

# CMV: THE UNITARY ANALOGUE OF JACOBI MATRICES

ROWAN KILLIP<sup>1</sup> AND IRINA NENCIU

*Happy 60th birthday, Percy Deift.*

ABSTRACT. We discuss a number of properties of CMV matrices, by which we mean the class of unitary matrices recently introduced by Cantero, Moral, and Velazquez. We argue that they play an equivalent role among unitary matrices to that of Jacobi matrices among all Hermitian matrices. In particular, we describe the analogues of well-known properties of Jacobi matrices: foliation by co-adjoint orbits, a natural symplectic structure, algorithmic reduction to this shape, Lax representation for an integrable lattice system (Ablowitz-Ladik), and the relation to orthogonal polynomials.

As offshoots of our analysis, we will construct action/angle variables for the finite Ablowitz-Ladik hierarchy and describe the long-time behaviour of this system.

## 1. INTRODUCTION

For many reasons, it is natural to regard Jacobi matrices as occupying a certain privileged place among all Hermitian matrices. By ‘Jacobi matrix’ we mean a tri-diagonal matrix

$$J = \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 & \ddots & & \\ & \ddots & \ddots & & \\ & & & a_{n-1} & \\ & & & a_{n-1} & b_n \end{bmatrix} \quad (1)$$

with  $a_j > 0$ ,  $b_j \in \mathbb{R}$ . The main purpose of this paper is explain why we consider a family of unitary matrices introduced recently by Cantero, Moral, and Velazquez (see [6]) as playing the corresponding role among unitary matrices.

**Definition 1.1.** Given coefficients  $\alpha_0, \dots, \alpha_{n-2}$  in  $\mathbb{D}$  and  $\alpha_{n-1} \in S^1$ , let  $\rho_k = \sqrt{1 - |\alpha_k|^2}$ , and define  $2 \times 2$  matrices

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

for  $0 \leq k \leq n-2$ , while  $\Xi_{-1} = [1]$  and  $\Xi_{n-1} = [\bar{\alpha}_{n-1}]$  are  $1 \times 1$  matrices. From these, form the  $n \times n$  block-diagonal matrices

$$\mathcal{L} = \text{diag}(\Xi_0, \Xi_2, \Xi_4, \dots) \quad \text{and} \quad \mathcal{M} = \text{diag}(\Xi_{-1}, \Xi_1, \Xi_3, \dots).$$

The *CMV matrix* associated to the coefficients  $\alpha_0, \dots, \alpha_{n-1}$  is  $\mathcal{C} = \mathcal{L}\mathcal{M}$ .

---

*Date:* August 5, 2005.

<sup>1</sup> Supported in part by NSF grant DMS-0401277 and a Sloan Foundation Fellowship.

Following [29], we will refer to the numbers  $\alpha_k$  as Verblunsky coefficients. A related system of matrices was discovered independently by Tao and Thiele, [31], in connection with the non-linear Fourier transform. Their matrices are bi-infinite and correspond to setting odd-indexed Verblunsky coefficients to zero. (This has the effect of doubling the spectrum; cf. [29, p. 84].)

Expanding out the matrix product  $\mathcal{LM}$  is rather labourious. In Figure 1, we show the result for  $n = 8$ . As can be seen from this example, CMV matrices have a rather rigid structure:

**Definition 1.2.** We say that an  $n \times n$  matrix has *CMV shape* if the entries have the following pattern of horizontal  $2 \times 4$  blocks:

$$\begin{bmatrix} * & * & + & 0 & 0 & 0 & 0 & 0 \\ + & * & * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & * & + & 0 & 0 & 0 \\ 0 & + & * & * & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & * & + & 0 \\ 0 & 0 & 0 & + & * & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & + & * & * \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} * & * & + & 0 & 0 & 0 & 0 \\ + & * & * & 0 & 0 & 0 & 0 \\ 0 & * & * & * & + & 0 & 0 \\ 0 & + & * & * & * & 0 & 0 \\ 0 & 0 & 0 & * & * & * & + \\ 0 & 0 & 0 & + & * & * & * \\ 0 & 0 & 0 & 0 & 0 & * & * \end{bmatrix} \quad (2)$$

where  $+$  represents a positive entry and  $*$  represents a possibly non-zero entry. The top left corner of a CMV matrix always has the  $2 \times 3$  structure depicted above, whereas the bottom right corner will consist of a  $2 \times 3$  block if  $n$  is even (left matrix above) or a  $1 \times 2$  block (right matrix above) if  $n$  is odd.

Naturally, CMV matrices have CMV shape; a proof of this can be found in the original paper of Cantero, Moral, and Velázquez, [6]. Conversely, [7] shows that a unitary matrix with CMV shape must be a CMV matrix. (Alternate proofs of these results can be found at the end of Section 3.) This will be very useful for us since shape and unitarity properties are easier to check than comparing all matrix entries. To facilitate describing the shape of matrices we will use the following terms:

**Definition 1.3.** The *upper staircase* of a matrix consists of the first non-zero entry in every column (or last in every row). Conversely, the *lower staircase* consists of the first non-zero entries in the rows.

The entries marked  $+$  are particularly important so we give them a name too.

**Definition 1.4.** The entries marked  $+$  are precisely  $(2, 1)$  and those of the form  $(2j - 1, 2j + 1)$  and  $(2j + 2, 2j)$  with  $j \geq 1$ . We will refer to these as the *exposed* entries of the CMV matrix.

We will now give a brief overview of the contents of the paper. We will structure this presentation around the order of the sections.

In Section 2, we describe the origin of CMV matrices in the theory of orthogonal polynomials on the unit circle. Nothing said there is new; it is provided for completeness.

In Section 3, we show how a general unitary matrix can be reduced to a (direct sum) of CMV matrices with a simple efficient algorithm. Our model here is the famous Householder implementation of the Lanczos reduction of a general symmetric matrix to tri-diagonal shape by orthogonal conjugation [20, §6.4]. This significantly reduces the storage requirements during eigenvalue computations.

$$\begin{bmatrix} \bar{\alpha}_0 & \rho_0 \bar{\alpha}_1 & \rho_0 \rho_1 & 0 & 0 & 0 & 0 & 0 \\ \rho_0 & -\alpha_0 \bar{\alpha}_1 & -\alpha_0 \rho_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho_1 \bar{\alpha}_2 & -\alpha_1 \bar{\alpha}_2 & \rho_2 \bar{\alpha}_3 & \rho_2 \rho_3 & 0 & 0 & 0 \\ 0 & \rho_1 \rho_2 & -\alpha_1 \rho_2 & -\alpha_2 \bar{\alpha}_3 & -\alpha_2 \rho_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_3 \bar{\alpha}_4 & -\alpha_3 \bar{\alpha}_4 & \rho_4 \bar{\alpha}_5 & \rho_4 \rho_5 & 0 \\ 0 & 0 & 0 & \rho_3 \rho_4 & -\alpha_3 \rho_4 & -\alpha_4 \bar{\alpha}_5 & -\alpha_4 \rho_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_5 \bar{\alpha}_6 & -\alpha_5 \bar{\alpha}_6 & \rho_6 \bar{\alpha}_7 \\ 0 & 0 & 0 & 0 & 0 & \rho_5 \rho_6 & -\alpha_5 \rho_6 & -\alpha_6 \bar{\alpha}_7 \end{bmatrix}$$

 FIGURE 1. An  $8 \times 8$  CMV matrix in terms of the Verblunsky Coefficients

In Section 4, we describe a symplectic structure on the set of CMV matrices of fixed determinant. Note that in the definition of CMV matrices,  $\det(\Xi_k) = -1$  for  $0 \leq k \leq n-2$  and hence

$$\det \mathcal{C} = (-1)^{n-1} \bar{\alpha}_{n-1}.$$

The specific symplectic structure is defined by algebraic means. This work is inspired by corresponding work on Jacobi matrices stimulated by their appearance in the solution of the Toda lattice, [14]. Kostant, [22], noticed that by altering the natural Lie bracket on real  $n \times n$  matrices, the set of Jacobi matrices with fixed trace becomes a co-adjoint orbit. As a consequence, this manifold is a symplectic leaf of the Lie–Poisson bracket. Moreover, he uncovered an algebraic interpretation of the complete integrability of the Toda lattice, which he extended to other finite dimensional Lie algebras.

As noted in Section 4, the natural algebraic setting for CMV matrices is a Lie group rather than a Lie algebra. We show that the set of CMV matrices with fixed determinant forms a symplectic leaf for a certain Poisson structure on  $GL(n; \mathbb{C})$ . We will refer to this as the Gelfand–Dikij bracket, though the name Sklyanin bracket would be equally appropriate.

In a concurrent paper, [23], Luen-Chau Li has independently derived the main results of this section. In a sense, his approach is the reverse of ours: he studies the action of certain dressing transformations, while we arrive at their existence only after studying the problem by other means.

Section 4 also contains a few simple remarks regarding the Ablowitz-Ladik hierarchy. The original (defocusing) Ablowitz Ladik equation [1, 2] is a space-discretization of the cubic nonlinear Schrödinger equation:

$$-i\dot{\beta}_k = \rho_k^2(\beta_{k+1} + \beta_{k-1}) - 2\beta_k, \quad (3)$$

where  $\beta_n$  is a sequence of numbers in the unit disk indexed over  $\mathbb{Z}$  and  $\rho_k = (1 - |\beta_k|^2)^{1/2}$ . If we change variables to  $\alpha_k(t) = e^{2it} \beta_k(t)$ , this system becomes

$$-i\dot{\alpha}_k = \rho_k^2(\alpha_{k+1} + \alpha_{k-1}), \quad (4)$$

which is a little simpler. If we then choose  $\alpha_{-1}$  and  $\alpha_{n-1}$  to lie on the unit circle, then they do not move and we obtain a finite system of ODEs for  $\alpha_k$ ,  $0 \leq k \leq n-2$ , which is the specific case we treat.

These equations form a completely integrable Hamiltonian system; indeed they were introduced with this very property in mind. By considering all possible functions of the commuting Hamiltonians, one is immediately lead to a hierarchy of

equations containing (4) as a special case. Similar considerations lead from the original Toda equation to the full Toda hierarchy.

As presented in Section 4, the matters discussed under the rubric ‘Ablowitz–Ladik hierarchy’ are not obviously linked to the equations just discussed. We make this connection in Section 5.

As a Hamiltonian system, the Ablowitz–Ladik equation comes with a symplectic form. In Section 5, we prove that this agrees with the bracket we introduced from algebraic considerations. Viewed from the opposite perspective, the goal of this section is to write the Gelfand–Dikij bracket tensorially using the Verblunsky coefficients as a system of coordinates.

In Sections 6 and 7, we study further properties of the Ablowitz–Ladik system following the work of Moser, [24], on the Toda lattice. Specifically, we study how the spectral measure naturally associated to a CMV matrix evolves under the Hamiltonians of the Ablowitz–Ladik hierarchy. This information is then used to derive long-time asymptotics and determine the scattering map.

The last topic we treat is the construction of action/angle coordinates for the finite Ablowitz–Ladik system. This is done in Section 8.

*Acknowledgements:* We are delighted to dedicate this paper to Percy Deift on the occasion of his sixtieth birthday. Our understanding of the matters discussed here owes much to his lectures, [12]. We are also grateful to him for encouragement along the way.

## 2. ORTHOGONAL POLYNOMIALS.

As CMV matrices arose in the study of orthogonal polynomials, it is natural that we begin there. We will first describe the relation of orthogonal polynomials to Jacobi matrices and then explain the connection to CMV matrices.

Given a probability measure  $d\nu$  supported on a finite subset of  $\mathbb{R}$ , say of cardinality  $n$ , we can apply the Gram–Schmidt procedure to  $\{1, x, x^2, \dots, x^{n-1}\}$  and so obtain an orthonormal basis for  $L^2(d\nu)$  consisting of polynomials,  $\{p_j(x) : j = 0, \dots, n-1\}$ , with positive leading coefficient. In this basis, the linear transformation  $f(x) \mapsto xf(x)$  is represented by a Jacobi matrix. An equivalent statement is that the orthonormal polynomials obey a three-term recurrence:

$$xp_j(x) = a_j p_{j+1}(x) + b_j p_j(x) + a_{j-1} p_{j-1}(x)$$

where  $a_{-1} = 0$  and  $p_n \equiv 0$ . A third equivalent statement is the following:  $\lambda$  is an eigenvalue of  $J$  if and only if  $\lambda \in \text{supp}(d\nu)$ ; moreover, the corresponding eigenvector is  $[p_0(\lambda), p_1(\lambda), \dots, p_{n-1}(\lambda)]^T$ .

We have just shown how measures on  $\mathbb{R}$  lead to Jacobi matrices; in fact, there is a one-to-one correspondence between them. Given a Jacobi matrix,  $J$ , let  $d\nu$  be the spectral measure associated to  $J$  and the vector  $e_1 = [1, 0, \dots, 0]^T$ . Then  $J$  represents  $x \mapsto xf(x)$  in the basis of orthonormal polynomials associated to  $d\nu$ .

Before explaining the origin of CMV matrices, it is necessary to delve a little into the theory of orthogonal polynomials on the unit circle. For a more complete description of what follows, the reader should turn to [29].

Given a finitely-supported probability measure  $d\mu$  on  $S^1$ , the unit circle in  $\mathbb{C}$ , we can construct an orthonormal system of polynomials,  $\phi_k$ , by applying the Gram–Schmidt procedure to  $\{1, z, \dots\}$ . These obey a recurrence relation; however, to

simplify the formulae, we will present the relation for the monic orthogonal polynomials  $\Phi_k(z)$ :

$$\Phi_{k+1}(z) = z\Phi_k(z) - \bar{\alpha}_k\Phi_k^*(z). \quad (5)$$

Here  $\alpha_k$  are recurrence coefficients, which are called Verblunsky coefficients, and  $\Phi_k^*$  denotes the reversed polynomial:

$$\Phi_k(z) = \sum_{l=0}^k c_l z^l \quad \Rightarrow \quad \Phi_k^*(z) = \sum_{l=0}^k \bar{c}_{k-l} z^l. \quad (6)$$

When  $d\mu$  is supported at exactly  $n$  points,  $\alpha_k \in \mathbb{D}$  for  $0 \leq k \leq n-2$  while  $\alpha_{n-1}$  is a unimodular complex number. (Incidentally, if  $d\mu$  has infinite support, then there are infinitely many Verblunsky coefficients and all lie inside the unit disk.)

To recover the relation between the orthonormal polynomials, one need only apply the following relation, which can be deduced from (5):

$$\|\Phi_k\|_{L^2(d\mu)} = \prod_{l=0}^{k-1} \rho_l, \quad \text{where } \rho_l = \sqrt{1 - |\alpha_l|^2}. \quad (7)$$

The Verblunsky coefficients completely describe the measure  $d\mu$ :

**Theorem 2.1** (Verblunsky). *There is a 1-to-1 correspondence between probability measures on the unit circle supported at  $n$  points and Verblunsky coefficients  $(\alpha_0, \dots, \alpha_{n-1})$  with  $\alpha_k \in \mathbb{D}$  for  $0 \leq k \leq n-2$  and  $\alpha_{n-1} \in S^1$ .*

From the discussion of Jacobi matrices, it would be natural to consider the matrix representation of  $f(z) \mapsto zf(z)$  in  $L^2(d\mu)$  with respect to the basis of orthonormal polynomials. This is *not* a CMV matrix; rather it is what Simon, [29], has dubbed a GGT matrix, from the initials of Geronimus, Gragg, and Teplyaev. Perhaps the most striking difference from a CMV (or Jacobi) matrix is that a GGT matrix is very far from sparse—generically, all entries above and including the sub-diagonal are non-zero. (In a sense, CMV matrices are optimally sparse; see Theorem 3.6.)

Cantero, Moral, and Velazquez had the simple and ingenious idea of applying the Gram–Schmidt procedure to  $\{1, z, z^{-1}, z^2, z^{-2}, \dots\}$  rather than  $\{1, z, \dots\}$ . The resulting functions,  $\chi_k(z)$  ( $0 \leq k \leq n-1$ ), are easily expressed in terms of the orthonormal polynomials:

$$\chi_k(z) = \begin{cases} z^{-k/2} \phi_k^*(z) & : k \text{ even} \\ z^{-(k-1)/2} \phi_k(z) & : k \text{ odd.} \end{cases} \quad (8)$$

In this basis, the map  $f(z) \mapsto zf(z)$  is represented in an especially simple form:

**Theorem 2.2.** *In the orthonormal basis  $\{\chi_k(z)\}$  of  $L^2(d\mu)$ , the operator  $f(z) \mapsto zf(z)$  is represented by the CMV matrix associated to the Verblunsky coefficients of the measure  $d\mu$ .*

The  $\mathcal{C} = \mathcal{LM}$  factorization presented in the introduction originates as follows: Let us write  $x_k$ ,  $0 \leq k \leq n-1$ , for the orthonormal basis constructed by applying the Gram–Schmidt procedure to  $\{1, z^{-1}, z, z^{-2}, z^2, \dots\}$ . Then the matrix elements of  $\mathcal{L}$  and  $\mathcal{M}$  are given by

$$\mathcal{L}_{j+1,k+1} = \langle \chi_j(z) | z x_k(z) \rangle, \quad \mathcal{M}_{j+1,k+1} = \langle x_j(z) | \chi_k(z) \rangle.$$

See [29] for further discussion.

