k(j) = 0,
\]
satisfying the normalization condition \( W(h_k, g_k)(N) = 1 \).

**Proof.** We use a limiting procedure first introduced in [16] for the nonlinear Schrödinger equation and subsequently used for the KdV equation in [11], [12]. We approximate \((b, a) \in M \) with \( \lambda_{2k}(b, a) = \mu_k(b, a) < \lambda_{2k+1}(b, a) \) by \((b', a') \in \text{Iso}(b, a) \) satisfying \( \lambda_{2k}(b, a) < \mu_k(b', a') < \lambda_{2k+1}(b, a) \). For such \((b', a')\), using the substitution \( \lambda = \lambda_{2k} + z \) in the integral of (91), we obtain
\[
\beta_k^{b'}(b', a') = \int_{\lambda_{2k}}^{\mu_k} \Psi_{\mu_k}(\lambda) \lambda^{\Delta_{\mu_k}} \sqrt{\lambda} d\lambda = \int_{0}^{\mu_k - \lambda_{2k}} \Psi_{\mu_k}(\lambda + z) \lambda^{\Delta_{\mu_k}} \sqrt{\lambda} d\lambda,
\]
where \( D(z) \equiv D(\lambda_{2k}, z) = (\Delta^2(\lambda_{2k} + z) - 4)/z \). Taking the gradient, undoing the substitution, and recalling the definition \((94)\) of the starred square root then leads to
\[
\nabla_{b,a}^n \beta_k^{b'} = \frac{\Psi_{\mu_k}(\mu_k)}{\sqrt{\Delta_{\mu_k}^2 - 4}} (\nabla_{b,a}^n \mu_k - \nabla_{b,a}^n \lambda_{2k}) + E(b', a'),
\]
with the remainder term \( E(b', a') \) given by
\[
E(b', a') = \int_{0}^{\mu_k - \lambda_{2k}} \nabla_{b',a'} \left( \frac{\Psi_{\mu_k}(\lambda + z)}{\sqrt{\Delta_{\mu_k}^2(\lambda_{2k}, z)}} \right) \frac{dz}{\sqrt{z}}.
\]
As the gradient in the latter integral is a bounded function in \( z \) near \( z = 0 \), locally uniformly in \((b', a')\), it follows by the dominated convergence theorem that
\[
\lim_{(b', a') \to (b, a)} E(b', a') = 0.
\]
The gradient \( \nabla_{b,a}^n \beta_k^{b'} \) depends continuously on \((b, a) \in M \) hence we can conclude by \((94)\) that it can be written as
\[
\nabla_{b,a}^n \beta_k^{b'} = \lim_{(b', a') \to (b, a)} \frac{\Psi_{\mu_k}(\mu_k)}{\sqrt{\Delta_{\mu_k}^2 - 4}} (\nabla_{b',a'} \mu_k - \nabla_{b',a'} \lambda_{2k}) \cdot
\]
The gradient of both sides of the Wronskian identity \((13)\),
\[
y_1(N, \lambda)y_2(N + 1, \lambda) - y_1(N + 1, \lambda)y_2(N, \lambda) = 1,
\]
leads to
\[
y_1(N + 1)\nabla_{b,a}^n y_2(N) + y_2(N)\nabla_{b,a}^n y_1(N + 1) = y_2(N + 1)\nabla_{b,a}^n \Delta + (y_1(N) - y_2(N + 1))\nabla_{b,a}^n y_2(N + 1),
\]
The \( \lambda \)-derivative can then be computed to be
\[
y_1(N + 1)\dot{y}_2(N) + y_1(N + 1)y_2(N) = y_2(N + 1)\dot{\Delta} + (y_1(N) - y_2(N + 1))\dot{y}_2(N + 1).
\]
Using (96), (97), and \( y_1(N + 1, \mu_k) = 0 \), formula (89) for \( \nabla_{b,a}\mu_k \) leads to

\[
\nabla_{b,a}\mu_k = \left. \frac{y_2(N + 1) \nabla_{b,a} \Delta + (y_1(N) - y_2(N + 1)) \nabla_{b,a} y_2(N + 1)}{y_2(N + 1) \Delta + (y_1(N) - y_2(N + 1)) \dot{y}_2(N + 1)} \right|_{\mu_k}.
\]

Further by (88),

\[
\nabla_{b,a} \lambda_{2k} = -\left. \frac{\nabla_{b,a} \Delta}{\Delta} \right|_{\lambda = \lambda_{2k}}.
\]

Now substitute (98) and (99) into \((\nabla_{b',a',\mu_k} - \nabla_{b',a',\ldots} \lambda_{2k})\) and use that by (29),

\[
\sqrt{\Delta_{\mu_k}^2 - 4} = (y_1(N) - y_2(N + 1))|_{\mu_k}.
\]

We claim that

\[
\lim_{(b',a') \to (b,a)} \frac{\nabla_{b,a}\mu_k - \nabla_{b,a}\lambda_{2k}}{\sqrt{\Delta_{\mu_k}^2 - 4}} = \frac{y_2(N + 1) \nabla_{b,a} y_1(N) - \dot{y}_1(N) \nabla_{b,a} y_2(N + 1)}{\Delta \dot{y}_1(N + 1) y_2(N)} |_{\lambda_{2k}}.
\]

Indeed, to obtain (100) after the above mentioned substitutions, we split the fraction \( \frac{1}{\sqrt{\Delta_{\mu_k}^2 - 4}}(\nabla_{b',a',\mu_k} - \nabla_{b',a',\lambda_{2k}}) \) into two parts which are treated separately. In the first part we collect all terms in \( \frac{1}{\sqrt{\Delta_{\mu_k}^2 - 4}}(\nabla_{b',a',\mu_k} - \nabla_{b',a',\lambda_{2k}}) \) which contain \((y_1(N) - y_2(N + 1))|_{\mu_k}\) in the nominator,

\[
I(a', b') := \frac{-\Delta|\lambda_{2k} \cdot \nabla_{b',a'} y_2(N + 1)|_{\mu_k} + \nabla_{b',a'} \Delta|\lambda_{2k} \cdot \dot{y}_2(N + 1)|_{\mu_k}}{\Delta|\lambda_{2k} \cdot (y_2(N + 1) \Delta + (y_1(N) - y_2(N + 1)) \dot{y}_2(N + 1))|_{\mu_k}}.
\]

Using again (97) we then get

\[
\lim_{(b',a') \to (b,a)} I(b', a') = \frac{\dot{y}_2(N + 1) \nabla_{b,a} y_1(N) - \dot{y}_1(N) \nabla_{b,a} y_2(N + 1)}{\Delta \dot{y}_1(N + 1) y_2(N)} |_{\lambda_{2k}}.
\]

The second term is then given by

\[
II(b', a') = \frac{y_2(N + 1)|_{\mu_k} \cdot (\nabla_{b',a'} \Delta|\lambda_{2k} \cdot \dot{\Delta}|_{\mu_k} - \hat{\Delta}|\lambda_{2k} \cdot \nabla_{b',a'} \Delta|_{\mu_k})}{\Delta|\lambda_{2k} \cdot (y_1(N) - y_2(N + 1))|_{\mu_k} \cdot (\dot{y}_1(N + 1) y_2(N))|_{\mu_k}}.
\]

Note that the nominator of \( II(b', a') \) is of the order \( O(\mu_k - \lambda_{2k}) \). In view of (29), we have

\[
(y_1(N) - y_2(N + 1))|_{\mu_k} = O(\sqrt{\mu_k - \lambda_{2k}})
\]

whereas the other terms in the denominator of \( II(b', a') \) are bounded away from zero. Indeed, \( \lambda_{2k} \) being a simple eigenvalue for \((b, a)\) means \( \hat{\Delta}|_{\lambda_{2k}} \neq 0 \) for \((b', a')\) near \((b, a)\). Further, use a version of (97) in the case \( \lambda_{2k} = \mu_k \) to conclude that

\[
\dot{y}_1(N + 1) y_2(N) = y_2(N + 1) \dot{\Delta}_{\lambda_{2k}}.
\]

