

CMV matrices in random matrix theory and integrable systems: a survey

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2006 J. Phys. A: Math. Gen. 39 8811

(<http://iopscience.iop.org/0305-4470/39/28/S04>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.105

The article was downloaded on 03/06/2010 at 04:40

Please note that [terms and conditions apply](#).

CMV matrices in random matrix theory and integrable systems: a survey

Irina Nenciu

Courant Institute, 251 Mercer St, New York, NY 10012, USA

E-mail: nenciu@cims.nyu.edu

Received 11 October 2005

Published 27 June 2006

Online at stacks.iop.org/JPhysA/39/8811

Abstract

We present a survey of recent results concerning a remarkable class of unitary matrices, the CMV matrices. We are particularly interested in the role they play in the theory of random matrices and integrable systems. Throughout the paper we also emphasize the analogies and connections to Jacobi matrices.

PACS numbers: 02.10.Yn, 02.30.Ik

1. Introduction

From many points of view, Jacobi matrices play a central role among all Hermitian matrices. We wish to discuss a class of unitary matrices recently introduced by Cantero, Moral and Velázquez [[CanMorVel1](#)] which play a similar role among unitary matrices. While this analogy may be extended in many directions, we investigate here their occurrence in the theory of random matrices and integrable systems.

By Jacobi matrix, we mean an $n \times n$ tri-diagonal matrix

$$J = \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 & \ddots & & \\ & \ddots & \ddots & a_{n-1} & \\ & & & a_{n-1} & b_n \end{bmatrix} \quad (1.1)$$

with $a_j > 0$, $b_j \in \mathbb{R}$.

Given coefficients $\alpha_0, \dots, \alpha_{n-2} \in \mathbb{D}$, the unit disc in \mathbb{C} and $\alpha_{n-1} \in S^1$, let $\rho_k = \sqrt{1 - |\alpha_k|^2}$, and define 2×2 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 1×1 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 with the coefficients $\alpha_0, \dots, \alpha_{n-1}$ is $\mathcal{C} = \mathcal{L}\mathcal{M}$. It is a 5-diagonal unitary matrix, given by

$$\mathcal{C} = \begin{pmatrix} \bar{\alpha}_0 & \rho_0 \bar{\alpha}_1 & \rho_0 \rho_1 & 0 & 0 & 0 & \dots \\ \rho_0 & -\alpha_0 \bar{\alpha}_1 & -\alpha_0 \rho_1 & 0 & 0 & 0 & \dots \\ 0 & \rho_1 \bar{\alpha}_2 & -\alpha_1 \bar{\alpha}_2 & \rho_2 \bar{\alpha}_3 & \rho_2 \rho_3 & 0 & \dots \\ 0 & \rho_1 \rho_2 & -\alpha_1 \rho_2 & -\alpha_2 \bar{\alpha}_3 & -\alpha_2 \rho_3 & 0 & \dots \\ 0 & 0 & 0 & \rho_3 \bar{\alpha}_4 & -\alpha_3 \bar{\alpha}_4 & \rho_4 \bar{\alpha}_5 & \dots \\ 0 & 0 & 0 & \rho_3 \rho_4 & -\alpha_3 \rho_4 & -\alpha_4 \bar{\alpha}_5 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{1.2}$$

Following [Sim1], we will refer to the numbers α_k as Verblunsky coefficients. A related system of matrices was discovered independently by Tao and Thiele [TaoThi] in connection with the nonlinear 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 [Sim1, p 84].)

Both Jacobi and CMV matrices arose in the study of orthogonal polynomials, so 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 $p_n \equiv 0$.

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 with J and the vector $e_1 = [1, 0, \dots, 0]^T$. Then J represents $f(x) \mapsto xf(x)$ in the basis of orthonormal polynomials associated with $d\nu$.

To explain 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 [Sim1].

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, ϕ_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 $\Phi_k(z)$

$$\Phi_{k+1}(z) = z\Phi_k(z) - \bar{\alpha}_k \Phi_k^*(z). \tag{1.3}$$

Here, Φ_k^* denotes the reversed polynomial:

$$\Phi_k(z) = \sum_{l=0}^k c_l z^l \quad \Rightarrow \quad \Phi_k^*(z) = \sum_{l=0}^k \bar{c}_{k-l} z^l = z^k \overline{\Phi_k\left(\frac{1}{z}\right)}, \tag{1.4}$$

and α_k are recurrence coefficients. They are in fact the same as the Verblunsky coefficients which appeared above, in the definition of the CMV matrix. When $d\mu$ is supported at exactly n points, $\alpha_k \in \mathbb{D}$ for $0 \leq k \leq n - 2$, while α_{n-1} is a unimodular complex number.

