# Matrix valued orthogonal polynomials

by

Luisa Miranian

B.A. (UC Berkeley) 1999

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

MATHEMATICS

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor F. Alberto Grünbaum, Chair
Professor James Sethian
Professor Deborah Nolan

Spring 2005

# Matrix valued orthogonal polynomials

# Abstract

Matrix valued orthogonal polynomials

by

Luisa Miranian
Doctor of Philosophy in MATHEMATICS

University of California at Berkeley

Professor F. Alberto Grünbaum, Chair

In the work presented below the classical subject of orthogonal polynomials is discussed in the matrix setting. An analog of the determinant definition of orthogonal polynomials is presented for the cases of the real line and the unit circle; the classical properties such as the recurrence relation, the kernel polynomials, the Christoffel-Darboux formulas 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. Also, Bochner's problem is considered in the matrix setting, and some equivalent formulations are derived.

Professor F. Alberto Grünbaum
Dissertation Committee Chair

To my family and friends,

# Acknowledgements

I am very grateful to my advisor Professor Alberto Grünbaum for being an exceptional mentor, for his support and encouragement, for sharing his wisdom and knowledge in matters scientific and non-scientific.

# Contents

# Chapter 1

# Matrix valued orthogonal polynomials

## 1.1 An overview

Since the fundamental work of Szegö [51], Akhiezer [1] and many others, orthogonal polynomials have been a major tool in the analysis of many problems in mathematics and physics, such as the moment problem, numerical quadrature, rational and polynomial interpolation and approximation, and applications of these techniques in engineering problems. The development of special and important examples goes much further back, see for instance Lebedev [44]. Matrix valued orthogonal polynomials supported on the real line are extensively used in the areas of rational approximations and in system theory, see [25]; in the Lanczos method for block matrices, see [29, 30]; in the spectral theory of the doubly infinite Jacobi matrices, see [5, 49] as well as [50]; in the analysis of sequences of polynomials satisfying higher order recurrence relations, see [16, 11] and more. Applications of matrix valued orthogonal polynomials supported on the unit circle include linear estimation theory, where finite block Toeplitz covariance matrix of a multivariate stationary stochastic process needs to be inverted, see [47]; the analysis of sequences of polynomials orthogonal with respect to scalar measure supported on equipotential curves in the complex plane, see [45]; frequency estimation of a stationary harmonic process $(X_n)$ in time series analysis, where

$$X_n = \sum_{k=1}^{n} [A_k \cos nw_k + B_k \sin nw_k] + Z_n,$$

and $(A_k)$, $(B_k)$ are matrices of dimension $p$ and $Z_n$ is a white noise. The frequencies $w_k$ are unknown and need to be estimated from the data. They can be given in terms of zeros of matrix valued orthogonal polynomials associated with some purely discrete measure supported on the unit circle, see [50]. Zeros of orthogonal polynomials are used in the areas of spectral analysis, digital filter design, quadrature formulas, etc. In the matrix setting, the zeros of orthogonal polynomials arise as nodes in quadrature formulas and as eigenvalues of block Jacobi matrices. Another application of orthogonal polynomials is the subject of rational approximations. Truncating a continued fraction gives an approximant for the function to which it converges and the link with orthogonal polynomials is that continued fractions are essentially equivalent to a three term recurrence relations, and orthogonal polynomials on the real line are known to satisfy such a recurrence.

Starting with the work of M. G. Krein [41, 42] as well as [3, 10, 11, 12, 13, 14, 23, 16, 27, 46, 50] there is a general theory of matrix valued orthogonal polynomials. Some very important results of the theory of scalar valued orthogonal polynomials, like Favard's theorem and Markov's theorem have been extended to the matrix valued case, see [10, 11, 12, 21, 16], and many more still need to be investigated in the new context of the matrix valued orthogonal polynomials.

The systems of orthogonal polynomials associated with the names of Hermite, Laguerre, Bessel and Jacobi (including the special cases named after Chebyshev, Legendre, and Gegenbauer) are the most extensively studied and widely applied systems. These four families of orthogonal polynomials are called collectively the "*classical orthogonal polynomials.*" There are many characterizations of the classical orthogonal polynomials, some of them are listed below.

- *Second order differential equation.* In 1929, Bochner posed a problem (called *Bochner's problem* below) of determining all families of scalar valued orthogonal polynomials that are eigenfunctions of some fixed second order differential operator. This problem was solved by Bochner in the original paper [5], and was rediscovered many times again, for example by Grünbaum and Haine in [32]. The only families of orthogonal polynomials that are eigenfunctions of some fixed second order differential operator are the classical ones with the classical weights:

  $$\text{Hermite weight: } W(x) = e^{-x^2},\ x \in [-\infty, \infty];$$

  $$\text{Laguerre weight: } W(x) = x^{\alpha}e^{-x},\ x \in [0, \infty],\ \alpha > -1;$$

  $$\text{Jacobi weight: } W(x) = (1+x)^{\alpha}(1-x)^{\beta},\ x \in [-1, 1],\ \alpha, \beta > -1.$$

  The orthogonality measure for Bessel polynomials is in the complex plane, which makes Bessel polynomials somewhat different from the ones above. A signed measure supported on the real line was found by Duran, see [15].

  In the matrix case Bochner's problem was first considered in [14]. Examples of such families of matrix polynomials can be found in [35, 36, 37, 38, 18]. For example, in [18] the matrix inner product and the symmetric second order differential operator $\mathbf{D}$ are defined as

  $$\langle P, Q \rangle = \int P(x)W(x)Q^*(x)dx$$

  $$\mathbf{D} = D^2A_2(x) + D^1A_1(x) + D^0A_0(x),$$

where $A_2(x),\ A_1(x),\ A_0(x)$ are matrix polynomials (which do not depend on $n$) of degrees not bigger than 2, 1 and 0 respectively and $D$ stands for the usual differential operator. The following weights were discovered

- Hermite weight

$$W(x) = \begin{cases} e^{-x^2} e^{Ax} e^{A^* x} \\ e^{-x^2} e^{Ax^2} e^{A^* x^2}; \end{cases}$$

- Laguerre weights

$$W(x) = \begin{cases} x^{\alpha} e^{-x} e^{Ax} e^{A^* x} \\ x^{\alpha} e^{-x} x^{A} x^{A^*}; \end{cases}$$

- Jacobi weights

$$W(x) = \begin{cases} (1+x)^{\alpha}(1-x)^{\beta}(1+x)^{A}(1+x)^{A^*} \\ (1+x)^{\alpha}(1-x)^{\beta}(1-x)^{A}(1-x)^{A^*}, \end{cases}$$

where $A$ is, in each of these cases, a matrix of a certain specific form which depends on number of parameters (see sections 5, 6 and 7 of [18]).

- *Derivatives.* The following property of the classical orthogonal polynomials can be observed: the derivatives of the classical polynomials are again orthogonal polynomials. It was first proved by W. Hahn in [39] that this property also characterizes the classical polynomials.

  In the matrix case, families of orthogonal polynomials were discovered with the property of satisfying a second order differential equation, but not having derivatives as orthogonal polynomials, see [18].

- *Rodrigues' type formula.* This characterization was suggested by Tricomi in [52] and a complete proof was supplied by Ebert in [24] and Cryer in [8]. It states that Jacobi, Laguerre and Hermite are the only polynomial sequences that have Rodrigues' type formulas

$$P_n(x) = K_n^{-1}\big[W(x)\big]^{-1}\frac{d^n}{dx^n}\big[\rho^n(x)W(x)\big],\ n = 0, 1, 2\cdots,$$

  where $K_n$ depends on $n$, $\rho(x)$ is a polynomial independent of $n$ and of degree 2, 1 or 0, $W(x)$ is positive and integrable over $(a, b)$, where $(a, b)$ is $(-1, 1)$, $(0, \infty)$ or $(-\infty, \infty)$.

  In the matrix case this characterization has not been discovered yet, but some progress in this direction is reported in [18, 19].

- *Differentiation formula.* This characterization was suggested in a conjecture of Karlin and Szegö in [43] and proved by Al-Salam and Chihara in [2]. It states that the only families of orthogonal polynomials for which there exists a differentiation formula

$$\pi(x)P_n'(x) = (\alpha_n x + \beta_n)P_n(x) + \gamma_n P_{n-1}(x),$$

where $\pi(x)$ is a fixed polynomial, are the classical ones. An analog of this characterization has not been yet discovered in the matrix case.

In this work we present some classical properties of scalar valued orthogonal polynomials along with appropriate extensions in the matrix setting:

- *Determinant formulation.* In sections 2.1 and 3.1 the matrix analogs of determinant formulas for the real line and unit circle are presented.
- *Orthogonality.* In sections 2.2 and 3.2 the orthogonality of the matrix valued polynomials is discussed.
- *Recurrence relation.* In sections 2.3 and 3.3 the recurrence relations satisfied by the polynomials are considered.
- *Gramm-Schmidt orthogonalization.* Section 2.4 concerns the Gramm-Schmidt orthogonalization process in the matrix setting.
- *Christoffel-Darboux formulas.* Section 2.5 and 3.4 discuss the kernel polynomials and the Christoffel-Darboux formulas in the matrix setting.
- *$\tau$-function.* In section 2.6 a $\tau$-function for the system of polynomials on the real line is introduced in the matrix setting. Expressions connecting polynomials of the first and second kinds with the $\tau$-function are presented and compared to these in the scalar case.
- *Bochner's problem.* In the last chapter Bochner's problem in the matrix setting is considered and several equivalent formulations are derived.

## 1.2 Future directions

The role of the classical orthogonal polynomials in numerical analysis, mathematical physics, engineering and many other disciplines cannot be overestimated. The subject of matrix valued orthogonal polynomials is relatively young, but has already proved to be rich with applications and new insights. Numerous connections with modern fields such as random matrix theory inspire a lot of interest in the future development of the subject of matrix valued orthogonal polynomials.

One of the most important future directions is the solution of the Bochner's problem in the matrix setting. There exist a lot of very interesting examples demonstrating how much richer than the scalar case the matrix setting is.

Another important aspect is the search for matrix valued analogs of Rodrigues' type formula.

It is very important to find more applications of the "classical" matrix valued orthogonal polynomials in disciplines like physics and engineering, for example, analogs of the applications of scalar valued polynomials to electrostatics and many other fields.

# Chapter 2

# Matrix valued orthogonal polynomials on the real line

In this chapter the classical notion of scalar valued orthogonal polynomials on the real line is extended to matrix valued ones; some new properties are presented and some known ones are investigated in a new context.

## 2.1 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 [7].

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

1. The $n$th moment of the measure $\mu(dx)$ $\mu_n \in \mathbf{R}^{k\times k}$, where
$$\mu_n = \int x^n \mu(dx) = \int x^n W(x)dx, \quad n = 0,\ 1, \cdots .$$
and $x \in \mathbf{R}$. Note that $\mu_n = \mu_n^*$. In this text "$*$" denotes transposition;

2. The matrix $T_n$, where $I$ is $k \times k$ identity matrix
$$T_n = \begin{pmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_n & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \dots & \mu_{2n-2} & \mu_{2n-1} \\ I & xI & \dots & x^{n-1}I & x^nI \end{pmatrix} \quad \in \mathbf{R}^{k(n+1)\times k(n+1)},$$
with $n \geq 1$;

3. A Hankel matrix
$$H_n = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} \\ \mu_1 & \mu_2 & \cdots & \mu_n \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} \end{pmatrix} \in \mathbf{R}^{kn\times kn},$$
$n \geq 1$;

4. The matrix $H$ which is the semi-infinite version of $H_n$ for $n \to \infty$;

5. The vector
$$v_{n,2n-1} = \begin{pmatrix} \mu_n & \mu_{n+1} & \cdots & \mu_{2n-1} \end{pmatrix}^*,$$
for $n \geq 1$;

6. 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 $\mu_{2n}$
$$S_n = \mu_{2n} - v^*_{n,2n-1} H_n^{-1} v_{n,2n-1}, \text{ with } S_0 = \mu_0. \tag{2.1}$$
Also, introduce the diagonal matrix
$$S = \operatorname{diag}\left[\mathrm{S}_0,\ \mathrm{S}_1, \cdots\right].$$

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.*

$$P_n(x) = x^n I - \begin{bmatrix} I & xI & \ldots & x^{n-1}I \end{bmatrix} \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} \\ \mu_1 & \mu_2 & \cdots & \mu_n \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-2} \end{pmatrix}^{-1} \begin{pmatrix} \mu_n \\ \mu_{n+1} \\ \vdots \\ \mu_{2n-1} \end{pmatrix}, \tag{2.2}$$

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

$$P = \left[P_0(x), P_1(x), \cdots\right]. \tag{2.3}$$

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

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

*which is exactly what we obtain using definition (2.2) in the scalar case. This is because for any matrix with partitioning* $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *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 2.2.*

**Note 2** *In the definition 2.2 it is assumed that matrices* $H_n$ *are invertible for all* $n \geq 1$, *which is a restriction on the measure* $\mu(dx)$. *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*

$$Q_n(x) = x\int \mu(dz)\frac{P_n(z)}{x - z}, \tag{2.4}$$

*where the polynomials* $P_n(x)$ *are defined in 2.2. 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 2.4, see [48].*

Note that 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 $i, j \geq 0$.

## 2.2 Orthogonality via the moments of the measure

In this section it will be shown that a family of monic polynomials $\{P_n(x)\}_{n=0}^{\infty}$ defined in 2.2 forms a set of orthogonal polynomials for any symmetric measure $\mu(dx) = W(x)dx$.