Hence \((y_1(N + 1) y_2(N))|_{\mu_k} \neq 0 \) for \((b', a')\) near \((b, a)\) and \( II(b', a') \) vanishes in the limit of \( \mu_k \to \lambda_{2k} \).
Substituting (80) and (81) into (100), we obtain
\[
\dot{y}_2(N+1)\nabla_{b,a} y_1(N) - \dot{y}_1(N)\nabla_{b,a} y_2(N+1) = - \frac{1}{a_N\Delta} y_1 \cdot s y_0
\]
with \(y_0 = \frac{\dot{y}_2(N+1)}{\dot{y}_1(N+1)} y_1 - y_2\). Hence
\[
\nabla_{b,a} \beta_n^k = - \frac{\psi_n(\mu_k)}{a_N \Delta(\mu_k)} y_1 \cdot s y_0.
\]
Since \(\beta_n^k\) is invariant under the translation \(b \mapsto b + t(1, \ldots, 1)\), the scalar product
\[
\langle \nabla_{b,a} \beta_n^k, (1, 0) \rangle_{\mathbb{R}^{2N}}
\]
vanishes. Hence
\[
0 = \sum_{j=1}^{N} \frac{\partial \beta_n^k}{\partial b_j} = - \frac{\psi_n(\mu_k)}{a_N \Delta(\mu_k)} \sum_{j=1}^{N} y_1(j) y_0(j).
\]
It means that \(y_1\) and \(y_0\) are orthogonal to each other. Finally we introduce
\(h_k := \|y_1\|y_0\) and verify that
\[
W(h_k, g_k) = W\left(\|y_1\|y_0, \frac{y_1}{\|y_1\|}\right) = W(y_0, y_1) = W\left(\frac{\dot{y}_2(N+1)}{\dot{y}_1(N+1)} y_1 - y_2, y_1\right).
\]
By (13), it then follows that
\[
W(h_k, g_k) = -W(y_2, y_1) = W(y_1, y_2).
\]
Hence by (14),
\[
W(h_k, g_k)(N) = W(y_1, y_2)(N) = 1.
\]
This completes the proof of Proposition 5.4.

6 Orthogonality relations

In Propositions 5.1, 5.3, and 5.4 we have expressed the gradients of \(\Delta \lambda, \mu_n\), and, on a subset of \(\mathcal{M}\), of \(\beta_n^k\) in terms of products of fundamental solutions of the difference equation (11). In this section we establish orthogonality relations between such products - see [2] for similar computations. Recall that in (72) we have introduced for arbitrary sequences \((v_j)_{j \in \mathbb{Z}}, (w_j)_{j \in \mathbb{Z}}\) the 2N-vector \(v \cdot s w\).

**Lemma 6.1.** For any \((b, a) \in \mathcal{M}\), let \(v_1, w_1\) and \(v_2, w_2\) be pairs of solutions of (11) for arbitrarily given real numbers \(\mu\) and \(\lambda\), respectively. Then
\[
\frac{2(\lambda - \mu)}{a_1 a_N} \langle v_1 \cdot s w_1, J(v_2 \cdot s w_2) \rangle = V + B,
\]
where
\[
V := (W_1 \cdot S W_2)^N_{i_0} + (S W_1 \cdot W_2)^N_{i_0}
\]
(102)
with $W_1$ and $W_2$ denoting the Wronskians $W_1 := W(v_1, w_2)$, $W_2 := W(w_1, v_2)$, and where $B$ is given by

$$B := \frac{(\lambda - \mu)}{\alpha_1} \left( (v_1 \cdot w_1)|_1^{N+1} (v_2 \cdot s w_2)(2N) - (v_2 \cdot w_2)|_1^{N+1} (v_1 \cdot s w_1)(2N) \right). \quad (103)$$

**Proof.** We prove (103) by a straightforward calculation, using the recurrence property (15) of the Wronskian sequences $W$. By the definition (13) of $J$ we can write

$$2 \langle v_1 \cdot s w_1, J(v_2 \cdot s w_2) \rangle = E_1 + B_1,$$

where

$$E_1 := \sum_{k=1}^{N} a_k [(v_1 \cdot w_1)(k)(v_2 \cdot s w_2)(N + k) - (v_1 \cdot s w_1)(k + 1)(v_2 \cdot s w_2)(N + k)]$$

and

$$B_1 := a_N \left((v_1 \cdot s w_1)(N + 1) - (v_1 \cdot s w_1)(1)\right)(v_2 \cdot s w_2)(2N) + a_N \left((v_2 \cdot s w_2)(1) - (v_2 \cdot s w_2)(N + 1)\right)(v_1 \cdot s w_1)(2N).$$

Let us first consider $E_1$. Calculating the products $v_j \cdot w_j$ according to (72), we obtain, after regrouping,

$$E_1 = \sum_{k=1}^{N} a_k [(v_2(k)w_1(k) + v_2(k + 1)w_1(k + 1))W_1(k) + (v_1(k)w_2(k) + v_1(k + 1)w_2(k + 1))W_2(k)] + B_2$$

with

$$B_2 := a_N \left(v_1(N + 1)w_1(N + 1) - (v_1 \cdot s w_1)(N + 1)\right)(v_2 \cdot s w_2)(2N) + a_N \left((v_2 \cdot s w_2)(N + 1) - v_2(N + 1)w_2(N + 1)\right)(v_1 \cdot s w_1)(2N).$$

Multiply $E_1$ by $(\lambda - \mu)$ and use the recurrence relation (15) to express $(\lambda - \mu)v_2(k)w_1(k)$, $(\lambda - \mu)v_2(k + 1)w_1(k + 1)$, $(\lambda - \mu)v_1(k)w_2(k)$, and $(\lambda - \mu)v_1(k + 1)w_2(k + 1)$ in terms of the Wronskians $W_1$ and $W_2$ to get

$$(\lambda - \mu)E_1 = \sum_{k=1}^{N} a_k a_{k+1} (W_1(k)W_2(k + 1) + W_1(k + 1)W_2(k))$$

$$- a_{k-1} a_k (W_1(k - 1)W_2(k) + W_1(k)W_2(k - 1))$$

$$+ (\lambda - \mu)B_2.$$ 

The sum on the right hand side of the latter identity is a telescoping sum and equals the term $a_1 a_n V$ with $V$ defined in (102). In a straightforward way one sees that $\frac{(\lambda - \mu)}{\alpha_1 a_n} (B_1 + B_2)$ equals the expression $B$ defined by (103), hence formula (103) is established. 

\(\square\)
Corollary 6.2. For any $\lambda, \mu \in \mathbb{C}$,
\[
\{\Delta_\lambda, \Delta_\mu\}_J = 0.
\] (104)

Proof. By the formula (76) for the gradient of $\Delta_\lambda$,
\[
\{\Delta_\lambda, \Delta_\mu\}_J = \langle \nabla_{b,a} \Delta_\lambda, J \nabla_{b,a} \Delta_\mu \rangle
\]
is a linear combination of terms of the form $\langle v_1 \cdot s w_1, J(v_2 \cdot s w_2) \rangle$ for pairs of fundamental solutions $v_1, w_1$ and $v_2, w_2$ of (11) for $\mu$ and $\lambda$, respectively. In view of (77) and (78), $\nabla_b \Delta_\lambda$ and $\nabla_a \Delta_\lambda$ are both $N$-periodic. In the case $\lambda \neq \mu$ we use Lemma 6.1 and note that the boundary terms (102) and (103) in Lemma 6.1 vanish, hence $\{\Delta_\lambda, \Delta_\mu\}_J = 0$. In the case $\lambda = \mu$ the identity (104) follows from the skew-symmetry of $\{\cdot, \cdot\}_J$.