The Verblunsky coefficients completely describe the measure $d\mu$: there is a one-to-one 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 a Hessenberg matrix which Simon [Sim1] 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.

Cantero, Moral and Velázquez 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, $\chi_k(z)$ ($0 \leq k \leq n - 1$), form a basis and are easily expressible in terms of the orthonormal polynomials. In this basis, the map $f(z) \mapsto zf(z)$ is represented in an especially simple form:

Theorem 1 (Cantero, Moral and Velázquez (2003)). *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 \mathcal{C} associated with the Verblunsky coefficients of the measure $d\mu$.*

The $\mathcal{C} = \mathcal{LM}$ factorization presented above 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 [Sim1] for further discussion.

The measure $d\mu$ can be reconstructed from \mathcal{C} in a manner analogous to the Jacobi case. Let $d\mu$ be the spectral measure associated with a CMV matrix, \mathcal{C} , and the vector e_1 . Then, \mathcal{C} is the CMV matrix associated with the measure $d\mu$.

Proofs of these statements can be found in [CanMorVel1] or [Sim1].

An interesting special case is when the measure $d\mu$ is symmetric with respect to complex conjugation. This occurs if and only if all Verblunsky coefficients are real. It is a famous observation of Szegő (see [Sze, section 11.5]) that the polynomials orthogonal with respect to this measure are intimately related to the polynomials orthogonal with respect to the measure $d\nu$ on $[-2, 2]$ defined by

$$\int_{S^1} f(z + z^{-1}) d\mu(z) = \int_{-2}^2 f(x) d\nu(x). \tag{1.5}$$

Moreover, the recurrence coefficients for these measures are related by the Geronimus relations:

$$\begin{aligned} b_{k+1} &= (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2} \\ a_{k+1} &= \{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})\}^{1/2}. \end{aligned} \tag{1.6}$$

The classical proof of (1.6) relies on the relation between the orthogonal polynomials associated with the two measures. An alternate proof was given by Killip and the author, [KilNen1, proposition B.3], who derive (1.6) from an explicit relation between the CMV and Jacobi matrices of the two measures $d\mu$ and $d\nu$, respectively.

2. Circular β -ensembles

In 1962, Dyson [Dys] introduced three ensembles of random unitary matrices with a view to simplifying the study of energy level behaviour in complex quantum systems. Earlier work in this direction, pioneered by Wigner, focused on ensembles of Hermitian matrices. In both of these situations, the eigenvalue distributions coincide with the Gibbs distribution for particles of Coulomb gas at inverse temperature $\beta = 1, 2, 4$ on the circle or the real line, respectively. Dumitriu and Edelman [DumEde] constructed tri-diagonal matrix models for two of the three standard examples of the Coulomb gas on the real line for *any* inverse temperature: the β -Hermite and the β -Laguerre. Killip and the author obtained the analogous result on the circle, using CMV matrices.

One of the main features of all of these models is that they are defined in terms of $O(n)$ independent random variables which appear in matrix entries via very simple combinations. Moreover, the matrices are all sparse, having $O(n)$ non-zero entries. Finally, these are not asymptotic models (see [ForRai1] and [Lip]), but exact ones. (Throughout this section, n denotes the number of particles.)

Let us elaborate: the easiest way to obtain a normalizable Gibbs measure on the real line is to add an external harmonic potential $V(x) = \frac{1}{2}x^2$. This gives rise to the probability measure

$$\mathbb{E}(f) \propto \int \cdots \int f(x_1, \dots, x_n) |\Delta(x_1, \dots, x_n)|^\beta \prod_j e^{-V(x_j)} dx_1 \cdots dx_n \quad (2.1)$$

on \mathbb{R}^n , where, as usual, Δ denotes the Vandermonde determinant. This is known as the Hermite ensemble, because of its intimate connection to the orthogonal polynomials of the same name, and when $\beta = 1, 2$ or 4 , arises as the eigenvalue distribution in the classical Gaussian ensembles of random matrix theory. Dumitriu and Edelman showed that (2.1) is the distribution of eigenvalues for a symmetric tri-diagonal matrix with independent entries (modulo symmetry). The diagonal entries have standard Gaussian distribution and the lower diagonal entries are $2^{-1/2}$ times a χ -distributed random variable with the number of degrees of freedom equal to β times the number of the row.

The second example treated by Dumitriu and Edelman is the Laguerre ensemble. In statistical circles, this is known as the Wishart ensemble, special cases of which arise in the empirical determination of the covariance matrix of a multivariate Gaussian distribution. For this ensemble, one needs to modify the distribution given in (2.1) in two ways: each particle x_j is confined to lie in $[0, \infty)$ and is subject to the external potential $V(x) = -a \log(x) + x$, where $a > -1$ is a parameter. In [DumEde] it is shown that if B is a certain $n \times n$ matrix with independent χ -distributed entries on the main and sub-diagonal (the number of degrees of freedom depends on a, β and the element in question) and zeros everywhere else, then the eigenvalues of $L = BB^T$ follow this distribution.

