# Lax Pairs for the Ablowitz-Ladik System via Orthogonal Polynomials on the Unit Circle

Thesis by

Irina Nenciu

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2005
(Defended May 9, 2005)

# Acknowledgements

I wish to express my deep gratitude to my advisor, Barry Simon, for the wealth of opportunities he has given me – the chance to write this thesis under his supervision being just one of many. His guidance, encouragement, and unflinching enthusiasm for mathematics made this work possible. I also wish to thank Percy Deift for suggesting the original problem from which all the questions answered in this thesis arose.

I would like to acknowledge all of those who contributed to my mathematical education. Most of all, I want to thank Rowan Killip for sharing so much of his time and knowledge with me. His friendship and support have been invaluable to me. My thanks go also to the Caltech math department for creating such a great atmosphere for doing research, and to the department staff, in particular Cherie Galvez, for all of their help.

Finally, and most of all, I want to thank my family. My parents have always been a tireless source of love and support; I dedicate this thesis to them.

# Abstract

We investigate the existence and properties of an integrable system related to orthogonal polynomials on the unit circle. We prove that the main evolution of the system is defocusing Ablowitz-Ladik (also known as the integrable discrete nonlinear Schrödinger equation). In particular, we give a new proof of complete integrability for this system.

Furthermore, we use the CMV and extended CMV matrices defined in the context of orthogonal polynomials on the unit circle by Cantero, Moral, and Velázquez, and Simon, respectively, to construct Lax pair representations for the Ablowitz-Ladik hierarchy in the periodic, finite, and infinite settings.

# Contents

# Chapter 1

# Introduction

The purpose of the work presented here is to introduce the connection between the theory of orthogonal polynomials on the unit circle and a classical integrable system, the defocusing Ablowitz-Ladik (AL) system. In particular, we shall use the former to introduce the full AL hierarchy and to obtain Lax pair representations for all the Hamiltonians of the hierarchy.

This is not the first instance when such a connection appeared and was used to recast a system in Lax pair form. Flaschka [13] proved complete integrability for the celebrated Toda lattice by recasting it as a Lax equation for Jacobi matrices. Later, van Moerbeke [36], following similar work of McKean and van Moerbeke [23] on Hill's equation, used Jacobi matrices to define the Toda hierarchy for the periodic Toda lattice and to find the corresponding Lax pairs.

But, as we shall explain in Section 3.1, Jacobi matrices can be viewed as a part of the theory of orthogonal polynomials on the real line. From this perspective, our results are complex analogues of the corresponding results on the real line, and the approach we use to prove our main result, the Lax pair formulation of the evolution equations for the AL hierarchy, is heavily indebted to van Moerbeke's ideas [36].

The role of complex analogue to the theory of orthogonal polynomials on the real line is naturally played by the theory of orthogonal polynomials on the unit circle. So the question that started this investigation was exactly the question of finding an analogous scheme to the one described above: Is there an integrable system related to orthogonal polynomials on the unit circle in the same way that Toda relates to orthogonal polynomials on the real line?

The answer is “yes”, and the results can be summarized as follows:

$$
\begin{array}{ccccc}
\text{Orthogonal} & & \text{Lax} & & \text{Integrable} \\
\text{polynomials} & & \text{operators} & & \text{systems} \\
& & & & \\
\text{OPRL} & \rightarrow & \text{Jacobi matrices} & \leftrightarrow & \text{Toda lattice} \\
\uparrow & & \uparrow & & \uparrow \\
\text{OPUC} & \rightarrow & \text{CMV matrices} & \leftrightarrow & \text{defocusing AL}
\end{array}
\tag{1.0.1}
$$

The arrows here represent conceptual connections. In the rest of the introduction we briefly introduce the different notions that appear in (1.0.1) and formalize the connections between them, while at the same time presenting the structure of the thesis.

As explained in Section 2.1, one of the most important impulses to the theory of integrable systems was the discovery of solitons. This in turn led to the development of the inverse scattering transform (IST) as an extremely powerful tool for solving nonlinear PDEs. While we do not wish to go into any detail concerning the theory of direct and inverse scattering (an extremely rich subject, still very much at the center of the field), let us just say here that the IST should be thought of as a nonlinear Fourier transform. In particular, it allows one to linearize the flow of the corresponding nonlinear PDE, and its existence can be used as a definition of complete integrability in the infinite-dimensional setting.

At the heart of this effort was the well-known KdV equation, followed closely by the cubic 1-dimensional NLS (focusing and defocusing). Almost simultaneously two discrete analogues also attracted a great deal of attention and interest. One is the Toda lattice, certainly the simplest and most studied differential-difference equation. In Section 2.3 we present those aspects of the very rich theory of the Toda lattice that we are most interested in.

Another differential-difference equation which emerged in the mid-’70s is the

Ablowitz-Ladik equation [1],[2] (also known as integrable discrete cubic NLS). It appeared in the general form

$$\begin{aligned} -i\dot{\alpha}_n - \big(\alpha_{n+1} - 2\alpha_n + \alpha_{n-1}\big) + \alpha_n\beta_n\big(\alpha_{n+1} + \alpha_{n-1}\big) &= 0 \\ -i\dot{\beta}_n + \big(\beta_{n+1} - 2\beta_n + \beta_{n-1}\big) - \alpha_n\beta_n\big(\beta_{n+1} + \beta_{n-1}\big) &= 0. \end{aligned}$$

In particular, taking

$$\beta_n = \bar{\alpha}_n$$

for all $n$, one gets the space discretization of NLS. It reads:

$$-i\dot{\alpha}_n = \rho_n^2(\alpha_{n+1} + \alpha_{n-1}) - 2\alpha_n, \tag{1.0.2}$$

where $\alpha = \{\alpha_n\} \subset \mathbb{D}$ is a sequence of complex numbers inside the unit disk and

$$\rho_n^2 = 1 - |\alpha_n|^2.$$

The analogy with the continuous NLS becomes transparent if we rewrite (1.0.2) as

$$-i\dot{\alpha}_n = \alpha_{n+1} - 2\alpha_n + \alpha_{n-1} - |\alpha_n|^2(\alpha_{n+1} + \alpha_{n-1}).$$

Here, and throughout the thesis, $\dot{f}$ will denote the time derivative of the function $f$.

Note that the condition that all the $\alpha$'s be inside the unit disk $\mathbb{D}$ is not unreasonable: If $\alpha_n(0) \in \mathbb{D}$ for all $n$, then this remains true for all time. Moreover, if, for a certain $N \in \mathbb{Z}$, we have $\alpha_N \in S^1$ at time $t = 0$, then $\alpha_N$ remains on the unit circle for all time.

Another important observation is that there are three types of boundary conditions that one can impose in (1.0.2):

- Periodic: $\alpha_{n+p} = \alpha_n$ for a fixed period $p \geq 1$ and all $n \in \mathbb{Z}$;
- Finite: $\alpha_{-1} = \alpha_N = -1$ for a fixed $N \geq 1$, and we are interested in the evolution of $\alpha_0, \ldots, \alpha_{N-1} \in \mathbb{D}$;

- Infinite: $\alpha_{-1} = -1$, and we investigate the evolution of $\{\alpha_j\}_{j\geq 0} \subset \mathbb{D}$.

We will study all of these cases, as they correspond to the same situations for orthogonal polynomials. The analogous conditions for the Toda lattice go under the names "periodic," "open," and "closed" Toda, respectively. Also, in the finite and infinite AL, we can choose any points on $S^1$ as boundary conditions. The value -1 is chosen by analogy with the theory of orthogonal polynomials on the unit circle.

Ablowitz and Ladik proved that (1.0.2) is completely integrable (in the infinite setting) by associating it with a discrete 2×2 scattering problem:

$$\begin{aligned} v_{1,n+1} &= zv_{1,n} - \sigma \bar{u}_n v_{2,n} \\ v_{2,n+1} &= z^{-1} v_{2,n} + u_n v_{1,n} \end{aligned}$$

with $\sigma = \pm 1$.

The AL system has been extensively studied over the past thirty years. Until very recently, most of the results were concerned with properties of the discrete system which are preserved in the continuous limit. More recently, algebro-geometric solutions were also studied. See [16] and the references therein.

At the other end of our scheme (1.0.1) are the two orthogonal polynomial theories: OPRL (for "orthogonal polynomials on the real line") and OPUC (for "orthogonal polynomials on the unit circle"). The basic results of these theories are sketched in Chapter 3, with more emphasis on OPUC, since these are the results we use.

For both OPRL and OPUC the beginning of the theory is the same: Consider a probability measure $\nu$ with the support contained in $\mathbb{R}$, or $\mu$, supported inside

$$S^1 = \{z \in \mathbb{C} \mid |z| = 1\},$$

respectively. Construct the orthonormal polynomials by applying the Gram-Schmidt procedure to the monomials $\{x^j\}_{j\geq 0}$, $x \in \mathbb{R}$ to obtain $\{p_j\}_{j\geq 0}$, or, correspondingly, to $\{z^j\}_{j\geq 0}$, $z \in S^1$ and get $\{\phi_j\}_{j\geq 0}$. The remarkable feature of these theories is that in both cases the orthonormal polynomials obey recurrence relations: (3.1.1) on $\mathbb{R}$

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_{n+1}p_n(x) + a_n p_{n-1}(x)$$

and (3.2.6) on $S^1$

$$\phi_{n+1}(z) = \frac{1}{\rho_n}\big[z\phi_n(z) - \bar{\alpha}_n\phi_n^*(z)\big],$$

where

$$\phi_n^*(z) = z^n\overline{\phi_n(1/\bar{z})}.$$

These relations provide the connection of the orthogonal polynomial theories to Jacobi and CMV matrices.

Jacobi matrices are real, symmetric, tri-diagonal matrices. Much of their study is driven by the fact that they represent a generalization of the 1-dimensional discrete Schrödinger operator. Indeed, note that, for a sequence $u = \{u(n)\}_{n\in\mathbb{Z}}$, the discrete Laplacian is given by

$$\begin{aligned}(\Delta u)(n) &= \big(u(n+1) - u(n)\big) - \big(u(n) - u(n-1)\big)\\ &= u(n+1) + u(n-1) - 2u(n).\end{aligned}$$

Therefore the discrete Schrödinger operator is

$$(Hu)(n) = u(n+1) + u(n-1) + V(n)u(n)$$

with $V = \{V(n)\}_{n\in\mathbb{Z}}$ being the potential (traditionally, the $-2u(n)$ term from the Laplacian is absorbed in the potential). In other words, if we write $H$ in the basis of delta functions $\delta_j = \{\delta_{j,n}\}_{n\in\mathbb{Z}}$, we obtain

$$H = \begin{bmatrix} \ddots & \ddots & & & \\ \ddots & V(1) & 1 & & \\ & 1 & V(2) & 1 & \\ & & 1 & V(3) & \ddots \\ & & & \ddots & \ddots \end{bmatrix}.$$

This is just a doubly-infinite Jacobi matrix

$$J = \begin{bmatrix} \ddots & \ddots & & & \\ \ddots & b_1 & a_1 & & \\ & a_1 & b_2 & a_2 & \\ & & a_2 & b_3 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}$$

with

$$a_n = 1 \quad \text{and} \quad b_n = V(n)$$

for all $n \in \mathbb{Z}$.

We are interested in Jacobi matrices from a different perspective. For us, they represent the link between the Toda lattice and orthogonal polynomials on the real line. Jacobi matrices are matrix representations of the operator of multiplication by $x$ in $L^2(\mathbb{R}, d\nu)$. While details can be found in Section 3.1, this is immediately apparent from the reccurence relation (3.1.1).

By the same token, we are interested in a matrix representation of the operator of multiplication by $z$ in $L^2(S^1, d\mu)$. Here things are more complicated than in the real case, and the relevant details can be found in Section 3.3. Nonetheless, it turns out that the correct matrix representation was found by Cantero, Moral, and Velázquez [7]. In terms of the Verblunsky coefficients $\{\alpha_n\}_{n\geq 0}$ which appear in the circle recurrence relation (3.2.6), the CMV matrix has the form (see (3.7))

$$\mathcal{C} = \begin{pmatrix} \bar{\alpha}_0 & \rho_0\bar{\alpha}_1 & \rho_0\rho_1 & 0 & 0 & \dots \\ \rho_0 & -\alpha_0\bar{\alpha}_1 & -\alpha_0\rho_1 & 0 & 0 & \dots \\ 0 & \rho_1\bar{\alpha}_2 & -\alpha_1\bar{\alpha}_2 & \rho_2\bar{\alpha}_3 & \rho_2\rho_3 & \dots \\ 0 & \rho_1\rho_2 & -\alpha_1\rho_2 & -\alpha_2\bar{\alpha}_3 & -\alpha_2\rho_3 & \dots \\ 0 & 0 & 0 & \rho_3\bar{\alpha}_4 & -\alpha_3\bar{\alpha}_4 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Now we easily understand the first line of our scheme (1.0.1):

$$\text{OPRL} \quad \rightarrow \quad \text{Jacobi matrices} \quad \leftrightarrow \quad \text{Toda lattice}$$

The Toda evolution is equivalent to an evolution equation for Jacobi matrices in Lax pair form:

$$\dot{J} = [J, P]$$

for some antisymmetric matrix $P$. On the other hand, Jacobi matrices arise as matrix representations related to the theory of orthogonal polynomials on the real line.

We can now return to our original question: To what is this line transformed if we replace OPRL by OPUC? The answer is given by the second line in (1.0.1):

$$\text{OPUC} \quad \rightarrow \quad \text{CMV matrices} \quad \leftrightarrow \quad \text{defocusing AL}$$

These are, in short, the new results of this thesis.

That CMV matrices play, on the unit circle, the role that Jacobi matrices play on the line has already been shown, to the extent that they both are natural matrix representations associated to the respective orthogonal polynomial theories. In fact, the analogies run much deeper, as can be seen from [21]. In a certain sense (that can be made rigorous), this reflects the fact that real symmetric matrices can be reduced to Jacobi matrices, while unitary matrices can be reduced to CMV matrices.

Yet for our purposes here, the main question remains: Which integrable system fills the diagram? In other words, how does one find that defocusing AL is the system in question.

In principle, one can follow exactly the ideas from the real line case: write down abstractly the time evolution for CMV matrices, and from there deduce the evolution equation obeyed by the Verblunsky coefficients. It will turn out to be the defocusing Ablowitz-Ladik evolution.

In fact, such a mathematical picture is only very clear in hindsight. The road that we followed was slightly more complicated, but it also provided more information. As explained above, one can impose different boundary conditions to the AL equation; in particular, we can consider the case of periodic coefficients. Equally well, one can investigate the theory of OPUC with periodic Verblunsky coefficients. In the course of this investigation, Simon found that the space of Verblunsky coefficients naturally decomposed into tori that were also level sets of certain functions. This behavior is typical of integrable systems, as explained in Section 2.2.

This led Nenciu and Simon to the results presented in Chapter 4. More precisely, we found a symplectic form on the space $\mathbb{D}^p$ of Verblunsky coefficients, and proved that a certain set of functions Poisson commute. Moreover, Simon [29] proved that these functions are independent almost everywhere (in the sense of Section 2.2). So we found a completely integrable system naturally associated to OPUC. But what is the evolution equation of the coefficients under the first Hamiltonian in the hierarchy? A simple computation revealed that it is exactly defocusing AL with periodic boundary conditions.

Having found this, the next natural question was whether this connection between Ablowitz-Ladik and OPUC can be used to provide a Lax pair formulation for the evolution equation. This turns out to work, even though it is necessary to change the Hamiltonians that are being considered. In Chapter 5 we introduce the new Hamiltonians, and prove the main result: the Lax pairs for the whole AL hierarchy. All of this is done in the periodic setting. A few relatively simple observations further allow us to deduce the analogous statements for the other two types of boundary conditions: finite and infinite. We achieve this in Chapter 6. And, as scheme (1.0.1) claims, the Lax operators are indeed the CMV matrices associated to the coefficients which obey the AL evolution.

There are only a few aspects of (1.0.1) left to clarify, and they refer to the meaning of the vertical arrows. We proceed from left to right. If the measure $\mu$ on the circle is invariant with respect to complex conjugation, then it can naturally be mapped into a measure on the interval $[-2, 2]$, as in (3.6.1). This correspondence between measures translates into relations between the orthogonal polynomials and between the recurrence coefficients (see (3.6.2)). A particularly short proof of the Geronimus relations (3.6.2) was given by Killip and Nenciu [20], and in the process shows how to recover the associated Jacobi matrix from the CMV matrix. This is presented in Section 3.6.

As for the integrable systems side of the picture, Section 6.3 shows how the Toda lattice is part of the AL hierarchy. The "translation" of the Toda flow to Verblunsky coefficients has appeared previously in the literature under the name Schur flow, but the connection with the defocusing AL hierarchy is new. The details of this correspondence can be found in Section 6.3.

The organization of the thesis follows the summary given above. Chapter 2 presents background information on integrable systems. As the subject is huge, we give only the minimum information required to make sense of our claims. The Toda lattice is also given some attention, not only as the technically simplest case of a completely integrable differential-difference equation, but also since it plays an important role in the mathematical picture that we are presenting.

Chapter 3 sketches the basics of the theory of orthogonal polynomials. While OPRL are briefly introduced in Section 3.1, OPUC are given more space and attention, and most of the relevant facts are proven. The last section of the chapter gives details about the connection between OPRL and OPUC.

From here on, we focus on the task at hand. Chapter 4 introduces the relevant symplectic structure in the periodic setting, and gives a first proof of complete integrability. Still in the periodic case, Chapter 5 contains the main theorem about Lax pairs for the defocusing AL hierarchy. In particular, complete integrability is recovered as an easy corollary.

The last chapter, Chapter 6, completes the answer to the big, initial question, by finding Lax pair representations involving the CMV matrix for the finite and infinite defocusing AL hierarchy. Moreover, Section 6.3 shows how to recover the Toda evolution from the second AL Hamiltonian.

# Chapter 2

# Completely Integrable Systems

## 2.1 Brief History

The mathematical modeling of a great variety of nonlinear phenomena arising in physics leads to certain nonlinear equations. It is quite remarkable that many of these equations are integrable. While the notion of integrability cannot be easily and universally defined, one can think of an infinite-dimensional system as being integrable if there exists a change of variables which linearizes the flow. What makes (nonlinear) integrable systems fascinating is the fact that they exhibit a richer phenomenology than linear systems, while still being approachable to mathematical investigation. In particular, many of them have solitons, that is, localized solutions with particle-like behavior.

The fascinating new world of solitons and integrable behavior was discovered by Kruskal and Zabursky, who were trying to explain the curious numerical results of Fermi, Pasta, and Ulam [12]. They were thus led to study the Korteweg-de Vries equation, and discovered that the localized traveling-wave solutions of KdV had an unexpected behavior: After interaction, these waves regained their initial amplitude and velocity, the only effect of the interaction being a phase-shift. This particle-like behavior led them to call these waves "solitons."

The next challenge was to search for additional conservation laws believed to be responsible for the stability properties of solitons. This led Gardner, Greene, Kruskal, and Miura [14] to the connection between KdV and the time-independent

Schrödinger scattering problem: Let KdV describe the propagation of a water wave and suppose that this wave is frozen at a given instant in time. By bombarding this wave with quantum particles, one can reconstruct its shape from knowledge of how these particles scatter. In other words, the scattering data provide an alternative description of the wave at a given time. The time evolution of the water wave satisfies KdV, which is a nonlinear equation. The above alternative description of the shape of the wave would be useful if the evolution of the scattering data were linear. This is indeed the case, and hence this highly nontrivial change of variables provides a linearization of the KdV equation.

The essence of this discovery was soon grasped by Lax, who in [22] introduced the so-called Lax pair formulation of KdV. Following Lax's formulation, Zakharov and Shabat [37] solved the nonlinear Schrödinger equation. Immediately following this, many of the well-known nonlinear PDEs were rewritten in Lax pair form. Besides allowing one to do inverse scattering, they are also isospectral deformations, and provide a qualitative, geometric understanding of the evolution. This is the aspect of Lax pairs that we show in the next section. All the results given in this chapter are classical, and the study of Lax pairs has evolved very much from this stage. But even only these very simple considerations show that they are an interesting and powerful tool in integrable systems.

## 2.2 General Results

In this section we wish to introduce some of the basics of the theory of finite-dimensional integrable systems. We will follow the excellent presentation of Deift [9]. For more details see also [5].

We begin by introducing the notion of a Hamiltonian system.

**Definition 2.1.** *Let $M$ be a $(2n)$-dimensional manifold. Assume that there exists a 2-form $\omega$ on $M$ which is*

- *nondegenerate: $\omega(u, v) = 0$ for all $v \in T_m M$ impliesu $= 0$;*

- *closed:* $d\omega = 0$.

*In this case, the pair $(M, \omega)$ is called a symplectic manifold, and $\omega$ a symplectic form.*

The simplest example of a symplectic manifold is $M = \mathbb{R}^{2n}$ with the standard 2-form

$$\omega = \sum_{j=1}^{n} dx_j \wedge dy_j \, .$$

On a symplectic manifold $(M, \omega)$, we consider Hamiltonians, that is, smooth, real-valued functions $H : M \to \mathbb{R}$. Fix $m \in M$. Then

$$dH_m : T_m M \to \mathbb{R}$$

is a linear functional. Since $\omega$ is nondegenerate, there exists a unique vector $X_H(m) \in T_m M$ so that

$$dH_m(v) = \omega(X_H(m), v), \quad v \in T_m M.$$

So every Hamiltonian $H$ gives rise to a so-called Hamiltonian vector field $X_H$ on $M$.

*Example* 2.2. In the case $(\mathbb{R}^{2n}, \omega = \sum_{j=1}^{n} dx_j \wedge dy_j)$, let

$$v = \sum_{j=1}^{n} \left( a_j \frac{\partial}{\partial x_j} + b_j \frac{\partial}{\partial y_j} \right) \equiv \begin{bmatrix} a \\ b \end{bmatrix}$$

and

$$v' = \sum_{j=1}^{n} \left( a'_j \frac{\partial}{\partial x_j} + b'_j \frac{\partial}{\partial y_j} \right) \equiv \begin{bmatrix} a' \\ b' \end{bmatrix}$$

be tangent vectors. Then

$$\omega(v', v) = \left( \begin{bmatrix} a' \\ b' \end{bmatrix}, J \begin{bmatrix} a \\ b \end{bmatrix} \right)$$

where $(\cdot,\cdot)$ is the usual inner product in $\mathbb{R}^{2n}$ and

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

If

$$H : \mathbb{R}^{2n} \to \mathbb{R}$$

is a Hamiltonian, then

$$\begin{aligned} dH(v) &= \sum_{j=1}^{n} \left( \frac{\partial H}{\partial x_j} a_j + \frac{\partial H}{\partial y_j} b_j \right) \\ &= \left( \begin{bmatrix} H_x \\ H_y \end{bmatrix}, \begin{bmatrix} a \\ b \end{bmatrix} \right) \\ &= \left( X_H, J \begin{bmatrix} a \\ b \end{bmatrix} \right), \end{aligned}$$

with

$$X_H = J^t \begin{bmatrix} H_x \\ H_y \end{bmatrix} = \begin{bmatrix} H_y \\ -H_x \end{bmatrix}.$$

An evolution equation is called Hamiltonian if it is given by a Hamiltonian vector field:

$$\dot{m} = X_H(m).$$

Note that in $(\mathbb{R}^{2n}, \omega = \sum_{j=1}^{n} dx_j \wedge dy_j)$ this just recovers the classical notion of a Hamiltonian system:

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial y} \\ \dot{y} &= -\frac{\partial H}{\partial x} \end{aligned}$$

**Definition 2.3.** *In the general case, let $H$ and $K$ be two Hamiltonians. We define* *the Poisson bracket of $H$ and $K$ as*

$$\{H, K\} = \omega(X_H, X_K).$$

Then note that a Hamiltonian equation can be rewritten as

$$\dot{K} = X_H(K) = dK(X_H) = \{K, H\},$$

where $\dot{K}$ is the derivative of $K$ in the direction of $X_H$. Henceforth, we will express Hamiltonian equations as

$$\dot{K} = \{K, H\}.$$

Let $H, K$, and $L$ be three Hamiltonians on $(M, \omega)$. Then

$$d\omega(X_H, X_K, X_L)$$

is proportional to

$$\{\{H, K\}, L\} + \{\{K, L\}, H\} + \{\{L, H\}, K\}.$$

So while nondegeneracy of $\omega$ allows us to correctly define Hamiltonian vector fields, closedness of $\omega$ implies that the corresponding Poisson bracket obeys the Jacobi identity,

$$\{\{H, K\}, L\} + \{\{K, L\}, H\} + \{\{L, H\}, K\} = 0. \tag{2.2.1}$$

In particular, (2.2.1) implies

$$\begin{aligned}
[X_H, X_K](L) &= X_H(X_K(L)) - X_K(X_H(L)) \\
&= X_H(\{L, K\}) - X_K(\{L, H\}) \\
&= \{\{L, K\}, H\} - \{\{L, H\}, K\} \\
&= -\{\{K, L\}, H\} - \{\{L, H\}, K\} \\
&= \{\{H, K\}, L\} \\
&= X_{\{K,H\}}(L).
\end{aligned}$$

So we can conclude that the commutator of two Hamiltonian vector fields is also a Hamiltonian vector field,

$$[X_H, X_K] = X_{\{K,H\}}.$$

Moreover, two Hamiltonian vector fields commute if and only if the associated Hamiltonians Poisson commute.

The most interesting feature of Hamiltonian systems is that their flows can be linearized using relatively few conserved quantities.

Indeed, let

$$\dot{x} = \{x, H\} \tag{2.2.2}$$

be a Hamiltonian system on $(M, \omega)$, with the dimension of $M$ equal to $2n$.

**Definition 2.4.** *A function*

$$\phi : M \to \mathbb{R}$$

*that remains unchanged under the flow generated by $H$ is called an integral of motion (or conserved quantity).*

Note that such a function obeys

$$\dot{\phi} = 0 \iff \{\phi, H\} = 0.$$

Let $D \subset M$ be a domain that is invariant under the flow generated by $H$ for all time.

**Definition 2.5.** *We say that the system* (2.2.2) *is completely integrable on $D$ if there exist $n$ integrals of motion $H_1 = H, H_2, \ldots, H_n$ that are independent (meaning that their derivatives $dH_1, \ldots, dH_n$ are independent at each point in $D$), and Poisson commute:*

$$\{H_j, H_k\} = 0 \quad \textit{for all} 1 \leq j, k \leq n.$$

Then the following result describes the motion of the system:

**Theorem 2.6 (Liouville-Arnold-Jost).** *Assume that $H$ is completely integrable on a domain $D$ with integrals $H_1 = H, H_2, \ldots, H_n$ and suppose that*

$$N = \bigcap_{j=1}^{n} H_j^{-1}(0)$$

*is compact and connected. Then*

(a) *$N$ is an imbedded $n$-dimensional torus $\mathbb{T}^n$.*

(b) *There exists an open neighborhood $U \subset M$ of $N$ which can be coordinatized as follows: If $(\phi_1, \ldots, \phi_n)$ are coordinates on the torus $\mathbb{T}^n$ and $(x_1, \ldots, x_n) \in D_1$, where $D_1 \subset \mathbb{R}^n$ is a domain which contains the origin, then there exists a diffeomorphism*

$$\psi : D_1 \times \mathbb{T}^n \to U$$

*so that*

$$(H \circ \psi)(x_1, \ldots, x_n, \phi_1, \ldots, \phi_n) = h(x_1, \ldots, x_n)$$

*for some function $h$, and $\psi$ is symplectic, that is,*

$$\psi^* \omega = \sum_{j=1}^{n} dx_j \wedge dy_j.$$

*Sketch of proof.* The idea is to immerse $\mathbb{R}^{2n}$ into $M^{2n}$ using the flows generated by the commuting Hamiltonians. Indeed, let $\psi_j^{t_j}(m) = \psi_j(t_j, m)$ be the flow induced on

$M^{2n}$ by $H_j$, with $\psi_j(0, m) = m$. Then, fixing $m_0 \in N$, the map

$$t = (t_1, \ldots, t_n) \mapsto \Gamma(t) = \psi_1^{t_1} \circ \cdots \circ \psi_n^{t_n}(m_0)$$

takes $\mathbb{R}^n$ into the level set

$$N = \{m \in M : \, H_j(m) = H_j(m_0), \, j = 1, \ldots, n\}.$$

This is because

$$\frac{d}{dt_j} H_i(\psi_j^{t_j}(m)) = \{H_i, H_j\}(\psi_j^{t_j}(m)) = 0.$$

One can then prove that $\Gamma$ is onto $N$, and that

$$\Lambda = \{t \in \mathbb{R}^n : \, \Gamma(t) = m_0\}$$

is a lattice in $\mathbb{R}^n$. On the other hand, $\mathbb{R}^n/\Lambda$ is mapped diffeomorphically onto $N$. Since $N$ is compact by assumption, it follows that $\Lambda$ must have $n$ generators, and so $\mathbb{R}^n/\Lambda$, and consequently also $N$, is an $n$-torus. □

The question now becomes how can one find a large enough number of conserved quantities for a given system. A possible answer was given by Peter Lax [22], and it consists of the realization that if one can recast the system in the form of a Lax pair, then the evolution is isospectral, and so the spectrum of the Lax operator provides invariant quantities.

Let us be more specific. A Lax pair is an evolution equation for a (traditionally) self-adjoint operator $L$ which has the specific form

$$\dot{L} = [P, L],$$

where both $P$ and $L$ are time-dependent, and $P$ is anti-symmetric. Its main advantage consists of the fact that such an evolution is isospectral.