Corollary 6.3. Let $1 \leq k \leq 2N$ and $\lambda \in \mathbb{C}$. On the open subset of $\mathcal{M}$ where $\lambda_k$ is a simple eigenvalue of $Q(b,a)$ one has
\[
\{\lambda_k, \Delta_\lambda\}_J = 0.
\]

Proof. Using formula (88) for $\nabla_{b,a} \lambda_k$, we conclude from Corollary 6.2 that
\[
\{\lambda_k, \Delta_\lambda\}_J = -\frac{1}{\Delta_{\lambda_k}} \{\Delta_\mu, \Delta_\lambda\}_J|_{\mu = \lambda_k} = 0.
\]

Corollary 6.4. Let $\mu_n$ be the $n$-th Dirichlet eigenvalue of $L(b,a)$ and $\lambda \neq \mu_n$ a real number. Then
\[
(\lambda - \mu_n)(y_1^2(\mu_n), Jy_2^2(\lambda)) = \left( a_N \frac{y_1(N + 1, \lambda)}{y_2(N + 1, \mu_n)} \right)^2
\]
\[
(\lambda - \mu_n)(y_1^2(\mu_n), Jy_1(\lambda) \cdot s y_2(\lambda)) = a_N^2 \frac{y_1(N + 1, \lambda)y_2(N + 1, \lambda)}{y_2(N + 1, \mu_n)^2}
\]
\[
(\lambda - \mu_n)(y_2^2(\mu_n), Jy_2^2(\lambda)) = a_N^2 \left( \frac{y_2(N + 1, \lambda)}{y_2(N + 1, \mu_n)} \right)^2 - 1
\]

Proof. The three stated identities follow from Lemma 6.1 using that $y_1(N + 1, \mu_n) = 0$, $y_1(2, \mu_n) = -a_N/a_1$, and, by the Wronskian identity (26), $y_1(N, \mu_n) \cdot y_2(N + 1, \mu_n) = 1$.

Corollary 6.5. Let $\mu_n$ be the $n$-th Dirichlet eigenvalue of $L(b,a)$ and $\lambda \neq \mu_n$ a real number. Then
\[
\{\mu_n, \Delta_\lambda\}_J = \frac{y_1(N + 1, \lambda)}{y_1(N + 1, \mu_n)} \sqrt{\Delta_{\mu_n}^2 - 4} \frac{\Delta_{\mu_n} - 4}{\lambda - \mu_n}.
\] (108)
Proof. By \((90)\), combined with \((26)\), we get

\[
\{\mu_n, \Delta_\lambda\} J = \frac{y_2(N + 1, \mu_n)}{a_N y_1(N + 1, \mu_n)} y_1^2(\mu_n), J \nabla_{b,a} \Delta_\lambda.
\]  \hspace{1cm} (109)

Substituting the formula \((76)\) for \(J \nabla_{b,a} \Delta_\lambda\) we obtain

\[
\langle y_1^2(\mu_n), J \nabla_{b,a} \Delta_\lambda \rangle = \frac{-1}{a_N} y_2(N, \lambda) \langle y_1^2(\mu_n), J y_1^2(\lambda) \rangle
\]
\[
- \frac{1}{a_N} (y_2(N + 1, \lambda) - y_1(N, \lambda)) \langle y_1^2(\mu_n), J y_1(\lambda) \rangle y_2(\lambda)
\]
\[
+ \frac{1}{a_N} y_1(N + 1, \lambda) \langle y_1^2(\mu_n), J y_2^2(\lambda) \rangle.
\]  \hspace{1cm} (110)

To evaluate the right side of \((110)\), we apply Corollary \(6.3\) and get

\[
\frac{\lambda - \mu_n}{a_N} y_1^2(\mu_n), J \nabla_{b,a} \Delta_\lambda = \frac{1}{a_1 y_2(N + 1, \mu_n)^2} \left( - y_2(N, \lambda) y_1(N + 1, \lambda)^2 \right)
\]
\[
- (y_2(N + 1, \lambda) - y_1(N, \lambda)) y_1(N + 1, \lambda) y_2(N + 1, \lambda)
\]
\[
+ y_1(N + 1, \lambda) y_2(N + 1, \lambda)^2 \right) - \frac{y_1(N + 1, \lambda)}{a_1}.
\]

Using the Wronskian identity \((14)\), the sum of the terms in the square bracket of the latter expression simplifies, and one obtains

\[
\frac{\lambda - \mu_n}{a_1 a_N} y_1^2(\mu_n), J \nabla_{b,a} \Delta_\lambda = \frac{y_1(N + 1, \lambda)}{a_1 y_2(N + 1, \mu_n)^2} \frac{y_1(N + 1, \lambda)}{a_1}
\]
\[
= \frac{y_1(N + 1, \lambda)}{a_1} (y_1(\lambda, \mu_n)^2 - 1), \hspace{1cm} (111)
\]

where for the latter equality we again used \((14)\). Substituting \((111)\) into \((109)\), we get

\[
\frac{\lambda - \mu_n}{a_1 a_N} \{\mu_n, \Delta_\lambda\} J = \frac{y_2(N + 1, \mu_n) y_1(N + 1, \lambda)}{a_1 a_N y_1(N + 1, \mu_n)} (y_1(\lambda, \mu_n)^2 - 1)
\]
\[
= \frac{y_1(N + 1, \lambda)}{a_1 a_N y_1(N + 1, \mu_n)} \sqrt{\Delta_{\mu_n}^2 - 4},
\]

where we used that, by the definition of the starred square root \((29)\),

\[
\sqrt{\Delta_{\mu_n}^2 - 4} = y_1(N, \mu_n) - y_2(N + 1, \mu_n) = y_2(N + 1, \mu_n)(y_1(\lambda, \mu_n)^2 - 1).
\]

This proves \((108)\). \(\square\)

**Proposition 6.6.** For any \(\lambda \in \mathbb{R}, 1 \leq n \leq N - 1, \) and \((b,a) \in M \setminus D_n,\)

\[
\{\theta_n, \Delta_\lambda\} J = \psi_n(\lambda).
\]  \hspace{1cm} (112)
Proof. Recall that \( \theta_n = \sum_{k=1}^{N-1} \beta_k^n \) (mod \( 2\pi \)) with \( \beta_k^n \) given by (91)-(92). To compute \( \{ \beta_k^n, \Delta \lambda \}_J \), we first consider the case where \((b, a) \notin \bigcup_{k=1}^{N-1} D_k \) and \( \lambda_{2k} < \mu_k < \lambda_{2k+1} \) for any \( 1 \leq k \leq N - 1 \). Then \( \lambda_{2k} \) and \( \mu_k \) are smooth near \((b, a)\) and, by Leibniz’s rule, we get

\[
\{ \beta_k^n, \Delta \lambda \}_J = \left( \int_{\lambda_{2k}}^{\mu_k} \frac{\psi_n(\mu)}{\sqrt{\Delta_\mu^2} - 4} \Delta \lambda \, d\mu \right)_{J} + \frac{\psi_n(\mu_k)}{\sqrt{\Delta_{\mu_k}^2} - 4} \{ \mu_k, \Delta \lambda \}_J - \frac{\psi_n(\lambda_{2k})}{\sqrt{\Delta_{\lambda_{2k}}^2} - 4} \{ \lambda_{2k}, \Delta \lambda \}_J. 
\]