The measure  $d\mu$  can be reconstructed from  $\mathcal{C}$  in a manner analogous to the Jacobi case:

**Theorem 2.3.** *Let  $d\mu$  be the spectral measure associated to a CMV matrix,  $\mathcal{C}$ , and the vector  $e_1$ . Then  $\mathcal{C}$  is the CMV matrix associated to the measure  $d\mu$ .*

Proofs of these two Theorems can be found in [6] or [29]. The second result also follows from Corollary 3.4 below.

### 3. REDUCTION TO CMV SHAPE.

The purpose of this section is to explain a simple analytic (and numeric) algorithm for the reduction of a unitary matrix to CMV form. At the core of the algorithm is the Householder method for reducing the number of nonzero entries of a matrix by conjugating it with reflections. These are chosen in such a way as to allow for successive applications of the reduction.

The Householder algorithm is the method of choice to reduce any real symmetric matrix to tridiagonal form or any non-symmetric matrix to Hessenberg shape. Here we implement the successive reductions differently, by alternating columns and rows (see Figure 2) to reduce any unitary matrix for which  $e_1 = [1, 0, \dots, 0]^T$  is cyclic to CMV shape. With the obvious modification, it reduces other unitary matrices to a direct sum of CMV matrices.

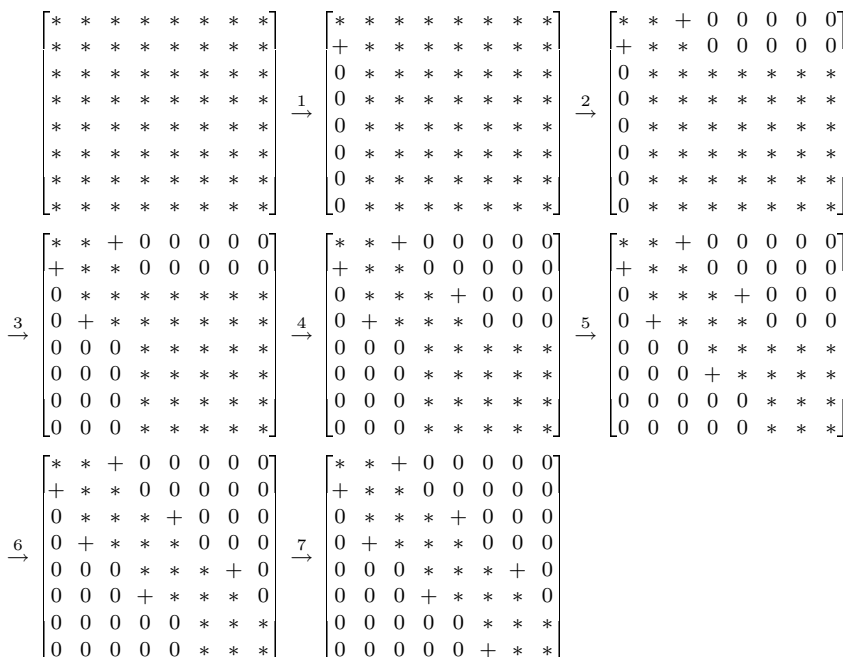


FIGURE 2. The CMV-ification algorithm in action.

Let us briefly describe the construction of Householder reflections; what we present is a very slight modification of what one will find in most textbooks. Given

$u \in \mathbb{C}^n$  and  $0 \leq m < n$ , let

$$v = [0, \dots, 0, \alpha, u_{m+2}, \dots, u_n]^T \quad \text{with} \quad \alpha = u_{m+1} - \frac{u_{m+1}}{|u_{m+1}|} \left( \sum_{j=m+1}^n |u_j|^2 \right)^{1/2}.$$

The reflection through the plane perpendicular to  $v$  is given by

$$R = I - 2 \frac{vv^\dagger}{\|v\|^2}.$$

Naturally,  $R^\dagger = R^{-1} = R$ ; moreover,

$$Ru = u - v = [u_1, \dots, u_m, u_{m+1} - \alpha, 0, \dots, 0]^T$$

and since the first  $m$  entries of  $v$  are identically zero,  $R$  has the block structure

$$R = \begin{bmatrix} I_m & 0 \\ 0 & * \end{bmatrix}, \quad (9)$$

where  $I_m$  is the  $m \times m$  identity matrix, and  $*$  denotes an unspecified  $(n-m) \times (n-m)$  matrix.

When working over  $\mathbb{R}$ , one typically chooses  $\alpha$  slightly differently to guarantee that  $u_{m+1} - \alpha$  is positive (or zero). This is not possible in our case. To remedy the situation, we multiply by a suitably chosen rotation  $D = \text{diag}(1, \dots, 1, e^{i\theta}, 1, \dots, 1)$  where  $e^{i\theta}$  sits in the  $(m+1)$ th position.

Replacing  $R$  by  $DR$  gives us the desired modification of the standard Householder reflection. We summarize its properties as follows:

**Lemma 3.1.** *Given  $u \in \mathbb{C}^n$  and  $0 \leq m < n$  there is a Householder reflection  $R$  that has the block structure (9) and*

$$R : u \mapsto [u_1, \dots, u_m, \sqrt{\sum_{l=m+1}^n |u_l|^2}, 0, \dots, 0]^T.$$

*We will refer to  $R$  as the reflection at level  $m$  for  $u$ .*

**Theorem 3.2** (A Householder algorithm for unitary matrices). *Any  $n \times n$  unitary matrix,  $U$ , for which  $e_1$  is cyclic can be reduced to CMV shape by successive conjugations with Householder reflections. Moreover, this can be done without altering the spectral measure associated to  $e_1$ .*

*Proof.* During the proof, we will refer the reader to Figure 2. The small numbers above each arrow enumerate the steps in the algorithm.

As we will see, at every step  $U$  is conjugated by a matrix that leaves  $e_1$  invariant. Thus the spectral measure associated to  $e_1$  will not change. In particular, this vector will remain cyclic throughout.

For the first step, we choose  $R$  to be the reflector at level two for the first column of  $U$  and then form  $RUR^\dagger$ . Left multiplication by  $R$  creates zeros in the places indicated and these are not destroyed by right multiplication by  $R^\dagger$  because of its block structure. One detail remains: why the entry marked  $+$  is not zero. The answer is simple. This entry would only vanish if the first column of the original matrix were a multiple of  $e_1$ . This would make  $e_1$  an eigenvector for  $U$  and so not cyclic.

Next we describe Step 3 as an archetype for all subsequent odd-numbered steps. Even-numbered steps differ from odd-numbered steps only by interchanging the

roles of rows and columns. We will write  $U$  for the matrix output by step two (not the original one).

Let  $R$  be the reflector at level 3 associated to the second column of  $U$ , then the effect of Step 3 is  $U \mapsto RUR^\dagger$ . For the reasons explained in Step 1, this will produce zeros in the second column as shown in Figures 2 and 3. This state of knowledge is represented by the centre matrix in Figure 3. Two questions remain however: why is the entry newly marked  $+$  not zero, and why do additional zeros appear in the fourth column. We answer them in reverse order.

$$\begin{bmatrix} * & * & + & 0 & 0 & 0 & 0 & 0 \\ + & * & * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \end{bmatrix} \xrightarrow{3} \begin{bmatrix} * & * & + & 0 & 0 & 0 & 0 & 0 \\ + & * & * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & * & * & * & * & * \\ 0 & + & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} * & * & + & 0 & 0 & 0 & 0 & 0 \\ + & * & * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & * & * & * & * & * \\ 0 & + & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \end{bmatrix}$$

FIGURE 3. Matrices in the discussion of Step 3.

The rows of a unitary matrix must be orthogonal; in particular, all rows must be perpendicular to the first. As the third entry in the top row is non-zero (indeed positive), it follows that the third column must have zeros in the places indicated (cf. the third matrix in Figure 3).

If the entry marked  $+$  in the second column were actually zero, then by the reasoning of the previous paragraph, its right-hand neighbour would also vanish. This contradicts the cyclicity of  $e_1$ ; specifically, the linear span of  $\{e_1, e_2, e_3\}$  would be an invariant subspace.

With obvious modifications, these arguments apply to any odd step ( $\geq 3$ ). As noted even steps are analogous but with rows and columns reversed; indeed one can just apply the discussion above to  $U^\dagger$ . (As  $U$  is unitary,  $e_1$  is also cyclic for  $U^\dagger$ .)

Note that in the last step, no additional zeros are produced. This does not result in any modifications to the argument, we merely wish to point out that in this case,  $R$  is a diagonal matrix.  $\square$

Theorem 3.2 tells us how to reduce any unitary matrix having  $e_1$  as a cyclic vector to CMV shape. Now we wish to show that any unitary matrix  $U$  in CMV shape is actually a CMV matrix, i.e., there exists a set of coefficients  $\alpha_0, \dots, \alpha_{n-2} \in \mathbb{D}$  and  $\alpha_{n-1} \in S^1$  so that  $U$  is the matrix given by (1). This result is not new, and can be found in [7]. We prove this result differently, and obtain it as a corollary of the following

**Proposition 3.3.** *Let  $B$  and  $C$  be two unitary matrices in CMV shape that have the same spectral measure with respect to the vector  $e_1$ . Then  $B = C$ .*

*Proof.* Since  $B$  and  $C$  have the same spectral measure with respect to  $e_1$ , there exists a unitary matrix  $V$  so that  $Ve_1 = e_1$  and

$$C = V^\dagger B V. \tag{10}$$

We inductively prove that  $Ve_k = e_k$  for all  $1 \leq k \leq n$ , which shows that  $V$  is the identity and so  $B = C$ .



Assume that  $Ve_k = e_k$  holds for all  $k \leq m$ . (From the definition of  $V$ , it is true when  $m = 1$ .) This implies that the top left  $m \times m$  blocks of  $B$  and  $C$  coincide:

$$\langle e_j | Ce_k \rangle = \langle Ve_j | B(Ve_k) \rangle = \langle e_j | Be_k \rangle$$

for all  $1 \leq j, k \leq m$ . We now proceed differently depending on  $m$ .

If  $m \geq 3$  is odd, consider the  $(m-1)^{\text{th}}$  columns of  $C$  and  $B$ . As both matrices are of CMV shape and agree on the principle  $m \times m$  minor, the only place that these columns may differ is in the  $(m+1)^{\text{th}}$  position. In fact, they must agree here too because CMV shape requires that this entry be positive and then it is uniquely determined by the fact that the columns of a unitary matrix have norm one.

From  $VC = BV$ ,  $Ve_{m-1} = e_{m-1}$ , and  $Ce_{m-1} = Be_{m-1}$ , which we just proved, we have

$$(V-1)Ce_{m-1} = (BV-C)e_{m-1} = 0. \quad (11)$$

But now by the shape of the  $(m-1)^{\text{th}}$  column of  $C$  described above, we know that

$$Ce_{m-1} = C_{m+1,m-1}e_{m+1} + \sum_{k=m-2}^m C_{k,m-1}e_k$$

with  $C_{m+1,m-1} > 0$ . Substituting this into (11) gives  $(1-V)e_{m+1} = 0$ , which concludes the inductive step in this case.

If  $m = 1$ , one should repeat the above analysis using the first column.

If  $m$  is even, then one must repeat the above argument using rows instead of columns. Alternatively, one may proceed as above after first taking the (conjugate) transpose in (10).  $\square$

**Corollary 3.4.** *Let  $U$  be an  $n \times n$  unitary matrix in CMV shape. Then there is a unique  $(n-1)$ -tuple of Verblunsky coefficients  $\alpha_0, \dots, \alpha_{n-2} \in \mathbb{D}$  and  $\alpha_{n-1} \in S^1$  so that  $U$  is the CMV matrix associated to these coefficients.*

*Proof.* For a unitary matrix in CMV shape, let  $\mu$  be the spectral measure associated to  $e_1$ , and  $\alpha_0, \dots, \alpha_{n-1}$  the corresponding Verblunsky coefficients. The CMV matrix  $\mathcal{C}$  given by Definition 1.1 has  $\mu$  as its spectral measure. So, by Proposition 3.3, the original matrix must equal  $\mathcal{C}$ , which is indeed a CMV matrix. Uniqueness of the coefficients is a direct consequence of Proposition 3.3 and Verblunsky's Theorem 2.1.  $\square$

**Remark 3.5.** There is another way of proving the corollary, which involves recursively identifying the  $\alpha$ 's in a unitary matrix of CMV shape. This is essentially the proof of Theorem 3.8 from [7].

Further note that one can read the Verblunsky coefficients from a CMV matrix. In other words, for any unitary matrix  $C$  in CMV shape there exists a *unique* set of Verblunsky coefficients so that  $C$  is the CMV matrix with these coefficients.

We close this section by presenting a result of Cantero, Moral, and Velazquez (see Theorem 3.9 of [7]) which implies that, in terms of sparseness, CMV matrices are minimal among unitary matrices having cyclic vectors.

We call a unitary matrix,  $U$ ,  $(p, q)$ -diagonal if it has only  $p$  nonzero subdiagonals and  $q$  nonzero superdiagonals. In other words, we require that  $U_{j,k} = 0$  for  $j - k \geq p + 1$  and  $k - j \geq q + 1$ . Note that CMV matrices are  $(2, 2)$ -diagonal. With this definition, the statement is the following:

**Theorem 3.6.** *A unitary  $(p, 1)$ -diagonal matrix is a sum of diagonal blocks of order no greater than  $p + 1$ . The same is true of any unitary  $(1, p)$ -diagonal matrix.*

In particular, note that any 4-diagonal unitary matrix is a direct sum of  $3 \times 3$  blocks. As breaking a matrix into a direct sum is tantamount to factoring the characteristic polynomial, no exact algorithm can perform this task for generic input. Similarly, it is impossible to reduce a unitary matrix to four diagonals whilst preserving the spectral measure associated to  $e_1$ .

#### 4. LIE STRUCTURE.

The manifold of Jacobi matrices with fixed trace forms a co-adjoint orbit associated to a particular Lie algebra structure on the  $n \times n$  matrices (although not the one induced by  $GL(n; \mathbb{R})$ ). This gives the manifold a symplectic structure. These matters are described in detail in [12, 27, 28], for example.

In this section, we will show that the manifold of CMV matrices with fixed determinant is a symplectic leaf for a natural Poisson structure on  $GL(n, \mathbb{C})$  and (what is basically equivalent) it is exactly the orbit of a natural group action on this space. and so also have a natural symplectic structure. For pedagogical reasons, we will present the symplectic structure first and then describe how it relates to a group action. (As described in the Introduction, [23], takes the opposite approach to arrive at the same conclusions.)

In deference to those new to the subject, we will endeavour to use the notation and terminology of [27]. We will also mostly refer to this book for proofs that we omit; we hope that the authors of the original articles will forgive this indirect reference to their work.

Self-adjoint matrices form a vector space, hence it is natural that the orbit and symplectic structures on Jacobi matrices arises from the study of a Lie algebra (actually, its vector-space dual). Unitary matrices form a group and hence the natural backdrop for CMV is a Lie group. Specifically, the construction we give below is that of a Sklyanin bracket on  $GL(n, \mathbb{C})$ ; see [27, §2.12]. However, we choose to give a presentation in which the algebra of matrices takes centre stage; an analogous construction for KdV using the algebra of pseudo-differential operators was given by Gelfand and Dickij [17]. This approach is described in Section 2.12.6 of [27]. (Note that here we are referring to the second symplectic structure associated with KdV, which was originally proposed by Adler [4, §4].)

Let  $\mathfrak{g}$  denote the (associative) algebra of  $n \times n$  complex matrices. The algebra structure gives rise to a natural Lie algebra structure:

$$[B, C] = BC - CB.$$

Of course, this also results from viewing  $\mathfrak{g}$  as the Lie algebra of  $GL(n; \mathbb{C})$ .