Indeed, let $Q(t)$ be the solution of

$$\frac{dQ}{dt} = -QP, \quad Q(t=0) = I.$$

Then note that, by anti-symmetry of $P$, we also have

$$\frac{dQ^*}{dt} = (-QP)^* = PQ^*.$$

Hence,

$$\frac{d}{dt}(QQ^*) = -(QP)Q^* + Q(PQ^*) = 0$$

and, similarly,

$$\frac{d}{dt}(Q^*Q) = 0.$$

We can then conclude that $Q$ is unitary.

Moreover, note that

$$\frac{d}{dt}(Q^*L(0)Q) = (PQ^*)L(0)Q - Q^*L(0)(QP) = [P, Q^*L(0)Q]$$

and

$$(Q^*L(0)Q)(t=0) = L(0).$$

Thus we get

$$Q^*(t)L(0)Q(t) = L(t),$$

and so $L(t)$ is unitarily equivalent to $L(0)$. In particular, if $\lambda$ is an eigenvalue of $L(0)$, then it is also an eigenvalue of $L(t)$. In other words, the eigenvalues of the Lax operator $L$ represent integrals of motion.

We are more interested here in the situation when $L$ is unitary, rather than self-adjoint. The above proof works without change as long as $P$ is anti-Hermitian. In particular, we see that a unitary operator evolving according to a Lax pair remains unitary at all time and the evolution is isospectral.

One can prove directly, by methods very similar to those above, preservation of unitarity. Assume

$$\dot{L} = [P, L],$$

with $L(0)$ unitary, and $P(t)^* = -P(t)$ for any $t$. Then it is also true that

$$\frac{d}{dt} L^* = [P, L^*]$$

and so

$$\frac{d}{dt}(LL^*) = [P, LL^*], \quad \frac{d}{dt}(L^*L) = [P, L^*L]$$

with $(LL^*)(0) = (L^*L)(0) = I$. By uniqueness of the solution we see

$$LL^* = L^*L = I$$

for all time $t$.

One can prove conservation of eigenvalues directly, without investigating the evolution of the eigenvectors (as the method presented before actually does). Indeed, assume that $L$ is a unitary matrix obeying a Lax equation, and let $\lambda \in S^1$ be an eigenvalue of $L$ and $\phi$ the corresponding unit eigenvector. Then

$$\lambda = (L\phi, \phi)$$

and hence,

$$\dot{\lambda} = (\dot{L}\phi, \phi) + (L\dot{\phi}, \phi) + (L\phi, \dot{\phi}).$$

Note that

$$\begin{aligned}
(L\dot{\phi}, \phi) + (L\phi, \dot{\phi}) &= (\dot{\phi}, L^*\phi) + (L\phi, \dot{\phi}) \\
&= (\dot{\phi}, \bar{\lambda}\phi) + (\lambda\phi, \dot{\phi}) \\
&= \lambda[(\dot{\phi}, \phi) + (\phi, \dot{\phi})] \\
&= \lambda(\phi, \phi)^{\cdot} \\
&= 0,
\end{aligned}$$

as $(\phi, \phi) \equiv 1$. Here we also used the fact that $L$ is unitary for all time, and so $L^*\phi = \bar{\lambda}\phi$ always holds.

So,

$$\begin{aligned}\dot{\lambda} &= (\dot{L}\phi, \phi) = ([L, P]\phi, \phi) \\ &= (P\phi, L^*\phi) - (PL\phi, \phi) \\ &= (P\phi, \bar{\lambda}\phi) - (\lambda P\phi, \phi) = 0,\end{aligned}$$

as claimed.

Both situations presented above are interesting. The Toda lattice, which we present in the next section, can be rewritten in terms of (real-)symmetric Lax matrices, as we will explain. The main result of this thesis consists of finding unitary Lax matrices for another classical integrable system, defocusing Ablowitz-Ladik.

## 2.3 The Toda Equation

Consider the classical mechanics problem of a 1-dimensional chain of particles with nearest neighbor interactions. Assume that the system is uniform (contains no impurities) and that the mass of each particle is 1. Then the equation that governs the evolution is

$$\frac{d^2 y_n}{dt^2} = V'(y_{n+1} - y_n) - V'(y_n - y_{n-1}), \tag{2.3.1}$$

where $y_n$ denotes the displacement of the $n^{\text{th}}$ particle, and $V$ is the interaction potential between neighboring particles.

If $V'(r)$ is proportional to $r$, then the interaction is linear, and the solutions are given by linear superpositions of the normal modes

$$y_n^{(l)} = C_n \sin\left(\frac{\pi l}{N+1}\right) \cos\left(\omega_l t + \delta_l\right).$$

In this case there is no transfer of energy between the modes.

The general belief in the early 1950s was that if a nonlinearity is introduced, then energy would flow between the different modes, eventually leading to a stable state of statistical equilibrium. Fermi, Pasta, and Ulam [12] set out to numerically investigate this phenomenon through a computer experiment performed at Los Alamos in 1955. What they found instead was quasiperiodic motion of the system. This phenomenon was explained by the connection to solitons, and by the discovery by Morikazu Toda [33] of what is now called the Toda lattice.

Before proceeding to describe this particular system, let us note that the equation (2.3.1) can be recast as a Hamiltonian system: Set

$$\mathcal{H} : \mathbb{R}^{2n} \to \mathbb{R},$$

$$\mathcal{H}(p, q) = \sum_{j=1}^{n} \left( \frac{p_j^2}{2} + V(q_{k+1} - q_k) \right).$$

Then the Hamiltonian system generated by $\mathcal{H}$ in $(\mathbb{R}^{2n}, \omega = \sum_{j=1}^{n} dx_j \wedge dy_j)$ is

$$\begin{cases} \dot{p}_j = \frac{\partial \mathcal{H}}{\partial q_j} = V'(q_j - q_{j-1}) - V'(q_{j+1} - q_j) \\ \dot{q}_j = -\frac{\partial \mathcal{H}}{\partial p_j} = -p_j. \end{cases}$$

This is equivalent to (2.3.1) if we set

$$y_k = q_k.$$

The Toda lattice is given by setting

$$V(r) = e^{-r} + r - 1.$$

It was introduced in 1972 by Toda, and its main interest at the time was that it had solitons.

Complete integrability of the system was proved by Flaschka in 1974 by introducing a change of variables that allowed him to set the system in Lax pair form.

Flaschka's change of variables is given by

$$\begin{cases} a_k = \frac{1}{2} e^{-\frac{q_{k+1}-q_k}{2}} \\ b_k = -\frac{1}{2} p_k \, . \end{cases} \tag{2.3.2}$$

The new variables obey the evolution equations

$$\dot{b}_k = -\tfrac{1}{2}\dot{p}_k = \tfrac{1}{2}\left[e^{-(q_k-q_{k-1})} - e^{-(q_{k+1}-q_k)}\right] = 2(a_{k-1}^2 - a_k^2), \tag{2.3.3}$$

with $a_0 = a_n = 0$, and

$$\dot{a}_k = \tfrac{1}{4} e^{-\frac{q_{k+1}-q_k}{2}} [-\dot{q}_{k+1} + \dot{q}_k] = a_k(b_k - b_{k+1}). \tag{2.3.4}$$

One can also rewrite the Poisson bracket from the $p, q$ variables

$$\{p_j, p_k\} = 0 \quad \{q_j, q_k\} = 0 \quad \{p_j, q_k\} = \delta_{j,k}$$

into the new variables:

$$\{b_k, a_k\} = -\tfrac{1}{4} a_k,$$

$$\{b_k, a_{k-1}\} = \tfrac{1}{4} a_{k-1}$$

for all $k = 1, \ldots, n$, and all the other brackets are zero.

Now set $J$ to be the Jacobi matrix with these $a$ and $b$, that is,

$$J = \begin{bmatrix} b_1 & a_1 & & \\ a_1 & b_2 & a_2 & \\ & a_2 & b_3 & \ddots \\ & & \ddots & \ddots \end{bmatrix}.$$

Flaschka's main observation was that the system of equations (2.3.3) and (2.3.4) is equivalent to the Lax pair

$$\dot{J} = [J, P]$$

with

$$P = J_+ - J_- = \begin{bmatrix} 0 & a_1 & & \\ -a_1 & 0 & a_2 & \\ & -a_2 & 0 & \ddots \\ & & \ddots & \ddots \end{bmatrix}.$$

One can also use the Lax pairs to prove that if $\lambda$ and $\mu$ are two distinct eigenvalues of $J$, then

$$\{\lambda, \mu\} = 0. \tag{2.3.5}$$

This proof can be found in [13], for example, but may have been known before that.

First note that (2.3.2) implies

$$\{b_k, a_k\} = -\tfrac{1}{4}a_k,$$

$$\{b_k, a_{k-1}\} = \tfrac{1}{4}a_{k-1}$$

for all $k = 1, \ldots, n$, and all the other brackets are zero. (Bear in mind that we imposed the conditions $a_0 = a_n = 0$.)

Then, if $\phi$ is a unit eigenvector of $J$ with eigenvalue $\lambda$, we get

$$\lambda = (L\phi, \phi).$$

Differentiating with respect to $b_k$, we have

$$\begin{aligned} \frac{\partial \lambda}{\partial b_k} &= \left(\frac{\partial J}{\partial b_k}\phi, \phi\right) + \left(J\frac{\partial \phi}{\partial b_k}, \phi\right) + \left(J\phi, \frac{\partial \phi}{\partial b_k}\right) \\ &= \left(\frac{\partial J}{\partial b_k}\phi, \phi\right) + 2\lambda\frac{\partial}{\partial b_k}(\phi, \phi) \\ &= \left(\frac{\partial J}{\partial b_k}\phi, \phi\right) \\ &= \phi_k^2 \end{aligned}$$

for any $1 \leq k \leq n$, as $b_k$ appears exactly once in $J$, in the position $(k, k)$. Similarly, we get that

$$\frac{\partial \lambda}{\partial a_j} = 2\phi_j \phi_{j+1}$$

for any $1 \le j \le n-1$.

Now let $\mu$ be another eigenvalue of $J$, and $\psi$ the associated unit eigenvector. Then we have

$$\begin{aligned}
\{\lambda, \mu\} &= \sum_{k=1}^{n-1} \{b_k, a_k\} \Big[ \frac{\partial \lambda}{\partial b_k} \cdot \frac{\partial \mu}{\partial a_k} - \frac{\partial \lambda}{\partial a_k} \cdot \frac{\partial \mu}{\partial b_k} \Big] \\
&\quad + \sum_{k=2}^{n} \{b_k, a_{k-1}\} \Big[ \frac{\partial \lambda}{\partial b_k} \cdot \frac{\partial \mu}{\partial a_{k-1}} - \frac{\partial \lambda}{\partial a_{k-1}} \cdot \frac{\partial \mu}{\partial b_k} \Big] \\
&= \tfrac{1}{2} \sum_{k=1}^{n} a_{k-1} \big[ \phi_k^2 \psi_{k-1} \psi_k - \psi_k^2 \phi_{k-1} \phi_k \big] \\
&\quad - \tfrac{1}{2} \sum_{k=1}^{n} a_k \big[ \phi_k^2 \psi_k \psi_{k+1} - \psi_k^2 \phi_k \phi_{k+1} \big] \\
&= \tfrac{1}{2} \sum_{k=1}^{n} \phi_k \psi_k \big[ R_k + R_{k+1} \big],
\end{aligned}$$

where

$$R_k = a_k \big( \phi_{k+1} \psi_k - \phi_k \psi_{k+1} \big).$$

Recall

$$J\phi = \lambda \phi.$$

This is the same as

$$a_k \phi_{k+1} + b_k \phi_k + a_k \phi_{k-1} = \lambda \phi_k.$$

If we multiply this by $\psi_k$ and then subtract the expression in which we interchange the roles of $\phi$ and $\psi$, we obtain

$$a_k \big( \phi_{k+1} \psi_k - \psi_{k+1} \phi_k \big) + a_{k-1} \big( \phi_{k-1} \psi_k - \psi_{k-1} \phi_k \big) = (\lambda - \mu) \phi_k \psi_k,$$

or, in terms of $R$,

$$\phi_k \psi_k = \frac{1}{\lambda - \mu} (R_k - R_{k-1}).$$

In other words, if we plug this into our expression for the Poisson bracket of the eigenvalues, we see that

$$\{\lambda, \mu\} = \frac{1}{2(\lambda-\mu)} \sum_{k=1}^{n} \left(R_k^2 - R_{k-1}^2\right) = \frac{1}{2(\lambda-\mu)} \left(R_n^2 - R_0^2\right) = 0$$

as $R_0 = R_n = 0$.

While this is a relatively easy proof, it relies heavily on the particular structure of the Jacobi matrix $J$, without explaining in any way why the computations miraculously turn out to give the right answer.

A step in the direction of this explanation is the following: If we rewrite the standard Poisson bracket from the $p$ and $q$ variables into the $a$ and $b$ variables, we will see that, in fact,

$$\{J, H_2\} = \dot{J} = [J, P],$$

where

$$H_2 = \tfrac{1}{2}\mathrm{Tr}(J^2).$$

Now define

$$H_n = \tfrac{1}{n}\mathrm{Tr}(J^n)$$

for all $n \geq 1$. Then van Moerbeke proved in [36] that

$$\{J, H_{n+1}\} = [J, (J^n)_+ - (J^n)_-].$$

(In fact, he considers the periodic Toda lattice, but his result implies the one claimed here by taking one of the $a$'s to 0.) This implies, by the general theory, that the eigenvalues of $J$ are conserved by all of these flows. Putting it differently,

$$\{\lambda, H_n\} = 0$$

for any eigenvalue $\lambda$ and any $n \geq 2$. But since the $H_n$'s are essentially traces of powers of $J$, it follows that this is another (more complicated) proof for (2.3.5); we also get

$$\{H_n, H_m\} = 0$$

for any $n, m$.

*Remark* 2.7. That the $H_n$'s are the correct Hamiltonians to consider, and the role that the special structure of $J$ plays, is explained by the underlying Lie algebra and by the identification of the Poisson bracket defined here with the so-called Kostant-Kirilov bracket. As it turns out, a similar statement is true in the Ablowitz-Ladik case, where the Poisson bracket is the Gelfand-Dikij bracket on an appropriately chosen Lie algebra. This is part of work in progress jointly with Rowan Killip [21].

# Chapter 3

# Orthogonal Polynomials

## 3.1 Orthogonal Polynomials on the Real Line

In the presentation of the Toda lattice in Section 2.3 we claimed that there exists a close connection between Jacobi matrices and orthogonal polynomials on the real line. In this section we will present some of the very basic results of this theory.

Let $\nu$ denote a nontrivial (i.e., with infinite support) measure on $\mathbb{R}$ so that

$$\int_{\mathbb{R}} |x|^n \, d\nu(x) < \infty$$

for all $n \geq 0$. Since $\nu$ is nontrivial, the monomials

$$1, x, x^2, \ldots, x^n, \ldots$$

are linearly independent in $L^2(\mathbb{R}, d\nu)$. So we can define monic orthogonal polynomials $\{P_n\}_{n\geq 0}$ and orthonormal polynomials $\{p_n\}_{n\geq 0}$ using the Gram-Schmidt procedure. We get for $n, m \geq 0$ that

$$(P_n, P_m) = \gamma_n^{-2} \delta_{nm}, \qquad (p_n, p_m) = \delta_{nm},$$

where

$$P_n(x) = x^n + \text{lower order}, \qquad p_n(x) = \gamma_n x^n + \text{lower order},$$

$\gamma_n > 0$ and $(\cdot,\cdot)$ denotes the usual inner product in $L^2(\mathbb{R}, d\nu)$,

$$(f, g) = \int_{\mathbb{R}} f(x)g(x)\, d\nu(x).$$

Note that

$$(xf, g) = \int_{\mathbb{R}} xf(x)g(x)\, d\nu(x) = (f, xg).$$

This simple observation allows one to prove the recurrence relations for orthogonal polynomials on the real line. Indeed, for $n \geq 0$, $xP_n(x)$ is a monic polynomial of degree $n+1$. Therefore,

$$xP_n(x) - P_{n+1}(x)$$

is a polynomial of degree at most $n$. Moreover, for $j \leq n-2$,

$$(xP_n, P_j) = (P_n, xP_j) = 0 = (P_{n+1}, P_j)$$

since $n > j + 1 = \deg(xP_j)$. Also,

$$(xP_n, P_{n-1}) = (P_n, xP_{n-1}) = (P_n, x^n) = \|P_n\|^2 > 0.$$

So there exist numbers $a_k > 0$ and $b_k \in \mathbb{R}$, $k \geq 1$, so that

$$xP_n(x) = P_{n+1}(x) + b_{n+1}P_n(x) + a_n^2 P_{n-1}(x)$$

for all $n \geq 0$, with $a_0 = 0$ and

$$a_n = \frac{\|P_n\|}{\|P_{n-1}\|}$$

for $n \geq 1$.

This implies

$$\gamma_n^{-1} = \|P_n\| = \prod_{j=1}^{n} a_j\, .$$

Therefore,

$$p_n(x) = \frac{1}{a_1 \cdot \ldots \cdot a_n} P_n(x)$$

and the orthonormal polynomials obey the recurrence relation

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_{n+1}p_n(x) + a_n p_{n-1}(x). \tag{3.1.1}$$

In other words, the operator of multiplication by $x$ in $L^2(d\nu)$ can be represented in the basis of orthonormal polynomials by the Jacobi matrix

$$J_\nu = \begin{bmatrix} b_1 & a_1 & & \\ a_1 & b_2 & a_2 & \\ & a_2 & b_3 & \ddots \\ & & \ddots & \ddots \end{bmatrix}.$$

Favard's theorem says that, given any Jacobi matrix $J$, there exists a measure $\nu$ on the real line for which

$$J = J_\nu.$$

In general, $\nu$ is not unique. (Contrast this with Verblunsky's Theorem 3.5 on the unit circle.)

## 3.2 Orthogonal Polynomials on the Unit Circle

In this section we present some of the basic notions and results related to the theory of orthogonal polynomials on the unit circle. The reader interested in more details can check Szegő's classic book [31]. In our presentation, we follow the two-volume treatise by Simon [28, 29]. For a shorter presentation of the subject, see [30].

Consider a probability measure $\mu$ on $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$. If $\mu$ is supported at infinitely many points, then the monomials $\{z^n\}_{n\geq 0}$ are independent in $L^2(d\mu)$ and one can apply the Gram-Schmidt procedure to produce the monic orthogonal polynomials $\{\Phi_n\}_{n\geq 0}$ and the orthonormal polynomials

$$\phi_n(z) = \frac{\Phi_n(z)}{\|\Phi_n\|_{L^2(d\mu)}}.$$

For each $n \geq 0$, we define the operator $R_n$ on $L^2(d\mu)$ by

$$(R_n f)(z) = z^n \overline{f(z)} = z^n \overline{f\left(1/\bar{z}\right)}.$$

The second equality holds since $z = 1/\bar{z}$ on $S^1$. Note that $R_n$ is anti-unitary; moreover $R_n^2 = I$, and if $f$ is a polynomials of degree at most $n$, then so is $R_n f$.

For orthogonal polynomials $\Phi_n$ and $\phi_n$, we shall use the standard (but somewhat ambiguous) notation

$$\Phi_n^* = R_n \Phi_n, \qquad \phi_n^* = R_n \phi_n.$$

Recall that $\Phi_n$ is, up to multiplication by a constant, the unique polynomial of degree at most $n$ orthogonal to $1, z, \ldots, z^{n-1}$. Since $R_n$ is anti-unitary and

$$R_n z^j = z^{n-j},$$

we obtain that $\Phi_n^*$ is, again up to multiplicative constants, the unique polynomial of degree less than or equal to $n$ that is orthogonal to $z, z^2, \ldots, z^n$.

Further note that $\Phi_n^*(0) = 1$ (since $\Phi_n$ is monic), and

$$\|\Phi_n\|^2 = \|R_n \Phi_n\|^2 = \|\Phi_n^*\|^2 = \int \Phi_n^*(e^{i\theta})\, d\mu(e^{i\theta}), \tag{3.2.1}$$

where all norms are taken in $L^2(d\mu)$.

**Theorem 3.1 (Szegö Recursion).** *For any nontrivial measure $\mu$ on $S^1$, there exists a sequence of complex numbers $\{\alpha_n\}_{n\geq 0} \subset \mathbb{D}$ so that the monic orthogonal polynomials obey the Szegö recursion formulae*

$$\Phi_{k+1}(z) = z\Phi_k(z) - \bar{\alpha}_k \Phi_k^*(z), \tag{3.2.2}$$

$$\Phi_{k+1}^*(z) = \Phi_k^*(z) - \alpha_k z \Phi_k(z). \tag{3.2.3}$$

*Moreover,*

$$\|\Phi_{n+1}\|_{L^2(d\mu)}^2 = \prod_{j=0}^{n} (1 - |\alpha_j|^2). \tag{3.2.4}$$

*Proof.* Since $\Phi_j$ is a monic polynomial of degree $j$, we obtain that

$$\Phi_{n+1}(z) - z\Phi_n(z)$$

is a polynomial of degree at most $n$. Furthermore, for $1 \leq j \leq n$, we have

$$(z\Phi_n, z^j) = (\Phi_n, \bar{z}z^j) = (\Phi_n, z^{j-1}) = 0 = (\Phi_{n+1}, z^j).$$

By the previous discussion, we see that $\Phi_{n+1}(z) - z\Phi_n(z)$ must be a constant multiple of $\Phi_n^*(z)$,

$$\Phi_{n+1}(z) - z\Phi_n(z) = -\bar{\alpha}_n\Phi_n^*(z)$$

for some $\alpha_n \in \mathbb{C}$. This is (3.2.2); equation (3.2.3) follows by applying $R_{n+1}$.

Rewrite (3.2.2) as

$$z\Phi_n(z) = \Phi_{n+1}(z) + \bar{\alpha}_n\Phi_n^*(z)$$

and take norms. As $\deg(\Phi_n^*) \leq n$, we will have that $\Phi_{n+1}$ is orthogonal to $\Phi_n^*$, and hence,

$$\|\Phi_n\|^2 = \|z\Phi_n\|^2 = \|\Phi_{n+1}\|^2 + |\alpha_n|^2\|\Phi_n^*\|^2.$$

But we know that $\|\Phi_n^*\|^2 = \|\Phi_n\|^2$, and hence,

$$\|\Phi_{n+1}\|^2 = (1 - |\alpha_n|^2)\|\Phi_n\|^2. \tag{3.2.5}$$

Since $\Phi_j \neq 0$ in $L^2(d\mu)$ for all $j \geq 0$, we obtain

$$1 - |\alpha_n|^2 > 0 \iff \alpha_n \in \mathbb{D}, \quad \text{for all}\, n \geq 0,$$

and, by iterating (3.2.5), we recover (3.2.4):

$$\|\Phi_{n+1}\|^2 = \prod_{j=0}^{n}(1 - |\alpha_j|^2).$$

$\square$

*Remark* 3.2. Following Simon [28, 29], we call the $\alpha_n$'s Verblunsky coefficients. The choice of using $-\bar{\alpha}_n$ in (3.2.2) was made so that the $\alpha$'s also represent the coefficients in the Schur algorithm for the Schur function associated to the measure $\mu$. For more details, see [28].

Recall

$$\phi_n(z) = \frac{\Phi_n(z)}{\|\Phi_n\|} = \Big(\prod_{j=0}^{n-1} \rho_j\Big)^{-1} \Phi_n(z),$$

where

$$\rho_j = \sqrt{1 - |\alpha_j|^2}.$$

Hence,

$$\begin{aligned} \phi_{n+1}(z) &= \Big(\prod_{j=0}^{n} \rho_j\Big)^{-1} \Phi_{n+1}(z) \\ &= \frac{1}{\rho_n}\Big[z\Big(\prod_{j=0}^{n-1} \rho_j\Big)^{-1} \Phi_n(z) - \bar{\alpha}_n \Big(\prod_{j=0}^{n-1} \rho_j\Big)^{-1} \Phi_n^*(z)\Big] \\ &= \frac{1}{\rho_n}\big[z\phi_n(z) - \bar{\alpha}_n \phi_n^*(z)\big]. \end{aligned}$$

Here we used the fact that $\rho_j \in \mathbb{R}$ for all $j \geq 0$, and so

$$R_n\Big(\Big(\prod_{j=0}^{n-1} \rho_j\Big)^{-1} \Phi_n\Big) = \Big(\prod_{j=0}^{n-1} \rho_j\Big)^{-1} R_n \Phi_n.$$

So the orthonormal polynomials obey the recurrence relations

$$\phi_{n+1}(z) = \frac{1}{\rho_n}\big[z\phi_n(z) - \bar{\alpha}_n \phi_n^*(z)\big] \tag{3.2.6}$$

and

$$\phi_{n+1}^*(z) = \frac{1}{\rho_n}\big[\phi_n^*(z) - \alpha_n z \phi_n(z)\big]. \tag{3.2.7}$$

These recursion relations for the orthonormal polynomials can be summarized as

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = A(\alpha_{n-1}, z) \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix},$$

where

$$A(\alpha_k, z) = \frac{1}{\rho_k} \begin{bmatrix} z & -\bar{\alpha}_k \\ -\alpha_k z & 1 \end{bmatrix}.$$

We define the transfer matrix

$$T_n(z) = A(\alpha_{n-1}, z) \ldots A(\alpha_0, z) \tag{3.2.8}$$

for all $n \geq 1$; hence,

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = T_n(z) \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

**Lemma 3.3.** *For each $n \geq 0$, there exist polynomials $A_n$ and $B_n$ of degree $n$, called Wall polynomials, so that*

$$T_n(z) = \Big(\prod_{j=0}^{n-1} \rho_j\Big)^{-1} \begin{bmatrix} zB_{n-1}^*(z) & -A_{n-1}^*(z) \\ -zA_{n-1}(z) & B_{n-1}(z) \end{bmatrix}. \tag{3.2.9}$$

*Moreover, these polynomials obey the recurrence relations*

$$A_{n+1}(z) = A_n(z) + \alpha_{n+1} z B_n^*(z) \tag{3.2.10}$$

$$B_{n+1}(z) = B_n(z) + \alpha_{n+1} z A_n^*(z) \tag{3.2.11}$$

*with $A_0(z) = \alpha_0$ and $B_0(z) = 1$.*

*Proof.* We proceed by induction.

For $n = 0$,

$$T_1(z) = \frac{1}{\rho_0} \begin{bmatrix} z & -\bar{\alpha}_0 \\ -\alpha_0 z & 1 \end{bmatrix}$$

and hence (3.2.9) holds with $A_0(z) = \alpha_0$ and $B_0(z) = 1$.

Assume that (3.2.9) holds for some $n \geq 0$. Then

$$\begin{aligned}
T_{n+2}(z) &= A(\alpha_{n+1}, z)T_{n+1}(z) \\
&= \frac{1}{\rho_{n+1}} \begin{bmatrix} z & -\bar{\alpha}_{n+1} \\ -\alpha_{n+1}z & 1 \end{bmatrix} \cdot \frac{1}{\prod_{j=0}^{n} \rho_j} \begin{bmatrix} zB_n^*(z) & -A_n^*(z) \\ -zA_n(z) & B_n(z) \end{bmatrix} \\
&= \Big(\prod_{j=0}^{n+1} \rho_j\Big)^{-1} \begin{bmatrix} z(zB_n^*(z) + \bar{\alpha}_{n+1}A_n(z)) & -(zA_n^*(z) + \bar{\alpha}_{n+1}B_n(z)) \\ -z(\alpha_{n+1}zB_n^*(z) + A_n(z)) & \alpha_{n+1}zA_n^*(z) + B_n(z) \end{bmatrix}.
\end{aligned}$$

Denote

$$B_{n+1}(z) = B_n(z) + \alpha_{n+1}zA_n^*(z)$$

and

$$A_{n+1}(z) = A_n(z) + \alpha_{n+1}zB_n^*(z).$$

Then

$$\begin{aligned}
B_{n+1}^*(z) &= R_{n+1}B_{n+1}(z) \\
&= R_{n+1}B_n(z) + \bar{\alpha}_{n+1}R_{n+1}(zA_n^*(z)) \\
&= zR_nB_n(z) + \bar{\alpha}_{n+1}R_nA_n^*(z) \\
&= zB_n^*(z) + \bar{\alpha}_{n+1}A_n(z)
\end{aligned}$$

and similarly,

$$A_{n+1}^*(z) = zA_n^*(z) + \bar{\alpha}_{n+1}B_n(z).$$

This proves (3.2.9), (3.2.10), and (3.2.11) for $n + 1$.

Further note that

$$A_{n+1}^*(z) = zA_n^*(z) + \bar{\alpha}_{n+1}B_n(z)$$

$$B_{n+1}^*(z) = zB_n^*(z) + \bar{\alpha}_{n+1}A_n(z).$$

Since $B_0^*(z) = 1$ and $A_0^*(z) = \bar{\alpha}_0$ have degree 0, these recurrence relations also imply that $B_k^*$ and $A_k^*$ are polynomials of degree exactly $k$, for all $0 \leq k \leq n + 1$. This, together with (3.2.10) and (3.2.11), implies that $A_k$ and $B_k$ are also polynomials of degree (exactly equal to) $k$, as claimed. □

*Remark* 3.4. The Wall polynomials and the recurrence relations that they obey will play a vital role in Chapter 4 in proving complete integrability for the Ablowitz-Ladik system.