By Corollary 6.3 \( \{ \lambda_{2k}, \Delta \lambda \}_J = 0 \). Moreover, as the gradient \( \nabla_{b,a} \sqrt{\Delta_{\lambda}^2 - 4} \) is orthogonal to \( T_{b,a} \) Iso \((b, a)\) and \( J \nabla_{b,a} \Delta \lambda \in T_{b,a} \) Iso \((b, a)\) it follows that the Poisson bracket \( \{ \psi_\alpha(\mu), \Delta \lambda \}_J \) vanishes for any \( \mu \) in the isolating neighborhood \( U_n \) of \( G_n \). Hence

\[
\{ \beta_k^n, \Delta \lambda \}_J = \frac{\psi_n(\mu_k)}{\sqrt{\Delta_{\mu_k}^2} - 4} \{ \mu_k, \Delta \lambda \}_J.
\]

By (108), we then obtain

\[
\{ \theta_n, \Delta \lambda \}_J = \sum_{k=1}^{N-1} \frac{\psi_n(\mu_k)}{y_1(N + 1, \mu_k)} \frac{y_1(N + 1, \lambda)}{\lambda - \mu_k} = \psi_n(\lambda),
\]

where for the latter equality we used that \( \sum_{k=1}^{N-1} \frac{\psi_n(\mu_k)}{y_1(N + 1, \mu_k)} \frac{y_1(N + 1, \lambda)}{\lambda - \mu_k} \) and \( \psi_n(\lambda) \) are both polynomials in \( \lambda \) of degree at most \( N - 2 \) which agree at the \( N - 1 \) points \( (\mu_k)_{1 \leq k \leq N - 1} \).

In the general case, where \((b, a) \in M \setminus D_n \) and the Dirichlet eigenvalues are arbitrary, \( \lambda_{2k} \leq \mu_k \leq \lambda_{2k+1} \) for any \( 1 \leq k \leq N - 1 \), the claimed result follows from the case treated above by continuity.

\[ \square \]

Proposition 6.7. Let \( 1 \leq n, m, k, l \leq N - 1 \) and let \((b, a) \in M \) with \( \lambda_{2i}(b, a) = \mu_i(b, a) \) for \( i = k, l \). Then

\( \{ \beta_k^n, \beta_l^m \}_J = 0. \)

Proof. In view of Proposition 6.4 this amounts to showing that the scalar product \( \langle (g_k h_k), (g_l h_l) \rangle \) vanishes. For \( k = l \), this follows from the skew-symmetry of the Poisson bracket, hence we can assume \( k \neq l \). We apply Lemma 6.1 with \( v_1 := g_k, w_1 := h_k, v_2 := h_l \) and \( w_2 := g_l \), which implies that \( W_1 = W(g_k, g_l) \) and \( W_2 = W(h_k, h_l) \). Since \( g_k(1), g_l(1), g_k(N + 1) \) and \( g_l(N + 1) \) all vanish, we conclude \( W_1(N) = W_1(0) = 0 \) and \( (SW_1)(N) = (SW_1)(0) = 0 \), hence the expressions \( V \) and \( E \), defined in (102) and (103), vanish. This proves the claim. \( \square \)
7 Canonical relations

In this section we complete the proof of Theorem 1.1 and Corollary 1.2. In particular we show that the variables \((I_n)_{1 \leq n \leq N-1}, (\theta_n)_{1 \leq n \leq N-1}\) satisfy the canonical relations stated in Theorem 1.1.

Using the results of the preceding sections, we can now compute the Poisson brackets among the action and angle variables introduced in section 3.

Theorem 7.1. The action-angle variables \((I_n)_{1 \leq n \leq N-1}\) and \((\theta_n)_{1 \leq n \leq N-1}\) satisfy the following canonical relations for \(1 \leq n, m \leq N-1\):

(i) on \(M\),
\[
\{I_n, I_m\}_J = 0;
\]

(ii) on \(M \setminus D_n\),
\[
\{\theta_n, I_m\}_J = -\{I_m, \theta_n\}_J = -\delta_{nm}.
\]

Proof. Recall that \(\frac{d}{dt} \arccosh(t) = (t^2 - 1)^{-\frac{1}{2}}\). Hence for any \((b,a) \in M\)
\[
I_n = \frac{1}{2\pi} \int_{\Gamma_n} \lambda \frac{d}{d\lambda} \arccosh \left(\frac{\Delta\lambda}{2}\right) d\lambda
\]
and therefore
\[
\nabla_{b,a} I_n = \frac{1}{2\pi} \int_{\Gamma_n} \lambda \frac{d}{d\lambda} \nabla_{b,a} \frac{\Delta\lambda}{\sqrt{\Delta^2 - 4}} d\lambda.
\]
Integrating by parts we get
\[
\nabla_{b,a} I_n = -\frac{1}{2\pi} \int_{\Gamma_n} \frac{\nabla_{b,a} \frac{\Delta\lambda}{\sqrt{\Delta^2 - 4}}}{\sqrt{\Delta^2 - 4}} d\lambda.
\]
As \(\{\Delta, \Delta_n\}_J = 0\) for all \(\lambda, \mu \in \mathbb{C}\) by Corollary 6.2 it follows that \(\{I_n, I_m\}_J = 0\) on \(M\) for any \(1 \leq n, m \leq N - 1\).

To prove (114), use (115) and then Proposition 6.6 to get
\[
\{\theta_n, I_m\}_J = -\frac{1}{2\pi} \int_{\Gamma_m} \frac{\theta_n(\Delta\lambda)}{\sqrt{\Delta^2 - 4}} d\lambda = -\frac{1}{2\pi} \int_{\Gamma_m} \frac{\psi_n(\lambda)}{\sqrt{\Delta^2 - 4}} d\lambda = -\delta_{nm},
\]
by the normalizing condition (36) of \(\psi_n\).

To prove that the angles \((\theta_n)_{1 \leq n \leq N-1}\) pairwise Poisson commute we need the following lemma. We denote by \(K = K(b,a)\) the index set of the open gaps, i.e.
\[
K(b,a) = \{1 \leq n \leq N - 1 : \gamma_n(b,a) > 0\}.
\]

Lemma 7.2. At every point \((b,a)\) in \(M\), the set of vectors

(i) \(\{(\nabla_{b,a} I_n)_{n \in K}, \nabla_{b,a} C_1, \nabla_{b,a} C_2\}\)

and
(ii) \((J \nabla_{b,a} I_n)_{n \in K}\)

are both linearly independent.

Proof. The claimed statements follow from the orthogonality relations stated in Theorem 7.3. Let \((b, a) \in \mathcal{M}\) and suppose that for some real coefficients \((r_n)_{n \in K} \subseteq \mathbb{R}\) and \(s_1, s_2 \in \mathbb{R}\) we have

\[\sum_{n \in K} r_n \nabla_{b,a} I_n + s_1 \nabla_{b,a} C_1 + s_2 \nabla C_2 = 0.\]

For any \(m \in K\), take the scalar product of this identity with \(J \nabla_{b,a} \theta_m\). Using that \(\{I_n, \theta_m\}_J = \delta_{nm}\) and that \(C_1\) and \(C_2\) are Casimir functions of \(\{\cdot, \cdot\}_J\) one obtains

\[0 = \sum_{n \in K} r_n \{I_n, \theta_m\}_J = \sum_{n \in K} r_n \delta_{nm} = r_m.\]

Thus \(r_m = 0\) for all \(m \in K\), and it follows that \(s_1 \nabla_{b,a} C_1 + s_2 \nabla_{b,a} C_2 = 0\). By \(\mathcal{S}\) and \(\mathcal{G}\), \(\nabla_{b,a} C_1\) and \(\nabla_{b,a} C_2\) are linearly independent, hence \(s_1 = s_2 = 0\). This shows (i). The proof of (i) also shows that (ii) holds.