As a vector space,  $\mathfrak{g} = \mathfrak{l} \oplus \mathfrak{a}$ , where

$$\mathfrak{a} = \{A : A = -A^\dagger\},$$

which is the Lie algebra of the group  $\mathbb{U}(n)$  of  $n \times n$  unitary matrices, and

$$\mathfrak{l} = \{A \in \mathfrak{g} : L_{i,j} = 0 \text{ for } i > j \text{ and } L_{i,i} \in \mathbb{R}\}$$

which is the Lie algebra of the group  $\mathbb{L}(n)$  of  $n \times n$  lower triangular matrices with positive diagonal entries. We will write  $\pi_{\mathfrak{a}}$  and  $\pi_{\mathfrak{l}}$  for the projections into these summands.

This vector-space splitting of  $\mathfrak{g}$  permits us to give it a second Lie algebra structure. First we define  $R : \mathfrak{g} \rightarrow \mathfrak{g}$  by either

$$\begin{aligned} R(X) &= \pi_{\mathfrak{l}}(X) - \pi_{\mathfrak{a}}(X), & \text{for all } X \in \mathfrak{g}, \text{ or} \\ R(L + A) &= L - A & \text{for all } A \in \mathfrak{a} \text{ and } L \in \mathfrak{l}. \end{aligned} \quad (12)$$

The second Lie bracket can then be written as either

$$\begin{aligned} [X, Y]_R &= \frac{1}{2}[R(X), Y] + \frac{1}{2}[X, R(Y)] \quad \forall X, Y \in \mathfrak{g}, \text{ or} \\ [L + A, L' + A']_R &= [L, L'] - [A, A'] \quad \forall L, L' \in \mathfrak{l}, \text{ and } A, A' \in \mathfrak{a}. \end{aligned} \quad (13)$$

The second definition also makes it transparent that the  $R$ -bracket obeys the Jacobi identity.

Whether one treats Jacobi or CMV matrices, the (vector space) dual  $\mathfrak{g}^*$  of  $\mathfrak{g}$  plays an important role. We can identify it with  $\mathfrak{g}$  using the pairing

$$\langle X, Y \rangle = \text{Im tr}(XY). \quad (14)$$

This is not an inner product; however, it is non-degenerate—just choose  $Y = iX^\dagger$ . It is also symmetric and ad-invariant:

$$\langle X, [Z, Y] \rangle = \text{Im tr}(XZY - XYZ) = \langle [X, Z], Y \rangle. \quad (15)$$

This is equivalent to Ad-invariance:

$$\langle BXB^{-1}, BYB^{-1} \rangle = \langle X, Y \rangle, \quad \text{for any } B \in GL(n, \mathbb{C}). \quad (16)$$

Concomitant with the pairing is an identification of 1-forms and vector fields on  $\mathfrak{g}$ :

**Definition 4.1.** Given  $\phi : \mathfrak{g} \rightarrow \mathbb{R}$  and  $B \in \mathfrak{g}$ , define  $\nabla\phi : \mathfrak{g} \rightarrow \mathfrak{g}$  by

$$\left. \frac{d}{dt} \right|_{t=0} \phi(B + tC) = \langle \nabla\phi|_B, C \rangle. \quad (17)$$

Equivalently, if we write  $b_{k,l} = u_{k,l} + iv_{k,l}$  for the matrix entries of  $B$ , then

$$[\nabla\phi]_{k,l} = \frac{\partial\phi}{\partial v_{l,k}} + i \frac{\partial\phi}{\partial u_{l,k}} \quad (18)$$

(notice the reversal of the order of the indices).

We can now define the desired Poisson bracket on  $\mathfrak{g}$ . In fact, by first checking some elementary properties of the objects set out above, one can show that it obeys the Jacobi identity.

**Proposition 4.2.** Let  $R : \mathfrak{g} \rightarrow \mathfrak{g}$  and  $\langle \cdot, \cdot \rangle$  be as above.

- (i)  $\mathfrak{a}^\perp = \mathfrak{a}$  and  $\mathfrak{l}^\perp = \mathfrak{l}$ .
- (ii)  $R$  is antisymmetric:  $\langle X, R(Y) \rangle = -\langle R(X), Y \rangle$ .
- (iii)  $R$  obeys the modified classical Yang-Baxter equation:

$$[R(X), R(Y)] - R([R(X), Y] + [X, R(Y)]) = -[X, Y]$$

- (iv) Given  $\phi, \psi : \mathfrak{g} \rightarrow \mathbb{R}$ , let  $X = \nabla\phi$  and  $Y = \nabla\psi$ . Then

$$\{\phi, \psi\}|_B = \frac{1}{2}\langle R(XB), YB \rangle - \frac{1}{2}\langle R(BX), BY \rangle \quad (19)$$

$$= \frac{1}{2}\langle R(XB), YB \rangle + \frac{1}{2}\langle BX, R(BY) \rangle. \quad (20)$$

defines a Poisson structure on  $\mathfrak{g}$ .

*Proof.* (i) For  $A, B \in \mathfrak{a}$ ,

$$\operatorname{Im} \operatorname{tr}(AB) = \frac{1}{2i} \operatorname{tr}(AB - B^\dagger A^\dagger) = \frac{1}{2i} \operatorname{tr}(AB - BA) = 0.$$

For  $L, L' \in \mathfrak{l}$ , the diagonal entries of  $LL'$  are products of the corresponding entries in  $L$  and  $L'$ . In particular, they are real; therefore,  $\operatorname{Im} \operatorname{tr}(LL') = 0$ .

We have just shown that  $\mathfrak{a}^\perp \subseteq \mathfrak{a}$  and  $\mathfrak{l}^\perp \subseteq \mathfrak{l}$ . As the pairing is non-degenerate, equality follows by dimension counting:  $\dim(\mathfrak{a}) = \dim(\mathfrak{l}) = n^2 = \frac{1}{2} \dim(\mathfrak{g})$ .

(ii) Antisymmetry follows from part (i) by simple computation:

$$\langle R(L+A), L'+A' \rangle = \langle L, L' \rangle - \langle A, A' \rangle = \langle L+A, R(L'+A') \rangle.$$

(iii) This can be checked by direct computation. However, this belies the important role of the mCYB equation in the theory of double Lie algebras and thence in the theory of integrable systems. For more information, see [27, §2.2] or [12, 28].

(iv) This is an example of a much more general statement, namely, these formulae define a Poisson bracket on  $\mathfrak{g}$  whenever  $\langle \cdot, \cdot \rangle$  is symmetric and ad-invariant and  $R$  has properties (ii) and (iii). The only axiom of a Poisson bracket that is not immediate from the definition is the Jacobi identity. Unfortunately, checking this consumes more space than we can justify here. Details can be found in [17, §4], but readers should not be discouraged from simply doing it themselves. We hope that the following two hints make this more palatable:

$$\nabla\{\phi, \psi\}|_B = \frac{1}{2}R(XB)Y - \frac{1}{2}R(YB)X + \frac{1}{2}XR(BY) - \frac{1}{2}YR(BX) \quad (21)$$

and given any trio of matrices  $A, B$ , and  $C$ ,

$$\begin{aligned} & \sum \langle R(B)A - R(A)B, R(C) \rangle \\ &= \sum \frac{2}{3} \langle R(R(B)A - R(A)B), C \rangle + \frac{1}{3} \langle R(B)R(A) - R(A)R(B), C \rangle \\ &= \frac{1}{3} \sum \langle [A, B], C \rangle \end{aligned}$$

where all sums are over cyclic permutations of  $(A, B, C)$ .  $\square$

The Poisson bracket given in the proposition is often referred to as the ‘quadratic bracket’ because the point  $B \in \mathfrak{g}$  where it is evaluated appears quadratically. The Lie-Poisson (or Kirillov) bracket associated to the Lie algebra  $(\mathfrak{g}, [\cdot, \cdot]_R)$  is linear in  $B$ :

$$\{\phi, \psi\}_{LP}|_B = \frac{1}{2} \langle [X, Y]_R, B \rangle.$$

This bracket is known to be compatible (in the sense of Magri-Lenard) with the bracket above; however, it is not pertinent to the study of CMV matrices. It is relevant to Jacobi matrices: under the embedding  $J \mapsto iJ$ , the manifolds of Jacobi matrices with fixed trace are symplectic leaves. This is just the usual construction in  $Gl(n, \mathbb{R})$  in disguise.

**Lemma 4.3.** *The Hamiltonian vector field on  $\mathfrak{g}$  associated to  $\phi : \mathfrak{g} \rightarrow \mathbb{R}$  is*

$$\dot{B} = \frac{1}{2} [BR(XB) - R(BX)B] \quad (22)$$

where  $X = \nabla\phi$ . Equivalently,

$$\dot{B} = B\pi_{\mathfrak{l}}(XB) - \pi_{\mathfrak{l}}(BX)B = \pi_{\mathfrak{a}}(BX)B - B\pi_{\mathfrak{a}}(XB). \quad (23)$$

In particular, if  $B$  is unitary, then

$$\dot{B} = -B\pi_{\mathfrak{a}}(B^{-1}LB) \quad (24)$$

where  $L = \pi_{\mathfrak{l}}(BX)$ .

*Proof.* In terms of the Poisson bracket, the defining property of the Hamiltonian vector field is  $\{\phi, \psi\} = \dot{\psi} = \langle \dot{B}, \nabla\psi \rangle$ . Equation (22) now follows by cycling the trace: if  $Y = \nabla\psi$ , then

$$\{\phi, \psi\} = \frac{1}{2}\langle R(XB), YB \rangle - \frac{1}{2}\langle R(BX), BY \rangle \quad (25)$$

$$= \frac{1}{2}\langle BR(XB) - R(BX)B, Y \rangle. \quad (26)$$

From the definition,  $R = \text{Id} - 2\pi_{\mathfrak{a}} = 2\pi_{\mathfrak{l}} - \text{Id}$ . Equation (23) follows by substituting these relations into (22).

To obtain the special case, let us write  $BX = L + A$ . As  $A$  and  $B^{-1}AB$  are anti-Hermitian,  $\pi_{\mathfrak{a}}(A)B = B\pi_{\mathfrak{a}}(B^{-1}AB)$ . This implies

$$\dot{B} = \pi_{\mathfrak{a}}(L + A)B - B\pi_{\mathfrak{a}}(B^{-1}(L + A)B) = -B\pi_{\mathfrak{a}}(B^{-1}LB), \quad (27)$$

which is exactly (24).  $\square$

**Proposition 4.4.** *The symplectic leaf passing through a particular CMV matrix contains only CMV matrices and all have the same determinant.*

*Proof.* Proving that the determinant is a Casimir is easy, so let us start there. Given  $\theta$ , let  $\phi_{\theta}(B) = \text{Im}[e^{i\theta} \log \det(B)]$  on an open neighbourhood of  $\mathcal{C}$ . Then  $\nabla\phi_{\theta} = e^{i\theta}B^{-1}$  and so  $R(B\nabla\phi_{\theta}) = R(\nabla\phi_{\theta}B) = e^{-i\theta} \text{Id}$ . Therefore,

$$\{\phi_{\theta}, \psi\} = \frac{1}{2}\langle e^{-i\theta} \text{Id}, \nabla\psi B \rangle - \frac{1}{2}\langle e^{-i\theta} \text{Id}, B\nabla\psi \rangle = \frac{1}{2} \text{Im tr}([\nabla\psi, B]) = 0 \quad (28)$$

for any function  $\psi$  and any angle  $\theta$ .

Let  $\mathcal{C}$  be a unitary matrix and  $\phi : \mathfrak{g} \rightarrow \mathbb{R}$ . From (24) in Lemma 4.3, we see that there is an anti-hermitian matrix  $A$  so that under the Hamiltonian flow generated by  $\phi$ ,  $\dot{\mathcal{C}} = -\mathcal{C}A$ . Thus  $\mathcal{C}$  remains unitary.

Now let us restrict our attention to the case of  $\mathcal{C}$  a CMV matrix. We will show that under the  $\phi$ -flow,  $\mathcal{C}$  remains in CMV shape, which implies that  $\mathcal{C}$  remains a CMV matrix by Corollary 3.4.

By (23) from Lemma 4.3, there are  $L, L' \in \mathfrak{l}$  so that

$$\dot{\mathcal{C}} = L'\mathcal{C} - \mathcal{C}L. \quad (29)$$

Elementary calculations show that (left or right) multiplication by a lower triangular matrix does not change entries above the upper staircase; they remain zero. Moreover, exposed entries are simply multiplied by the corresponding diagonal entry in the lower triangular matrix. The first fact shows that  $\dot{\mathcal{C}}$  vanishes above the upper staircase of  $\mathcal{C}$ . The second shows that the exposed entries in the upper staircase obey an equation of the form  $\dot{\mathcal{C}}_{i,j} = \gamma(t)\mathcal{C}_{i,j}$  for some real-valued function  $\gamma$ ; therefore, they remain positive.

This reasoning can be transferred to the lower staircase by noting that

$$\partial_t \mathcal{C}^{\dagger} = -\mathcal{C}^{\dagger} \dot{\mathcal{C}} \mathcal{C}^{\dagger} = \mathcal{C}^{\dagger} L' - L \mathcal{C}^{\dagger}. \quad (30)$$

This completes the proof that  $\mathcal{C}$  remains in CMV shape and hence, of the proposition.  $\square$

It remains for us to show that the symplectic leaf actually fills out the set of CMV matrices with fixed determinant. As this manifold has dimension  $2(n-1)$ , one solution to this problem would be to find  $n-1$  functions that Poisson commute and have linearly independent Hamiltonian vector fields. While appearing roundabout, this approach is actually rather efficient for us because there is just such a family of Hamiltonians that we wish to study anyway, namely, those of the form

$B \mapsto \operatorname{Im} \operatorname{tr}\{f(B)\}$ , for some polynomial  $f$ . Our interest stems from their relevance to the Ablowitz-Ladik hierarchy. We begin with the simplest abstract properties of these Hamiltonians.

**Proposition 4.5.** *Given a polynomial  $f$ , let  $\phi : \mathfrak{g} \rightarrow \mathbb{R}$  by*

$$\phi(B) = \operatorname{Im} \operatorname{tr}\{f(B)\}. \quad (31)$$

- (i)  $\nabla\phi|_B = f'(B)$ .
- (ii) *Functions of the type (31) Poisson commute.*
- (iii) *Under the flow generated by  $\phi$ ,*

$$\dot{B} = [B, \tfrac{1}{2}R(Bf'(B))] = [B, \pi_1(Bf'(B))] = -[B, \pi_a(Bf'(B))]. \quad (32)$$

- (iv) *There is a unique factorization*

$$\exp\{tBf'(B)\} = L(t)Q^{-1}(t) \quad (33)$$

with  $L(t) \in \mathbb{L}(n)$  and  $Q(t) \in \mathbb{U}(n)$ .

- (v) *The integral curve  $B(t)$  with  $B(0) = B$  is*

$$B(t) = L^{-1}(t)BL(t) = Q^{-1}(t)BQ(t). \quad (34)$$

*Proof.* Part (i) follows for monomials by cycling the trace. This extends to polynomials by linearity.

Part (ii) is readily deduced from part (i): if  $\psi(B) = \operatorname{Im} \operatorname{tr}\{g(B)\}$ , then  $[B, \nabla\phi] = [B, \nabla\psi] = 0$  and so  $\{\phi, \psi\} = 0$  follows immediately from the definition, (19).

The first inequality in (32) follows immediately from (22) and the fact that  $[B, \nabla\phi] = 0$ . The second and third equalities follow from (23).

After taking the adjoint of both sides, (33) becomes the QR factorization.

We will prove part (v) by direct computation. As (34) is certainly true when  $t = 0$ , it suffices to check that all three formulae obey the same differential equation. We will just prove this for  $B_1(t) = L^{-1}(t)BL(t)$  as  $Q^{-1}(t)BQ(t)$  can be treated identically. From the definition,

$$\dot{B}_1(t) = L^{-1}(t)B\dot{L}(t) - L^{-1}(t)\dot{L}(t)L^{-1}(t)BL(t) = [B_1(t), L^{-1}(t)\dot{L}(t)]$$

so by (32), it suffices to show that for every  $t$ ,  $\pi_1(B_1f'(B_1)) = L^{-1}\dot{L}$ . This can be demonstrated by differentiating (33):

$$Bf'(B) \exp\{tBf'(B)\} = \partial_t \exp\{tBf'(B)\} = \dot{L}Q^{-1} - LQ^{-1}\dot{Q}Q^{-1}, \quad (35)$$

which implies  $Bf'(B)LQ^{-1} = \dot{L}Q^{-1} - LQ^{-1}\dot{Q}Q^{-1}$  and so

$$B_1f'(B_1) = L^{-1}Bf'(B)L = L^{-1}\dot{L}(t) - Q^{-1}\dot{Q}. \quad (36)$$

This shows not only that  $\pi_1(B_1f'(B_1)) = L^{-1}\dot{L}$ , but also  $\pi_a(B_1f'(B_1)) = -Q^{-1}\dot{Q}$ .  $\square$

Part (iii) of this proposition gives Lax pair representations for the flows generated by Hamiltonians of this form. From the abstract theory, this is to be expected:  $B \mapsto \operatorname{Im} \operatorname{tr}\{f(B)\}$  are central functions on  $\mathfrak{g}$ —that is, they are constant on conjugacy classes—cf. [27, §2.12.4]. It is not difficult to see that these Lax pairs are precisely those discovered in [26].