Finally, we give without proof the theorem that is the analogue on the unit circle of Favard's theorem from the real line:

**Theorem 3.5 (Verblunsky).** *There exists a one-to-one correspondence between nontrivial probability measures on the unit circle and sequences of Verblunsky coefficients* $\{\alpha_n\}_{n\geq 0} \subset \mathbb{D}$.

## 3.3 The CMV Matrix

Consider the operator

$$f(z) \mapsto zf(z)$$

in $L^2(d\mu)$. We want to represent this operator as a matrix. In order to achieve this we first need to choose an appropriate basis in $L^2(d\mu)$.

A first, natural choice would be the set of orthonormal polynomials, $\{\phi_n\}_{n\geq 0}$. But there are several reasons why this does not represent the best choice. On the one hand, the orthonormal polynomials form a basis in $L^2(d\mu)$ if and only if the Verblunsky coefficients are not square-summable: $\sum_{j=0}^{\infty} |\alpha_j|^2 = \infty$ (for a proof of this statement see [28]).

On the other hand, even in the case when $\{\phi_j\}$ is a basis in $L^2(d\mu)$, the matrix that we obtain is a Hessenberg matrix:

$$\mathcal{G}_{kl} = (\phi_k, z\phi_l) = \begin{cases} -\bar{\alpha}_l \alpha_{k-1} \prod_{j=k}^{l-1} \rho_j, & 0 \leq k \leq l, \\ \rho_l, & k = l+1, \\ 0, & k \geq l+2. \end{cases}$$

Generically, all the entries above the main diagonal and the first subdiagonal are nonzero, and they depend on an unbounded number of Verblunsky coefficients. Consequently, this matrix representation, known as the GGT matrix, is somewhat difficult to manipulate.

A more useful orthonormal basis was recently discovered by Cantero, Moral, and Velázquez [7]. Indeed, they define two such bases: Applying the Gram-Schmidt procedure to

$$1, z, z^{-1}, z^2, z^{-2}, \dots$$

in $L^2(d\mu)$ produces the orthonormal basis $\{\chi_n\}_{n\geq 0}$. Similarly, we obtain a second orthonormal basis $\{x_n\}_{n\geq 0}$ from

$$1, z^{-1}, z, z^{-2}, z^2, \dots .$$

Denote

$$\chi_n^{(0)}(z) = \begin{cases} z^{-k}, & n = 2k, \\ z^{k+1}, & n = 2k+1, \end{cases}$$

and $P^{(n)}$ the orthonormal projection in $L^2(d\mu)$ onto

$$\mathcal{H}^{(n)} = \begin{cases} \mathcal{H}_{(-k,k)}, & n = 2k, \\ \mathcal{H}_{(-k,k+1)}, & n = 2k+1, \end{cases}$$

where $\mathcal{H}_{(k,l)}$ is the subspace of Laurent polynomials spanned by $z^k, z^{k+1}, \dots, z^l$. Then note that

$$\chi_n = \frac{(1 - P^{(n)})\chi_n^{(0)}}{\|(1 - P^{(n)})\chi_n^{(0)}\|} .$$

**Proposition 3.6.** *We have*

$$\chi_n(z) = \begin{cases} z^{-k+1}\phi_{2k-1}(z), & n = 2k-1, \\ z^{-k}\phi_{2k}^*(z), & n = 2k, \end{cases} \tag{3.3.1}$$

*and*

$$x_n(z) = \begin{cases} z^{-k}\phi^*_{2k-1}(z), & n = 2k-1, \\ z^{-k}\phi_{2k}(z), & n = 2k. \end{cases} \tag{3.3.2}$$

*In particular,*

$$x_n(z) = \overline{\chi_n\,(1/\bar{z})}. \tag{3.3.3}$$

*Proof.* Note that

$$\phi_{2n-1}(z) = \frac{(1 - P_{(0,2n-2)})z^{2n-1}}{\|(1 - P_{(0,2n-2)})z^{2n-1}\|},$$

where $P_{(k,l)}$ is the orthogonal projection in $L^2(d\mu)$ onto $\mathcal{H}_{(k,l)}$. Also note that

$$z^l P_{(k,m)} z^{-l} = P_{(k+l,m+l)}.$$

Given these two facts, we obtain

$$\begin{aligned} z^{-k+1}\phi_{2k-1} &= \frac{\left[z^{-k+1}(1 - P_{(0,2k-2)})z^{k-1}\right]z^k}{\|\left[z^{-k+1}(1 - P_{(0,2k-2)})z^{k-1}\right]z^k\|} \\ &= \frac{(1 - P_{(-k+1,k-1)})z^k}{\|(1 - P_{(-k+1,k-1)})z^k\|} \\ &= \frac{(1 - P^{(2k-2)})\chi^{(0)}_{2k-1}}{\|(1 - P^{(2k-2)})\chi^{(0)}_{2k-1}\|} \\ &= \chi_{2k-1}, \end{aligned}$$

which proves half of (3.3.1). The other part of (3.3.1), as well as (3.3.2), can be proved by similar calculations.

Finally, (3.3.3) follows immediately from (3.3.1) and (3.3.2):

$$\begin{aligned} \overline{\chi_{2k-1}\,(1/\bar{z})} &= \overline{\bar{z}^{-(-k+1)}\phi_{2k-1}\,(1/\bar{z})} \\ &= z^{k-1}\overline{\phi_{2k-1}\,(1/\bar{z})} \\ &= z^{-k}\cdot z^{2k-1}\overline{\phi_{2k-1}\,(1/\bar{z})} \\ &= z^{-k}\phi^*_{2k-1}(z) \\ &= x_{2k-1}(z), \end{aligned}$$

and

$$\begin{aligned}
\overline{\chi_{2k}\left(1/\bar{z}\right)} &= \overline{\bar{z}^{-(-k)}\phi_{2k}^{*}\left(1/\bar{z}\right)} \\
&= z^{k}\overline{\phi_{2k}^{*}\left(1/\bar{z}\right)} \\
&= z^{-k}\cdot z^{2k}\overline{\phi_{2k}^{*}\left(1/\bar{z}\right)} \\
&= z^{-k}\phi_{2k}(z) \\
&= x_{2k}(z).
\end{aligned}$$

□

The CMV matrix representation is given by

$$\mathcal{C}_{k,l} = (\chi_k, z\chi_l).$$

for all $k, l \geq 0$.

**Proposition 3.7.** *The CMV matrix is given by*

$$\mathcal{C} = \begin{pmatrix}
\bar{\alpha}_0 & \rho_0\bar{\alpha}_1 & \rho_0\rho_1 & 0 & 0 & \dots \\
\rho_0 & -\alpha_0\bar{\alpha}_1 & -\alpha_0\rho_1 & 0 & 0 & \dots \\
0 & \rho_1\bar{\alpha}_2 & -\alpha_1\bar{\alpha}_2 & \rho_2\bar{\alpha}_3 & \rho_2\rho_3 & \dots \\
0 & \rho_1\rho_2 & -\alpha_1\rho_2 & -\alpha_2\bar{\alpha}_3 & -\alpha_2\rho_3 & \dots \\
0 & 0 & 0 & \rho_3\bar{\alpha}_4 & -\alpha_3\bar{\alpha}_4 & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots
\end{pmatrix}.$$

*In other words, for any $j \geq 0$ even, the (generically) nonzero entries on the $j^{th}$ and $(j+1)^{st}$ rows of $\mathcal{C}$ are given by:*

$$\begin{aligned}
\mathcal{C}_{j,j-1} &= \rho_{j-1}\bar{\alpha}_j\,, \\
\mathcal{C}_{j,j} &= -\alpha_{j-1}\bar{\alpha}_j\,, \\
\mathcal{C}_{j,j+1} &= \rho_j\bar{\alpha}_{j+1}\,, \\
\mathcal{C}_{j,j+2} &= \rho_j\rho_{j+1}\,,
\end{aligned}$$

*for the $j^{th}$ row, and*

$$\begin{aligned}
\mathcal{C}_{j+1,j-1} &= \rho_{j-1}\rho_j \,, \\
\mathcal{C}_{j+1,j} &= -\alpha_{j-1}\rho_j \,, \\
\mathcal{C}_{j+1,j+1} &= -\alpha_j\bar{\alpha}_{j+1} \,, \\
\mathcal{C}_{j+1,j+2} &= -\alpha_j\rho_{j+1}
\end{aligned}$$

*for the $(j+1)^{st}$ row.*

*Remark* 3.8. Note that, in the previous statement, as well as everywhere from this point on, we consider all the indices to be greater than or equal to 0, and we set $\alpha_{-1} = -1$. In order to better understand how this boundary condition influences the form of $\mathcal{C}$, see Section 3.4.

*Proof.* Since

$$\mathcal{C}_{k,l} = (\chi_k, z\chi_l),$$

and the $\chi$'s can be expressed in terms of the $\phi$'s and $\phi^*$'s as in (3.3.1), the statement of this proposition reduces to computing certain inner products of polynomials.

For example, for $j \geq 0$,

$$\begin{aligned}
\mathcal{C}_{2j,2j+2} &= (\chi_{2j}, z\chi_{2j+2}) = (z^{-j}\phi^*_{2j}, zz^{-j-1}\phi^*_{2j+2}) \\
&= (\phi^*_{2j}, \phi^*_{2j+2}) = (\phi^*_{2j}(0), \phi^*_{2j+2}) \\
&= \Big(\prod_{l=0}^{2j-1}\rho_l\Big)^{-1}\Big(\prod_{l=0}^{2j+1}\rho_l\Big)^{-1}(1, \Phi^*_{2j+2}) \\
&= \frac{1}{\rho_{2j}\rho_{2j+1}}\Big(\prod_{l=0}^{2j-1}\rho_l\Big)^{-2}\|\Phi_{2j+2}\|^2 \\
&= \frac{1}{\rho_{2j}\rho_{2j+1}}\Big(\prod_{l=0}^{2j-1}\rho_l\Big)^{-2}\Big(\prod_{l=0}^{2j+1}\rho_l\Big)^{2} \\
&= \rho_{2j}\rho_{2j+1}.
\end{aligned}$$

Here we used the fact that $\phi^*_k$ is orthogonal to $z, \ldots, z^k$, as well as relations (3.2.1) and (3.2.4).

All the other expressions for the entries of $\mathcal{C}$ are proved in very much the same way, and using the recurrence relations (3.2.2) and (3.2.3), together with the identities that we used above. $\square$

There is an alternate way of describing $\mathcal{C}$, which is particularly illuminating. We will present it here without proof, mainly since proving it just means running several arguments similar to the ones in Proposition 3.7.

For $i, j \geq 0$, define

$$\mathcal{L}_{i,j} = (\chi_i(z), zx_j(z)) \quad \text{and} \quad \mathcal{M}_{i,j} = (x_i(z), \chi_j(z)).$$

Set

$$\Theta_k = \begin{bmatrix} \bar{\alpha}_k & \rho_k \\ \rho_k & -\alpha_k \end{bmatrix}.$$

Then

$$\mathcal{L} = \operatorname{diag}\big(\Theta_0, \Theta_2, \Theta_4, \dots\big) \quad \text{and} \quad \mathcal{M} = \operatorname{diag}\big([1], \Theta_1, \Theta_3, \dots\big), \tag{3.3.4}$$

and

$$\mathcal{C} = \mathcal{L}\mathcal{M}.$$

*Remark* 3.9. Let us note here that throughout the thesis we index rows and columns of matrices starting with 0: for example, $\mathcal{L}_{jj} = \bar{\alpha}_j$ for all $j \geq 0$, $j$ even. The (infinite) CMV matrix $\mathcal{C}$ is the matrix that we use in Section 6.1 to define Lax pairs for the flows generated by the Ablowitz-Ladik Hamiltonians on the coefficients $\alpha_j$, $j \geq 0$.

## 3.4 Measures with Finite Support

Throughout Sections 3.2 and 3.3 we assumed that the measure $\mu$ we were starting with had infinite support. Assume now that $\mu$ is a probability measure on $S^1$ with finite support; more precisely, let

$$\mu = \sum_{j=1}^{n} \mu_j \delta_{z_j}$$

be supported at $n$ points $z_1, \ldots, z_n \in S^1$, with

$$\sum_{j=1}^{n} \mu_j = 1.$$

Then the monomials $1, z, \ldots, z^{n-1}$ form a basis in $L^2(d\mu)$. To these we can apply the Gram-Schmidt procedure, and define the monic orthogonal polynomials $\{\Phi_j\}_{j=0}^{n-1}$ and the orthonormal polynomials

$$\phi_j = \frac{\Phi_j}{\|\Phi_j\|}$$

for $0 \leq j \leq n-1$.

Furthermore, we can consider $z^n \in L^2(d\mu)$ and define

$$\Phi_n(z) = z^n - \sum_{j=0}^{n-1} (z^n, \phi_j)\phi_j(z).$$

We obtain a monic polynomial of degree $n$, which is equal to zero in $L^2(d\mu)$, since $z^n \in \operatorname{span}(1, z, \ldots, z^{n-1}) = L^2(d\mu)$. In particular, this means that $\Phi_n$ is the unique monic polynomial of degree $n$ with zeroes at the mass points $z_1, \ldots, z_n$ of $\mu$.

The same reasoning as in the proof of Theorem 3.1 allows us to define $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{C}$ so that

$$\Phi_{j+1}(z) = z\Phi_j(z) - \bar{\alpha}_j \Phi_j^*(z)$$

for $0 \leq j \leq n-1$, and

$$\|\Phi_{j+1}\|^2 = (1 - |\alpha_j|^2)\|\Phi_j\|^2.$$

Since $\Phi_0, \ldots, \Phi_{n-1}$ are nonzero in $L^2(d\mu)$, we again obtain

$$\alpha_0, \ldots, \alpha_{n-2} \in \mathbb{D}.$$

But $\Phi_n = 0$ in $L^2(d\mu)$, and hence,

$$1 - |\alpha_{n-1}|^2 = \frac{\|\Phi_n\|^2}{\|\Phi_{n-1}\|^2} = 0$$

or

$$\alpha_{n-1} \in S^1.$$

If, as in the infinite case, we represent the operator of multiplication by $z$ in the basis considered by Cantero, Moral, and Velázquez, we obtain a finite CMV matrix

$$\mathcal{C}_f = \mathcal{L}_f \mathcal{M}_f \, .$$

Note that, since $|\alpha_{n-1}| = 1$,

$$\Theta_{n-1} = \begin{bmatrix} \bar{\alpha}_{n-1} & 0 \\ 0 & -\alpha_{n-1} \end{bmatrix}$$

decomposes as the direct sum of two $1 \times 1$ matrices. Hence, if we replace $\Theta_{n-1}$ by the $1 \times 1$ matrix that is its top left entry, $\bar{\alpha}_{n-1}$, and discard all $\Theta_m$ with $m \geq n$, we find that $\mathcal{L}_f$ and $\mathcal{M}_f$ are naturally $n \times n$ block-diagonal matrices. As in the infinite case, the finite CMV matrix $\mathcal{C}_f$ allows us to recast the Ablowitz-Ladik hierarchy of equations in Lax pair form.

## 3.5 Periodic Verblunsky Coefficients

The theory of periodic Verblunsky coefficients was first studied by Geronimus, and, more recently, by Peherstorfer and collaborators, and Golinskii and collaborators (for detailed references to their work, see [29]). Simon used the analogy with Hill's equation to fully develop the theory for periodic Verblunsky coefficients in [29, Chapter 11].

We are interested in sequences of Verblunsky coefficients $\{\alpha_j\}_{j\geq 0}$ which are periodic with period $p$:

$$\alpha_{j+p} = \alpha_j \quad \text{for all}\, j \geq 0.$$

These are completely described by their first $p$ terms, so from now on, whenever we talk about periodic Verblunsky coefficients, we will think of finite sets $\{\alpha_j\}_{j=0}^{p-1} \in \mathbb{D}^p$.

Let us first observe that in this case, besides the usual CMV matrix $\mathcal{C}$, one can also define a so-called extended CMV matrix that we shall denote by $\mathcal{E}$. If the $\alpha$'s are periodic with period $p$ even, that is, they obey $\alpha_{j+p} = \alpha_j$ for all $j \geq 0$, then we can define a two-sided infinite sequence of coefficients by periodicity. The extended CMV matrix is

$$\mathcal{E} = \tilde{\mathcal{L}}\tilde{\mathcal{M}},$$

where

$$\tilde{\mathcal{L}} = \bigoplus_{j \text{ even}} \Theta_j \qquad \text{and} \qquad \tilde{\mathcal{M}} = \bigoplus_{j \text{ odd}} \Theta_j, \tag{3.5.1}$$

with $\Theta_j$ defined on $l^2(\mathbb{Z})$ by

$$\Theta_j = \begin{bmatrix} \bar{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{bmatrix}$$

on the span of $\delta_j$ and $\delta_{j+1}$, and identically 0 otherwise. The extended CMV matrix $\mathcal{E}$ will play an important role in determining the Lax pairs associated with the Hamiltonian flows of the periodic Ablowitz-Ladik system.

Another very important notion defined in [29] is the *discriminant* $\Delta(z)$ naturally associated to this periodic problem. It is given by

$$\Delta(z) = z^{-p/2}\mathrm{Tr}(T_p(z)),$$

where $T_p(z)$ is the transfer matrix defined in (3.2.8). The form of $\Delta$ that we will use follows from (3.2.9):

$$\Delta(z) = z^{-p/2}M^{-1}[B_{p-1}(z) + zB^*_{p-1}(z)], \tag{3.5.2}$$

where

$$M = \prod_{j=0}^{p-1} \rho_j$$

is called the modulus. We note here that a related quantity, $K_0 = M^2$, will play a very important role for the periodic Ablowitz-Ladik system.

It will be very important for our study to link the discriminant $\Delta$ to the extended CMV matrix. In order to achieve this, note first that $\mathcal{E}$ acts boundedly on the space of bounded sequences $l^\infty$. Moreover, if $S$ is the $p$-shift

$$(Su)_m = u_{m+p}, \quad \text{for } u \in l^\infty,$$

then, by periodicity of the $\alpha$'s, we see that

$$S\mathcal{E} = \mathcal{E}S.$$

In particular, if $\beta \in S^1$ and we consider

$$X_\beta = \{u \in l^\infty \,|\, Su = \beta u\},$$

then $\mathcal{E}$ takes $X_\beta$ to itself:

$$\mathcal{E}(X_\beta) \subset X_\beta.$$

We can therefore define

$$\mathcal{E}(\beta) = \mathcal{E} \restriction X_\beta.$$

Moreover, if for $0 \leq j \leq p-1$ we consider $\delta_j(\beta) \in X_\beta$ given by

$$(\delta_j(\beta))_m = \beta^l \quad \text{for } m = lp + j$$

and 0 otherwise, then $\{\delta_j(\beta)\}_{j=0}^{p-1}$ is a basis in $X_\beta$ and we can represent $\mathcal{E}(\beta)$ in this basis as

$$\mathcal{E}(\beta) = \mathcal{L}_p \mathcal{M}_p(\beta),$$

with

$$\mathcal{L}_p = \begin{pmatrix} \Theta_0 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots & \\ & & & & \Theta_{p-2} \end{pmatrix}$$

and

$$\mathcal{M}_p(\beta) = \begin{pmatrix} -\alpha_{p-1} & & & & \rho_{p-1}\beta^{-1} \\ & \Theta_1 & & & \\ & & \ddots & & \\ & & & \Theta_{p-3} & \\ \rho_{p-1}\beta & & & & \bar{\alpha}_{p-1} \end{pmatrix}.$$

Note that we will do something very similar in Section 5.1, where we will define a restriction of $\mathcal{E}$, but considering $\beta = 1$ and the period a multiple of $p$ (for more details, see the explanations preceding equation (5.1.4)).

The relation between the discriminant $\Delta$ and the extended CMV matrix $\mathcal{E}$ is given by the following relation:

$$\det(z - \mathcal{E}(\beta)) = \Big(\prod_{j=0}^{p-1} \rho_j\Big) z^{p/2} \big[\Delta(z) - (\beta + \beta^{-1})^{-1}\big] \tag{3.5.3}$$

for all $\beta \in S^1$. For a proof of this statement, see Section 11.2 of [29].

Moreover:

**Proposition 3.10 (Simon).** *Let $p$ (the period of the coefficients) be even. Let $\{\alpha_j\}_{j=0}^{p-1}$ and $\{\gamma_j\}_{j=0}^{p-1}$ be two elements of $\mathbb{D}^p$. Then, the following are equivalent:*

1. $\Delta(z; \{\alpha_j\}) = \Delta(z; \{\gamma_j\})$.
2. $\prod_j (1-|\alpha_j|^2) = \prod_j (1-|\gamma_j|^2)$, *and the eigenvalues of* $\mathcal{E}_{(\beta)}(\{\alpha_j\}_{j=0}^{p-1})$ *and* $\mathcal{E}_{(\beta)}(\{\gamma_j\}_{j=0}^{p-1})$ *coincide for one* $\beta \in \partial\mathbb{D}$.
3. *The eigenvalues of* $\mathcal{E}_{(\beta)}(\{\alpha_j\}_{j=0}^{p-1})$ *and* $\mathcal{E}_{(\beta)}(\{\gamma_j\}_{j=0}^{p-1})$ *are equal for all* $\beta \in \partial\mathbb{D}$.

4. $\operatorname{spec}(\mathcal{E}(\{\alpha_j\}_{j=0}^{p-1})) = \operatorname{spec}(\mathcal{E}(\{\gamma_j\}_{j=0}^{p-1}))$.

When these conditions hold, we say that $\{\alpha_j\}_{j=0}^{p-1}$ and $\{\gamma_j\}_{j=0}^{p-1}$ are isospectral.

Next, we present two examples which represented a first step in establishing the connection between OPUC and the AL system. For the full computations which justify our claims, see Examples 11.1.4 and 11.1.5 in [29].

*Example* 3.11 *(Geronimus).* Let $\alpha \in \mathbb{D}$ and define $\alpha_j \equiv \alpha$ for all $j \geq 0$.

The isospectral manifold in this case is a circle

$$\{\alpha = (1-\rho^2)^{1/2} e^{i\theta} : \theta \in [0, 2\pi]\}$$

if $|\alpha| \neq 0$, and a point (or a zero-dimensional torus), $\alpha = 0$, if $|\alpha| = 0$.

*Example* 3.12 *(Akhiezer).* Consider $\alpha_{2j} = \alpha$ and $\alpha_{2j+1} = \alpha'$, with $\alpha, \alpha' \in \mathbb{D}$ and $j \geq 0$, to be periodic Verblunsky coefficients with period $p = 2$. Again, the discriminant is easily computable and

$$\Delta(e^{i\theta}) = \frac{2}{\rho\rho'}[\cos(\theta) + \operatorname{Re}(\bar{\alpha}\alpha')]. \tag{3.5.4}$$

Let $\theta_\pm \in [0, \pi)$ solve $\cos(\theta_\pm) = -\operatorname{Re}(\bar{\alpha}\alpha') \pm \rho\rho'$. Note that $|\operatorname{Re}(\bar{\alpha}\alpha')| + \rho\rho' \leq 1$, and hence, there are always solutions, with $0 \leq \theta_+ < \theta_- \leq \pi$. Thus, $|\Delta(e^{i\theta})| \leq 2$ if and only if $\pm\theta \in [\theta_+, \theta_-]$. We are interested in finding the set of pairs $(\alpha, \alpha') \in \mathbb{D}^2$ which lead to a given $\Delta$ of the form (3.5.4). This can be done explicitly, and the conclusion is:

- There are no open gaps for $|\alpha| = |\alpha'| = 0$, and so the isospectral manifold is a point (0-dimensional torus).
- There is exactly one open gap when $\alpha = \pm\alpha' \neq 0$, which leads to the isospectral manifold being a circle.
- There are two open gaps if and only if the isospectral manifold is a two-dimensional torus.

The two examples above suggest that $\mathbb{D}^p$ fibers into tori, generically of real dimension $p$, half of the real dimension of $\mathbb{D}^p$. This was proved by Simon in [29].

## 3.6 The Connection Between OPUC and OPRL

Let us now consider the case where the measure $d\mu$ is symmetric with respect to complex conjugation, or what is equivalent, where all Verblunsky coefficients are real. It is a famous observation of Szegő (see [31, §11.5]) that the polynomials orthogonal with respect to this measure are intimately related to the polynomials orthogonal with respect to the measure $d\nu$ on $[-2, 2]$ defined by

$$\int_{S^1} f(z + z^{-1})\, d\mu(z) = \int_{-2}^{2} f(x)\, d\nu(x). \tag{3.6.1}$$

The recurrence coefficients for these systems of orthogonal polynomials are related by Geronimus relations (see [17] and [18, Section 30]):

$$\begin{aligned} b_{k+1} &= (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2} \\ a_{k+1} &= \left\{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})\right\}^{1/2}. \end{aligned} \tag{3.6.2}$$

We will now present the short proof of these formulae given by Killip and Nenciu in [20]. As an off-shoot of our method, we also recover relations to the recurrence coefficients for $(4 - x^2)\, d\nu(x)$ and $(2 \pm x)d\nu$. The former appears in the proposition below, the latter in the remark that follows it. These formulae also appear in [6].

**Proposition 3.13.** *Let $\alpha_k$ be the system of real Verblunsky coefficients associated to a symmetric measure $d\mu$ and let $L$ and $M$ denote the matrices of* (3.3.4)*. Then $LM + ML$ is unitarily equivalent to the direct sum of two Jacobi matrices:*

$$J = \begin{bmatrix} b_1 & a_1 & 0 & \\ a_1 & b_2 & \ddots & \\ 0 & \ddots & \ddots & \end{bmatrix} \qquad \tilde{J} = \begin{bmatrix} \tilde{b}_1 & \tilde{a}_1 & 0 & \\ \tilde{a}_1 & \tilde{b}_2 & \ddots & \\ 0 & \ddots & \ddots & \end{bmatrix}$$

*where $a_k$ and $b_k$ are as in* (3.6.2) *and*

$$\begin{aligned}\tilde{b}_{k+1} &= (1-\alpha_{2k+1})\alpha_{2k} - (1+\alpha_{2k+1})\alpha_{2k+2} \\ \tilde{a}_{k+1} &= \big\{(1+\alpha_{2k+1})(1-\alpha_{2k+2}^2)(1-\alpha_{2k+3})\big\}^{1/2}.\end{aligned}$$

*Moreover, the spectral measure for* $(J, e_1)$ *is precisely the* $d\nu$ *of* (3.6.1). *The spectral measure for* $(\tilde{J}, e_1)$ *is* $\frac{1}{2(1-\alpha_0^2)(1-\alpha_1)}(4-x^2)\,d\nu(x)$.

*Proof.* Let $S$ denote the following unitary block matrix

$$S = \operatorname{diag}([1], S_1, S_3, \dots) \quad \text{where} \quad S_k = \frac{1}{\sqrt{2}}\begin{bmatrix} -\sqrt{1-\alpha_k} & \sqrt{1+\alpha_k} \\ \sqrt{1+\alpha_k} & \sqrt{1-\alpha_k}\end{bmatrix},$$

which is easily seen to diagonalize $M$. Indeed, $S^\dagger M S = \operatorname{diag}(+1, -1, +1, -1, \dots)$. We will denote this matrix by $R$.

The matrix $LM + ML$ is unitarily equivalent to $A = S^\dagger(LM+ML)S = S^\dagger LSR + RS^\dagger LS$, which we will show is the direct sum of two Jacobi matrices. We begin by showing that even-odd and odd-even entries of $A$ vanish, from which it follows that $A$ is the direct sum of its even-even and odd-odd submatrices.

Left multiplication by $R$ changes the sign of the entries in each even-numbered row, while right multiplication by $R$ reverses the sign of each even-numbered column. In this way, $RB + BR$ has the stated direct sum structure for any matrix $B$ and hence, in particular, for $B = S^\dagger LS$.

It remains only to calculate the nonzero entries of $A$. As $S$ and $L$ are both tri-diagonal, $A$ must be hepta-diagonal and so the direct sum of tri-diagonal matrices. Moreover, $A$ is symmetric (because $L$ is) so there are only four categories of entries to calculate: the odd/even diagonals and the odd/even off-diagonals. We begin with the diagonals:

$$\begin{aligned}
A_{2k+1,2k+1} &= \begin{bmatrix}\sqrt{1+\alpha_{2k-1}} & \sqrt{1-\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}-\alpha_{2k-2} & 0\\ 0 & \alpha_{2k}\end{bmatrix}\begin{bmatrix}\sqrt{1+\alpha_{2k-1}}\\ \sqrt{1-\alpha_{2k-1}}\end{bmatrix}\\
&= (1-\alpha_{2k-1})\alpha_{2k} - (1+\alpha_{2k-1})\alpha_{2k-2}\\
A_{2k,2k} &= -\begin{bmatrix}-\sqrt{1-\alpha_{2k-1}} & \sqrt{1+\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}-\alpha_{2k-2} & 0\\ 0 & \alpha_{2k}\end{bmatrix}\begin{bmatrix}-\sqrt{1-\alpha_{2k-1}}\\ \sqrt{1+\alpha_{2k-1}}\end{bmatrix}\\
&= (1-\alpha_{2k-1})\alpha_{2k-2} - (1+\alpha_{2k-1})\alpha_{2k}.
\end{aligned}$$

Note that the factor of 2 resulting from $A$ being the sum of two terms is cancelled by the factors of $2^{-1/2}$ coming from $S$ and $S^\dagger$. The calculation of the off-diagonal terms proceeds in a similar fashion:

$$\begin{aligned}
A_{2k+1,2k+3} &= \begin{bmatrix}\sqrt{1+\alpha_{2k-1}} & \sqrt{1-\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}0 & 0\\ \rho_{2k} & 0\end{bmatrix}\begin{bmatrix}\sqrt{1+\alpha_{2k+1}}\\ \sqrt{1-\alpha_{2k+1}}\end{bmatrix}\\
&= \sqrt{(1-\alpha_{2k-1})(1-\alpha_{2k}^2)(1+\alpha_{2k+1})}\\
A_{2k,2k+2} &= -\begin{bmatrix}-\sqrt{1-\alpha_{2k-1}} & \sqrt{1+\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}0 & 0\\ \rho_{2k} & 0\end{bmatrix}\begin{bmatrix}-\sqrt{1-\alpha_{2k+1}}\\ \sqrt{1+\alpha_{2k+1}}\end{bmatrix}\\
&= \sqrt{(1+\alpha_{2k-1})(1-\alpha_{2k}^2)(1-\alpha_{2k+1})}.
\end{aligned}$$

That $d\nu$ is the spectral measure for $(J, e_1)$ is an immediate consequence of the spectral theorem, $LM + ML = LM + (LM)^{-1}$, and the fact that $S$ leaves the vector $[1, 0, \ldots, 0]$ invariant.

