SCHUR FLOWS AND ORTHOGONAL POLYNOMIALS ON THE UNIT CIRCLE

LEONID GOLINSKII

Abstract. The relation between the Toda lattices and similar nonlinear chains and orthogonal polynomials on the real line has been elaborated immensely for the last decades. We examine another system of differential-difference equations known as the Schur flow, within the framework of the theory of orthogonal polynomials on the unit circle. This system can be displayed in equivalent form as the Lax equation, and the corresponding spectral measure undergoes a simple transformation. The long time behavior of the solution is also studied.

1. Introduction

In 1975 J. Moser [11, 12] suggested a method for solution of the finite Toda lattice equations (specifically, the Cauchy problem for such lattices) based on the spectral theory of finite Jacobi matrices. Later on Yu.M. Berezanskii [5] adapted this method to semi-infinite Toda lattices

$$\begin{aligned} a_n' &= a_n(b_{n+1}-b_n), \\ b_n' &= 2(a_n^2-a_{n-1}^2), \qquad n\in\mathbb{Z}_+=\{0,1,\dots\}, \quad a_{-1}=0, \end{aligned} \tag{1.1}$$

where $'$ means differentiation with respect to t, in the class of bounded real b's and positive a's with the initial data $\{b_n(0)=\overline{b_n(0)},\ a_n(0)>0\}$. The key idea is to compose a semi-infinite Jacobi matrix

$$J=J(\{a_n\},\{b_n\})=\begin{pmatrix} b_0(t) & a_0(t) & & & \\ a_0(t) & b_1(t) & a_1(t) & & \\ & a_1(t) & b_2(t) & a_2(t) & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \tag{1.2}$$

and trace the evolution of the matrix-valued function $J=J(t)$ and its spectral characteristics. It turned out that (1.1) can be paraphrased in equivalent forms in terms of J itself (the Lax equation)

$$J'(t) = [A,J]=AJ-JA, \tag{1.3}$$

$$A = \begin{pmatrix} 0 & a_0(t) & & & \\ -a_0(t) & 0 & a_1(t) & & \\ & -a_1(t) & 0 & a_2(t) & \\ & & \ddots & \ddots & \ddots \end{pmatrix} = \pi(J):=J_+-J_- \tag{1.4}$$

1991 *Mathematics Subject Classification.* 42C05, 37K10, 37K15.

Key words and phrases. Verblunsky coefficients, Lax equations, CMV matrices, modification of measures, Bessel orthogonal polynomials.

with the standard notation $X_\pm$ for the upper (lower) projection of a matrix X, as well as the corresponding spectral (orthogonality) measure $d\mu(x,t)$ which undergoes a simple modification

$$d\mu(x,t) = e^{-xt} d\mu(x,0). \tag{1.5}$$

Hence the solution of (1.1) boils down to a combination of the direct spectral problem (from $\{a_n(0), b_n(0)\}$ to $d\mu(x,0)$) at $t=0$, plus (1.5), plus the inverse spectral problem (from $d\mu(x,t)$ to $\{a_n(t), b_n(t)\}$) at $t>0$.

The theory of orthogonal polynomials on the real line plays one of the first fiddles in the performance (albeit not entering the final result directly). For instance, it furnishes a nice setting for solving the inverse spectral problem. There is a parallel theory of orthogonal polynomials on the unit circle (OPUC) which has experienced a splash of activity lately thanks to primarily Simon's disquisition [18, 19]. So the question arises naturally whether there exist nonlinear chains (so to say, the "Toda lattices for the unit circle") which can be handled by the similar method. The main goal of the present paper is to develop the "Moser–Berezanskii scheme for the unit circle" based on the spectral theory of a certain class of unitary matrices in application to a system of nonlinear differential-difference equations known as the *Schur flow*.

We begin with some basics on orthogonal polynomials on the unit circle. Given a nontrivial (i.e., not a finite combination of delta functions) probability measure μ on the unit circle $\mathbb{T}$ with the moments

$$\mu_k := \int_{\mathbb{T}} \zeta^{-k} d\mu, \qquad k \in \mathbb{Z} = \{0, \pm 1, \dots\},$$

we define the monic orthogonal polynomials $\Phi_n(z,\mu)$ (or just Φ_n if μ is understood) by

$$\int_{\mathbb{T}} \zeta^{-k} \Phi_n(\zeta) d\mu = 0, \quad k = 0, \dots, n-1; \quad \Phi_n(z) = z^n + l_n z^{n-1} + \dots + \Phi_n(0). \tag{1.6}$$

Clearly such system is uniquely determined and

$$\int_{\mathbb{T}} \Phi_m(\zeta) \overline{\Phi_n(\zeta)}\, d\mu = 0, \qquad m \neq n. \tag{1.7}$$

The orthonormal polynomials $\varphi_n = \kappa_n \Phi_n$, $\kappa_n > 0$ enjoy the property

$$\int_{\mathbb{T}} \varphi_m(\zeta) \overline{\varphi_n(\zeta)}\, d\mu = \delta_{mn}.$$

A key role throughout the whole OPUC theory is played by the sequences of complex numbers $\{\alpha_n\}_{n\ge 0}$, $|\alpha_n| < 1$,

$$\alpha_n = \alpha_n(\mu) = -\overline{\Phi_{n+1}(0)}, \quad n \in \mathbb{Z}_+, \quad \alpha_{-1} := -1, \tag{1.8}$$

known as the *Verblunsky coefficients* or parameters of OPUC system. Firstly, due to the celebrated Verblunsky theorem, there is one-one correspondence between the class $\mathcal{P}$ of all nontrivial probability measures on $\mathbb{T}$ and the set $\mathbb{D}^\infty$, so each sequence of complex numbers $\{\gamma_n\}_{n\ge 0}$ from the open unit disk $\mathbb{D}$ comes up as a system of parameters for uniquely determined measure $\mu \in \mathcal{P}$. Secondly, Verblunsky coefficients (1.8) enter the Szegő recurrence relations given in the vector form by

$$\begin{bmatrix} \Phi_{n+1}(z) \\ \Phi^*_{n+1}(z) \end{bmatrix} = T_n(z) \begin{bmatrix} \Phi_n(z) \\ \Phi^*_n(z) \end{bmatrix}, \qquad T_n(z) = \begin{pmatrix} z & -\overline{\alpha}_n \\ -\alpha_n z & 1 \end{pmatrix} \tag{1.9}$$

is called the Szegő matrix, and so

$$\begin{bmatrix} \Phi_{n+1}(z) \\ \Phi^*_{n+1}(z) \end{bmatrix} = \mathcal{T}_n(z) \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \mathcal{T}_n(z) = T_n(z)T_{n-1}(z)\dots T_0(z)$$

is the transfer matrix. So both monic orthogonal and orthonormal polynomials are completely determined by the sequence α_n, the latter because of the equality

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

To complete with the basic properties of OPUC let us mention explicit determinant formulae for both monic polynomials and Verblunsky coefficients in terms of the moments of the orthogonality measure:

$$\Phi_n(z) = \frac{1}{D_n}\begin{vmatrix} \mu_0 & \mu_{-1} & \dots & \mu_{-n} \\ \mu_1 & \mu_0 & \dots & \mu_{-n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_{n-2} & \dots & \mu_{-1} \\ 1 & z & \dots & z^n \end{vmatrix}, \quad D_{n+1} := \det \|\mu_{k-j}\|_{k,j=0}^n,$$

$$\Phi_n(0) = -\overline{\alpha}_{n-1} = \frac{(-1)^n}{D_n}\begin{vmatrix} \mu_{-1} & \dots & \mu_{-n} \\ \mu_0 & \dots & \mu_{-n+1} \\ \vdots & \vdots & \vdots \\ \mu_{n-2} & \dots & \mu_{-1} \end{vmatrix}. \tag{1.10}$$

