Matrix-valued orthogonal polynomials on the real line: some extensions of the classical theory



2005 J. Phys. A: Math. Gen. 38 5731

(http://iopscience.iop.org/0305-4470/38/25/009)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) 5731–5749 doi:10.1088/0305-4470/38/25/009

# Matrix-valued orthogonal polynomials on the real line: some extensions of the classical theory

**L Miranian**

Department of Mathematics, University of California, Berkeley, CA 94720, USA

E-mail: luiza@math.berkeley.edu

Received 16 February 2005
Published 8 June 2005
Online at stacks.iop.org/JPhysA/38/5731

**Abstract**

In the work presented below the classical subject of orthogonal polynomials on the real line is discussed in the matrix setting. An analogue of the determinant definition of orthogonal polynomials is presented; the classical properties such as the recurrence relation, the kernel polynomials, and the Christoffel–Darboux formula are discussed. A $\tau$-function for the system of matrix-valued orthogonal polynomials on the real line is presented. Some properties of the $\tau$-functions are investigated.

PACS numbers: 02.30.Gp, 02.30.Tb

## 1. Introduction

Since the fundamental works of Akhiezer [1], Szegö [31], and many others, orthogonal polynomials have been a major tool in the analysis of many problems in mathematics, such as the rational and polynomial approximation and interpolation, the moment problem, and numerical quadrature. The development of special and important examples goes much further back, see for instance Lebedev [24]. Matrix-valued orthogonal polynomials supported on the real line are used extensively in the areas of rational approximations and in system theory, see [14]; in the Lanczos method for block matrices, see [17, 19]; in the spectral theory of the doubly infinite Jacobi matrices, see [3, 29] as well as [30]; in the analysis of sequences of polynomials satisfying higher order recurrence relations, see [6, 10] and more. Applications of matrix valued orthogonal polynomials supported on the unit circle include linear estimation theory, where finite block Toeplitz matrices need to be inverted, see [27]; the analysis of sequences of polynomials orthogonal with respect to scalar measure supported on equipotential curves in the complex plane, see [25]; frequency estimation in time series analysis and many more. Zeros of orthogonal polynomials are used in the areas of spectral analysis, digital filter design, quadrature formulae, etc. In the matrix setting, the zeros of orthogonal polynomials arise as nodes in quadrature formulae and as eigenvalues of block Jacobi matrices.

Starting with the work of Krein [22, 23] as well as [2, 5–10, 13, 16, 26, 30] there is a general theory of matrix-valued orthogonal polynomials. Some very important results of the theory of scalar-valued orthogonal polynomials, such as Favard's theorem and Markov's theorem have been extended to the matrix-valued case, see [5–7, 10, 12], and many more still need to be investigated in the new context of the matrix-valued orthogonal polynomials.

This paper is organized as follows. In section 2, notations are introduced and the matrix analogue of the determinant formula for the polynomials of the first kind on the real line is presented. Section 3 concerns orthogonality of the polynomials introduced in section 2. In section 4, the Gramm–Schmidt orthogonalization procedure is discussed in the matrix setting. The recurrence relation in the matrix case is presented in section 5. Section 6 concerns the matrix-valued version of the kernel polynomials and the Christoffel–Darboux formula. In section 7 a $\tau$-function for the system of matrix-valued polynomials on the real line is introduced. Expressions connecting polynomials of the first and second kinds with the $\tau$-function are presented and compared to these in the scalar case.

## 2. Definitions

In this section we introduce the notations and present a definition of the scalar/matrix-valued orthogonal polynomials on the real line which is a natural extension of the classical determinant definition discussed in numerous books and articles, for example, see [4].

Given a measure $\mu(\mathrm{d}x) = W(x)\,\mathrm{d}x$ with symmetric weight function $W(x) \in \mathbf{R}^{k\times k}$ for $k \geqslant 1$ and supported on the real line $x \in \mathbf{R}$, introduce

- The $n$th moment of the measure $\mu(\mathrm{d}x)\mu_n \in \mathbf{R}^{k\times k}$, where

$$\mu_n = \int x^n \mu(\mathrm{d}x) = \int x^n W(x)\,\mathrm{d}x, \qquad n = 0, 1, \ldots.$$

  Note that $\mu_n = \mu_n^*$. In this text '*' denotes transposition.
- The matrix $T_n$ for $n \geqslant 1$, where $I$ is $k \times k$ identity matrix

$$T_n = \begin{pmatrix} \mu_0 & \mu_1 & \ldots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_2 & \ldots & \mu_n & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \ldots & \mu_{2n-2} & \mu_{2n-1} \\ I & xI & \ldots & x^{n-1}I & x^n I \end{pmatrix} \in \mathbf{R}^{k(n+1)\times k(n+1)}.$$

- A Hankel matrix $H_n$ for $n \geqslant 1$

$$H_n = \begin{pmatrix} \mu_0 & \mu_1 & \ldots & \mu_{n-1} \\ \mu_1 & \mu_2 & \ldots & \mu_n \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \ldots & \mu_{2n-2} \end{pmatrix} \in \mathbf{R}^{kn\times kn}.$$

- The matrix $H$ which is the semi-infinite version of $H_n$ for $n \to \infty$.
- The vector $v_{n,2n-1}$ for $n \geqslant 1$

$$v_{n,2n-1} = (\mu_n \quad \mu_{n+1} \quad \ldots \quad \mu_{2n-1})^*.$$

  Introduce a shift operator $v_{n,2n-1}^{(m)} \equiv v_{n+m,2n-1+m}$. In section 6 this 'shift' notation will be interpreted as a certain $m$th derivative.

- In the matrix

$$H_{n+1} = \begin{pmatrix} H_n & v_{n,2n-1} \\ v^*_{n,2n-1} & \mu_{2n} \end{pmatrix}$$

denote the Schur complement of $\mu_{2n}$

$$S_n = \mu_{2n} - v^*_{n,2n-1} H_n^{-1} v_{n,2n-1}, \qquad \text{with} \quad S_0 = \mu_0. \tag{1}$$

Introduce the diagonal matrix

$$S = \operatorname{diag}[S_0, S_1, \ldots]. \tag{2}$$

Using the notations above we introduce the following definition:

**Definition 1** (Monic matrix-valued polynomials on the real line)**.** *Define a family of polynomials* $\{P_n(x)\}_{n=0}^{\infty}$ *as the Schur complement of* $x^n I$ *in the matrix* $T_n$*, i.e.*

$$P_n(x) = x^n I - [I \quad xI \quad \ldots \quad x^{n-1} I] \begin{pmatrix} \mu_0 & \mu_1 & \ldots & \mu_{n-1} \\ \mu_1 & \mu_2 & \ldots & \mu_n \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \ldots & \mu_{2n-2} \end{pmatrix}^{-1} \begin{pmatrix} \mu_n \\ \mu_{n+1} \\ \vdots \\ \mu_{2n-1} \end{pmatrix}, \tag{3}$$

*with* $P_0(x) = I$*. Denote by* $P$ *the row vector of matrix-valued polynomials*

$$P = [P_0(x), P_1(x), \ldots]. \tag{4}$$

**Note 1.** In the classical theory of scalar-valued orthogonal polynomials, monic polynomials are defined as (for example, see [1])

$$p_n(x) = \frac{\det(T_n)}{\det(H_n)}, \qquad \text{with} \quad p_0(x) = 1,$$

which is exactly what we obtain using definition (3) in the scalar case. This is because for any matrix with partitioning $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ its determinant can be computed as $\det(M) = \det(A)\det(D - CA^{-1}B)$, hence $\det(T_n) = \det(H_n)\det(P_n(x)) = \det(H_n)P_n(x)$, where $P_n$ is the one defined in (3).

**Note 2.** In definition (3) it is assumed that matrices $H_n$ are invertible for all $n \geqslant 1$, which is a restriction on the measure $\mu(\mathrm{d}x)$. In particular, all matrices $H_n$ being invertible implies that the matrices $S_n$ are invertible, since $\det(H_{n+1}) = \det(H_n)\det(S_n)$.