Tracing back through the definitions, we find that the spectral measure for $(\tilde{J}, e_1)$ is equal to that for the operator $f(z) \mapsto (z + z^{-1})f(z)$ in $L^2(d\mu)$ and the vector

$$f(z) = \left(\tfrac{1+\alpha_1}{2}\right)^{\frac{1}{2}} \chi_2(z) - \left(\tfrac{1-\alpha_1}{2}\right)^{\frac{1}{2}} \chi_1(z) = \left(\tfrac{1+\alpha_1}{2}\right)^{\frac{1}{2}} z^{-1}\phi_2^*(z) - \left(\tfrac{1-\alpha_1}{2}\right)^{\frac{1}{2}} \phi_1(z).$$

From the relations (3.2.2), (3.2.3), and (3.2.4), we find $\rho_1\phi_2^*(z) = \phi_1^*(z) - \alpha_1 z\phi_1(z)$,

$\rho_0\phi_1(z) = z - \alpha_0$, and $\rho_0\phi_1^*(z) = 1 - \alpha_0 z$. These simplify the formula considerably:

$$f(z) = \frac{z^{-1} - z}{\rho_0\sqrt{2(1-\alpha_1)}}.$$

The expression for the spectral measure for $(\tilde{J}, e_1)$ now follows from the simple calculation $|z^{-1} - z|^2 = 4 - (z + z^{-1})^2$. □

*Remark* 3.14. In the above proof, we conjugated $LM + ML$ by the unitary matrix which diagonalizes $M$. One may instead use the matrix $\mathrm{diag}(S_0, S_2, \dots)$, which diagonalizes $L$. This also conjugates $LM + ML$ to the direct sum of two Jacobi matrices. In this way, we learn that the recurrence coefficients for $\frac{1}{2(1\pm\alpha_0)}(2 \pm x)d\nu(x)$ are given by

$$\begin{aligned} b_{k+1} &= \pm(1 \mp \alpha_{2k})\alpha_{2k+1} \mp (1 \pm \alpha_{2k})\alpha_{2k-1} \\ a_{k+1} &= \left\{(1 \mp \alpha_{2k})(1 - \alpha_{2k+1}^2)(1 \pm \alpha_{2k+2})\right\}^{1/2}. \end{aligned}$$

# Chapter 4

# The Periodic Ablowitz-Ladik System

## 4.1 Commutativity of Discriminants

We begin by defining the symplectic structure. We are considering the problem of periodic Verblunsky coefficients with period $p$, so we are interested in a symplectic form on $\mathbb{D}^p$, which has real dimension $2p$. Let $\underline{\alpha} = (\alpha_0, \dots, \alpha_{p-1}) \in \mathbb{D}^p$, and let $u_j = \operatorname{Re} \alpha_j$ and $v_j = \operatorname{Im} \alpha_j$ for all $0 \leq j \leq p-1$. Then we define our symplectic form by

$$\omega = \frac{1}{2} \sum_{j=0}^{p-1} \frac{1}{\rho_j^2} \, du_j \wedge dv_j. \tag{4.1.1}$$

As all of the subsequent computations will involve only the corresponding Poisson bracket, let us note that, for $f$ and $g$ functions on $\mathbb{D}^p$, we have

$$\{f, g\} = \frac{1}{2} \sum_{j=0}^{p-1} \rho_j^2 \left[ \frac{\partial f}{\partial u_j} \frac{\partial g}{\partial v_j} - \frac{\partial f}{\partial v_j} \frac{\partial g}{\partial u_j} \right] \tag{4.1.2}$$

$$= i \sum_{j=0}^{p-1} \rho_j^2 \left[ \frac{\partial f}{\partial \bar{\alpha}_j} \frac{\partial g}{\partial \alpha_j} - \frac{\partial f}{\partial \alpha_j} \frac{\partial g}{\partial \bar{\alpha}_j} \right], \tag{4.1.3}$$

where for $z = u + iv \in \mathbb{D}$ we use the standard notation

$$\frac{\partial}{\partial z} = \frac{1}{2} \left( \frac{\partial}{\partial u} - i \frac{\partial}{\partial v} \right) \quad \text{and} \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left( \frac{\partial}{\partial u} + i \frac{\partial}{\partial v} \right).$$

**Lemma 4.1.** *The 2-form defined by* (4.1.1) *is a symplectic form. Equivalently, the bracket* (4.1.2) *obeys the Jacobi identity and is nondegenerate.*

*Proof.* The form $\omega$ is a sum of 2-forms, each of which acts only on one of the variables $\alpha_j$ for $0 \le j \le p-1$. But any 2-form is closed in $\mathbb{R}^2$, and hence $\omega$ is closed. It is also nondegenerate, since the function $\rho_j^{-2}$ is positive on $\mathbb{D}^p$ for each $j$. □

The first result is

**Theorem 4.2 (Nenciu–Simon).** *With the above Poisson bracket, we have*

$$\{\Delta(z), \Delta(w)\} = 0 \tag{4.1.4}$$

*for any $z, w \in \mathbb{C}$.*

In particular, one has

**Corollary 4.3.** *The Hamiltonian flows generated by $\Delta(z)$ for $z \in \partial\mathbb{D}$ and by $\prod_{j=0}^{p-1} \rho_j$ all commute with each other and leave $\Delta(w)$ invariant.*

Theorem 4.2 is proved using the expression of $\Delta$ in terms of Wall polynomials and the recurrence relations that they obey. More precisely, note first that, under the flow generated by

$$K_0 = \prod_{j=0}^{p-1} \rho_j^2,$$

the $\alpha$'s evolve according to

$$\{\alpha_j, K_0\} = iK_0\alpha_j$$

or

$$\alpha_j(t) = e^{iK_0 t}\alpha_j(0).$$

This immediately shows that

$$\{\Delta(z), K_0\} = 0.$$

Since we know from (3.5.2) that

$$\Delta(z) = z^{-p/2} K_0^{-1/2}[B_{p-1}(z) + zB_{p-1}^*(z)],$$

we can conclude (4.1.4) once we prove that

$$\{B_n(z), B_n(w)\} = \{zB_n^*(z), wB_n^*(w)\} = 0 \tag{4.1.5}$$

and

$$\{B_n(z), wB_n^*(w)\} + \{zB_n^*(z), B_n(w)\} = 0. \tag{4.1.6}$$

*Remark* 4.4. We will prove these statements by induction on $n$. Before we begin the actual proofs, note that such a strategy makes sense: Let us be more precise in our notation and use

$$\{f, g\}_{(m)} = i \sum_{j=0}^{m} \rho_j^2 \left[ \frac{\partial f}{\partial \bar{\alpha}_j} \frac{\partial g}{\partial \alpha_j} - \frac{\partial f}{\partial \alpha_j} \frac{\partial g}{\partial \bar{\alpha}_j} \right]$$

for the Poisson bracket on $\mathbb{D}^{m+1}$. It is a simple observation that if $f$ and $g$ depend only on $\alpha_0, \dots, \alpha_n$ and $m \geq n$, then

$$\{f, g\}_{(m)} = \{f, g\}_{(n)}.$$

In particular, as $B_n(z)$ is a polynomial in $\alpha_0, \dots, \alpha_n$, we can use in (4.1.5) and (4.1.6) any bracket that involves "enough" $\alpha$'s.

**Proposition 4.5.** *For $n \geq 0$ and $z, w \in \mathbb{C}$, define*

$$F_n(z, w) = -i\{A_n^*(z), B_n(w)\} \tag{4.1.7}$$

*and*

$$Q_n(z, w) = zw \cdot \frac{z^{n-1} - w^{n-1}}{z - w}.$$

*Then the following statements hold for all $n \geq 0$ and $z, w \in \mathbb{C}$:*

$(\alpha_n)$ $\quad \{B_n(z), B_n(w)\} = 0$

$(\beta_n)$ $\quad \{A_n^*(z), A_n^*(w)\} = 0$

$(\gamma_{n,q})$ $\quad$ *For* $q \geq 1$,

$$z^q F_n(z, w) - w^q F_n(w, z) = Q_q(z, w)[A_n^*(z)B_n(w) - A_n^*(w)B_n(z)]$$

*Remark* 4.6. Note that $(\alpha_n)$ implies (4.1.5). Indeed, if $f$ and $g$ are two complex-valued functions, then

$$\begin{aligned}\overline{\{f,g\}} &= \overline{\frac{1}{2}\sum_{j=0}^{p-1}\rho_j^2\left[\frac{\partial f}{\partial u_j}\frac{\partial g}{\partial v_j} - \frac{\partial f}{\partial v_j}\frac{\partial g}{\partial u_j}\right]} \\ &= \{\bar{f},\bar{g}\}.\end{aligned}$$

In particular,

$$\begin{aligned}\{zB_n^*(z), wB_n^*(w)\} &= z^{n+1}w^{n+1}\left\{\overline{B_n(\tfrac{1}{\bar{z}})}, \overline{B_n(\tfrac{1}{\bar{w}})}\right\} \\ &= z^{n+1}w^{n+1}\overline{\{B_n(\tfrac{1}{\bar{z}}), B_n(\tfrac{1}{\bar{w}})\}} \\ &= 0\end{aligned}$$

by $(\alpha_n)$ applied to $\frac{1}{\bar{z}}$ and $\frac{1}{\bar{w}}$.

*Proof.* For $n = 0$, recall that $B_0(z) = 1$ and $A_0^*(z) = \bar{\alpha}_0$. Since $B_0$ is constant (as a function of the Verblunsky coefficients), statement $(\alpha_0)$ is immediate. $(\beta_0)$ is merely stating that $\{\alpha_0, \alpha_0\} = 0$, while $(\gamma_{0,q})$ holds since $F_0(z,w) = 0$ and neither $A_0^*$ nor $B_0$ depend on $z$ or $w$.

Assume that $n \geq 1$ and the statements $(\alpha_{n-1})$–$(\gamma_{n-1,q})$ are known. Recall that the Wall polynomials obey recurrence relations (see (3.2.10) and (3.2.11)):

$$B_n(z) = B_{n-1}(z) + \alpha_n z A_{n-1}^*(z)$$

and

$$A_n^*(z) = zA_{n-1}^*(z) + \bar{\alpha}_n B_{n-1}(z).$$

We will prove $(\alpha_n)$–$(\gamma_{n,q})$ one by one, by plugging in these recurrence relations and using the induction hypothesis.

<u>$(\alpha_n)$ :</u> We use the recurrence relation for $B_n$, and in the expression for the Poisson

bracket we separate the $\alpha_n$-derivatives from the other $\alpha_0$- through $\alpha_{n-1}$-derivatives:

$$\begin{aligned}\{B_n(z), B_n(w)\} =&\{B_{n-1}(z), B_{n-1}(w)\} + (\alpha_n z)\cdot(\alpha_n w)\{A^*_{n-1}(z), A^*_{n-1}(w)\}\\ &+ \alpha_n\big[z\{A^*_{n-1}(z), B_{n-1}(w)\} + w\{B_{n-1}(z), A^*_{n-1}(w)\}\big].\end{aligned}$$

Here we use the fact that all the other terms in the expansion of the right-hand side contain factors of the type $\{\alpha_n, B_{n-1}\}$, $\{\alpha_n, A^*_{n-1}\}$, which are zero as both $A^*_{n-1}$ and $B_{n-1}$ depend only on $\alpha_k$ with $k \le n-1$, or $\{\alpha_n, \alpha_n\}$. Considering the right-hand side now, observe that the first two terms are zero by $(\alpha_{n-1})$ and $(\beta_{n-1})$, while the third term equals

$$i\alpha_n\big[zF_{n-1}(z,w) - wF_{n-1}(w,z)\big] = 0$$

by (4.1.7) and $(\gamma_{n-1,1})$.

<u>$(\beta_n)$ :</u> We prove the statement in the same way as $(\alpha_n)$. Indeed, from the recurrence we get

$$\begin{aligned}\{A^*_n(z), A^*_n(w)\} =&zw\{A^*_{n-1}(z), A^*_{n-1}(w)\} + \bar\alpha_n^2\{B_{n-1}(z), B_{n-1}(w)\}\\ &+ \bar\alpha_n\big[z\{A^*_{n-1}(z), B_{n-1}(w)\} + w\{B_{n-1}(z), A^*_{n-1}(w)\}\big],\end{aligned}$$

and we use $(\alpha_{n-1})$, $(\beta_{n-1})$, and $(\gamma_{n-1,1})$ to conclude $(\beta_n)$.

Before proceeding to prove the last statement, we will deduce a recurrence formula for $F_n$:

$$F_n(z,w) = \rho_n^2\big[zF_{n-1}(z,w) + wA^*_{n-1}(w)B_{n-1}(z)\big]. \tag{4.1.8}$$

Indeed, note that

$$\begin{aligned}F_n(z,w) &= -i\{A^*_n(z), B_n(w)\}\\ &= -i\{\bar\alpha_n, \alpha_n\}wB_{n-1}(z)A^*_{n-1}(w) + zF_{n-1}(z,w) + w|\alpha_n|^2F_{n-1}(w,z)\\ &= \rho_n^2 wA^*_{n-1}(w)B_{n-1}(z) + \rho_n^2 zF_{n-1}(z,w),\end{aligned}$$

where we used the recurrence formulae for $A^*_n$ and $B_n$, and induction hypotheses $(\alpha_{n-1})$, $(\beta_{n-1})$, and $(\gamma_{n-1,1})$.

$\underline{(\gamma_{n,q})}$ : Using the recurrence formula (4.1.8), we see that

$$\begin{aligned}\rho_n^{-2}\big[z^q F_n(z,w) - w^q F_n(w,z)\big] =& z^{q+1}F_{n-1}(z,w) - w^{q+1}F_{n-1}(w,z)\\ &+ z^q w A_{n-1}^*(w)B_{n-1}(z) - w^q z A_{n-1}^*(z)B_{n-1}(w).\end{aligned}$$

From the induction hypothesis $(\gamma_{n-1,q+1})$, we have

$$z^{q+1}F_{n-1}(z,w) - w^{q+1}F_{n-1}(w,z) = zw\frac{z^q - w^q}{z-w}\big[A_{n-1}^*(z)B_{n-1}(w) - A_{n-1}^*(w)B_{n-1}(z)\big].$$

Plugging this into the previous formula, we get

$$\begin{aligned}\rho_n^{-2}\big[z^q F_n(z,w) - w^q F_n(w,z)\big] =& A_{n-1}^*(z)B_{n-1}(w)\left(zw\frac{z^q - w^q}{z-w} - w^q z\right)\\ &- A_{n-1}^*(w)B_{n-1}(z)\left(zw\frac{z^q - w^q}{z-w} - z^q w\right).\end{aligned}$$

Note that

$$zw\frac{z^q - w^q}{z-w} - w^q z = z^2 w\frac{z^{q-1} - w^{q-1}}{z-w} = zQ_q(z,w)$$

and

$$zw\frac{z^q - w^q}{z-w} - z^q w = w^2 z\frac{z^{q-1} - w^{q-1}}{z-w} = wQ_q(z,w).$$

So we can conclude that

$$\rho_n^{-2}\big[z^q F_n(z,w) - w^q F_n(w,z)\big] = Q_q(z,w)\big[zA_{n-1}^*(z)B_{n-1}(w) - wA_{n-1}^*(w)B_{n-1}(z)\big].$$

It remains to treat the right-hand side. Here we use the recurrence relations (3.2.10) and (3.2.11) to get

$$A_n^*(z)B_n(w) - A_n^*(w)B_n(z) = \rho_n^2\big[zA_{n-1}^*(z)B_{n-1}(w) - wA_{n-1}^*(w)B_{n-1}(z)\big].$$

The last two equations imply the statement of $(\gamma_{n,q})$:

$$z^q F_n(z,w) - w^q F_n(w,z) = Q_q(z,w)[A_n^*(z)B_n(w) - A_n^*(w)B_n(z)].$$

□

We now turn our attention to proving (4.1.6). This is also achieved by induction, but a more involved one. First, we introduce some extra notation:

$$\begin{aligned} R_n(z,w) &= i\{B_n^*(z), B_n(w)\} \\ S_n(z,w) &= i\{A_n(z), A_n^*(w)\} \\ X_n(z,w) &= i\{B_n^*(z), A_n^*(w)\} \\ Y_n(z,w) &= i\{A_n(z), B_n(w)\} \end{aligned}$$

and deduce recurrence relations for these quantities. For all $n \geq 1$,

$$\begin{aligned} R_n(z,w) &= i\{B_n^*(z), B_n(w)\} \\ &= i\{zB_{n-1}^*(z) + \bar{\alpha}_n A_{n-1}(z), B_{n-1}(w) + \alpha_n w A_{n-1}^*(w)\} \\ &= zR_{n-1}(z,w) + |\alpha_n|^2 w S_{n-1}(z,w) + \alpha_n zw X_{n-1}(z,w) \\ &\quad + \bar{\alpha}_n Y_{n-1}(z,w) - \rho_n^2 w A_{n-1}(z) A_{n-1}^*(w), \end{aligned} \tag{4.1.9}$$

$$\begin{aligned} S_n(z,w) &= i\{A_n(z), A_n^*(w)\} \\ &= i\{A_{n-1}^*(z) + \alpha_n z B_{n-1}^*(z), wA_{n-1}^*(w) + \bar{\alpha}_n B_{n-1}(w)\} \\ &= |\alpha_n|^2 z R_{n-1}(z,w) + wS_{n-1}(z,w) + \alpha_n zw X_{n-1}(z,w) \\ &\quad + \bar{\alpha}_n Y_{n-1}(z,w) + \rho_n^2 z B_{n-1}^*(z) B_{n-1}(w), \end{aligned} \tag{4.1.10}$$

$$\begin{aligned} X_n(z,w) &= i\{B_n^*(z), A_n^*(w)\} \\ &= i\{zB_{n-1}^*(z) + \bar{\alpha}_n A_{n-1}(z), wA_{n-1}^*(w) + \bar{\alpha}_n B_{n-1}(w)\} \\ &= \bar{\alpha}_n z R_{n-1}(z,w) + \bar{\alpha}_n w S_{n-1}(z,w) \\ &\quad + zw X_{n-1}(z,w) + \bar{\alpha}_n^2 Y_{n-1}(z,w), \end{aligned} \tag{4.1.11}$$

$$
\begin{aligned}
Y_n(z,w) &= i\{A_n(z), B_n(w)\} \\
&= i\{A_{n-1}(z) + \alpha_n z B^*_{n-1}(z), B_{n-1}(w) + \alpha_n w A^*_{n-1}(w)\} \\
&= \alpha_n z R_{n-1}(z,w) + \alpha_n w S_{n-1}(z,w) \\
&\quad + \alpha_n^2 zw X_{n-1}(z,w) + Y_{n-1}(z,w).
\end{aligned}
\tag{4.1.12}
$$

We can now state our result:

**Proposition 4.7.** *The following statements hold for all $n \geq 0$, $q \in \mathbb{Z}$, and $z, w \neq 0$*

$$
\begin{aligned}
&(r_{n,q}) && z^q R_n(z,w) - w^q R_n(w,z) = Q_q(z,w)\big[A_n(z)A^*_n(w) - A_n(w)A^*_n(z)\big] \\
&(s_{n,q}) && z^q S_n(z,w) - w^q S_n(w,z) = -\big[z^q A_n(z)A^*_n(w) - w^q A_n(w)A^*_n(z)\big] \\
& && \qquad + \frac{z^q - w^q}{z - w}\big[zB^*_n(z)B_n(w) - wB^*_n(w)B_n(z)\big] \\
&(x_{n,q}) && z^q X_n(z,w) - w^q X_n(w,z) = Q_q(z,w)\big[B^*_n(z)A^*_n(w) - B^*_n(w)A^*_n(z)\big] \\
&(y_{n,q}) && z^q Y_n(z,w) - w^q Y_n(w,z) = Q_q(z,w)\big[A_n(z)B_n(w) - A_n(w)B_n(z)\big].
\end{aligned}
$$

*Remark* 4.8. As with the previous proposition, note that relation (4.1.6), which we set out to prove, is exactly $(r_{n,1})$. As the proof will show, all the other relations in the statement of the proposition are necessary in order to prove $(r_{n,1})$ by induction on $n$.

*Proof.* First, we deal with $n = 0$: Recall that $B_0(z) = B^*_0(z) = 1$, $A_0(z) = \alpha_0$, and $A^*_0(z) = \bar{\alpha}_0$. Therefore, we immediately find

$$
R_0(z,w) = X_0(z,w) = Y_0(z,w) = 0 \quad \text{and} \quad S_0(z,w) = \rho_0^2.
$$

The right-hand side of relations $(r_{0,q})$, $(x_{0,q})$, and $(y_{0,q})$ is identically equal to zero, as $B_0$, $B^*_0$, $A_0$, and $A^*_0$ are all constant (as polynomials in $z$).

It remains to consider $(s_{0,q})$. It follows by a simple computation that

$$\begin{aligned}\operatorname{rhs}(s_{0,q}) &= -\big[z^q|\alpha_0|^2 - w^q|\alpha_0|^2\big] + \frac{z^q - w^q}{z-w}\big[z\cdot 1 - w\cdot 1\big] \\ &= z^q\rho_0^2 - w^q\rho_0^2 = \operatorname{lhs}(s_{0,q}).\end{aligned}$$

Now assume that all the induction statements hold for some $n$, $n \geq 0$. We prove them for $n+1$. We use the recurrence formulae (4.1.9), (4.1.10), (4.1.11), and (4.1.12) throughout these computations without mentioning them.

We start with relation $(r_{n+1,q})$. In this case we get, using the induction hypotheses,

$$\begin{aligned}\operatorname{lhs}(r_{n+1,q}) &= z^q R_{n+1}(z,w) - w^q R_{n+1}(w,z) \\ &= \operatorname{lhs}(r_{n,q+1}) + \alpha_{n+1} zw \cdot \operatorname{lhs}(x_{n,q}) \\ &\quad + \bar{\alpha}_{n+1}\cdot \operatorname{lhs}(y_{n,q}) + |\alpha_{n+1}|^2 zw \cdot \operatorname{lhs}(s_{n,q-1}) \\ &\quad - \rho_{n+1}^2 zw\big[z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)\big] \\ &= Q_{q+1}(z,w)\big[A_n(z)A_n^*(w) - A_n(w)A_n^*(z)\big] \\ &\quad + \alpha_{n+1} zw Q_q(z,w)\big[B_n^*(z)A_n^*(w) - B_n^*(w)A_n^*(z)\big] \\ &\quad + \bar{\alpha}_{n+1} Q_q(z,w)\big[A_n(z)B_n(w) - A_n(w)B_n(z)\big] \\ &\quad + |\alpha_{n+1}|^2 zw\big[z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)\big] \\ &\quad - \rho_{n+1}^2 zw\big[z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)\big].\end{aligned}$$

Group together the terms in the last identity that contain $A_n(z)A_n^*(w)$, recalling that

$$\rho_{n+1}^2 = 1 - |\alpha_{n+1}|^2 \quad \text{and} \quad Q_{q+1}(z,w) = zw\frac{z^q - w^q}{z-w}.$$

Doing this, we obtain that the coefficient that multiplies $A_n(z)A_n^*(w)$ is

$$zw\frac{z^q - w^q}{z-w} - zwz^{q-1} = zw\cdot w\frac{z^{q-1} - w^{q-1}}{z-w} = wQ_q(z,w).$$

Similarly, the coefficient multiplying $A_n(w)A_n^*(z)$ is $zQ_q(z,w)$. Plugging this into the identity above, we obtain

$$\begin{aligned}
\mathrm{lhs}(r_{n+1,q}) &= z^q R_{n+1}(z,w) - w^q R_{n+1}(w,z) \\
&= Q_q(z,w)\big[wA_n(z)A_n^*(w) - zA_n(w)A_n^*(z) \\
&\qquad + \alpha_{n+1}zwB_n^*(z)A_n^*(w) - \alpha_{n+1}zwB_n^*(w)A_n^*(z) \\
&\qquad + \bar{\alpha}_{n+1}A_n(z)B_n(w) - \bar{\alpha}_{n+1}A - n(w)B_n(z) \\
&\qquad + |\alpha_{n+1}|^2 zB_n^*(z)B_n(w) - |\alpha_{n+1}|^2 B_n^*(w)B_n(z)\big] \\
&= Q_q(z,w)\big[A_{n+1}(z)A_{n+1}^*(w) - A_{n+1}(w)A_{n+1}^*(z)\big] \\
&= \mathrm{rhs}(r_{n+1,q}),
\end{aligned}$$

as claimed. In the last sequence of identities we have used the recurrence relations for the Wall polynomials without giving any details.

We now turn to proving $(s_{n,q})$, which requires somewhat more involved computations. From the recurrence relation (4.1.10), we get

$$\begin{aligned}
\mathrm{lhs}(s_{n+1,q}) &= z^q S_{n+1}(z,w) - w^q S_{n+1}(w,z) \\
&= |\alpha_{n+1}|^2 \cdot \mathrm{lhs}(r_{n,q+1}) + \alpha_{n+1}zw \cdot \mathrm{lhs}(x_{n,q}) \\
&\quad + \bar{\alpha}_{n+1} \cdot \mathrm{lhs}(y_{n,q}) + zw \cdot \mathrm{lhs}(s_{n,q-1}) \\
&\quad + \rho_{n+1}^2\big[z^{q+1}B_n^*(z)B_n(w) - w^{q+1}B_n^*(w)B_n(z)\big] \\
&= Q_q(z,w)\big[zB_n^*(z)B_n(w) - wB_n^*(w)B_n(z)\big] \\
&\quad - zw\big[z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)\big] \\
&\quad + \alpha_{n+1}zwQ_q(z,w)\big[B_n^*(z)A_n^*(w) - B_n^*(w)A_n^*(z)\big] \\
&\quad + \bar{\alpha}_{n+1}Q_q(z,w)\big[A_n(z)B_n(w) - A_n(w)B_n(z)\big] \\
&\quad + |\alpha_{n+1}|^2 Q_{q+1}(z,w)\big[A_n(z)A_n^*(w) - A_n(w)A_n^*(z)\big] \\
&\quad + \rho_{n+1}^2\big[z^{q+1}B_n^*(z)B_n(w) - w^{q+1}B_n^*(w)B_n(z)\big].
\end{aligned}$$

Using the recurrence relations for $B_{n+1}$ and $B_{n+1}^*$, we find that the right-hand side rhs$(s_{n+1,q})$ equals

$$\begin{aligned}
\frac{z^q - w^q}{z - w}\big[&z^2 B_n^*(z)B_n(w) - w^2 B_n^*(w)B_n(z) \\
&+ \alpha_n z^2 w B_n^*(z)A_n^*(w) - \alpha_n z w^2 B_n^*(w)A_n^*(z) \\
&+ \bar{\alpha}_n z A_n(z)B_n(w) - \bar{\alpha}_n w A_n(w)B_n(z) \\
&+ |\alpha_n|^2 zw(A - n(z)A_n^*(w) - a_n(w)A_n^*(z))\big] \\
-\big[&zw(z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)) \\
&+ \alpha_n zw(z^q B_n^*(z)A_n^*(w) - w^q B_n^*(w)A_n^*(z)) \\
&+ \bar{\alpha}_n(z^q A_n(z)B_n(w) - w^q A_n(w)B_n(z)) \\
&+ |\alpha_n|^2(z^{q+1}B_n^*(z)B_n(w) - w^{q+1}B_n^*(w)B_n(z))\big].
\end{aligned}$$

Moreover, if we group terms together according to which of the Wall polynomials they contain, we get that rhs$(s_{n+1,q})$ further equals

$$\begin{aligned}
&- zw\big[z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)\big] \\
&+ \alpha_{n+1} zw\Big[B_n^*(z)A_n^*(w)\big(z \cdot \frac{z^q - w^q}{z - w} - z^q\big) - B_n^*(w)A_n^*(z)\big(w \cdot \frac{z^q - w^q}{z - w} - w^q\big)\Big] \\
&+ \bar{\alpha}_{n+1}\Big[A_n(z)B_n(w)\big(z \cdot \frac{z^q - w^q}{z - w} - z^q\big) - A_n(w)B_n(z)\big(w \cdot \frac{z^q - w^q}{z - w} - w^q\big)\Big] \\
&+ \Big[B_n^*(z)B_n(w)\big(z^2 \cdot \frac{z^q - w^q}{z - w} - z^{q+1}\big) - B_n^*(w)B_n(z)\big(w^2 \cdot \frac{z^q - w^q}{z - w} - w^{q+1}\big)\Big] \\
&+ |\alpha_{n+1}|^2 Q_{q+1}(z, w)\big[A_n(z)A_n^*(w) - A_n(w)A_n^*(z)\big] \\
&+ \rho_{n+1}^2\big[z^{q+1}B_n^*(z)B_n(w) - w^{q+1}B_n^*(w)B_n(z)\big].
\end{aligned}$$

If we also recall that

$$z \cdot \frac{z^q - w^q}{z - w} - z^q = w \cdot \frac{z^q - w^q}{z - w} - w^q = Q_q(z, w),$$

we can immediately conclude that rhs$(s_{n+1,q}) =$ lhs$(s_{n+1,q})$.