**Proposition 1** *Let* $\{P_n(x)\}_{n=0}^{\infty}$ *be a family of polynomials defined in 2.2 and* $S_n$ *be defined in 2.1. Define an inner product on* $\boldsymbol{L}_2(\boldsymbol{R}^k)$ *by means of*

$$\langle P, Q\rangle = \int P^*(x)W(x)Q(x)dx, \tag{2.5}$$

*then*

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

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

***Proof:*** Observe first that for any $0 \leq m \leq n-1$

$$v_{m,m+n-1}H_n^{-1} = \begin{bmatrix}\mu_m & \mu_{m+1} & \mu_{m+2} & \dots & \mu_{m+n-1}\end{bmatrix} H_n^{-1} = \begin{bmatrix}0 & \dots & I & \dots & 0\end{bmatrix},$$

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

$$v_{m,m+n-1}H_n^{-1}v_{n,2n-1} = \mu_{m+n}. \tag{2.6}$$

It is enough to show that $P_n$ is orthogonal to all $x^mI$ for $0 \leq m \leq n-1$, i.e.

$$\begin{aligned}
\int x^m W(x)P_n(x)dx &= \int x^m W(x)\left(x^nI - \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix} H_n^{-1}v_{n,2n-1}\right)dx \\
&= \int x^m W(x)x^nI - x^m W(x) \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix} H_n^{-1}v_{n,2n-1}dx \\
&= \mu_{m+n} - \int \begin{bmatrix} x^m W(x) & x^{m+1}W(x) & \dots & x^{m+n-1}W(x) \end{bmatrix} H_n^{-1}v_{n,2n-1}dx \\
&= \mu_{m+n} - \begin{bmatrix} \mu_m & \mu_{m+1} & \dots & \mu_{m+n-1} \end{bmatrix} H_n^{-1}v_{n,2n-1} \\
&= \mu_{m+n} - \mu_{m+n} = 0.
\end{aligned}$$

This proves that for any $m < n$

$$\int P_m^*(x)W(x)P_n(x)dx = 0.$$

If $m = n$ then

$$\begin{aligned}
\int P_n^*(x)W(x)P_n(x)dx &= \int x^n W(x)\left(x^nI - \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix} H_n^{-1}v_{n,2n-1}\right)dx \\
&= \int x^n W(x)x^n - x^n W(x) \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix} H_n^{-1}v_{n,2n-1}dx \\
&= \mu_{2n} - \begin{bmatrix} \mu_n & \mu_{n+1} & \dots & \mu_{2n-1} \end{bmatrix} H_n^{-1}v_{n,2n-1} \\
&= \mu_{2n} - v_{n,2n-1}^* H_n^{-1}v_{n,2n-1} = S_n,
\end{aligned}$$

which concludes the proof of the proposition. ■

**Note 4** *The inner product introduced in 2.5 is different from the one used in many papers on this subject, e.g. [38, 31, 18, 14, 50] and many others. The standard inner product used is called "left inner product"*

$$\langle P, Q \rangle_L = \int P(x)W(x)Q^*(x)dx,$$

*which is different from the one defined in 2.5 by* $\langle P, Q \rangle_R = \int P^*(x)W(x)Q(x)dx$, *called "right inner product". The left inner product has the following properties (see [50]):*

1. $\langle P, Q \rangle_L = \langle P, Q \rangle_R^T$;
2. *if* $C_1$, $C_2 \in \boldsymbol{R}^{k \times k}$, *then* $\langle C_1P_1 + C_2P_2, Q \rangle_L = C_1 \langle P_1, Q \rangle_L + C_2 \langle P_2, Q \rangle_L$;
3. $\langle xP, Q \rangle_L = \langle P, xQ \rangle_L$;
4. $\langle P, P \rangle_L = 0$ *if and only if* $P = 0$.

*Similar properties are valid for the right inner product, but property* (2) *becomes*

2′ *if* $C_1$, $C_2 \in \boldsymbol{R}^{k \times k}$, *then* $\langle P, Q_1C_1 + Q_2C_2 \rangle_R = \langle P, Q_1 \rangle_R C_1 + \langle P, Q_2 \rangle_R C_2$.

## 2.3 Orthogonality via the recurrence relation

In this section the recurrence relation for the case of matrix valued orthogonal polynomials is investigated.

In lemma 1 it is proven that matrix valued orthogonal polynomials obey an appropriate three term recurrence relation. The proof can also be found in [50].

**Lemma 1** *Let* $\{P_n(x)\}_{n=0}^{\infty}$ *be a family of matrix valued orthogonal polynomials, where* $n$ *corresponds to the degree of the polynomial. Then, for some matrices* $c_n^*$, $b_n^*$ *and* $a_n^*$

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

***Proof:*** Since the polynomials $\{P_n(x)\}_{n=0}^{\infty}$ form a basis for the space of matrix valued polynomials, one can write

$$xP_n(x) = \sum_{i=0}^{n+1} P_i(x)\alpha_i,$$

where $\alpha_i$ are some constant matrices. Multiplying the expression above by $P_m^*(x)W(x)$ from the left and integrating leads to the following:

$$\int P_m^*(x)W(x)xP_n(x)dx = \alpha_m.$$

From the orthogonality property it follows that $\alpha_m = 0$ for $m \leq n-2$, hence

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

for some matrices $c_n^*$, $b_n^*$ and $a_n^*$. ■

In lemma 2 we present expressions for $c_n^*$, $b_n^*$ and $a_n^*$ in terms of the moments of the measure.

**Lemma 2** *The monic orthogonal polynomials defined as in 2.2 obey the following recurrence relation*

$$xP_n(x) = P_{n+1}(x) + P_n(x)b_n^* + P_{n-1}(x)a_n^*, \tag{2.7}$$

*with*

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

*where*

$$u^{n-1} = \begin{pmatrix} u_0^{n-1} \\ u_1^{n-1} \\ \vdots \\ u_{n-1}^{n-1} \end{pmatrix} = H_n^{-1}v_{n,2n-1}; \quad u_n^n = S_n^{-1}\left(\mu_{2n+1} - v_{n,2n-1}^*H_n^{-1}v_{n+1,2n}\right). \tag{2.8}$$

***Proof:*** In the previous lemma we obtained the following recurrence relation

$$xP_n(x) = P_{n+1}c_n^* + P_n(x)b_n^* + P_{n-1}(x)a_n^* \tag{2.9}$$

After multiplying 2.9 by $P^*_{n+1}(x)W(x)$ from the left and integrating we obtain

$$\int P^*_{n+1}(x)W(x)xP_n(x)dx = S_{n+1}c^*_n.$$

Note that from the orthogonality property it follows that

$$\int P^*_{n+1}(x)W(x)xP_n(x)dx = \int P^*_{n+1}(x)W(x)x^{n+1}dx = S_{n+1},$$

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

$$\begin{aligned} S_{n+1} &= \int xP^*_{n+1}(x)W(x)P_n(x)dx \\ &= \int \Big(P^*_{n+2}(x) + b_{n+1}P^*_{n+1}(x) + a_{n+1}P^*_n(x)\Big)W(x)P_n(x)dx \\ &= 0 + 0 + \int a_{n+1}P^*_n(x)W(x)P_n(x)dx = a_{n+1}S_n, \end{aligned}$$

implying

$$a_n = S_nS^{-1}_{n-1}, \text{ or } a^*_n = S^{-1}_{n-1}S_n. \tag{2.10}$$

After multiplying 2.9 by $P^*_n(x)W(x)$ from the left and integrating we obtain

$$\int xP^*_n(x)W(x)P_n(x)dx = \left(\int P^*_n(x)W(x)P_n(x)dx\right)b^*_n = S_nb^*_n,$$

which implies that $S_nb^*_n = b_nS_n$.

In order to compute $b_n$ in terms of the moments let us compare powers of $x$ in the recursion relation 2.7, which can be written out as

$$\begin{aligned} x\left(x^nI - \begin{bmatrix} I & xI & \dots & x^{n-1}I\end{bmatrix}u^{n-1}\right) &= \left(x^{n+1}I - \begin{bmatrix} I & xI & \dots & x^nI\end{bmatrix}u^n\right) \\ &+ \left(x^nI - \begin{bmatrix} I & xI & \dots & x^{n-1}I\end{bmatrix}u^{n-1}\right)b^*_n \\ &+ \left(x^{n-1}I - \begin{bmatrix} I & xI & \dots & x^{n-2}I\end{bmatrix}u^{n-2}\right)a^*_n \end{aligned}$$

Equating the coefficients of front of $x^n$ we get $-u^{n-1}_{n-1} = -u^n_n + b^*_n$, hence

$$b^*_n = u^n_n - u^{n-1}_{n-1}.$$

Let us compute $u^n_n$ in terms of the moments of the orthogonality measure. In order to compute the inverse of $H^{-1}_{n+1}$ we partition $H_{n+1}$ and $H^{-1}_{n+1}$ in the following way:

$$H_{n+1} = \begin{pmatrix} H_n & v_{n,2n-1} \\ v^*_{n,2n-1} & \mu_{2n}\end{pmatrix} \quad \text{and} \quad H^{-1}_{n+1} = \begin{pmatrix} A & \gamma \\ \gamma^* & \alpha\end{pmatrix}, \tag{2.11}$$

and after some simple calculations the following expressions are derived:

$$\alpha = S^{-1}_n, \quad \gamma = -H^{-1}_nv_{n,2n-1}S^{-1}_n, \tag{2.12}$$

and

$$A = H_n^{-1} + H_n^{-1} v_{n,2n-1} S_n^{-1} v_{n,2n-1}^* H_n^{-1} = (H_n - v_{n,2n-1} \mu_{2n}^{-1} v_{n,2n-1}^*)^{-1}. \tag{2.13}$$

From the definition $u^n = H_{n+1}^{-1} v_{n+1,2n+1}$ and the formula for $H_{n+1}^{-1}$ above it follows that

$$\begin{aligned} u_n^n &= \gamma^* v_{n+1,2n} + \alpha\ \mu_{2n+1} = -S_n^{-1} v_{n,2n-1}^* H_n^{-1} v_{n+1,2n} + S_n^{-1} \mu_{2n+1} \\ &= S_n^{-1} \left( \mu_{2n+1} - v_{n,2n-1}^* H_n^{-1} v_{n+1,2n} \right), \end{aligned}$$

which concludes the proof of the lemma. ■

**Note 5** *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 2.10, since in the scalar case* $S_n = \dfrac{\det(H_{n+1})}{\det(H_n)}$.

**Note 6** *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 7** *Matrix polynomials of the second kind defined in 2.4 satisfy the same recursion relation as these of the first kind defined in 2.2, since*

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

**Note 8 (Matrix form of the recurrence relation)** *In the matrix form relations 2.7 for all* $n \geq 0$ *can be written as*

$$L P^*(x) = x P^*(x), \tag{2.14}$$

*where* $L$ *is given by the block tridiagonal matrix*

$$L = \begin{pmatrix} b_0 & I & 0 & 0 & \dots \\ a_1 & b_1 & I & 0 & \dots \\ 0 & a_2 & b_2 & I & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{2.15}$$

**Note 9 (Favard's theorem)** *In the classical theory of scalar valued orthogonal polynomials the following result is known as Favard's theorem, see [7], page 21.*

- *Let* $\{a_n,\ b_n\}_{n=0}^{\infty}$ *be arbitrary sequences of complex numbers and let a family of polynomials* $\{P_n(x)\}_{n=0}^{\infty}$ *be defined by the recurrence formula*

  $$P_n(x) = (x - b_{n-1})P_{n-1}(x) - a_{n-1}P_{n-2},\ P_{-1}(x) = 0,\ P_0(x) = 1.$$

  *Then there is a unique moment functional* $\mathcal{L}$ *such that*

  $$\mathcal{L}[1] = a_0,\ \mathcal{L}[P_n(x), P_m(x)] = 0,\ \textit{for}\ \ m \neq n,\ m, n = 0, 1, \cdots .$$

  *Define* $\mathcal{L}$ *to be quasi-definite (positive definite) if and only if* $\det(H_n) \geq 0$ $\big(\det(H_n) > 0\big)$ *for all* $n \geq 0$. $\mathcal{L}$ *is quasi-definite and* $\{P_n(x)\}_{n=0}^{\infty}$ *are monic orthogonal polynomials if and only if* $a_n \neq 0$ *while* $\mathcal{L}$ *is positive definite if and only if* $b_n$ *are real and* $a_n > 0$.

- *The matrix valued analog of Favard's theorem was presented and proved in papers [11, 16, 21, 22]. In [16] it was shown that every system of polynomials satisfying some* $(2N + 1)$*-term recurrence relation can be expressed in terms of orthonormal matrix polynomials for which coefficients are* $N \times N$ *matrices. In [11] it was proved that any sequence of polynomials satisfying* $(2N + 1)$*-term recurrence relation is orthogonal with respect to a positive definite* $N \times N$ *matrix of measures. Results of these two papers combined allow us to conclude that for any sequence of orthonormal matrix polynomials there exists a positive definite weight matrix with respect to which the polynomials are orthonormal.*

Below the matrix valued orthonormal polynomials are presented.

**Note 10 (Orthonormal matrix valued polynomials)** *Along with monic orthogonal polynomials, one can try to introduce orthonormal matrix valued polynomials. Define a family* $\big\{\bar{P}_n(x)\big\}_{n=0}^{\infty}$ *by means of*

$$\bar{P}_n(x) = P_n(x)S_n^{-1/2}, \quad n \geq 0, \tag{2.16}$$

*then*

$$\begin{aligned}\int \bar{P}_n^*(x)W(x)\bar{P}_n(x)dx &= S_n^{-1/2}\Big(\int P_n^*(x)W(x)P_n(x)dx\Big)S_n^{-1/2} \\ &= S_n^{-1/2}S_nS_n^{-1/2} = I.\end{aligned}$$

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

$$x\bar{P}_n(x) = \bar{P}_{n+1}(x)\bar{a}_{n+1} + \bar{P}_n(x)\bar{b}_n + \bar{P}_{n-1}(x)\bar{a}_n^*, \quad \textit{where} \tag{2.17}$$

$$\bar{b}_n = \int x\bar{P}_n(x)W(x)\bar{P}_n(x)dx = \bar{b}_n^* = S_n^{-1/2}b_nS_n^{1/2};$$

$$\bar{a}_n = S_n^{1/2}S_{n-1}^{-1/2},$$

*or, in the matrix form,* $\bar{L}\bar{P}^*(x) = x\bar{P}^*(x)$, *where* $\bar{L} = S^{1/2}LS^{-1/2}$.