**Definition 2** (matrix-valued polynomials of the second kind)**.** *Define a family of matrix-valued polynomials* $\{Q_n(x)\}_{n=0}^{\infty}$ *as*

$$Q_n(x) = x \int \mu(\mathrm{d}z) \frac{P_n(z)}{x - z}, \tag{5}$$

*where the polynomials* $P_n(x)$ *are defined in (3). The polynomials* $Q_n(x)$ *are called polynomials of the second kind.*

**Note 3.** In the classical theory of scalar-valued orthogonal polynomials, the polynomials of the second kind are defined in the same way as in (5), see [28].

In what follows our matrix indexing starts from zero, i.e. $M_{i,j}$ refers to $\{i, j\}$th $k \times k$ block of matrix $M$, where $i, j \geqslant 0$.

## 3. Orthogonality via the moments of the measure

The following proposition shows that a family of monic polynomials $\{P_n(x)\}_{n=0}^{\infty}$ defined in (3) forms a set of orthogonal polynomials for any symmetric measure $\mu(\mathrm{d}x) = W(x)\,\mathrm{d}x$.

**Proposition 1.** *Let $\{P_n(x)\}_{n=0}^{\infty}$ be a family of polynomials defined in (3) and $S_n$ be defined in (1). Define an inner product on $\mathbf{L}_2(\mathbf{R}^k)$ by means of*

$$\langle P, Q\rangle = \int P^*(x)W(x)Q(x)\,\mathrm{d}x, \tag{6}$$

*then*

$$\langle P_i, P_j\rangle = \delta_{ij}S_i,$$

*for any $i, j \geqslant 0$.*

**Proof.** Observe first that for any $0 \leqslant m \leqslant n-1$

$$v_{m,m+n-1}H_n^{-1} = [\mu_m \quad \mu_{m+1} \quad \mu_{m+2} \quad \ldots \quad \mu_{m+n-1}]H_n^{-1} = [0 \quad \ldots \quad I \quad \ldots \quad 0],$$

where $I$ is at the $m$th location. Hence,

$$v_{m,m+n-1}H_n^{-1}v_{n,2n-1} = \mu_{m+n}. \tag{7}$$

It is enough to show that $P_n(x)$ is orthogonal to all $x^m I$ for $0 \leqslant m \leqslant n-1$, i.e.

$$\begin{aligned}\int x^m W(x)P_n(x)\,\mathrm{d}x &= \int x^m W(x)\big(x^n I - [I \quad xI \quad \ldots \quad x^{n-1}I]H_n^{-1}v_{n,2n-1}\big)\,\mathrm{d}x\\ &= \mu_{m+n} - [\mu_m \quad \mu_{m+1} \quad \ldots \quad \mu_{m+n-1}]H_n^{-1}v_{n,2n-1}\\ &= \mu_{m+n} - \mu_{m+n} = 0.\end{aligned}$$

This proves that $\langle P_m(x), P_n(x)\rangle = 0$ for any $m < n$. If $m = n$ then

$$\begin{aligned}\int P_n^*(x)W(x)P_n(x)\,\mathrm{d}x &= \int x^n W(x)\big(x^n I - [I \quad xI \quad \ldots \quad x^{n-1}I]H_n^{-1}v_{n,2n-1}\big)\,\mathrm{d}x\\ &= \mu_{2n} - v_{n,2n-1}^* H_n^{-1}v_{n,2n-1} = S_n.\end{aligned}$$

□

The inner product introduced in (6) is different from the one used in many papers on this subject, e.g. [9, 11, 18, 20, 30] and others. The standard inner product used is called 'left inner product'

$$\langle P, Q\rangle_L = \int P(x)W(x)Q^*(x)\,\mathrm{d}x,$$

which is different from the one defined in (6) by $\langle P, Q\rangle_R = \int P^*(x)W(x)Q(x)\,\mathrm{d}x$, called 'right inner product'.

## 4. Orthogonality via Gramm–Schmidt

A family of orthogonal polynomials (either scalar- or matrix-valued) can be obtained in at least three ways: the method of moments introduced in section 2, the familiar recursion relation and the Gramm–Schmidt orthogonalization procedure, which will be discussed in this section.

Let us obtain a family of monic matrix-valued orthogonal polynomials by performing the Gramm–Schmidt procedure on the space of matrix-valued $k \times k$ polynomials. Define the row vector

$$\Omega = [I \quad xI \quad x^2I \quad \ldots] \tag{8}$$

and a unit semi-infinite block upper triangular matrix

$$R = \begin{pmatrix} I & r_{01} & r_{02} & \cdots \\ 0 & I & r_{12} & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix}, \qquad \text{where} \quad r_{i,j} \in \mathbf{R}^{k,k}. \tag{9}$$

To find an orthogonal basis for the space spanned by $\{x^n I\}_{n=0}^{\infty}$ with respect to the given measure we perform the Gramm–Schmidt orthogonalization procedure on $\Omega$, obtaining

$$\Omega = PR, \tag{10}$$

where the matrix $R$ depends on the moments of the measure, and the vector $P$ is defined in (4). In the next proposition elements $r_{i,j}$ of the matrix $R$ are computed.

**Proposition 2.** *Let matrices $\Omega$, $P$ and $R$ be defined as in (8), (4) and (9). Assume that equation (10) holds and the polynomials in the vector $P$ are defined in (3), then*

$$S_n r_{n,n+k} = \mu_{2n+k} - v_{n,2n-1}^* H_n^{-1} v_{n,2n-1}^{(k)}, \tag{11}$$

*for all $k \geqslant 0$, where notation $v_{n,2n-1}^{(k)}$ is defined in section 2.*
*In particular, for $n = 0$*

$$r_{0,m} = \mu_0^{-1} \mu_m. \tag{12}$$

*In the matrix form,*

$$H = R^* S R, \tag{13}$$

*where matrices $S$ and $H$ are defined in section 2.*

**Proof.** From (10) it follows that

$$P_0(x) r_{0,m} + P_1(x) r_{1,m} + \cdots + P_k(x) r_{k,m} + \cdots + P_m(x) I = x^m I.$$

After multiplying the expression above by $P_n^*(x)W(x)$ from the left and integrating we obtain

$$S_n r_{n,m} = \int P_n^*(x) W(x) x^m \,\mathrm{d}x$$

and (11) easily follows from writing out $P_n^*(x)$ as defined in (3) and integrating. To prove (13) observe that

$$\left( \int \Omega^* W(x) \Omega \,\mathrm{d}x \right)_{i,j} = \int x^i W(x) x^j \,\mathrm{d}x = \mu_{i+j} = H_{i,j}$$

and

$$\left( \int P^* W(x) P \,\mathrm{d}x \right)_{i,j} = \int P_i^*(x) W(x) P_j(x) \,\mathrm{d}x = \delta_{i,j} S_i,$$

hence

$$H = \int \Omega^* W(x) \Omega \,\mathrm{d}x = R^* \left( \int P^* W(x) P \,\mathrm{d}x \right) R = R^* S R. \qquad \square$$

The following lemma expresses $H_{n+1}^{-1}$ in terms of $H_n^{-1}$, $v_{n,2n-1}$ and $S_n$ and will be very useful throughout the paper.

**Lemma 1.** *Given the following partitioning of the matrices $H_{n+1}$ and $H_{n+1}^{-1}$*

$$H_{n+1} = \begin{pmatrix} H_n & v_{n,2n-1} \\ v_{n,2n-1}^* & \mu_{2n} \end{pmatrix} \qquad \text{and} \qquad H_{n+1}^{-1} = \begin{pmatrix} A & \gamma \\ \gamma^* & \alpha \end{pmatrix},$$

*then*

$$\alpha = S_n^{-1}, \qquad \gamma = -H_n^{-1} v_{n,2n-1} S_n^{-1},$$

*and*

$$A = H_n^{-1} + H_n^{-1} v_{n,2n-1} S_n^{-1} v_{n,2n-1}^* H_n^{-1} = \left(H_n - v_{n,2n-1}\mu_{2n}^{-1} v_{n,2n-1}^*\right)^{-1}.$$

*In particular,*

$$H_{n+1}^{-1} v_{n+1,2n+1}^{(m)} = \begin{pmatrix} H_n^{-1}\left(v_{n,2n-1}^{(m+1)} - v_{n,2n-1} r_{n,n+m+1}\right) \\ r_{n,n+m+1} \end{pmatrix}. \tag{14}$$

**Proof.** The formulae above can be easily verified by direct computation. Expression (14) is obtained using (11). □

**Lemma 2.** *Let polynomials $P_n(x)$ be defined in (3). Then*

$$x^n = \sum_{i=0}^{m} P_{n-i}(x) r_{n-i,n} + [I \quad xI \quad \cdots \quad x^{n-m-1} I] H_{n-m}^{-1} v_{n-m,2(n-m)-1}^{(m)}.$$

*To illustrate,*

$$P_{n+1}(x) = x^{n+1} I - P_n(x) r_{n,n+1} - [I \quad xI \quad \cdots \quad x^{n-1} I] H_n^{-1} v_{n,2n-1}.$$

**Proof.** Follows from (14). □

## 5. Orthogonality via the recurrence relation

It is well known that matrix-valued orthogonal polynomials obey an appropriate three-term recurrence relation, for example, see [30]. In the following lemma we present expressions for the coefficients of the recursion relation in terms of the moments of the measure.