The proofs of $(x_{n+1,q})$ and $(y_{n+1,q})$ follow the same pattern, while being even easier than the ones above. We present them without giving too many explanations:

$$
\begin{aligned}
\mathrm{lhs}(x_{n+1,q}) = {} & z^q X_{n+1}(z,w) - w^q X_{n+1}(w,z) \\
= {} & \bar{\alpha}_{n+1} \cdot \mathrm{lhs}(r_{n,q+1}) + \bar{\alpha}_{n+1} zw \cdot \mathrm{lhs}(s_{n-1,q-1}) \\
& + zw \cdot \mathrm{lhs}(x_{n,q}) + \bar{\alpha}_{n+1}^2 \cdot \mathrm{lhs}(y_{n,q}) \\
= {} & zw Q_q(z,w)\big[B_n^*(z)A_n^*(w) - B_n^*(w)A_n^*(z)\big] \\
& + \bar{\alpha}_{n+1} Q_{q+1}(z,w)\big[A_n(z)A_n^*(w) - A_n(w)A_n^*(z)\big] \\
& + \bar{\alpha}_{n+1}^2 Q_q(z,w)\big[A_n(z)B_n(w) - A_n(w)B_n(z)\big] \\
& + \bar{\alpha}_{n+1} Q_q(z,w)\big[zB_n^*(z)B_n(w) - wB_n^*(w)B_n(z)\big] \\
& - \bar{\alpha}_{n+1} zw\big[z^{q-1}A_n(z)A_n^*(w) - w^{q-1}A_n(w)A_n^*(z)\big] \\
= {} & Q_q(z,w)\big[zwB_n^*(z)A_n^*(w) - zwB_n^*(w)A_n^*(z) \\
& \qquad + \bar{\alpha}_{n+1} wA_n(z)A_n^*(w) - \bar{\alpha}_{n+1} zA_n(w)A_n^*(z) \\
& \qquad + \bar{\alpha}_{n+1}^2 A_n(z)B_n(w) - \bar{\alpha}_{n+1}^2 A_n(w)B_n(z) \\
& \qquad + \bar{\alpha}_{n+1} zB_n^*(z)B_n(w) - \bar{\alpha}_{n+1}^2 wB_n^*(w)B_n(z)\big] \\
= {} & Q_q(z,w)\big[B_{n+1}^*(z)A_{n+1}^*(w) - B_{n+1}^*(w)A_{n+1}^*(z)\big] \\
= {} & \mathrm{rhs}(x_{n+1,q})
\end{aligned}
$$

and

$$\begin{aligned}
\operatorname{lhs}(y_{n+1,q}) = & z^q Y_{n+1}(z,w) - w^q Y_{n+1}(w,z) \\
= & \alpha_{n+1} \cdot \operatorname{lhs}(r_{n,q+1}) + \alpha_{n+1} zw \cdot \operatorname{lhs}(s_{n-1,q-1}) \\
& + \alpha_{n+1}^2 zw \cdot \operatorname{lhs}(x_{n,q}) + \operatorname{lhs}(y_{n,q}) \\
= & \alpha_{n+1}^2 zw Q_q(z,w) \big[ B_n^*(z) A_n^*(w) - B_n^*(w) A_n^*(z) \big] \\
& + \alpha_{n+1} Q_{q+1}(z,w) \big[ A_n(z) A_n^*(w) - A_n(w) A_n^*(z) \big] \\
& + Q_q(z,w) \big[ A_n(z) B_n(w) - A_n(w) B_n(z) \big] \\
& + \alpha_{n+1} Q_q(z,w) \big[ z B_n^*(z) B_n(w) - w B_n^*(w) B_n(z) \big] \\
& - \alpha_{n+1} zw \big[ z^{q-1} A_n(z) A_n^*(w) - w^{q-1} A_n(w) A_n^*(z) \big] \\
= & \, Q_q(z,w) \big[ \alpha_{n+1}^2 zw B_n^*(z) A_n^*(w) - \alpha_{n+1}^2 zw B_n^*(w) A_n^*(z) \\
& \qquad + \alpha_{n+1} w A_n(z) A_n^*(w) - \alpha_{n+1} z A_n(w) A_n^*(z) \\
& \qquad + A_n(z) B_n(w) - A_n(w) B_n(z) \\
& \qquad + \alpha_{n+1} z B_n^*(z) B_n(w) - \alpha_{n+1}^2 w B_n^*(w) B_n(z) \big] \\
= & \, Q_q(z,w) \big[ A_{n+1}(z) B_{n+1}(w) - A_{n+1}(w) B_{n+1}(z) \big] \\
= & \operatorname{rhs}(y_{n+1,q}),
\end{aligned}$$

as claimed. □

## 4.2 Independence of the Hamiltonians

We want to discuss some issues related to the independence of the commuting Hamiltonians defined in the previous section. All the results that we present here have been proved by Simon, and can be found in [29, Chapter 11].

As explained before, one of our main reasons for investigating Hamiltonian systems related to orthogonal polynomials on the unit circle was to try to explain the existence of so-called isospectral tori of Verblunsky coefficients, observed in Examples 3.11 and 3.12. Recall that the name "isospectral" is justified by Proposition 3.10.

In fact, more is true:

**Theorem 4.9.** *The isospectral manifold is a torus of dimension equal to the number of open gaps.*

*Remark* 4.10. Note that, if $\{\alpha_j\}_{j=0}^{p-1}$ and $\{\gamma_j\}_{j=0}^{p-1}$ are two isospectral sequences of periodic Verblunsky coefficients, then they define the same discriminant $\Delta(z)$, and hence the associated measure has the same (and so the same number of) open gaps.

Therefore, the statement of Theorem 4.9 makes sense: The number of open gaps is constant on any given isospectral manifold.

This theorem is proved in [29] by other methods than the ones we are concerned with here. Nonetheless, we believe this phenomenon to be the consequence of Conjecture 4.13 below. The analogous statement on the real line holds, and was proved in the continuous setting by McKean and van Moerbeke [23], and in the discrete setting (Toda lattice) by van Moerbeke [36].

**Theorem 4.11.** *Let $p$ be even. Any set $z_1, \ldots, z_p$ of distinct points in $S^1$ which obey*

$$\prod_{j=1}^{p} z_j = 1 \tag{4.2.1}$$

*is a possible spectrum of some $\mathcal{Q}$.*

*Define*

$$Q_-(z_1, \ldots, z_p) = -\max\left[\tilde{Q}(\theta)\,\middle|\,\frac{\partial\tilde{Q}}{\partial\theta} = 0, \tilde{Q}(\theta) < 0\right],$$

*where*

$$\tilde{Q}(\theta) = e^{-ip\theta/2}\prod_{j=1}^{p}(e^{i\theta} - z_j).$$

*Then the allowed moduli consistent with $(z_1, \ldots, z_p)$ as spectrum for $\mathcal{Q}$ is $(0, \frac{1}{4}Q_-]$. In the interior, the set of $\{\alpha_j\}_{j=0}^{p-1}$ with that set of eigenvalues and modulus is a $p$-dimensional torus. At the end-point where the modulus is $\frac{1}{4}Q_-$, the dimension is strictly less than $p$.*

We do not prove this result here, but instead refer the interested reader to [29, Section 11.4].

*Remark* 4.12. It is convenient to parameterize the $(p-1)$-dimensional $z$'s obeying (4.2.1) by using the fact that

$$\prod_{j\in C} z_j = \prod_{j\notin C} \bar{z}_j$$

for any $C \subset \{1, 2, \ldots, p\}$.

Denote

$$I_q = \sum_{\{j_1,\ldots,j_q\}\subset\{1,\ldots,p\}} z_{j_1}\cdots z_{j_q}.$$

Then

$$I_0 = I_p = 1,$$

$$I_{p-q} = \bar{I}_q$$

for $1 \leq q \leq \frac{p}{2}$, and so

$$I_{p/2} \text{ is real.}$$

Define

$$\mathcal{F} : \mathbb{D}^p \to \mathbb{R}^p$$

by

$$\mathcal{F}(\alpha) = (\operatorname{Re} I_1, \operatorname{Im} I_1, \ldots, \operatorname{Re} I_{p/2-1}, \operatorname{Im} I_{p/2-1}, I_{p/2}, M^2).$$

Then $\mathcal{F}$ is a polynomial in $\{\operatorname{Re}\alpha_j\}_{j=0}^{p-1}$ and $\{\operatorname{Im}\alpha_j\}_{j=0}^{p-1}$, and by relation (3.5.3) for $\beta = 1$, we have

$$\Big(\prod_{j=0}^{p-1}\rho_j\Big) z^{p/2}[\Delta(z) - 2] = \det(z - \mathcal{Q}) = \sum_{q=0}^{p}(-1)^q I_q z^q.$$

Note that Theorem 4.2 and its corollary prove that for any nonzero $z, w \in \mathbb{C}$, we have

$$\Big\{\Big(\prod_{j=0}^{p-1}\rho_j\Big) z^{p/2}[\Delta(z) - 2], \prod_{j=0}^{p-1}\rho_j\Big\} = 0$$

and

$$\Big\{\Big(\prod_{j=0}^{p-1}\rho_j\Big) z^{p/2}[\Delta(z) - 2], \Big(\prod_{j=0}^{p-1}\rho_j\Big) w^{p/2}[\Delta(w) - 2]\Big\} = 0.$$

In particular, this implies

$$\{M^2, I_q\} = 0 \quad \text{and} \quad \{I_q, I_{q'}\} = 0$$

for any $1 \leq q, q' \leq p-1$. If we also remember that

$$I_q = \bar{I}_{p-q}$$

for all $q$, we obtain that

$$\operatorname{Re} I_1, \operatorname{Im} I_1, \ldots, \operatorname{Re} I_{p/2-1}, \operatorname{Im} I_{p/2-1}, I_{p/2}, M^2$$

are a set of $p$ Poisson commuting Hamiltonians on $\mathbb{D}^p$. So independence of these Hamiltonians at a point in $\mathbb{D}^p$ is equivalent to rank($\mathcal{F}$)=$p$ at that point.

**Conjecture 4.13.** *The rank of $\mathcal{F}$ at a point $\alpha = (\alpha_0, \ldots, \alpha_{p-1}) \in \mathbb{D}^p$ equals the number of open gaps of the measure associated with the sequence of periodic Verblunsky coefficients $\{\alpha_j\}_{j\geq 0}$, $\alpha_{k+np} = \alpha_k$ for any $0 \leq k \leq p-1$ and $n \geq 0$.*

Here we will prove the weaker result:

**Theorem 4.14.** *The image of $\mathbb{D}^p$ under $\mathcal{F}$ contains an open subset of $\mathbb{R}^p$.*

*Proof.* Notice that, as $\alpha_{p-1} \to -1$, the periodized CMV matrix $\mathcal{Q}$ converges to a finite CMV matrix

$$\mathcal{C}_f = \mathcal{C}_f(\alpha_0, \ldots, \alpha_{p-2}; \alpha_{-1} = \alpha_{p-1} = -1),$$

since

$$\Theta_{p-1} = \begin{bmatrix} \bar{\alpha}_{p-1} & \rho_{p-1} \\ \rho_{p-1} & -\alpha_{p-1} \end{bmatrix} \to \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So, in this limit,

$$\begin{aligned}\det(\mathcal{Q}-z) &= \det(\mathcal{C}_f - z)\\ &= \Phi_p(z;\alpha_0,\dots,\alpha_{p-2},\alpha_{p-1}=-1)\\ &= z\Phi_{p-1}(z;\alpha_0,\dots\alpha_{p-2}) + \Phi^*_{p-1}(z;\alpha_0,\dots,\alpha_{p-2}).\end{aligned}$$

But then, by Theorem 2.2.13 of [28], we get that

$$\operatorname{ran}\bigl(\lim_{\alpha_{p-1}\to -1} \mathcal{F}\restriction (M=0)\bigr)$$

is open, and hence, by continuity, for $m_0$ small, there are points where $\mathcal{F}\restriction (M^2=m_0)$ has maximal rank $p-1$. If we parameterize $\mathbb{D}^p$ so that $M^2$ is the last coordinate, then in these coordinates we get

$$d\mathcal{F} = \begin{bmatrix} d\bigl(\mathcal{F}\restriction (M^2=m_0)\bigr) & \star \\ 0 & 1\end{bmatrix}.$$

Thus $d\mathcal{F}$ has maximal rank $p$ at points where $\mathcal{F}\restriction (M^2=m_0)$ has maximal rank $(p-1)$. By the implicit function theorem, this means that the range $\operatorname{ran}(\mathcal{F})$ contains an open set. □

We can now use this result to prove that the Hamiltonians

$$\operatorname{Re} I_1, \operatorname{Im} I_1, \dots, \operatorname{Re} I_{p/2-1}, \operatorname{Im} I_{p/2-1}, I_{p/2}, M^2$$

are independent on a dense set of $\alpha\in\mathbb{D}^p$. The argument that we give here also represents the first part in proving Theorem 4.9.

Let $\delta_1,\dots,\delta_p$ be the canonical basis in $\mathbb{R}^p$ and define

$$G:\mathbb{D}^p\to \wedge^p(\mathbb{R}^{2p})$$

by

$$G(\alpha) = \wedge^p [d\mathcal{F}_\alpha^t](\delta_1 \wedge \cdots \wedge \delta_p),$$

where $t$ denotes the transpose. Then note that

$$\dim(\operatorname{ran}(d\mathcal{F}_\alpha)) = p \quad \text{if and only if} \quad G(\alpha) \neq 0.$$

But $G$ is a vector-valued polynomial in $\operatorname{Re}(\alpha_j)$ and $\operatorname{Im}(\alpha_j)$, and so it is real-analytic. We conclude that either $G$ is identically zero, or else its nonzero points are dense.

Assume that $G$ is identically zero. This is the same as saying that no point in $\mathbb{D}^p$ is regular for $\mathcal{F}$, and hence the range of $\mathcal{F}$ consists only of singular values. But Sard's theorem says that the set of singular values of $\mathcal{F}$ has measure zero. We therefore find that in this case, the range of $\mathcal{F}$ is a set of measure zero in $\mathbb{R}^p$, which contradicts the result of Theorem 4.14.

Hence $G$ cannot be identically zero, and so

$$\dim(\operatorname{ran}(d\mathcal{F}_\alpha)) = p$$

on a dense set of $\alpha \in \mathbb{D}^p$, as claimed.

# Chapter 5

# Lax Pairs for the Defocusing Ablowitz-Ladik System

## 5.1 Lax Pairs in the Periodic Setting: Main Results

We must first define our Hamiltonians $K_n$. Essentially, they are traces per volume of the powers of the extended CMV matrix $\mathcal{E}$.

Consider the periodic Ablowitz-Ladik problem with period $p$. If $p$ is even, then let $\mathcal{E}$ be the extended CMV matrix associated to these $\alpha$'s; if $p$ is odd, think of the sequence of Verblunsky coefficients as having period $2p$ and thus define the extended CMV matrix $\mathcal{E}$.

For each $n \geq 1$, we define the Hamiltonians we will be working with as

$$K_n = \frac{1}{n} \sum_{k=0}^{p-1} \mathcal{E}^n_{kk} \,. \tag{5.1.1}$$

For $n = 0$, we set

$$K_0 = \prod_{j=0}^{p-1} \rho_j^2.$$

Finally, for $\mathcal{A}$ a doubly-infinite matrix, we set $\mathcal{A}_+$ as the matrix with entries

$$(\mathcal{A}_+)_{jk} = \begin{cases} \mathcal{A}_{jk}, & \text{if } j < k, \\ \frac{1}{2}\mathcal{A}_{jj}, & \text{if } j = k, \\ 0, & \text{if } j > k. \end{cases}$$

Our central result is:

**Theorem 5.1.** *The Lax pairs for the $n^{\text{th}}$ Hamiltonian of the periodic defocusing Ablowitz-Ladik system are given by*

$$\{\mathcal{E}, K_n\} = [\mathcal{E}, i\mathcal{E}_+^n] \tag{5.1.2}$$

*and*

$$\{\mathcal{E}, \bar{K}_n\} = [\mathcal{E}, i(\mathcal{E}_+^n)^*] \tag{5.1.3}$$

*for all $n \geq 1$.*

Here we use $\{\mathcal{E}, f\}$ to denote the doubly-infinite matrix with $(j,k)$ entry $\{\mathcal{E}_{jk}, f\}$; also, $\mathcal{E}_+^n$ denotes $(\mathcal{E}^n)_+$ .

*Remark* 5.2. The form of Theorem 5.1 and the main idea of the proof were inspired by the analogous result of van Moerbeke [36] for the periodic Toda lattice. But neither of the two results implies the other.

Moreover, in the case of the Toda lattice, the necessary calculations are very simple due to the tri-diagonal, symmetric nature of the Jacobi matrices naturally associated with that problem. The analogue on the circle are CMV matrices, whose more complicated structure makes proving this result computationally much more involved.

But we are dealing with a finite-dimensional problem, so we are interested in finding appropriate finite-dimensional spaces to which we can restrict the operators in (5.1.2) and (5.1.3). Also, we want to express the Hamiltonians $K_n$ in terms of these restrictions.

The following lemma is an immediate consequence of the structure of $\mathcal{E}$; it can easily be proved by induction whenever $\mathcal{E}$ can be defined.

**Lemma 5.3.** *Let $n \geq 1$ be an integer. Then $\mathcal{E}^n_{j,k}$ is identically zero as a function of the Verblunsky coefficients if one of the following holds:*

$$|j-k| \geq 2n+1$$

*or $j-k = 2n$ and $j$ and $k$ are even*

*or $j-k = -2n$ and $j$ and $k$ are odd.*

*In particular, the number of entries which are not identically zero (as functions of the $\alpha$'s) on any row of $\mathcal{E}^n$ is bounded by $4n$.*

Recall that the definition of $\mathcal{E}$ depends on the parity of the period $p$. This explains why we need to study the cases $p$ even and $p$ odd separately.

Let us first consider the case of the period $p$ being even. We denote by $X_{(d)}$ the subspace of $l^\infty(\mathbb{Z})$

$$X_{(d)} = \{u \in l^\infty(\mathbb{Z}) \,|\, u_{m+dp} = u_m\}$$

of sequences of period $dp$. As the Verblunsky coefficients are periodic with period $p$, we find that $\mathcal{E}_{j+p,k+p} = \mathcal{E}_{j,k}$ for any $j,k \in \mathbb{Z}$, and hence $\mathcal{E}^n$ restricts to $X_{(d)}$ for all $n \in \mathbb{Z}$ and $d \geq 1$. Moreover, if we denote by $\xi^{(d)}_k$, $k = 0, \ldots, dp-1$, the $l^\infty(\mathbb{Z})$ vector given by

$$(\xi^{(d)}_k)_j = 1 \text{ when } j \equiv k \pmod{dp}, \text{ and } 0 \text{ otherwise},$$

we have that $\{\xi^{(d)}_0, \xi^{(d)}_1, \ldots, \xi^{(d)}_{dp-1}\}$ is a basis in $X_{(d)}$, and

$$(\mathcal{E}^n \xi^{(d)}_k)_{j+p} = (\mathcal{E}^n \xi^{(d)}_k)_j = \sum_{l \in \mathbb{Z}} \mathcal{E}^n_{j,k+ldp}\,.$$

Notice that this sum has only a finite number of nonzero terms for any choice of $n, j, k, p$, and $d$.

Let us denote by $\mathcal{Q}_{(d)}$ the matrix representation of the restriction $\mathcal{E} \restriction X_{(d)}$ in the basis $\{\xi_0^{(d)}, \xi_1^{(d)}, \ldots, \xi_{dp-1}^{(d)}\}$. Then the matrix representing $\mathcal{E}^n \restriction X_{(d)}$ in the same basis is $\mathcal{Q}_{(d)}^n$, whose entries are given by

$$\mathcal{Q}^n_{(d),jk} = \sum_{l \in \mathbb{Z}} \mathcal{E}^n_{j,k+ldp} \tag{5.1.4}$$

for $0 \le j, k \le dp - 1$.

**Lemma 5.4.** *For $dp \ge 2n + 1$, we have that*

$$\frac{1}{d} \mathrm{Tr}(\mathcal{Q}^n_{(d)})$$

*is independent of $d$ and equals $K_n$.*

*Proof.*

$$\frac{1}{d} \mathrm{Tr}(\mathcal{Q}^n_{(d)}) = \frac{1}{d} \sum_{k=0}^{dp-1} \mathcal{Q}^n_{(d),kk} \,.$$

From Lemma 5.3 we know that $\mathcal{E}^n_{jk} = 0$ for $|j - k| > 2n$. So for $dp \ge 2n + 1$ we get

$$\mathcal{Q}^n_{(d),kk} = \sum_{l \in \mathbb{Z}} \mathcal{E}^n_{k,k+ldp} = \mathcal{E}^n_{kk} \,.$$

From this and periodicity, we can conclude that

$$\frac{1}{d} \mathrm{Tr}(\mathcal{Q}^n_{(d)}) = \frac{1}{d} \sum_{k=0}^{dp-1} \mathcal{E}^n_{kk} = \sum_{k=0}^{p-1} \mathcal{E}^n_{kk} = nK_n$$

is indeed independent of $d$. □

If $p$ is odd, we consider the same objects as above, with the extra constraint that $dp$, and hence $d$, must always be even. Recall that, in this case, we define $\mathcal{E}$ by thinking of the Verblunsky coefficients as having period $2p$. For $d$ even, we can then define $X_{(d)}$ and $\mathcal{Q}_{(d)}$ as above, while always keeping in mind that we can use the results we just proved for $dp = \frac{d}{2} \cdot 2p$.

Therefore, if $d$ is even and large enough, we have

$$\frac{2}{d}\mathrm{Tr}(\mathcal{Q}^n_{(d)}) = \frac{2}{d}\sum_{k=0}^{dp-1}\mathcal{E}^n_{kk} = \sum_{k=0}^{2p-1}\mathcal{E}^n_{kk}\,.$$

The last observation we need to make is that, in this case, the entries of $\mathcal{E}$ obey

$$\mathcal{E}_{jk} = \mathcal{E}_{k+p,j+p}\,.$$

This comes from the fact that

$$\mathcal{L}_{j+p,k+p} = \mathcal{M}_{jk}$$

and that $\mathcal{L}$ and $\mathcal{M}$ are symmetric. Hence,

$$\mathcal{E}_{k+p,j+p} = \sum_{l\in\mathbb{Z}}\mathcal{L}_{k+p,l}\mathcal{M}_{l,j+p} = \sum_{l\in\mathbb{Z}}\mathcal{L}_{kl}\mathcal{M}_{lj} = \sum_{l\in\mathbb{Z}}\mathcal{L}_{lk}\mathcal{M}_{jl} = \mathcal{E}_{jk}\,,$$

as claimed. A straightforward induction shows that

$$\mathcal{E}^n_{k+p,j+p} = \mathcal{E}^n_{jk}$$

for all $n$, and hence,

$$\frac{1}{d}\mathrm{Tr}(\mathcal{Q}^n_{(d)}) = \frac{1}{2}\sum_{k=0}^{2p-1}\mathcal{E}^n_{kk} = \sum_{k=0}^{p-1}\mathcal{E}^n_{kk} = nK_n$$

also holds for $p$ odd, as long as $dp$ is even and $dp \geq 2n+1$.

So we proved that, with $K_n$ defined as in (5.1.1), we have

$$K_n = \frac{1}{dn}\mathrm{Tr}(\mathcal{Q}^n_{(d)}) \tag{5.1.5}$$

for $dp$ even and greater than $2n+1$.

Let us note that relations (5.1.2) and (5.1.3) hold in the sense of bounded operators on $l^\infty(\mathbb{Z})$. Moreover, all the matrices in these relations obey the same periodicity conditions as $\mathcal{E}$, so it makes sense to restrict (5.1.2) and (5.1.3) to $X_{(d)}$ for $d \geq 1$. By doing this we get

**Corollary 5.5.** *For all $d \geq 1$, with $dp$ even, and $n \geq 1$, we have*

$$\{\mathcal{Q}_{(d)}, K_n\} = [\mathcal{Q}_{(d)}, i\mathcal{Q}^n_{(d),+}]$$

*and*

$$\{\mathcal{Q}_{(d)}, \bar{K}_n\} = [\mathcal{Q}_{(d)}, i(\mathcal{Q}^n_{(d),+})^*],$$

*where we denote by $\mathcal{Q}^n_{(d),+}$ the matrix representation of $(\mathcal{E}^n)_+ \restriction X_{(d)}$ in the basis $\{\xi^{(d)}_0, \xi^{(d)}_1, \ldots, \xi^{(d)}_{dp-1}\}$.*

Note that $\mathcal{Q}^n_{(d),+}$ is not an upper triangular matrix, as it contains entries which are generically nonzero in its lower left corner.

Let us make an observation that will explain why we cannot simply use the traces of powers of $\mathcal{Q}_{(1)}$ even if $p$ is even, but also that we are not changing by much the Hamiltonians we are most interested in:

**Proposition 5.6.** *For $p$ even and $1 \leq n \leq \frac{p}{2} - 1$, we have*

$$K_n = \frac{1}{n}\mathrm{Tr}(\mathcal{Q}^n_{(1)}),$$

*but*

$$\frac{2}{p}\mathrm{Tr}(\mathcal{Q}^{p/2}_{(1)}) = K_{p/2} + 2K_0^{1/2}.$$

*Proof.* From formula (5.1.4) and Lemma 5.3 we see that, for $n \leq \frac{p}{2} - 1$,

$$\mathcal{Q}^n_{(1),jj} = \sum_{l \in \mathbb{Z}} \mathcal{E}^n_{j,j+lp} = \mathcal{E}^n_{jj}$$

for all $j = 0, \ldots, p-1$. This follows since, for $|l| \geq 1$,

$$|j - (j + lp)| \geq p \geq 2n + 1.$$

Hence, using (5.1.1),

$$\frac{1}{n}\mathrm{Tr}(\mathcal{Q}_{(1)}^{n}) = \frac{1}{n}\sum_{j=0}^{p-1}\mathcal{E}_{jj}^{n} = K_n.$$

If $n = \frac{p}{2}$ and $j$ even, the formulae (5.1.4), (6.2.2), (6.2.3), and periodicity of the Verblunsky coefficients imply that

$$\begin{aligned}\mathcal{Q}_{(1),jj}^{n} &= \sum_{l\in\mathbb{Z}}\mathcal{E}_{j,j+lp}^{n}\\ &= \mathcal{E}_{jj}^{n} + \mathcal{E}_{j,j+p}^{n}\\ &= \mathcal{E}_{jj}^{n} + \prod_{k=0}^{p-1}\rho_k\end{aligned}$$

and

$$\begin{aligned}\mathcal{Q}_{(1),j+1,j+1}^{n} &= \sum_{l\in\mathbb{Z}}\mathcal{E}_{j+1,j+1+lp}^{n}\\ &= \mathcal{E}_{j+1,j+1}^{n} + \mathcal{E}_{j+1,j+1-p}^{n}\\ &= \mathcal{E}_{j+1,j+1}^{n} + \prod_{k=0}^{p-1}\rho_k.\end{aligned}$$

Therefore,

$$\frac{2}{p}\mathrm{Tr}(\mathcal{Q}_{(1)}^{p/2}) = \frac{2}{p}\sum_{j=0}^{p-1}\mathcal{E}_{jj}^{n} + \frac{2p}{p}\prod_{k=0}^{p-1}\rho_k = K_{p/2} + 2K_0^{1/2},$$

as claimed. □

*Remark* 5.7. An easy computation shows

$$\{\alpha_j, 2\,\mathrm{Re}(K_1)\} = i\rho_j^2(\alpha_{j-1} + \alpha_{j+1})$$

and

$$\{\alpha_j, \log(K_0)\} = i\alpha_j$$

for all $0 \leq j \leq p-1$. Hence (1.0.2), the periodic defocusing Ablowitz-Ladik equation, is the evolution of the Verblunsky coefficients under the flow generated by the

Hamiltonian $2\,\mathrm{Re}(K_1) - 2\log(K_0)$.