*Observe, that*

$$\bar{L} = \bar{L}^*,$$

*since*

$$\begin{aligned}\int x\bar{P}^*(x)W(x)\bar{P}(x)dx &= \int \bar{L}\bar{P}^*(x)W(x)\bar{P}(x)dx = \bar{L}\int \bar{P}^*(x)W(x)\bar{P}(x)dx \\ &= \bar{L} = \int \bar{P}^*(x)W(x)\bar{P}(x)\bar{L}^*dx = \left(\int \bar{P}^*(x)W(x)\bar{P}(x)dx\right)\bar{L}^* \\ &= \bar{L}^*.\end{aligned}$$

*Also,*

$$LS = SL^*,$$

*since*

$$\bar{L} = S^{1/2}LS^{-1/2} = \bar{L}^* = S^{-1/2}LS^{1/2}.$$

*Orthonormal polynomials will be used in section 2.5 to derive the so called kernel polynomial and Christoffel-Darboux formula.*

**Note 11** *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 $x \in \boldsymbol{R} \Leftrightarrow$*

  *for any vector $v \in \boldsymbol{R}^k$ and $n \geq 0$ $[v^*P_n^*(x)]\,W(x)\,[P_n(x)v] > 0 \Leftrightarrow$*

  *for any vector $v \in \boldsymbol{R}^k$ and $n \geq 0$ $[v^*S_nv] > 0$, which implies that $S_n$ are positive definite for all $n \geq 0$.*

- *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 $Q(x) = \sum_i P_i(x)\alpha_i$. It is easy to see that*

  $$\int Q^*(x)W(x)Q(x)dx = \sum_i \alpha_i^* S_i \alpha_i.$$

  *The above expression implies that $S_n$ being positive definite for all $n \geq 0$ is equivalent to $\int Q^*(x)W(x)Q(x)dx$ being positive definite for all polynomials $Q(x)$, which, in turn, is equivalent to the weight matrix $W(x)$ being positive definite.*

## 2.4 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 2.1, the Gramm-Schmidt orthogonalization procedure and the familiar recursion relation. In this section we will try to establish a connection between the last two.

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

$$\Omega = \begin{bmatrix} I & xI & x^2 I & \dots \end{bmatrix} \tag{2.18}$$

and a unit semi-infinite block upper triangular matrix

$$R = \begin{pmatrix} I & r_{01} & r_{02} & \dots \\ 0 & I & r_{12} & \dots \\ \vdots & \vdots & \vdots & \dots \end{pmatrix}, \quad \text{where} \quad r_{ij} \in \mathbf{R}^{k,k}. \tag{2.19}$$

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

$$\Omega = PR, \tag{2.20}$$

where the matrix $R$ depends on the moments of the measure, and the vector $P$ is defined in 2.3.

**Note 12** *Observe that in order to be able to perform the Gramm-Schmidt procedure, the norm* $\langle P, Q \rangle$ *needs to have the following property: for any nonzero polynomial* $P \neq 0$, $\langle P, P \rangle$ *is non-singular.*

**Proposition 2** *Let matrices* $\Omega$, $P$ *and* $R$ *be defined as in 2.18, 2.3 and 2.19. Assume that equation 2.20 holds and the polynomials in the vector* $P$ *are defined in 2.2, then*

$$S_n r_{n,m} = \mu_{n+m} - v^*_{n,2n-1} H_n^{-1} v_{m,m+n-1}, \tag{2.21}$$

*for all* $n + 1 < m$. *In particular, for* $n = 0$

$$r_{0,m} = \mu_0^{-1} \mu_m. \tag{2.22}$$

*This can be put in matrix form,*

$$H = R^* S R, \tag{2.23}$$

*where* $S = \operatorname{diag}\left[\mathrm{S}_0, \mathrm{S}_1, \mathrm{S}_2, \cdots\right]$.

***Proof:*** From 2.20 it follows that

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

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

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

and 2.21 easily follows from writing out $P_n^*(x)$ as defined in 2.2 and integrating.

In order to show 2.23 consider the matrix $\int \Omega^* W(x)\Omega dx$. Observe that

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

Hence

$$\begin{aligned} H &= \int \Omega^* W(x)\Omega dx = \int R^* P^* W(x) P R dx = R^* \left(\int P^* W(x) P dx\right) R \\ &= R^* S R. \end{aligned}$$

In the last equation we used the fact that

$$S = \int P^* W(x) P dx, \tag{2.24}$$

which is true since

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

which concludes the proof of the proposition. ■

The following proposition establishes the connection between the matrix $L$ defined in 2.15 and the matrix $R$ defined in 2.19.

**Proposition 3** *Denote $R_k$ to be kth column of the matrix $R$, where $k \geq 0$. Then*

$$L^* R_k = R_{k+1} \quad \textit{or} \quad R_k = \left(L^*\right)^k R_0. \tag{2.25}$$

*In particular,*

$$\left(L^*\right)^k_{0,0} = \mu_0^{-1} \mu_k. \tag{2.26}$$

***Proof:*** After multiplying $\Omega = PR$ by $P^* W(x)$ from the left and integrating, we obtain

$$SR = \int P^* W(x) \Omega dx, \quad \text{or} \quad \left(SR\right)_{i,j} = \int P_i^*(x) W(x) x^j dx. \tag{2.27}$$

Also,

$$LSR = \int x P^* W(x) \Omega dx, \quad \text{or} \quad \left(LSR\right)_{i,j} = \int P_i^*(x) W(x) x^{j+1} dx. \tag{2.28}$$

From 2.27 and 2.28 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 2.25.

Observe that since $R$ is a unit upper triangular, $R_0$ is a column of all zeros except the identity at the block $\{0, 0\}$, so $R_k$ is equal to the first column of $(L^*)^k$. From 2.22 we know that $(R_k)_0 = R_{0,k} = \mu_0^{-1} \mu_k$, hence

$$(L^*)^k_{0,0} = (R_k)_0 = \mu_0^{-1} \mu_k,$$

which concludes the proof of the proposition. ■

**Note 13** *In the classical theory of scalar valued orthogonal polynomials expression 2.26 can be found in Akhiezer [1].*

## 2.5 The Christoffel-Darboux formula: the case of the real line

In this section a matrix valued form of 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 2.16, denote the kernel polynomial of degree $n$ to be*

$$K_n(x,y) = \sum_{i=0}^{n} \bar{P}_i(y)\bar{P}_i^*(x). \tag{2.29}$$

*Then*

$$K_n(x,y) = \begin{bmatrix} I & yI & \dots & y^nI \end{bmatrix} \begin{pmatrix} \mu_0 & \mu_1 & \dots & \mu_n \\ \mu_1 & \mu_2 & \dots & \mu_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \dots & \mu_{2n} \end{pmatrix}^{-1} \begin{bmatrix} I \\ xI \\ \vdots \\ x^nI \end{bmatrix} \tag{2.30}$$

***Proof by induction:*** For $n = 0$ we have $K_0(x,y) = \bar{P}_0(x)\bar{P}_0^*(y) = \mu_0^{-1}$ which agrees with formula 2.30. To simplify the notation denote the right hand side of expression 2.30 as $RHS(n)$. For the inductive step $(n-1) \to n$ we use the notation and partitioning in 2.11 as well as formulas 2.12 and 2.13 to rewrite $RHS(n)$ as

$$\begin{aligned} RHS(n) &= \begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} A \begin{bmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{bmatrix} + y^n\gamma^* \begin{bmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{bmatrix} \\ &+ \begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} \gamma_n x^n + y^n x^n \alpha \\ &= \begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} \left( H_n^{-1} + H_n^{-1} v_{n,2n-1} S_n^{-1} v_{n,2n-1}^* H_n^{-1} \right) \begin{bmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{bmatrix} \\ &+ y^n S_n^{-1} v_{n,2n-1}^* H_n^{-1} \begin{bmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{bmatrix} + \begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} H_n^{-1} v_{n,2n-1} S_n^{-1} x^n + y^n x^n \alpha. \end{aligned}$$

From 2.2 and 2.16 we have that for both $x$ and $y$ the following holds

$$\begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} H_n^{-1} v_{n,2n-1} S_n^{-1/2} = y^n S_n^{-1/2} - \bar{P}_n(y)$$

and substituting this into the above expression we obtain

$$\begin{aligned}
RHS(n) &= \begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} H_n^{-1} \begin{bmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{bmatrix} + \left( y^n S_n^{-1/2} - \bar{P}_n(y) \right) \left( x^n S_n^{-1/2} - \bar{P}_n^*(x) \right) \\
&- y^n S_n^{-1/2} \left( x^n S_n^{-1/2} - \bar{P}_n^*(x) \right) - x^n \left( y^n S_n^{-1/2} - \bar{P}_n(y) \right) S_n^{-1/2} + y^n x^n S_n^{-1} \\
&= \begin{bmatrix} I & yI & \dots & y^{n-1}I \end{bmatrix} H_n^{-1} \begin{bmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{bmatrix} + \bar{P}_n(y)\bar{P}_n^*(x) \\
&= RHS(n-1) + \bar{P}_n(y)\bar{P}_n^*(x),
\end{aligned}$$

which completes the proof by induction. ■

**Note 14** *In the classical theory of scalar valued orthogonal polynomials (see [7]), 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 the formula 2.30 derived above.*

In the next lemma the derivation of the Christoffel-Darboux formula is presented.

**Lemma 4** *Let a family of orthonormal polynomials* $\left\{\bar{P}_n(x)\right\}_{n=0}^{\infty}$ *be defined by (2.2) and (2.16). Then*

$$\sum_{m=0}^{n} \bar{P}_m(y)\bar{P}_m^*(x) = \frac{\bar{P}_n(y)\bar{a}_{n+1}^*\bar{P}_{n+1}^*(x) - \bar{P}_{n+1}(y)\bar{a}_{n+1}\bar{P}_n^*(x)}{x-y}. \tag{2.31}$$

***Proof:*** From the recursion relation 2.17, we have

$$x\bar{P}_n(y)\bar{P}_n^*(x) = \bar{P}_n(y)\bar{a}_{n+1}^*\bar{P}_{n+1}^*(x) + \bar{P}_n(y)\bar{b}_n\bar{P}_n^*(x) + \bar{P}_n(y)\bar{a}_n\bar{P}_{n-1}^*(x),$$

$$y\bar{P}_n(y)\bar{P}_n^*(x) = \bar{P}_{n+1}(y)\bar{a}_{n+1}\bar{P}_n^*(x) + \bar{P}_n(y)\bar{b}_n\bar{P}_n^*(x) + \bar{P}_{n-1}(y)\bar{a}_n^*\bar{P}_n^*(x).$$

Subtracting the second equation from the first yields

$$\begin{aligned}
(x-y)\bar{P}_n(y)\bar{P}_n^*(x) &= \left( \bar{P}_n(y)\bar{a}_{n+1}^*\bar{P}_{n+1}^*(x) - \bar{P}_{n+1}(y)\bar{a}_{n+1}\bar{P}_n^*(x) \right) \\
&- \left( \bar{P}_{n-1}(y)\bar{a}_n^*\bar{P}_n^*(x) - \bar{P}_n(y)\bar{a}_n\bar{P}_{n-1}^*(x) \right).
\end{aligned}$$

If we denote

$$F_n(x,y) = \frac{\bar{P}_n(y)\bar{a}^*_{n+1}\bar{P}^*_{n+1}(x) - \bar{P}_{n+1}(y)\bar{a}_{n+1}\bar{P}^*_n(x)}{x-y},$$

then the last equation can be rewritten

$$\bar{P}_m(y)\bar{P}^*_m(x) = F_m(x,y) - F_{m-1}(x,y).$$

Summing the latter from 0 to $n$ and noticing that $F_{-1}(x,y) = 0$, we obtain (2.31), which concludes the proof of the lemma. ■

**Note 15** *In the classical theory of scalar valued monic orthogonal polynomials (see [7]) 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 the formula 2.31 derived above.*

## 2.6 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 its properties.

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

$$\mu_t(dx) = e^{\sum_{i=1}^{\infty} t_i x^i I}\mu(dx), \tag{2.32}$$

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

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

Observe that

$$\frac{d\mu_n(t)}{dt_m} = \int x^{n+m} e^{\sum_{i=1}^{\infty} t_i x^i I}\mu(dx) = \mu_{n+m}(t). \tag{2.33}$$

In what follows prime " $'$ " denotes differentiation with respect to $t_1$, in particular $\mu'_n(t) = \mu_{n+1}(t)$ and $v'_{n,2n-1} = v_{n+1,2n}$.

**Definition 3** *Define*

$$\tau_n(t) \equiv S_n(t) = \mu_{2n}(t) - v^*_{n,2n-1}(t)H_n^{-1}(t)v_{n,2n-1}(t), \tag{2.34}$$

*where $H_n(t)$ and $v_{n,2n-1}(t)$ are defined in section 2.1, but with "time" dependence.*

**Note 16** *In the classical theory of scalar valued orthogonal polynomials, the $\tau(t)$-function is defined in the following fashion*

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

*whereas our new definition in the scalar case is*

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

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

**Lemma 5**

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

*This could also be written as*

$$\Big(H_n\Big)' H_n^{-1} v_{n,2n-1} = v'_{n,2n-1} - \begin{pmatrix} 0 \\ 0 \\ \vdots \\ S_n \end{pmatrix}.$$

***Proof:***

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

where identity 2.6 was used. ■

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.