The third standard ensemble on the real line is the Jacobi ensemble. The distribution is as in (2.1), but now the particles are confined to lie within $[-2, 2]$ and are subject to the external potential $V(x) = -a \log(2-x) - b \log(2+x)$, where $a, b > -1$ are parameters. This corresponds to the probability measure on $[-2, 2]^n$ that is proportional to

$$|\Delta(x_1, \dots, x_n)|^\beta \prod_j (2-x_j)^a (2+x_j)^b dx_1 \cdots dx_n. \quad (2.2)$$

Dumitriu and Edelman did not give a matrix model for this ensemble, listing it as an open problem.

On the unit circle, one does not need a confining potential in order to define the Gibbs distribution for n particles of the Coulomb gas at inverse temperature β ; it is given by

$$\mathbb{E}_n^\beta(f) \propto \int \cdots \int f(e^{i\theta_1}, \dots, e^{i\theta_n}) |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi} \tag{2.3}$$

for any symmetric function f . CMV matrices realize a random matrix model with this eigenvalue distribution.

We say that a complex random variable, X , with values in the unit disc, \mathbb{D} , is Θ_ν -distributed (for $\nu > 1$) if

$$\mathbb{E}\{f(X)\} = \frac{\nu-1}{2\pi} \int \int_{\mathbb{D}} f(z) (1 - |z|^2)^{(\nu-3)/2} d^2z. \tag{2.4}$$

For $\nu \geq 2$ an integer, this has the following geometric interpretation: if ν is chosen from the unit sphere S^ν in $\mathbb{R}^{\nu+1}$ at random according to the usual surface measure, then $v_1 + iv_2$ is Θ_ν -distributed. As a continuation of this geometric picture, we shall say that X is Θ_1 -distributed if it is uniformly distributed on the unit circle in \mathbb{C} . With this definition, we have the following family of matrix models:

Theorem 2 (Killip and Nenciu (2004)). *Given $\beta > 0$, let $\alpha_k \sim \Theta_{\beta(n-k-1)+1}$ be independent random variables for $0 \leq k \leq n - 1$. Then the CMV matrix $\mathcal{C} = \mathcal{LM}$ defined by these Verblunsky coefficients gives a (sparse) matrix model for the Coulomb gas at inverse temperature β . That is, its eigenvalues are distributed according to the Gibbs distribution (2.3).*

Finally, we present a tri-diagonal matrix model for the Jacobi ensemble (2.2). The independent parameters follow a beta distribution: a real-valued random variable X is said to be beta-distributed with parameters $s, t > 0$, which we denote by $X \sim B(s, t)$, if

$$\mathbb{E}\{f(X)\} = \frac{2^{1-s-t} \Gamma(s+t)}{\Gamma(s)\Gamma(t)} \int_{-1}^1 f(x) (1-x)^{s-1} (1+x)^{t-1} dx. \tag{2.5}$$

Note that $B(\frac{\nu}{2}, \frac{\nu}{2})$ is the distribution of the first component of a random vector from the ν -sphere.

Theorem 3 (Killip and Nenciu (2004)). *Given $\beta > 0$, let $\alpha_k, 0 \leq k \leq 2n - 2$, be distributed as follows:*

$$\alpha_k \sim \begin{cases} B\left(\frac{2n-k-2}{4}\beta+a+1, \frac{2n-k-2}{4}\beta+b+1\right) & k \text{ even,} \\ B\left(\frac{2n-k-3}{4}\beta+a+b+2, \frac{2n-k-1}{4}\beta\right) & k \text{ odd.} \end{cases} \tag{2.6}$$

Let $\alpha_{2n-1} = \alpha_{-1} = -1$ and define

$$b_{k+1} = (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2} \tag{2.7}$$

$$a_{k+1} = \{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})\}^{1/2} \tag{2.8}$$

for $0 \leq k \leq n - 1$; then the eigenvalues of the tri-diagonal matrix

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

are distributed according to the Jacobi ensemble (2.2).

Let us note that one of the essential ideas behind the proof of this theorem is to construct the distributions on Verblunsky coefficients and use the Geronimus relations (1.6) to transfer them to the Jacobi setting.