From Theorem 5.1 and Corollary 5.5, we can immediately conclude

**Corollary 5.8.** *The Lax pairs for the Hamiltonians* $\mathrm{Re}(K_n)$ *and* $\mathrm{Im}(K_n)$*,* $n \geq 1$*, are given by*

$$\{\mathcal{E}, 2\,\mathrm{Re}(K_n)\} = [\mathcal{E}, i\mathcal{E}_+^n + i(\mathcal{E}_+^n)^*] \tag{5.1.6}$$

*and*

$$\{\mathcal{E}, 2\,\mathrm{Im}(K_n)\} = [\mathcal{E}, \mathcal{E}_+^n - (\mathcal{E}_+^n)^*], \tag{5.1.7}$$

*while the corresponding statements for* $\mathcal{Q}_{(d)}$*,* $d \geq 1$ *and* $dp$ *even, are given by*

$$\{\mathcal{Q}_{(d)}, 2\,\mathrm{Re}(K_n)\} = [\mathcal{Q}_{(d)}, i\mathcal{Q}_{(d),+}^n + i(\mathcal{Q}_{(d),+}^n)^*] \tag{5.1.8}$$

*and*

$$\{\mathcal{Q}_{(d)}, 2\,\mathrm{Im}(K_n)\} = [\mathcal{Q}_{(d)}, \mathcal{Q}_{(d),+}^n - (\mathcal{Q}_{(d),+}^n)^*]\,. \tag{5.1.9}$$

In particular, relations (5.1.8) and (5.1.9), together with the observations on isospectrality of Lax operators from Section 2.2 and (5.1.5), imply

**Corollary 5.9.** *We have*

$$\{K_n, \mathrm{Re}(K_m)\} = \{K_n, \mathrm{Im}(K_m)\} = 0,$$

*and hence,*

$$\{K_n, K_m\} = \{K_n, \bar{K}_m\} = 0.$$

Define the doubly-infinite matrix $\mathcal{P}$ by

$$\mathcal{P}_{lm} = (-1)^l \delta_{lm} \frac{i}{2} \Big( \prod_{k=0}^{p-1} \rho_k^2 \Big).$$

**Proposition 5.10.** *The Lax pair representation for the flow generated by* $K_0 = \prod_{j=0}^{p-1} \rho_j^2$ *is given by*

$$\{\mathcal{E}, K_0\} = [\mathcal{E}, \mathcal{P}]. \tag{5.1.10}$$

*In particular, we can conclude*

$$\{K_0, K_n\} = \{K_0, \bar{K}_n\} = 0, \tag{5.1.11}$$

*or, equivalently,*

$$\{K_0, 2\,\mathrm{Re}(K_n)\} = \{K_0, 2\,\mathrm{Im}(K_n)\} = 0. \tag{5.1.12}$$

*Proof.* The Lax pair representation (5.1.10) is verified by a straightforward computation. It is based on the fact that the flow generated by $K_0$ rotates all the $\alpha$'s by the same angle

$$\{\alpha_j, K_0\} = iK_0\alpha_j\,,$$

while

$$[\mathcal{E}, \mathcal{P}]_{j,k} = \mathcal{E}_{j,k}(\mathcal{P}_{k,k} - \mathcal{P}_{j,j}).$$

The Poisson commutation relations (5.1.11) and (5.1.12) follow, as in the previous cases, by restricting the Lax pair to periodic subspaces and concluding that the flow preserves eigenvalues, and hence traces. □

From (3.5.3), Corollary 5.9, and Proposition 5.10, we immediately get that $\prod_{j=0}^{p-1}\rho_j$ and the coefficients $c_k$ of $z^{p/2}\Big(\prod_{j=0}^{p-1}\rho_j\Big)[\Delta(z)-2]$ Poisson commute. Note also that, by (3.5.3), we see that the connection between the $K$'s and the $c$'s cannot be explicitly written down. Hence, one cannot write simple Lax pairs in terms of $\mathcal{E}$ for the flows generated by the $c$'s.

## 5.2 Proof of Theorem 5.1: Main Ideas

The main technical ingredient in the proof of Theorem 5.1 is the following:

**Lemma 5.11.** *For all $n \geq 0$ and $j$ even, we have*

$$\begin{aligned}\frac{\partial K_{n+1}}{\partial \alpha_j} = &-\frac{\bar{\alpha}_j\bar{\alpha}_{j+1}}{2\rho_j}\mathcal{E}^n_{j+1,j} - \frac{\bar{\alpha}_j\rho_{j+1}}{2\rho_j}\mathcal{E}^n_{j+2,j} - \frac{\bar{\alpha}_j\rho_{j-1}}{2\rho_j}\mathcal{E}^n_{j-1,j+1} \\ &+ \frac{\bar{\alpha}_j\alpha_{j-1}}{2\rho_j}\mathcal{E}^n_{j,j+1} - \bar{\alpha}_{j+1}\mathcal{E}^n_{j+1,j+1} - \rho_{j+1}\mathcal{E}^n_{j+2,j+1}\end{aligned} \tag{5.2.1}$$

$$\begin{aligned}\frac{\partial K_{n+1}}{\partial \bar{\alpha}_j} =& -\frac{\alpha_j\rho_{j+1}}{2\rho_j}\mathcal{E}^n_{j+2,j} - \frac{\alpha_j\rho_{j-1}}{2\rho_j}\mathcal{E}^n_{j-1,j+1} + \frac{\alpha_j\alpha_{j-1}}{2\rho_j}\mathcal{E}^n_{j,j+1} \\ &+ \rho_{j-1}\mathcal{E}^n_{j-1,j} - \alpha_{j-1}\mathcal{E}^n_{j,j} - \frac{\alpha_j\bar{\alpha}_{j+1}}{2\rho_j}\mathcal{E}^n_{j+1,j}\end{aligned} \tag{5.2.2}$$

$$\begin{aligned}\frac{\partial K_{n+1}}{\partial \alpha_{j-1}} =& -\frac{\bar{\alpha}_{j-1}\rho_{j-2}}{2\rho_{j-1}}\mathcal{E}^n_{j,j-2} + \frac{\bar{\alpha}_{j-1}\alpha_{j-2}}{2\rho_{j-1}}\mathcal{E}^n_{j,j-1} - \frac{\bar{\alpha}_{j-1}\bar{\alpha}_j}{2\rho_{j-1}}\mathcal{E}^n_{j-1,j} \\ &- \frac{\bar{\alpha}_{j-1}\rho_j}{2\rho_{j-1}}\mathcal{E}^n_{j-1,j+1} - \bar{\alpha}_j\mathcal{E}^n_{j,j} - \rho_j\mathcal{E}^n_{j,j+1}\end{aligned} \tag{5.2.3}$$

$$\begin{aligned}\frac{\partial K_{n+1}}{\partial \bar{\alpha}_{j-1}} =& \rho_{j-2}\mathcal{E}^n_{j-1,j-2} - \alpha_{j-2}\mathcal{E}^n_{j-1,j-1} - \frac{\alpha_{j-1}\rho_{j-2}}{2\rho_{j-1}}\mathcal{E}^n_{j,j-2} \\ &- \frac{\alpha_{j-1}\bar{\alpha}_j}{2\rho_{j-1}}\mathcal{E}^n_{j-1,j} - \frac{\alpha_{j-1}\rho_j}{2\rho_{j-1}}\mathcal{E}^n_{j-1,j+1} + \frac{\alpha_{j-1}\alpha_{j-2}}{2\rho_{j-1}}\mathcal{E}^n_{j,j-1}\,.\end{aligned} \tag{5.2.4}$$

*Remark* 5.12. Note that, for any $n \geq 1$ and $0 \leq j \leq p-1$, we have

$$\frac{\partial \bar{K}_n}{\partial \beta_j} = \overline{\left(\frac{\partial K_n}{\partial \bar{\beta}_j}\right)}$$

and hence one can easily find the derivatives of $\bar{K}_n$ with respect to $\alpha_j$ and $\bar{\alpha}_j$ from Lemma 5.11.

*Proof.* The proof reduces to direct computations once one notices that, by invariance of the trace under circular permutations,

$$\frac{\partial K_{n+1}}{\partial \beta_j} = \frac{1}{d}\mathrm{Tr}\left(\frac{\partial \mathcal{Q}_{(d)}}{\partial \beta_j}\mathcal{Q}^n_{(d)}\right). \tag{5.2.5}$$

We give here the complete proof of (5.2.1); (5.2.2) through (5.2.4) can be found in a similar way.

Notice that, for $j$ even, $\alpha_j$ appears in exactly $6d$ entries of $\mathcal{Q}_{(d)}$. So (5.2.1) follows by periodicity and by a straightforward computation from (5.2.5):

$$\begin{aligned}\frac{\partial K_{n+1}}{\partial \alpha_j} =& \frac{1}{d}\sum_{k,l}\frac{\partial \mathcal{Q}_{(d),kl}}{\partial \alpha_j}\mathcal{Q}^n_{(d),lk} \\ =& -\frac{\bar{\alpha}_j}{\rho_j}\bar{\alpha}_{j+1}\mathcal{E}^n_{j+1,j} - \frac{\bar{\alpha}_j}{\rho_j}\rho_{j+1}\mathcal{E}^n_{j+2,j} - \frac{\bar{\alpha}_j}{\rho_j}\rho_{j-1}\mathcal{E}^n_{j-1,j+1} \\ &+ \frac{\bar{\alpha}_j}{\rho_j}\alpha_{j-1}\mathcal{E}^n_{j,j+1} - \bar{\alpha}_{j+1}\mathcal{E}^n_{j+1,j+1} - \rho_{j+1}\mathcal{E}^n_{j+2,j+1}.\end{aligned}$$

□

Before we embark on the proof of the main theorem, we provide another preliminary result; while the statement is almost certainly not new, we give a proof for the reader's convenience.

Consider an $N \times N$ matrix $A$ having the following *stair-shape*

$$A = \begin{pmatrix} \star & 0 & 0 & \cdots & 0 \\ \star & \star & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \star & \star & \star & \cdots & 0 \\ \star & \star & \star & \cdots & 0 \end{pmatrix},$$

where the stars and 0's represent rectangular matrix blocks. Formally, that means that for any row number $i$, there exists a column number $j(i)$ so that $A_{ij} = 0$ for all $j > j(i)$, and the function $i \mapsto j(i)$ is nondecreasing. In particular, it is also true that for any column $j$, there exists a row $i(j)$ so that $A_{ij} = 0$ for $i < i(j)$. We note in passing that $j(i)$ and $i(j)$ are not equal.

We will say, somewhat informally, that another matrix $\tilde{A}$ *has the same shape as* $A$ if $\tilde{A}_{ij} = 0$ whenever $j > j(i)$ for all $i$.

**Lemma 5.13.** *Let $A$ be a matrix as above and $B$ be an arbitrary $N \times N$ matrix. Then*

$$[A, B_+]_{ij} = [A, B]_{ij}$$

*for all $(i, j)$ with $j > j(i)$. This implies that, for the same indices $(i, j)$ with $j > j(i)$, we have*

$$[A, B_-]_{ij} = 0.$$

*Remark* 5.14. Note that:

- If $A$ and $B$ commute, then the commutators $[A, B_+]$ and $[A, B_-]$ have the same shape as $A$.

- Also, by transposing these equations, we obtain the same type of result for lower triangle shapes.

- The same type of result holds for doubly-infinite matrices. In particular, if $\mathcal{A}$ and $\mathcal{B}$ are two doubly-infinite, stair-shaped matrices such that the commutator $[\mathcal{A}, \mathcal{B}]$ makes sense and equals 0, then the commutators $[\mathcal{A}, \mathcal{B}_+]$ and $[\mathcal{A}, \mathcal{B}_-]$ are themselves stair-shaped.

*Proof.* We proceed by direct computation: Let $(i, j)$ be an index so that $j > j(i)$; equivalently, $i < i(j)$. Then

$$\begin{aligned}
[A, B_+]_{ij} &= \sum_k A_{ik} B_{+,kj} - \sum_k B_{+,ik} A_{kj} \\
&= \sum_{k \le j(i) < j} A_{ik} B_{+,kj} - \sum_{i < i(j) \le k} B_{+,ik} A_{kj} \\
&= \sum_k A_{ik} B_{kj} - \sum_k B_{ik} A_{kj} \\
&= [A, B]_{ij}.
\end{aligned}$$

Since $B_- = B - B_+$, we get

$$[A, B_-] = [A, B] - [A, B_+]$$

and so the second relation is just a consequence of the first one. □

We are now ready to prove Theorem 5.1.

*Proof.* We will first deal with relation (5.1.2) for $n+1$, $n \ge 0$,

$$\{\mathcal{E}, K_{n+1}\} = i[\mathcal{E}, \mathcal{E}_+^{n+1}].$$

The left-hand side matrix has two types of entries: the ones outside the shape of a CMV matrix, which are identically zero, and the ones inside the shape.

The entries outside the shape are dealt with immediately by applying Lemma 5.13.

Indeed, $\mathcal{E}$ and $\mathcal{E}^n$ are doubly-infinite matrices, and they commute; hence, by the third observation above, the commutator $[\mathcal{E}, \mathcal{E}_+^{n+1}]$ has the same shape as $\mathcal{E}$.

We are now left with the entries $(j,k)$ that are inside the shape. Before we start computing, we make a short observation. Consider the doubly-infinite matrix $\mathcal{U}$ given by

$$\mathcal{U}_{jk} = \delta_{j,k+1}$$

for all $j,k \in \mathbb{Z}$. In other words, $\mathcal{U}$ is the left-shift on $l^\infty(\mathbb{Z})$ in the usual basis. Note that for a doubly-infinite matrix $\mathcal{B}$ we have

$$(\mathcal{U}^*\mathcal{B}\mathcal{U})_{jk} = B_{j-1,k-1} \quad \text{and} \quad (\mathcal{U}\mathcal{B}\mathcal{U}^*)_{jk} = B_{j+1,k+1}.$$

Consider $\mathcal{E} = \mathcal{E}(\{\alpha_j\})$ to be a doubly-infinite CMV matrix. We know that $\mathcal{E} = \tilde{\mathcal{L}}\tilde{\mathcal{M}}$ with

$$\tilde{\mathcal{L}} = \operatorname{diag}\big(\dots, \Theta_0, \Theta_2, \Theta_4, \dots\big)$$

and

$$\tilde{\mathcal{M}} = \operatorname{diag}\big(\dots, \Theta_{-1}, \Theta_1, \Theta_3, \dots\big).$$

It is easily seen that

$$\mathcal{U}^*\tilde{\mathcal{L}}(\{\alpha_j\})^t\mathcal{U} = \tilde{\mathcal{M}}(\{\alpha_{j-1}\}) \quad \text{and} \quad \mathcal{U}^*\tilde{\mathcal{M}}(\{\alpha_j\})^t\mathcal{U} = \tilde{\mathcal{L}}(\{\alpha_{j-1}\}),$$

which implies that

$$\mathcal{U}^*\mathcal{E}(\{\alpha_j\})^t\mathcal{U} = \mathcal{E}(\{\alpha_{j-1}\})$$

is also a doubly-infinite CMV matrix. The same is true for

$$\mathcal{U}\mathcal{E}(\{\alpha_j\})^t\mathcal{U}^* = \mathcal{E}(\{\alpha_{j+1}\}).$$

We use the notation $(5.1.2)_{kl}$ for the $(k,l)$ entry of relation (5.1.2), and similarly for (5.1.3). Assume we know $(5.1.2)_{kl}$ for a fixed pair of indices $(k,l)$. As for any $\{\alpha_j\}$

the matrix $\mathcal{U}^*\mathcal{E}^t\mathcal{U}$ is a doubly-infinite CMV matrix, we know that

$$\{(\mathcal{U}^*\mathcal{E}^t\mathcal{U})_{kl}, K_{n+1}(\mathcal{U}^*\mathcal{E}^t\mathcal{U})\} = i[\mathcal{U}^*\mathcal{E}^t\mathcal{U}, (\mathcal{U}^*\mathcal{E}^t\mathcal{U})_+^{n+1}]_{kl}. \tag{5.2.6}$$

But

$$\begin{aligned} K_{n+1}(\mathcal{U}^*\mathcal{E}^t\mathcal{U}) &= \frac{1}{(n+1)d}\mathrm{Tr}\big((\mathcal{U}_{(d)}^*\mathcal{Q}_{(d)}^t\mathcal{U}_{(d)})^{n+1}\big) \\ &= \frac{1}{(n+1)d}\mathrm{Tr}\big(\mathcal{U}_{(d)}^*(\mathcal{Q}_{(d)}^t)^{n+1}\mathcal{U}_{(d)}\big) \\ &= K_{n+1}(\mathcal{E}) \end{aligned}$$

and $\mathcal{U}$ is a constant matrix. Therefore,

$$\begin{aligned} \{(\mathcal{U}^*\mathcal{E}^t\mathcal{U})_{kl}, K_{n+1}(\mathcal{U}^*\mathcal{E}^t\mathcal{U})\} &= \big(\mathcal{U}^*\{\mathcal{E}^t, K_{n+1}(\mathcal{E})\}\mathcal{U}\big)_{kl} \\ &= \{\mathcal{E}^t_{k-1,l-1}, K_{n+1}(\mathcal{E})\} \\ &= \{\mathcal{E}_{l-1,k-1}, K_{n+1}(\mathcal{E})\}. \end{aligned} \tag{5.2.7}$$

On the other hand,

$$\begin{aligned} i[\mathcal{U}^*\mathcal{E}^t\mathcal{U}, (\mathcal{U}^*\mathcal{E}^t\mathcal{U})_+^n]_{kl} &= i\big(\mathcal{U}^*[\mathcal{E}^t, (\mathcal{E}^t)_+^{n+1}]\mathcal{U}\big)_{kl} = i[\mathcal{E}^t, (\mathcal{E}^t)_+^{n+1}]_{k-1,l-1} \\ &= i[\mathcal{E}^t, (\mathcal{E}_-^{n+1})^t]_{k-1,l-1} = i[\mathcal{E}, \mathcal{E}_+^{n+1}]_{l-1,k-1}. \end{aligned} \tag{5.2.8}$$

Plugging (5.2.7) and (5.2.8) into (5.2.6), one gets relation $(5.1.2)_{l-1,k-1}$:

$$\{\mathcal{E}_{l-1,k-1}, K_{n+1}\} = i[\mathcal{E}, \mathcal{E}_+^{n+1}]_{l-1,k-1}.$$

If instead of considering $\mathcal{U}^*\mathcal{E}^t\mathcal{U}$ we consider $\mathcal{U}\mathcal{E}^t\mathcal{U}^*$, we obtain that $(5.1.2)_{kl}$ implies $(5.1.2)_{l+1,k+1}$. In particular, this means:

- $(5.1.2)_{kk} \quad \Leftrightarrow \quad (5.1.2)_{k+1,k+1}$
- $(5.1.2)_{k,k-1} \quad \Leftrightarrow \quad (5.1.2)_{k,k+1}$
- $(5.1.2)_{k+1,k-1} \quad \Leftrightarrow \quad (5.1.2)_{k,k+2}$

- $(5.1.2)_{k+1,k} \quad \Leftrightarrow \quad (5.1.2)_{k+1,k+2}$.

So the proof of relation (5.1.2) is complete once we prove it for the indices $(k,k)$, $(k,k-1)$, $(k+1,k-1)$, and $(k+1,k)$ with $k$ even.

We note here that we can apply the same reasoning as above to $\mathcal{E}^*$ instead of $\mathcal{E}^t$, but we do not obtain anything new.

Finally, these relations are proved using Lemma 5.11. We give the computational details in Section 5.3.

The second part of the proof deals with relation (5.1.3) for $n+1$, $n \geq 0$:

$$\{\mathcal{E}, \bar{K}_{n+1}\} = [\mathcal{E}, i(\mathcal{E}_+^{n+1})^*].$$

We shall proceed in very much the same way as with (5.1.2), while incorporating the necessary computational adjustments.

Let us first note that

$$(\mathcal{E}_+^{n+1})^* = ((\mathcal{E}^*)^{n+1})_- \,.$$

So Lemma 5.13 and the subsequent remarks apply here too and we can conclude that $[\mathcal{E}, i(\mathcal{E}_+^{n+1})^*]$ has the same shape as $\mathcal{E}$.

Turning our attention to the entries inside the shape of $\mathcal{E}$, we note that using exactly the same reasoning as for equation (5.1.2) shows that

- $(5.1.3)_{kk} \quad \Leftrightarrow \quad (5.1.3)_{k+1,k+1}$
- $(5.1.3)_{k,k-1} \quad \Leftrightarrow \quad (5.1.3)_{k,k+1}$
- $(5.1.3)_{k+1,k-1} \quad \Leftrightarrow \quad (5.1.3)_{k,k+2}$
- $(5.1.3)_{k+1,k} \quad \Leftrightarrow \quad (5.1.3)_{k+1,k+2}$.

So, again, we only have to check four relations; the only difference is that, in this case, $(5.1.3)_{k+1,k+1}$, $(5.1.3)_{k,k+1}$, $(5.1.3)_{k,k+2}$, and $(5.1.3)_{k+1,k+2}$ turn out to be computationally easier to verify. We do this in Section 5.3. $\square$

## 5.3 Proof of Theorem 5.1: The Full Computations

We prove relation (5.1.2) for the necessary indices.

First, let $k = l$ be even. Then

$$\begin{aligned}
i\{\mathcal{E}_{kk}, K_{n+1}\} &= \sum_j \rho_j^2 \Big[\frac{\partial(-\alpha_{k-1}\bar{\alpha}_k)}{\partial \alpha_j}\frac{\partial K_{n+1}}{\partial \bar{\alpha}_j} - \frac{\partial(-\alpha_{k-1}\bar{\alpha}_k)}{\partial \bar{\alpha}_j}\frac{\partial K_{n+1}}{\partial \alpha_j}\Big] \\
&= -\rho_{k-1}^2 \bar{\alpha}_k \frac{\partial K_{n+1}}{\partial \bar{\alpha}_{k-1}} + \rho_k^2 \alpha_{k-1}\frac{\partial K_{n+1}}{\partial \alpha_k} \\
&= -\rho_{k-1}^2 \bar{\alpha}_k \Big[\rho_{k-2}\mathcal{E}^n_{k-1,k-2} - \alpha_{k-2}\mathcal{E}^n_{k-1,k-1} - \frac{\alpha_{k-1}\rho_{-2}}{2\rho_{k-1}}\mathcal{E}^n_{k,k-2} \\
&- \frac{\alpha_{k-1}\bar{\alpha}_k}{2\rho_{k-1}}\mathcal{E}^n_{k-1,k} - \frac{\alpha_{k-1}\rho_k}{2\rho_{k-1}}\mathcal{E}^n_{k-1,k+1} + \frac{\alpha_{k-1}\alpha_{k-2}}{2\rho_{k-1}}\mathcal{E}^n_{k,k-1}\Big] \\
&+ \rho_k^2 \alpha_{k-1}\Big[ - \frac{\bar{\alpha}_k\bar{\alpha}_{k+1}}{2\rho_k}\mathcal{E}^n_{k+1,k} - \frac{\bar{\alpha}_k\rho_{k+1}}{2\rho_k}\mathcal{E}^n_{k+2,k} - \frac{\bar{\alpha}_k\rho_{k-1}}{2\rho_k}\mathcal{E}^n_{k-1,k+1} \\
&+ \frac{\bar{\alpha}_k\alpha_{k-1}}{2\rho_k}\mathcal{E}^n_{k,k+1} - \bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k+1} - \rho_{k+1}\mathcal{E}^n_{k+2,k+1}\Big].
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\big[\mathcal{E}, \mathcal{E}_+^{n+1}\big]_{k,k} &= \mathcal{E}_{k,k-1}\mathcal{E}^{n+1}_{k-1,k} - \mathcal{E}^{n+1}_{k,k+1}\mathcal{E}_{k+1,k} \\
&= \rho_{k-1}\bar{\alpha}_k\mathcal{E}^{n+1}_{k-1,k} + \alpha_{k-1}\rho_k\mathcal{E}^{n+1}_{k,k+1} \\
&= \rho_{k-1}\bar{\alpha}_k\Big[\rho_{k-2}\rho_{k-1}\mathcal{E}^n_{k-1,k-2} - \alpha_{k-2}\rho_{k-1}\mathcal{E}^n_{k-1,k-1} \\
&\qquad - \alpha_{k-1}\bar{\alpha}_k\mathcal{E}^n_{k-1,k} - \alpha_{k-1}\rho_k\mathcal{E}^n_{k-1,k+1}\Big] \\
&+\alpha_{k-1}\rho_k\Big[\rho_{k-1}\bar{\alpha}_k\mathcal{E}^n_{k-1,k+1} - \alpha_{k-1}\bar{\alpha}_k\mathcal{E}^n_{k,k+1} \\
&\qquad + \rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k+1} + \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1}\Big].
\end{aligned}$$

After a few simple manipulations, we find that $i\{\mathcal{E}_{kk}, K_{n+1}\} + \big[\mathcal{E}, \mathcal{E}_+^{n+1}\big]_{kk}$ equals

$$\begin{aligned}
&\frac{\alpha_{k-1}\bar{\alpha}_k}{2}\Big[\big(\mathcal{E}^n_{k,k-2}\rho_{k-2}\rho_{k-1} - \mathcal{E}^n_{k,k-1}\alpha_{k-2}\rho_{k-1} - \mathcal{E}^n_{k,k+1}\alpha_{k-1}\bar{\alpha}_k\big) \\
&\qquad - \big(\rho_{k-1}\bar{\alpha}_k\mathcal{E}^n_{k-1,k} + \rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k} + \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k}\big)\Big] \\
&= \frac{\alpha_{k-1}\bar{\alpha}_k}{2}\Big[\big(\mathcal{E}^n\mathcal{E}\big)_{kk} - \big(\mathcal{E}\mathcal{E}^n\big)_{kk}\Big] = 0,
\end{aligned}$$

which concludes the proof of $(5.1.2)_{kk}$.

The second case we must consider is $(5.1.2)_{k,k-1}$ with $k$ even. Again, we first look at

$$
\begin{aligned}
i\{\mathcal{E}_{k,k-1}, K_{n+1}\} &= \sum_j \rho_j^2 \Big[\frac{\partial(\rho_{k-1}\bar\alpha_k)}{\partial\alpha_j}\frac{\partial K_{n+1}}{\partial\bar\alpha_j} - \frac{\partial(\rho_{k-1}\bar\alpha_k)}{\partial\bar\alpha_j}\frac{\partial K_{n+1}}{\partial\alpha_j}\Big] \\
&= \rho_{k-1}^2\left[-\frac{\bar\alpha_{k-1}\bar\alpha_k}{2\rho_{k-1}}\frac{\partial K_{n+1}}{\partial\bar\alpha_{k-1}} + \frac{\alpha_{k-1}\bar\alpha_k}{2\rho_{k-1}}\frac{\partial K_{n+1}}{\partial\alpha_{k-1}}\right] - \rho_{k-1}\rho_k^2\frac{\partial K_{n+1}}{\partial\alpha_k} \\
&= \frac{\rho_{k-1}\bar\alpha_k}{2}\Big[\rho_{k-2}\bar\alpha_{k-1}\mathcal{E}^n_{k-1,k-2} - \alpha_{k-2}\bar\alpha_{k-1}\mathcal{E}^n_{k-1,k-1} \\
&\qquad - \alpha_{k-1}\bar\alpha_k\mathcal{E}^n_{k,k} - \alpha_{k-1}\rho_k\mathcal{E}^n_{k,k+1}\Big] \\
&\quad -\rho_{k-1}\rho_k\Big[\rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1} + \rho_k\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k+1} - \frac{\alpha_{k-1}\bar\alpha_k}{2}\mathcal{E}^n_{k,k+1} \\
&\qquad \frac{\rho_{k-1}\bar\alpha_k}{2}\mathcal{E}^n_{k-1,k+1} + \frac{\bar\alpha_k\rho_{k+1}}{2}\mathcal{E}^n_{k+2,k} + \frac{\bar\alpha_k\bar\alpha_{k+1}}{2}\mathcal{E}^n_{k+1,k}\Big].
\end{aligned}
$$

On the other hand,

$$
[\mathcal{E}, \mathcal{E}^{n+1}_+]_{k,k-1} = -\rho_{k-1}\rho_k\mathcal{E}^{n+1}_{k,k+1} + \frac{\rho_{k-1}\bar\alpha_k}{2}(\mathcal{E}^{n+1}_{k-1,k-1} - \mathcal{E}^{n+1}_{kk}).
$$

If we write $\mathcal{E}^{n+1} = \mathcal{E}\mathcal{E}^n$ and plug the appropriate entries into the expression above, we obtain

$$
\begin{aligned}
&i\{\mathcal{E}_{k,k-1}, K_{n+1}\} + [\mathcal{E}, \mathcal{E}^{n+1}_+]_{k,k-1} \\
&= -\frac{\rho_{k-1}\bar\alpha_k}{2}\Big[\alpha_{k-1}\rho_k\mathcal{E}^n_{k,k+1} + \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k} + \rho_k\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k} \\
&\qquad + \rho_{k-1}\bar\alpha_k\mathcal{E}^n_{k-1,k} - \rho_{k-2}\rho_{k-1}\mathcal{E}^n_{k,k-2} - \rho_{k-1}\bar\alpha_k\mathcal{E}^n_{k,k-1}\Big] \\
&= -\frac{\rho_{k-1}\bar\alpha_k}{2}\Big[(\mathcal{E}\mathcal{E}^n)_{kk} - (\mathcal{E}^n\mathcal{E})_{kk}\Big] = 0,
\end{aligned}
$$

which proves $(5.1.2)_{k,k-1}$.

Having done these two cases in some detail, we will just present the main steps in the computations for $(5.1.2)_{k+1,k-1}$ and $(5.1.2)_{k+1,k}$. A useful observation is that, since both sides of our identities are polynomials in the $\alpha$'s and $\bar\alpha$'s, one can more easily identify terms by keeping track of the powers of $\frac{1}{2}$ that occur.