Of course, the key to reconciling the concrete work of Nenciu with the abstract approach described here is precisely the discovery that CMV matrices have a Lie theoretic interpretation.

Similarly, (34) can be deduced from the abstract theory. Because of the sparsity of the CMV matrix and the ready availability of good implementations of the QR algorithm, this provides an easy method for computational studies of the Ablowitz-Ladik hierarchy. Those familiar with the work of Deift, Li, Nanda, and Tomei [8, 11] might immediately ask if we can say anything new about diagonalization algorithms for unitary matrices; this will be discussed at the end of Section 7.

We will continue our investigation of these Hamiltonians in Section 6. By borrowing a result from there, we can prove the main result of this section:

**Theorem 4.6.** *The manifold of CMV matrices with fixed determinant form a symplectic leaf in the Poisson manifold  $\mathfrak{g}$ .*

*Proof.* Let us write  $z_1, \dots, z_n$  for the eigenvalues of the CMV matrix  $\mathcal{C}$ . By Lagrange interpolation, we can find polynomials  $F_k$  so that  $F_k(0) = 0$  and  $F_k(z_j) = \delta_{jk}$ . It is then elementary to construct polynomials  $f_k$  with  $F_k(z) = 2zf'_k(z)$ .

Now consider the following functions on  $\mathfrak{g}$ :

$$\phi_k(B) = \operatorname{Im} \operatorname{tr}(f_k(B)). \quad (37)$$

By Corollary 6.4,  $\phi_1, \dots, \phi_{n-1}$  give rise to linearly independent Hamiltonian vector fields. Moreover, by Proposition 4.5, these Hamiltonians Poisson commute. Therefore, the symplectic leaf passing through  $\mathcal{C}$  must be of dimension no less than  $2n - 2$ .

From Proposition 4.4, we can deduce that the dimension must be exactly  $2n - 2$ . As the space of CMV matrices with fixed determinant is path connected (it is homeomorphic to  $\mathbb{D}^{n-1}$ ), a simple chain-of-balls argument shows that the symplectic leaf must exhaust this manifold.  $\square$

In the remainder of this section, we will describe dressing transformations and determine how they act on CMV matrices.

We need to consider three groups  $G = GL(n, \mathbb{C})$ ,  $D = G \times G$ , and  $G_R = \mathbb{L}(n) \times \mathbb{U}(n)$ . By  $\mathbb{U}(n)$  we mean the group of  $n \times n$  unitary matrices with the order of multiplication reversed:  $\omega_1 \star \omega_2 = \omega_2 \omega_1$ . (We will use lower-case Greek letters for group elements.)

The significance of the group  $G_R$  is that the corresponding Lie algebra is  $\mathfrak{g}$  with the bracket given by (13). It is also worth noting that the map  $(\lambda, \omega) \mapsto \lambda\omega^{-1}$  defines a diffeomorphism from  $G_R$  to  $G$ . The inverse mapping is given by  $\beta \mapsto (\beta_+, \beta_-)$  where  $\beta_+ \in \mathbb{L}(n)$  and  $\beta_- \in \mathbb{U}(n)$  are the solution of the factorization problem  $\beta = \beta_+ \beta_-^{-1}$ . This problem is always uniquely soluble; indeed, it is essentially the QR factorization of  $\beta^\dagger$ .

Both  $G$  and  $G_R$  can be regarded as subgroups (and submanifolds) of  $D$  via the following embeddings:

$$\begin{aligned} i : G &\hookrightarrow D & \text{by } \beta &\mapsto (\beta, \beta), & \text{and} \\ i' : G_R &\hookrightarrow D & \text{by } (\lambda, \omega) &\mapsto (\lambda, \omega^{-1}). \end{aligned}$$

Moreover, the product map

$$i \cdot i' : G \times G_R \rightarrow D \quad \text{by } (\beta, \lambda, \omega) \mapsto (\beta\lambda, \beta\omega^{-1})$$

defines a diffeomorphism (as does  $i' \cdot i$ ). The inverse mapping is easily seen to be  $\pi \oplus \pi'$  where

$$\begin{aligned} \pi' : D &\rightarrow G_R \quad \text{by} \quad (\xi, \eta) \mapsto ([(\xi^{-1}\eta)_+]^{-1}, (\xi^{-1}\eta)_-), \quad \text{and} \\ \pi : D &\rightarrow G \quad \text{by} \quad (\xi, \eta) \mapsto \xi(\xi^{-1}\eta)_+ = \eta(\xi^{-1}\eta)_-. \end{aligned}$$

These projections also permit us to identify the coset spaces  $D/G_R$  and  $G \setminus D$  with  $G$  and  $G_R$  respectively.

The group  $D$  acts by right multiplication on  $D/G_R$  and hence, by the identification just noted, on  $G$ . As the product map  $i \cdot i'$  is onto, we can get a complete understanding of this group action by studying its restriction to the subgroups  $i(G)$  and  $i'(G_R)$ .

The first subgroup,  $i(G)$ , leads to nothing new, just the action of  $G$  on itself by left multiplication; however  $i'(G_R)$  is different, it leads to dressing transformations:  $(\lambda, \omega) \in G_R$  maps  $G$  to itself via

$$\beta \mapsto \pi(\lambda\beta, \omega^{-1}\beta) = \lambda\beta(\beta^{-1}\lambda^{-1}\omega^{-1}\beta)_+ = \omega^{-1}\beta(\beta^{-1}\lambda^{-1}\omega^{-1}\beta)_-. \quad (38)$$

**Theorem 4.7.** *The orbit of a CMV matrix under the dressing transformations (38) is precisely the set of CMV matrices with the same determinant.*

*Proof.* As we have already proved Theorem 4.6, this follows from the general theory of Poisson Lie groups (cf. [27, §2.12]). However, as it is possible to give a quick and concrete proof that the orbit coincides with the symplectic leaf, we do so.

When  $\beta$  is unitary, the formula for the action can be simplified considerably:  $(\beta^{-1}\lambda^{-1}\omega^{-1}\beta)_- = \beta^{-1}\omega(\beta^{-1}\lambda^{-1})_-$  and so

$$(\lambda, \omega) : \beta \mapsto (\beta^{-1}\lambda^{-1})_-. \quad (39)$$

In particular,  $\omega$  plays no role. This formula also shows immediately that the orbit of a unitary matrix contains only unitary matrices.

As the group  $G_R$  is connected, we can prove the theorem by showing that the tangent vectors to the orbit coincide with those of the symplectic leaf.

Writing (39) with  $\lambda = e^{-tL}$  for some  $L \in \mathfrak{l}$  and using the fact that  $\beta$  is unitary shows

$$\beta \mapsto (\beta^{-1}e^{tL})_- = \beta(\beta^{-1}e^{tL}\beta)_- = \beta - t\beta\pi_{\mathfrak{a}}(\beta^{-1}L\beta) + O(t^2). \quad (40)$$

Thus, the tangent space to the orbit through  $\beta$  is given by  $-\beta\pi_{\mathfrak{a}}(\beta^{-1}L\beta)$  as  $L$  varies over  $\mathfrak{l}$ . This is exactly the tangent space to the symplectic leaf as given in (24) of Lemma 4.3.  $\square$

## 5. THE ABLOWITZ-LADIK BRACKET

As discussed earlier, the  $n$ -tuple  $(\alpha_0, \dots, \alpha_{n-2}, \alpha_{n-1}) \in \mathbb{D}^{n-1} \times S^1$  of Verblunsky coefficients gives a system of coordinates for the manifold of CMV matrices.

As shown in Section 4, the determinant,  $\det(\mathcal{C}) = (-1)^{n-1}\bar{\alpha}_{n-1}$  is a Casimir and so  $(\alpha_0, \dots, \alpha_{n-2})$  can be used as coordinates for the symplectic manifold of CMV matrices with fixed determinant. It is therefore natural to write the Gelfand–Dikij bracket in these coordinates. By doing so, we will recover a bracket introduced



earlier for the study of the Ablowitz–Ladik system:

$$\{f, g\}_2 = \sum_{j=0}^{n-2} \rho_j^2 \left[ \frac{\partial f}{\partial u_j} \frac{\partial g}{\partial v_j} - \frac{\partial f}{\partial v_j} \frac{\partial g}{\partial u_j} \right] \quad (41)$$

$$= 2i \sum_{j=0}^{n-2} \rho_j^2 \left[ \frac{\partial f}{\partial \bar{\alpha}_j} \frac{\partial g}{\partial \alpha_j} - \frac{\partial f}{\partial \alpha_j} \frac{\partial g}{\partial \bar{\alpha}_j} \right] \quad (42)$$

where  $\alpha_j = u_j + iv_j$  for all  $0 \leq j \leq n-2$ , and, as usual,

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

For clarity, we will call this second bracket the Ablowitz-Ladik bracket. Note that this differs by a factor of two from that used in [26]; we will adjust results quoted from this paper accordingly.

On the set of Verblunsky coefficients  $(\alpha_0, \dots, \alpha_{n-2}) \in \mathbb{D}^{n-1}$  (with fixed  $\alpha_{n-1} \in S^1$ ) we consider the Hamiltonians  $\text{Re}(K_m)$  and  $\text{Im}(K_m)$ , where

$$K_m = \frac{1}{m} \text{tr}(\mathcal{C}^m)$$

for  $m \geq 1$ . The evolutions under the flows generated by the real and imaginary parts of  $K_m$  in the Ablowitz-Ladik bracket were found by Nenciu [26]:

$$\dot{\mathcal{C}} = \{\text{Re}(K_m), \mathcal{C}\}_2 = -[\mathcal{C}, \pi_{\mathbf{a}}(i\mathcal{C}^m)], \quad (43)$$

$$\dot{\mathcal{C}} = \{\text{Im}(K_m), \mathcal{C}\}_2 = -[\mathcal{C}, \pi_{\mathbf{a}}(\mathcal{C}^m)]. \quad (44)$$

To obtain the formulae given above from those of [26], one should account for the difference in notation. In [26], a subscript  $+$  is used to indicate

$$(B_+)_{ij} = \begin{cases} B_{ij} & : i < j \\ \frac{1}{2} B_{ii} & : i = j \\ 0 & : i > j \end{cases}$$

from which it follows that  $B_+ - (B_+)^{\dagger} = \pi_{\mathbf{a}}(B)$ ; indeed both sides are antisymmetric and agree on the upper triangle.

These formulae also hold true if we replace the Ablowitz-Ladik bracket by the Gelfand-Dikij bracket (19), as one can immediately see from the second identity in (23). This suggests that the Ablowitz-Ladik bracket might be the one we seek.

We will indeed prove that the Gelfand–Dikij bracket defined in Proposition 4.2 agrees, in the  $\alpha$  coordinates, with that given above.

**Proposition 5.1.** *Consider the functions  $\phi_j^a(\mathcal{C}) = \text{Im}[i^a \mathcal{C}_{jj}]$  where  $1 \leq j \leq n$  and  $a \in \{0, 1\}$ . The Gelfand-Dikij brackets among these functions are the following:*

$$\{\phi_j^a, \phi_{j-1}^b\} = -\{\phi_{j-1}^b, \phi_j^a\} = \text{Im}[i^{b+a} \mathcal{C}_{j-1,j} \mathcal{C}_{j,j-1}] \quad (45)$$

$$= -\rho_{j-2}^2 \text{Im}[i^{b+a} \alpha_{j-3} \bar{\alpha}_{j-1}], \quad (46)$$

for all choices of  $a, b$  and  $j$ ,

$$\{\phi_j^0, \phi_j^1\} = -\{\phi_j^1, \phi_j^0\} = |\mathcal{C}_{j+1,j}|^2 - |\mathcal{C}_{j,j-1}|^2 = \rho_{j-1}^2 - \rho_{j-2}^2, \quad (47)$$

and all other pairs Poisson commute. Following the standard conventions,  $\alpha_{-1} = -1$ ,  $\rho_{-1} = 0$ , and matrix entries with indices laying outside the bounds of the matrix are zero.

*Proof.* By definition,  $\nabla\phi_j^a = i^a E_{jj}$  where  $E_{jj}$  is the elementary matrix with a 1 in the  $(j, j)$  position. Thus,

$$\{\phi_j^a, \phi_k^b\} = \frac{1}{2}\langle R(i^a E_{jj}\mathcal{C}), i^b E_{kk}\mathcal{C} \rangle - \frac{1}{2}\langle R(i^a \mathcal{C}E_{jj}), i^b \mathcal{C}E_{kk} \rangle \quad (48)$$

$$= \operatorname{Im} i^b \langle e_k | \mathcal{C} \pi_{\mathfrak{l}}(i^a E_{jj}\mathcal{C}) e_k \rangle - \operatorname{Im} i^b \langle e_k | \pi_{\mathfrak{l}}(i^a \mathcal{C}E_{jj}) \mathcal{C} e_k \rangle \quad (49)$$

$$= -\operatorname{Im} i^b \langle e_k | \mathcal{C} \pi_{\mathfrak{a}}(i^a E_{jj}\mathcal{C}) e_k \rangle + \operatorname{Im} i^b \langle e_k | \pi_{\mathfrak{a}}(i^a \mathcal{C}E_{jj}) \mathcal{C} e_k \rangle \quad (50)$$

where the last two lines follow from  $R = 2\pi_{\mathfrak{l}} - \operatorname{Id} = \operatorname{Id} - 2\pi_{\mathfrak{a}}$  and by observing that the identity terms cancel. The proof proceeds differently depending on the parity of  $j$ .

Fix  $j$ , even. Then both  $A = \pi_{\mathfrak{a}}(i^a E_{jj}\mathcal{C})$  and  $B = \pi_{\mathfrak{a}}(i^a \mathcal{C}E_{jj})$  have only three non-zero entries:

$$\begin{aligned} A_{j,j} &= i \operatorname{Im}(i^a \mathcal{C}_{j,j}), & A_{j,j+1} &= -\overline{A_{j+1,j}} = i^a \mathcal{C}_{j,j+1} \\ B_{j,j} &= i \operatorname{Im}(i^a \mathcal{C}_{j,j}), & B_{j-1,j} &= -\overline{B_{j,j-1}} = i^a \mathcal{C}_{j-1,j} \end{aligned}$$

Thus by (50), the bracket can only be non-zero for  $k \in \{j-1, j, j+1\}$ . Moreover, if  $k = j-1$  then only the first summand in (50) contributes, while for  $k = j+1$ , only the second summand contributes. Multiplying things out gives (45) and (47). Equation (46) and the last formula in (47) follow from the expressions for the matrix entries.

We now treat  $j$  odd using (49). Both  $L = \pi_{\mathfrak{l}}(i^a E_{jj}\mathcal{C})$  and  $M = \pi_{\mathfrak{l}}(i^a \mathcal{C}E_{jj})$  have only four non-zero entries:

$$\begin{aligned} L_{j,j} &= \operatorname{Re}(i^a \mathcal{C}_{j,j}), & L_{j,j-1} &= i^a \mathcal{C}_{j,j-1}, & L_{j+1,j} &= \overline{i^a \mathcal{C}_{j,j+1}}, & L_{j+2,j} &= \overline{i^a \mathcal{C}_{j,j+2}}, \\ M_{j,j} &= \operatorname{Re}(i^a \mathcal{C}_{j,j}), & M_{j+1,j} &= i^a \mathcal{C}_{j+1,j}, & M_{j,j-1} &= \overline{i^a \mathcal{C}_{j-1,j}}, & M_{j,j-2} &= \overline{i^a \mathcal{C}_{j-2,j}}. \end{aligned}$$

Thus the first summand in (49) is only non-zero for  $k \in \{j-1, j\}$ , while the second summand contributes only for  $k \in \{j, j+1\}$ . If  $k = j+1$  or  $j-1$  this leads quickly to the formulae given in the proposition. When  $k = j$  it is a little more involved:

$$\begin{aligned} \{\phi_j^a, \phi_j^b\} &= \operatorname{Im}(i^{b-a} [|\mathcal{C}_{j,j+1}|^2 + |\mathcal{C}_{j,j+2}|^2 - |\mathcal{C}_{j-1,j}|^2 - |\mathcal{C}_{j-2,j}|^2]) \\ &= \operatorname{Im}(i^{b-a} [-|\mathcal{C}_{j,j-1}|^2 + |\mathcal{C}_{j+1,j}|^2]) \end{aligned}$$

where we used the fact that the  $j$  row and column of  $\mathcal{C}$  are unit vectors. Taking  $a = 0$  and  $b = 1$  gives (47).  $\square$

To simplify our calculations, we extend the Poisson brackets defined above to be bilinear over  $\mathbb{C}$ . Note that  $\{\bar{\phi}, \bar{\psi}\}$  is then the complex conjugate of  $\{\phi, \psi\}$  and so the values of  $\{\phi, \psi\}$  and  $\{\bar{\phi}, \bar{\psi}\}$  suffice to determine all Poisson brackets between the real and imaginary parts of  $\phi$  and  $\psi$ .