3. The Ablowitz–Ladik system: structure

By analogy with the connection between the Toda lattice and Jacobi matrices, we investigate the existence and properties of an integrable system related to orthogonal polynomials on the unit circle and CMV matrices. The main evolution of the system is defocusing Ablowitz–Ladik (also known as the integrable discrete nonlinear Schrödinger equation). An excellent elementary introduction into integrable systems is given in [Dei]; [BabBerTal] is a detailed presentation of many aspects of the subject. For the more advanced notions that we shall use (especially regarding Lie and Poisson–Lie structures) we refer to [OlsPerReySem].

The celebrated Toda lattice is the central example of a discrete integrable Hamiltonian system. It models a one-dimensional chain of particles with exponential nearest neighbour interactions and it was introduced by Morikazu Toda [Tod] to explain the findings of the famous 1955 Los Alamos computer experiment of Fermi, Pasta and Ulam [FerPasUla]. Complete integrability of the system was proved by Flaschka in 1974 [Fla] by introducing a change of variables that allowed him to set the system in Lax pair form. Indeed, if Flaschka's variables a_k and b_k are set to be the entries of a Jacobi matrix, J , then the evolution of the system is given by the equation

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

with

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

Thus, the Toda lattice is intrinsically connected to Jacobi matrices and orthogonal polynomials on the real line \mathbb{R} .

In [Sim2, section 11.11], Barry Simon and the author found that the integrable system naturally related to the theory of orthogonal polynomials on the unit circle is the defocusing Ablowitz–Ladik system. It was introduced in 1975–1976 by Ablowitz and Ladik [AbiLad1, AbiLad2] as a space discretization of the cubic nonlinear Schrödinger equation (NLS). It reads

$$-i\dot{\beta}_n = \rho_n^2(\beta_{n+1} + \beta_{n-1}) - 2\beta_n, \quad (3.1)$$

where $\beta = \{\beta_n\} \subset \mathbb{D}$ is a sequence of complex numbers inside the unit disc and

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

The analogy with the continuous NLS becomes transparent if we rewrite (3.1) as

$$-i\dot{\beta}_n = \beta_{n+1} - 2\beta_n + \beta_{n-1} - |\beta_n|^2(\beta_{n+1} + \beta_{n-1}).$$

From this point onwards, \dot{f} denotes the time derivative of the function f .

If we change variables to $\alpha_k(t) = e^{2it} \beta_k(t)$, this system becomes

$$-i\dot{\alpha}_k = \rho_k^2(\alpha_{k+1} + \alpha_{k-1}), \quad (3.2)$$

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

Nenciu and Simon [Sim2, section 11.11] found that if periodic Verblunsky coefficients are allowed to evolve according to Hamiltonians in the Ablowitz–Ladik hierarchy, then their movement takes place on tori which are isospectral with respect to the CMV matrix. The

symplectic form they introduced extends naturally to the case of finite Verblunsky coefficients, which we present here. In this setting, it is given by

$$\{f, g\}_{AL} = \sum_{j=0}^{n-2} \rho_j^2 \left[\frac{\partial f}{\partial u_j} \frac{\partial g}{\partial v_j} - \frac{\partial f}{\partial v_j} \frac{\partial g}{\partial u_j} \right] \tag{3.3}$$

$$= 2i \sum_{j=0}^{n-2} \rho_j^2 \left[\frac{\partial f}{\partial \bar{\alpha}_j} \frac{\partial g}{\partial \alpha_j} - \frac{\partial f}{\partial \alpha_j} \frac{\partial g}{\partial \bar{\alpha}_j} \right] \tag{3.4}$$

for any smooth functions $f, g : \mathbb{D}^{n-1} \rightarrow \mathbb{R}$, 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 bracket the Ablowitz–Ladik bracket. Note that this differs by a factor of 2 from that used in [Sim2] and [Nen]; we will adjust the results 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 $m \geq 1$. These form the Ablowitz–Ladik hierarchy. Indeed, the evolution of the Verblunsky coefficients under the flow generated by $\text{Re}(K_1)$ is the Ablowitz–Ladik evolution:

$$\{\alpha_j, \text{Re}(K_1)\}_{AL} = \dot{\alpha}_j = i\rho_j^2(\alpha_{j-1} + \alpha_{j+1}).$$

The evolution equations for the CMV matrices under these flows in the Ablowitz–Ladik bracket were found by Nenciu [Nen]:

Theorem 4 (Nenciu (2005)). *The Lax pairs for the m th Hamiltonian of the finite defocusing Ablowitz–Ladik system are given by*

$$\{\mathcal{C}, \text{Re}(K_m)\}_{AL} = [\mathcal{C}, i\mathcal{C}_+^m + i(\mathcal{C}_+^m)^*] \tag{3.5}$$

and

$$\{\mathcal{C}, \text{Im}(K_m)\}_{AL} = [\mathcal{C}, \mathcal{C}_+^m - (\mathcal{C}_+^m)^*] \tag{3.6}$$

for all $m \geq 1$, where for a matrix \mathcal{A} we set \mathcal{A}_+ as the matrix with entries