Theorem 7.3. In addition to the canonical relations stated in Theorem 7.1, the angle variables \((\theta_n)_{1 \leq n \leq N-1}\) satisfy for any \(1 \leq n, m \leq N-1\) on \(\mathcal{M} \setminus (D_n \cup D_m)\)

\[\{\theta_n, \theta_m\}_J = 0.\] (116)

Proof. Let \(1 \leq n, m \leq N-1\). By continuity, it suffices to prove the identity (116) for \((b, a) \in \mathcal{M} \setminus (\bigcup_{i=1}^{N-1} D_i)\). Let \((b, a)\) be an arbitrary element in \(\mathcal{M} \setminus (\bigcup_{i=1}^{N-1} D_i)\). Recall that \(\text{Iso}(b, a)\) denotes the set of all elements \((b', a')\) in \(\mathcal{M}\) with \(\text{spec}(Q_{b', a'}) = \text{spec}(Q_{b, a})\).

\[\text{Iso}(b, a) = \{(b', a') \in \mathcal{M} : \Delta(\cdot, b', a') = \Delta(\cdot, b, a)\}.\]

Then \(\text{Iso}(b, a)\) is a torus contained in \(\mathcal{M} \setminus (\bigcup_{i=1}^{N-1} D_i)\), and as all eigenvalues of \(Q(b, a)\) are simple, its dimension is \(N-1\). By Lemma 7.2 at any point \((b', a') \in \text{Iso}(b, a)\), the vectors \((J \nabla_{b', a'} I_k)_{1 \leq k \leq N-1}\) are linearly independent. Using the formula (115) for the gradient of \(I_k\), one sees that, by Corollary 6.2, for any \(\mu \in \mathbb{R}, 1 \leq k \leq N-1\),

\[\langle \nabla_{b', a'} \Delta_{\mu}, J \nabla_{b', a'} I_k \rangle = -\frac{1}{2\pi} \int_{\Gamma_n} \frac{\{\Delta_{\mu}, \Delta_\lambda\}_J}{\sqrt{\Delta_\lambda^2 - 4}} d\lambda = 0.\]

Hence for any \((b', a') \in \text{Iso}(b, a)\),

\[(J \nabla_{b', a'} I_k)_{1 \leq k \leq N-1} \in T_{b', a'} \text{Iso}(b, a),\]

and therefore these vectors form a basis of \(T_{b', a'} \text{Iso}(b, a)\).

To prove the identity (116), we apply the Jacobi identity

\[\{F, \{G, H\}_J\}_J + \{G, \{H, F\}_J\}_J + \{H, \{F, G\}_J\}_J = 0\]
to the functions $I_k$, $\theta_n$ and $\theta_m$. Since by Theorem 7.1, $\{I_k, \theta_n\}_J = \delta_{kn}$, we obtain

$$\{I_k, \{\theta_n, \theta_m\}_J\}_J = 0 \quad \text{on } \mathcal{M} \setminus \bigcup_{l=1}^{N-1} D_l$$

for any $1 \leq k \leq N - 1$.

It then follows by the above considerations that $\nabla_{b',a'} \{\theta_n, \theta_m\}_J$ is orthogonal to $T_{b',a'} \text{Iso}(b,a)$ for all $(b',a') \in \text{Iso}(b,a)$, i.e. $\{\theta_n, \theta_m\}_J$ is constant on $\text{Iso}(b,a)$,

$$\{\theta_n, \theta_m\}_J(b',a') = \{\theta_n, \theta_m\}_J(b,a) \quad \forall (b',a') \in \text{Iso}(b,a).$$

By Theorem 2.1, there exists a unique element $(b',a') \in \text{Iso}(b,a)$ satisfying $\mu_k(b',a') = \lambda_{2k}(b,a)$ for all $1 \leq k \leq N - 1$. The claimed identity (116) then follows from Proposition 6.7.

**Proof of Theorem 7.7.** By Theorem 8.3 and Theorem 4.2, the action and angle variables introduced in Definitions 3.1 and 4.1 respectively, have the claimed analyticity properties. The canonical relations among these variables have been verified in Theorem 7.1 and Theorem 7.3 and the relations $\{C_i, I_n\}_J = 0$ (on $\mathcal{M}$) and $\{C_i, \theta_n\}_J = 0$ (on $\mathcal{M} \setminus D_n$) follow from the fact that $C_1$ and $C_2$ are Casimir functions. It remains to show that the actions Poisson commute with the Toda Hamiltonian. To this end note that that the Hamiltonian $H$ can be written as

$$H = \frac{1}{2} \sum_{i=1}^{N} b_i^2 + \sum_{n=1}^{N} a_n^2 = \frac{1}{2} \text{tr} (L(b,a)^2) = \frac{1}{2} \sum_{j=1}^{N} (\lambda_j^+)^2$$

where $(\lambda_j^+)_{1 \leq j \leq N}$ are the $N$ eigenvalues of $L(b,a)$. Recall that on the dense open subset $\mathcal{M} \setminus \bigcup_{k=1}^{N-1} D_k$ of $\mathcal{M}$, the $\lambda_i^+$'s $(1 \leq i \leq N)$ are simple eigenvalues and hence real analytic. It then follows by (115) that for any $1 \leq n \leq N - 1$,

$$\{H, I_n\}_J = \sum_{i=1}^{N} \lambda_i^+ \{\lambda_i^+, I_n\}_J = -\sum_{i=1}^{N} \frac{\lambda_i^+}{2\pi} \int_{\Gamma_n} \frac{\{\lambda_i^+, \Delta\lambda\}_J}{\sqrt{\Delta\lambda^2 - 4}} d\lambda = 0,$n

where for the latter identity we used Corollary 5.3. Hence for any $1 \leq n \leq N - 1$,

$$\{H, I_n\}_J = 0 \quad \text{on } \mathcal{M} \setminus \bigcup_{k=1}^{N-1} D_k.$$

By continuity it then follows that $\{H, I_n\}_J = 0$ everywhere on $\mathcal{M}$.

**Proof of Corollary 7.2.** Since for any $\beta \in \mathbb{R}$ and $\alpha > 0$ the symplectic leaf $\mathcal{M}_{\beta,\alpha}$ is a submanifold of $\mathcal{M}$ of dimension $2(N-1)$, there are at most $N - 1$ independent integrals in involution on $\mathcal{M}_{\beta,\alpha}$. For any given $(b,a) \in \mathcal{M}_{\beta,\alpha}$ let $\pi_{\beta,\alpha}$ denote the orthogonal projection $T_{b,a} \mathcal{M} \to T_{b,a} \mathcal{M}_{\beta,\alpha}$. Then the gradient of the restriction $I_n|_{\mathcal{M}_{\beta,\alpha}}$ of $I_n$ to $\mathcal{M}_{\beta,\alpha}$ $(1 \leq n \leq N - 1)$ is given by $\pi_{\beta,\alpha} \nabla_{b,a} I_n$. By Lemma 7.2 the vectors $(\pi_{\beta,\alpha} \nabla_{b,a} I_n)_{n \in K}$ are linearly independent. As $\mathcal{M}_{\beta,\alpha} \setminus \bigcup_{k=1}^{N-1} D_k$ is dense in $\mathcal{M}_{\beta,\alpha}$, it then follows that $(I_n|_{\mathcal{M}_{\beta,\alpha}})_{1 \leq n \leq N-1}$ are functionally independent.
Finally, as $C_1$ and $C_2$ are Casimir functions of $\{\cdot, \cdot\}_J$, it follows that for any $(b, a) \in \mathcal{M}_{\beta, \alpha}$

$$\{I_n|_{\mathcal{M}_{\beta, \alpha}}, I_m|_{\mathcal{M}_{\beta, \alpha}}\}_J(b, a) = \langle \pi_{\beta, \alpha} \nabla_{b,a} I_n, \pi_{\beta, \alpha} J \nabla_{b,a} I_m \rangle$$

$$= \langle \nabla_{b,a} I_n, J \nabla_{b,a} I_m \rangle = \{I_n, I_m\}_J = 0,$$

i.e. the restrictions $I_n|_{\mathcal{M}_{\beta, \alpha}}$ of $I_n$ $(1 \leq n \leq N - 1)$ are in involution. 

\[\square\]

### A Proof of Lemma 3.7

In this Appendix, we prove Lemma 3.7. It turns out that the proof in [1], p. 601-602 of the special case where the parameter $\alpha$ in (1) equals 1 can be adapted for arbitrary values.