One of the most interesting developments in the theory of OPUC in recent years is the discovery by Cantero, Moral, and Velázquez [6] of a matrix realization for multiplication by ζ on $L^2(\mathbb{T},\mu)$ which is of finite band size (i.e., $|\langle\zeta\chi_m,\chi_n\rangle| = 0$ if $|m-n| > k$ for some k; in this case, $k = 2$ to be compared with $k = 1$ for the real line case). Their basis (complete, orthonormal system) $\{\chi_n\}$ is obtained by orthonormalizing the sequence $1, \zeta, \zeta^{-1}, \zeta^2, \zeta^{-2}, \dots$. Remarkably, the χ's can be expressed in terms of φ's and φ^*'s (see [18, Proposition 4.2.2])

$$\chi_{2n}(z) = z^{-n}\varphi^*_{2n}(z), \qquad \chi_{2n+1}(z) = z^{-n}\varphi_{2n+1}(z), \quad n \in \mathbb{Z}_+, \tag{1.11}$$

and the matrix elements

$$\mathcal{C}(\mu) = \|c_{nm}\| = \langle\zeta\chi_m,\chi_n\rangle, \qquad m,n \in \mathbb{Z}_+$$

in terms of Verblunsky coefficients

$$\mathcal{C}(\{\alpha_n\}) = \begin{pmatrix} \overline{\alpha}_0 & \overline{\alpha}_1\rho_0 & \rho_0\rho_1 & 0 & 0 & \dots \\ \rho_0 & -\overline{\alpha}_1\alpha_0 & -\alpha_0\rho_1 & 0 & 0 & \dots \\ 0 & \overline{\alpha}_2\rho_1 & -\overline{\alpha}_2\alpha_1 & \overline{\alpha}_3\rho_2 & \rho_2\rho_3 & \dots \\ 0 & \rho_1\rho_2 & -\alpha_1\rho_2 & -\overline{\alpha}_3\alpha_2 & -\alpha_2\rho_3 & \dots \\ 0 & 0 & 0 & \overline{\alpha}_4\rho_3 & -\overline{\alpha}_4\alpha_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{1.12}$$

with $\rho_n^2 := 1-|\alpha_n|^2$, $0 < \rho \le 1$.

There is an important relation between CMV matrices and monic orthogonal polynomials akin to the well-known property of orthogonal polynomials on the real line:

$$\Phi_n(z) = \det(zI_n - \mathcal{C}^{(n)}), \tag{1.13}$$

where $\mathcal{C}^{(n)}$ is the principal $n \times n$ block of $\mathcal{C}$ (see, e.g., [18, p. 271, formula (4.2.47)]).

The CMV matrices $\mathcal{C}$ play much the same role in our study of the Schur flows that Jacobi matrices (1.2) in the Toda lattices setting.

We are in a position now to announce our main result.

Theorem 1. *Let $\alpha_n(t)$, $n \in \mathbb{Z}_+$, be a sequence of complex valued functions with $|\alpha_n(t)| < 1$ for $t \ge 0$ and let $\alpha_{-1} = -1$. The following three statements are equivalent.*

(1) *α_n solve the Schur flow equations*

$$\alpha_n'(t) = (1 - |\alpha_n(t)|^2)(\alpha_{n+1}(t) - \alpha_{n-1}(t)), \qquad t > 0; \tag{1.14}$$

(2) *The CMV matrices $\mathcal{C}(t)$ satisfy the Lax equation*

$$\mathcal{C}'(t) = [A, \mathcal{C}], \tag{1.15}$$

where A is an upper triangular and tridiagonal matrix

$$A = \|a_{mn}\|_0^\infty = \begin{pmatrix} \Re\overline{\alpha}_0 & \rho_0\overline{\Delta}_0 & \rho_0\rho_1 & & & \\ & -\Re\overline{\alpha}_1\alpha_0 & \rho_1\Delta_1 & \rho_1\rho_2 & & \\ & & -\Re\overline{\alpha}_2\alpha_1 & \rho_2\overline{\Delta}_2 & \rho_2\rho_3 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{1.16}$$

$\Delta_n := \alpha_{n+1}(t) - \alpha_{n-1}(t)$;

(3) *The orthogonality measure $\mu(\zeta, t)$, having $\alpha_n(t)$ as its Verblunsky coefficients, satisfies*

$$d\mu(\zeta, t) = C(t)e^{t(\zeta+\zeta^{-1})}\, d\mu(\zeta, 0), \tag{1.17}$$

where C is a normalizing factor.

We refer to (1.17) as the *Bessel modification* of the measure $d\mu = d\mu(\zeta, 0)$.

Remarks. 1. It is not hard to see that if $\{\alpha_n(t)\}$ solves (1.14) with the initial data $|\alpha_n(0)| < 1$, then $|\alpha_n(t)| < 1$ for each $t > 0$. Indeed, assume for the contrary that for some $n \ge 0$ there is $t_0 > 0$ such that $|\alpha_n(t_0)| = 1$ and $|\alpha_n(t)| < 1$ for all $0 < t < t_0$. It follows from (1.14) and $\rho_n^2 = 1 - |\alpha_n|^2$ that

$$(\rho_n^2)' = 2\rho_n\rho_n' = -\alpha_n'\overline{\alpha}_n - \alpha_n\overline{\alpha}_n' = -2\rho_n^2\Re(\overline{\alpha}_{n+1}\alpha_n - \overline{\alpha}_n\alpha_{n-1}),$$

and so

$$\rho_n' = -\rho_n\Re(\overline{\alpha}_{n+1}\alpha_n - \overline{\alpha}_n\alpha_{n-1}), \quad 0 < t < t_0. \tag{1.18}$$

Hence

$$\rho_n(t) = \rho_n(0)\exp\left\{-\int_0^t \Re(\overline{\alpha}_{n+1}(s)\alpha_n(s) - \overline{\alpha}_n(s)\alpha_{n-1}(s))ds\right\}, \tag{1.19}$$

and the right hand side is bounded away from zero as $t \to t_0$, whereas the left hand side tends to zero. The contradiction shows that $|\alpha_n(t)| < 1$ for all $n \in \mathbb{Z}_+$ and $t > 0$. Therefore, by (3) the solution of the initial-boundary-value (IBV) problem

$$\alpha_n'(t) = (1 - |\alpha_n(t)|^2)(\alpha_{n+1}(t) - \alpha_{n-1}(t)), \quad t > 0; \qquad |\alpha_n(0)| < 1, \tag{1.20}$$

$n \in \mathbb{Z}_+$, $\alpha_{-1} = -1$ exists and unique.

2. We can modify the second statement by observing that

$$A + A^* = \mathcal{C} + \mathcal{C}^* = \begin{pmatrix} 2\Re\overline{\alpha}_0 & \rho_0\overline{\Delta}_0 & \rho_0\rho_1 & & & \\ \rho_0\Delta_0 & -2\Re\overline{\alpha}_1\alpha_0 & \rho_1\Delta_1 & \rho_1\rho_2 & & \\ \rho_0\rho_1 & \rho_1\overline{\Delta}_1 & -2\Re\overline{\alpha}_2\alpha_1 & \rho_2\overline{\Delta}_2 & \rho_2\rho_3 & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

and so

$$\mathcal{C}'(t) = [B, \mathcal{C}], \tag{1.21}$$

$$B = \frac{(\mathcal{C}+\mathcal{C}^*)_+ - (\mathcal{C}+\mathcal{C}^*)_-}{2} = \frac{1}{2}\begin{pmatrix} 0 & \rho_0\overline{\Delta}_0 & \rho_0\rho_1 & & & \\ -\rho_0\Delta_0 & 0 & \rho_1\Delta_1 & \rho_1\rho_2 & & \\ -\rho_0\rho_1 & -\rho_1\overline{\Delta}_1 & 0 & \rho_2\overline{\Delta}_2 & \rho_2\rho_3 & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

$$= A - \frac{\mathcal{C}+\mathcal{C}^*}{2} = -B^*, \tag{1.22}$$

which makes it closer to its counerpart in the Toda lattices setting.

So, once again, the solution of IBV problem (1.20) amounts to a combination of the direct and inverse spectral problems with (1.17) in between. Note that the orthogonality measure $\mu(\zeta, 0)$ can be retrieved from the initial data $\alpha_n(0)$ by either the Spectral Theorem for the CMV matrix $\mathcal{C}(0)$ (1.12) or via orthonormal polynomials, since $\mu(\zeta, 0)$ arises as a *-weak limit of the sequence of measures $|\varphi_n|^{-2}dm$, dm being a normalized Lebesgue measure on $\mathbb{T}$ (Rakhmanov's theorem). In turn, the Verblunsky coefficients $\alpha_n(t)$ are recovered from the measure $\mu(\zeta, t)$ by (1.10).