**Proposition 3.** *The monic matrix-valued orthogonal polynomials defined in (3) obey the following recurrence relation,*

$$xP_n(x) = P_{n+1}(x) + P_n(x) b_n^* + P_{n-1}(x) a_n^*, \tag{15}$$

*with*

$$a_n^* = S_{n-1}^{-1} S_n, \qquad b_n^* = u_n^n - u_{n-1}^{n-1}, \tag{16}$$

*where*

$$u^{n-1} = \begin{pmatrix} u_0^{n-1} \\ u_1^{n-1} \\ \vdots \\ u_{n-1}^{n-1} \end{pmatrix} = H_n^{-1} v_{n,2n-1}; \qquad u_n^n = S_n^{-1}\left(\mu_{2n+1} - v_{n,2n-1}^* H_n^{-1} v_{n+1,2n}\right). \tag{17}$$

**Proof.** Let the orthogonal polynomials obey the following three-term recursion relation,

$$xP_n(x) = P_{n+1} c_n^* + P_n(x) b_n^* + P_{n-1}(x) a_n^* \tag{18}$$

for some matrices $a_n, b_n$ and $c_n$. After multiplying (18) by $P_{n+1}^*(x)W(x)$ from the left and integrating one arrives at

$$\int P_{n+1}^*(x) W(x) x^{n+1}\, dx = S_{n+1} = \int P_{n+1}^*(x) W(x) x P_n(x)\, dx = S_{n+1} c_n^*,$$

hence $c_n^* = I$. The above expression can also be written in the following way:

$$\begin{aligned} S_{n+1} &= \int x P_{n+1}^*(x) W(x) P_n(x)\,\mathrm{d}x \\ &= \int (P_{n+2}^*(x) + b_{n+1} P_{n+1}^*(x) + a_{n+1} P_n^*(x)) W(x) P_n(x)\,\mathrm{d}x \\ &= 0 + 0 + \int a_{n+1} P_n^*(x) W(x) P_n(x)\,\mathrm{d}x = a_{n+1} S_n, \end{aligned}$$

implying $a_n^* = S_{n-1}^{-1} S_n$. After multiplying (18) by $P_n^*(x)W(x)$ from the left and integrating one obtains

$$\int x P_n^*(x) W(x) P_n(x)\,\mathrm{d}x = \left(\int P_n^*(x) W(x) P_n(x)\,\mathrm{d}x\right) b_n^* = S_n b_n^* = b_n S_n.$$

In order to compute $b_n$ in terms of the moments compare powers of $xI$ in the recursion relation (15), which can be written out as

$$\begin{aligned} x(x^n I - [I \;\; xI \;\; \ldots \;\; x^{n-1} I] u^{n-1}) &= (x^{n+1} I - [I \;\; xI \;\; \ldots \;\; x^n I] u^n) \\ &\quad + (x^n I - [I \;\; xI \;\; \ldots \;\; x^{n-1} I] u^{n-1}) b_n^* \\ &\quad + (x^{n-1} I - [I \;\; xI \;\; \ldots \;\; x^{n-2} I] u^{n-2}) a_n^*. \end{aligned}$$

Equating the coefficients of front of $x^n I$ leads to $b_n^* = u_n^n - u_{n-1}^{n-1}$. From the definition $u^n = H_{n+1}^{-1} v_{n+1,2n+1}$ and using lemma 1 one concludes that $r_{n,n+1} \equiv u_n^n = S_n^{-1}\big(\mu_{2n+1} - v_{n,2n-1}^* H_n^{-1} v_{n+1,2n}\big)$. □

**Note 4.** In the classical theory of scalar-valued orthogonal polynomials, the expression for $a_n$ is given by (for example, see [1])

$$a_n = \frac{\det(H_{n+1})\det(H_{n-1})}{\det(H_n)^2} = \frac{S_n}{S_{n-1}},$$

which is equivalent to our formula (16), since in the scalar case $S_n = \frac{\det(H_{n+1})}{\det(H_n)}$. Also, as observed in note 2, invertibility of the matrices $H_n$ implies invertibility of the matrices $S_n$, hence the matrices $a_n$ are always well defined in this setting.

**Note 5.** Matrix polynomials of the second kind defined in (5) satisfy the same recursion relation as these of the first kind defined in (3), since

$$\begin{aligned} Q_{n+1}(x) + Q_n(x) b_n^* + Q_{n-1}(x) a_n^* &= x \int \mu(\mathrm{d}u) \frac{P_{n+1}(u) + P_n(u) b_n^* + P_{n-1}(u) a_n^*}{x-u} \\ &= x \int \mu(\mathrm{d}u) \frac{u P_n(u)}{x-u} \\ &= x \int \mu(\mathrm{d}u) \frac{(u-x) P_n(u)}{x-u} + x^2 \int \mu(\mathrm{d}u) \frac{P_n(u)}{x-u} \\ &= x Q_n(x). \end{aligned}$$

In the matrix form relations (15) for all $n \geqslant 0$ can be written as

$$L P^*(x) = x P^*(x), \tag{19}$$

where $L$ is given by the block tridiagonal matrix

$$L = \begin{pmatrix} b_0 & I & 0 & 0 & \ldots \\ a_1 & b_1 & I & 0 & \ldots \\ 0 & a_2 & b_2 & I & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{20}$$

The following proposition establishes the connection between the matrix $L$ defined in (20) and the matrix $R$ defined in (9).

**Proposition 4.** *Let $R_k$ to be kth column of the matrix R defined in (9), where $k \geqslant 0$. Then*

$$L^* R_k = R_{k+1}, \qquad \text{or} \qquad R_k = (L^*)^k R_0. \tag{21}$$

*In particular,*

$$((L^*)^k)_{0,0} = \mu_0^{-1}\mu_k. \tag{22}$$

**Proof.** After multiplying $\Omega = PR$ by $P^*W(x)$ from the left and integrating, one obtains

$$SR = \int P^* W(x)\Omega\,\mathrm{d}x, \qquad \text{or} \qquad (SR)_{i,j} = \int P_i^*(x)W(x)x^j\,\mathrm{d}x \tag{23}$$

and

$$LSR = \int xP^*W(x)\Omega\,\mathrm{d}x, \qquad \text{or} \qquad (LSR)_{i,j} = \int P_i^*(x)W(x)x^{j+1}\,\mathrm{d}x. \tag{24}$$

From (23) and (24) it follows

$$(LSR)_{i,j} = (SL^*R)_{i,j} = (SR)_{i,j+1},$$

and since $S$ is diagonal, we conclude that $(L^*R)_{i,j} = r_{i,j+1}$, which implies (21). Expression (22) follows from the facts that $R_0$ is a column of all zeros except the identity at the block $\{0, 0\}$ and $(R_k)_0 \equiv r_{0,k} = \mu_0^{-1}\mu_k$. □

**Note 6.** In the classical theory of scalar-valued orthogonal polynomials the expression (22) can be found in Akhiezer [1].