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

**Theorem 1** *Given a family of monic orthogonal polynomials as defined in 2.2 satisfying the recursion relation*

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

*Then*

$$b_n^* = \tau_n(t)^{-1}\tau_n'(t)|_{t_1=0} = \ln(\tau_n(t))'_{t_1=0}$$

*and*

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

*for the $\tau$-function defined in 2.34.*

***Proof:*** By definition, $\tau_n(t) = \mu_{2n}(t) - v^*(t)H_n^{-1}(t)v(t)$. According to (2.33), after we differentiate with respect to $t_1$ we obtain

$$\begin{aligned}\tau_n(t)' &= \mu_{2n+1} - \Big(v^* H_n^{-1} v\Big)' \\ &= \mu_{2n+1} - \Big(v^*\Big)' H_n^{-1} v - v^* \Big(H_n^{-1}\Big)' v - v^* H_n^{-1} \Big(v\Big)'\end{aligned}$$

Observe that

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

hence

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

Recall from 2.8 that $\mu_{2n+1} - v^* H_n^{-1} v' = S_n u_n^n$, implying that

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

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

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

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

$$\tau(t)' = \tau(t)\Big(u_n^n(t) - u_{n-1}^{n-1}(t)\Big) = \Big(u_n^n(t) - u_{n-1}^{n-1}(t)\Big)^* \tau(t),$$

which implies that

$$\ln(\tau_n(t))'_{t_1=0} := \tau(t)^{-1}\tau(t)'|_{t_1=0} = \Big(u_n^n(0) - u_{n-1}^{n-1}(0)\Big) = b_n^*.$$

From 2.10 we can see that

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

which concludes the proof of the theorem. ■

**Note 17** *In the classical theory of scalar valued orthogonal polynomials, 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 $\tau_n(t) = \dfrac{\det(H_{n+1}(t))}{\det(H_n(t))}$, the expression above is equivalent to the one obtained in the previous theorem, $b_n = \frac{\partial}{\partial t_1}\ln(\tau_n(t))$.*

The next theorem expresses the connection between the monic matrix valued orthogonal polynomials as defined in 2.2 and the $\tau$-function as defined in 2.34. The following two lemmas are used in the proof of the first part of the theorem. To ease the notation denote $X = \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix}^*$.

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

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

*where*

$$w = Xw_0, \text{ and } P_n^*(x,t)w_0 + S_n(t) = 0.$$

***Proof:*** Suppose

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

for some $w$, then using lemma 5 one obtains

$$\begin{aligned} v - \frac{v'}{x} &= \left(I - \frac{H_n' H_n^{-1}}{x}\right)(v+w) = v + w - \frac{H_n' H_n^{-1} v}{x} - \frac{H_n' H_n^{-1} w}{x} \\ &= v + w - \frac{v'}{x} + \frac{1}{x}\begin{pmatrix} 0 & 0 & \cdots & S_n \end{pmatrix}^* - \frac{H_n' H_n^{-1} w}{x}, \end{aligned}$$

which implies

$$H_n' H_n^{-1} w = xw + \begin{pmatrix} 0 & 0 & \cdots & S_n \end{pmatrix}^*. \tag{2.35}$$

Note that

$$H_n' H_n^{-1} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ v^* H_n^{-1} \end{pmatrix}, \tag{2.36}$$

where $e_i = [0 \cdots I \cdots 0]$ and $I$ is at the $i$th location. After applying 2.36 to 2.35 we obtain the following equation

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ v^* H_n^{-1} w \end{pmatrix} = \begin{pmatrix} w_0 x \\ w_1 x \\ \vdots \\ w_{n-1} x + S_n \end{pmatrix}, \tag{2.37}$$

where $w_i$ denotes $i$th element of $w$. Expression 2.37 implies that $w_i = x^i w_0$ for $i = 0, \cdots, n-1$, and

$$v^* H_n^{-1} X w_0 = x^n w_0 + S_n,$$

leading to

$$0 = \left( x^n - v^* H_n^{-1} X \right) w_0 + S_n = P_n^*(x,t) w_0 + S_n,$$

which concludes the proof of the lemma. ■

**Lemma 7**

$$P_{n+1}^* = x^{n+1} I - u_n^{n*} P_n^* - v'^* H_n^{-1} X.$$

***Proof:*** Writing

$$P_{n+1}^*(x) = x^{n+1} - v_{n+1,2n+1}^* H_{n+1}^{-1} \begin{pmatrix} X \\ x^n \end{pmatrix} = x^{n+1} - [v'^* \ \mu_{2n+1}] H_{n+1}^{-1} \begin{pmatrix} X \\ x^n \end{pmatrix},$$

and using the partition of $H_{n+1}$ as discussed in 2.11, 2.12, 2.13 one obtains the statement of the lemma. ■

The following several lemmas are used in the proof of the second part of the theorem that will follow. To ease the notation, $v^{(i)}$ and $H_n^{(i)}$ denote the $i$th derivatives with respect to $t_1$.

**Lemma 8** *Let $Q_n(x,t)$ be defined as in 2.4. Then,*

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

*where $r_{n,m}^*$ is defined in 2.21. Denote*

$$R(n,x) = \sum_{i=0}^{\infty} \frac{r_{n,n+i}^*}{x^i}, \tag{2.38}$$

*hence*

$$Q_n(x,t) = \frac{1}{x^n} R(n,x) S_n(t), \tag{2.39}$$

*which concludes the proof of the lemma.* ■

**Lemma 9** *For any $x \in \boldsymbol{R}$ let us consider the equation*

$$\left( \sum_{i=0}^{\infty} \frac{H_n^{(i)}}{x^i} \right)^{-1} \left( \sum_{i=0}^{\infty} \frac{v^{(i)}}{x^i} \right) = H_n^{-1}(v+w). \tag{2.40}$$

*Then*

$$w = [0\ 0\ \cdots\ w_{n-1}^*]^*; \quad w_{n-1} = \frac{R(n,x)S_n - U}{x}, \tag{2.41}$$

*where*

$$U = \sum_{i=0}^{\infty} \frac{v^{*(i)} H_n^{-1} w}{x^i}. \tag{2.42}$$

***Proof:*** Rewriting expression 2.21 and keeping in mind that for any vector $\xi$

$$v_{i,n-1+i}^* H_n^{-1} \xi = \xi_i,$$

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

$$v^{*(i)} H_n^{-1} v = v_{n+i,2n+i-1}^* H_n^{-1} v = \mu_{n+i} - r_{n,n+i}^* S_n. \tag{2.43}$$

Rearranging 2.40 one arrives at

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

or, element-wise

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

The last equation implies $w_{n-1} = \dfrac{R(n,x)S_n - U}{x}$. 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 $0 \leq i \leq n-2$, which concludes the proof of the lemma. ■

**Lemma 10** *Suppose $U$ and $R(n,x)$ are as defined in 2.42 and 2.38, then*

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

***Proof:*** Using formulas 2.11 and 2.12 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

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

which concludes the proof of the lemma. ■

**Theorem 2** *Let $\tau(t)$ be as defined in 2.34.*

- *Let $\{P_n(x,t)\}_{n=0}^{\infty}$ be a family monic matrix valued orthogonal polynomials as defined in 2.2 with space $x$ and "time" dependence $t$. Then*

$$P_{n+1}(x,t) = xP_n(x,t)\tau_n^{-1}(t)\tau_n\big(t - [x^{-1}]\big), \tag{2.44}$$

*where*

$$\begin{aligned} \mu_n\big(t - [x^{-1}]\big) &= \int z^n e^{\sum_{i=1}^{\infty}\left(t_i - \frac{x^{-i}}{i}\right)z^i} W(z)dz \\ &= \mu_n(t) - \frac{\mu_{n+1}(t)}{x}. \end{aligned}$$

- *Let $\{Q_n(x,t)\}_{n=0}^{\infty}$ be a family matrix valued orthogonal polynomials as defined in 2.4 with space $x$ and "time" dependence $t$. Then*

$$xQ_{n+1}(x,t) = Q_n(x,t)\tau_n^{-1}(t)\tau_{n+1}\big(t + [x^{-1}]\big), \tag{2.45}$$

*where*

$$\begin{aligned} \mu_n\big(t + [x^{-1}]\big) &= \int z^n e^{\sum_{i=1}^{\infty}\left(t_i + \frac{x^{-i}}{x}\right)z^i} W(z)dz \\ &= \sum_{i=0}^{\infty} \frac{\mu_{n+i}(t)}{x^i}. \end{aligned}$$

***Proof:***

- Observe that

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

Then

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

where lemma 6 was used in the fourth equation and lemma 7 was used in the ninth equation.

- Observe that

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

Then

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

where lemma 9 was used in the second equation. From lemma 9 we know that $R(n,x)S_n(t) - U = xw_{n-1}$, hence

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

Using the result of lemma 10 we can 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 8 the expression above becomes

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

implying

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

which concludes the proof of the theorem. ■

**Note 18** *In the classical theory of scalar valued orthogonal polynomials, the expression 2.44 has the following form*

$$p_n(x,t) = x^n\frac{\tilde{\tau}_n\left(t-[x^{-1}]\right)}{\tilde{\tau}_n(t)}, \tag{2.46}$$

*where* $\tilde{\tau}_n(t) = \det\left(H_n(t)\right)$, *see [40] and [48]. Observe that in the scalar case 2.44 is equivalent to 2.46, since*

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

*Using the facts that*

$$\tau_n(t) = \frac{\det\big(H_{n+1}(t)\big)}{\det\big(H_n(t)\big)} \quad \textit{and} \quad \tau_0(t) = S_0(t) = \det\left(H_1(t)\right) = \mu_0(t)$$

*the expression above becomes*

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

*The scalar valued analog of expression 2.45 is*

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

*and its equivalence to 2.45 is proved similarly.*

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

**Proposition 4** *Let* $\{P_n(x,t)\}_{n=0}^{\infty}$ *be a family of monic orthogonal matrix valued polynomials as defined in 2.2 with "time" dependent moments. Let* $a_n$ *and* $b_n$ *be the coefficients of the recursion relation 2.7 with "time" dependence, and let* $\dfrac{\partial}{\partial t_1}$ *be denoted by "′". Then*

1. $P'_{n+1}(x,t) = -P_n(x,t)a^*_{n+1}$;
2. $(u^n_n)' = \tau_n^{-1}(t)\tau_{n+1}(t) = S_n^{-1}(t)S_{n+1}(t) = a^*_{n+1}$;
3. $(b^*_n)' = a^*_{n+1} - a^*_n$;
4. $(a^*_n)' = a^*_n b^*_n - b^*_{n-1}a^*_n$.

***Proof:***

1. Denote $X = \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix}^*$, then

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

Using partition and notation from 2.11, 2.12 and 2.13 we conclude that

$$\begin{aligned}
-\left[X^*\ x^nI\right]H^{-1}_{n+1}\begin{pmatrix}0 & 0 & \cdots & S_n\end{pmatrix}^* &= -\left[X^*\ x^nI\right]\begin{pmatrix}-H_n^{-1}v_{n,2n-1}S_n^{-1}S_{n+1} \\ S_n^{-1}S_{n+1}\end{pmatrix} \\
&= -(x^n - X^*H_n^{-1}v_{n,2n-1})a^*_{n+1} = -P_na^*_{n+1}.
\end{aligned}$$

2. By similar reasoning,

$$(u^n_n)' = e_n\left(H^{-1}_{n+1}v_{n+1,2n+1}\right)' = e_nH^{-1}_{n+1}\begin{pmatrix}0 & 0 & \cdots & S_n\end{pmatrix}^* = S_n^{-1}S_{n+1}.$$

3. Follows from the previous part and the fact that $b^*_n = u^n_n - u^{n-1}_{n-1}$.

4. $a^*_{n+1} = S_n^{-1}S_{n+1}$, hence $S_na^*_{n+1} = S_{n+1}$. After differentiating both sides we obtain $S'_na^*_{n+1} + S_n(a^*_{n+1})' = S'_{n+1}$. Since $S'_n = S_nb^*_n$, we obtain

$$\left(a^*_{n+1}\right)' = a^*_nb^*_n - b^*_{n-1}a^*_n.$$

**Note 19** *It is only natural to call* (3) *and* (4) *in the above proposition the non-Abelian Toda equations, see [26].*

# Chapter 3

# Matrix valued orthogonal polynomials on the unit circle

In this chapter the classical notion of scalar valued orthogonal polynomials on the unit circle is extended to matrix valued one; some new properties are presented and some known ones are investigated in a new context.

## 3.1 Definitions

In this section we introduce the notations and present a definition of the scalar/matrix valued orthogonal polynomials on the unit circle which is a natural extension of the classical determinant definition discussed in numerous books and articles, for example, see [7].

Throughout this section " $*$ " means transposition and conjugation.

Given a measure $\mu(d\theta) = f(\theta)d\theta$ with Hermitian weight function $f(\theta) \in \mathbf{R}^{k\times k}$, $k \geq 1$ supported and integrable on $[-\pi, \pi]$, introduce

1. The $n$th moment of the measure $\mu(d\theta)$ $\mu_n \in \mathbf{C}^{k\times k}$, where
$$\mu_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{in\theta}\mu(d\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{in\theta} f(\theta)d\theta; \quad n = 0, \pm 1, \pm 2...$$
Note that $\mu_{-n} = \mu_n^*$.

2. The matrices $M_n^r$ and $M_n^l$ in $\mathbf{C}^{k(n+1)\times k(n+1)}$, where $I$ is $k \times k$ identity matrix, $x = e^{i\theta} \in \mathbf{C}$,

$\theta \in [-\pi, \pi]$

$$M_n^r = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} & \mu_n \\ \mu_{-1} & \mu_0 & \cdots & \mu_{-n+2} & \mu_{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{-n+1} & \mu_{-n+2} & \cdots & \mu_0 & \mu_1 \\ I & xI & \cdots & x^{n-1}I & x^nI \end{pmatrix},$$

$$M_n^l = \begin{pmatrix} \mu_0 & \mu_{-1} & \cdots & \mu_{-n+1} & I \\ \mu_1 & \mu_0 & \cdots & \mu_{-n+2} & xI \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_{n-2} & \cdots & \mu_0 & x^{n-1}I \\ \mu_n & \mu_{n-1} & \cdots & \mu_1 & x^nI \end{pmatrix},$$

for $n \geq 1$;

3. Toeplitz matrices $T_n^r$ and $T_n^l \in \mathbf{C}^{kn \times kn}$

$$T_n^r = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{n-1} \\ \mu_{-1} & \mu_0 & \cdots & \mu_{n-2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{-n+1} & \mu_{-n+2} & \cdots & \mu_0 \end{pmatrix} \text{ and } T_n^l = \begin{pmatrix} \mu_0 & \mu_{-1} & \cdots & \mu_{-n+1} \\ \mu_1 & \mu_0 & \cdots & \mu_{-n+2} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_{n-2} & \cdots & \mu_0 \end{pmatrix},$$

for $n \geq 1$;

4. The vectors

$$\nu_n = \begin{pmatrix} \mu_n \\ \mu_{n-1} \\ \vdots \\ \mu_1 \end{pmatrix} \quad \text{and} \quad \xi_n = \begin{pmatrix} \mu_{-n} \\ \mu_{-n+1} \\ \vdots \\ \mu_{-1} \end{pmatrix},$$

for $n \geq 1$;

5. In the matrices

$$T_{n+1}^r = \begin{pmatrix} T_n^r & \nu_n \\ \nu_n^* & \mu_0 \end{pmatrix} \text{ and } T_{n+1}^l = \begin{pmatrix} T_n^l & \xi_n \\ \xi_n^* & \mu_0 \end{pmatrix}$$

denote the Schur complements of $\mu_0$

$$S_n^r = \mu_0 - \nu_n^* T_n^{-1} \nu_n; \quad S_n^l = \mu_0 - \xi_n^* T_n^{-1} \xi_n, \text{ with } S_n^l = S_n^r = \mu_0, \tag{3.1}$$

where $T_n^{-l}$ and $T_n^{-r}$ denote $\left(T_n^l\right)^{-1}$ and $(T_n^r)^{-1}$ correspondingly.