The Schur flow (1.14) emerged in [1, 2] under the name *discrete modified KdV equation*, as a spatial discretization of the modified Korteweg–de Vries equation

$$\partial_t f = 6f^2\partial_x f - \partial_x^3 f.$$

In [8] the authors deal with finite real Schur flows and suggest two more distinct Lax equations based on the Hessenberg matrix representation of the multiplication operator (see also [3]). In [13, 14] the Bessel modification of measures appeared and a part of our main result which concerns (3) ⇒ (1) is proved. In a recent paper [15] the author deals with the Poisson structure and Lax pairs for the Ablowits–Ladik systems closely related to the Schur flows. The latter can also be viewed as the zero-curvature equation for the Szegő matrices (cf. [9])

$$T_n'(z,t) + T_n(z,t)W_n(z,t) - W_{n+1}(z,t)T_n(z,t) = 0,$$

$$W_n(z,t) := \begin{pmatrix} z + 1 - \alpha_{n-1}\overline{\alpha}_n & -\overline{\alpha}_n - \overline{\alpha}_{n-1}z^{-1} \\ -\alpha_{n-1}z - \alpha_n & 1 - \overline{\alpha}_{n-1}\alpha_n + z^{-1} \end{pmatrix}.$$

We proceed as follows. In Sections 2 and 3 the proof of our main result is presented with some comments on the general IBV problem and doubly infinite systems. In Section 4 we study the long time behavior of the Schur flows and look into the modified Bessel polynomials on the unit circle, a nice example which corresponds to the zero initial conditions in our setting. In this case the long time behavior of the Verblunsky coefficients can be specified.

2. Proof of Theorem 1: (1) ⇔ (2).

(1) ⇒ (2). Once the Lax pair is enunciated, the proof goes through by brute force computation which is much more involved compared to the Toda case.

Let $2\varepsilon_m := 1-(-1)^m$, $m\in\mathbb{Z}_+$, and $\varepsilon_{-1}=1$, so $\{\varepsilon_m\}_{m\ge 0}=\{0,1,0,1,\dots\}$,

$$\varepsilon_m+\varepsilon_{m+1}=1, \qquad \varepsilon_m\varepsilon_{m+1}=0, \qquad \varepsilon_m-\varepsilon_{m+1}=(-1)^{m+1}.$$

It is instructive to write the diagonals of $\mathcal{C}$ (1.12) in a unique way

$$c_{mm} = -\overline{\alpha}_m\alpha_{m-1}, \tag{2.1}$$

$$c_{m+2,m} = \rho_m\rho_{m+1}\varepsilon_m, \qquad c_{m,m+2}=\rho_m\rho_{m+1}\varepsilon_{m+1}, \tag{2.2}$$

and

$$c_{m+1,m} = \overline{\alpha}_{m+1}\rho_m\varepsilon_m-\alpha_{m-1}\rho_m\varepsilon_{m+1}, \tag{2.3}$$

$$c_{m,m+1} = \overline{\alpha}_{m+1}\rho_m\varepsilon_{m+1}-\alpha_{m-1}\rho_m\varepsilon_m. \tag{2.4}$$

In the same vein for the matrix entries of A (1.16)

$$a_{mm}=-\Re\overline{\alpha}_m\alpha_{m-1}, \qquad a_{m,m+2}=\rho_m\rho_{m+1}, \tag{2.5}$$

and

$$a_{m,m+1}=\rho_m\overline{\Delta}_m\varepsilon_{m+1}+\rho_m\Delta_m\varepsilon_m. \tag{2.6}$$

Next, it follows from (1.14) and (1.18) that

$$\begin{aligned}(\rho_m\rho_{m+1})' &= -\rho_m\rho_{m+1}\Re(\overline{\alpha}_{m+2}\alpha_{m+1}-\overline{\alpha}_m\alpha_{m-1}),\\ (\overline{\alpha}_m\alpha_{m-1})' &= \alpha_{m-1}\rho_m^2\overline{\Delta}_m+\overline{\alpha}_m\rho_{m-1}^2\Delta_{m-1},\end{aligned}$$

and

$$\begin{aligned}(\alpha_{m-1}\rho_m)' &= \left[\rho_{m-1}^2\Delta_{m-1}-\alpha_{m-1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})\right]\rho_m,\\ (\overline{\alpha}_{m+1}\rho_m)' &= \left[\rho_{m+1}^2\overline{\Delta}_{m+1}-\overline{\alpha}_{m+1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})\right]\rho_m.\end{aligned}$$

Hence, for derivatives of the CMV matrix entries we have now

$$c'_{mm} = -(\overline{\alpha}_m\alpha_{m-1})'=-\alpha_{m-1}\rho_m^2\overline{\Delta}_m-\overline{\alpha}_m\rho_{m-1}^2\Delta_{m-1}, \tag{2.7}$$

$$c'_{m,m+2} = -\rho_m\rho_{m+1}\Re(\overline{\alpha}_{m+2}\alpha_{m+1}-\overline{\alpha}_m\alpha_{m-1})\varepsilon_{m+1}, \tag{2.8}$$

$$c'_{m,m-2} = -\rho_{m-2}\rho_{m-1}\Re(\overline{\alpha}_m\alpha_{m-1}-\overline{\alpha}_{m-2}\alpha_{m-3})\varepsilon_m, \tag{2.9}$$

and

$$\begin{aligned}c'_{m,m+1} &= \left[\rho_{m+1}^2\overline{\Delta}_{m+1}-\overline{\alpha}_{m+1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})\right]\rho_m\varepsilon_{m+1}\\ &- \left[\rho_{m-1}^2\Delta_{m-1}-\alpha_{m-1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})\right]\rho_m\varepsilon_m,\end{aligned} \tag{2.10}$$

$$\begin{aligned}c'_{m,m-1} &= \left[\rho_m^2\overline{\Delta}_m-\overline{\alpha}_m\Re(\overline{\alpha}_m\alpha_{m-1}-\overline{\alpha}_{m-1}\alpha_{m-2})\right]\rho_{m-1}\varepsilon_{m+1}\\ &- \left[\rho_{m-2}^2\Delta_{m-2}-\alpha_{m-2}\Re(\overline{\alpha}_m\alpha_{m-1}-\overline{\alpha}_{m-1}\alpha_{m-2})\right]\rho_{m-1}\varepsilon_m.\end{aligned} \tag{2.11}$$

Let $K=\|K_{mn}\|=[A,\mathcal{C}]$, $K_{mn}=\sum_i(a_{mi}c_{in}-c_{mi}a_{in})$. Since both A and $\mathcal{C}$ are of band size 2, i.e., $a_{mn}=c_{mn}=0$ for $|m-n|>2$, and A is an upper triangular, i.e., $a_{mn}=0$ for $m>n$, we actually have

$$\begin{aligned}K_{mn} &= a_{mm}c_{mn}+a_{m,m+1}c_{m+1,n}+a_{m,m+2}c_{m+2,n}\\ &- c_{mn}a_{nn}-c_{m,n-1}a_{n-1,n}-c_{m,n-2}a_{n-2,n}.\end{aligned} \tag{2.12}$$

We want to show that

$$c'_{m,m+j}(t)=K_{m,m+j}(t), \qquad m,m+j\in\mathbb{Z}_+. \tag{2.13}$$

To this end we will plug (2.1)–(2.6) into (2.12) and compare the outcome with (2.7)–(2.11). [1] For $j \ge 5$ and $j \le -3$ the equality holds for trivial reason, as the both sides in (2.13) vanish. For $j = 4$

$$\begin{aligned} K_{m,m+4} &= a_{m,m+2}c_{m+2,m+4} - c_{m,m+2}a_{m+2,m+4} \\ &= \rho_m\rho_{m+1}\rho_{m+2}\rho_{m+3}(\varepsilon_{m+3} - \varepsilon_{m+1}) = 0. \end{aligned}$$

For $j = 3$ by (2.2) and (2.4)–(2.6)

$$\begin{aligned} K_{m,m+3} &= a_{m,m+1}c_{m+1,m+3} + a_{m,m+2}c_{m+2,m+3} \\ &- c_{m,m+1}a_{m+1,m+3} - c_{m,m+2}a_{m+2,m+3} = \end{aligned}$$

$$\rho_m\rho_{m+1}\rho_{m+2}\left[\varepsilon_m(\Delta_m - \alpha_{m+1} + \alpha_{m-1}) + \varepsilon_{m+1}(\overline{\alpha}_{m+3} - \overline{\alpha}_{m+1} - \overline{\Delta}_{m+2})\right] = 0,$$

that is consistent with the banded structure of size 2 of $\mathcal{C}'$.

The main work begins when $|j| = 0, 1, 2$.

1. $j = -2$. We have by (2.2) and (2.5)

$$\begin{aligned} K_{m,m-2} &= c_{m,m-2}(a_{mm} - a_{m-2,m-2}) \\ &= \rho_{m-2}\rho_{m-1}\Re(\overline{\alpha}_{m-2}\alpha_{m-3} - \overline{\alpha}_m\alpha_{m-1})\varepsilon_m, \end{aligned} \tag{2.14}$$

and so (2.13) holds by (2.9).