Along with monic matrix-valued orthogonal polynomials, one can try to introduce orthonormal matrix-valued polynomials in the following fashion:

**Definition 3.** *Given a family of monic polynomials presented in (3) define a family $\{\bar{P}_n(x)\}_{n=0}^{\infty}$ by means of*

$$\bar{P}_n(x) = P_n(x)S_n^{-1/2}, \qquad \text{for} \quad n \geqslant 0. \tag{25}$$

It is easy to see that

$$\langle \bar{P}_n(x), \bar{P}_m(x)\rangle = S_n^{-1/2}\left(\int P_n^*(x)W(x)P_n(x)\,\mathrm{d}x\right)S_n^{-1/2} = S_n^{-1/2}S_nS_n^{-1/2} = I.$$

The recurrence relation for $\{\bar{P}_n(x)\}_{n=0}^{\infty}$ can be written as

$$\begin{aligned} &x\bar{P}_n(x) = \bar{P}_{n+1}(x)\bar{a}_{n+1} + \bar{P}_n(x)\bar{b}_n + \bar{P}_{n-1}(x)\bar{a}_n^*, \qquad \text{where} \\ &\bar{a}_n = S_n^{1/2}S_{n-1}^{-1/2}, \quad \bar{b}_n = \bar{b}_n^* = S_n^{-1/2}b_nS_n^{1/2}, \end{aligned} \tag{26}$$

or, in the matrix form,

$$\bar{L}\bar{P}^* = x\bar{P}^*, \qquad \text{where} \quad \bar{L} = S^{1/2}LS^{-1/2}.$$

Note that

$$\bar{L} = \bar{L}^*, \qquad \text{or} \qquad LS = SL^*.$$

In order to be able to define an orthonormal family in this fashion, the matrices $S_n$ have to be positive definite for all $n$. In general, the matrices $S_n$ being positive definite is equivalent to the weight matrix $W(x)$ being positive definite, and the reason for that is the following:

- $W(x)$ is positive definite for all $x \in \mathbf{R} \Leftrightarrow$ for any vector $v \in \mathbf{R}^k$ and $n \geqslant 0$ $[v^* P_n^*(x)] W(x)[P_n(x)v] > 0 \Leftrightarrow$ for any vector $v \in \mathbf{R}^k$ and $n \geqslant 0\,[v^* S_n v] > 0$, which implies that $S_n$ are positive definite for all $n \geqslant 0$.
- The polynomials $\{P_n(x)\}_{n=0}^{\infty}$ form a basis for the space of matrix polynomials, hence any polynomial $Q(x)$ can be written as $Q(x) = \sum_i P_i(x)\alpha_i$ for some matrices $\alpha_i$. It is easy to see that
$$\int Q^*(x)W(x)Q(x)\,\mathrm{d}x = \sum_i \alpha_i^* S_i \alpha_i.$$
The above expression implies that $S_n$ being positive definite for all $n \geqslant 0$ is equivalent to $\int Q^*(x)W(x)Q(x)\,\mathrm{d}x$ being positive definite for all polynomials $Q(x)$, which, in turn, is equivalent to the weight matrix $W(x)$ being positive definite.

Orthonormal polynomials will be used in the next section to derive the so-called kernel polynomials and the Christoffel–Darboux formula.

## 6. The Christoffel–Darboux formula

In this section a matrix-valued kernel polynomial will be introduced and the Christoffel–Darboux formula will be derived.
The following lemma introduces the matrix-valued kernel polynomial.

**Lemma 3.** *Given a family of orthonormal polynomials as defined in (25), denote the kernel polynomial of degree n to be*
$$K_n(x, y) = \sum_{i=0}^{n} \bar{P}_i(y)\bar{P}_i^*(x). \tag{27}$$
*Then*
$$K_n(x, y) = [I \quad yI \quad \ldots \quad y^n I] \begin{pmatrix} \mu_0 & \mu_1 & \ldots & \mu_n \\ \mu_1 & \mu_2 & \ldots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \ldots & \mu_{2n} \end{pmatrix}^{-1} \begin{bmatrix} I \\ xI \\ \vdots \\ x^n I \end{bmatrix}. \tag{28}$$

**Proof by induction.** To ease the notation, denote
$$Y = [I \quad yI \quad \ldots \quad y^{n-1}I]^* \qquad \text{and} \qquad X = [I \quad xI \quad \ldots \quad x^{n-1}I]^*.$$
For $n = 0$ we have $K_0(x, y) = \bar{P}_0(x)\bar{P}_0^*(y) = \mu_0^{-1}$ which agrees with formula (28). To simplify the notation denote the right-hand side of expression (28) as $RHS(n)$. For the inductive step $(n-1) \to n$ we use the notation and partitioning in lemma 1, as well as the fact that $X^* H_n^{-1} v S_n^{-1/2} = x^n S_n^{-1/2} - \bar{P}_n(x)$ to rewrite $RHS(n)$ as
$$\begin{aligned} RHS(n) &= Y^*AX + y^n\gamma^*X + Y^*\gamma_n x^n + y^n x^n \alpha \\ &= Y^*\big(H_n^{-1} + H_n^{-1}v_{n,2n-1}S_n^{-1}v_{n,2n-1}^* H_n^{-1}\big)X \\ &\quad - y^n S_n^{-1} v_{n,2n-1}^* H_n^{-1} X - Y^* H_n^{-1} v_{n,2n-1} S_n^{-1} x^n + y^n x^n S_n^{-1} \\ &= Y^* H_n^{-1} X + \big(y^n S_n^{-1/2} - \bar{P}_n(y)\big)\big(x^n S_n^{-1/2} - \bar{P}_n^*(x)\big) \\ &\quad - y^n S_n^{-1/2}\big(x^n S_n^{-1/2} - \bar{P}_n^*(x)\big) - x^n\big(y^n S_n^{-1/2} - \bar{P}_n(y)\big)S_n^{-1/2} + y^n x^n S_n^{-1} \\ &= Y^* H_n^{-1} X + \bar{P}_n(y)\bar{P}_n^*(x) = RHS(n-1) + \bar{P}_n(y)\bar{P}_n^*(x), \end{aligned}$$
which completes the proof by induction. □

**Note 7.** In the classical theory of scalar-valued orthogonal polynomials (see [4]), the kernel polynomial is given by

$$K_n(x,y) = -\det\begin{pmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-1} & I \\ \mu_1 & \mu_2 & \dots & \mu_n & xI \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n} & x^n I \\ I & yI & \cdots & y^n I & 0 \end{pmatrix},$$

which agrees with the scalar version of formula (28) derived above.

In the next lemma the matrix-valued version of the Christoffel–Darboux formula is presented. The proof of the lemma is very similar to the one in the scalar case.

**Lemma 4.** *Let a family of orthonormal polynomials $\{\bar{P}_n(x)\}_{n=0}^{\infty}$ be defined by (3) and (25). Then*

$$\sum_{m=0}^{n} \bar{P}_m(y)\bar{P}_m^*(x) = \frac{\bar{P}_n(y)\bar{a}_{n+1}^*\bar{P}_{n+1}^*(x) - \bar{P}_{n+1}(y)\bar{a}_{n+1}\bar{P}_n^*(x)}{x-y}. \tag{29}$$

**Note 8.** In the classical theory of scalar-valued monic orthogonal polynomials (see [4]) the Christoffel–Darboux identity has the following form:

$$\sum_{m=0}^{n} \frac{p_m(y)p_m(x)}{\langle p_m, p_m\rangle} = \frac{p_n(y)p_{n+1}(x) - p_{n+1}(y)p_n(x)}{\langle p_n, p_n\rangle(x-y)},$$

which agrees with the scalar version of formula (29) derived above.

## 7. A matrix-valued $\tau(t)$-function

In this section we define a $\tau(t)$-function for a system of matrix-valued orthogonal polynomials on the real line and investigate some of its properties.