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

The system of nonlinear differential-difference equations

$$\dot{\alpha}_n = (1 - \alpha_n^2)(\alpha_{n+1} - \alpha_{n-1}), \quad \{\alpha_j\} \subset (-1, 1), \tag{3.7}$$

is known as the discrete modified KdV equation (see [AbiLad1] and [Gek]) or the equation of the Schur flows (see [FayGek]). This system’s main interest lies in its connection to the Toda and Volterra (or Kac-van Moerbeke) lattices. More precisely, the Schur flows appear in the work of Ammar and Gragg [AmmGra] as an evolution equation on Verblunsky coefficients obtained by transferring the Toda equation via the Geronimus relations from the as and bs . As such, it is an evolution on real α s. One can then relax this condition and think of it as an evolution on complex coefficients, which preserves reality (that is, if the α s are real at time 0, then they remain real for all times).

As it turns out, this description applies to ‘half’ of the Ablowitz–Ladik flows, the ones generated by $\text{Im}(K_m)$ for all $n \geq 1$. Indeed, the evolution on the Verblunsky coefficients under these Hamiltonians is equivalent to the Lax pairs

$$\{\mathcal{C}, \text{Im}(K_m)\}_{AL} = [\mathcal{C}, (\mathcal{C}^m)_+ - ((\mathcal{C}^m)_+)^*],$$

which preserve reality of the α s. So it is natural to ask ourselves what are the corresponding evolutions of the a s and b s. Moreover, we recover (3.7) as the evolution generated by $-\text{Im}(K_1)$ under the Ablowitz–Ladik bracket (3.3), (3.4). In fact, in the finite case, all the evolutions induced by $\text{Im}(K_n)$ become, via the Geronimus relations, simple combinations of the evolutions in the Toda hierarchy (this can be deduced from the evolution of the corresponding spectral measures from [KilNen2]).

There exists an abstract way of defining Poisson brackets on Lie algebras. In particular, if \mathfrak{g} denotes the (associative) algebra of $n \times n$ complex matrices, then 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 the classical R -matrix $R : \mathfrak{g} \rightarrow \mathfrak{g}$ by $R = \pi_{\mathfrak{l}} - \pi_{\mathfrak{a}}$. The second Lie bracket can then be written as either

$$\begin{aligned} [X, Y]_R &= \frac{1}{2}[R(X), Y] + \frac{1}{2}[X, R(Y)] \quad \forall X, Y \in \mathfrak{g} \quad \text{or} \\ [L + A, L' + A']_R &= [L, L'] - [A, A'] \quad \forall L, L' \in \mathfrak{l} \quad \text{and} \quad A, A' \in \mathfrak{a}. \end{aligned} \tag{3.8}$$

Using this R -matrix, one can define several Poisson brackets on the Lie algebra \mathfrak{g} . The Lie–Poisson (or Kirillov) bracket associated with the Lie algebra $(\mathfrak{g}, [\cdot, \cdot]_R)$ is linear and it is relevant to Jacobi matrices: under the embedding $J \mapsto iJ$, the manifolds of Jacobi matrices with fixed trace are symplectic leaves. In our case, we focus on the Gelfand–Dikij (or quadratic) bracket, denoted here by $\{\cdot, \cdot\}_{\text{GD}}$. We refer to [OlsPerReySem] for the definitions and proofs of the general results of the theory.

Theorem 5 (Killip and Nenciu (2005)). *The manifold of CMV matrices with fixed determinant forms a symplectic leaf in the Poisson manifold $(\mathfrak{g}, \{\cdot, \cdot\}_{\text{GD}})$.*

The proof of this theorem relies on the fact that a unitary matrix having the same shape as a CMV matrix is exactly a CMV matrix. Moreover, one can show that the set of CMV matrices with fixed determinant is an orbit under the action of a certain group of dressing transformations (see theorem 4.7 of [KilNen2]). In a concurrent paper [Li], Li has independently derived this result. In a sense, his approach is the reverse of ours: he begins by investigating the action of certain dressing transformations, while we arrive at their existence only after studying the problem by other means.