2. $j = -1$. Write $K_{m,m-1} = K^{(1)}_{m,m-1} + K^{(2)}_{m,m-1}$ with

$$\begin{aligned} K^{(1)}_{m,m-1} &= c_{m,m-1}(a_{mm} - a_{m-1,m-1}) \\ &= (\overline{\alpha}_m\rho_{m-1}\varepsilon_{m+1} - \alpha_{m-2}\rho_{m-1}\varepsilon_m)\Re(\overline{\alpha}_{m-1}\alpha_{m-2} - \overline{\alpha}_m\alpha_{m-1}), \\ K^{(2)}_{m,m-1} &= a_{m,m+1}c_{m+1,m-1} - a_{m-2,m-1}c_{m,m-2} \\ &= \rho_{m-1}(\rho_m^2\overline{\Delta}_m\varepsilon_{m+1} - \rho_{m-2}^2\Delta_{m-2}\varepsilon_m) \end{aligned}$$

and hence

$$\begin{aligned} K_{m,m-1} &= \left[\rho_m^2\overline{\Delta}_m - \overline{\alpha}_m\Re(\overline{\alpha}_m\alpha_{m-1} - \overline{\alpha}_{m-1}\alpha_{m-2})\right]\rho_{m-1}\varepsilon_{m+1} \\ &- \left[\rho_{m-2}^2\Delta_{m-2} - \alpha_{m-2}\Re(\overline{\alpha}_m\alpha_{m-1} - \overline{\alpha}_{m-1}\alpha_{m-2})\right]\rho_{m-1}\varepsilon_m. \end{aligned} \tag{2.15}$$

Now (2.13) follows from (2.11).

3. $j = 0$. Write $K_{mm} = K^{(1)}_{mm} + K^{(2)}_{mm}$ with

$$\begin{aligned} K^{(1)}_{mm} &= a_{m,m+2}c_{m+2,m} - c_{m,m-2}a_{m-2,m} = \left[(\rho_m\rho_{m+1})^2 - (\rho_{m-2}\rho_{m-1})^2\right]\varepsilon_m, \\ K^{(2)}_{mm} &= a_{m,m+1}c_{m+1,m} - c_{m,m-1}a_{m-1,m} \\ &= (\overline{\alpha}_{m+1}\rho_m^2\Delta_m + \alpha_{m-2}\rho_{m-1}^2\overline{\Delta}_{m-1})\varepsilon_m \\ &- (\alpha_{m-1}\rho_m^2\overline{\Delta}_m + \overline{\alpha}_m\rho_{m-1}^2\Delta_{m-1})\varepsilon_{m+1}. \end{aligned}$$

But

$$(\rho_m\rho_{m+1})^2 - (\rho_{m-2}\rho_{m-1})^2 + \overline{\alpha}_{m+1}\rho_m^2\Delta_m + \alpha_{m-2}\rho_{m-1}^2\overline{\Delta}_{m-1} =$$

$$\rho_m^2(1 - \overline{\alpha}_{m+1}\alpha_{m-1}) - \rho_{m-1}^2(1 - \overline{\alpha}_m\alpha_{m-2}) = -\alpha_{m-1}\rho_m^2\overline{\Delta}_m - \overline{\alpha}_m\rho_{m-1}^2\Delta_{m-1},$$

and so by (2.7)

$$K_{mm} = -(\alpha_{m-1}\rho_m^2\overline{\Delta}_m + \overline{\alpha}_m\rho_{m-1}^2\Delta_{m-1})(\varepsilon_m + \varepsilon_{m+1}) = c'_{mm}.$$

[1] One needs some paper and patience to slug one's way through the lengthy calculation.

4. $j=1$. Now $K_{m,m+1}=K^{(1)}_{m,m+1}+K^{(2)}_{m,m+1}+K^{(3)}_{m,m+1}$ with

$$\begin{aligned}
K^{(1)}_{m,m+1} &= c_{m,m+1}(a_{mm}-a_{m+1,m+1}) \\
&= (\overline{\alpha}_{m+1}\varepsilon_{m+1}-\alpha_{m-1}\varepsilon_m)\rho_m\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1}), \\
K^{(2)}_{m,m+1} &= a_{m,m+1}(c_{m+1,m+1}-c_{mm}) \\
&= -(\overline{\Delta}_m\varepsilon_{m+1}-\Delta_m\varepsilon_m)\rho_m(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1}), \\
K^{(3)}_{m,m+1} &= a_{m,m+2}c_{m+2,m+1}-a_{m-1,m+1}c_{m,m-1} \\
&= \left[(\overline{\alpha}_{m+2}\rho^2_{m+1}-\overline{\alpha}_m\rho^2_{m-1})\varepsilon_{m+1}-(\alpha_m\rho^2_{m+1}-\alpha_{m-2}\rho^2_{m-1})\varepsilon_m\right]\rho_m.
\end{aligned}$$

Hence $\rho_m^{-1}K_{m,m+1}=u_m\varepsilon_m+v_m\varepsilon_{m+1}$ with

$$\begin{aligned}
u_m &= -\alpha_{m-1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})+\alpha_{m-2}\rho^2_{m-1}-\alpha_m\rho^2_{m+1} \\
&- \Delta_m(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1}), \\
v_m &= \overline{\alpha}_{m+1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})+\overline{\alpha}_{m+2}\rho^2_{m+1}-\overline{\alpha}_m\rho^2_{m-1} \\
&- \overline{\Delta}_m(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1}).
\end{aligned}$$

Next,

$$\begin{aligned}
u^{(1)}_m &:= \alpha_{m-2}\rho^2_{m-1}-\alpha_m\rho^2_{m+1}+(\alpha_{m-1}-\alpha_{m+1})(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1}) \\
&= \alpha_{m+1}\overline{\alpha}_m\alpha_{m-1}-\overline{\alpha}_m\alpha^2_{m-1}+\overline{\alpha}_{m+1}\alpha_m\alpha_{m-1}-\alpha_m+\alpha_{m-2}\rho^2_{m-1} \\
&= 2\alpha_{m-1}\Re\overline{\alpha}_{m+1}\alpha_m-2\alpha_{m-1}\Re\overline{\alpha}_m\alpha_{m-1}-(\alpha_m-\alpha_{m-2})\rho^2_{m-1},
\end{aligned}$$

and so

$$u_m=\alpha_{m-1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})-\rho^2_{m-1}\Delta_{m-1}.$$

In exactly the same way

$$v_m=-\overline{\alpha}_{m+1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})+\rho^2_{m+1}\overline{\Delta}_{m+1},$$

and finally

$$\begin{aligned}
K_{m,m+1} &= \left[\rho^2_{m+1}\overline{\Delta}_{m+1}-\overline{\alpha}_{m+1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})\right]\rho_m\varepsilon_{m+1} \\
&- \left[\rho^2_{m-1}\Delta_{m-1}-\alpha_{m-1}\Re(\overline{\alpha}_{m+1}\alpha_m-\overline{\alpha}_m\alpha_{m-1})\right]\rho_m\varepsilon_m.
\end{aligned} \tag{2.16}$$

We come to (2.13) on account of (2.10).

5. $j=2$. Write $K_{m,m+2}=K^{(1)}_{m,m+2}+K^{(2)}_{m,m+2}+K^{(3)}_{m,m+2}$ with

$$\begin{aligned}
K^{(1)}_{m,m+2} &= a_{m,m+2}(c_{m+2,m+2}-c_{mm})=\rho_m\rho_{m+1}(\overline{\alpha}_m\alpha_{m-1}-\overline{\alpha}_{m+2}\alpha_{m+1})(\varepsilon_m+\varepsilon_{m+1}), \\
K^{(2)}_{m,m+2} &= c_{m,m+2}(a_{mm}-a_{m+2,m+2})=\rho_m\rho_{m+1}\varepsilon_{m+1}\Re(-\overline{\alpha}_m\alpha_{m-1}+\overline{\alpha}_{m+2}\alpha_{m+1}), \\
K^{(3)}_{m,m+2} &= a_{m,m+1}c_{m+1,m+2}-c_{m,m+1}a_{m+1,m+2} \\
&= \rho_m\rho_{m+1}\left[(\overline{\alpha}_{m+2}\Delta_m+\alpha_{m-1}\overline{\Delta}_{m+1})\varepsilon_m-(\alpha_m\overline{\Delta}_m+\overline{\alpha}_{m+1}\Delta_{m+1})\varepsilon_{m+1}\right].
\end{aligned}$$

We have

$$\begin{aligned}
\overline{\alpha}_{m+2}\Delta_m+\alpha_{m-1}\overline{\Delta}_{m+1} &= -\overline{\alpha}_m\alpha_{m-1}+\overline{\alpha}_{m+2}\alpha_{m+1} \\
\alpha_m\overline{\Delta}_m+\overline{\alpha}_{m+1}\Delta_{m+1} &= -\alpha_m\overline{\alpha}_{m-1}+\alpha_{m+2}\overline{\alpha}_{m+1}
\end{aligned}$$

and it follows from (2.8) that

$$K_{m,m+2}=-\rho_m\rho_{m+1}\Re(\overline{\alpha}_{m+2}\alpha_{m+1}-\overline{\alpha}_m\alpha_{m-1})\varepsilon_{m+1}, \tag{2.17}$$

so (2.13) holds again. The proof is complete.