In the previous proposition, we calculated the Gelfand-Dikij brackets of the real and imaginary parts of the diagonal entries. After a little computation, we obtain the following equivalent information:

$$\{\mathcal{C}_{jj}, \bar{\mathcal{C}}_{jj}\} = 2i(\rho_{j-1}^2 - \rho_{j-2}^2), \quad (51)$$

$$\{\mathcal{C}_{jj}, \mathcal{C}_{j-1,j-1}\} = -2i\rho_{j-2}^2 \alpha_{j-3} \bar{\alpha}_{j-1}, \quad (52)$$

$$\{\mathcal{C}_{jj}, \bar{\mathcal{C}}_{j-1,j-1}\} = 0, \quad (53)$$

and  $\mathcal{C}_{jj}$  commutes with all other  $\mathcal{C}_{kk}$  and  $\bar{\mathcal{C}}_{kk}$ . These equations are the key to showing that the two brackets are the same:

**Theorem 5.2.** *For any  $0 \leq k, l \leq n-2$  the Gelfand-Dikij brackets of the Verblunsky coefficients are given by*

$$\{\alpha_k, \alpha_l\} = 0 \quad \text{and} \quad \{\alpha_k, \bar{\alpha}_l\} = -2i\delta_{kl}\rho_k^2. \quad (54)$$

*That is, the Gelfand-Dikij and Ablowitz-Ladik brackets agree.*

*Proof.* We will prove (54) by induction on  $k+l$ . As  $\mathcal{C}_{11} = \bar{\alpha}_0$ , equation (51) settles the case  $k = l = 0$ . Using the  $k \leftrightarrow l$  antisymmetry, we divide the inductive step into three cases:

The case  $l = k$ : As  $\mathcal{C}_{k+1, k+1} = -\alpha_{k-1}\bar{\alpha}_k$ , equation (51) implies

$$-2i|\alpha_k|^2\rho_{k-1}^2 + |\alpha_{k-1}|^2\{\bar{\alpha}_k, \alpha_k\} = 2i(\rho_k^2 - \rho_{k-1}^2),$$

which simplifies to  $\{\alpha_k, \bar{\alpha}_k\} = -2i\rho_k^2$ , because  $\alpha_{k-1}$  is non-zero on a dense set.

The case  $l > k+1$ : As  $\mathcal{C}_{k+1, k+1} = -\alpha_{k-1}\bar{\alpha}_k$ , with the usual convention  $\alpha_{-1} = -1$ ,

$$0 = \{\mathcal{C}_{k+1, k+1}, \mathcal{C}_{l+1, l+1}\} = \alpha_{k-1}\alpha_{l-1}\{\bar{\alpha}_k, \bar{\alpha}_l\}.$$

Similarly  $\{\bar{\mathcal{C}}_{k+1, k+1}, \mathcal{C}_{l+1, l+1}\} = 0$  implies  $\{\alpha_k, \bar{\alpha}_l\} = 0$ .

The case  $l = k+1$ : By (53), we have  $\{\mathcal{C}_{k+1, k+1}, \bar{\mathcal{C}}_{k+2, k+2}\} = 0$  from which we deduce  $\{\bar{\alpha}_k, \bar{\alpha}_{k+1}\} = 0$  as above. Finally, by (52) with  $j = k+2$ ,

$$2i\rho_k^2\alpha_{k-1}\bar{\alpha}_{k+1} = \{\alpha_{k-1}\bar{\alpha}_k, \alpha_k\bar{\alpha}_{k+1}\} = 2i\alpha_{k-1}\bar{\alpha}_{k+1}\rho_k^2 + \alpha_{k-1}\alpha_k\{\bar{\alpha}_k, \bar{\alpha}_{k+1}\}.$$

Cancelling gives  $\{\alpha_k, \alpha_{k+1}\} = 0$ .  $\square$

**Remark 5.3.** The evolution of the Verblunsky coefficients in the Ablowitz-Ladik bracket under the flow generated by  $2\operatorname{Re}(K_1)$  is the Ablowitz-Ladik evolution (see [1] and [2]):

$$\{\alpha_j, 2\operatorname{Re}(K_1)\}_2 = i\rho_j^2(\alpha_{j-1} + \alpha_{j+1}).$$

The proof of Theorem 5.2 was based around the computation of the Gelfand-Dikij brackets of diagonal entries. While it is possible to directly compute the Gelfand-Dikij bracket of any pair of entries of a CMV matrix, even the formulae for the Hamiltonian vector fields are rather unpleasant. However, in the particular case of exposed entries, things are not too painful:

**Lemma 5.4.** *Fix  $k, l$  so that  $(k, l)$  is an exposed entry for CMV matrices. The evolution of CMV matrices under the Hamiltonian  $r_{kl}(B) = \operatorname{Re}(B_{kl})$  is given by*

$$i\dot{\mathcal{C}} = (-1)^k r_{kl}(\mathcal{C})[E_{kk}\mathcal{C} - \mathcal{C}E_{ll}]. \quad (55)$$

*Proof.* We will use the notation  $E_{kl}$  for the matrix with 1 in the  $(k, l)$  position and zeros everywhere else.

As  $\nabla r_{kl} = iE_{lk}$ , equation (23) implies

$$\dot{\mathcal{C}} = \mathcal{C}\pi_l(iE_{lk}\mathcal{C}) - \pi_l(i\mathcal{C}E_{lk})\mathcal{C}. \quad (56)$$

We proceed differently depending on the parity of  $k$ .

As noted in Definition 1.4, the exposed entries with  $k$  odd all take the form  $(2j-1, 2j+1)$  with  $j \geq 1$ . In this case, both  $E_{lk}\mathcal{C}$  and  $\mathcal{C}E_{lk}$  are lower triangular and the sole entry on the diagonal is positive (its value is  $\mathcal{C}_{kl}$ ). Thus we are interested in the value of  $\pi_l(B)$  where  $B$  is lower triangular with purely imaginary diagonal. It is easy to see that in this case,  $\pi_l(B) = B - D$ , where  $D$  is the diagonal part of  $B$ . Consequently, (56) reduces to

$$\dot{\mathcal{C}} = -i\mathcal{C}_{kl}[\mathcal{C}E_{ll} - E_{kk}\mathcal{C}]$$

in agreement with (55).

For  $k$  even,  $iE_{lk}\mathcal{C}$  and  $i\mathcal{C}E_{lk}$  are upper triangular with purely imaginary diagonal. For matrices  $B$  of this form,  $\pi_1(B) = B^\dagger + D$ , where  $D$  is again the diagonal part of  $B$ . As  $\mathcal{C}^\dagger\mathcal{C} = \mathcal{C}\mathcal{C}^\dagger = \text{Id}$  for CMV matrices, only the contribution from  $D$  survives. The fact that this now appears with a plus sign results in the  $(-1)^k$  factor in (55).  $\square$

## 6. ABLOWITZ–LADIK: EVOLUTION OF THE SPECTRAL MEASURE.

Shortly after the discovery of a Lax pair representation for the Toda Lattice [14], Moser gave a complete solution for the finite system, [24]. Specifically, he discovered the angle variables associated to the actions of Hénon [19] and Flaschka [14]. In addition, he studied the long-time asymptotics and determined the scattering map.

In this section, we will discuss some corresponding results for the Ablowitz-Ladik system. The remaining parts of the analogue of Moser's solution can be found in Section 7. A special case of Corollary 6.5 has already appeared, [25]. The approach used there was to begin with a special case of (70) and determine the induced evolution on the Verblunsky coefficients.

(To be precise, Moser did not check that his ‘angles’ Poisson commute. This follows as a special case of [10, Theorem 1].)

The key observation of Moser was that the spectral measure associated to the Jacobi matrix and the vector  $e_1$  has a very simple evolution. We will show that the same is true for the Ablowitz-Ladik system. To do this, we need a few lemmas; those eager to see the result should skip ahead to Proposition 6.3 and its Corollaries.

**Lemma 6.1.** *For any matrices  $B$  and  $C$ ,*

$$R(B^\dagger) = R(B) + B - B^\dagger \quad (57)$$

$$R(BP) = BP - P(B - B^\dagger)P \quad (58)$$

$$R([C, P]) = CP + PC - 2C^\dagger P - 2P(C - C^\dagger)P \quad (59)$$

where  $P$  is the rank-one projection  $|e_1\rangle\langle e_1|$ .

*Proof.* The first result is very simple:  $B - B^\dagger$  is anti-Hermitian and so  $R(B - B^\dagger) = -B + B^\dagger$ . The result now follows from the fact that  $R$  is linear over  $\mathbb{R}$ .

Next, notice that  $B|e_1\rangle\langle e_1|$  has non-zero entries only in the first column. Therefore, the anti-hermitian part is just  $\frac{1}{2}P(B - B^\dagger)P$ . Equation (58) now follows easily from (12).

The value of  $R(PC)$  can be deduced from (58) by first applying (57) with  $B = C^\dagger P$ . The result is  $R(PC) = 2C^\dagger P - PC + P(C - C^\dagger)P$ . Combining this with (58) gives (59).  $\square$

**Lemma 6.2.** *Given a polynomial  $g$ , let  $P = |e_1\rangle\langle e_1|$  and  $\psi : \mathfrak{g} \rightarrow \mathbb{R}$  by*

$$\psi(B) = \text{Re} \langle e_1 | g(B) e_1 \rangle = \langle ig(B), P \rangle \quad (60)$$

then  $[B, \nabla\psi] = i[g(B), P]$  and so

$$R([B, \nabla\psi]) = ig(B)P + iPg(B) + 2ig(B)^\dagger P - 4i\psi(B)P. \quad (61)$$

*Proof.* As the operations involved are linear over  $\mathbb{R}$ , it suffices to treat the case  $g(B) = i^l B^k$ . By cycling the trace,

$$\nabla\psi = i^{l+1} \sum_{q=0}^{k-1} B^q P B^{k-q-1}.$$

Thus, the sum in  $[B, \nabla\psi]$  telescopes to give  $i^{l+1}[B^k, P]$ . Lastly, (61) follows from this and Lemma 6.1.  $\square$

**Proposition 6.3.** *Given a polynomials  $f$  and  $g$ , let us define maps from  $\mathfrak{g}$  to  $\mathbb{R}$  by*

$$\phi(B) = \operatorname{Im} \operatorname{tr}\{f(B)\} \quad \text{and} \quad \psi(B) = \operatorname{Re} \langle e_1 | g(B) e_1 \rangle.$$

*Taking  $\phi$  as the Hamiltonian,*

$$\dot{\psi}(B) = \operatorname{Re} \langle e_1 | B f'(B) (g(B) + g(B)^\dagger) e_1 \rangle - 2\psi(B) \operatorname{Re} \langle e_1 | B f'(B) e_1 \rangle \quad (62)$$

*In particular, if  $d\mu$  is the spectral measure associated to  $e_1$  and a CMV matrix  $\mathcal{C}$ , then writing  $F(z) = 2 \operatorname{Re} z f'(z)$ , we have*

$$\partial_t \int G d\mu = \int F G d\mu - \int F d\mu \int G d\mu \quad (63)$$

*for any function  $G : S^1 \rightarrow \mathbb{R}$ .*

*Proof.* From Proposition 4.5 we have

$$\dot{\psi}(B) = \frac{1}{2} \langle [B, R(Bf'(B))], \nabla\psi \rangle = -\frac{1}{2} \langle R(Bf'(B)), [B, \nabla\psi] \rangle \quad (64)$$

$$= \frac{1}{2} \langle Bf'(B), R([B, \nabla\psi]) \rangle. \quad (65)$$

Now we can apply (61) from Lemma 6.2 to deduce

$$\dot{\psi}(B) = \langle iBf'(B)(g(B) + g(B)^\dagger), P \rangle - 2\psi(B) \langle iBf'(B), P \rangle, \quad (66)$$

which says the same thing as (62).  $\square$

**Corollary 6.4.** *Given a pair of polynomials  $f_1$  and  $f_2$ , the Hamiltonians  $\phi_j(B) = \operatorname{Im} \operatorname{tr}\{f_j(B)\}$  give rise to the same integral curve through the CMV matrix  $\mathcal{C}$  if and only if  $z \mapsto \operatorname{Re}[zf_1'(z) - zf_2'(z)]$  is constant on the spectrum of  $\mathcal{C}$ .*

*Proof.* From Proposition 4.4, we know that both integral curves remain in the set of CMV matrices. As a CMV matrix is entirely determined by its spectral measure (cf. Theorem 2.3), it suffices to understand whether the two Hamiltonians give rise to the same evolution for this measure. From (63), the answer is clear: the integral curves are the same if and only if

$$\int [F_1 - F_2] G d\mu = \int [F_1 - F_2] d\mu \int G d\mu \quad (67)$$

for all functions  $G : S^1 \rightarrow \mathbb{R}$ . That is, if and only if  $F_1 - F_2$  agrees with a constant  $d\mu$ -almost everywhere. Lastly, the support of  $d\mu$  is precisely the spectrum of  $\mathcal{C}$ .  $\square$

Since the vector  $e_1$  is cyclic for any  $n \times n$  CMV matrix  $\mathcal{C}$ , the associated spectral measure  $\mu$  is supported at  $n$  points. Conversely, if  $\mu$  is a probability measure on the circle which is supported at  $n$  points, then by the results in Section 2 it is the spectral measure for  $(\mathcal{C}, e_1)$ , with  $\mathcal{C}$  the CMV matrix representing multiplication by

$z$  in  $L^2(d\mu)$ . It is natural to parameterize the spectral measure,  $d\mu$ , in terms of the eigenvalues and the mass  $d\mu$  gives to them:

$$\int f(z) d\mu(z) = \sum_{j=1}^n f(e^{i\theta_j}) \mu_j. \quad (68)$$

By the observations of the previous paragraph,  $\theta_j$  and  $\mu_j$  are well-defined on the manifold of CMV matrices.

By Theorem 2.3, one may view  $\theta_1, \dots, \theta_{n-1}, \mu_1, \dots, \mu_{n-1}$  as a system of coordinates on the manifold of CMV matrices with fixed determinant. From part (ii) of Proposition 4.5 we know that the functions  $\theta_j$  Poisson commute. In particular, they do not change under Hamiltonians of the form  $\phi(B) = \text{Im tr}\{f(B)\}$ . The evolution of  $\mu_j$  is easily determined from the proposition above.

**Corollary 6.5.** *Under the flow generated by  $\phi(B) = \text{Im tr}\{f(B)\}$ ,*

$$\partial_t \log[\mu_j] = \{\phi, \log[\mu_j]\} = F(e^{i\theta_j}) - \sum_{l=1}^n F(e^{i\theta_l}) \mu_l \quad (69)$$

where  $F(z) = 2 \text{Re } z f'(z)$ . Consequently,

$$\mu_j(t) = \frac{\exp[F(e^{i\theta_j}) t] \mu_j(0)}{\sum \exp[F(e^{i\theta_l}) t] \mu_l(0)} \quad (70)$$

and for any  $j, l \in \{1, \dots, n-1\}$ ,

$$\{\theta_l, \frac{1}{2} \log[\mu_j / \mu_n]\} = \delta_{jl}. \quad (71)$$

*Proof.* Equation (69) amounts to nothing more than (63) when  $G(z)$  is the characteristic function of the set  $\{e^{i\theta_j}\}$ . Rather than bother solving this system of ODEs directly, it is simple enough to check that (70) is indeed the solution.

Consider (69) with  $f(z) = z^k$  and  $f(z) = iz^k$  for  $k = 1, 2, \dots, n$ . Combining pairs gives

$$\sum ik e^{ik\theta_l} \{\theta_l, \log[\mu_j]\} = 2ik \left[ e^{ik\theta_j} - \sum e^{ik\theta_p} \mu_p \right]$$

This is a system of equations for  $\{\theta_l, \log[\mu_j]\}$  that has a unique solution (the Vandermonde determinant does not vanish):

$$\{\theta_l, \log[\mu_j]\} = 2\delta_{lj} - 2\mu_l. \quad (72)$$

Subtraction now gives (71).  $\square$

Equation (71) gives an obvious candidate for a system of canonically conjugate (or Darboux) coordinates:

$$\theta_1, \dots, \theta_{n-1}, \frac{1}{2} \log[\mu_1 / \mu_n], \dots, \frac{1}{2} \log[\mu_{n-1} / \mu_n].$$

Moreover, the natural analogues for the Toda lattice are canonically conjugate, [10, Theorem 1]. In Section 8, we will show that the variables  $\log[\mu_k / \mu_n]$  do not commute and use the values of their brackets to find a system of Darboux coordinates.