Using the notations above we introduce the following definition:

**Definition 4 (Monic matrix valued polynomials on the unit circle)** *Define two families of polynomials* $\left\{P_n^r(x)\right\}_{n=0}^{\infty}$ *and* $\left\{P_n^l(x)\right\}_{n=0}^{\infty}$ *as Schur complements of* $x^nI$ *in the matrices* $M_{n+1}^r$ *and* $M_{n+1}^l$*, i.e.*

$$P_n^r(x) = x^nI - \begin{bmatrix} I & xI & \dots & x^{n-1}I \end{bmatrix} T_n^{-r} \begin{pmatrix} \mu_n \\ \mu_{n-1} \\ \vdots \\ \mu_1 \end{pmatrix} \tag{3.2}$$

*and*

$$P_n^l(x) = x^nI - \begin{bmatrix} \mu_n & \mu_{n-1} & \cdots & \mu_1 \end{bmatrix} T_n^{-l} \begin{pmatrix} I \\ xI \\ \vdots \\ x^{n-1}I \end{pmatrix}, \tag{3.3}$$

*where "r" and "l" stand for the right and the left polynomials.*

**Note 20** *Observe that*

$$T_n^r = PT_n^lP, \quad for \quad P = \begin{pmatrix} 0 & 0 & \cdots & I \\ 0 & 0 & I & 0 \\ \vdots & \vdots & \vdots & \vdots \\ I & 0 & 0 & 0 \end{pmatrix}.$$

**Note 21** *If* $k = 1$*, i.e. in the scalar case* $P_n^l(x) = P_n^r(x)$*. In the classical theory of scalar valued orthogonal polynomials on the unit circle (see [51]), monic polynomials are defined as*

$$p_n^r(x) = p_n^l(x) = \frac{\det(M_n^r)}{\det(T_n^r)},$$

*which is exactly what we obtain using definition 3.2 in the scalar case.*

## 3.2 Orthogonality via the moments of the measure

In this section it will be shown that families of monic polynomials $\left\{P_n^r(x)\right\}_{n=0}^{\infty}$ and $\left\{P_n^l(x)\right\}_{n=0}^{\infty}$ as defined in (3.2) and (3.3) form sets of monic orthogonal polynomials for any symmetric measure $\mu(d\theta) = f(\theta)d\theta$.

**Proposition 5** *Let* $\left\{P_n^r(x)\right\}_{n=0}^{\infty}$ *and* $\left\{P_n^l(x)\right\}_{n=0}^{\infty}$ *be families of monic polynomials as defined in 3.2 and 3.3. Define right and left inner products on the unit circle as*

$$\langle P, Q\rangle_r = \int P^*\left(e^{i\theta}\right) f\left(e^{i\theta}\right) Q\left(e^{i\theta}\right) d\theta, \quad \langle P, Q\rangle_l = \int P\left(e^{i\theta}\right) f\left(e^{i\theta}\right) Q^*\left(e^{i\theta}\right) d\theta.$$

*Let* $S_n^r$ *and* $S_n^l$ *be defined in 3.1. Then*

$$\left\langle P_i^r, P_j^r\right\rangle_r = \int_{-\pi}^{\pi} P_i^{r*}\left(e^{i\theta}\right) f(\theta) P_j^r\left(e^{i\theta}\right) d\theta = \delta_{ij}S_i^r,$$

$$\langle P_i^l, P_j^l \rangle_l = \int_{-\pi}^{\pi} P_i^l\left(e^{i\theta}\right) f(\theta) P_j^{l*}\left(e^{i\theta}\right) d\theta = \delta_{ij} S_i^l,$$

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

***Proof:*** We will prove the proposition for the right norm only, conclusion for the left norm can be proved similarly. Observe first that for any $0 \leq m \leq n-1$

$$\begin{bmatrix} \mu_{-m} & \mu_{-m+1} & \mu_{-m+2} & \dots & \mu_{-m+n-1} \end{bmatrix} T_n^{-r} = \begin{bmatrix} 0 & \dots & I & \dots & 0 \end{bmatrix},$$

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

$$\begin{bmatrix} \mu_{-m} & \mu_{-m+1} & \mu_{-m+2} & \dots & \mu_{-m+n-1} \end{bmatrix} T_n^{-r} \nu_n = \mu_{-m+n}. \tag{3.4}$$

It is enough to show that $P_n^r\left(e^{i\theta}\right)$ is orthogonal to all $e^{-im\theta}$ for $0 \leq m \leq n-1$, i.e.

$$\begin{aligned} \int_{-\pi}^{\pi} e^{-im\theta} f(\theta) P_n^r\left(e^{i\theta}\right) d\theta &= \int_{-\pi}^{\pi} e^{-im\theta} f(\theta) \left( I e^{in\theta} - \begin{bmatrix} I & e^{i\theta} I & \cdots & e^{i(n-1)\theta} I \end{bmatrix} T_n^{-r} \nu_n \right) d\theta \\ &= \mu_{n-m} - \begin{bmatrix} \mu_{-m} & \mu_{-m+1} & \dots & \mu_{-m+n-1} \end{bmatrix} T_n^{-r} \nu_n \\ &= \mu_{n-m} - \mu_{n-m} = 0. \end{aligned}$$

This proves that for any $m < n$

$$\int_{-\pi}^{\pi} P_m^{r*}\left(e^{i\theta}\right) f(\theta) P_n^r\left(e^{i\theta}\right) d\theta = 0.$$

If $m = n$ then

$$\begin{aligned} \int P_n^{r*}\left(e^{i\theta}\right) f(\theta) P_n^r\left(e^{i\theta}\right) d\theta &= \int e^{-in\theta} f(\theta) P_n^r\left(e^{i\theta}\right) d\theta \\ &= \mu_0 - \nu_n^* T_n^{-r} \nu_n = S_n^r. \end{aligned}$$

Statement for the left norm is proved similarly. ■

**Note 22** *In the classical theory of scalar valued orthogonal polynomials on the unit circle (for example, see [51]),*

$$\langle P_n^r, P_n^r \rangle_r = \frac{\det\left(T_{n+1}\right)}{\det\left(T_n\right)},$$

*which is identical to our formula applied for the scalar case.*

## 3.3 The recursion relations

In this section the recurrence relation for the case of matrix valued orthogonal polynomials is investigated.

**Proposition 6** *Let* $\{P_n^r(x)\}_{n=0}^{\infty}$ *and* $\left\{P_n^l(x)\right\}_{n=0}^{\infty}$ *be families of monic matrix valued orthogonal polynomials as defined in 3.2 and 3.3. Then the following recursion relations are equivalent:*

*1.* $P^r_{n+1}(x) = xP^r_n(x) + \hat{P}^l_n(x)P^r_{n+1}(0)$;

*2.* $\hat{P}^r_{n+1}(x) = \hat{P}^r_n(x) + xP^{r*}_{n+1}(0)P^l_n(x)$;

*3.* $P^l_{n+1}(x) = xP^l_n(x) + P^l_{n+1}(0)\hat{P}^r_n(x)$;

*4.* $\hat{P}^l_{n+1}(x) = \hat{P}^l_n(x) + xP^r_n(x)P^{l*}_{n+1}(0)$;

*5.* $P^r_{n+1}(x) = xP^r_n(x)\Big(I - P^{l*}_{n+1}(0)P^r_{n+1}(0)\Big) + \hat{P}^l_{n+1}(x)P^r_{n+1}(0)$;

*6.* $P^l_{n+1}(x) = x\Big(I - P^l_{n+1}(0)P^{r*}_{n+1}(0)\Big)P^l_n(x) + P^l_{n+1}(0)\hat{P}^r_{n+1}(x)$;

*7.* $\Big(I - P^{l*}_{n+1}(0)P^r_{n+1}(0)\Big) = S^{-r}_n S^r_{n+1}$;

*8.* $\Big(I - P^l_{n+1}(0)P^{r*}_{n+1}(0)\Big) = S^l_{n+1}S^{-l}_n$;

*9.* $S^l_n P^r_n(0) = P^l_n(0)S^r_n$.

*where* $\hat{P}^{r,l}_n(x) = x^n\left(P^{r,l}_n(x)\right)^*$.

**Proof:** We will prove the first recursion relation, and the rest of them can be obtained from the first one by trivial manipulations. Let us partition matrices $T^r_{n+1}$, $T^{-r}_{n+1}$ and $\nu_{n+1}$ in the following way:

$$T^r_{n+1} = \begin{pmatrix} \mu_0 & \phi^* \\ \phi & T^r_n \end{pmatrix}; \quad T^{-r}_{n+1} = \begin{pmatrix} \alpha & \gamma^* \\ \gamma & A \end{pmatrix}; \quad \nu_{n+1} = \begin{pmatrix} \mu_{n+1} \\ \nu_n \end{pmatrix}; \quad \phi = \begin{pmatrix} \mu_{-1} \\ \mu_{-2} \\ \vdots \\ \mu_{-n} \end{pmatrix}.$$

After some simple calculations we arrive at

$$\alpha = (\mu_0 - \phi^* T^{-r}_n \phi)^{-1}; \qquad \gamma = -T^{-r}_n \phi\alpha; \quad A = T^{-r}_n - T^{-r}_n \phi\gamma^*;$$

and

$$P^r_{n+1}(0) = -(\alpha\mu_{n+1} + \gamma^*\nu_n).$$

Observe that

$$\begin{aligned} \hat{P}^l_n(x) &= x^n\Big(x^{-n}I - [I\ x^{-1}I\ \cdots x^{-n+1}I]T^{-l}_n \xi_n\Big) \\ &= I - [x^n I\ x^{n-1}I\ \cdots I]T^{-l}_n \xi_n \\ &= I - [I\ xI\ \cdots x^n I]T^{-r}_n \phi, \end{aligned}$$

hence

$$\begin{aligned}
P^r_{n+1}(x) - xP^r_n(x) &= -\Big[I \quad xI \quad \dots \quad x^nI\Big]\, T^{-r}_{n+1}\nu_{n+1} + \Big[xI \quad x^2I \quad \dots \quad x^nI\Big]\, T^{-r}_n \nu_n \\
&= \Big[xI \quad x^2I \quad \dots \quad x^nI\Big]\Big(T^{-r}_n \nu_n - \gamma\mu_{n+1} - A\nu_n\Big) - \alpha\mu_{n+1} - \gamma^*\nu_n \\
&= \Big[xI \quad x^2I \quad \dots \quad x^nI\Big]\Big(T^{-r}_n \phi\gamma^*\nu_n + T^{-r}_n \phi\alpha\mu_{n+1}\Big) + P^r_{n+1}(0) \\
&= -\Big[xI \quad x^2I \quad \dots \quad x^nI\Big]\, T^{-r}_n \phi P^r_{n+1}(0) + P^r_{n+1}(0) \\
&= \hat{P}^l_n(x)P^r_{n+1}(0),
\end{aligned}$$

which proves identity (1) of the proposition.

By applying the "ˆ" operator (introduced at the end of the proposition above) to the identity (1) in the proposition we obtain (2). By partitioning the matrix $T^l_{n+1}$ and applying the same technique as above we obtain (3) and (4). Identity (5) is obtained by expressing $\hat{P}^l_n(x)$ from (4) and substituting into (1). Identity (6) is obtained similarly.

In order to prove (7) let us rewrite identity (5) from the proposition in the following way:

$$\frac{P^r_{n+1}(x)}{x^{n+1}} = \frac{P^r_n(x)}{x^n}\Big(I - P^{l*}_{n+1}(0)P^r_{n+1}(0)\Big) + P^{l*}_{n+1}(x)P^r_{n+1}(0).$$

After multiplying this expression by $P^l_n\left(e^{i\theta}\right) f(\theta)$ from the left, substituting $x = e^{i\theta}$ and integrating we arrive at the following three integrals:

$$\begin{aligned}
\int_{-\pi}^{\pi} P^l_n\left(e^{i\theta}\right) f(\theta) \frac{P^r_{n+1}\left(e^{i\theta}\right)}{e^{i\theta(n+1)}} d\theta &= \int_{-\pi}^{\pi} \frac{P^l_n\left(e^{i\theta}\right)}{e^{i\theta(n+1)}} f(\theta) P^r_{n+1}\left(e^{i\theta}\right) d\theta \\
&= \int_{-\pi}^{\pi} -e^{-i\theta(n+1)} P^l_n(0) f(\theta) P^r_{n+1}\left(e^{i\theta}\right) d\theta \\
&= -P^l_n(0)S^r_{n+1},
\end{aligned}$$

where we used the facts that

$$\int_{-\pi}^{\pi} e^{-mi\theta} f(\theta)P^r_{n+1}(e^{i\theta})d\theta = 0 \text{ for } \ m < n+1; \quad \int_{-\pi}^{\pi} e^{-i\theta(n+1)} f(\theta)P^r_{n+1}(e^{i\theta})d\theta = S^r_{n+1}.$$

Similarly,

$$\int_{-\pi}^{\pi} P^l_n\left(e^{i\theta}\right) f(\theta) \frac{P^r_n\left(e^{i\theta}\right)}{e^{i\theta n}} d\theta = \int_{-\pi}^{\pi} \frac{P^l_n\left(e^{i\theta}\right)}{e^{in\theta}} f(\theta) P^r_n\left(e^{i\theta}\right) d\theta = -P^l_n(0)S^r_n.$$

By orthogonality,

$$\int_{-\pi}^{\pi} P^l_n\left(e^{i\theta}\right) f(\theta) P^{l*}_{n+1}\left(e^{i\theta}\right) d\theta = 0.$$

Finally,

$$P^l_n(0)S^r_{n+1} = P^l_n(0)S^r_n\Big(I - P^{l*}_{n+1}(0)P^r_{n+1}(0)\Big),$$

and the expression (7) follows. Identity (8) is proved similarly.