(2) $\Rightarrow$ (1). The problem we are faced with is that, in contrast to the Toda lattices, no α's in a pure form appear among the matrix entries of $\mathcal{C}$.

Write $|\overline{\alpha}_{n+1}\rho_n|^2 = (1-\rho_{n+1}^2)\rho_n^2 = \rho_n^2 - (\rho_n\rho_{n+1})^2$ and so

$$(\rho_n^2)' = 2\rho_n\rho_{n+1}(\rho_n\rho_{n+1})' + (\overline{\alpha}_{n+1}\rho_n)'(\alpha_{n+1}\rho_n) + (\overline{\alpha}_{n+1}\rho_n)(\alpha_{n+1}\rho_n)'.$$

Hence

$$\rho_n' = \rho_{n+1}(\rho_n\rho_{n+1})' + \Re[\alpha_{n+1}(\overline{\alpha}_{n+1}\rho_n)']. \tag{2.18}$$

Next, $(\alpha_{n-1}\rho_n)' = \alpha_{n-1}'\rho_n + \alpha_{n-1}\rho_n'$, so that

$$\alpha_{n-1}' = \frac{1}{\rho_n}\left[(\alpha_{n-1}\rho_n)' - \alpha_{n-1}\rho_n'\right]. \tag{2.19}$$

The right hand side of (2.19) can be expressed in terms of derivatives of the CMV matrix entries and thereby, via the Lax equation, of α's themselves. First, by (2.14) and (2.17)

$$\begin{aligned} (\rho_n\rho_{n+1})' &= c_{n,n+2}' + c_{n+2,n}' = K_{n,n+2} + K_{n+2,n} && (2.20)\\ &= \rho_n\rho_{n+1}\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+2}\alpha_{n+1}). && (2.21)\end{aligned}$$

Next, it is immediate from (2.3), (2.4) that $\overline{\alpha}_{n+1}\rho_n = c_{n+1,n}\varepsilon_n + c_{n,n+1}\varepsilon_{n+1}$, and so by (2.15), (2.16)

$$\begin{aligned} (\overline{\alpha}_{n+1}\rho_n)' &= c_{n+1,n}'\varepsilon_n + c_{n,n+1}'\varepsilon_{n+1} = K_{n+1,n}\varepsilon_n + K_{n,n+1}\varepsilon_{n+1} && (2.22)\\ &= \overline{\alpha}_{n+1}\rho_n\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n) + \rho_n\rho_{n+1}^2\overline{\Delta}_{n+1}. && (2.23)\end{aligned}$$

Similarly

$$(\alpha_{n-1}\rho_n)' = \alpha_{n-1}\rho_n\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n) + \rho_n\rho_{n-1}^2\Delta_{n-1}.$$

Plugging (2.20) and (2.22) into (2.18) gives

$$\begin{aligned} \rho_n' &= \rho_n\rho_{n+1}^2\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+2}\alpha_{n+1} + \alpha_{n+1}\overline{\Delta}_{n+1}) + \rho_n|\alpha_{n+1}|^2\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n)\\ &= \rho_n\rho_{n+1}^2\Re(-\overline{\alpha}_{n+2}\alpha_{n+1} + \overline{\alpha}_{n+1}\alpha_n + \alpha_{n+1}\overline{\Delta}_{n+1}) + \rho_n\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n)\\ &= \rho_n\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n).\end{aligned}$$

In the upshot, the Schur flow equations emerge from (2.19):

$$\begin{aligned} \alpha_{n-1}'\rho_n &= \alpha_{n-1}\rho_n\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n) + \rho_n\rho_{n-1}^2\Delta_{n-1}\\ &= -\alpha_{n-1}\rho_n\Re(\overline{\alpha}_n\alpha_{n-1} - \overline{\alpha}_{n+1}\alpha_n) = \rho_n\rho_{n-1}^2\Delta_{n-1},\end{aligned}$$

as claimed. □

Remark. We could equally well have considered the general IBV problem, that, strictly speaking has nothing to do with OPUC:

$$\alpha_n'(t) = (1-|\alpha_n(t)|^2)(\alpha_{n+1}(t) - \alpha_{n-1}(t)), \qquad t>0$$

with a continuous boundary function $|\alpha_{-1}(t)| \le 1$. The above evaluation shows that the Lax form of such problem is $\mathcal{C}_g' = [A_g, \mathcal{C}_g]$ with

$$\mathcal{C}_g = \begin{pmatrix} -\overline{\alpha}_0\alpha_{-1} & \overline{\alpha}_1\rho_0 & \rho_0\rho_1 & 0 & 0 & \dots\\ -\alpha_{-1}\rho_0 & -\overline{\alpha}_1\alpha_0 & -\alpha_0\rho_1 & 0 & 0 & \dots\\ 0 & \overline{\alpha}_2\rho_1 & -\overline{\alpha}_2\alpha_1 & \overline{\alpha}_3\rho_2 & \rho_2\rho_3 & \dots\\ 0 & \rho_1\rho_2 & -\alpha_1\rho_2 & -\overline{\alpha}_3\alpha_2 & -\alpha_2\rho_3 & \dots\\ 0 & 0 & 0 & \overline{\alpha}_4\rho_3 & -\overline{\alpha}_4\alpha_3 & \dots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

and

$$A_g = \begin{pmatrix} -\Re\overline{\alpha}_0\alpha_{-1} & \rho_0\overline{\Delta}_0 & \rho_0\rho_1 & & & \\ & -\Re\overline{\alpha}_1\alpha_0 & \rho_1\Delta_1 & \rho_1\rho_2 & & \\ & & -\Re\overline{\alpha}_2\alpha_1 & \rho_2\overline{\Delta}_2 & \rho_2\rho_3 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

(to be compared with (1.12) and (1.16)). Furthermore, the doubly infinite system

$$\alpha_n'(t) = (1-|\alpha_n(t)|^2)(\alpha_{n+1}(t)-\alpha_{n-1}(t)), \qquad t>0, \quad n\in\mathbb{Z}$$

with the initial data $\{\alpha_n(0)\}_{k\in\mathbb{Z}} \subset \mathbb{D}$ has its equivalent Lax form

$$\hat{\mathcal{C}}' = [\hat{A}, \hat{\mathcal{C}}],$$

where $\hat{\mathcal{C}}$ and $\hat{A}$ are doubly infinite extensions of (1.12) and (1.16) given by the same expressions (2.1)–(2.4) and (2.5)–(2.6), respectively, with

$$\varepsilon_m = \frac{1-(-1)^{|m|}}{2}, \qquad m\in\mathbb{Z}.$$

3. Proof of Theorem 1: (2) ⇒ (3) ⇒ (1).

(2) ⇒ (3). Let

$$R_z(t) := (\mathcal{C} - zI)^{-1} = \|r_{mn}(z,t)\|_{m,n=0}^{\infty}$$

be a resolvent of the CMV matrix (1.12). It is easy to see that R_z obeys the same Lax equation (1.15). Indeed, differentiating the identity $(\mathcal{C}-zI)R_z = I$ with respect to t entails

$$\mathcal{C}'(t)R_z(t) + (\mathcal{C}(t)-zI)R_z'(t) = 0,$$

and so

$$R_z'(t) = -R_z(t)\mathcal{C}'(t)R_z(t) = -R_z(t)[A,\mathcal{C}]R_z(t) = [A, R_z], \tag{3.1}$$

as claimed.