The right-hand side of $(5.1.2)_{k+1,k-1}$ gives us

$$[\mathcal{E},\mathcal{E}_+^{n+1}]_{k+1,k-1} = \frac{\rho_{k-1}\rho_k}{2}\big(\mathcal{E}^{n+1}_{k-1,k-1} - \mathcal{E}^{n+1}_{k+1,k+1}\big).$$

So these are the terms we want to identify on the left-hand side:

$$\begin{aligned}
i\{\mathcal{E}_{k+1,k-1}, K_{n+1}\} &= \sum_j \rho_j^2\Big[\frac{\partial(\rho_{k-1}\rho_k)}{\partial\alpha_j}\frac{\partial K_{n+1}}{\partial\bar\alpha_j} - \frac{\partial(\rho_{k-1}\rho_k)}{\partial\bar\alpha_j}\frac{\partial K_{n+1}}{\partial\alpha_j}\Big] \\
&= \frac{\rho_{k-1}\rho_k}{2}\Big[ -\mathcal{E}^n_{k-1,k-2}\rho_{k-2}\bar\alpha_{k-1} - \mathcal{E}^n_{k-1,k-1}(-\alpha_{k-2}\bar\alpha_{k-1}) \\
&\qquad + (-\alpha_{k-1}\rho_k)\mathcal{E}^n_{k,k+1} - \mathcal{E}^n_{k-1,k}\rho_{k-1}\bar\alpha_k \\
&\qquad + (-\alpha_k\bar\alpha_{k+1})\mathcal{E}^n_{k+1,k+1} + (-\alpha_k\rho_{k+1})\mathcal{E}^n_{k+2,k+1}\Big] \\
&+ \frac{|\alpha_{k-1}|^2}{4}\Big[\rho_{k-2}\rho_k\mathcal{E}^n_{k,k-2} + \bar\alpha_k\rho_k\mathcal{E}^n_{k-1,k} \\
&\qquad + \rho_k^2\mathcal{E}^n_{k-1,k+1} - \alpha_{k-2}\rho_k\mathcal{E}^n_{k,k-1} \\
&\qquad - \rho_{k-2}\rho_k\mathcal{E}^n_{k,k-2} + \alpha_{k-2}\rho_k\mathcal{E}^n_{k,k-1} \\
&\qquad - \bar\alpha_k\rho_k\mathcal{E}^n_{k-1,k} - \rho_k^2\mathcal{E}^n_{k-1,k+1}\Big] \\
&+ \frac{|\alpha_k|^2}{4}\Big[\rho_{k-1}\rho_{k+1}\mathcal{E}^n_{k+2,k} + \rho_{k-1}\bar\alpha_{k+1}\mathcal{E}^n_{k-1,k} \\
&\qquad + \rho_{k-1}^2\mathcal{E}^n_{k-1,k+1} - \alpha_{k-1}\rho_{k-1}\mathcal{E}^n_{k,k+1} \\
&\qquad - \rho_{k-1}\rho_{k+1}\mathcal{E}^n_{k+2,k} - \rho_{k-1}\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k} \\
&\qquad - \alpha_{k-1}\rho_{k-1}\mathcal{E}^n_{k,k+1} - \rho_{k-1}^2\mathcal{E}^n_{k-1,k+1}\Big] \\
&= \frac{\rho_{k-1}\rho_k}{2}\Big[ -\big(\mathcal{E}^n\mathcal{E}\big)_{k-1,k-1} + \mathcal{E}^n_{k-1,k+1}\mathcal{E}_{k+1,k-1} \\
&\qquad + \big(\mathcal{E}\mathcal{E}^n\big)_{k+1,k+1} - \mathcal{E}_{k+1,k-1}\mathcal{E}^n_{k-1,k+1}\Big] \\
&= -[\mathcal{E},\mathcal{E}_+^{n+1}]_{k+1,k-1}\,.
\end{aligned}$$

Finally, we deal with $(5.1.2)_{k+1,k}$. As before, we notice that

$$[\mathcal{E},\mathcal{E}_+^{n+1}]_{k+1,k} = \rho_{k-1}\rho_k\mathcal{E}^{n+1}_{k-1,k} - \frac{\alpha_{k-1}\rho_k}{2}\big(\mathcal{E}^{n+1}_{kk} - \mathcal{E}^{n+1}_{k+1,k+1}\big).$$

The other side of the identity can be transformed as follows:

$$
\begin{aligned}
i\{\mathcal{E}_{k+1,k}, K_{n+1}\} &= -\sum_j \rho_j^2 \Big[\frac{\partial(\rho_k\alpha_{k-1})}{\partial\alpha_j}\frac{\partial K_{n+1}}{\partial\bar\alpha_j} - \frac{\partial(\rho_k\alpha_{k-1})}{\partial\bar\alpha_j}\frac{\partial K_{n+1}}{\partial\alpha_j}\Big] \\
&= -\rho_{k-1}\rho_k\Big[\mathcal{E}^n_{k-1,k-2}\rho_{k-2}\rho_{k-1} + \mathcal{E}^n_{k-1,k-1}(-\alpha_{k-2}\rho_{k-1})\Big] \\
&\quad + \frac{\alpha_{k-1}\rho_k}{2}\Big[\rho_{k-2}\rho_{k-1}\mathcal{E}^n_{k,k-2} + \rho_{k-1}\bar\alpha_k\mathcal{E}^n_{k-1,k} \\
&\qquad + \rho_{k-1}\rho_k\mathcal{E}^n_{k-1,k+1} - \alpha_{k-2}\rho_{k-1}\mathcal{E}^N_{k,k-1} \\
&\qquad + \rho_{k-1}\bar\alpha_k\mathcal{E}^n_{k-1,k} - \alpha_{k-1}\bar\alpha_k\mathcal{E}^n_{k,k} \\
&\qquad + \alpha_k\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k+1} + \alpha_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1}\Big] \\
&\quad + \frac{\alpha_{k-1}|\alpha_k|^2}{4}\Big[-\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k} - \rho_{k+1}\mathcal{E}^n_{k+2,k} - \rho_{k-1}\mathcal{E}^n_{k-1,k+1} \\
&\qquad + \alpha_{k-1}\mathcal{E}^n_{k,k+1} + \bar\alpha_{k+1}\mathcal{E}^n_{k+1,k} + \rho_{k+1}\mathcal{E}^n_{k+2,k} \\
&\qquad + \rho_{k-1}\mathcal{E}^n_{k-1,k+1} - \alpha_{k-1}\mathcal{E}^n_{k,k+1}\Big] \\
&= -\rho_{k-1}\rho_k\Big[\mathcal{E}^n_{k-1,k-2}\mathcal{E}_{k-2,k} + \mathcal{E}^n_{k-1,k-1}\mathcal{E}_{k-1,k}\Big] \\
&\quad + \frac{\alpha_{k-1}\rho_k}{2}\Big[\mathcal{E}^n_{k,k-2}\mathcal{E}_{k-2,k} + 2\rho_{k-1}\bar\alpha_k\mathcal{E}^n_{k-1,k} \\
&\qquad - \mathcal{E}_{k+1,k-1}\mathcal{E}^n_{k-1,k+1} + 2\rho_{k-1}\rho_k\mathcal{E}^n_{k-1,k+1} \\
&\qquad + \mathcal{E}^n_{k,k-1}\mathcal{E}_{k-1,k} + \mathcal{E}^n_{k,k}\mathcal{E}_{k,k} \\
&\qquad - \mathcal{E}_{k+1,k+1}\mathcal{E}^n_{k+1,k+1} - \mathcal{E}_{k+1,k+2}\mathcal{E}^n_{k+2,k+1}\Big] \\
&= -\rho_{k-1}\rho_k(\mathcal{E}^n\mathcal{E})_{k-1,k} + \frac{\alpha_{k-1}\rho_k}{2}\Big[(\mathcal{E}^n\mathcal{E})_{k,k} - (\mathcal{E}\mathcal{E}^n)_{k+1,k+1}\Big] \\
&= -[\mathcal{E}, \mathcal{E}^{n+1}_+]_{k+1,k}\,.
\end{aligned}
$$

This concludes the proof of $(5.1.2)_{k+1,k}$, and hence of relation (5.1.2).

The second part of the proof deals with relation (5.1.3):

$$\{\mathcal{E}, \bar K_{n+1}\} = [\mathcal{E}, i(\mathcal{E}^{n+1}_+)^*].$$

We shall proceed in very much the same way as with (5.1.2), while incorporating the necessary computational adjustments.

Again, we only have to check four relations; the only difference is that in this case $(5.1.3)_{k+1,k+1}$, $(5.1.3)_{k,k+1}$, $(5.1.3)_{k,k+2}$, and $(5.1.3)_{k+1,k+2}$ turn out to be computationally easier to verify.

As before, we start with the diagonal entry, $(5.1.3)_{k+1,k+1}$, that we shall prove in some detail.

We start by analyzing the left-hand side and observing that

$$\begin{aligned} i\{\mathcal{E}_{k+1,k+1}, \bar{K}_{n+1}\} &= \sum_j \rho_j^2 \Big[ \frac{\partial(-\alpha_k\bar{\alpha}_{k+1})}{\partial\alpha_j} \frac{\partial\bar{K}_{n+1}}{\partial\bar{\alpha}_j} - \frac{\partial(-\alpha_k\bar{\alpha}_{k+1})}{\partial\bar{\alpha}_j} \frac{\partial\bar{K}_{n+1}}{\partial\alpha_j} \Big] \\ &= -\rho_k^2 \bar{\alpha}_{k+1} \frac{\partial\bar{K}_{n+1}}{\partial\bar{\alpha}_k} + \rho_{k+1}^2 \alpha_k \frac{\partial\bar{K}_{n+1}}{\partial\alpha_{k+1}}. \end{aligned}$$

So by taking the complex conjugate in this relation, we get

$$\begin{aligned} \overline{i\{\mathcal{E}_{k+1,k+1}, \bar{K}_{n+1}\}} &= -\rho_k^2 \alpha_{k+1} \frac{\partial K_{n+1}}{\partial\alpha_k} + \rho_{k+1}^2 \bar{\alpha}_k \frac{\partial K_{n+1}}{\partial\bar{\alpha}_{k+1}} \\ &= \rho_k\alpha_{k+1} \Big[ \rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k+1} + \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1} \\ &\qquad + \frac{\bar{\alpha}_k\bar{\alpha}_{k+1}}{2}\mathcal{E}^n_{k+1,k} + \frac{\bar{\alpha}_k\rho_{k+1}}{2}\mathcal{E}^n_{k+2,k} \\ &\qquad + \frac{\bar{\alpha}_k\rho_{k-1}}{2}\mathcal{E}^n_{k-1,k+1} - \frac{\alpha_{k-1}\bar{\alpha}_k}{2}\mathcal{E}^n_{k,k+1} \Big] \\ &+ \bar{\alpha}_k\rho_{k+1} \Big[ \rho_k\rho_{k+1}\mathcal{E}^n_{k+1,k} - \alpha_k\rho_{k+1}\mathcal{E}^n_{k+1,k+1} \\ &\qquad - \frac{\rho_k\alpha_{k+1}}{2}\mathcal{E}^n_{k+2,k} - \frac{\alpha_{k+1}\bar{\alpha}_{k+2}}{2}\mathcal{E}^n_{k+1,k+2} \\ &\qquad - \frac{\alpha_{k+1}\rho_{k+2}}{2}\mathcal{E}^n_{k+1,k+3} + \frac{\alpha_k\alpha_{k+1}}{2}\mathcal{E}^n_{k+2,k+1} \Big]. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \overline{[\mathcal{E}, (\mathcal{E}_+^{n+1})^*]_{k+1,k+1}} &= \sum_{k-1\le j\le k+2} \overline{\mathcal{E}_{k+1,j}}(\mathcal{E}_+^{n+1})_{k+1,j} - \sum_{k\le j\le k+3} (\mathcal{E}_+^{n+1})_{j,k+1}\overline{\mathcal{E}_{j,k+1}} \\ &= \overline{\mathcal{E}_{k+1,k+2}}\mathcal{E}^{n+1}_{k+1,k+2} - \mathcal{E}^{n+1}_{k,k+1}\overline{\mathcal{E}_{k,k+1}} \\ &= -\bar{\alpha}_k\rho_{k+1}\mathcal{E}^{n+1}_{k+1,k+2} - \rho_k\alpha_{k+1}\mathcal{E}^{n+1}_{k,k+1}. \end{aligned}$$

Notice that

$$
\begin{aligned}
\overline{i\{\mathcal{E}_{k+1,k+1},\bar{K}_{n+1}\}} &= -\rho_k^2\alpha_{k+1}\frac{\partial K_{n+1}}{\partial\alpha_k} + \rho_{k+1}^2\bar{\alpha}_k\frac{\partial K_{n+1}}{\partial\bar{\alpha}_{k+1}} \\
&= \rho_k\alpha_{k+1}\Big[(\mathcal{E}\cdot\mathcal{E}^n)_{k,k+1} - \rho_{k-1}\bar{\alpha}_k\mathcal{E}^n_{k-1,k+1} + \alpha_{k-1}\bar{\alpha}_k\mathcal{E}^n_{k,k+1} \\
&\qquad + \frac{\bar{\alpha}_k\bar{\alpha}_{k+1}}{2}\mathcal{E}^n_{k+1,k} + \frac{\bar{\alpha}_k\rho_{k+1}}{2}\mathcal{E}^n_{k+2,k} \\
&\qquad + \frac{\bar{\alpha}_k\rho_{k-1}}{2}\mathcal{E}^n_{k-1,k+1} - \frac{\alpha_{k-1}\bar{\alpha}_k}{2}\mathcal{E}^n_{k,k+1}\Big] \\
&\quad + \bar{\alpha}_k\rho_{k+1}\Big[(\mathcal{E}^n\cdot\mathcal{E})_{k+1,k+2} + \alpha_{k+1}\bar{\alpha}_{k+2}\mathcal{E}^n_{k+1,k+2} + \alpha_{k+1}\rho_{k+2}\mathcal{E}^n_{k+1,k+3} \\
&\qquad - \frac{\rho_k\alpha_{k+1}}{2}\mathcal{E}^n_{k+2,k} - \frac{\alpha_{k+1}\bar{\alpha}_{k+2}}{2}\mathcal{E}^n_{k+1,k+2} \\
&\qquad - \frac{\alpha_{k+1}\rho_{k+2}}{2}\mathcal{E}^n_{k+1,k+3} + \frac{\alpha_k\alpha_{k+1}}{2}\mathcal{E}^n_{k+2,k+1}\Big].
\end{aligned}
$$

Therefore, we get that $\overline{i\{\mathcal{E}_{k+1,k+1},\bar{K}_{n+1}\} + [\mathcal{E},(\mathcal{E}_+^{n+1})^*]_{k+1,k+1}}$ equals

$$
\begin{aligned}
&\frac{\bar{\alpha}_k\alpha_{k+1}}{2}\Big[-\rho_{k-1}\rho_k\mathcal{E}^n_{k-1,k+1} + \alpha_{k-1}\rho_k\mathcal{E}^n_{k,k+1} + \alpha_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1} \\
&\qquad + \rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k} + \rho_{k+1}\bar{\alpha}_{k+2}\mathcal{E}^n_{k+1,k+2} + \rho_{k+1}\rho_{k+2}\mathcal{E}^n_{k+1,k+3}\Big] \\
&= \frac{\bar{\alpha}_k\alpha_{k+1}}{2}\Big[-(\mathcal{E}\cdot\mathcal{E}^n)_{k+1,k+1} + (\mathcal{E}^n\cdot\mathcal{E})_{k+1,k+1}\Big] = 0,
\end{aligned}
$$

which ends the proof of $(5.1.3)_{k+1,k+1}$.

We now turn to $(5.1.3)_{k,k+1}$:

$$
\overline{[\mathcal{E},(\mathcal{E}_+^{n+1})^*]}_{k,k+1} = \rho_k\rho_{k+1}\mathcal{E}^{n+1}_{k+1,k+2} + \frac{\rho_k\alpha_{k+1}}{2}\Big(\mathcal{E}^{n+1}_{k+1,k+1} - \mathcal{E}^{n+1}_{k,k}\Big).
$$

The left-hand side becomes

$$
\begin{aligned}
\overline{i\{\mathcal{E}_{k,k+1},\bar{K}_{n+1}\}} = \overline{i\{\rho_k\bar{\alpha}_{k+1},\bar{K}_{n+1}\}} \\
&= -\rho_k\rho_{k+1}\Big(\mathcal{E}^n_{k+1,k}\mathcal{E}_{k,k+2}+\mathcal{E}^n_{k+1,k+1}\mathcal{E}_{k+1,k+2}\Big) \\
&\quad -\frac{\rho_k\alpha_{k+1}}{2}\Big(-\mathcal{E}_{k,k+2}\mathcal{E}^n_{k+2,k}+\mathcal{E}^n_{k+1,k+3}\mathcal{E}_{k+3,k+1} \\
&\qquad -2\mathcal{E}^n_{k+1,k+3}\rho_{k+1}\rho_{k+2}+\mathcal{E}^n_{k+1,k+2}\mathcal{E}_{k+2,k+1} \\
&\qquad -2\mathcal{E}^n_{k+1,k+2}\rho_{k+1}\bar{\alpha}_{k+2}-\mathcal{E}_{k,k-1}\mathcal{E}^n_{k-1,k} \\
&\qquad -\mathcal{E}_{k,k}\mathcal{E}^n_{k,k}+\mathcal{E}^n_{k+1,k+1}\mathcal{E}_{k+1,k+1}\Big) \\
&= -\rho_k\rho_{k+1}(\mathcal{E}^n\mathcal{E})_{k,k+1}-\frac{\rho_k\alpha_{k+1}}{2}\Big((\mathcal{E}\mathcal{E}^n)_{k,k}+(\mathcal{E}^n\mathcal{E})_{k+1,k+1}\Big).
\end{aligned}
$$

So we find what we wanted:

$$
\{\mathcal{E}_{k,k+1},\bar{K}_{n+1}\} = i[\mathcal{E},(\mathcal{E}^{n+1}_+)^*]_{k,k+1}.
$$

The next entry that we analyze is $(5.1.3)_{k,k+2}$. Considering the right-hand side first, we get

$$
\begin{aligned}
\overline{[\mathcal{E},(\mathcal{E}^{n+1}_+)^*]_{k,k+2}} &= \sum_j \overline{\mathcal{E}_{k,j}}\cdot(\mathcal{E}^{n+1}_+)_{k+2,j}-\sum_j(\mathcal{E}^{n+1}_+)_{j,k}\cdot\overline{\mathcal{E}_{j,k+2}} \\
&= \frac{\rho_k\rho_{k+1}}{2}\big(\mathcal{E}^{n+1}_{k+2,k+2}-\mathcal{E}^{n+1}_{k,k}\big).
\end{aligned}
$$

If we look at the left-hand side now, we get

$$
\begin{aligned}
\overline{i\{\mathcal{E}_{k,k+2},\bar{K}_{n+1}\}} &= \overline{i\{\rho_k\rho_{k+1},\bar{K}_{n+1}\}} \\
&= \rho_k^2\Big[-\frac{\alpha_k\rho_{k+1}}{2\rho_k}\cdot\frac{\partial K_{n+1}}{\partial\alpha_k}+\frac{\bar\alpha_k\rho_{k+1}}{2\rho_k}\cdot\frac{\partial K_{n+1}}{\partial\bar\alpha_k}\Big] \\
&+ \rho_{k+1}^2\Big[-\frac{\alpha_{k+1}\rho_k}{2\rho_{k+1}}\cdot\frac{\partial K_{n+1}}{\partial\alpha_{k+1}}+\frac{\bar\alpha_{k+1}\rho_k}{2\rho_{k+1}}\cdot\frac{\partial K_{n+1}}{\partial\bar\alpha_{k+1}}\Big] \\
&= \frac{\rho_k\rho_{k+1}}{2}\Big[\alpha_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1}+\bar\alpha_k\rho_{k-1}\mathcal{E}^n_{k-1,k} \\
&\qquad -\alpha_{k-1}\bar\alpha_k\mathcal{E}^n_{k,k}+\alpha_{k+1}\bar\alpha_{k+2}\mathcal{E}^n_{k+2,k+2} \\
&\qquad +\alpha_{k+1}\rho_{k+2}\mathcal{E}^n_{k+2,k+3}+\bar\alpha_{k+1}\rho_k\mathcal{E}^n_{k+1,k}\Big] \\
&+ \frac{|\alpha_k|^2\rho_{k+1}}{4}\Big[\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k}+\rho_{k+1}\mathcal{E}^n_{k+2,k} \\
&\qquad +\rho_{k-1}\mathcal{E}^n_{k-1,k+1}-\alpha_{k-1}\mathcal{E}^n_{k,k+1} \\
&\qquad -\bar\alpha_{k+1}\mathcal{E}^n_{k+1,k}-\rho_{k+1}\mathcal{E}^n_{k+2,k} \\
&\qquad -\rho_{k-1}\mathcal{E}^n_{k-1,k+1}+\alpha_{k-1}\mathcal{E}^n_{k,k+1}\Big] \\
&+ \frac{|\alpha_{k+1}|^2\rho_k}{4}\Big[\rho_k\mathcal{E}^n_{k+2,k}-\alpha_k\mathcal{E}^n_{k+2,k+1} \\
&\qquad +\bar\alpha_{k+2}\mathcal{E}^n_{k+1,k+2}+\rho_{k+2}\mathcal{E}^n_{k+1,k+3} \\
&\qquad -\rho_k\mathcal{E}^n_{k+2,k}+\alpha_k\mathcal{E}^n_{k+2,k+1} \\
&\qquad -\bar\alpha_{k+2}\mathcal{E}^n_{k+1,k+2}-\rho_{k+2}\mathcal{E}^n_{k+1,k+3}\Big] \\
&= \frac{\rho_k\rho_{k+1}}{2}\Big[(\mathcal{E}\mathcal{E}^n)_{k,k}-\mathcal{E}_{k,k+2}\mathcal{E}^n_{k+2,k} \\
&\qquad -(\mathcal{E}^n\mathcal{E})_{k+2,k+2}+\mathcal{E}^n_{k+2,k}\mathcal{E}_{k,k+2}\Big] \\
&= -\overline{[\mathcal{E},(\mathcal{E}^{n+1}_+)^*]_{k,k+2}}\,,
\end{aligned}
$$

which immediately implies the equation $(5.1.3)_{k,k+2}$.

Finally, we turn to the last relation we have to prove, equation $(5.1.3)_{k+1,k+2}$. As above, we start with the right-hand side and observe that

$$
\overline{[\mathcal{E},(\mathcal{E}^n_+)^*]}_{k+1,k+2} = \frac{\bar\alpha_k\rho_{k+1}}{2}\big(\mathcal{E}^{n+1}_{k+1,k+1}-\mathcal{E}^{n+1}_{k+2,k+2}\big)-\rho_k\rho_{k+1}\mathcal{E}^{n+1}_{k,k+1}.
$$

Considering the left-hand side now, we have

$$
\begin{aligned}
&\overline{i\{\mathcal{E}_{k+1,k+2},\bar{K}_{n+1}\}}\\
&\quad= -\rho_k^2\rho_{k+1}\frac{\partial K_{n+1}}{\partial \alpha_k} - \bar{\alpha}_k\rho_{k+1}^2\Big[-\frac{\alpha_{k+1}}{2\rho_{k+1}}\frac{\partial K_{n+1}}{\partial \alpha_{k+1}} + \frac{\bar{\alpha}_{k+1}}{2\rho_{k+1}}\frac{\partial K_{n+1}}{\partial \bar{\alpha}_{k+1}}\Big]\\
&\quad= \rho_k\rho_{k+1}\Big[\rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k+1} + \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1}\Big]\\
&\qquad -\frac{\bar{\alpha}_k\rho_{k+1}}{2}\Big[-\rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k} - \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k}\\
&\qquad\qquad -\rho_k\rho_{k-1}\mathcal{E}^n_{k-1,k+1} + \alpha_{k-1}\rho_k\mathcal{E}^n_{k,k+1}\\
&\qquad\qquad +\alpha_{k+1}\bar{\alpha}_{k+2}\mathcal{E}^n_{k+2,k+2} + \alpha_{k+1}\rho_{k+2}\mathcal{E}^n_{k+2,k+3}\\
&\qquad\qquad +\bar{\alpha}_{k+1}\rho_k\mathcal{E}^n_{k+1,k} - \alpha_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k+1}\Big]\\
&\qquad -\frac{\bar{\alpha}_k|\alpha_{k+1}|^2}{4}\Big[\rho_k\mathcal{E}^n_{k+2,k} - \alpha_k\mathcal{E}^n_{k+2,k+1}\\
&\qquad\qquad +\bar{\alpha}_{k+2}\mathcal{E}^n_{k+1,k+2} + \rho_{k+2}\mathcal{E}^n_{k+1,k+3}\\
&\qquad\qquad -\rho_k\mathcal{E}^n_{k+2,k} - \bar{\alpha}_{k+2}\mathcal{E}^n_{k+1,k+2}\\
&\qquad\qquad -\rho_{k+2}\mathcal{E}^n_{k+1,k+3} + \alpha_k\mathcal{E}^n_{k+2,k+1}\Big]\\
&\quad= \rho_k\rho_{k+1}\Big[\rho_k\bar{\alpha}_{k+1}\mathcal{E}^n_{k+1,k+1} + \rho_k\rho_{k+1}\mathcal{E}^n_{k+2,k+1}\Big]\\
&\qquad -\frac{\bar{\alpha}_k\rho_{k+1}}{2}\Big[-(\mathcal{E}^n\mathcal{E})_{k+2,k+2} + (\mathcal{E}\mathcal{E}^n)_{k+1,k+1}\\
&\qquad\qquad -2\rho_{k-1}\rho_k\mathcal{E}^n_{k-1,k+1} + 2\alpha_{k-1}\rho_k\mathcal{E}^n_{k,k+1}\Big]\\
&\quad= \rho_k\rho_{k+1}\mathcal{E}^{n+1}_{k,k+1} - \frac{\bar{\alpha}_k\rho_{k+1}}{2}\big(\mathcal{E}^{n+1}_{k+1,k+1} - \mathcal{E}^{n+1}_{k+2,k+2}\big)\\
&\quad= -\overline{[\mathcal{E},(\mathcal{E}^n_+)^*]}_{k+1,k+2}\,,
\end{aligned}
$$

which implies

$$
i\{\mathcal{E}_{k+1,k+2},\bar{K}_{n+1}\} = -[\mathcal{E},(\mathcal{E}^n_+)^*]_{k+1,k+2},
$$

and hence $(5.1.3)_{k+1,k+2}$ holds.

# Chapter 6

# The Finite and Infinite Settings

## 6.1 Lax Pairs in the Finite Case

In this section we prove Lax pair representations for the finite Ablowitz-Ladik system.

We are interested in studying the system

$$-i\dot{\alpha}_j = \rho_j^2(\alpha_{j+1} + \alpha_{j-1})$$

for $0 \leq j \leq k-2$, with boundary conditions $\alpha_{-1} = \alpha_{k-1} = -1$. The idea behind finding Lax pairs for this system is to take one of the $\alpha$'s in the appropriate periodic problem to the boundary, and identify all the objects obtained in this way. As it turns out, they are all naturally related to both the Ablowitz-Ladik system and orthogonal polynomials on the circle, and can be defined independently of the periodic setting.

Let us elaborate. As presented in Section 3.4, if we start with a finitely supported measure $\mu$ on $S^1$, the associated Verblunsky coefficients are $\alpha_0, \ldots, \alpha_{k-2} \in \mathbb{D}$, $\alpha_{k-1} \in S^1$. The CMV matrix is, in this case, a unitary $k \times k$ matrix,

$$\mathcal{C}_f = \mathcal{L}_f \mathcal{M}_f$$

with

$$\mathcal{L}_f = \begin{pmatrix} \bar{\alpha}_0 & \rho_0 & & & \\ \rho_0 & -\alpha_0 & & & \\ & & \ddots & & \\ & & & \bar{\alpha}_{k-2} & \rho_{k-2} \\ & & & \rho_{k-2} & -\alpha_{k-2} \end{pmatrix}$$

and

$$\mathcal{M}_f = \begin{pmatrix} 1 & & & & \\ & \bar{\alpha}_1 & \rho_1 & & \\ & \rho_1 & -\alpha_1 & & \\ & & & \ddots & \\ & & & & \bar{\alpha}_{k-1} \end{pmatrix}.$$

If, in addition, we restrict our attention to the case when $\alpha_{k-1} = -1$, then we obtain the following connection between the finite and the periodic cases:

**Lemma 6.1.** *Let $k$ be even and $\mathcal{C}_f$ as above with $\alpha_{k-1} = -1$. Define a doubly-infinite set of Verblunsky coefficients by periodicity: $\alpha_{nk+j} = \alpha_j$ for all $n \in \mathbb{Z}$ and $0 \leq j \leq k-1$. Then the extended CMV matrix $\mathcal{E}$ associated to these $\alpha$'s has the direct sum decomposition*

$$\mathcal{E} = \bigoplus_{r \in \mathbb{Z}} S^r(\mathcal{C}_f), \tag{6.1.1}$$

*where $S : l^\infty(\mathbb{Z}) \to l^\infty(\mathbb{Z})$ is the right $k$-shift.*

*In particular, the following also hold:*

$$\mathcal{Q}_{(d)} = \bigoplus_{r=0}^{d-1} S^r(\mathcal{C}_f) \tag{6.1.2}$$

*and*

$$K_n(\mathcal{E}) = \frac{1}{n} \mathrm{Tr}(\mathcal{C}_f^n) \tag{6.1.3}$$

*for all $d \geq 1$ and $n \geq 1$.*

*Proof.* Relation (6.1.1) follows immediately if we observe that $\rho_{rk-1} = 0$ for all $r \in \mathbb{Z}$, and this implies that (see equation (3.5.1))

$$\tilde{\mathcal{M}} = \bigoplus_{r\in\mathbb{Z}} S^r(\mathcal{M}_f).$$

By periodicity, we always have

$$\tilde{\mathcal{L}} = \bigoplus_{r\in\mathbb{Z}} S^r(\mathcal{L}_f).$$

So (6.1.1) follows from the definition of $\mathcal{C}_f = \mathcal{L}_f\mathcal{M}_f$. Likewise, (6.1.2) is just the restriction of (6.1.1) to $X_{(d)}$. So then

$$\mathcal{Q}^n_{(d)} = \bigoplus_{r=0}^{d-1} S^r(\mathcal{C}^n_f)$$

and, by taking the trace, we get (6.1.3). □

Note also that the Poisson bracket (4.1.2) separates the $\alpha$'s, and hence it naturally restricts to the space of $(\alpha_0, \ldots, \alpha_{k-2}, \alpha_{k-1} = -1) \in \mathbb{D}^{k-1}$. If two functions $f$ and $g$ depend only on $\alpha_0, \ldots, \alpha_{k-2}$, then

$$\begin{aligned}\{f, g\} &= \frac{1}{2}\sum_{j=0}^{k-2} \rho_j^2 \left[\frac{\partial f}{\partial u_j}\frac{\partial g}{\partial v_j} - \frac{\partial f}{\partial v_j}\frac{\partial g}{\partial u_j}\right] \\ &= i\sum_{j=0}^{k-2} \rho_j^2 \left[\frac{\partial f}{\partial \bar{\alpha}_j}\frac{\partial g}{\partial \alpha_j} - \frac{\partial f}{\partial \alpha_j}\frac{\partial g}{\partial \bar{\alpha}_j}\right],\end{aligned}$$

where, as before, $\alpha_j = u_j + iv_j$ for all $0 \le j \le k-2$.