## 7. ABLOWITZ–LADIK: ASYMPTOTICS AND SCATTERING.

In this section, we study the asymptotics of solutions to the equations in the Ablowitz-Ladik hierarchy. These asymptotics will be expressed both in terms of the spectral measure  $d\mu$  and the Verblunsky coefficients. In the former case, the answer is an immediate corollary of the results in the previous section.

Let us fix a Hamiltonian from the Ablowitz-Ladik hierarchy:

$$\phi(\mathcal{C}) = \operatorname{Im} \operatorname{tr}(f(\mathcal{C})), \quad (73)$$

where  $f$  is a polynomial. As in the previous section, we will use the shorthand  $F(z) = 2 \operatorname{Re}[zf'(z)]$ .

We will write the spectral measure as

$$\int f(z) d\mu(z) = \sum_{k=1}^n f(z_k) \mu_k. \quad (74)$$

Naturally,  $\mathcal{C}$ , the spectral measure, and the Verblunsky coefficients evolve under the flow generated by  $\phi$ ; we use  $\mathcal{C}(t)$ ,  $\mu_k(t)$ , and  $\alpha_j(t)$  to denote these quantities at time  $t$ . By Proposition 4.5 part (ii), the eigenvalues do not change under this flow; hence there is no need consider  $z_k(t)$ . When we omit the time dependence, we refer to the initial data.

In defining  $\mu_k$  and  $z_k$ , we can choose any ordering we please; however, there is a particular condition on this choice that simplifies the formulae below. Namely, we require that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n, \quad (75)$$

where we use the shorthand  $\lambda_k = F(z_k) = 2 \operatorname{Re}[z_k f'(z_k)]$ . Of course generically, the ordering will be strict. Note that by the continuity of  $F$  this labeling is well-defined on an open set which is invariant under the flow of  $\phi$ .

**Proposition 7.1.** *Under the flow generated by the Hamiltonian  $\phi$ , the masses have the following asymptotics*

$$\log[\mu_k(t)] = -(\lambda_1 - \lambda_k)t + \log \left[ \frac{\mu_k}{\mu_1 + \cdots + \mu_\nu} \right] + O(e^{-at}) \quad (76)$$

as  $t \rightarrow \infty$ . Here  $\nu$  is defined by  $\lambda_1 = \cdots = \lambda_\nu > \lambda_{\nu+1}$  and  $a = \lambda_1 - \lambda_{\nu+1} > 0$ . In particular if  $k > \nu$ , then  $\mu_k \rightarrow 0$  exponentially fast.

*Proof.* Equation (76) is an immediate consequence of (70), which says

$$\log[\mu_k(t)] = \lambda_k t + \log[\mu_k] - \log \left[ \sum_{l=1}^n e^{\lambda_l t} \mu_l \right].$$

The definition of  $\nu$  implies

$$\sum_{m=1}^n \mu_m e^{\lambda_m t} = e^{\lambda_1 t} (\mu_1 + \cdots + \mu_\nu) [1 + O(e^{-at})].$$

Combining these two formulae completes the proof.  $\square$

We now turn to studying the asymptotics of the Verblunsky coefficients. In order to simplify the formulae that follow, we will use multi-index notation. All

our multi-indices will be ordered. Given a multi-index  $I = (i_1 < \dots < i_l)$  of length  $l$ , we write

$$\mu_I = \prod_{k=1}^l \mu_{i_k}, \quad z_I = \prod_{k=1}^l z_{i_k}, \quad (77)$$

and we abbreviate Vandermonde determinants as follows:

$$\Delta(\zeta_1, \dots, \zeta_m) = \det(\zeta_l^{k-1}) = \prod_{1 \leq j < k \leq m} [\zeta_k - \zeta_j] \quad (78)$$

$$\Delta_I = \Delta(z_{i_1}, \dots, z_{i_l}). \quad (79)$$

We follow the natural convention that the Vandermonde determinant of a single number is 1.

The key to converting asymptotic information on the measure into asymptotics for the Verblunsky coefficients is the following lemma.

**Lemma 7.2.** *Let  $d\mu = \sum_{k=1}^n \mu_k \delta_{z_k}$  be a discrete measure on the unit circle and let  $\hat{\mu}(p) = \int z^p d\mu(z)$ . Then for each  $1 \leq m \leq n$ ,*

$$\det[\hat{\mu}(k-l)] = \sum_I |\Delta_I|^2 \mu_I = \prod_{j=0}^{m-2} \rho_j^{2(m-j-1)}, \quad (80)$$

$$\det[\hat{\mu}(k-l-1)] = \sum_I |\Delta_I|^2 \mu_I \bar{z}_I = (-1)^{m-1} \alpha_{m-1} \prod_{j=0}^{m-2} \rho_j^{2(m-j-1)}, \quad (81)$$

where the determinants are taken over  $1 \leq k, l \leq m$  and both sums are over ordered multi-indices  $I = (1 \leq i_1 < i_2 < \dots < i_m \leq n)$ .

*Proof.* The proof is an elementary application of a special case of the Cauchy–Binet Formula (see [16, p. 9] or [32, Lemma 36.2]): Suppose  $m \leq n$ . Given an  $m \times n$  matrix  $A$  and a  $n \times n$  diagonal matrix  $D$ ,

$$\det(ADA^\dagger) = \sum_{i_1 < \dots < i_m} \left\{ |\det(A_{k, i_l})|^2 \prod_{r=1}^m D_{i_r, i_r} \right\}. \quad (82)$$

To prove the first equalities in (80) and (81), we choose

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_n \\ \vdots & \vdots & & \vdots \\ z_1^{m-1} & z_2^{m-1} & \dots & z_n^{m-1} \end{bmatrix}$$

with  $D = \text{diag}[\mu_1, \dots, \mu_n]$  and  $D = \text{diag}[\mu_1 \bar{z}_1, \dots, \mu_n \bar{z}_n]$ , respectively. The right-hand side of (82) reduces to the form given above because of (78). Evaluating the matrix products  $ADA^\dagger$  leads to the Toeplitz matrices given on the left-hand sides of (80) and (81).

To prove the second equality in (80), we apply row operations to  $A$ :

$$EA = \begin{bmatrix} \Phi_0(z_1) & \dots & \Phi_0(z_n) \\ \vdots & & \vdots \\ \Phi_{m-1}(z_1) & \dots & \Phi_{m-1}(z_n) \end{bmatrix}$$



for some lower-triangular matrix  $E$  of determinant one. The orthogonality of  $\Phi_j$  with respect to  $d\mu$  shows that  $EADA^\dagger E^\dagger$  is a diagonal matrix with entries  $\|\Phi_j\|^2$ . The product of the squared norms of  $\Phi_j$  reduces to (80) via (7).

The second part of (81) follows from (80) by the Heine Formula, [29, Theorem 1.5.11], which says  $(-1)^{m-1}\alpha_{m-1}$  is the ratio of the left-hand sides of (80) and (81).  $\square$

**Theorem 7.3.** *Fix  $1 \leq k \leq n-1$ . Let  $B_k = \{l : \lambda_l = \lambda_k\}$ , let  $s(k)$  be the largest index not contained in  $B_k$ . Then, as  $t \rightarrow \infty$ ,*

$$\alpha_{k-1}(t) \rightarrow (-1)^{k-1} \bar{z}_J \frac{\sum |\Delta_{J \cup I}|^2 \mu_I \bar{z}_I}{\sum |\Delta_{J \cup I}|^2 \mu_I}, \quad (83)$$

where  $J = (1 < 2 < \dots < s(k))$  and both sums are over all ordered multi-indices  $I \subseteq B_k$  of length  $k - s(k)$ . In particular,  $|\alpha_{k-1}| \rightarrow 1$  if and only if  $\lambda_k > \lambda_{k+1}$ .

If all  $\lambda_j$  are distinct,

$$\alpha_{k-1}(t) = (-1)^{k-1} \bar{z}_1 \cdots \bar{z}_k \left[ 1 + \xi_{k-1} e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t}) \right] \quad (84)$$

where

$$\xi_{k-1} = (z_k \bar{z}_{k+1} - 1) \frac{\mu_{k+1}}{\mu_k} \prod_{l=1}^{k-1} \left| \frac{z_{k+1} - z_l}{z_k - z_l} \right|^2 \quad (85)$$

and  $\gamma > (\lambda_k - \lambda_{k+1}) > 0$ .

*Proof.* By taking the ratio of (80) and (81), we obtain

$$\alpha_{k-1}(t) = (-1)^{k-1} \frac{\sum |\Delta_L|^2 \mu_L(t) \bar{z}_L}{\sum |\Delta_L|^2 \mu_L(t)} \quad (86)$$

where the sums are over all ordered multi-indices,  $L$ , of length  $k$ .

The key to proving (83) is determining the asymptotically dominant terms in these sums. By Proposition 7.1, if  $\lambda_p > \lambda_q$ , then the ratio  $\mu_q(t)/\mu_p(t)$  converges to zero exponentially fast. Therefore, the dominant terms arise from those multi-indices for which  $\lambda_{l_1} + \dots + \lambda_{l_k}$  is maximal. As the  $\lambda$ s are decreasing, the maximal value is  $\lambda_1 + \dots + \lambda_k$  and the set of multi-indices that achieve this value is exactly the collection of  $J \cup I$  stated in the theorem. This shows that

$$\alpha_{k-1}(t) \approx (-1)^{k-1} \frac{\sum |\Delta_{J \cup I}|^2 \mu_{J \cup I}(t) \bar{z}_{J \cup I}}{\sum |\Delta_{J \cup I}|^2 \mu_{J \cup I}(t)}, \quad (87)$$

where  $A \approx B$  means that the ratio  $A/B \rightarrow 1$  exponentially fast. (We will maintain this convention throughout of the proof.) Note that  $z_J$  can be factored out, leaving  $z_I$ .

By Proposition 7.1,

$$\mu_{J \cup I}(t) \approx \frac{\mu_J \mu_I}{(\mu_1 + \dots + \mu_\nu)^k} \exp[(\lambda_1 + \dots + \lambda_k)t - k\lambda_1 t]$$

for the multi-indices  $I$  under consideration. Substituting this into (87) and cancelling common factors in the numerator and denominator gives (83).

From (83) we see that the limiting value of  $(-1)^{k-1} z_J \alpha_{k-1}(t)$  is a convex combination of points on the unit circle. This sum contains exactly one term if and only if  $\lambda_k > \lambda_{k+1}$ . Therefore under this condition,  $|\alpha_{k-1}(t)| \rightarrow 1$ . If the sum contains more than one term, then it contains two multi-indices that differ at only one index.

Thus the limiting value of  $\alpha_{k-1}(t)$  is a non-trivial convex combination of points on the unit circle and so lies strictly inside the unit disk.

We will now consider the case of distinct  $\lambda_j$ . This amounts to studying the asymptotics of  $\mu_L(t)$  for different choices of ordered multi-indices,  $L = (l_1 < \dots < l_k)$ . By Proposition 7.1,

$$\log[\mu_L(t)] \approx -t \sum_{j=1}^k (\lambda_1 - \lambda_{l_j}) + \log[\mu_L] - k \log[\mu_1]$$

and so the question reduces to finding the largest two values of  $\lambda_{l_1} + \dots + \lambda_{l_k}$ . Because  $\lambda_1 > \lambda_2 > \dots > \lambda_n$ , the largest value is  $\lambda_1 + \dots + \lambda_k$ ; the second largest is  $\lambda_1 + \dots + \lambda_{k-1} + \lambda_{k+1}$ . All other choices of  $L$  lead to strictly smaller values for the sum. We write  $\tilde{\gamma}$  for the difference between the largest and third largest values of this sum.

Returning to (87) and factoring out the dominant term in both the numerator and the denominator gives

$$\begin{aligned} (-1)^{k-1} z_1 \dots z_k \alpha_{k-1} &= \frac{1 + z_k \bar{z}_{k+1} P_k e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\tilde{\gamma}t})}{1 + P_k e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\tilde{\gamma}t})} \\ &= 1 + (z_k \bar{z}_{k+1} - 1) P_k e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t}), \end{aligned}$$

where  $\gamma = \min\{\tilde{\gamma}, 2(\lambda_k - \lambda_{k+1})\}$  and

$$P_k = \frac{\mu_{k+1}}{\mu_k} \left| \frac{\Delta(z_1, \dots, z_{k-1}, z_{k+1})}{\Delta(z_1, \dots, z_k)} \right|^2.$$

Cancelling common factors in the Vandermonde determinants leads to the formula given in the theorem.  $\square$

If all  $\lambda_j$  are distinct, then viewed as a curve in the disk,  $\alpha_{k-1}(t)$  approaches the boundary in a fixed non-tangential direction. This simply amounts to the statement that  $\xi_{k-1}$  is non-zero and  $\arg(\xi_{k-1}) = \arg(z_k \bar{z}_{k+1} - 1)$  belongs to  $(-\pi/2, \pi/2)$ . Let us also note that the asymptotics of  $\rho_{k-1}$  are easily deduced from (84):

$$\rho_{k-1}^2(t) = -2 \operatorname{Re}(\xi_{k-1}) e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t}) \quad (88)$$

$$= |z_{k+1} - z_k|^2 \frac{\mu_{k+1}}{\mu_k} \prod_{l=1}^{k-1} \left| \frac{z_{k+1} - z_l}{z_k - z_l} \right|^2 e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t}). \quad (89)$$

This shows that the factors  $\mathcal{L}(t)$  and  $\mathcal{M}(t)$  of the CMV matrix  $\mathcal{C}(t)$  diagonalize as  $t \rightarrow \infty$  and hence so does  $\mathcal{C}(t)$ . Moreover, the eigenvalues are ordered by the corresponding value of  $F(z)$ . This is a well known phenomenon for the Toda Lattice.

When the  $\lambda_j$  are not all distinct,  $\mathcal{C}(t)$  converges to a direct sum of CMV matrices and their adjoints. Specifically, if

$$\lambda_{k-1} > \lambda_k = \dots = \lambda_{k+m} > \lambda_{k+m+1},$$

then  $\alpha_{k-1}, \dots, \alpha_{k+m-2}$  do not approach the boundary and  $\mathcal{C}(\infty)$  has a non-trivial block of size  $m$  beginning at row/column  $k$ . If  $k$  is odd, this will be a CMV matrix; if  $k$  is even, it will be the adjoint of a CMV matrix.

While this phenomenon cannot occur for the Toda Lattice, it can occur for Hamiltonians in the same hierarchy. Some examples of non-diagonalization are discussed on page 389 of [9]. However, we do not know if anyone has troubled to write down the analogue of Theorem 7.3.

We now turn our attention to scattering; we will only consider the case of distinct  $\lambda$ s. Thus far, we have only discussed the behaviour as  $t \rightarrow +\infty$ ; however the behaviour for  $t \rightarrow -\infty$  can easily be deduced from this by reversing the sign of the Hamiltonian. This results in the following changes to the  $t \rightarrow \infty$  formulae:

$$\lambda_k \mapsto -\lambda_{n-k+1}, \quad z_k \mapsto z_{n-k+1}, \quad \text{and} \quad \mu_k \mapsto \mu_{n-k+1}$$

In particular,

$$\alpha_{k-1}(t) = (-1)^{k-1} \bar{z}_n \cdots \bar{z}_{n-k+1} \left[ 1 + \zeta_{k-1} e^{(\lambda_{n-k} - \lambda_{n-k+1})t} + O(e^{\gamma t}) \right]$$

as  $t \rightarrow -\infty$  where

$$\zeta_{k-1} = (z_{n-k+1} \bar{z}_{n-k} - 1) \frac{\mu_{n-k}}{\mu_{n-k+1}} \prod_{l=n-k+2}^n \left| \frac{z_{n-k} - z_l}{z_{n-k+1} - z_l} \right|^2 \quad (90)$$

and  $\gamma > (\lambda_{n-k} - \lambda_{n-k+1}) > 0$ .