In order to prove (9) we rewrite expressions (1) and (3) from the proposition in the following way:

$$\frac{P^r_{n+1}(x)}{x^{n+1}} = \frac{P^r_n(x)}{x^n} + P^{l*}_n(x)P^r_{n+1}(0);$$

$$\frac{P^l_{n+1}(x)}{x^{n+1}} = \frac{P^r_n(x)}{x^n} + P^l_{n+1}(0)P^{r*}_n(x).$$

After multiplying the first expression by $P^l_n\left(e^{i\theta}\right)f(\theta)$ from the left and second expression by $f(\theta)P^r_{n+1}\left(e^{i\theta}\right)$ from the right, substituting $x = e^{i\theta}$ and integrating we arrive at:

$$\begin{aligned}\int_{-\pi}^{\pi} P^l_{n+1}\left(e^{i\theta}\right)f(\theta)\frac{P^r_{n+1}\left(e^{i\theta}\right)}{e^{i(n+1)\theta}}d\theta &= \int_{-\pi}^{\pi} P^l_{n+1}\left(e^{i\theta}\right)f(\theta)\frac{P^r_n\left(e^{i\theta}\right)}{e^{in\theta}}d\theta \\ &+ \left(\int_{-\pi}^{\pi} P^l_{n+1}\left(e^{i\theta}\right)f(\theta)\frac{P^{l*}_n\left(e^{i\theta}\right)}{e^{i\theta}}d\theta\right) \\ &= 0 - S^l_{n+1}P^r_{n+1}(0)\end{aligned}$$

and

$$\begin{aligned}\int_{-\pi}^{\pi} P^l_{n+1}\left(e^{i\theta}\right)f(\theta)\frac{P^r_{n+1}\left(e^{i\theta}\right)}{e^{i(n+1)\theta}}d\theta &= \int_{-\pi}^{\pi} P^l_n\left(e^{i\theta}\right)f(\theta)\frac{P^r_{n+1}\left(e^{i\theta}\right)}{e^{in\theta}}d\theta \\ &+ P^l_{n+1}(0)\int_{-\pi}^{\pi} P^{r*}_n\left(e^{i\theta}\right)f(\theta)\frac{P^r_{n+1}\left(e^{i\theta}\right)}{e^{i\theta}}d\theta \\ &= 0 - P^l_{n+1}(0)S^r_{n+1}.\end{aligned}$$

Hence, $P^l_{n+1}(0)S^r_{n+1} = S^l_{n+1}P^r_{n+1}(0)$, which concludes the proof of the proposition. ■

**Note 23** *Formulas similar to the ones in the proposition above are obtained in a different way and presented in [50].*

**Note 24** *In the classical theory of orthogonal polynomials on the unit circle, for example see [28], a family of monic orthogonal polynomials* $\{p_n(x)\}_{n=0}^{\infty}$ *as defined in 3.2 satisfies the following recursion relations:*

1. $p_n(x) = xp_{n-1}(x) + p_n(0)p^*_{n-1}(x)$, *(forward recurrence relation);*
2. $p_n(x) = \left(1 - |p_n(0)|^2\right)xp_{n-1}(x) + p_n(0)p^*_n(x)$, *(backward recurrence relation), where* $p^*_n(x) = x^n\overline{p_n(1/\overline{x})}$ *and "‾" denotes complex conjugation.*

## 3.4 The Christoffel-Darboux formula: the case of the unit circle

In this section a matrix valued form of the Christoffel-Darboux formula will be derived.

Along with monic orthogonal polynomials, one can introduce orthonormal matrix valued polynomials on the unit circle.

Define families $\{Q_n^r(x)\}_{n=0}^{\infty}$ and $\left\{Q_n^l(x)\right\}_{n=0}^{\infty}$ such that

$$Q_n^r(x) = P_n^r(x) S_n^{-r/2} \text{ and } Q_n^l(x) = S_n^{-l/2} P_n^l(x), \tag{3.5}$$

then

$$\left\langle Q_n^{r,l}, Q_n^{r,l} \right\rangle_{r,l} = S_n^{-r,l/2} \left\langle P_n^{r,l}, P_n^{r,l} \right\rangle S_n^{-r,l/2} = S_n^{-r,l/2} S_n^{r,l} S_n^{-r,l/2} = I.$$

The following lemma introduces the matrix valued kernel polynomial on the unit circle.

**Lemma 11** *Given two families of orthonormal polynomials on the unit circle as defined in 3.5, denote the right and the left kernel polynomials of degree* $n$ *to be* $K_n^r(x,y)$ *and* $K_n^l(x,y)$*. Then*

1. $K_n^r(x,y) = \displaystyle\sum_{i=0}^{n} Q_i^r(y) Q_i^{r*}(x) = \begin{bmatrix} I & yI & \dots & y^n I \end{bmatrix} T_{n+1}^{-r} \begin{bmatrix} I \\ x^{-1} I \\ \vdots \\ x^{-n} I \end{bmatrix}$;

2. $K_n^l(x,y) = \displaystyle\sum_{i=0}^{n} Q_i^{l*}(y) Q_i^l(x) = \begin{bmatrix} I & y^{-1} I \dots & y^{-n} I \end{bmatrix} T_{n+1}^{-l} \begin{bmatrix} I \\ xI \\ \vdots \\ x^n I \end{bmatrix}$;

3. *Christoffel-Darboux formula*

$$K_n^r(x,y) = \frac{\hat{Q}_{n+1}^l(x) \hat{Q}_{n+1}^{l*}(y) - Q_{n+1}^r(x) Q_{n+1}^{r*}(y)}{I - x\bar{y}};$$

4. *Christoffel-Darboux formula*

$$K_n^l(x,y) = \frac{\hat{Q}_{n+1}^{r*}(x) \hat{Q}_{n+1}^r(y) - Q_{n+1}^{l*}(x) Q_{n+1}^l(y)}{I - x\bar{y}}.$$

**Proof:** Proofs for identities (1) and (2) are very similar to the proof presented in lemma 2.29 for the case of the real line.

In order to derive the Christoffel-Darboux formula (3) we write the following two recursion relations for orthonormal polynomials:

$$Q_{n+1}^r(t) = t Q_n^r(t) a + \hat{Q}_n^l(t) b;$$

$$\hat{Q}_{n+1}^l(t) = \hat{Q}_n^l(t) c + t Q_n^r(t) d;$$

where $a = S_n^{r/2} S_{n+1}^{-r/2}$, $b = S_n^{l/2} P_{n+1}^r(0) S_{n+1}^{-r/2}$, $c = S_n^{l/2} S_{n+1}^{-l/2}$ and $d = S_n^{r/2} P_{n+1}^{l*}(0) S_{n+1}^{-l/2}$. In the matrix form this could be written as

$$\Phi_{n+1}(t) = C(t) \Phi_n(t);$$

where $\Phi_{n+1}(t) = \left[Q^r_{n+1}(t); \hat{Q}^l_{n+1}(t)\right]$ and $C(t) = \begin{pmatrix} ta & td \\ b & c \end{pmatrix}$.

Define $J = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$, and note that

$$\begin{aligned} dc^* - ab^* &= S_n^{r/2} P^{l*}_{n+1}(0) S_{n+1}^{-l/2} S_{n+1}^{-l/2} S_n^{l/2} - S_n^{r/2} S_{n+1}^{-r/2} S_{n+1}^{-r/2} P^{r*}_{n+1}(0) S_n^{l/2} \\ &= S_n^{r/2} \left( P^{l*}_{n+1}(0) S^{-l}_{n+1} - S^{-r}_{n+1} P^{r*}_{n+1}(0) \right) S_n^{l/2} = 0 \end{aligned}$$

according to the identity (9) in the proposition in the previous section. Similarly,

$$\begin{aligned} dd^* - aa^* &= S_n^{r/2} P^{l*}_{n+1}(0) S_{n+1}^{-l/2} S_{n+1}^{-l/2} P^{l}_{n+1}(0) S_n^{r/2} - S_n^{r/2} S_{n+1}^{-r/2} S_{n+1}^{-r/2} S_n^{r/2} \\ &= S_n^{r/2} \left( P^{l*}_{n+1}(0) S^{-l}_{n+1} P^{l}_{n+1}(0) - S^{-r}_{n+1} \right) S_n^{r/2} = -I \end{aligned}$$

and

$$\begin{aligned} cc^* - bb^* &= S_n^{l/2} S_{n+1}^{-l/2} S_{n+1}^{-l/2} S_n^{l/2} - S_n^{l/2} P^{r}_{n+1}(0) S_{n+1}^{-r/2} S_{n+1}^{-r/2} P^{r*}_{n+1}(0) S_n^{l/2} \\ &= S_n^{l/2} \left( S^{-l}_{n+1} - P^{r}_{n+1}(0) S^{-r}_{n+1} P^{r*}_{n+1}(0) \right) S_n^{l/2} = I. \end{aligned}$$

Hence

$$C(x) J C^*(y) = \begin{pmatrix} x\bar{y}(dd^* - aa^*) & x(dc^* - ab^*) \\ \bar{y}(cd^* - ba^*) & cc^* - bb^* \end{pmatrix} = \begin{pmatrix} -x\bar{y}I & 0 \\ 0 & I \end{pmatrix},$$

which implies that

$$\begin{aligned} \Phi_{n+1}(x) J \Phi^*_{n+1}(y) &= \hat{Q}^l_{n+1}(x) \hat{Q}^{l*}_{n+1}(y) - Q^r_{n+1}(x) Q^{r*}_{n+1}(y) \\ &= \Phi_n(x) C(x) J C^*(y) \Phi^*_n(y) = \hat{Q}^l_n(x) \hat{Q}^{l*}_n(y) - x\bar{y} Q^r_n(x) Q^{r*}_n(y). \end{aligned}$$

Thus,

$$\sum_{k=0}^{n} Q^r_k(x) Q^{r*}_k(y) = \frac{\hat{Q}^l_{n+1}(x) \hat{Q}^{l*}_{n+1}(y) - Q^r_{n+1}(x) Q^{r*}_{n+1}(y)}{I - x\bar{y}}.$$

Identity (4) is proved similarly. ■

# Chapter 4

# Bochner's problem

## 4.1 Introduction

In the classical paper [5], Bochner characterized all families of polynomials $\{P_n(x)\}_{n=0}^{\infty}$ with the following property:

$$LP^* = xP^* \quad \text{and} \quad \mathbf{D}P = P\Lambda \tag{4.1}$$

where $L$ is defined in 2.15; $P = [P_0(x),\ P_1(x),\ \cdots]$; $\Lambda = [\Lambda_0,\ \Lambda_1,\ \cdots]$;

$$\mathbf{D} = B_2(x)\frac{d^2}{dx^2} + B_1(x)\frac{d}{dx} + B_0,$$

$$B_2(x) = \alpha_2 x^2 + \alpha_1 x + \alpha_0;$$

$$B_1(x) = \beta_1 x + \beta_0;$$

$$B_0 = \gamma_0,$$

with $\alpha_2,\ \alpha_1,\ \alpha_0,\ \beta_1,\ \beta_0,\ \gamma_0,\ \Lambda_n \in \mathbf{R}^{k,k}$, and $\Lambda_n$ depends on $n$, but not on $x$.

Let us call the problem of finding families of polynomials which satisfy both a three term recursion relation as well as a second order differential equation a "**Bochner's problem.**"

**Note 25** *In the classical theory of scalar valued polynomials, Bochner's problem is solved and the families of polynomials which possess the two properties mentioned above are called the "classical orthogonal polynomials" (Jacobi, Hermite, Laguerre and Bessel).*

## 4.2 Matrix formulation: "ad" condition

In this section we will derive several equivalent formulations of Bochner's problem involving the "ad" condition, which was first introduced by Duistermaat and Grünbaum in the continuous-continuous version, see [9], and adapted in the continuous-discrete case by Grünbaum and Haine in [33]. These conditions were used in the matrix valued case in [34]. For a collection of examples obtained by solving the equations resulting from the "ad" condition in the case of differential operators of order one see [6].

**Proposition 7** *Given a family of monic matrix valued polynomials* $\{P_n(x)\}_{n=0}^{\infty}$ *which satisfy a three term recursion relation* $LP^* = xP^*$, *the following conditions are equivalent:*

1. $\boldsymbol{D}P = P\Lambda$, *where* $\boldsymbol{D}$ *is defined in 4.1;*
2. $2B_2P' + B_1P = P(\Lambda L^* - L^*\Lambda) = P\Lambda(L^* - xI) = P\ \mathrm{ad}(L^*)(\Lambda)$;
3. $2B_2P = P\Big((L^*)^2\Lambda - 2L^*\Lambda L^* + \Lambda(L^*)^2\Big) = P\ \mathrm{ad}(L^*)^2(\Lambda) = P\Lambda(xI - L^*)^2$;
4. $(\mathrm{ad}L^*)^3(\Lambda) = 0$, *where* $\mathrm{ad}(A)(B) \equiv AB - BA$.

***Proof:***

**Note 26** *If* $PL^* = xP$, *where* $P$ *is a vector containing the monic orthogonal polynomials, then any other vector* $Q = [Q_0(x),\ Q_1(x),\ \cdots]$ *(where* $i$ *denotes the degree of the polynomial) such that* $QL^* = xQ$ *can be expressed as* $Q = MP$, *where* $M$ *is some matrix, i.e.* $Q_i(x) = MP_i(x)$ *for all* $i \geq 0$.