Take the equation for $(0,0)$-entry of (3.1):

$$\begin{aligned} r_{00}' &= \Re\overline{\alpha}_0 r_{00} + \rho_0\overline{\Delta}_0 r_{10} + \rho_0\rho_1 r_{20} - r_{00}\Re\overline{\alpha}_0 \\ &= \rho_0(\overline{\alpha}_1 + 1)r_{10} + \rho_0\rho_1 r_{20}. \end{aligned}$$

As it follows from $(\mathcal{C}-zI)R_z = I$

$$(\overline{\alpha}_0 - z)r_{00} + \overline{\alpha}_1\rho_0 r_{10} + \rho_0\rho_1 r_{20} = 1,$$

which allows to eliminate r_{20} in favor of r_{10}, r_{00}, and so

$$r_{00}' = \rho_0 r_{10} + 1 - (\overline{\alpha}_0 - z)r_{00}. \tag{3.2}$$

By the Spectral Theorem, the resolvent entries can be found from

$$r_{m,n}(z,t) = \int_{\mathbb{T}} \frac{\chi_n(\zeta)\overline{\chi_m(\zeta)}}{\zeta - z}\, d\mu(\zeta,t)$$

with χ_n (1.11), which is particularly simple for the first two elements

$$r_{00}(z,t) = \int_{\mathbb{T}} \frac{d\mu(\zeta,t)}{\zeta - z}, \qquad r_{10}(z,t) = \int_{\mathbb{T}} \frac{\overline{\varphi_1(\zeta)}}{\zeta - z}\, d\mu(\zeta,t),$$

$\varphi_1(z) = \rho_0^{-1}(z-\overline{\alpha}_0)$. Hence for the right hand side of (3.2)

$$\rho_0 r_{10} + 1 - (\overline{\alpha}_0 - z)r_{00} = \int_{\mathbb{T}} \frac{\zeta + \overline{\zeta} - 2\Re\alpha_0}{\zeta - z}\, d\mu(\zeta,t)$$

holds, and we end up with a differential equation for orthogonality measures

$$d\mu'(\zeta,t) = (\zeta+\bar\zeta-2\Re\alpha_0)d\mu(\zeta,t).$$

Finally,

$$\begin{aligned} d\mu(\zeta,t) &= \exp\left\{\int_0^t (\zeta+\bar\zeta-2\Re\alpha_0(s))ds\right\} d\mu(\zeta,0) \\ &= C(t)e^{t(\zeta+\zeta^{-1})}\, d\mu(\zeta,0), \end{aligned}$$

as needed.

(3) $\Rightarrow$ (1). We start out from the modification of the orthogonality measure $d\mu(\zeta,t)=\varphi(\zeta,t)\,d\mu(\zeta,0)$ and derive a differential equation for the moments

$$\mu_k(t) = \int_{\mathbb{T}} \zeta^{-k}\, d\mu(\zeta,t) = \int_{\mathbb{T}} \zeta^{-k}\varphi(\zeta,t)\, d\mu(\zeta,0).$$

Now

$$\varphi(\zeta,t) = \frac{\exp(t(\zeta+\zeta^{-1}))}{f(t)}, \qquad f(t) = \int_{\mathbb{T}} e^{t(\zeta+\zeta^{-1})}\, d\mu(\zeta,0),$$

and so

$$\varphi'(\zeta,t) = (\zeta+\zeta^{-1})\varphi(\zeta,t) - \frac{f'(t)}{f(t)}\varphi(\zeta,t).$$

As $f'f^{-1}=c_1+c_{-1}$, we have

$$\mu_k'(t) = \mu_{k-1}(t)+\mu_{k+1}(t)-g(t)\mu_k(t), \qquad g(t)=c_1(t)+c_{-1}(t), \qquad k\in\mathbb{Z}. \tag{3.3}$$

The rest is based heavily on (1.10) which relates Verblunsky coefficients and moments of the orthogonality measure. The idea to differentiate determinants and take into account (3.3) goes back to [5], see also [13, lemma 1]. For a set of integers $k_1<k_2<\ldots<k_n$ denote

$$T(k_1,\ldots,k_n) := \begin{pmatrix} \mu_{k_1} & \mu_{k_1-1} & \cdots & \mu_{k_1-n+1} \\ \mu_{k_2} & \mu_{k_2-1} & \cdots & \mu_{k_2-n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{k_n} & \mu_{k_n-1} & \cdots & \mu_{k_n-n+1} \end{pmatrix},$$

$D(k_1,\ldots,k_n) := \det T(k_1,\ldots,k_n)$, and so $D_n = D(0,1,\ldots,n-1)$. Put $G_n = D(-1,0,\ldots,n-2)$ and write (1.10) as $\Phi_n(0,t)=(-1)^nG_nD_n^{-1}$. Then

$$\Phi_n'(0,t) = (-1)^n\frac{G_n'D_n-G_nD_n'}{D_n^2}. \tag{3.4}$$

It is clear from (3.3) that intermediate determinants in the sum

$$D_n' = \begin{vmatrix} \mu_0' & \mu_{-1}' & \cdots & \mu_{-n+1}' \\ \mu_1 & \mu_0 & \cdots & \mu_{-n+2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_{n-2} & \cdots & \mu_0 \end{vmatrix} + \ldots + \begin{vmatrix} \mu_0 & \mu_{-1} & \cdots & \mu_{-n+1} \\ \mu_1 & \mu_0 & \cdots & \mu_{-n+2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1}' & \mu_{n-2}' & \cdots & \mu_0' \end{vmatrix}$$

have the same value $-gD_n$, whereas the first and the last ones equal, respectively,

$$\begin{vmatrix} \mu'_0 & \mu'_{-1} & \cdots & \mu'_{-n+1} \\ \mu_1 & \mu_0 & \cdots & \mu_{-n+2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_{n-2} & \cdots & \mu_0 \end{vmatrix} = D(-1,1,2,\dots,n-1) - g(t)D_n,$$

$$\begin{vmatrix} \mu_0 & \mu_{-1} & \cdots & \mu_{-n+1} \\ \mu_1 & \mu_0 & \cdots & \mu_{-n+2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu'_{n-1} & \mu'_{n-2} & \cdots & \mu'_0 \end{vmatrix} = D(0,1,\dots,n-2,n) - g(t)D_n.$$

Hence

$$D'_n(t) = -ng(t)D_n(t) + D(-1,1,\dots,n-2,n-1) + D(0,1,\dots,n-2,n). \tag{3.5}$$

Similarly,

$$G'_n(t) = -ng(t)G_n(t) + D(-2,0,\dots,n-3,n-2) + D(-1,0,\dots,n-3,n-1). \tag{3.6}$$

After plugging (3.5) and (3.6) into (3.4) we come to

$$\begin{aligned} \Phi'_n(0,t) &= \frac{(-1)^n}{D_n^2}\left\{D_n\left[D(-2,0,\dots,n-2) + D(-1,\dots,n-3,n-1)\right]\right. \\ &- \left. G_n\left[D(-1,1,\dots,n-1) + D(0,\dots,n-2,n)\right]\right\}. \end{aligned}$$

Let us now go over to the right hand side of (1.14), written for $\Phi_n(0)$:

$$\left(1-|\Phi_n(0)|^2\right)\left(\Phi_{n+1}(0) - \Phi_{n-1}(0)\right) = (-1)^{n+1}\frac{D_n^2 - |G_n|^2}{D_n^2}\left\{\frac{G_{n+1}}{D_{n+1}} - \frac{G_{n-1}}{D_{n-1}}\right\}.$$

The standard Silvester identity applied to the matrix $T(0,1,\dots,n)$ gives

$$D_n^2 - |G_n|^2 = D_{n+1}D_{n-1},$$

and so

$$\left(1-|\Phi_n(0)|^2\right)\left(\Phi_{n+1}(0) - \Phi_{n-1}(0)\right) = \frac{(-1)^{n+1}}{D_n^2}(G_{n+1}D_{n-1} - D_{n+1}G_{n-1}).$$

Another application of the Silvester identity (in a bit modified form) shows that

$$\begin{aligned} G_{n+1}D_{n-1} &= G_nD(-1,1,\dots,n-1) - D_nD(-2,0,\dots,n-2) \\ D_{n+1}G_{n-1} &= D_nD(-1,0,\dots,n-3,n-1) - G_nD(0,1,\dots,n-2,n), \end{aligned}$$

and we arrive at the Schur flow (1.14). That completes the proof of Theorem 1. □

There is yet another way to prove (3) ⇒ (1), which gives not only (1.14), but the differential equations for the monic orthogonal polynomials. I learned it from [10, Section 8.3].