As is natural in the theory of Poisson–Lie algebras, we focus on Hamiltonians given by $\phi(B) = \text{Im tr } f(B)$, where f is a polynomial. The general theory guarantees that the evolution

equations they generate will be expressible as Lax pairs. In our case, these Lax pairs at a CMV matrix are the same as those obtained from theorem 4 by taking the appropriate linear combinations. Given this result, the following is not surprising:

Theorem 6 (Killip and Nenciu (2005)). *For any $0 \leq k, l \leq n - 2$, the Gelfand–Dikij brackets of the Verblunsky coefficients are given by*

$$\{\alpha_k, \alpha_l\}_{\text{GD}} = 0 \quad \text{and} \quad \{\alpha_k, \bar{\alpha}_l\}_{\text{GD}} = -2i\delta_{kl}\rho_k^2. \tag{3.9}$$

That is, the restriction of the Gelfand–Dikij bracket to CMV matrices agrees with the Ablowitz–Ladik bracket when written in the coordinates given by the Verblunsky coefficients. From now on we denote both brackets by $\{\cdot, \cdot\}$.

4. The Ablowitz–Ladik system: asymptotics

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

The identification of the Ablowitz–Ladik bracket as an abstract Poisson–Lie bracket allows one to follow the same path in the unitary setting. More precisely, we begin by investigating the evolution under a Hamiltonian of the Ablowitz–Ladik hierarchy of the spectral measure $\mu = \sum \mu_j \delta_{z_j}$ associated with a CMV matrix \mathcal{C} and the vector $e_1 = [1, 0, \dots, 0]^T$:

Theorem 7 (Killip and Nenciu (2005)). *Under the flow generated by $\phi(B) = \text{Im tr}\{f(B)\}$,*

$$\partial_t(\log[\mu_j(t)]) = \{\phi, \log[\mu_j]\} = F(e^{i\theta_j}) - \sum_{l=1}^n F(e^{i\theta_l})\mu_l(t), \tag{4.1}$$

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

$$\mu_j(t) = \frac{\exp[F(e^{i\theta_j})t]\mu_j(0)}{\sum \exp[F(e^{i\theta_l})t]\mu_l(0)} \tag{4.2}$$

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

$$\{\theta_l, \frac{1}{2} \log[\mu_j/\mu_n]\} = \delta_{jl}. \tag{4.3}$$

A special case of this result (for Schur flows) has already appeared, [MukNak]. The approach used there was to begin with a special case of (4.2) and determine the induced evolution on the Verblunsky coefficients.

In defining μ_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

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n, \tag{4.4}$$

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

In what follows, we will assume that we are indeed in a situation where the ordering described above is strict. This simplifies the formulae, but is not necessary; the interested reader can find the full results in [KilNen2].

Under the flow generated by the Hamiltonian ϕ , the masses have the following asymptotics:

$$\log[\mu_k(t)] = -(\lambda_1 - \lambda_k)t + \log \left[\frac{\mu_k}{\mu_1} \right] + O(e^{-at}) \tag{4.5}$$

as $t \rightarrow \infty$. (The quantities for which we do not specify time dependence are assumed to be evaluated to $t = 0$.) In particular $\mu_1(t) \rightarrow 1$ and, if $k > 1$, then $\mu_k(t) \rightarrow 0$ exponentially fast. We now turn to the asymptotics of the Verblunsky coefficients: fix $1 \leq k \leq n - 1$. As all λ_j are distinct,

$$\alpha_{k-1}(t) = (-1)^{k-1} \bar{z}_1 \cdots \bar{z}_k [1 + \xi_{k-1} e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t})] \tag{4.6}$$

where

$$\xi_{k-1} = (z_k \bar{z}_{k+1} - 1) \frac{\mu_{k+1}}{\mu_k} \prod_{l=1}^{k-1} \left| \frac{z_{k+1} - z_l}{z_k - z_l} \right|^2 \tag{4.7}$$

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

In other words, if all λ_j are distinct, then viewed as a curve in the disc, $\alpha_{k-1}(t)$ approaches the boundary in a fixed non-tangential direction. This simply amounts to the statement that ξ_{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 ρ_{k-1} are easily deduced from (4.6)

$$\begin{aligned} \rho_{k-1}^2(t) &= -2 \operatorname{Re}(\xi_{k-1}) e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t}) \\ &= |z_{k+1} - z_k|^2 \frac{\mu_{k+1}}{\mu_k} \prod_{l=1}^{k-1} \left| \frac{z_{k+1} - z_l}{z_k - z_l} \right|^2 e^{-(\lambda_k - \lambda_{k+1})t} + O(e^{-\gamma t}). \end{aligned}$$

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.

Let us note here that when the λ_j are not all distinct, $\mathcal{C}(t)$ converges to a direct sum of CMV matrices and their adjoints. Specifically, if

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

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 in page 389 of [DeiLiTom].