So the next theorem is an immediate consequence of Theorem 5.1:

**Theorem 6.2.** *Let*

$$K^f_n = K_n(\mathcal{C}_f) = \frac{1}{n}\mathrm{Tr}(\mathcal{C}^n_f)$$

*for all $n \ge 1$. Then the Lax pairs associated to these Hamiltonians are given by*

$$\{\mathcal{C}_f, K^f_n\} = [\mathcal{C}_f, i(\mathcal{C}^n_f)_+]$$

*and*

$$\{\mathcal{C}_f, \bar{K}_n^f\} = [\mathcal{C}_f, i((\mathcal{C}_f^n)_+)^*]$$

*for all* $n \geq 1$.

*Or, in terms of real-valued flows, we have*

$$\{\mathcal{C}_f, 2\,\mathrm{Re}(K_n^f)\} = [\mathcal{C}_f, i(\mathcal{C}_f^n)_+ + i((\mathcal{C}_f^n)_+)^*]$$

*and*

$$\{\mathcal{C}_f, 2\,\mathrm{Im}(K_n^f)\} = [\mathcal{C}_f, (\mathcal{C}_f^n)_+ - ((\mathcal{C}_f^n)_+)^*]$$

*for all* $n \geq 1$.

As in the periodic case, since $K_n^f$ is the trace of $\mathcal{C}_f^n$, we obtain Poisson commutativity of the Hamiltonians:

**Corollary 6.3.** *For all* $m, n \geq 1$, *we have*

$$\{K_n^f, \mathrm{Re}(K_m^f)\} = \{K_n^f, \mathrm{Im}(K_m^f)\} = 0$$

*and*

$$\{K_n^f, K_m^f\} = \{K_n^f, \bar{K}_m^f\} = 0.$$

*Remark* 6.4. Note that, since $\alpha_{k-1} \equiv -1$, we get $\rho_{k-1} \equiv 0$, and so $K_0 = \prod_{j=0}^{k-1} \rho_j^2 \equiv 0$ on $\mathbb{D}^{k-1}$. But if we define

$$K_0^f = \prod_{j=0}^{k-2} \rho_j^2,$$

then

$$\{\alpha_m, K_0^f\} = iK_0^f \alpha_m$$

or

$$\{\alpha_m, \log(K_0^f)\} = i\alpha_m.$$

But, even though $K_0^f$ acts on the $\alpha$'s in the finite case in the same way as $K_0$ does in the periodic case, there exists no Lax pair representation for $K_0^f$ in terms of $\mathcal{C}_f$. The reason is that

$$\mathrm{Tr}\{\mathcal{C}_f, K_0^f\} = -iK_0^f(\bar{\alpha}_0 - \alpha_{k-2}),$$

which is not identically zero on $\mathbb{D}^{k-1}$, while the trace of a commutator is always zero.

## 6.2 Lax Pairs in the Infinite Case

Finally, we deal with the infinite defocusing Ablowitz-Ladik system. By this, we mean that we consider the system whose first equation is

$$i\dot{\alpha}_j = \rho_j^2(\alpha_{j+1} + \alpha_{j-1})$$

for all $j \geq 0$, with the boundary condition $\alpha_{-1} = 0$. The idea behind constructing Lax pairs for this system is to use the finite AL result. Since each entry in a fixed power of the CMV matrix depends on only a bounded number of $\alpha$'s, extending the finite Lax pairs to the infinite case only requires an appropriate definition of the "infinite" Hamiltonians $K_n^i$ for all $n \geq 1$.

Let us explain these claims: Fix $n_0 \geq 1$ and $j_0, m_0 \geq 0$. Consider the finite problem with $k$ very large ($k \geq 20(j_0 + k_0 + n_0)$ is sufficient, though a much more precise bound can be found). In this case,

$$\mathcal{C}_{j,m}^n = (\mathcal{C}_f)_{j,m}^n$$

for all $0 \leq j, m \leq j_0 + 4, m_0 + 4$ respectively, and $1 \leq n \leq n_0$. Say we can define a $K_n^i$ such that its dependence on the first $k$ $\alpha$'s is the same as that of $K_n^f$. Then, for $0 \leq j, m \leq j_0 + 4, m_0 + 4$ respectively, we can replace "finite" by "infinite" in $(6.2)_{j_0,m_0}$ and $(6.2)_{j_0,m_0}$.

The last element we need is $K_n^i$, the $n^{\text{th}}$ Hamiltonian for the infinite problem. It is a function defined on sequences $\{\alpha_j\}_{j\geq 0}$ of numbers inside the unit disk, having a certain decay. The condition it must satisfy is

$$K_n^i(\{\alpha_0, \alpha_1 \ldots, \alpha_{k-1} = -1, 0, 0, \ldots\}) = K_n^f(\{\alpha_0, \alpha_1, \ldots, \alpha_{k-1} = -1\}).$$

Given that

$$K_n^f(\mathcal{C}_f) = \frac{1}{n}\mathrm{Tr}(\mathcal{C}_f^n),$$

a natural guess for $K_n^i$ would be

$$K_n^i(\mathcal{C}) = \frac{1}{n}\text{``Tr''}(\mathcal{C}^n).$$

But recall that the CMV matrix is unitary, so it is not trace class. Nonetheless, given the special structure of $\mathcal{C}$, we can define our Hamiltonian $K_n^i$ following this intuition as the sum of the diagonal entries of $\mathcal{C}^n$. While this statement will be rigorously proved in the following lemma, the reason why one can sum the series of diagonal entries is that all of these entries have the same structure for shifted $\alpha$'s: They are the sum of a bounded number of "monomials." By "monomial" we mean a finite product of $\alpha$'s and $\rho$'s. All the monomials that appear as terms in the diagonal entries contain at least one $\alpha$ factor. Since all the $\alpha$'s and $\rho$'s have absolute values less than 1, and if we assume $l^1$-decay of the sequence of coefficients, the one $\alpha$ factor in each monomial will ensure convergence of the whole series.

The next lemma and its proof explore in more detail the structure of the entries of powers of the CMV matrix and its consequences for the definition of Hamiltonians in the infinite case.

*Remark* 6.5. We must observe that Theorem 4.2.14 from [28] proves that, if

$$\{\alpha_j\}_{j\geq 0} \in l^2,$$

then the operator $\mathcal{C}^n - \mathcal{C}_0^n$ is trace class for all $n \geq 2$, and

$$\mathrm{Tr}(\mathcal{C}^n - \mathcal{C}_0^n) = \sum_{j=0}^{\infty}(\mathcal{C}^n)_{jj}.$$

Here $\mathcal{C}_0$ denotes the free CMV matrix, given by all Verblunsky coefficients zero. Also in this case note that, by Cauchy-Schwarz, the series

$$\sum_{j\geq 0} \mathcal{C}_{jj}$$

is absolutely convergent.

While this proves that we can define our Hamiltonians in the case of $l^2$ decay of the coefficients, we are also interested in knowing how a Hamiltonian depends on any $\alpha$. More precisely, in order to apply the symplectic form to a Hamiltonian which is the sum of a series (for which we do not know uniform convergence), we will prove that, for each $n \geq 1$, there exists a constant $c(n)$, depending only on $n$, such that for any $j \geq 0$ there are at most $c(n)$ terms in the series defining $K_n^i$ which depend on $\alpha_j$. Our proof is direct and fairly brutal, but it also describes the general structure of the entries of a power of the CMV matrix. In particular, it emphasizes the repetitive structure of CMV matrices.

**Lemma 6.6.** *Let $\{\alpha_j\}_{j\geq 0} \in l^1(\mathbb{N})$ be a sequence of coefficients with $\alpha_j \in \mathbb{D}$ for all $j \geq 0$. Let $\mathcal{C}$ be the CMV matrix associated to these coefficients. Then the series*

$$\sum_{k\geq 0} \mathcal{C}^n_{k,k} \tag{6.2.1}$$

*converges absolutely for any $n \geq 1$.*

*Moreover, for any $k \geq 0$, we have that $\mathcal{C}^n_{k,k}$ depends only on $\alpha_{k-(2n-1)}, \ldots, \alpha_{k+2n-1}$, where all the $\alpha$'s with negative indices are assumed to be identically zero.*

*Proof.* We prove these statements by making two important observations.

The first refers to the general, doubly-infinite case. Let $\{\alpha_j\}_{j\in\mathbb{Z}}$ be a sequence of complex numbers in $\mathbb{D}$, and $\mathcal{E}$ be the associated extended CMV matrix. Notice that the structure of $\mathcal{E}$ is such that there exist functions $f^e_{1,d_1}$ and $f^o_{1,d_1}$ defined on $\mathbb{D}^3$ with

$$\mathcal{E}_{j,k} = f^e_{1,d_1}(\alpha_{j-1}, \alpha_j, \alpha_{j+1})$$

for all $j$ even and $j - k = d_1$, and

$$\mathcal{E}_{j+1,k} = f^{o}_{1,d_1}(\alpha_{j-1}, \alpha_j, \alpha_{j+1})$$

for all $j$ even and $(j+1) - k = d_1$. Here $e$ and $o$ are used to denote "even" or "odd" respectively, and $-2 \leq d_1 \leq 1$.

Using this simple remark, one can prove by induction that for all $n \geq 1$, there exist functions

$$f^{e}_{n,d_n}, f^{o}_{n,d_n} : \mathbb{D}^{4n-1} \to \mathbb{C}$$

with $-2n \leq d_n \leq 2n - 1$ such that

$$\mathcal{E}^{n}_{j,k} = f^{e}_{n,j-k}(\alpha_{j-(2n-1)}, \dots, \alpha_{j+(2n-1)})$$

for $j$ even and $-2n \leq j - k \leq 2n - 1$,

$$\mathcal{E}^{n}_{j+1,k} = f^{o}_{n,j-k+1}(\alpha_{j-(2n-1)}, \dots, \alpha_{j+(2n-1)})$$

for $j$ even and $-2n \leq j + 1 - k \leq 2n - 1$, and

$$\mathcal{E}^{n}_{l,m} = 0$$

for all the other indices $(l, m)$.

Moreover, for $|d_n| \leq 2n - 1$, each such function $f^{e/o}_{n,d_n}$ is a sum of at most $4^n$ monomials, that is, products of $\alpha$'s and $\rho$'s, and each monomial contains at least one $\alpha$ factor. The only entries containing only $\rho$'s are the extreme ones:

$$\begin{aligned} \mathcal{E}^{n}_{j,j+2n} &= f^{e}_{n,-2n}(\alpha_{j-(2n-1)}, \dots, \alpha_{j+(2n-1)}) \\ &= \rho_j \rho_{j+1} \cdots \rho_{j+2n-1} \end{aligned} \tag{6.2.2}$$

and

$$\begin{aligned} \mathcal{E}^n_{j+1,j-(2n-1)} &= f^o_{n,-2n}(\alpha_{j-(2n-1)}, \cdots, \alpha_{j+(2n-1)}) \\ &= \rho_{j-(2n-1)}\rho_{j-(2n-2)} \cdots \rho_j \end{aligned} \tag{6.2.3}$$

for all $j$ even.

Fix $n \geq 1$. Each monomial in $f^{e/o}_{n,d_n}$ is bounded by the absolute value of one of the $\alpha$'s involved, and there are $4^n$ such monomial terms in each sum. Putting all of this together, we get that, for all $j$ even,

$$|\mathcal{E}^n_{j,j}|\,,\, |\mathcal{E}^n_{j+1,j+1}| \leq 4^n(|\alpha_{j-(2n-1)}| + \cdots + |\alpha_{j+2n-1}|).$$

The second observation we need to make in order to conclude the convergence of the series (6.2.1) concerns what changes in all of these formulae when we introduce a boundary condition $\alpha_{-1} = -1$.

From the discussion above, we see that actually

$$\mathcal{C}^n_{j,k} = \mathcal{E}^n_{j,k}$$

for $j, k \geq 4n$, as these entries only depend on $\alpha$'s with positive indices. (As we remarked earlier, these bounds are not optimal, but they are certainly sufficient for our purposes.) Hence we also get

$$|\mathcal{C}^n_{j,j}|\,,\, |\mathcal{C}^n_{j+1,j+1}| \leq 4^n(|\alpha_{j-(2n-1)}| + \cdots + |\alpha_{j+2n-1}|)$$

for $j \geq 4n$ even. So, since the sequence of $\alpha$'s is in $l^1$, we get that, for any $n \geq 1$, the series (6.2.1) converges absolutely. □

We can now define our Hamiltonians as

$$K^i_n = K^i_n(\mathcal{C}) = \sum_{k=0}^{\infty} \mathcal{C}^n_{k,k}. \tag{6.2.4}$$

By Remark 6.5, they are well-defined for $\{\alpha_j\}_{j\geq 0} \in l^2$, and, for any fixed $j \geq 0$, only a finite number of terms in the series depends on $\alpha_j$. We can therefore state our main theorem in the infinite case:

**Theorem 6.7.** *Let $\{\alpha_j\}_{j\geq 0}$ be an $l^2(\mathbb{N})$ sequence of complex numbers inside the unit disk, $\mathcal{C}$ the associated CMV matrix, and $K_n^i$ the function defined by (6.2.4). Then the Lax pairs associated to these Hamiltonians are given by*

$$\{\mathcal{C}, K_n^i\} = [\mathcal{C}, i\mathcal{C}_+^n] \tag{6.2.5}$$

*and*

$$\{\mathcal{C}, \bar{K}_n^i\} = [\mathcal{C}, i(\mathcal{C}_+^n)^*] \tag{6.2.6}$$

*for all $n \geq 1$.*

*Or, in terms of real-valued flows, we have*

$$\{\mathcal{C}, 2\,\mathrm{Re}(K_n^i)\} = [\mathcal{C}, i\mathcal{C}_+^n + i(\mathcal{C}_+^n)^*] \tag{6.2.7}$$

*and*

$$\{\mathcal{C}, 2\,\mathrm{Im}(K_n^i)\} = [\mathcal{C}, \mathcal{C}_+^n - (\mathcal{C}_+^n)^*] \tag{6.2.8}$$

*for all $n \geq 1$.*

*Proof.* For each fixed $n$ and entry $(j, l)$, there exists a $k$ large enough such that all the entries of $\mathcal{C}$ and $\mathcal{C}^n$ that appear in $(6.2.5)_{j,l}$ and $(6.2.6)_{j,l}$ are equal to the entries of $\mathcal{C}_f$ and $\mathcal{C}_f^n$, respectively, in the corresponding finite Lax pairs.

Moreover, since $\mathcal{C}_{j,l}$ depends on two $\alpha$'s, and these appear in only finitely many of the terms in $K_n^i$, the Poisson brackets on the left-hand side are well-defined finite sums and equal the corresponding Poisson brackets in the finite case.

Therefore, the results of Theorem 6.7 follow directly from Theorem 6.2 and the observations in the proof of Lemma 6.6. □

*Remark* 6.8. As in the finite case, we define

$$K_0^i = \prod_{j=0}^{\infty} \rho_j^2.$$

Recall that

$$\rho_j^2 = 1 - |\alpha_j|^2 \leq 2(1 - |\alpha_j|)$$

and $\{\alpha_j\}_{j\geq 0} \in l^1(\mathbb{N})$. Therefore, $K_0^i$ is well-defined and positive; also the following Poisson bracket makes sense:

$$\{\alpha_j, \log(K_0^i)\} = -2i\alpha_j.$$

But, as in the finite case, we cannot hope to find a Lax pair representation for the flow generated by $K_0^i$ in terms of $\mathcal{C}$. The dependence of $\sum_{j\geq 0}\{\mathcal{C}_{jj}, K_0^i\}$ on $\bar{\alpha}_0$ is nontrivial, while $\sum_{j\geq 0}[\mathcal{C}, \mathcal{A}]_{jj}$ is identically zero for any infinite matrix $\mathcal{A}$ for which the commutator makes sense.

## 6.3 The Schur Flows and Their Relation to the Toda Lattice

In this section we want to present another system of nonlinear differential-difference equations

$$\dot{\alpha}_n = (1 - \alpha_n^2)(\alpha_{n+1} - \alpha_{n-1}), \quad \{\alpha_j\} \subset (-1, 1), \tag{6.3.1}$$

known as the discrete modified KdV equation (see [1] and [15]) or the equation of the Schur flows (see [11]). This system's main interest lies in its connection to the Toda and Volterra (or Kac-van Moerbeke) lattices.

More precisely, the Schur flows appear in the work of Ammar and Gragg [4] as an evolution equation on the Verblunsky coefficients obtained by transferring the Toda equation via the Geronimus relations from the $a$'s and $b$'s. As such, it is an evolution on real $\alpha$'s. One can then relax this condition and think of it as an evolution on complex coefficients, which preserves reality (that is, if the $\alpha$'s are real at time 0, then they remain real for all times).

As it turns out, this description applies to "half" of the Ablowitz-Ladik flows, the ones generated by $\mathrm{Im}(K_n)$ for all $n \geq 1$. Indeed, the evolution on the Verblunsky coefficients under these Hamiltonians is equivalent to the Lax pairs

$$\{\mathcal{C}, 2\,\mathrm{Im}(K_n)\} = [\mathcal{C}, (\mathcal{C}^n)_+ - ((\mathcal{C}^n)_+)^*],$$

which preserve reality of the $\alpha$'s. (We consider here both the finite and infinite AL.) So it is natural to ask ourselves what are the corresponding evolutions of the $a$'s and $b$'s.

Here we will only show that (6.3.1) is equivalent (up to a multiplicative constant) via the Geronimus relations (3.6.2) to the Toda evolution. In other words, as the notation suggests, we think of the $\alpha$'s as real Verblunsky coefficients. Moreover, we recover (6.3.1) as the evolution generated by $-2\,\mathrm{Im}(K_1)$ under the AL Poisson bracket (4.1.2). In fact, in joint work with Rowan Killip [21], we prove that, in the finite case, all the evolutions induced by $\mathrm{Im}(K_n)$ become, via the Geronimus relations, simple combinations of the evolutions in the Toda hierarchy.

Let us first see what the flow generated by $-2\,\mathrm{Im}(K_1)$ is. Recall that

$$K_1 = -\sum_{k=0}^{p-1} \alpha_k \bar{\alpha}_{k+1}$$

and

$$\{\alpha_j, K_1\} = i\rho_j^2 \alpha_{j-1},$$

$$\{\alpha_j, \bar{K}_1\} = i\rho_j^2 \alpha_{j+1}.$$

Combining these two equations, we easily find

$$\dot{\alpha}_j = \{\alpha_j, -2\,\mathrm{Im}(K_1)\} = \rho_j^2(\alpha_{j+1} - \alpha_{j-1}).$$

If $\alpha_j(0) \in \mathbb{R}$ for all $j$, then $\alpha_j \in \mathbb{R}$ for all $j$ and for all time. Since in this case

$$\rho_j^2 = 1 - |\alpha_j|^2 = 1 - \alpha_j^2,$$

we recover exactly (6.3.1).

Recall that, as presented in Section 3.6, measures on the circle which are invariant under complex conjugation correspond via (3.6.1) to measures on $[-2, 2]$. This happens exactly when the Verblunsky coefficients are real, and in this case they are related to the recurrence coefficients on the interval via the Geronimus relations (3.6.2):

$$\begin{aligned} b_{k+1} &= (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2} \\ a_{k+1}^2 &= (1 - \alpha_{2k-1})(1 - \alpha_k^2)(1 + \alpha_{2k+1}). \end{aligned}$$

Let the Verblunsky coefficients evolve according to the Schur flows (6.3.1). We want to deduce the evolution equations for the $a$'s and $b$'s.

$$\begin{aligned} \dot{b}_{k+1} &= (1 - \alpha_{2k-1})\dot{\alpha}_{2k} - \dot{\alpha}_{2k-1}\alpha_{2k} - \dot{\alpha}_{2k-1}\alpha_{2k-2} - (1 + \alpha_{2k-1}\dot{\alpha}_{2k-2}) \\ &= (1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(\alpha_{2k+1} - \alpha_{2k-1}) \\ &\quad - (1 - \alpha_{2k-1}^2)(\alpha_{2k} - \alpha_{2k-2})(\alpha_{2k} + \alpha_{2k-2}) \\ &\quad - (1 + \alpha_{2k-1})(1 - \alpha_{2k-2}^2)(\alpha_{2k-1} - \alpha_{2k-3}) \\ &= (1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1}) \\ &\quad - (1 - \alpha_{2k-3})(1 - \alpha_{2k-2}^2)(1 + \alpha_{2k-1}) \\ &\quad - (1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k-1}) \\ &\quad + (1 + \alpha_{2k-1})(1 - \alpha_{2k-2}^2)(1 - \alpha_{2k-1}) \\ &\quad - (1 - \alpha_{2k-1}^2)(\alpha_{2k}^2 - \alpha_{2k-2}^2) \\ &= a_{k+1}^2 - a_k^2 \end{aligned}$$

and, by a similar, but somewhat tedious calculation,

$$2a_{k+1}\dot{a}_{k+1} = a_{k+1}^2[b_{k+2} - b_{k+1}].$$

So we find

$$\begin{aligned} \dot{b}_{k+1} &= a_{k+1}^2 - a_k^2 \\ \dot{a}_{k+1} &= \frac{1}{2}a_{k+1}(b_{k+2} - b_{k+1}). \end{aligned}$$

Let

$$\mathcal{B}_k = -\frac{1}{2}b_k, \quad \mathcal{A}_k = -\frac{1}{2}a_k.$$

Then we get

$$\begin{aligned} \dot{\mathcal{B}}_{k+1} &= 2(\mathcal{A}_k^2 - \mathcal{A}_{k+1}^2) \\ \dot{\mathcal{A}}_{k+1} &= \mathcal{A}_{k+1}(\mathcal{B}_{k+1} - \mathcal{B}_{k+2}), \end{aligned}$$

which are exactly the equations of the Toda lattice (2.3.3) and (2.3.4) written for $\mathcal{B}$ and $\mathcal{A}$.

# Bibliography

[1] M. J. Ablowitz, J. F. Ladik, Nonlinear differential-difference equations. *J. Math. Phys.* **16** (1975), 598–603.

[2] M. J. Ablowitz, J. F. Ladik, Nonlinear differential-difference equations and Fourier analysis. *J. Math. Phys.* **17** (1976), 1011–1018.

[3] M. J. Ablowitz, B. Prinari, A. D. Trubach, *Discrete and Continuous Nonlinear Schrödinger Systems.* London Mathematical Society Lecture Note Series, Vol. 302, Cambridge University Press, Cambridge, 2004.

[4] G. S. Ammar, W. B. Gragg, Schur flows for orthogonal Hessenberg matrices, *Hamiltonian and Gradient Flows, Algorithms and Control*, 27–34, Fields Inst. Commun., **3**, American Mathematical Society, Providence, RI, 1994.

[5] V. I. Arnold, V. V. Kozlov, A. I. Neishtandt, Mathematical aspects of classical and celestial mechanics, *Dynamical Systems, III*, pp. vii–xiv and 1–291, Encyclopaedia Math. Sci., **3**, Springer, Berlin, 1993.

[6] E. Berriochoa, A. Cachafeiro, J. García-Amor, Generalizations of the Szegö transformation interval-unit circle, preprint.

[7] M. J. Cantero, L. Moral, L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.* **362** (2003), 29–56.

[8] M. J. Cantero, L. Moral, L. Velázquez, Minimal representations of unitary operators and orthogonal polynomials on the unit circle, preprint, arXiv:math.CA/0405246.

[9] P. Deift, Integrable Hamiltonian systems, *Dynamical Systems and Probabilistic Methods in Partial Differential Equations* (Berkeley, CA, 1994), pp. 103–138, Lectures in Appl. Math., **31**, American Mathematical Society, Providence, RI, 1996.

[10] L. Fadeev, L. Takhtajan, *Hamiltonian Methods in the Theory of Solitons,* Springer, Berlin, 1987.

[11] L. Faybusovich, M. Gekhtman, On Schur flows. *J. Phys. A: Math. Gen.* **32** (1999), 4671–4680.

[12] E. Fermi, J. Pasta, S. Ulam, Studies of nonlinear problems, *Collected Works of Enrico Fermi,* Vol. II, pp. 978–988, University of Chicago Press, Chicago, 1965.

[13] H. Flaschka, Discrete and periodic illustrations of some aspects of the inverse method, *Dynamical Systems, Theory and Applications* (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), pp. 441–466, Lecture Notes in Phys., **38**, Springer, Berlin, 1975.

[14] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura, Korteweg de Vries equation and generalizations, *Phys. Rev. Letters* **19** (1967), 1095–1097.

[15] M. Gekhtman, Non-Abelian nonlinear lattice equations on finite interval, *J. Phys. A: Math. Gen.* **26** (1993), 6303–6317.

[16] J. Geronimo, F. Gesztesy, H. Holden, Algebro-geometric solutions of a discrete system related to the trigonometric moment problem, *Comm. Math. Phys.* (to appear), arXiv:math.SP/0408073.

[17] Ya. L. Geronimus, On the trigonometric moment problem, *Ann. of Math. (2)* **47** (1946), 742–761.

[18] Ya. L. Geronimus, *Polynomials Orthogonal on a Circle and Their Applications,* Amer. Math. Soc. Translation **1954** (1954), no. 104, 79pp.

[19] F. Gesztesy, H. Holden, *Soliton Equations and Their Algebro-Geometric Solutions. Volume II: (1+1)-Dimensional Discrete Models,* Cambridge Studies in Adv. Math., Cambridge University Press, Cambridge, in preparation.

[20] R. Killip, I. Nenciu, Matrix models for circular ensembles, *Int. Math. Res. Not.* **50** (2004), 2665–2701.

[21] R. Killip, I. Nenciu, CMV: The unitary analogue of Jacobi matrices, in preparation.

[22] P. Lax, Integrals of nonlinear equations of evolution and solitary waves, *Comm. Pure Appl. Math.* **21** (1968), 467–490.

[23] H. McKean, P. van Moerbeke, The spectrum of Hill's equation, *Invent. Math.* **30** (1975), no. 3, 217–274.

[24] P. D. Miller, N. M. Ercolani, I. M. Krichever, C. D. Levermore, Finite genus solutions to the Ablowitz-Ladik equations, *Comm. Pure Appl. Math.* **48** (1995), 1369–1440.

[25] P. D. Miller, A. C. Scott, J. Carr, J. C. Eilbeck, Binding energies for discrete nonlinear Schrödinger equations, *Phys. Scripta* **44** (1991), 509–516.

[26] I. Nenciu, Lax pairs for the Ablowitz-Ladik system via orthogonal polynomials on the unit circle, *Int. Math. Res. Not.* **11** (2005), 647–686.

[27] D. E. Rourke, Elementary Bäcklund transformations for a discrete Ablowitz-Ladik eigenvalue problem, *J. Phys. A* **37** (2004), 2693–2708.

[28] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[29] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[30] B. Simon, OPUC on one foot, preprint, arXiv:math.SP/0502485.

[31] G. Szegő, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, Rhode Island, 1975.

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

[33] M. Toda, *Theory of Nonlinear Lattices*, second edition, Springer Series in Solid-State Sciences, **20**, Springer-Verlag, Berlin, 1989.

[34] K. L. Vaninsky, Symplectic structures and volume elements in the function space for the cubic Schrödinger equation, *Duke Math. J.* **92** (1998), 381–402.

[35] K. L. Vaninsky, An additional Gibbs state for the cubic Schrödinger equation on the circle, *Comm. Pure Appl. Math.* **54** (2001), 537–582.

[36] P. van Moerbeke, The spectrum of Jacobi matrices. *Invent. Math.* **37** (1976), 45–81.

[37] V. E. Zakharov, A. B. Shabat, Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media, *Soviet Physics JETP* **34** (1972), no. 1, 62–69; *translated from Z. Èksper. Teoret. Fiz.* **61** (1971), no. 1, 118–134 (Russian).