The scattering map is the transformation that links the asymptotics of solutions as  $t \rightarrow -\infty$  to those as  $t \rightarrow \infty$ . The result follows easily from the formulae given above:

**Proposition 7.4.** *Suppose all  $\lambda_j$  are distinct, then*

$$\alpha_{k-1}(+\infty) \alpha_{n-k-1}(-\infty) = (-1)^n \bar{z}_1 \cdots \bar{z}_n = (-1)^n \det(\mathcal{C}^\dagger) = -\alpha_{n-1},$$

which is independent of  $t$ . Furthermore,

$$\xi_{k-1} \zeta_{n-k-1} = |z_{k+1} - z_k|^2 \prod_{l=1}^{k-1} \left| \frac{z_{k+1} - z_l}{z_k - z_l} \right|^2 \prod_{m=k+2}^n \left| \frac{z_k - z_m}{z_{k+1} - z_m} \right|^2.$$

The eigenvalues  $z_j$  are time independent and can be recovered from the asymptotic values of the Verblunsky coefficients as follows:

$$z_j = -\frac{\alpha_{j-2}(+\infty)}{\alpha_{j-1}(+\infty)} = -\frac{\alpha_{n-j-1}(-\infty)}{\alpha_{n-j}(-\infty)}.$$

The corresponding result for the Toda Lattice can be found in [24, §4].

As noted earlier, the Toda flow diagonalizes Jacobi matrices. One may ask if it suggests a good algorithm for doing this. Indeed it does, namely the venerable QR algorithm. See [11, 12].

For unitary matrices, the naive QR algorithm is useless: the QR factorization of  $U$  is  $U \cdot \text{Id}$ . However, by incorporating shifting and deflation one does obtain an algorithm that works. In view of parts (iv) and (v) of Proposition 4.5, any algorithm based on the Ablowitz–Ladik flows will merely amount to the application of the QR method to a (possibly step dependent) function of  $U$ . We do not wish to be entirely pessimistic, particularly in light of the storage savings afforded by CMV matrices. One may try the following algorithm: Iterate

$$\mathcal{C} \mapsto Q\mathcal{C}Q^\dagger \quad \text{where} \quad \mathcal{C} \pm \mathcal{C}^\dagger = LQ.$$

The choice of sign may be fixed or vary. Whether this has any real numerical benefit, is a matter for experiment.

## 8. CANONICAL COORDINATES

In this section, we extend  $\theta_1, \dots, \theta_{n-1}$  to a system of Darboux coordinates; these Poisson commute by Proposition 4.5. Based on (71) it is natural to imagine that

$$\frac{1}{2} \log[\mu_1/\mu_n], \dots, \frac{1}{2} \log[\mu_{n-1}/\mu_n] \quad (91)$$

are the missing variables as is true for Jacobi matrices (a special case of [10, Theorem 1]). As we will see, they are not, because they do not Poisson commute. The remedy is well known (see [5, Ch. 4]): the conjugate variables can be written in the form

$$\varphi_l = \frac{1}{2} \log[\mu_l/\mu_n] + f_l(\theta_1, \dots, \theta_{n-1})$$

where  $f_l$  are determined by the system of differential equations

$$\frac{\partial f_l}{\partial \theta_k} - \frac{\partial f_k}{\partial \theta_l} = \left\{ \frac{1}{2} \log[\mu_k/\mu_n], \frac{1}{2} \log[\mu_l/\mu_n] \right\}. \quad (92)$$

This system is soluble via the Poincaré lemma.

We begin by calculating the right hand side of (92). There doesn't seem to be any easy way to do this. Because of the permutation symmetry in the indices, it suffices to compute  $\{\log[\mu_2/\mu_1], \log[\mu_3/\mu_1]\}$ .

**Lemma 8.1.** *For any enumeration of the eigenvalues  $e^{i\theta_1}, \dots, e^{i\theta_n}$ , the functions*

$$\Psi_{j,k}(\mathcal{C}) = \{\log[\mu_j/\mu_1], \log[\mu_k/\mu_1]\}$$

*Poisson commute with all Hamiltonians of the form  $\text{Im tr } f(\mathcal{C})$  and so remain constant under the corresponding flows.*

*Proof.* By the Leibnitz rule, it suffices to prove that  $\Psi_{j,k}$  commutes with  $\theta_l$  for each  $l = 1, \dots, n$ . For this case, we merely need to look to the Jacobi identity and (72):

$$\begin{aligned} & -\left\{ \left\{ \log[\mu_j/\mu_1], \log[\mu_k/\mu_1] \right\}, \theta_l \right\} \\ &= \left\{ \left\{ \log[\mu_k/\mu_1], \theta_l \right\}, \log[\mu_j/\mu_1] \right\} + \left\{ \left\{ \theta_l, \log[\mu_j/\mu_1] \right\}, \log[\mu_k/\mu_1] \right\} \\ &= \left\{ -2\delta_{lk}, \log[\mu_j/\mu_1] \right\} + \left\{ 2\delta_{lj}, \log[\mu_k/\mu_1] \right\} \\ &= 0, \end{aligned}$$

which completes the proof.  $\square$

**Lemma 8.2.** *Fix distinct points  $\zeta_1, \dots, \zeta_n$  on the unit circle and consider the collection of CMV matrices with determinant  $\prod \zeta_k$  that lie within the  $\epsilon$ -neighbourhood of  $\text{diag}(\zeta_1, \dots, \zeta_n)$  as measured in the Hilbert–Schmidt norm. For  $\epsilon$  sufficiently small, there are analytic functions  $G$  and  $G_j$  of the  $3(n-1)$  variables  $\alpha_k, \bar{\alpha}_k, \rho_k$ ,  $0 \leq k \leq n-2$ , so that*

$$\mu_1 = 1 - \rho_0^2 G(\boldsymbol{\alpha}, \bar{\boldsymbol{\alpha}}, \boldsymbol{\rho}) \quad (93)$$

$$\log\left[\frac{\mu_2}{\mu_1}\right] = \log[\rho_0^2] + \log\left[\frac{\bar{\alpha}_1}{(\bar{\alpha}_0 + \bar{\alpha}_1 \alpha_0)^2}\right] + \sum \rho_j G_j(\boldsymbol{\alpha}, \bar{\boldsymbol{\alpha}}, \boldsymbol{\rho}), \quad (94)$$

where  $\mu_j$  is the mass associated to the eigenvalue nearest  $\zeta_j$ .

*Proof.* Let  $D$  denote the diagonal of the CMV matrix and  $E$ , the remainder:  $\mathcal{C} = D + E$ . We will also write  $P_j$  for the orthogonal projection onto the space spanned by  $e_j$ . These are, of course, the eigenprojections for  $D$ .

As the  $\zeta_j$  are distinct, we can construct disjoint circular contours around each of these points. We orient these in the anti-clockwise direction and label them  $\Gamma_j$ .

For  $\epsilon$  sufficiently small, each contour encircles exactly one eigenvalue of  $\mathcal{C}$  and the corresponding eigenprojection is given by

$$\begin{aligned} \tilde{P}_j &= \oint_{\Gamma_j} (z - D - E)^{-1} \frac{dz}{2\pi i} = P_j + \oint_{\Gamma_j} (z - D)^{-1} E (z - D)^{-1} \frac{dz}{2\pi i} \\ &\quad + \oint_{\Gamma_j} (z - D)^{-1} E (z - D - E)^{-1} E (z - D)^{-1} \frac{dz}{2\pi i}. \end{aligned}$$

The second equality follows from two applications of the resolvent identity.

Now  $\mu_j = \text{tr}(P_1 \tilde{P}_j) = \langle e_1 | \tilde{P}_j e_1 \rangle$  and so using the fact that  $D_{11} = \bar{\alpha}_0$ , we obtain

$$\mu_j = \delta_{1j} + \frac{1}{2\pi i} \oint_{\Gamma_j} \langle e_1 | E (z - D - E)^{-1} E e_1 \rangle \frac{dz}{(z - \bar{\alpha}_0)^2}.$$

The term that is linear in  $E$  disappears because it has only a double pole when  $j = 1$  and is analytic when  $j \neq 1$ . As  $E e_1 = \rho_0 e_2$  and  $E^\dagger e_1 = \alpha_1 \rho_0 e_2 + \rho_0 \rho_1 e_3$ ,

$$\mu_j = \delta_{1j} + \frac{\rho_0^2}{2\pi i} \oint_{\Gamma_j} \langle \alpha_1 e_2 + \rho_1 e_3 | (z - D - E)^{-1} e_2 \rangle \frac{dz}{(z - \bar{\alpha}_0)^2}. \quad (95)$$

To see that the integral gives an analytic function of  $\alpha$ ,  $\bar{\alpha}$ , and  $\rho$ , we need merely write  $(z - D - E)^{-1} = \sum (z - D)^{-1} [E (z - D)^{-1}]^l$ , which is norm convergent in a neighbourhood of  $\Gamma_j$  provided  $\|E\| < \epsilon$  is sufficiently small. Taking  $j = 1$  immediately gives (93).

Elementary manipulations show that (94) follows from

$$\mu_2 = \rho_0^2 \frac{\bar{\alpha}_1}{(\bar{\alpha}_0 + \bar{\alpha}_1 \alpha_0)^2} + \rho_0^2 \sum \rho_j \tilde{G}_j(\alpha, \bar{\alpha}, \rho), \quad (96)$$

for some analytic functions  $\tilde{G}_j$ . To prove this, we simply need to apply the resolvent formula one more time. From (95) and  $D_{22} = -\alpha_0 \bar{\alpha}_1$ ,

$$\mu_2 = \frac{\rho_0^2 \bar{\alpha}_1}{(\alpha_0 \bar{\alpha}_1 + \bar{\alpha}_0)^2} + \rho_0^2 \oint_{\Gamma_2} \frac{\langle \alpha_1 e_2 + \rho_1 e_3 | (z - D - E)^{-1} E e_2 \rangle dz}{2\pi i (z + \alpha_0 \bar{\alpha}_1)(z - \bar{\alpha}_0)^2}.$$

As  $E e_2 = \rho_0 \bar{\alpha}_1 e_1 + \rho_1 \bar{\alpha}_2 e_3 + \rho_1 \rho_2 e_4$ , an additional factor of  $\rho_0$  or  $\rho_1$  can be extracted from the integral to give (96).  $\square$

**Lemma 8.3.** *Under the hypotheses of Lemma 8.2, there exist analytic functions  $H_j$  of the  $3(n-1)$  variables  $\alpha_k, \bar{\alpha}_k, \rho_k$ ,  $0 \leq k \leq n-2$ , so that*

$$\log\left[\frac{\mu_3}{\mu_1}\right] = \log[\rho_0^2 \rho_1^2] + \log\left[\frac{-\alpha_1 \bar{\alpha}_2^2}{\Phi_2(-\alpha_1 \bar{\alpha}_2)^2}\right] + \sum \rho_j H_j(\alpha, \bar{\alpha}, \rho), \quad (97)$$

where (consistent with the notation of Section 2)

$$\Phi_2(z) = z^2 + (\alpha_0 \bar{\alpha}_1 - \bar{\alpha}_0)z - \bar{\alpha}_1$$

and  $\mu_j$  is the mass associated to the eigenvalue nearest  $\zeta_j$ .

*Proof.* In view of (93), it suffices to show that

$$\mu_3 = \rho_0^2 \rho_1^2 \left[ \frac{-\alpha_1 \bar{\alpha}_2^2}{\Phi_2(-\alpha_1 \bar{\alpha}_2)^2} + \sum \rho_j \tilde{H}_j(\alpha, \bar{\alpha}, \rho) \right], \quad (98)$$

for some analytic functions  $\tilde{H}_j$ . This result does not follow by using the resolvent formula once more in the previous proof; instead, we will use a block decomposition.

Let us write  $\mathcal{C} = \begin{bmatrix} A & B \\ C & D+E \end{bmatrix}$  where

$$A = \begin{bmatrix} \bar{\alpha}_0 & \rho_0 \bar{\alpha}_1 \\ \rho_0 & -\alpha_0 \bar{\alpha}_1 \end{bmatrix}, \quad B = \begin{bmatrix} \rho_0 \rho_1 & 0 & \cdots \\ -\alpha_0 \rho_1 & 0 & \cdots \end{bmatrix}, \quad C^T = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ \rho_1 \bar{\alpha}_2 & \rho_1 \rho_2 & 0 & \cdots \end{bmatrix},$$

and  $D + E$  is the splitting of the remaining block into its diagonal and off-diagonal parts. By the well known formulae for block matrix inversion, the top left  $2 \times 2$  block of  $(z - \mathcal{C})^{-1}$  is given by

$$(z - A)^{-1} + (z - A)^{-1} B [z - D - E - C(z - A)^{-1} B]^{-1} C (z - A)^{-1},$$

which is relevant because as in the previous lemma,

$$\mu_3 = \frac{1}{2\pi i} \oint_{\Gamma_3} \langle e_1 | (z - \mathcal{C})^{-1} e_1 \rangle dz.$$

For  $\epsilon$  sufficiently small,  $A$  has no eigenvalues inside  $\Gamma_3$  and so the first summand above does not contribute.

To continue, we calculate

$$(z - A)^{-1} = \frac{1}{\Phi_2(z)} \begin{bmatrix} z + \alpha_0 \bar{\alpha}_1 & \rho_0 \bar{\alpha}_1 \\ \rho_0 & z - \bar{\alpha}_0 \end{bmatrix},$$

where  $\det(z - A) = \Phi_2(z) = z^2 + (\alpha_0 \bar{\alpha}_1 - \bar{\alpha}_0)z - \bar{\alpha}_1$ . This immediately implies that

$$C(z - A)^{-1} e_1 = \frac{\rho_0 \rho_1}{\Phi_2(z)} [\bar{\alpha}_2 e_1 + \rho_2 e_2]$$

and

$$e_1^T (z - A)^{-1} B = \frac{z \rho_0 \rho_1}{\Phi_2(z)} e_1^T.$$

Putting all these formulae together gives

$$\mu_3 = \rho_0^2 \rho_1^2 \oint_{\Gamma_3} \langle e_1 | [z - D - E - C(z - A)^{-1} B]^{-1} (\bar{\alpha}_2 e_1 + \rho_2 e_2) \rangle \frac{z dz}{2\pi i \Phi_2(z)^2}.$$

From here we proceed as usual: let  $\tilde{E} = E + C(z - A)^{-1} B$  and apply the resolvent identity  $(z - D - \tilde{E})^{-1} = (z - D)^{-1} + (z - D)^{-1} \tilde{E} (z - D - \tilde{E})^{-1}$  in the equation above. As  $D_{11} = \mathcal{C}_{33} = -\alpha_1 \bar{\alpha}_2$ , the first term gives rise to

$$\rho_0^2 \rho_1^2 \bar{\alpha}_2 \oint_{\Gamma_3} \langle e_1 | (z - D)^{-1} e_1 \rangle \frac{z dz}{2\pi i \Phi_2(z)^2} = \rho_0^2 \rho_1^2 \frac{-\alpha_1 \bar{\alpha}_2^2}{\Phi_2(-\alpha_1 \bar{\alpha}_2)^2},$$

which is exactly the dominant term in (98). The remaining term can be rewritten

$$\rho_0^2 \rho_1^2 \oint_{\Gamma_3} \langle e_1 | \tilde{E} (z - D - \tilde{E})^{-1} (\bar{\alpha}_2 e_1 + \rho_2 e_2) \rangle \frac{z dz}{2\pi i (z + \alpha_1 \bar{\alpha}_2) \Phi_2(z)^2}$$

which is of the form  $\rho_0^2 \rho_1^2 \sum \rho_j \tilde{H}_j(\alpha, \bar{\alpha}, \rho)$ , because every entry of  $\tilde{E}$  contains at least one factor  $\rho_j$ .  $\square$

**Proposition 8.4.** *For any labelling of the eigenvalues,*

$$\{\log[\mu_2/\mu_1], \log[\mu_3/\mu_1]\} = 2 \cot\left(\frac{\theta_1 - \theta_2}{2}\right) + 2 \cot\left(\frac{\theta_2 - \theta_3}{2}\right) + 2 \cot\left(\frac{\theta_3 - \theta_1}{2}\right) \quad (99)$$

*in the Gelfand–Dikij (or Ablowitz–Ladik) bracket.*

*Proof.* We can choose a polynomial  $f$  so that  $F(z_j)$  are all distinct and

$$F(z_1) > F(z_2) > F(z_3)$$

are the three largest values. (We remind the reader that  $F(z) = 2 \operatorname{Re} z f'(z)$ .) By Lemma 8.1,  $\{\frac{1}{2} \log[\mu_2/\mu_1], \frac{1}{2} \log[\mu_3/\mu_1]\}$  is conserved under the flow generated by  $\phi(\mathcal{C}) = \operatorname{Im} \operatorname{tr} f(\mathcal{C})$ . The key idea is to use results from Section 7 to evaluate this bracket as  $t \rightarrow \infty$ .