**Proof of Lemma 3.7.** Let $(b, a)$ be an arbitrary element in $\mathcal{M}$ and $1 \leq n \leq N-1$. First note that $I_n = \frac{1}{\pi} \int_{\lambda_{2n}}^{\lambda_{2n+1}} \frac{\int |\Delta(\lambda)|}{\frac{\sqrt{\lambda_{2n+1} - \lambda}}{2}} d\lambda$ and use $\frac{d}{dt} \text{arcosh}(t) = \frac{1}{\sqrt{t^2 - 1}}$ to obtain

$$I_n = \frac{1}{\pi} \int_{\lambda_{2n}}^{\lambda_{2n+1}} \int_1^{\frac{|\Delta(\lambda)|}{2}} \frac{1}{\sqrt{t^2 - 1}} dt d\lambda.$$

Since the integrand of the inner integral is nonincreasing, we estimate it from below by its value at $|\Delta(\lambda)|/2$. This leads to

$$I_n \geq \frac{1}{\pi} \int_{\lambda_{2n}}^{\lambda_{2n+1}} \frac{|\Delta(\lambda)| - 2}{\sqrt{|\Delta(\lambda)| + 2}} d\lambda. \quad (117)$$

We will show below that for $\lambda_{2n} \leq \lambda \leq \lambda_{2n+1}$

$$\frac{|\Delta(\lambda)| - 2}{\sqrt{|\Delta(\lambda)| + 2}} \geq \frac{\sqrt{\lambda - \lambda_{2n}} \sqrt{\lambda_{2n+1} - \lambda}}{\lambda_{2n+1} - \lambda_{2n}}. \quad (118)$$

We then substitute (118) into the integral (117) and split the integration interval into two equal parts,

$$I_n \geq \frac{2}{\pi} \frac{1}{\lambda_{2n+1} - \lambda_{2n}} \int_{\lambda_{2n}}^{\lambda_{2n+1}} \sqrt{\lambda - \lambda_{2n}} \sqrt{\lambda_{2n+1} - \lambda} d\lambda,$$

where $\tau_n = (\lambda_{2n} + \lambda_{2n+1})/2$. For $\lambda_{2n} \leq \lambda \leq \tau_n$ we estimate the quantity $\lambda_{2n+1} - \lambda$ from below by $\gamma_n/2$, yielding

$$I_n \geq \frac{2}{\pi} \frac{1}{\lambda_{2n+1} - \lambda_{2n}} \int_{\lambda_{2n}}^{\tau_n} \frac{\gamma_n}{2} \sqrt{\lambda - \lambda_{2n}} d\lambda = \frac{1}{3\pi(\lambda_{2n+1} - \lambda_{2n})} \gamma_n^2.$$  

It remains to verify (118). Recall that $\lambda_{2n}$ and $\lambda_{2n+1}$ are either both periodic or both antiperiodic eigenvalues of $L$. If $\lambda_{2n}$ and $\lambda_{2n+1}$ are periodic eigenvalues, we have $\Delta(\lambda) \geq 2$ for $\lambda_{2n} \leq \lambda \leq \lambda_{2n+1}$, i.e. $|\Delta(\lambda)| = \Delta(\lambda)$. In order to make writing easier, let us assume that $N$ is even - the case where $N$ is odd is treated...
in the same way. Then by \( (19) \), \( \lambda_1 \) and \( \lambda_{2N} \) are periodic eigenvalues of \( L \) and thus for any \( \lambda_{2n} \leq \lambda \leq \lambda_{2n+1} \), the left side of \( (118) \) can be estimated from below by

\[
\sqrt{\Delta(\lambda)} - 2 \sqrt{\Delta(\lambda)} + 2 = \frac{\sqrt{\lambda - \lambda_{2n}} \sqrt{\lambda_{2n+1} - \lambda}}{\lambda_{2N} - \lambda_1} \cdot R,
\]

where

\[
R \equiv R(\lambda) = \frac{\sqrt{\lambda - \lambda_1} \cdots \sqrt{\lambda - \lambda_{2n-3}} \sqrt{\lambda_{2n+4} - \lambda} \cdots \sqrt{\lambda_{2N} - \lambda}}{\lambda - \lambda_{2n-2} - \lambda_{2n+2} - \lambda_2N - \lambda_1}.
\]

As each of the the fractions under the square root in \( (119) \) can be estimated from below by 1, for any \( \lambda_{2n} \leq \lambda \leq \lambda_{2n+1} \) it follows that \( R(\lambda) \geq 1 \) on \( [\lambda_{2n}, \lambda_{2n+1}] \), leading to the claimed estimate \( (118) \).

\[\Box\]

B Proof of Theorem 3.9

In this Appendix we prove Theorem 3.9 using estimates derived in [1]. Let \((b,a)\) be in \( M_{\beta,\alpha} \) with \( \beta \in \mathbb{R} \) and \( \alpha > 0 \) arbitrary.

To show Theorem 3.9 we need the following

**Proposition B.1.** For any \((b,a)\) in \( M_{\beta,\alpha} \) with \( \beta \in \mathbb{R} \), \( \alpha > 0 \) arbitrary and any \( 1 \leq n \leq N \),

\[
\lambda_{2n}(b,a) - \lambda_{2n-1}(b,a) \leq \frac{2\pi\alpha}{N}.
\]

Before proving Proposition B.1 we show how to use it to prove Theorem 3.9.

**Proof of Theorem 3.9.** We begin by adding up the inequalities \((60)\) and get

\[
\sum_{n=1}^{N-1} \gamma_n^2 \leq 3\pi (\lambda_{2N} - \lambda_1) \left( \sum_{n=1}^{N-1} I_n \right).
\]

Note that

\[
\lambda_{2N} - \lambda_1 = \sum_{n=1}^{N-1} \gamma_n + \sum_{n=1}^{N} (\lambda_{2n} - \lambda_{2n-1}).
\]

By the estimate of Proposition B.1 we get for any \((b,a)\) in \( M_{\beta,\alpha} \)

\[
\lambda_{2N} - \lambda_1 \leq 2\pi\alpha + \sum_{n=1}^{N-1} \gamma_n
\]

which we substitute into \((121)\) to yield

\[
\sum_{n=1}^{N-1} \gamma_n^2 \leq 6\pi^2 \left( \sum_{n=1}^{N-1} I_n \right) + 3\pi \left( \sum_{n=1}^{N-1} \gamma_n \right) \left( \sum_{n=1}^{N-1} I_n \right).
\]
Using the inequality
\[ 2ab \leq \epsilon^2 a^2 + \frac{1}{\epsilon^2} b^2 \quad (a, b \in \mathbb{R}, \epsilon > 0) \]
with \( a = \sum_{n=1}^{N-1} \gamma_n, b = \sum_{n=1}^{N-1} I_n, \) and \( \epsilon^2 = \frac{1}{3\pi(N-1)} \), one gets
\[
\sum_{n=1}^{N-1} \gamma_n^2 \leq 6\pi^2 \alpha \left( \sum_{n=1}^{N-1} I_n \right) + \frac{3\pi}{2} \left( \frac{1}{3\pi(N-1)} \left( \sum_{n=1}^{N-1} \gamma_n \right)^2 + 3\pi(N-1) \left( \sum_{n=1}^{N-1} I_n \right)^2 \right).
\]
As \( \left( \sum_{n=1}^{N-1} \gamma_n \right)^2 \leq (N-1) \left( \sum_{n=1}^{N-1} \gamma_n^2 \right) \), one then concludes that
\[
\frac{1}{2} \sum_{n=1}^{N-1} \gamma_n^2 \leq 6\pi^2 \alpha \left( \sum_{n=1}^{N-1} I_n \right) + \frac{9\pi^2}{2} (N-1) \left( \sum_{n=1}^{N-1} I_n \right)^2,
\]
which is the claimed estimate (61).