Finally, our understanding of the asymptotic behaviour of the relevant quantities allows us to find canonical coordinates for the system. We know from the Lax pairs that the eigenvalues Poisson commute: $\{\theta_j, \theta_k\} = 0$, for all $1 \leq j, k \leq n$, where θ_j is the argument of the eigenvalue z_j . Moreover, the evolution of the spectral measure given in theorem 7 implies that

$$\left\{ \theta_l, \frac{1}{2} \log[\mu_j / \mu_n] \right\} = \delta_{jl}.$$

It would be natural to imagine that

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

are canonical coordinates. This is the case for the Toda lattice (as can be concluded from [DeiLiNanTom]). The same does not happen here; instead, the following holds:

Theorem 8 (Killip and Nenciu (2005)). *For any labelling of the eigenvalues,*

$$\{\log[\mu_2 / \mu_1], \log[\mu_3 / \mu_1]\} = 2 \cot \left(\frac{\theta_1 - \theta_2}{2} \right) + 2 \cot \left(\frac{\theta_2 - \theta_3}{2} \right) + 2 \cot \left(\frac{\theta_3 - \theta_1}{2} \right) \tag{4.8}$$

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

While the previous results concerning asymptotics for the Ablowitz–Ladik system are proved using the abstract (Gelfand–Dikij) form of the bracket, we were unable to do the same here. Instead note that $\{\log[\mu_j/\mu_l], \log[\mu_k/\mu_l]\}$ is constant under the flow generated by a Hamiltonian ϕ defined as above, and so one can work in a regime where the masses are already exponentially ordered. We then use perturbation theory to obtain more precise asymptotics for $\log[\mu_j/\mu_1]$ for $j = 2, 3$ in this regime and plug these into the explicit (Ablowitz–Ladik) bracket to obtain the result. It would be interesting to find a proof using only the Gelfand–Dikij bracket and to better understand the appearance of the cotangent function in formula (4.8).

In an earlier paper [KilNen1] Killip and the author considered the map

$$(\theta_1, \mu_1, \dots, \theta_{n-1}, \mu_{n-1}, \theta_n) \mapsto (u_0, v_0, \dots, u_{n-1}, v_{n-2}, \phi) \quad (4.9)$$

where $\alpha_k = u_k + iv_k$ and $\alpha_{n-1} = e^{i\phi}$. We found that

Proposition 9. *The Jacobian of the change of variables (4.9) is given by*

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

The analogous result for Jacobi matrices is due to Dumitriu and Edelman, [DumEde], which served as our guide. Both of these proofs are probabilistic in nature.

We also asked for a simpler, more direct derivation of the Jacobian. A recent preprint of Forrester and Rains [ForRai2] gives a very direct, computational solution. Percy Deift showed us a derivation of the Jacobi matrix result using the symplectic structure naturally associated with the Toda lattice. The same can be done in the circular case (see [KilNen2, corollary 8.6]). The key idea is the following: as we can write the underlying symplectic form in either set of variables, we can view (4.9) as a symplectomorphism between two concrete symplectic manifolds. In particular, it must preserve the Liouville volume, which allows one to prove proposition 9.

Acknowledgments

The author wishes to thank the organizers of the CRM 2005 Meeting on Random Matrices, Random Processes and Integrable Systems for the opportunity to attend and to speak there. She is also grateful to Percy Deift and Barry Simon for their support and helpful discussions. Much of the author’s recent work on different aspects of CMV matrices is joint with Rowan Killip.

Note added in proof. It turns out that CMV matrices had already appeared in the literature before the work of Cantero, Moral, and Valázquez; see, for example, [Wat].

References

- [AblLad1] Ablowitz M J and Ladik J F 1975 Nonlinear differential-difference equations *J. Math. Phys.* **16** 598–603
- [AblLad2] Ablowitz M J and Ladik J F 1976 Nonlinear differential-difference equations and Fourier analysis *J. Math. Phys.* **17** 1011–8
- [AmmGra] Ammar G S and Gragg W B 1994 Schur flows for orthogonal Hessenberg matrices *Hamiltonian and Gradient Flows, Algorithms and Control* (Providence, RI: American Mathematical Society) pp 27–34 (*Fields Inst. Commun.* **3**)
- [BabBerTal] Babelon O, Bernard D and Talon M 2003 Introduction to classical integrable systems *Cambridge Monographs on Mathematical Physics* (Cambridge: Cambridge University Press)