Let us introduce 'times' into the measure the following way:

$$\mu_t(\mathrm{d}x) = \mathrm{e}^{\sum_{i=1}^{\infty} t_i x^i I}\mu(\mathrm{d}x), \tag{30}$$

where $I$ is the $k \times k$ identity matrix. The new moments are defined as

$$\mu_n(t) = \int x^n\, \mathrm{e}^{\sum_{i=1}^{\infty} t_i x^i I}\mu(\mathrm{d}x).$$

Observe that

$$\frac{\partial \mu_n(t)}{\partial t_m} = \int x^{n+m}\, \mathrm{e}^{\sum_{i=1}^{\infty} t_i x^i I}\mu(\mathrm{d}x) = \mu_{n+m}(t). \tag{31}$$

**Definition 4.** *Define $\tau_0(t) \equiv S_0(t)$, and for $n \geqslant 1$*

$$\tau_n(t) \equiv S_n(t) = \mu_{2n}(t) - v_{n,2n-1}^*(t)H_n^{-1}(t)v_{n,2n-1}(t), \tag{32}$$

*where $H_n(t)$ and $v_{n,2n-1}(t)$ are defined in section 2, but now with 'time' dependence.*

**Note 9.** In the classical theory of scalar-valued orthogonal polynomials, the $\tau(t)$-function is defined in the following fashion,

$$\tilde{\tau}_n(t) = \det(H_n(t)),$$

whereas our new definition in the scalar case becomes

$$\tau_n(t) = \frac{\det(H_{n+1}(t))}{\det(H_n(t))}.$$

In what follows prime '′' denotes differentiation with respect to $t_1$, in particular $\mu_n'(t) = \mu_{n+1}(t)$ and $v'_{n,2n-1} = v_{n+1,2n}$. To ease the notation in the discussion below, the 'times' $t$ will be dropped when not essential, and $v_{n,2n-1}(t)$ will be substituted with $v$, i.e.

$$v := v_{n,2n-1}(t) \qquad \text{and} \qquad v' := v_{n+1,2n}(t).$$

In theorem 1 the connection between the $\tau(t)$-function and the coefficients of the recursion relation is established. The following lemma is needed in the proof of the theorem:

**Lemma 5.**

$$\begin{pmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \\ \mu_2 & \mu_3 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} \end{pmatrix} H_n^{-1} v = \begin{pmatrix} \mu_{n+1} \\ \vdots \\ \mu_{2n-2} \\ \mu_{2n} - S_n \end{pmatrix}.$$

*This could also be written as*

$$(H_n)' H_n^{-1} v = v' - (0 \;\; 0 \;\; \cdots \;\; S_n)^*.$$

**Proof 8.** See the appendix. □

**Theorem 1.** *Given a family of monic matrix-valued orthogonal polynomials as defined in (3), which satisfies the recursion relation*

$$x P_n(x) = P_{n+1}(x) + P_n(x) b_n^* + P_{n-1}(x) a_n^*,$$

*and a $\tau(t)$-function defined in (32), then*

$$b_n^* = \frac{\partial}{\partial t_1} \ln(\tau_n(t))|_{t_1=0} \qquad \text{and} \qquad a_n^* = \tau_{n-1}(0)^{-1} \tau_n(0).$$

**Proof.** By definition, $\tau_n(t) = \mu_{2n}(t) - v^*(t) H_n^{-1}(t) v(t)$. According to (31), after differentiating with respect to $t_1$ one obtains

$$\tau_n'(t) = \mu_{2n+1} - \left(v^* H_n^{-1} v\right)' = \mu_{2n+1} - v'^* H_n^{-1} v - v^* \left(H_n^{-1}\right)' v - v^* H_n^{-1} v'.$$

Observe that

$$0 = I' = \left(H_n^{-1} H_n\right)' = \left(H_n^{-1}\right)' H_n + H_n^{-1} H_n',$$

hence

$$\left(H_n^{-1}\right)' = -H_n^{-1} H_n' H_n^{-1}.$$

Recall from (17) that $\mu_{2n+1} - v^* H_n^{-1} v' = S_n u_n^n$, implying that

$$\begin{aligned} \tau_n'(t) &= S_n(t) u_n^n(t) - v'^* H_n^{-1} v - v^* \left(H_n^{-1}\right)' v \\ &= S_n(t) u_n^n(t) - v'^* H_n^{-1} v - v^* H_n^{-1} H_n' H_n^{-1} v \\ &= S_n(t) u_n^n(t) - \left(v'^*(t) - v^* H_n^{-1} H_n'\right) H_n^{-1} v \\ &= S_n(t) u_n^n(t) - [0 \;\; \ldots \;\; 0 \;\; S_n] H_n^{-1} v \\ &= S_n(t) u_n^n(t) - S_n(t) u_{n-1}^{n-1}(t) = S_n(t)\left(u_n^n(t) - u_{n-1}^{n-1}(t)\right). \end{aligned}$$

In the fourth equation lemma 5 was used. Observe that one could use

$$\mu_{2n+1}(t) - v'^{*}(t)H_n^{-1}(t)v(t) = \left(u_n^n(t)\right)^* S_n(t),$$

and obtain $\tau_n'(t) = \left(u_n^n(t) - u_{n-1}^{n-1}(t)\right)^* S_n(t)$. Thus,

$$\tau_n'(t) = \tau_n(t)\left(u_n^n(t) - u_{n-1}^{n-1}(t)\right) = \left(u_n^n(t) - u_{n-1}^{n-1}(t)\right)^* \tau_n(t),$$

which implies that

$$\frac{\partial}{\partial t_1}\ln(\tau_n(t))|_{t_1=0} = \tau_n(t)^{-1}\frac{\partial}{\partial t_1}\tau_n(t)|_{t_1=0} = \left(u_n^n(0) - u_{n-1}^{n-1}(0)\right) = b_n^*.$$

From (16) we can see that

$$a_n = \tau_n(0)\tau_{n-1}^{-1}(0),$$

which concludes the proof of the theorem. □

**Note 10.** In the classical theory of scalar-valued orthogonal polynomials, the expression for $b_n$ is

$$b_n = \frac{\partial}{\partial t_1}\ln\left(\frac{\det(H_{n+1}(t))}{\det(H_n(t))}\right).$$

With our definition of $\tau_n(t) = \frac{\det(H_{n+1}(t))}{\det(H_n(t))}$, the above expression is equivalent to that obtained in the above theorem.

The next several lemmas are used in the proof of theorem 2, where the connection between the monic matrix-valued orthogonal polynomials as defined in (3) and (5) and the $\tau$-function as defined in (32) is established. To ease the notation introduce $X = [I \;\; xI \;\; \ldots \;\; x^{n-1}I]^*$; denote $v^{(i)}$ and $H_n^{(i)}$ to be the $i$th derivatives with respect to $t_1$. Note that the 'shift' notation $v_{n,2n-1}^{(i)}$ introduced in section 2 can now be interpreted as the $i$th derivative of $v_{n,2n-1}$ with respect to $t_1$.

**Lemma 6.** *For any $x \in \mathbf{R}$*

$$\left(I - \frac{H_n' H_n^{-1}}{x}\right)^{-1}\left(v - \frac{v'}{x}\right) = v + \xi$$

*where*

$$\xi = X\xi_0, \qquad \xi_0 = (\xi)_{0,0} \qquad \text{and} \qquad P_n^*(x,t)\xi_0 + S_n(t) = 0.$$

**Proof.** See the appendix. □

**Lemma 7.** *Let $Q_n(x,t)$ be defined as in (5). Then*

$$\begin{aligned}
Q_n(x,t) &= x\int \mu(\mathrm{d}u)\frac{P_n(u,t)}{x-u} \\
&= \int \mu(\mathrm{d}u)\frac{u^n - [1\; u \cdots u^{n-1}]H_n^{-1}v}{1-u/x} \\
&= \int \mu(\mathrm{d}u)u^n\sum_{i=0}^{\infty}(u/x)^i - \int \mu(\mathrm{d}u)[1\; u\cdots u^{n-1}]\sum_{i=0}^{\infty}(u/x)^i H_n^{-1}v \\
&= \sum_{i=0}^{\infty}\frac{\mu_{n+i}}{x^i} - \left(\sum_{i=0}^{\infty}\frac{v_{0,n-1}^{*(i)}}{x^i}\right)H_n^{-1}v \\
&= \sum_{i=0}^{\infty}\frac{\mu_{n+i} - v_{0,n-1}^{*(i)}H_n^{-1}v}{x^i} = \frac{1}{x^n}\sum_{i=0}^{\infty}\frac{r_{n,n+i}^*}{x^i}S_n,
\end{aligned}$$