For  $t$  sufficiently large, we can apply Lemmas 8.2 and 8.3 because  $\mathcal{C}(t)$  converges to a diagonal matrix under the  $\phi$ -flow. Thus we need only show that

$$\left\{ \log[\rho_0^2] + \log\left[\frac{\bar{\alpha}_1}{(\bar{\alpha}_0 + \bar{\alpha}_1 \alpha_0)^2}\right] + \sum \rho_j G_j, \log[\rho_0^2 \rho_1^2] + \log\left[\frac{-\alpha_1 \bar{\alpha}_2^2}{\Phi_2(-\alpha_1 \bar{\alpha}_2)^2}\right] + \sum \rho_j H_j \right\}$$

converges to the right-hand side of (99) as  $t \rightarrow \infty$ . To do this we employ linearity, the Leibnitz rule, and the following consequences of (42):

$$\begin{aligned} \{\log[\rho_m], \rho_m h(\alpha_m)\} &= \{\rho_m, h(\alpha_m)\} = -i \rho_m h'(\alpha_m) \rightarrow 0 \\ \{\log[\rho_m], \rho_m h(\bar{\alpha}_m)\} &= \{\rho_m, h(\bar{\alpha}_m)\} = i \rho_m h'(\bar{\alpha}_m) \rightarrow 0 \\ \{h(\alpha_m), g(\bar{\alpha}_m)\} &= -2i \rho_m^2 h'(\alpha_m), g'(\bar{\alpha}_m) \rightarrow 0 \end{aligned}$$

for any analytic functions  $h$  and  $g$ . Of the nine terms in the original bracket, this kills off all but two:

$$\left\{ \log[\rho_0^2], \log\left[\frac{-\alpha_1 \bar{\alpha}_2^2}{\Phi_2(-\alpha_1 \bar{\alpha}_2)^2}\right] \right\} + \left\{ \log\left[\frac{\bar{\alpha}_1}{(\bar{\alpha}_0 + \bar{\alpha}_1 \alpha_0)^2}\right], \log[\rho_0^2 \rho_1^2] \right\}.$$

These can be computed using (42). The result is

$$-4i \frac{\alpha_1 \bar{\alpha}_2 (\alpha_0 \bar{\alpha}_1 + \bar{\alpha}_0)}{\Phi_2(-\alpha_1 \bar{\alpha}_2)} - 2i \frac{\alpha_0 \bar{\alpha}_1 - \bar{\alpha}_0}{\alpha_0 \bar{\alpha}_1 + \bar{\alpha}_0}.$$

From Theorem 7.3,

$$\begin{aligned} \bar{\alpha}_0 &\rightarrow e^{i\theta_1}, & -\alpha_0 \bar{\alpha}_1 &\rightarrow e^{i\theta_2}, \\ -\alpha_1 \bar{\alpha}_2 &\rightarrow e^{i\theta_3}, & \Phi_2(z) &\rightarrow (z - e^{i\theta_1})(z - e^{i\theta_2}) \end{aligned}$$

as  $t \rightarrow \infty$ . By using these formulae,

$$\left\{ \log[\mu_2/\mu_1], \log[\mu_3/\mu_1] \right\} = 4i \frac{e^{i\theta_3} (e^{i\theta_1} - e^{i\theta_2})}{(e^{i\theta_3} - e^{i\theta_1})(e^{i\theta_3} - e^{i\theta_2})} + 2i \frac{e^{i\theta_1} + e^{i\theta_2}}{e^{i\theta_1} - e^{i\theta_2}}.$$

The rest is trigonometry.  $\square$

While the proposition is stated with fixed indices  $(1, 2, 3)$ , the arbitrariness of the labelling immediately gives the result for any triple of distinct indices. The result is

$$\begin{aligned} \Psi_{q,r,s} &:= \{\log[\mu_q/\mu_s], \log[\mu_r/\mu_s]\} \\ &= 2 \cot\left(\frac{\theta_q - \theta_r}{2}\right) + 2 \cot\left(\frac{\theta_r - \theta_s}{2}\right) + 2 \cot\left(\frac{\theta_s - \theta_q}{2}\right). \end{aligned} \quad (100)$$

Note that  $\Psi_{q,r,s}$  is invariant under cyclic permutations of  $(q, r, s)$  and odd under transpositions. We adopt the convention that  $\Psi_{q,r,s}$  vanishes if two indices coincide.

We will now give formulae for the functions  $f_l$  discussed at the beginning of this section. Let  $e^{i\eta} = \det(\mathcal{C})$  and  $G(x) = 2 \int_0^1 t \cot(xt/2) dt$ . Then

$$f_l(\theta_1, \dots, \theta_{n-1}) = \sum_{\substack{k=1 \\ k \neq l}}^{n-1} \theta_k [G(\theta_k - \theta_l) + G(\theta_l + \eta - \theta_n) + G(\theta_n - \eta - \theta_k)] \quad (101)$$

solve the system (92). Given any scalar function  $g$ , another solution would be  $f_l + \frac{\partial g}{\partial \theta_l}$ ; however, we have found no choice of  $g$  that leads to a simpler formula. It is possible to compute the integral defining  $G$ , but this is not an elementary function—it involves dilogarithms. Verifying that (101) do indeed solve (92) is unenlightening; the key observations are that  $G$  is odd,  $G'$  is even, and that  $xG'(x) + 2G(x) = 2 \cot(x/2)$ .

While it is not immediately obvious, Proposition 8.4 does permit us to determine  $\{\mu_q, \mu_r\}$ .

**Proposition 8.5.** *With  $\Psi$  as in (100),*

$$\{\log[\mu_q], \log[\mu_r]\} = \sum_{k=1}^n \mu_k \Psi_{q,r,k}. \quad (102)$$

*Proof.* Let us write  $M_{kl} = \{\log[\mu_k], \log[\mu_l]\}$ . Regarding  $\mu_1$  as a function of the remaining masses, the Leibnitz rule gives

$$\Psi_{k,l,1} = M_{kl} + \sum_{p=2}^n \frac{\mu_p}{\mu_1} M_{pl} + \sum_{p=2}^n \frac{\mu_p}{\mu_1} M_{kp}.$$

Multiplying through by  $\mu_k$  and summing over  $k = 2, \dots, n$  gives

$$\sum_{k=2}^n \mu_k \Psi_{k,l,1} = \sum_{k=2}^n \mu_k M_{kl} + (1 - \mu_1) \sum_{p=2}^n \frac{\mu_p}{\mu_1} M_{pl} + \sum_{k,p=2}^n \frac{\mu_p \mu_k}{\mu_1} M_{kp}.$$

The last (double) sum vanishes because  $M_{kp}$  is antisymmetric and so we can deduce

$$\sum_{k=2}^n \mu_k \Psi_{k,l,1} = \frac{1}{\mu_1} \sum_{p=2}^n \mu_p M_{pl} = -\{\log[\mu_1], \log[\mu_l]\}.$$

The result now follows by relabelling indices.  $\square$

In an earlier paper, [21], we determined the Jacobian of the map

$$(\theta_1, \mu_1, \dots, \theta_{n-1}, \mu_{n-1}, \theta_n) \mapsto (u_0, v_0, \dots, u_{n-1}, v_{n-2}, \phi) \quad (103)$$

where  $\alpha_k = u_k + iv_k$  and  $\alpha_{n-1} = e^{i\phi}$ . The analogous result for Jacobi matrices is due to Dumitriu and Edelman, [13], which served as our guide.

We asked for a simpler, more direct derivation of the Jacobian. A recent preprint of Forrester and Rains, [15], gives a very direct solution. Percy Deift showed us a derivation of the Jacobi matrix result using the symplectic structure naturally associated to the Toda lattice. We would like to close with the corresponding proof for CMV matrices. The key idea is the following: As we can write the underlying symplectic form in either set of variables, we can view (103) as a symplectomorphism between two concrete symplectic manifolds. In particular, it must preserve the Liouville volume.

**Corollary 8.6.** *The Jacobian of the change of variables (103) is given by*

$$\det \left[ \frac{\partial(u_0, v_0, \dots, \phi)}{\partial(\theta_1, \mu_1, \dots, \theta_n)} \right] = -2^{1-n} \frac{\rho_0^2 \cdots \rho_{n-2}^2}{\mu_1 \cdots \mu_n}.$$



*Proof.* The whole proof amounts to writing the symplectic volume in each set of coordinates. In terms of the Verblunsky coefficients, the answer can be read off from (41):

$$\omega \wedge \cdots \wedge \omega = \left[ \prod_{k=0}^{n-2} \rho_k^{-2} \right] du_0 \wedge dv_0 \wedge \cdots \wedge du_{n-2} \wedge dv_{n-2}. \quad (104)$$

To determine the volume in terms of  $\theta_k$  and  $\mu_k$  we begin by using the Darboux coordinates  $\theta_k, \phi_k$  constructed in this section:

$$\begin{aligned} \omega \wedge \cdots \wedge \omega &= d\theta_1 \wedge d\phi_1 \wedge \cdots \wedge d\theta_{n-1} \wedge d\phi_{n-1} \\ &= \det(J) d\theta_1 \wedge d\mu_1 \wedge \cdots \wedge d\theta_{n-1} \wedge d\mu_{n-1} \end{aligned}$$

where  $J$  is the  $2(n-1) \times 2(n-1)$  block matrix

$$J = \begin{bmatrix} \text{Id} & 0 \\ \frac{\partial f_k}{\partial \theta_l} & K \end{bmatrix} \quad \text{with} \quad K_{kl} = \frac{\partial}{\partial \mu_l} \frac{1}{2} \log[\mu_k/\mu_n] = \frac{1}{2} \mu_k^{-1} \delta_{kl} + \frac{1}{2} \mu_n^{-1}.$$

It is not difficult to evaluate  $\det(K)$  by Gaussian elimination: first subtract the first column from all other columns, then for each  $2 \leq k \leq n-1$ , subtract  $\mu_k/\mu_n$  times the  $k$ th column from the first. The result is an upper triangular matrix whose diagonal is  $[\frac{1}{2\mu_1} + \frac{1}{2\mu_n} + \sum_{k=2}^{n-1} \frac{\mu_k}{2\mu_1\mu_n}, \frac{1}{2\mu_2}, \dots, \frac{1}{2\mu_{n-1}}]$ . Recalling that  $\mu_1 + \cdots + \mu_n = 1$ , the product simplifies to give

$$\omega \wedge \cdots \wedge \omega = 2^{1-n} \left[ \prod_{k=1}^n \mu_k^{-1} \right] d\theta_1 \wedge d\mu_1 \wedge \cdots \wedge d\theta_{n-1} \wedge d\mu_{n-1}$$

In order to add the last coordinate, recall that  $(-1)^{n-1} \bar{\alpha}_{n-1} = \det(\mathcal{C}) = \prod e^{i\theta_k}$ . Therefore  $-d\phi = d\theta_1 + \cdots + d\theta_n$  and so

$$\begin{aligned} & \left[ \prod_{k=0}^{n-2} \rho_k^{-2} \right] du_0 \wedge dv_0 \wedge \cdots \wedge du_{n-2} \wedge dv_{n-2} \wedge d\phi \\ &= -2^{1-n} \left[ \prod_{k=1}^n \mu_k^{-1} \right] d\theta_1 \wedge d\mu_1 \wedge \cdots \wedge d\theta_{n-1} \wedge d\mu_{n-1} \wedge d\theta_n, \end{aligned}$$

which implies the claim by the change of variables formula of differential geometry.  $\square$

**Remark 8.7.** By placing the result at the end of such a long paper, we have perhaps given the impression that this derivation of the Jacobian must be the most complicated of all. However, all that we really used was the commutativity of the eigenvalues and (71). If one remains single-minded, these can be derived very quickly from either definition of the bracket (19) or (42).

## REFERENCES

- [1] M. J. Ablowitz and J. F. Ladik, Nonlinear differential-difference equations, *J. Math. Phys.* **16** (1975), 598–603.
- [2] M. J. Ablowitz and J. F. Ladik, Nonlinear differential-difference equations and Fourier analysis, *J. Math. Phys.* **17** (1976), 1011–1018.
- [3] M. J. Ablowitz, B. Prinari, and A. D. Trubach, *Discrete and Continuous Nonlinear Schrödinger Systems*, London Mathematical Society Lecture Note Series, vol. 302, Cambridge University Press, Cambridge, 2004.

- [4] M. Adler, On a trace functional for formal pseudo differential operators and the symplectic structure of the Korteweg-de Vries type equations. *Invent. Math.* **50** (1979), 219–248.
- [5] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, Mathematical aspects of classical and celestial mechanics. *Dynamical systems, III*, Encycl. Math. Sci., **3**, Springer, Berlin, 1993.
- [6] M. J. Cantero, L. Moral, and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.* **362** (2003), 29–56.
- [7] M. J. Cantero, L. Moral, and L. Velázquez, Minimal representations of unitary operators and orthogonal polynomials on the unit circle. Preprint [math.CA/0405246](#).
- [8] P. Deift, T. Nanda, and C. Tomei, Ordinary differential equations and the symmetric eigenvalue problem. *SIAM J. Numer. Anal.* **20** (1983), 1–22.
- [9] P. Deift, L. C. Li, and C. Tomei, Toda flows with infinitely many variables. *J. Funct. Anal.* **64** (1985), no. 3, 358–402.
- [10] P. Deift, L. C. Li, T. Nanda, and C. Tomei, The Toda flow on a generic orbit is integrable. *Comm. Pure Appl. Math.* **39** (1986), 183–232.
- [11] P. Deift, L. C. Li, T. Nanda, and C. Tomei, Symplectic aspects of some eigenvalue algorithms. *Important developments in soliton theory*, Springer, Berlin, 1993, 511–536.
- [12] P. Deift, Integrable Hamiltonian systems. *Dynamical systems and probabilistic methods in partial differential equations*, Lectures in Appl. Math., **31**, Amer. Math. Soc., Providence, RI, 1996, 103–138.
- [13] I. Dumitriu and A. Edelman, Matrix models for beta ensembles. *J. Math. Phys.* **43** (2002), 5830–5847.
- [14] H. Flaschka, The Toda lattice. I. Existence of integrals. *Phys. Rev. B* **9** (1974), 1924–1925.
- [15] P. J. Forrester and E. M. Rains, Jacobians and rank 1 perturbations relating to unitary Hessenberg matrices. Preprint [math.PR/0505552](#).
- [16] F. R. Gantmacher, *The theory of matrices. Vol. 1*. AMS Chelsea Publishing, Providence, RI, 1998.
- [17] I. M. Gelfand and L. A. Dikij, A family of Hamiltonian structures related to nonlinear integrable differential equations. *Izrail M. Gelfand Collected Papers, vol. 1*, Springer-Verlag 1987, 625–646.
- [18] Ya. L. Geronimus, *Orthogonal Polynomials*. Consultants Bureau, New York, 1961.
- [19] M. Hénon, Integrals of the Toda lattice. *Phys. Rev. B* **9** (1974), 1921–1923.
- [20] A. S. Householder, *The Theory of Matrices in Numerical Analysis*. Blaisdell Publishing Company, New York, 1965.
- [21] R. Killip and I. Nenciu, Matrix models for circular ensembles. *Int. Math. Res. Not.* **50** (2004), 2665–2701.
- [22] B. Kostant, The solution to a generalized Toda lattice and representation theory. *Adv. in Math.* **34** (1979), 195–338.
- [23] L.-C. Li, Some remarks on CMV matrices and dressing orbits. Preprint [math.SG/0507299](#).
- [24] J. Moser, Finitely many mass points on the line under the influence of an exponential potential—an integrable system. *Dynamical systems, theory and applications.*, pp. 467–497. Lecture Notes in Phys., Vol. 38, Springer, Berlin, 1975.
- [25] A. Mukaihira and Y. Nakamura, Schur flow for orthogonal polynomials on the unit circle and its integrable discretization. *J. Comput. Appl. Math.* **139** (2002), 75–94.
- [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] M. A. Olshanetsky, A. M. Perelomov, A. G. Reyman and M. A. Semenov-Tian-Shansky, Integrable systems. II. *Dynamical systems. VII*. Encycl. Math. Sci. **16**, 83–259.
- [28] A. M. Perelomov, *Integrable systems of classical mechanics and Lie algebras. Vol. I*. Translated from the Russian by A. G. Reyman. Birkhäuser Verlag, Basel, 1990.
- [29] B. Simon, *Orthogonal Polynomials on the unit circle, part 1: classical theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.
- [30] G. Szegő, *Orthogonal polynomials*. AMS Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, Rhode Island, 1975.
- [31] C. Thiele and T. Tao, The Dirac non-linear Fourier transform on  $\ell^2(\mathbb{Z})$ . Unpublished.
- [32] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, Cambridge University Press, Cambridge, 1992.

ROWAN KILLIP, UCLA MATHEMATICS DEPARTMENT, BOX 951555, LOS ANGELES, CA 90095  
*E-mail address:* [killip@math.ucla.edu](mailto:killip@math.ucla.edu)

IRINA NENCIU, MATHEMATICS 253-37, CALTECH, PASADENA, CA 91125  
*E-mail address:* [nenciu@caltech.edu](mailto:nenciu@caltech.edu)