To prove Proposition B.1 we first need to make some preparations. Note that for an element of the form \((b,a) = (\beta_1 N, \alpha_1 N)\) one has, by Lemma 2.6,
\[
\lambda_{2n}(\beta_1 N, \alpha_1 N) - \lambda_{2n-1}(\beta_1 N, \alpha_1 N) = 2\alpha \left( \cos \left( \frac{(n-1)\pi}{N} \right) - \cos \frac{n\pi}{N} \right)
= 4\alpha \sin \frac{(2n-1)\pi}{2N} \sin \frac{\pi}{2N}
< \frac{2\pi\alpha}{N}.
\]
Hence to prove Proposition B.1 it suffices to show that for any \((b,a) \in \mathcal{M}_{\beta,\alpha}\) and any \(1 \leq n \leq N\)
\[
\lambda_{2n}(b,a) - \lambda_{2n-1}(b,a) \leq \lambda_{2n}(\beta_1 N, \alpha_1 N) - \lambda_{2n-1}(\beta_1 N, \alpha_1 N).
\]
(123)

To this end, following [15] (cf. also [7]), we introduce the conformal map
\[
\delta(\lambda) := (-1)^N \int_{\lambda_i}^{\lambda} \frac{\hat{\Delta}(\mu)}{\sqrt{4 - \Delta^2(\mu)}} \, d\mu,
\]
(124)
where the sign of the square root is chosen such that for \(\mu < \lambda_1, \sqrt{4 - \Delta^2(\mu)}\) has positive imaginary part. It is defined on the upper half plane \(U := \{ \text{Im} \, z > 0 \}\) and its image is the spike domain
\[
\Omega(b,a) := \{ x + iy : 0 < x < N\pi, y > 0 \} \setminus \bigcup_{n=1}^{N-1} T_n
\]
where for \(1 \leq n \leq N - 1\), \(T_n\) denotes the spike
\[
T_n := \left\{ n\pi + it : 0 < t \leq \text{arcosh} \left( (-1)^{N+n} \frac{\Delta(\lambda_n)}{2} \right) \right\}.
\]
To see that $\delta(U) = \Omega(b, a)$, note that for any $(b, a) \in M$ and $\lambda \in U$ the discriminant $\Delta(\lambda)$ and the function $\delta(\lambda)$ are related by the formula

$$\Delta(\lambda) = 2(-1)^N \cos \delta(\lambda). \quad (125)$$

To prove (125), recall that for $-1 < t < 1$, one has $\frac{d}{dt} \arccos t = \frac{-1}{\sqrt{1-t^2}}$. This formula remains valid for any $t$ in $\mathbb{C} \setminus ((-\infty, -1] \cup [1, \infty))$. Thus

$$\delta(\lambda) = (-1)^N \int_{\lambda_1}^{\lambda} \frac{\Delta(\mu)}{\sqrt{4 - \Delta^2(\mu)}} d\mu = \int_{\lambda_1}^{\lambda} \frac{d}{d\mu} \arccos \left( (-1)^N \frac{\Delta(\mu)}{2} \right).$$

Since by (19), $\Delta(\lambda_1) = 2(-1)^N$, we then get

$$\delta(\lambda) = \arccos \left( (-1)^N \frac{\Delta(\lambda)}{2} \right) \bigg|_{\lambda_1} = \arccos \left( (-1)^N \frac{\Delta(\lambda)}{2} \right), \quad (126)$$

leading to formula (125) and the claimed statement that $\delta(U) = \Omega(b, a)$.

The map $\delta$ can be extended continuously to the closure $\{ \text{Im} z \geq 0 \}$ of the upper half plane. This extension, again denoted by $\delta$, is 2-1 over each nontrivial spike $T_n$ and 1-1 otherwise. Since the $n$-th spike $T_n$ is the image under $\delta$ of the $n$-th gap $(\lambda_{2n}, \lambda_{2n+1})$, all spikes are empty iff all gaps are collapsed. By Lemma 2.6, all gaps are collapsed for $(b, a) = (-\beta_1 N, \alpha_1 N)$, hence $\Omega(b, a) \subset \Omega(-\beta_1 N, \alpha_1 N)$ for any $(b, a) \in M_{\beta, \alpha}$.

Note that

$$\lambda_{2n}(b, a) - \lambda_{2n-1}(b, a) = \delta^{-1}(n\pi-) - \delta^{-1}((n-1)\pi+) = \int_{-\infty}^{\infty} u^{(n)}(\delta(\lambda)) d\lambda, \quad (127)$$

where $u^{(n)} : \Omega(b, a) \to \mathbb{R}$, $z \mapsto u^{(n)}(z; b, a)$ is the harmonic measure of the open subset $(n-1)\pi, n\pi)$ of $\partial \Omega(b, a)$ (see e.g. [6] for the notion of the harmonic measure).

We need two lemmas from complex and harmonic analysis, respectively.

**Lemma B.2.** For $(b, a) = (-\beta_1 N, \alpha_1 N)$ with $\beta \in \mathbb{R}$, $\alpha > 0$ arbitrary, the map $\delta(\lambda)$ defined by (124) is given by

$$\delta(\lambda) = N \arccos \left( -\frac{\lambda + \beta}{2\alpha} \right). \quad (128)$$

For arbitrary $(b, a) \in M_{\beta, \alpha}$ and $\xi \in \mathbb{R}$, the following asymptotic estimate holds as $\eta \to \infty$

$$\delta(\xi + i\eta) = N \arccos \left( -\frac{\xi + i\eta + \beta}{2\alpha} \right) + O(\eta^{-2}), \quad (129)$$

locally uniformly in $\xi$.

**Proof of Lemma B.2.** In view of the formulas (51) and (52) for the fundamental solutions $y_1$ and $y_2$ for $(b, a) = (-\beta_1 N, \alpha_1 N)$, the discriminant $\Delta(\lambda) =$
\( \Delta(\lambda, -\beta_1 N, \alpha_1 N) \) is given by

\[
\Delta(\lambda) = y_1(N, \lambda) + y_2(N + 1, \lambda) = -\sin(\rho(N - 1)) + \sin(\rho(N + 1))\sin\rho = 2\cos(\rho N),
\]

where \( \pi < \rho < 2\pi \) is determined by \( \cos \rho = \frac{\lambda + \beta}{2\alpha} \). Hence

\[
\Delta(\lambda) = 2T_N \left( \frac{\lambda + \beta}{2\alpha} \right)
\]

(130)

where for any \( z \in U \),

\[
T_N(z) = \cos(N \arccos z).
\]

(131)

Actually, \( T_N(z) \) is a polynomial in \( z \) of degree \( N \), referred to as Chebychev polynomial of the first kind. Substituting (130) into (126), we obtain

\[
\delta(\lambda) = \arccos \left( (-1)^N \frac{\Delta(\lambda)}{2} \right) = \arccos \left( (-1)^N T_N \left( \frac{\lambda + \beta}{2\alpha} \right) \right).
\]

The claimed identity (128) now follows from the elementary symmetry

\[
T_N(z) = (-1)^N T_N(-z) \quad \forall z \in \mathbb{C}.
\]