*where* $r^*_{n,m}$ *is defined in (11). Denote*

$$R(n,x)=\sum_{i=0}^{\infty}\frac{r^*_{n,n+i}}{x^i}, \tag{33}$$

*hence*

$$Q_n(x,t)=\frac{1}{x^n}R(n,x)S_n(t). \tag{34}$$

**Lemma 8.** *For any* $x\in\mathbf{R}$ *let us consider the equation*

$$\left(\sum_{i=0}^{\infty}\frac{H_n^{(i)}}{x^i}\right)^{-1}\left(\sum_{i=0}^{\infty}\frac{v^{(i)}}{x^i}\right)=H_n^{-1}(v+w). \tag{35}$$

*Then*

$$w=[0\ \ 0\cdots\ \ w^*_{n-1}]^*;\quad w_{n-1}=\frac{R(n,x)S_n-U}{x}, \tag{36}$$

*where*

$$U=\sum_{i=0}^{\infty}\frac{v^{*(i)}H_n^{-1}w}{x^i}. \tag{37}$$

**Proof.** See the appendix. □

**Lemma 9.** *Suppose* $U$ *and* $R(n,x)$ *are as defined in (37) and (33), then*

$$U=(xR(n-1,x)-xI)w_{n-1}.$$

**Proof.** See the appendix. □

**Theorem 2.** *Let* $\tau_n(t)$ *be as defined in (32).*

- *Let* $\{P_n(x,t)\}_{n=0}^{\infty}$ *be a family of monic matrix-valued orthogonal polynomials as defined in (3) with the space variable* $x$ *and 'time' dependence* $t$*. Then*

$$P_{n+1}(x,t)=xP_n(x,t)\tau_n^{-1}(t)\tau_n(t-[x^{-1}]), \tag{38}$$

*where*

$$\mu_n(t-[x^{-1}])=\int z^n\,\mathrm{e}^{\sum_{i=1}^{\infty}\left(t_i-\frac{x^{-i}}{i}\right)z^i}W(z)\,\mathrm{d}z$$
$$=\mu_n(t)-\frac{\mu_{n+1}(t)}{x}.$$

- *Let* $\{Q_n(x,t)\}_{n=0}^{\infty}$ *be a family of matrix-valued orthogonal polynomials as defined in (5) with the space variable* $x$ *and 'time' dependence* $t$*. Then*

$$xQ_{n+1}(x,t)=Q_n(x,t)\tau_n^{-1}(t)\tau_{n+1}(t+[x^{-1}]), \tag{39}$$

*where*

$$\mu_n(t+[x^{-1}])=\int z^n\,\mathrm{e}^{\sum_{i=1}^{\infty}\left(t_i+\frac{x^{-i}}{x}\right)z^i}W(z)\,\mathrm{d}z$$
$$=\sum_{i=0}^{\infty}\frac{\mu_{n+i}(t)}{x^i}.$$

**Proof.**

- Observe that

$$H_n(t-[x^{-1}]) = H_n(t) - \frac{H_n'}{x} \qquad \text{and} \qquad v(t-[x^{-1}]) = v - \frac{v'}{x}.$$

Then

$$\begin{aligned}
\tau_n(t-[x^{-1}]) &= \mu_{2n}(t-[x^{-1}]) - v^*(t-[x^{-1}])H_n^{-1}(t-[x^{-1}])v(t-[x^{-1}]) \\
&= \mu_{2n} - \frac{1}{x}\mu_{2n+1} - \left(v - \frac{v'}{x}\right)^* \left(H_n - \frac{H_n'}{x}\right)^{-1} \left(v - \frac{v'}{x}\right) \\
&= \mu_{2n} - \frac{1}{x}\mu_{2n+1} - \left(v - \frac{v'}{x}\right)^* H_n^{-1} \left(I - \frac{H_n' H_n^{-1}}{x}\right)^{-1} \left(v - \frac{v'}{x}\right) \\
&= \mu_{2n} - \frac{1}{x}\mu_{2n+1} - \left(v - \frac{v'}{x}\right)^* H_n^{-1}(v+\xi) \\
&= \mu_{2n} - v^* H_n^{-1} v - \frac{1}{x}\mu_{2n+1} + \frac{v'^* H_n^{-1} v}{x} - v^* H_n^{-1}\xi + \frac{v'^* H_n^{-1}\xi}{x} \\
&= S_n - \frac{u_n^{n*} S_n}{x} + v^* H_n^{-1} X (P_n^*)^{-1} S_n - \frac{v'^* H_n^{-1} X (P_n^*)^{-1} S_n}{x} \\
&= \left(P_n^* - \frac{u_n^{n*} P_n^*}{x} - \frac{v'^* H_n^{-1} X}{x} + v^* H_n^{-1} X\right)(P_n^*)^{-1} S_n \\
&= \frac{\left(x^{n+1} I - u_n^{n*} P_n^* - v'^* H_n^{-1} X\right)(P_n^*)^{-1} S_n}{x} \\
&= \frac{P_{n+1}^*(x,t)(P_n^*(x,t))^{-1}\tau_n(t)}{x},
\end{aligned}$$

where lemma 6 was used in the fourth equation and lemma 2 was used in the ninth equation.

- Observe that

$$H_n(t+[x^{-1}]) = \sum_{i=0}^{\infty} \frac{H_n^{(i)}(t)}{x^i} \qquad \text{and} \qquad v(t+[x^{-1}]) = \sum_{i=0}^{\infty} \frac{v^{(i)}(t)}{x^i}.$$

Then

$$\begin{aligned}
\tau_n(t+[x^{-1}]) &= \mu_{2n}(t+[x^{-1}]) - v^*(t+[x^{-1}])H_n^{-1}(t+[x^{-1}])v(t+[x^{-1}]) \\
&= \sum_{i=0}^{\infty} \frac{\mu_{n+i}(t)}{x^i} - \sum_{i=0}^{\infty} \frac{v^{*(i)}(t)}{x^i} H_n^{-1}(v+w) \\
&= \sum_{i=0}^{\infty} \frac{\mu_{n+i}(t) - v^{*(i)} H_n^{-1} v}{x^i} - \sum_{i=0}^{\infty} \frac{v^{*(i)}(t)}{x^i} H_n^{-1} w \\
&= R(n,x)S_n(t) - U,
\end{aligned}$$

where lemma 8 was used in the second equation. From lemma 8 we know that $R(n,x)S_n(t) - U = xw_{n-1}$, hence

$$\tau_n(t+[x^{-1}]) = xw_{n-1}.$$

Using the result of lemma 9 write

$$R(n,x)S_n(t) - U = R(n,x)S_n - xR(n-1,x)w_{n-1} + xw_{n-1} = xw_{n-1},$$

which implies

$$R(n,x)S_n = xR(n-1,x)w_{n-1} = xR(n-1,x)S_{n-1}S_{n-1}^{-1}w_{n-1}.$$

Using the result of lemma 7 the above expression becomes

$$Q_n(x,t)x^n = Q_{n-1}(x,t)x^{n-1}S_{n-1}^{-1}\tau_n(t+[x^{-1}]),$$

implying

$$Q_n(x,t)x = Q_{n-1}(x,t)\tau_{n-1}^{-1}(t)\tau_n(t+[x^{-1}])$$

which concludes the proof of the theorem. □

**Note 11.** In the classical theory of scalar-valued orthogonal polynomials, expression (38) has the following form,

$$p_n(x,t) = x^n \frac{\tilde{\tau}_n(t-[x^{-1}])}{\tilde{\tau}_n(t)}, \tag{40}$$

where $\tilde{\tau}_n(t) = \det(H_n(t))$, see [21, 28]. Observe that in the scalar case (38) is equivalent to (40), since

$$\begin{aligned} P_{n+1}(x,t) &= xP_n(x,t)\tau_n^{-1}(t)\tau_n(t-[x^{-1}]) \\ &= x^2P_{n-1}(x,t)\tau_{n-1}^{-1}(t)\tau_{n-1}(t-[x^{-1}])\tau_n^{-1}(t)\tau_n(t-[x^{-1}]) \\ &= \vdots \\ &= x^{n+1}(\tau_n(t)\cdots\tau_0(t))^{-1}(\tau_n(t-[x^{-1}])\cdots\tau_0(t-[x^{-1}])). \end{aligned}$$