The equation $PL^* = xP$ implies

$$P'L^* = xP' + P \text{ and } P''L^* = xP'' + 2P'. \tag{4.2}$$

To prove (1) ⇒ (2) we substitute 4.2 into $\mathbf{D}P = P\Lambda$ and after simple manipulations we obtain statement (2) in the proposition.

In order to prove (2) ⇒ (1) substitute

$$P' = (P''L^* - xP'')/2 \quad \text{and} \quad P = P'L^* - xP'$$

into (2) and obtain the expression

$$\big(B_2P'' + B_1P' - P\Lambda\big)L^* = \big(B_2P'' + B_1P' - P\Lambda\big)x.$$

As observed in the note at the beginning of the proof, there must exist a matrix $B_0$ such that

$$B_2P'' + B_1P' - P\Lambda = -B_0P.$$

Observe that by comparing powers of $x$ in the expression above we conclude that $B_0$ has to be a constant matrix, and statement (1) of the proposition follows.

In order to show (2) ⇔ (3) and (3) ⇔ (4) the same technique is used. To illustrate, let us show that (4) ⇒ (3). Observe that

$$(adL^*)^3(\Lambda) = 0 \Leftrightarrow P\Lambda(xI - L^*)^3 = 0 \Leftrightarrow P\Lambda(x^3I - 3x^2L^* + 3xL^{*2} - L^{*3}) = 0.$$

Denote $Q = P\Lambda(x^2 I - 2xL^* + L^{*2})$, then one can see that $QL^* = xQ$, which implies that $Q = 2B_2 P$, for some matrix $B_2$. By comparing powers of $x$ we conclude that $B_2$ has to be quadratic in $x$, i.e. $B_2 = x^2\alpha_2 + x\alpha_1 + \alpha_0$, where $\alpha_2,\ \alpha_1$ and $\alpha_0$ are some coefficients independent on $x$. We have proved that $(adL^*)^3(\Lambda) = 0$ implies

$$P\Lambda(x^2 I - 2xL^* + L^{*2}) = P\Lambda(xI - L^*)^2 = B_2 P,$$

which is exactly condition (3) of the proposition. ■

**Note 27** *Let us consider condition* (4) *in more detail. Introduce a seven diagonal matrix* $M$, *such that*

$$M = L^{*3}\Lambda - 3L^{*2}\Lambda L^* + 3L^*\Lambda L^{*2} - \Lambda L^{*3}. \tag{4.3}$$

*Since* $LS = SL^*$, *condition 4.3 can be written as*

$$SMS^{-1} = L^3\tilde{\Lambda} - 3L^2\tilde{\Lambda}L + 3L\tilde{\Lambda}L^2 - \tilde{\Lambda}L^3 \tag{4.4}$$

*where* $\tilde{\Lambda} = S\Lambda S^{-1}$. *If we transpose condition 4.4 and change the sign, we obtain*

$$-(SMS^{-1})^* = L^{*3}\tilde{\Lambda}^* - 3L^{*2}\tilde{\Lambda}^*L^* + 3L^*\tilde{\Lambda}^*L^{*2} - \tilde{\Lambda}^*L^{*3}. \tag{4.5}$$

*From equations 4.3 and 4.5 above we conclude the following:*

1. *The equations resulting from equating the upper three diagonals of* $M$ *to* 0 *are the same as the equations for the lower three diagonals, but with* $\Lambda$ *substituted with* $\tilde{\Lambda}^*$.

2. *If* $\tilde{\Lambda}^* = \Lambda$, *(which is equivalent to* ***D*** *being symmetric, will be discussed later), then* $M = -(SMS^{-1})^* = -S^{-1}M^*S$. *This implies that when the upper three diagonals of* $M$ *are zero, the lower three diagonals of* $M$ *will be zero automatically.*

3. *In the scalar case* $\tilde{\Lambda}^* = \Lambda$, *hence the top three diagonals of the matrix* $M$ *being zero implies that the bottom three diagonals are also zero. The diagonal entries* $M_{n,n}$ *can be easily checked to be identically zero in the scalar case. This implies that in the scalar case the matrix* $M$ *is identically zero once its top three diagonals are zero. In the matrix case this is known not to be true, i.e. it is possible to have the top three diagonals of* $M$ *to be zero and have non-zero entries in the rest of the matrix* $M$.

## 4.3 Matrix formulation: direct computation

Given a family of monic orthogonal polynomials $\{P_n(x)\}_{n=0}^{\infty}$ as defined in 2.2, consider a second order differential operator

$$\Big(\alpha_2 x^2 + \alpha_1 x + \alpha_0\Big)P_n''(x) + \Big(\beta_1 x + \beta_0\Big)P_n'(x) + \gamma_0 P_n(x) = P_n(x)\Lambda_n, \tag{4.6}$$

where $\alpha_2,\ \alpha_1,\ \alpha_0,\ \beta_1,\ \beta_0,\ \gamma_0,\ \Lambda_n \in \mathbf{R}^{k\times k}$. By comparing powers of $x$ in 4.6 we arrive at the set of conditions summarized below:

1. $\Lambda(n) = \alpha_2 n^2 + (-\alpha_2 + \beta_1)n + \gamma_0 = n(n-1)\alpha_2 + n\beta_1 + \gamma_0;$
   $\gamma_0 = \Lambda_0; \quad \beta_1 = \Lambda_1 - \Lambda_0; \quad \alpha_2 = \dfrac{\Lambda_2 - 2\Lambda_1 + \Lambda_0}{2};$
   $\Lambda_{n+1} - 3\Lambda_n + 3\Lambda_{n-1} - \Lambda_{n-2} = 0;$

2. $\beta_0 = \Lambda_0 u_0^0 - u_0^0 \Lambda_1;$
3. $\alpha_1 = \dfrac{\Lambda_{n-1}u_{n-1}^{n-1} - u_{n-1}^{n-1}\Lambda_n}{n(n-1)} - \dfrac{n\beta_0}{n(n-1)};$
4. $\alpha_0 = \dfrac{(n-1)\Big((n-2)\alpha_1 + \beta_0\Big)}{n(n-1)} u_{n-1}^{n-1} + \dfrac{\Lambda_{n-2}u_{n-2}^{n-1} - u_{n-2}^{n-1}\Lambda_n}{n(n-1)};$
5. $\alpha_0 u_{n-i+2}^{n-1} = -\dfrac{\Lambda_{n-i}u_{n-i}^{n-1} - u_{n-i}^{n-1}\Lambda_n}{(n-i+2)(n-i+1)} - \dfrac{(n-i+1)\Big((n-i)\alpha_1 + \beta_0\Big)}{(n-i+2)(n-i+1)} u_{n-i+1}^{n-1}$, for $i = 3, ...n-1$.

The formulas above can be written more compactly in a matrix form. Using the previous notation write $P = \Omega R^{-1}$ and substitute this expression into the differential equation 4.6. The following expression is obtained

$$B_2(x)\Omega'' + B_1(x)\Omega' + B_0\Omega = \Omega R^{-1}\Lambda R = \Omega K \quad \text{and} \quad K = R^{-1}\Lambda R.$$

By looking at the matrix $\Omega$ we arrive at the following formulas for the matrix $K$ in terms of the coefficients of the differential equation:

$$\begin{aligned} K_{n-2,n} &= n(n-1)\alpha_0; \\ K_{n-1,n} &= n(n-1)\alpha_1 + n\beta_0; \\ K_{nn} &= n(n-1)\alpha_2 + n\beta_1 + \gamma_0 \end{aligned} \tag{4.7}$$

for $n \geq 0$. The matrix $K$ is upper tridiagonal, and

$$RK = \Lambda R. \tag{4.8}$$

From the equation above we see that the problem of generating polynomials that satisfy a given differential equation is equivalent to the problem of diagonalizing a tridiagonal matrix $K$, which is always possible to do by solving the following Sylvester's equations with respect to $r_{i,j}$ :

$$r_{i,j-2}K_{j-2,j} + r_{i,j-1}K_{j-1,j} = \Lambda_i r_{i,j} - r_{i,j}\Lambda_j.$$

This will have a unique solution provided all $\Lambda_n$'s have disjoint spectrums. Note that expression 4.8 is equivalent to formulas 1-5 above.

We can now reformulate problem 4.6 in the following way:

**Summary 1**

- $R^*SR = H$ : *this equation describes the situation when we have a family of orthogonal polynomials defined by a matrix* $R$ *as a result of the Gramm-Schmidt procedure on the space spanned by* $\{x^n I\}_{n=0}^{\infty}$ .
- $RK = \Lambda R$ : *this equation describes the situation when the family of polynomials defined by* $R$ *satisfies a second order differential equation defined by* $K$.

*Problem 4.1 is now equivalent to both of the equations above being true for some symmetric measure* $\mu(dx) = W(x)dx$ *defining a matrix of moments* $H$ *and a matrix* $K$ *defining the differential operator.*

The matrix formulation above is used in the next section to derive moment equations.

## 4.4 The case of a symmetric operator D

Let us consider the specific case of a symmetric differential operator $\mathbf{D}$, i.e. for any polynomials $P$ and $Q$

$$\langle Q, \mathbf{D}P \rangle = \langle \mathbf{D}Q, P \rangle, \tag{4.9}$$

where the norm $\langle , \rangle$ is defined in 2.5.

**Note 28** *It is clear that the orthonormalized polynomials satisfy the same differential equation* $\boldsymbol{D}\bar{P}_n = \bar{P}_n \bar{\Lambda}_n$ *as the monic ones, except for the eigenvalue matrices which now are* $\bar{\Lambda}_n = S_n^{1/2} \Lambda_n S_n^{-1/2}$. *From the two equations below*

$$\left\langle \bar{P}, \boldsymbol{D}\bar{P} \right\rangle = \int \bar{P}^* W(x) \boldsymbol{D}\bar{P} dx = \int \bar{P}^* W(x) \bar{P} \bar{\Lambda} dx = \bar{\Lambda};$$

$$\left\langle \boldsymbol{D}\bar{P}, \bar{P} \right\rangle = \int \left(\boldsymbol{D}\bar{P}\right)^* W(x) \bar{P} dx = \int \bar{\Lambda}^* \bar{P}^* W(x) \bar{P} dx = \bar{\Lambda}^*$$

*it is clear that the operator* $\boldsymbol{D}$ *being symmetric is equivalent to* $\bar{\Lambda}$ *being Hermitian, i.e.*

$$\boldsymbol{D} \text{ is symmetric if and only if } S_n \Lambda_n = \Lambda_n^* S_n,$$

*or*

$$S\Lambda = (S\Lambda)^* = \Lambda^* S. \tag{4.10}$$

In the summary 1 at the end of the last section we derived the equivalence of the problem 4.1 to the two conditions, which imply the following:

$$\begin{cases} \Lambda = RKR^{-1} \\ H = R^*SR \end{cases} \Rightarrow S\Lambda = SRKR^{-1} = (R^*)^{-1} H R^{-1} R K R^{-1} = (R^*)^{-1} H K R^{-1}.$$

From (4.10) it follows that

$$S\Lambda = (S\Lambda)^* = (R^*)^{-1}HKR^{-1} = \left((R^*)^{-1}HKR^{-1}\right)^* = (R^*)^{-}K^*HR^{-1},$$

which implies that

$$HK = K^*H \tag{4.11}$$

Problem 4.1 in the case of the symmetric differential operator $\mathbf{D}$ is now equivalent to the matrix $HK$ being Hermitian, as we see below.

**Lemma 12** *Given very particularly structured matrices $K$ and $H$ such that*

1. *$H$ is a Hankel matrix*
2. *$K$ is defined in (4.7)*
3. *The matrix $HK$ is Hermitian,*

*then there exists a family of orthonormal polynomials defined by the matrix $H$ satisfying a second order differential equation defined by the matrix $K$.*

***Proof:*** After performing the Cholesky decomposition on matrix $H$ one obtains

$$H = \bar{R}^*\bar{R} = R^*SR,$$

where $R$ is unit block upper triangular, $\bar{R} = S^{1/2}R$ and a family of orthogonal polynomials is defined by $\Omega R^{-1} = P$. The matrix $HK$ being Hermitian implies

$$\bar{R}^*\bar{R}K = K^*\bar{R}^*\bar{R},$$

hence

$$\bar{R}K\bar{R}^{-1} = (\bar{R}^*)^{-1}K^*\bar{R}^*. \tag{4.12}$$

Matrices $R$ and $K$ are upper triangular, hence $\bar{R}K\bar{R}^{-1}$ is upper triangular. For the same reason $(\bar{R}^*)^{-1}K^*\bar{R}^*$ is lower triangular, which implies that equation (4.12) in fact can be written as

$$\bar{R}K\bar{R}^{-1} = \Lambda,$$

where $\Lambda$ is the diagonal of $K$. The expression above is equivalent to having a family of orthogonal polynomials $P$ satisfy a second order differential equation defined by the matrix $K$. ■

In the next proposition element-wise conditions for the matrix $HK$ to be symmetric are derived.

**Proposition 8** *Given a semi-infinite Hankel matrix $H$ and a matrix $K$ with structure described in (4.7), then the matrix $HK$ being Hermitian is equivalent to the following set of equations:*

$$A_2 = A_2^*; \tag{4.13}$$

$$-2(k+1)A_2 = A_1 + A_1^*; \tag{4.14}$$

$$A_2(k+1)(k+2) + A_1(k+2) + A_0 = A_0^*; \tag{4.15}$$

*where*

$$A_2 = \mu_{k+2}\alpha_2 + \mu_{k+1}\alpha_1 + \mu_k\alpha_0,$$

$$A_1 = \mu_{k+2}\beta_1 + \mu_{k+1}\beta_0;$$

$$A_0 = \mu_{k+2}\gamma_0,$$

*for* $k \geq 0$.

*In other words, having a family of matrix valued orthogonal polynomials that are eigenfunctions of a symmetric second order differential operator is equivalent to formulas (4.13), (4.14) and (4.15).*

**Note 29** *The formulas 4.13 , 4.14 and 4.15 were first discovered by A. Duran in [14]. See also [18].*

***Proof:*** Denote the matrix $Z = HK - K^*H$ and consider the "right-to-left" diagonals of $Z$, i.e. all $i$ and $j$ such that $i + j = m$, where $m \geq 0$.