Now let \( (b, a) \in M_{\beta, \alpha} \). The asymptotic estimate (129) follows by comparing the polynomials \( \Delta(\lambda) \) corresponding to \( (b, a) \) and the one corresponding to \( (-\beta_1 N, \alpha_1 N) \). By (17) (and the discussion following it), in both cases,

\[
\Delta(\lambda) = \alpha^{-N} \lambda^N \left( 1 + N \beta \lambda^{-1} + O(\lambda^{-2}) \right) \text{ as } |\lambda| \to \infty.
\]

This implies that

\[
\Delta_{b,a}(\lambda) = \Delta_{-\beta_1 N, \alpha_1 N}(\lambda) \cdot (1 + O(\lambda^{-2})),
\]

hence by (128) and (130)

\[
\delta_{b,a}(\lambda) = \arccos \left( (-1)^N \frac{\Delta_{b,a}(\lambda)}{2} \right) = \arccos \left( (-1)^N T_N \left( \frac{\lambda + \beta}{2\alpha} \right) \cdot (1 + O(\lambda^{-2})) \right).
\]

\[
= \arccos \left( (-1)^N T_N \left( \frac{\lambda + \beta}{2\alpha} \right) \cdot (1 + O(\lambda^{-2})) \right).
\]

Substituting formula (131) for \( T_N \), one then concludes that

\[
\delta_{b,a}(\lambda) = \arccos \left( \cos \left( N \arccos \left( -\frac{\lambda + \beta}{2\alpha} \right) \right) \cdot (1 + O(\lambda^{-2})) \right)
\]

\[
= N \arccos \left( -\frac{\lambda + \beta}{2\alpha} \right) + O(\lambda^{-2}), \quad (132)
\]
where in the last step we used that \( \arccos z = -i \log(z + i \sqrt{1 - z^2}) \) for any \( z \) in \( \mathbb{C} \setminus ((-\infty, -1] \cup [1, \infty)) \).

**Lemma B.3.** Let \( u : \Omega(b, a) \to \mathbb{R} \) be a bounded harmonic function such that the nontangential limit of \( u(z) \) on \( \partial \Omega(b, a) \) has compact support, and let \( U(\lambda) := u(\delta(\lambda)) \), where \( \delta(\lambda) \) is the function defined by (124). Then for almost every \( t \in \mathbb{R} \), the limit \( U(t) := \lim_{\eta \to 0} U(t + i\eta) \) exists and is integrable, and

\[
\int_{-\infty}^{\infty} U(t) dt = \lim_{x \to \infty} 2\pi \alpha \sinh \left( \frac{x}{N} \right) u \left( \frac{N\pi}{2} + ix \right). \quad (133)
\]

**Proof of Lemma B.3.** Again by Fatou’s theorem, for a.e. \( t \) the (nontangential) limit \( \lim_{\eta \to 0} U(t + i\eta) \) exists, since \( U \) is a bounded harmonic function on \( \{ \text{Im} \ z > 0 \} \). Since \( u \) is bounded on \( \Omega(b, a) \) and its nontangential limit to \( \partial \Omega \) has compact support, \( U(t) \) is bounded and compactly supported and thus in particular integrable. For \( \lambda = \xi + i\eta \), one then has the Poisson representation

\[
U(\xi + i\eta) = \frac{\eta}{\pi} \int_{-\infty}^{\infty} \frac{U(t)}{(t - \xi)^2 + \eta^2} dt. \quad (134)
\]

and by dominated convergence we conclude that

\[
\int_{-\infty}^{\infty} U(t) dt = \lim_{\eta \to \infty} \pi \eta U(\xi + i\eta). \quad (135)
\]

(In particular, the limit in the latter expression exists.) In order to compute the right hand side of (135), let \( \xi + i\eta \) be given by

\[
\xi + i\eta = \delta^{-1} \left( \frac{N\pi}{2} + ix \right),
\]

for \( x \) sufficiently large. Then \( U(\xi + i\eta) = u \left( \frac{N\pi}{2} + ix \right) \), and by (132), it follows that

\[
\frac{\pi}{2} + i x = \arccos \left( -\frac{(\xi + \beta) + i \eta}{2\alpha} \right) + O \left( (\xi + i\eta)^{-2} \right) \quad (x \to \infty). \quad (136)
\]

Taking the cosine of both sides of (136), multiplying by \(-2\alpha \) and using that \( \cos \left( \frac{\pi}{2} + it \right) = -i \sin t \) for \( t \in \mathbb{R} \), we obtain

\[
2i\alpha \sinh \frac{x}{N} = ((\xi + \beta) + i \eta) \left( 1 + O \left( (\xi + i\eta)^{-2} \right) \right) \quad (x \to \infty).
\]

Hence, as \( x \to \infty \),

\[
\xi = O(1), \quad \eta = 2\alpha \sinh \frac{x}{N} + O(1). \quad (137)
\]

Substituting (137) into (135) leads to the claimed formula (133). \( \square \)
Proof of Proposition B.1. Let \((b, a) \in \mathcal{M}_{\beta, \alpha}\). Besides the harmonic measure \(u^{(n)}\) of the set \(E := ((n-1)\pi, n\pi) \subset \partial \Omega(b, a)\) we also consider the harmonic measure \(u^{(n)}_{\beta, \alpha}\) of \(E \subset \partial \Omega(-\beta N_1, \alpha N_1)\); note that

\[
\Omega(-\beta N_1, \alpha N_1) = \{ x + iy | 0 < x < N\pi, y > 0 \}
\]

and hence \(\Omega(b, a) \subseteq \Omega(-\beta N_1, \alpha N_1)\). According to [6], both \(u^{(n)}\) and \(u^{(n)}_{\beta, \alpha}\) satisfy the hypotheses of Lemma B.3. Let us recall the monotonicity property of the harmonic measures \(u(z, E, \Omega)\) with respect to \(\Omega\) (see e.g. [6]): If \(\Omega_1 \subseteq \Omega_2\), \(E \subset \partial \Omega_1 \cap \partial \Omega_2\), and \(u(z, E, \Omega_i) (i = 1, 2)\) denotes the harmonic measure of \(E \subset \partial \Omega_i\), then for any \(z \in \Omega_1\), \(u(z, E, \Omega_1) \leq u(z, E, \Omega_2)\). Apply this general principle to \(\Omega_1 := \Omega(b, a)\) and \(\Omega_2 := \Omega(-\beta N_1, \alpha N_1)\) to get

\[
u^{(n)}(x) \leq u^{(n)}_{\beta, \alpha}(x).
\] (138)

Writing \(U^{(n)}(\lambda) := u^{(n)}(\delta(\lambda))\) as well as \(U^{(n)}_{\beta, \alpha}(\lambda) := u^{(n)}_{\beta, \alpha}(\delta(\lambda))\) and combining (127), (133), and (138), we conclude that

\[
\lambda_{2n}(b, a) - \lambda_{2n-1}(b, a) = \int_{-\infty}^{\infty} U^{(n)}(\lambda)d\lambda
\]

\[
= \lim_{x \to \infty} 2\pi \alpha \sinh \left( \frac{x}{N} \right) u^{(n)} \left( \frac{N\pi}{2} + ix \right)
\]

\[
\leq \lim_{x \to \infty} 2\pi \alpha \sinh \left( \frac{x}{N} \right) u^{(n)}_{\beta, \alpha} \left( \frac{N\pi}{2} + ix \right)
\]

\[
= \int_{-\infty}^{\infty} U^{(n)}_{\beta, \alpha}(\lambda)d\lambda
\]

\[
= \lambda_{2n}(-\beta N_1, \alpha N_1) - \lambda_{2n-1}(-\beta N_1, \alpha N_1).
\]

This completes the proof of the estimate (120) and therefore of Proposition B.1. \\

References


REFERENCES


Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
E-mail address: andreas.henrici@math.unizh.ch

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
E-mail address: thomas.kappeler@math.unizh.ch