Using the facts that

$$\tau_n(t) = \frac{\det(H_{n+1}(t))}{\det(H_n(t))} \qquad \text{and} \qquad \tau_0(t) = S_0(t) = \det(H_1(t)) = \mu_0(t)$$

the above expression becomes

$$\begin{aligned} P_{n+1}(x,t) &= x^{n+1}\left(\frac{\det(H_{n+1}(t))}{\det(H_n(t))}\cdots\frac{\det(H_2(t))}{\det(H_1(t))}\det(H_1(t))\right)^{-1} \\ &\quad\times\left(\frac{\det(H_{n+1}(t-[x^{-1}]))}{\det(H_n(t-[x^{-1}]))}\cdots\frac{\det(H_2(t-[x^{-1}]))}{\det(H_1(t-[x^{-1}]))}\det(H_1(t-[x^{-1}]))\right) \\ &= x^{n+1}\frac{\det(H_{n+1}(t-[x^{-1}])}{\det(H_{n+1}(t))} = x^{n+1}\frac{\tilde{\tau}_{n+1}(t-[x^{-1}])}{\tilde{\tau}_{n+1}(t)}. \end{aligned}$$

The scalar-valued analogue of expression (39) is

$$q_n(x,t) = x^{-n}\frac{\tilde{\tau}_{n+1}(t+[x^{-1}])}{\tilde{\tau}_n(t)},$$

and its equivalence to (39) is proved similarly.

The next proposition is a collection of facts about the recursion relation coefficients.

**Proposition 5.** *Let $\{P_n(x,t)\}_{n=0}^{\infty}$ be a family of monic orthogonal matrix-valued polynomials as defined in (3) with 'time' dependent moments. Let $a_n$ and $b_n$ be the coefficients of the recursion relation (15) with 'time' dependence, and let $\partial/\partial t_1$ be denoted by '′'. Then*

*(i)* $P'_{n+1}(x,t) = -P_n(x,t)a^*_{n+1}$;
*(ii)* $\left(u^n_n\right)' = \tau_n^{-1}(t)\tau_{n+1}(t) = S_n^{-1}(t)S_{n+1}(t) = a^*_{n+1}$;
*(iii)* $(b^*_n)' = a^*_{n+1} - a^*_n$;
*(iv)* $(a^*_n)' = a^*_nb^*_n - b^*_{n-1}a^*_n$.

**Proof.**

(i) Denote $X = [I \quad xI \quad \ldots \quad x^{n-1}I]^*$, then

$$\begin{aligned}
P'_{n+1}(x,t) &= -[X^* x^n I]\big(H_{n+1}^{-1} v_{n+1,2n+1}\big)' \\
&= -[X^* x^n I]\big(\big(H_{n+1}^{-1}\big)' v_{n+1,2n+1} + H_{n+1}^{-1}(v_{n+1,2n+1})'\big) \\
&= -[X^* x^n I]\big(-H_{n+1}^{-1} H'_{n+1} H_{n+1}^{-1} v_{n+1,2n+1} + H_{n+1}^{-1} v'_{n+1,2n+1}\big) \\
&= -[X^* x^n I]\left(-H_{n+1}^{-1} v'_{n+1,2n+1} + H_{n+1}^{-1}\begin{pmatrix}0\\0\\ \vdots \\ S_n\end{pmatrix} + H_{n+1}^{-1} v'_{n+1,2n+1}\right) \\
&= -[X^* x^n I] H_{n+1}^{-1}(0 \;\; 0 \;\; \cdots \;\; S_n)^*.
\end{aligned}$$

Using partition and notation from lemma 1 we conclude that

$$\begin{aligned}
-[X^* x^n I] H_{n+1}^{-1}(0 \;\; 0 \;\; \cdots \;\; S_n)^* &= -[X^* x^n I]\begin{pmatrix}-H_n^{-1} v_{n,2n-1} S_n^{-1} S_{n+1} \\ S_n^{-1} S_{n+1}\end{pmatrix} \\
&= -\big(x^n - X^* H_n^{-1} v_{n,2n-1}\big) a_{n+1}^* = -P_n a_{n+1}^*.
\end{aligned}$$

(ii) Denote $e_n = [0 \;\; 0 \;\; \cdots \;\; I]$. By the reasoning similar to that above one arrives at

$$\big(u_n^n\big)' = e_n\big(H_{n+1}^{-1} v_{n+1,2n+1}\big)' = e_n H_{n+1}^{-1}(0 \;\; 0 \;\; \cdots \;\; S_n)^* = S_n^{-1} S_{n+1}.$$

(iii) Follows from the previous part and the fact that $b_n^* = u_n^n - u_{n-1}^{n-1}$.

(iv) $a_{n+1}^* = S_n^{-1} S_{n+1}$, hence $S_n a_{n+1}^* = S_{n+1}$. After differentiating both sides we obtain $S'_n a_{n+1}^* + S_n (a_{n+1}^*)' = S'_{n+1}$. Since $S'_n = S_n b_n^*$, we obtain

$$(a_{n+1}^*)' = a_n^* b_n^* - b_{n-1}^* a_n^*.$$

□

**Note 12.** It is only natural to call $(iii)$ and $(iv)$ in the above proposition the non- Abelian Toda equations, see [15].

## Acknowledgment

The author is very grateful to Professor F A Grünbaum for many invaluable discussions on this subject.

## Appendix

**Proof of lemma 5.** Using identity (7) one arrives at

$$\begin{aligned}
\begin{pmatrix}\mu_1 & \mu_2 & \ldots & \mu_n \\ \mu_2 & \mu_3 & \ldots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \ldots & \mu_{2n-1}\end{pmatrix} H_n^{-1} v_{n,2n-1} &= \begin{pmatrix} v_{1,n}^* H_n^{-1} v_{n,2n-1} \\ v_{2,n+1}^* H_n^{-1} v_{n,2n-1} \\ \vdots \\ v_{n,2n-1}^* H_n^{-1} v_{n,2n-1}\end{pmatrix} \\
&= \begin{pmatrix}\mu_{n+1} \\ \mu_{n+2} \\ \vdots \\ \mu_{2n} - S_n\end{pmatrix} = v'_{n,2n-1} - \begin{pmatrix}0 \\ 0 \\ \vdots \\ S_n\end{pmatrix}.
\end{aligned}$$

□

**Proof of lemma 6.** Suppose

$$\left(I - \frac{H_n' H_n^{-1}}{x}\right)^{-1}\left(v - \frac{v'}{x}\right) = v + \xi$$

for some $\xi$, then using lemma 5 one obtains

$$v - \frac{v'}{x} = \left(I - \frac{H_n' H_n^{-1}}{x}\right)(v+\xi) = v + \xi - \frac{H_n' H_n^{-1} v}{x} - \frac{H_n' H_n^{-1}\xi}{x}$$
$$= v + \xi - \frac{v'}{x} + \frac{1}{x}(0 \;\; 0 \;\; \cdots \;\; S_n)^* - \frac{H_n' H_n^{-1}\xi}{x},$$

which implies

$$H_n' H_n^{-1}\xi = x\xi + (0 \;\; 0 \;\; \cdots \;\; S_n)^*. \tag{A.1}$$

Note that

$$H_n' H_n^{-1} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ v^* H_n^{-1} \end{pmatrix}, \tag{A.2}$$

where $e_i = [0 \cdots I \cdots 0]$ and $I$ is at the $i$th location. After applying (A.2) to (A.1) we obtain the following equation,

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ v^* H_n^{-1}\xi \end{pmatrix} = \begin{pmatrix} \xi_0 x \\ \xi_1 x \\ \vdots \\ w_{n-1}x + S_n \end{pmatrix}, \tag{A.3}$$

where $\xi_i$ denotes the $i$th element of $\xi$. Expression (A.3) implies that $\xi_i = x^i \xi_0$ for $i = 0, \ldots, n-1$, and

$$v^* H_n^{-1} X \xi_0 = x^n \xi_0 + S_n,$$

leading to

$$0 = \left(x^n - v^* H_n^{-1} X\right)\xi_0 + S_n = P_n^*(x,t)\xi_0 + S_n,$$

which concludes the proof of the lemma. □

**Proof of lemma 7.** Rewriting expression (11) and keeping in mind that for any vector $\psi$

$$v_{i,n-1+i}^* H_n^{-1}\psi = \psi_i,$$

for all $i < n$, and $v_{n+m,2n-1+m} = v^{(m)}$ for all $m \geqslant 0$ one obtains

$$v^{*(i)} H_n^{-1} v = v_{n+i,2n+i-1}^* H_n^{-1} v = \mu_{n+i} - r_{n,n+i}^* S_n. \tag{A.4}$$