Theorem 2. *The monic polynomials* $\Phi_n(\cdot,t)$ *orthogonal with respect to the Bessel modification* (1.17) *satisfy the differential equation*

$$\Phi'_n(z,t) = \Phi_{n+1}(z,t) - (z+\overline{\alpha}_n\alpha_{n-1})\Phi_n(z,t) - (1-|\alpha_{n-1}|^2)\Phi_{n-1}(z,t). \tag{3.8}$$

Proof. The idea is to differentiate orthogonality relations (1.7) with respect to t. Now $w' = (\zeta+\overline{\zeta})w$, and we have for $m<n$

$$\int_{\mathbb{T}}\Phi'_m\overline{\Phi}_n\,d\mu(\zeta,t) + \int_{\mathbb{T}}\Phi_m\overline{\Phi'_n}\,d\mu(\zeta,t) + \int_{\mathbb{T}}(\zeta+\overline{\zeta})\Phi'_m\overline{\Phi}_n\,d\mu(\zeta,t) = 0. \tag{3.9}$$

It is clear that Φ'_m is a polynomial of degree at most $m-1$, and so the first integral in the above sum is zero. If $m \le n-2$ then

$$\int_{\mathbb{T}} \Phi_m \overline{(\Phi'_n + \zeta\Phi_n)}\, d\mu(\zeta,t) = 0,$$

and hence

$$\Phi'_n(z) + z\Phi_n(z) = x_n\Phi_{n+1}(z) + y_n\Phi_n(z) + z_n\Phi_{n-1}(z) \tag{3.10}$$

with some parameters x_n, y_n, z_n depending on t. By matching the coefficients for z^{n+1} and z^n in (3.10) and using (1.6) we find

$$x_n = 1, \qquad y_n = l_n(t) - l_{n+1}(t).$$

To get z_n take (3.9) with $m = n-1$ and apply the Szegő recurrences (1.9):

$$0 = \int_{\mathbb{T}} \Phi_{n-1}\overline{(\Phi'_n + \zeta\Phi_n)}\, d\mu(\zeta,t) + \int_{\mathbb{T}} (\Phi_n + \overline{\alpha}_{n-1}\Phi^*_{n-1})\overline{\Phi}_n\, d\mu(\zeta,t),$$

and so

$$\begin{aligned}\int_{\mathbb{T}} \Phi_{n-1}\overline{(\Phi'_n + \zeta\Phi_n)}\, d\mu(\zeta,t) &= -\|\Phi_n\|^2_\mu, \\ z_n &= -\frac{\|\Phi_n\|^2_\mu}{\|\Phi_{n-1}\|^2_\mu} = -\frac{\kappa^2_{n-1}}{\kappa^2_n} = |\alpha_{n-1}|^2 - 1.\end{aligned}$$

To find the expression for l_n we turn to (1.13)

$$\Phi_n(z) = z^n + l_n z^{n-1} + \ldots = \prod_{j=0}^{n-1}(z + \overline{\alpha}_j\alpha_{j-1}) + \ldots,$$

so that $l_n = \sum_{j=0}^{n-1}\overline{\alpha}_j\alpha_{j-1}$, and we come to (3.8). Putting $z=0$ yields (1.14), as was to be proved. □

It might be worth pointing out that some properties of the modified Verblunsky coefficients (such as the rate of decay) are inherited from those of the initial data.

Theorem 3. *Let $\alpha_n(0)$ enjoy either of the properties*

(1) $\{\alpha_n(0)\} \in \ell^p$, $p = 1, 2$;
(2) $|\alpha_n(0)| \le Ce^{-\alpha n}$, $\alpha > 0$.

Then the same holds for $\alpha_n(t)$ for each $t > 0$.

Proof. It is obvious from (1.17) that $\mu(\zeta,t)$ belongs to the Szegő class (i.e., $\log\mu' \in \Lambda^1(\mathbb{T})$ if and only if $\mu(\zeta,0)$ does, and so the first statement with $p=2$ follows from fundamental Szegő's Theorem [18, Theorem 2.3.1]. As for the case $p=1$, note that by Baxter's theorem (see, e.g., [18, Theorem 5.2.1]) $\mu(\zeta,0) = wdm$ with $w > 0$ and $w \in W$, class of absolutely convergent Fourier series. It is clear from (1.17) and the Wiener–Levy theorem that

$$d\mu(\zeta,t) = w(\zeta,t)dm, \qquad w(\zeta,t) = e^{t(\zeta+\zeta^{-1})}w(\zeta) \in W,$$

$w(\zeta,t) > 0$, and so the repeated application of Baxter's theorem does the job.

To prove the second statement, let us introduce the Szegő function

$$D(z,\mu) := \exp\left(\frac{1}{2}\int_{\mathbb{T}} \frac{\zeta+z}{\zeta-z}\,\log w(\zeta)\, dm\right),$$

defined for an arbitrary measure $\mu = wdm + \mu_s$ from the Szegő class. A straightforward computation gives for the modified Bessel measure (1.17)

$$D(\zeta, \mu(t)) = \exp\left(\frac{1}{2}\int_{\mathbb{T}} \frac{\zeta + z}{\zeta - z}\,[t(\zeta + \zeta^{-1}) + \log w(\zeta)]\,dm\right) = e^{tz} D(z, \mu(0)). \tag{3.11}$$

Denote by R_t the radius of convergence of the Taylor series for $D^{-1}(z, \mu(t))$ about the origin. By the Nevai–Totik theorem [16]

$$\limsup_{n\to\infty} |\alpha_n(t)|^{\frac{1}{n}} = R_t^{-1},$$

and it is clear from (3.11) that $R_t = R_0$ for all $t > 0$, as claimed. □

Note that under assumptions of Theorem 3 the series

$$\mathcal{K} = -\sum_{j=0}^{\infty} \overline{\alpha}_j(t)\alpha_{j-1}(t)$$

converges absolutely and the Schur flow can be written in the form

$$\alpha_j' = \{\alpha_j, H\}, \qquad H = -2\Im\mathcal{K},$$

where the (formal) Poisson brackets are defined by

$$\{f, g\} = i\sum_{j\ge 0} \rho_j^2 \left[\frac{\partial f}{\partial\overline{\alpha}_j}\frac{\partial g}{\partial\alpha_j} - \frac{\partial f}{\partial\alpha_j}\frac{\partial g}{\partial\overline{\alpha}_j}\right].$$

So (1.14) is the evolution of the Verblunsky coefficients under the flow generated by the Hamiltonian $-2\Im\mathcal{K}$ (cf. [15]).

4. Long time behavior of the Schur flow and modified Bessel polynomials

The first part of this section is obviously very strongly influenced by Deift, Li and Tomei paper [7]. We want to show that the solution $\alpha_n(t)$ tends to the unit circle for each $n \in \mathbb{Z}_+$ as $t \to \infty$.

Theorem 4. *For the solution α_n of the Schur flow* (1.14) *the limit relations*

$$\lim_{t\to\infty} \alpha_n(t) = \lambda_n \in \mathbb{T}, \qquad n \in \mathbb{Z}_+ \tag{4.1}$$

hold.

Proof. To begin with, let us focus on the form of the Lax equation given in (1.21)–(1.22). As $B^* = -B$, it is clear that

$$L' = [\pi(L), L], \qquad L := \frac{\mathcal{C} + \mathcal{C}^*}{2}, \quad \pi(L) = L_+ - L_-. \tag{4.2}$$

The latter is exactly what is called in [7] the *Toda flow*. By [7, Proposition 5] $L(t)$ converges strongly to a diagonal operator $\mathrm{diag}(\gamma_0, \gamma_1, \ldots)$. Furthermore, for any real bounded measurable function G on $[-1, 1] \supset \sigma(L)$, $\sigma(L)$ being a spectrum of the bounded and self-adjoint operator L, we also have [7, remark 1, p. 363]

$$\tilde{L}' = [\pi(\tilde{L}), \tilde{L}], \qquad \tilde{L} = G(L(t)),$$

and so $\tilde{L}(t)$ converges strongly to a diagonal operator $\mathrm{diag}(\tilde{\gamma}_0, \tilde{\gamma}_1, \ldots)$. Our particular choice of G is $G(x) = \sqrt{1 - x^2}$, which gives

$$\tilde{L}(t) = \frac{\mathcal{C} - \mathcal{C}^{-1}}{2i} = \frac{\mathcal{C} - \mathcal{C}^*}{2i},$$

as $\mathcal{C}$ is the unitary operator. Therefore $L + i\tilde{L} = \mathcal{C}$ converges strongly to the diagonal operator $\operatorname{diag}(\Gamma_0, \Gamma_1, \ldots)$, $\Gamma_k := \gamma_k + i\tilde{\gamma}_k$. In particular, $c_{kj} \to 0$ for $k \neq j$, so

$$\alpha_n(t)\rho_{n+1}(t) \to 0, \qquad \rho_n(t)\rho_{n+1}(t) \to 0, \qquad n \in \mathbb{Z}_+$$

as $t \to \infty$. But

$$|\alpha_n(t)\rho_{n+1}|^2 = \rho_n^2(t) - \rho_n^2(t)\rho_{n+1}^2(t),$$

and hence $|\alpha_n| \to 1$ as $t \to \infty$. Next, taking the diagonal entries gives

$$c_{nn} = -\overline{\alpha}_n(t)\alpha_{n-1}(t) \to \Gamma_n \in \mathbb{T}, \qquad t \to \infty, \tag{4.3}$$

that entails ($\alpha_{-1} = -1$!) the existence of the limit (4.1) for all $n \geq 0$ and the equality

$$\overline{\lambda}_n = (-1)^n \Gamma_0 \Gamma_1 \ldots \Gamma_n \in \mathbb{T}, \tag{4.4}$$

as claimed.