- [CanMorVel1] Cantero M J, Moral L and Velázquez L 2003 Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle *Linear Algebra Appl.* **362** 29–56
- [CanMorVel2] Cantero M J, Moral L and Velázquez L 2004 Minimal representations of unitary operators and orthogonal polynomials on the unit circle *Preprint math.CA/0405246*
- [Dei] Deift P 1996 Integrable Hamiltonian systems *Dynamical Systems and Probabilistic Methods in Partial Differential Equations (Lectures in Applied Mathematics vol 31)* (Providence, RI: American Mathematical Society) pp 103–38
- [DeiLiNanTom] Deift P, Li L-C, Nanda T and Tomei C 1986 The Toda flow on a generic orbit is integrable *Commun. Pure Appl. Math.* **39** 183–232
- [DeiLiTom] Deift P, Li L-C and Tomei C 1985 Toda flows with infinitely many variables *J. Funct. Anal.* **64** 358–402
- [DumEde] Dumitriu I and Edelman A 2002 Matrix models for beta ensembles *J. Math. Phys.* **43** 5830–47
- [Dys] Dyson F 1962 Statistical theory of the energy levels of complex systems: I, II, and III *J. Math. Phys.* **3** 140–56, 157–65 and 166–75
- [FayGek] Faybusovich L and Gekhtman M 1999 On Schur flows *J. Phys. A: Math. Gen.* **32** 4671–80
- [FerPasUla] Fermi E, Pasta J and Ulam S 1965 Studies of nonlinear problems *Collected Works of Enrico Fermi vol II* (Chicago: University of Chicago Press) pp 978–88
- [Fla] Flaschka H 1974 The Toda lattice: I. Existence of integrals *Phys. Rev. B* **9** 1924–5
- [ForRai1] Forrester P J and Rains E M 2005 Interpretations of some parameter dependent generalizations of classical matrix ensembles *Probab. Theory Relat. Fields* **131** 1–61
- [ForRai2] Forrester P J and Rains E M 2005 Jacobians and rank 1 perturbations relating to unitary Hessenberg matrices *Preprint math.PR/0505552*
- [Gek] Gekhtman M 1993 Non-Abelian nonlinear lattice equations on finite interval *J. Phys. A: Math. Gen.* **26** 6303–17
- [GelDik] Gelfand I M and Dikij L A 1987 A family of Hamiltonian structures related to nonlinear integrable differential equations *Izrail M. Gelfand Collected Papers vol 1* (Berlin: Springer) pp 625–46
- [Ger1] Geronimus Ya L 1944 On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathodory and Schur functions *Mat. Sb.* **15** 99–130 (in Russian)
- [Ger2] Geronimus Ya L 1961 *Orthogonal Polynomials* (New York: Consultants Bureau)
- [Hen] Hénon M 1974 Integrals of the Toda lattice *Phys. Rev. B* **9** 1921–23
- [KilNen1] Killip R and Nenciu I 2004 Matrix models for circular ensembles *Int. Math. Res. Not.* **50** 2665–701
- [KilNen2] Killip R and Nenciu I CMV: the unitary analogue of Jacobi matrices *Preprint (Commun. Pure Appl. Math. at press)*
- [Kos] Kostant B 1979 The solution to a generalized Toda lattice and representation theory *Adv. Math.* **34** 195–338
- [Li] Li L-C 2005 Some remarks on CMV matrices and dressing orbits *Preprint math.SG/0507299*
- [Lip] Lippert R A 2002 A matrix model for the β -Jacobi ensemble *J. Math. Phys.* **44** 4807–16
- [Mos] Moser J 1975 Finitely many mass points on the line under the influence of an exponential potential—an integrable system *Dynamical systems, Theory and Applications (Lecture Notes in Physics vol 38)* (Berlin: Springer) pp 467–97
- [MukNak] Mukaihira A and Nakamura Y 2002 Schur flow for orthogonal polynomials on the unit circle and its integrable discretization *J. Comput. Appl. Math.* **139** 75–94
- [Nen] Nenciu I 2005 Lax pairs for the Ablowitz–Ladik system via orthogonal polynomials on the unit circle *Int. Math. Res. Not.* **11** 647–86
- [OlsPerReySem] Olshanetsky M A, Perelomov A M, Reyman A G and Semenov-Tian-Shansky M A Integrable systems: II. Dynamical systems: VII *Encycl. Math. Sci.* **16** 83–259
- [Sim1] Simon B 2005 Orthogonal polynomials on the unit circle: part 1. Classical theory *AMS Colloquium Series* (Providence, RI: American Mathematical Society)
- [Sim2] Simon B 2005 Orthogonal polynomials on the unit circle: part 2. Spectral theory *AMS Colloquium Series* (Providence, RI: American Mathematical Society)
- [Sze] Szegő G 1975 Orthogonal polynomials *American Mathematical Society Colloquium Publications vol XXIII* (Providence, RI: American Mathematical Society)
- [TaoThi] Thiele C and Tao T The Dirac non-linear Fourier transform on $\ell^2(\mathbb{Z})$ (unpublished)
- [Tod] Toda M 1989 *Theory of Nonlinear Lattices (Springer Series in Solid-State Sciences 20)* 2nd edn (Berlin: Springer)
- [Wat] Watkins D S 1993 Some perspectives on the eigenvalue problem *SIAM Rev.* **35** no. 3 430–71