Rearranging (35) one arrives at

$$\sum_{i=0}^{\infty} \frac{v^{(i)}}{x^i} = \left(\sum_{i=0}^{\infty} \frac{H_n^{(i)}}{x^i}\right) H_n^{-1}(v+w) = \sum_{i=0}^{\infty} \frac{H_n^{(i)} H_n^{-1}(v+w)}{x^i},$$

or, element-wise

$$\begin{cases} \sum_{i=0}^{\infty} \frac{\mu_{n+i}}{x^i} = \sum_{i=0}^{\infty} \frac{\mu_{n+i}}{x^i} - \frac{R(n,x)S_n}{x^n} + \sum_{i=0}^{n-1} \frac{w_i}{x^i} + \frac{U}{x^n} \\ \sum_{i=0}^{\infty} \frac{\mu_{n+1+i}}{x^i} = \sum_{i=0}^{\infty} \frac{\mu_{n+1+i}}{x^i} - \frac{R(n,x)S_n}{x^{n-1}} + \sum_{i=0}^{n-2} \frac{w_{i+1}}{x^i} + \frac{U}{x^{n-1}} \\ \vdots \\ \sum_{i=0}^{\infty} \frac{\mu_{2n-1+i}}{x^i} = \sum_{i=0}^{\infty} \frac{\mu_{2n-1+i}}{x^i} - \frac{R(n,x)S_n}{x} + w_{n-1} + \frac{U}{x}. \end{cases}$$

The last equation implies $w_{n-1} = \frac{R(n,x)S_n - U}{x}$. Substituting this into the previous one we obtain that $w_{n-2} = 0$. By continuing this process we arrive at $w_i = 0$ for all $0 \leqslant i \leqslant n-2$, which concludes the proof of the lemma. □

**Proof of lemma 8.** Using the result of lemma 7 and the notation from lemma 1 we conclude that

$$H_n^{-1} w = \begin{pmatrix} -H_{n-1}^{-1} v_{n-1,2n-3} \\ I \end{pmatrix} S_{n-1}^{-1} w_{n-1}.$$

To ease the notation, denote $p = v_{n-1,2n-3}$, then

$$\begin{aligned} \left( \sum_{i=0}^{\infty} \frac{v^{(i)}}{x^i} \right) H_n^{-1} w &= \left[ \sum_{i=0}^{\infty} \frac{p^{*(i+1)}}{x^i} \sum_{i=0}^{\infty} \frac{\mu_{2n-1+i}}{x^i} \right] \begin{pmatrix} -H_{n-1}^{-1} p \\ I \end{pmatrix} S_{n-1}^{-1} w_{n-1} \\ &= \left( \sum_{i=0}^{\infty} \frac{\mu_{2n-1+i} - p^{*(i+1)} H_{n-1}^{-1} p}{x^i} \right) S_{n-1}^{-1} w_{n-1} \\ &= (xR(n-1,x) - xI) w_{n-1}, \end{aligned}$$

which concludes the proof of the lemma. □

## References


[1] Akhiezer N I 1965 *The Classical Moment Problem and Some Related Questions in Analysis* (Edinburgh: Oliver and Boyd) (Engl. Transl.)

[2] Aptekarev A I and Nikishin E M 1984 The scattering problem for discrete Sturm–Liouville operator *Math USSR-Sb* **49** 325–55

[3] Berezanskii Yu M 1968 *Expansions of Eigenfunctions of Self-adjoined Operators (Transl. Math. Monographs* vol 17*)* (Providence, RI: American Mathematical Society)

[4] Chihara T S 1978 *An Introduction to Orthogonal Polynomials* (London: Gordon and Breach)

[5] Duran A J 1997 A generalization of Favard's theorem for polynomials satisfying a recurrence relation *J. Approx. Theory* **74** 7818–29

[6] Duran A J 1995 On orthogonal polynomials with respect to positive definite matrix of measures *Can. J. Math.* **47** 88–112

[7] Duran A J 1996 Markov's theorem for orthogonal matrix polynomials *Can. J. Math.* **48** 1180–95

[8] Duran A J 1999 Ratio asymptotics for orthogonal matrix polynomials *J. Approx. Theory* **100** 304–44

[9] Duran A J 1997 Matrix inner product having a matrix symmetric second order differential operator *Rocky Mount. J. Math.* **27** 585–600

[10] Duran A J and Van Assche W 1995 Orthogonal matrix polynomials and higher order recurrence relations *Linear Algebra Appl.* **219** 261–80

[11] Duran A J and Grünbaum F A 2004 Orthogonal matrix polynomials satisfying second order differential equations *Int. Math. Res. Not.* **10** 461–84

[12] Duran A and Lopez P 1996 Orthogonal matrix polynomials: zeros and blumenthal's theorem *J. Approx. Theory* **84** 96–118

[13] Duran A J and Polo B 2002 Gaussian formulae for matrix weights *Linear Algebra Appl.* **355** 119–46

[14] Fuhrmann P A 1988 Orthogonal matrix polynomials and system theory *Rend. Sem. Mat. Univ. Politec. Torino* 68–124

[15] Gekhtman M 1998 Hamiltonian structure of non-Abelian Toda lattice *Lett. Math. Phys.* **46** 189–205

[16] Geronimo J S 1982 Scattering theory and matrix orthogonal polynomials on the real line *Circuits Syst. Signal Process.* **1** 471–95

[17] Golub G and Loan C V 1989 *Matrix Computations* 2nd edn (Baltimore: The Johns Hopkins University Press)

[18] Grünbaum F A 2003 Matrix-valued Jacobi polynomials *Bull. Sci. Math.* **127** 207–14

[19] Golub G and Underwood R 1977 The block Lanczos methods for computing eigenvalues *Mathematical Software III* ed J R Rice (New York: Academic) pp 364–77

[20] Grünbaum F A, Pacharoni I and Tirao J 2003 Matrix-valued orthogonal polynomials of the Jacobi type *Indag. Mathem. V.* **14** 353–66

[21] Haine L and Horozov E 1993 Toda orbits of Laguerre polynomials and representations of the Virasoro algebra *Bull. Sci. Math.* **117** 485–518

[22] Krein M G 1971 *Fundamental Aspects of the Representation Theory of Hermitian Operators with Deficiency Index (m, m) (AMS Translations* Series 2, vol 97*)* (Providence, RI: American Mathematical Society) pp 75–143

[23] Krein M G 1949 Infinite J-matrices and a matrix moment problem *Dokl. Akad. Nauk SSSR* **69** 125–8

[24] Lebedev N N 1965 *Special Functions and their Applications* (Englewood, NJ: Prentice-Hall)

[25] Marcellán F and Gonzàles I R 1989 A class of matrix orthogonal polynomials on the unit circle *Linear Algebra Appl.* **121** 233–41

[26] Marcellán F and Sansigre G 1993 On a class of matrix orthogonal polynomials on the real line *Linear Algebra Appl.* **181** 97–109

[27] Mignolet M P 1989 Matrix polynomials orthogonal on the unit circle and accuracy of autoregressive models *J. Comp. Appl. Math.* **62** 229–38

[28] van Moerbeke P 2001 Integrable lattices: random matrices and random permutations *Random Matrices and Their Applications* MSRI publications p 40

[29] Nikishin E M 1984 Discrete Sturm–Liouville operators and some problems of function theory *Trudy Sem. Petrovsk.* **10** 3–77
Nikishin E M 1986 *J. Sov. Math.* **35** 2679–744

[30] Sinap A and Van Assche W 1996 Orthogonal matrix polynomials and applications *J. Comput. Appl. Math.* **66** 27–52

[31] Szegö G 1975 *Orthogonal Polynomials* 4th edn (Providence, RI: American Mathematical Society)