Remark. It is easy to see from (1.19) and (4.3) that the sequence $\{\Re\Gamma_n\}$ decreases monotonically. Hence, if, for instance, all α's are real and $\alpha_0 \to \Gamma_0 = -1$, then

$$\Gamma_n = -1, \qquad \lim_{t\to\infty} \alpha_n(t) = -1, \qquad n \in \mathbb{Z}_+.$$

Example (modified Bessel polynomials on the unit circle). Because of the boundary condition $\alpha_{-1} = -1$ IBV problem (1.20) with zero initial conditions

$$\alpha_0(0) = \alpha_1(0) = \ldots = 0$$

has a nontrivial solution. Theorem 1, (3), shows that we are dealing now with the Bessel modification of the Lebesgue measure

$$d\mu(\zeta, t) = C(t)e^{t(\zeta+\zeta^{-1})}\, dm,$$

with α_n being the Verblunsky coefficients of this measure. The corresponding system of orthogonal polynomials has arisen from studies of the length of longest increasing subsequences of random words [4] and matrix models [17] (see [10, example 8.3.4] for more detail about the Bessel OPUC).

Note first that $C(t)$ can be easily computed

$$\begin{aligned} C^{-1}(t) &= \int_{\mathbb{T}} e^{t(\zeta+\zeta^{-1})} = \frac{1}{2\pi}\int_0^{2\pi} e^{2t\cos x}\, dx = \frac{1}{2\pi}\sum_{n=0}^{\infty} \frac{(2t)^n}{n!}\int_0^{2\pi} (\cos x)^n\, dx \\ &= \sum_{n=0}^{\infty} \frac{t^{2n}}{(n!)^2} = I_0(2t), \end{aligned}$$

where I_k is the modified Bessel function of order k. Similarly, for the moments of the measure we have

$$\mu_p(t) = \int_{\mathbb{T}} \zeta^{-p}\, d\mu(\zeta, t) = \frac{I_p(2t)}{I_0(2t)}, \quad p \in \mathbb{Z}_+, \qquad \mu_{-p} = \mu_p.$$

There is an important feature of the modified Bessel polynomials, namely their Verblunsky coefficients satisfy the recurrence relation (see [10, lemma 8.3.5]) [2]

$$-(n+1)\frac{\alpha_n(t)}{t(1-\alpha_n^2(t))} = \alpha_{n+1}(t) + \alpha_{n-1}(t), \quad n \in \mathbb{Z}_+, \tag{4.5}$$

with $\alpha_{-1} = -1$, $\alpha_0 = I_1(2t)/I_0(2t)$. Clearly, all α's are real now.

We can refine the general result of Theorem 4 in this particular case as follows.

[2] In the notation of [10] $r_n(t) = -\alpha_{n-1}(t/2)$.

Theorem 5. *For the modified Bessel OPUC the limit relations*

$$\lim_{t\to\infty} \alpha_n(t) = (-1)^n, \tag{4.6}$$

$$\lim_{t\to\infty} t(1-\alpha_n^2(t)) = \frac{n+1}{2} \tag{4.7}$$

hold for all $n \ge 0$.

Proof. Note first that (4.7) is an immediate consequence of (4.6) and (4.5). To prove (4.6) we proceed in two steps.

1. As we know (see remark after Theorem 4) the sequence $\Gamma_n \downarrow$. Let us show that in fact all Γ_n's are equal. Assume on the contrary that $\Gamma_k > \Gamma_{k+1}$ for some k. It follows from (1.19) that

$$\rho_k^2(t) \le Ce^{-\delta t}, \qquad \delta > 0.$$

But then

$$\frac{1}{t(1-\alpha_k^2(t))} = \frac{1}{\rho_k^2(t)} \ge \frac{e^{\delta t}}{Ct},$$

that contradicts (4.5), since $|\alpha_k| \to 1$ and the right hand side is bounded. So we need only find the value of Γ_n.

2. The case $n = 0$ can be handled directly. As is well known,

$$I_k(t) = \frac{e^t}{\sqrt{2\pi t}} \left(1 + O\left(\frac{1}{t}\right)\right), \quad t \to \infty,$$

and so

$$\lim_{t\to\infty} \alpha_0(t) = \lim_{t\to\infty} \frac{I_1(2t)}{I_0(2t)} = \Gamma_0 = 1,$$

that is, $\Gamma_n = 1$ for all n. The desired result drops out immediately from (4.4). $\square$

Remark. It might be a challenging problem to give a direct proof of (4.6) based on the explicit formula (1.10), which now takes on the form

$$\alpha_n(t) = (-1)^n \frac{\det \|I_{k-j-1}(2t)\|_{0\le k,j\le n}}{\det \|I_{k-j}(2t)\|_{0\le k,j\le n}}, \qquad n \in \mathbb{Z}_+,$$

and the complete asymptotic series expansion for the modified Bessel function

$$I_k(t) \simeq \frac{e^t}{\sqrt{2\pi t}} \sum_{j=0}^{\infty} (-1)^j \frac{(4k^2-1^2)\dots(4k^2-(2j-1)^2)}{j!(8t)^j}, \quad t \to \infty.$$

I managed to carry out the computation for $n = 1$, and it seems like one needs n terms of this series for α_n.

Acknowledgement. I thank Yu.M. Berezanskii for drawing my attention to the problem discussed in the paper, and M. Ismail for giving a chance to get acquainted with the manuscript of his ongoing book [10]. The work was partially supported by INTAS Research Network NeCCA 03-51-6637 and NATO Collaborative linkage grant PST. CLG. 979738.

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] G.S. Ammar and W.B. Gragg, Schur flows for orthogonal Hessenberg matrices. Hamiltonian and gradient flows, algorithms and control, *Fields Inst. Commun.* V.3 (1994), American Math. Soc., Providence, RI, 27-34.

[4] J. Baik, P. Deift and K. Johansson, On the distribution of the lenght of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* **12** (1999), 1119-1178.

[5] Yu.M. Berezanskii, The integration of semi-infinite Toda chain by means of inverse spectral problem, *Rep. in Math. Phys.* **24** No.1, (1986), 21-47.

[6] M.J. Cantero, L.Moral and L.Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, *Lin. Algebra Appl.* **362** (2003), 29-56

[7] P. Deift, L.C. Li and C. Tomei, Toda flows with infinitely many variables, *J. of Func. Anal.* **64** (1985), 358-402.

[8] L. Faybusovich and M. Gekhtman, On Schur flows, *J. Phys. A: Math. Gen.* **32** (1999), 4671-4680.

[9] G.S. Geronimo, F. Gesztesy and H. Holden, Algebro-geometric solution of the Baxter–Szegő difference equation, to appear in Comm. Math. Phys.

[10] M. E. H. Ismail, Classical and quantum orthogonal polynomials in one variable, Encyclopedia in Mathematics, Cambridge University Press, 2005.

[11] J. Moser, Three integrable Hamiltonian systems connected with isospectral deformations, *Adv. Math.* **16** (1975), 197-220.

[12] J. Moser, Finitely many mass points on the line under the influence of an exponential potential – an integrable system, *Lecture Notes in Physics: Dynamical Systems. Theory and Application*, Berlin: Springer, J. Ehlers, K. Hepp and H.A. Weidenmüller (eds), **38** (1975), 467-497.

[13] A. Mukaihira and Y. Nakamura, Schur flow for orthogonal polynomials on the unit circle and its integrable discretization, *J. of Comp. Appl. Math.* **139** (2002), 75-94.

[14] A. Mukaihira and Y. Nakamura, Integrable discretization of the modified KdV equation and applications, *Inverse Problems* **16** (2000), 413-424.

[15] I. Nenciu, Lax pairs for the Ablowitz–Ladik system via orthogonal polynomials on the unit circle, to appear in IMRN.

[16] P. Nevai and V. Totik, Orthogonal polynomials and their zeros, *Acta Sci Math.* (Szeged) **53** (1989), 99-104.

[17] V. Perival and D. Shevitz, Unitary-matrix models as exactly solvable string theories, *Phys. Rev. Lett.* **64** (1990), 1326-1329.

[18] B. Simon, "Orthogonal Polynomials on the Unit Circle, Vol. 1", AMS Colloquium Series, American Mathematical Socity, Providence, RI, 2005.

[19] B. Simon, "Orthogonal Polynomials on the Unit Circle, Vol. 2", AMS Colloquium Series, American Mathematical Socity, Providence, RI, 2005.

Institute for Low Temperature Physics and Engineering, 47, Lenin Ave, Kharkov, 61103, Ukraine

E-mail address: golinskii@ilt.kharkov.ua