Denote $J_1(n) = n(n-1)$, $J_2(n) = n$, then

$$\begin{aligned}
Z_{i,j} &= \mu_{m-2}J_1(j)\alpha_0 + \mu_{m-1}\Big(J_1(j)\alpha_1 + J_2(j)\beta_1\Big) + \mu_m\Big(J_1(j)\alpha_2 + J_2(j)\beta_1 + \gamma_0\Big) \\
&- \alpha_0^* J_1(i)\mu_{m-2} - \Big(\alpha_1^* J_1(i) + \beta_0^* J_2(i)\Big)\mu_{m-1} + \Big(\alpha_2^* J_1(j) + \beta_1^* J_2(i) + \gamma_0^*\Big)\mu_m \\
&= \Big(\mu_m\alpha_2 + \mu_{m-1}\alpha_1 + \mu_{m-2}\alpha_0\Big)J_1(j) + \Big(\mu_m\beta_1 + \mu_{m-1}\beta_0\Big)J_2(j) + \mu_m\gamma_0 \\
&- \Big(\alpha_0^*\mu_m + \alpha_1^*\mu_{m-1} + \alpha_0^*\mu_{m-2}\Big)J_1(i) - \Big(\beta_1^*\mu_m + \beta_0^*\mu_{m-1}\Big)J_2(i) - d_0^*\mu_{m-2}.
\end{aligned}$$

For simplicity, lets denote

$$A_2 = \mu_m\alpha_2 + \mu_{m-1}\alpha_1 + \mu_{m-2}\alpha_0;$$

$$A_1 = \mu_m\beta_1 + \mu_{m-1}\beta_0;$$

$$A_0 = \mu_m\gamma_0.$$

Using this new notation we put $Z = HK - K^*H = 0$ and write the following system of equations

$$\begin{cases} Z_{i,j} = A_2J_1(j) + A_1J_2(j) + A_0 - A_2^*J_1(i) - A_1^*J_2(i) - A_0^* = 0 \\ Z_{i',j'} = A_2J_1(j') + A_1J_2(j') + A_0 - A_2^*J_1(i') - A_1^*J_2(i') - A_0^* = 0 \end{cases}$$

for some $(i', j')$ such that $i' + j' = m$. After subtracting the first equation from the second one we arrive at the expression

$$A_2(j + j' - 1) + A_2^*(2m - j - j' - 1) + A_1 + A_1^* = 0, \tag{4.16}$$

which must also hold for any pair $(i'', j'')$ such that $i'' + j'' = m$, i.e.

$$A_2(j + j'' - 1) + A_2^*(2m - j - j'' - 1) + A_1 + A_1^* = 0. \tag{4.17}$$

Subtracting (4.16) from (4.17) we obtain

$$A_2(j' - j'') - A_2^*(j' - j'') = 0,$$

which implies that

$$A_2 = A_2^*.$$

Substituting this into (4.16) gives

$$2(1 - m)A_2 = A_1 + A_1^*;$$

and plugging this into the equation for $Z_{i,j}$ we obtain

$$A_2(m - 1)m + A_1 m + A_0 = A_0^*,$$

which concludes the proof of the proposition. ■

**Note 30** *Rewriting equation (4.13) in the following way*

$$\begin{aligned} A_2 - A_2^* &= \int \Big( W(x)\left(x^{k+2}\alpha_2 + x^{k+1}\alpha_1 + x^k\alpha_0\right) - \left(x^{k+2}\alpha_2^* - x^{k+1}\alpha_1^* - x^k\alpha_0^*\right) W(x) \Big) dx \\ &= \int x^k \Big( W(x)B_2 - B_2^* W(x) \Big) dx = 0, \text{ for all } \ k \geq 0, \end{aligned}$$

*implies, under appropriate conditions,*

$$W(x)B_2 = B_2^* W(x). \tag{4.18}$$

*Equation (4.14) can similarly be rewritten as*

$$\begin{aligned} 2A_2(k+1) + A_1 + A_1^* &= \int \Big( 2(k+1)x^k B_2 W(x) + x^{k+1} W(x) B_1 + x^{k+1} B_1^* W(x) \Big) dx \\ &= \int x^{k+1} \Big( -2\big(W(x)B_2\big)' + W(x)B_1 + B_1^* W(x) \Big) dx = 0, \end{aligned}$$

*which implies*

$$2\Big(W(x)B_2\Big)' = W(x)B_1 + B_1^* W(x), \tag{4.19}$$

*with $W(x)B_2(x)$ vanishing at the boundary of the support of the measure. Equation (4.15) can be rewritten as*

$$\Big(W(x)B_2\Big)'' - \Big(W(x)B_1\Big)' + W(x)B_0 = B_0^* W(x), \tag{4.20}$$

*with $\big(W(x)B_2\big)' - W(x)B_1$ vanishing at the boundary.*

*The equations 4.18, 4.19, 4.20 and the boundary conditions mentioned above first appeared in [18] and [38].*

Further research is being done in the direction of better understanding the Bochner's problem in the matrix setting. The matrix case is richer in examples and hopefully will help to describe a variety of physical phenomena. There are many promising connections of this subject with other disciplines and hopefully more applications will be discovered.

# Bibliography

[1] N.I. Akhiezer, *The classical moment problem and some related questions in analysis*, English translation, Oliver and Boyd, Edinburgh, 1965.

[2] W. A. Al-Salam, T. S. Chihara, *Another characterization of the classical orthogonal polynomials,* SIAM J. Math, Anal., 3 (1972), pp. 65 - 70.

[3] A. I. Aptekarev, E. M. Nikishin, *The scattering problem for discrete Sturm-Liouville operator*, Math., USSR-Sb, 49, pp. 325-355.

[4] Yu. M. Berezanskii, *Expansions of eigenfunctions of self-adjoined operators,* Transl. Math. Monographs, Vol 17. (AMS, Prividence, RI, 1968).

[5] S. Bochner, *Über Sturm-Liouvillesche Polynomsysteme,* Math. Z. 29 (1929), 730–736

[6] M. Castro, A. Grünbaum, *Orthogonal matrix polynomials satisfying first order differential equations: a collection of instructive examples*, to appear in Journal of Nonlinear Mathematical Physics.

[7] T.S. Chihara, *An introduction to orthogonal polynomials*, Gordon and Breach Science Publishers, 1978.

[8] C. W. Cryer, *Rodrigues' formula and the classical orthogonal polynomials,* Boll. Un. Mat. Ital., (3) 25 (1970), pp. 1-11.

[9] J. J. Duistermaat, F. A. Grünbaum, *Differential equations in the spectral parameter,* Comm. Math. Phys. 103, 2 (1986), pp. 177–240. Math.Sci.Net. Zentralblatt für Mathematik.

[10] A. J. Duran, *A generalization of Favard's theorem for polynomials satisfying a recurrence relation,* J. Approximation theory, 74 (1997), pp. 7818-7829.

[11] A. J. Duran, *On orthogonal polynomials with respect to positive definite matrix of measures,* Can. J. Math., 47 (1995), pp. 88-112.

[12] A. J. Duran, *Markov's theorem for orthogonal matrix polynomials,* Can. J. Math., 48 (1996), pp. 1180-1195.

[13] A. J. Duran, *Ratio asymptotics for orthogonal matrix polynomials,* J. Approx. Th. 100, (1999) pp. 304-344.

[14] A. J. Duran, *Matrix inner product having a matrix symmetric second order differential operator,* Rocky Mount. J. Math. 27 (1997), pp. 585-600.

[15] A. J. Duran, *Functions with given moments and weight functions for orthogonal polynomials,* Rocky Mount. J. Math. 23 (1993), pp. 87-104.

[16] A. J. Duran, W. Van Assche, *Orthogonal matrix polynomials and higher order recurrence relations,* Linear Algebra and Applications, 219 (1995), pp. 261-280.

[17] A.J. Duran, F.A. Grünbaum, *A characterization for a class of weight matrices with orthogonal matrix polynomials satisfying second order differential equations,* submitted for publication.

[18] A. J. Duran, F. A. Grünbaum, *Orthogonal matrix polynomials satisfying second order differential equations*, Int. Math. Res. Not., 10:461-484, 2004.

[19] A. J. Duran, F. A. Grünbaum, *Orthogonal matrix polynomials, scalar type Rodrigues' formula and Pearson type equations*, to appear in Journal of Approximation Theory.

[20] A. J. Duran, F. A. Grünbaum, *Structural Formulas for Orthogonal Matrix Polynomials Satisfying Second Order Differential Equations*, to appear in I. Constr Approx. 2005.

[21] A. Duran, P. Lopez, *Orthogonal matrix polynomials: zeros and Blumenthal's theorem,* J. Approx. Theory, 84 (1996) pp. 96-118.

[22] A. Duran, P. Lopez, *Orthogonal matrix polynomials,* Laredo's SIAG Lecture Notes, R. Alvarrez-Nodarse et. al. editors, Advances in the theory of special functions and orthogonal polynomials, Nova Science Publishers, 1 (2003), pp. 11-43.

[23] A. J. Duran, B. Polo, *Gaussian formulae for matrix weights,* Linear Algebra and Applications, 355, (2002). pp. 119-146.

[24] R. Ebert, *Über Polynomsysteme mit Rodriguesscher Darstellung,* Dissertation, Cologne, 1964.

[25] P.A. Fuhrmann, *Orthogonal matrix polynomials and system theory,* Rend. Sem. Mat. Università Politecnica. Torino, (1988), pp. 68-124.

[26] M. Gekhtman, *Hamiltonian structure of non-abelian Toda lattice*, Letters in Math. Physics, 46 (1998), pp. 189-205.

[27] J.S. Geronimo, *Scattering theory and matrix orthogonal polynomials on the real line,* Circuits Systems Signal Process 1 (1982), pp. 471-495.

[28] Ya. L. Geronimus, *Polynomials orthogonal on a circle and their applications,* Amer. Math. Soc. Translations, 1954

[29] G. Golub, C. V. Loan, *Matrix computations,* The Johns Hopkins University Press, Baltimore, second ed., 1989.

[30] G. Golub, R. Underwood, *The block Lanczos methods for computing eigenvalues,* in Mathematical Software III, J. R. Rice, ed., Academic Press, New York, 1977, pp. 364-377.

[31] F. A. Grünbaum, *Matrix valued Jacobi polynomials*, Bull. Sci. math. 127 (2003). pp. 207-214.

[32] F.A. Grünbaum and L. Haine, *Orthogonal Polynomials Satisfying Differential Equations: The Role of the Darboux Transformation*, CRM Proceedings, Lecture Notes (9), AMS Providence, RI 1996, pp. 143-154.

[33] F.A. Grünbaum and L. Haine, *A theorem of Bochner revisited,* A.S. Fokas and I. M. Gelfand (eds.), Algebraic Aspects of integrable systems; Progr. Nonlinear Differential equations 26, Birkhauser, Boston, (1997), pp. 143-172.

[34] F. A. Grünbaum, P. Iliev, *A noncommutative version of the bispectral problem,* J. Comput. Appl. Math., 161 (2003), 99–118.

[35] F. A. Grünbaum, I. Pacharoni, J. Tirao, *Matrix valued spherical functions associated to the complex projective plane,* J. Functional Analysis, 188 (2002), pp. 350-441.

[36] F. A. Grünbaum, I. Pacharoni, J. Tirao, *A matrix valued solution to Bochner's problem,* J. Physics A: Math. Gen. 34 (2001), 10647-10656.

[37] F. A. Grünbaum, I. Pacharoni, J. Tirao, *An invitation to matrix matrix valued spherical functions: Linearization of products in the case of the complex projective space* $P_2(\boldsymbol{C})$, arXiv math. RT/0202304.

[38] F. A. Grünbaum, I. Pacharoni, J. Tirao, *Matrix valued orthogonal polynomials of the Jacobi type,* Indag. Mathem. V. 14: 3,4 (2003), pp. 353-366.

[39] W. Hahn, *Über die Jacobischen Polynome und zwei verwandte Polynomklassen,* Math. Zeit., 39 (1935), pp. 634-638.

[40] L. Haine, E. Horozov, *Toda orbits of Laguerre polynomials and representations of the Virasoro algebra,* Bull. Sci. Math. 117 (1993), no. 4, 485-518.

[41] M. G. Krein, *Fundamental aspects of the representation theory of Hermitian operators with deficiency index* $(m.m)$, AMS Translations, Series 2, 97, Providence, Rhode Island, (1971), 75-143.

[42] M.G. Krein, *Infinite J-matrices and a matrix moment problem,* Dokl. Akad. Nauk SSSR, 69 nr. 2, 2 (1949), pp. 125-128.

[43] S. Karlin, G. Szegó, *On certain determinants whose elements are orthogonal polynomials,* J. a'Analyse Math., 8 (1961) pp. 1-157.

[44] N. N. Lebedev, *Special Functions and their Applications*, Prentice-Hall, Englewood, N.J., 1965.

[45] F. Marcellán, I. R. Gonzàles, *A class of matrix orthogonal polynomials on the unit circle,* Linear Algebra and Appl., 121 (1989), pp. 233-241.

[46] F. Marcellán, G. Sansigre, *On a class of matrix orthogonal polynomials on the real line,* Linear Algebra and Applications, 181 (1993), pp. 97-109.

[47] M. P. Mignolet, *Matrix polynomials orthogonal on the unit circle and accuracy of autoregressive models,* J. Comp. Appl. Math., 62 (1989), pp. 229-238.

[48] P. van Moerbeke, *Integrable lattices: random matrices and random permutations,* Random matrices and their applications, MSRI publications, (2001), 40.

[49] E. M. Nikishin, *Discrete Sturm-Liouville operators and some problems of function theory,* Trudy Sem. Petrovsk., 10 (1984) 3-77; J. Soviet Math., 35 (1986), pp. 2679-2744.

[50] A. Sinap, W. Van Assche, *Orthogonal matrix polynomials and applications,* Journal of Computational and Applied Mathematics, 66(1996), pp. 27-52.

[51] G. Szegó, *Orthogonal polynomials,* Fourth Edition, Amer. math. Soc., Providence, 1975.

[52] F. Tricomi, *Equazioni differentiali,* Torino, 1948.