Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 21 (2002), No. 3, 611–626

Tensor Algebras and Displacement Structure II: Non-Commutative Szegö Polynomials

T. Constantinescu and J. L. Johnson

Abstract. In this paper we continue to explore the connection between tensor algebras and displacement structure. We focus on recursive orthonormalization and we develop an analogue of the Szegö-type theory of orthogonal polynomials in the unit circle for several non-commuting variables. Thus we obtain recurrence equations and Christoffel-Darboux formulas for Szegö polynomials in several non-commuting variables, as well as a Favard type result. Also, we continue to study a Szegö-type kernel for the N-dimensional unit ball of an infinite-dimensional Hilbert space.

Keywords: *Displacement structure, tensor algebras, Szegö polynomials*

AMS subject classification: 15A69, 47A57

1. Introduction

In the first part of this paper [5] we explored the connection between tensor algebras and displacement structure. The displacement structure theory was initiated in [13] as a recursive factorization theory for matrices whose implicit structure is encoded by a so-called displacement equation. This has been useful in several directions including constrained and unconstrained rational interpolation, maximum entropy, inverse scattering, H^∞-control, signal detection, digital filter design, nonlinear Riccati equations, certain Fredholm and Wiener-Hopf equations, etc. (see [14]). Aspects of the Szegö theory can be also revealed within the displacement structure theory. Our main goal is to develop an analogue for polynomials in several non-commuting variables of the Szegö theory of orthogonal polynomials on the unit circle. An analogue of the Szegö theory of orthogonal polynomials on the real line is being developed in the companion paper [6].

The paper is organized as follows. In Section 2 we review notation and several results from [5]. In this way, this paper can be read independently of [5]. In Section 3 we introduce orthogonal polynomials in several non-commuting variables associated to certain representations of the free semigroup and discuss their algebraic properties, mostly related to the recursions that they satisfy. In Section 4 we consider several positive definite kernels on the N-dimensional unit ball of an infinite-dimensional Hilbert space. In particular, we prove a basic property of the Szegö-type kernel studied in [4] by

T. Constantinescu: Univ. of Texas at Dallas, Dept. Math., Richardson, TX 75083, USA
J. L. Johnson: Wagner Coll., Dept. Math. & Comp. Sci., Staten Island, NY 10301, USA
tiberiu@utdallas.edu and jlj@utdallas.edu

ISSN 0232-2064 / $ 2.50

characterizing its Kolmogorov decomposition. In Section 5 we discuss the problem of recovering the representation from orthogonal polynomials and we prove a Favard-type result.

We plan a more detailed study of applications to multiscale systems in a sequel of this paper.

2. Preliminaries

We briefly review several constructions of the tensor algebra and introduce necessary notation. We also review the connection with displacement structure theory as established in [7].

2.1 Tensor algebras. The tensor algebra over $\mathbb{C}^N$ is defined by the algebraic direct sum

$$\mathcal{T}_N = \oplus_{k\geq 0}(\mathbb{C}^N)^{\otimes k}$$

where $(\mathbb{C}^N)^{\otimes k}$ denotes the k-fold tensor product of $\mathbb{C}^N$ with itself. The addition is taken componentwise and the multiplication is defined by juxtaposition as

$$(x \otimes y)_n = \sum_{k+l=n} x_k \otimes y_l.$$

If $\{e_1, \ldots, e_N\}$ is the standard basis of $\mathbb{C}^N$, then $\left\{e_{i_1} \otimes \cdots \otimes e_{i_k} : 1 \leq i_1, \ldots, i_k \leq N\right\}$ is a basis of $\mathcal{T}_N$. Let $\mathbb{F}_N^+$ be the unital free semigroup on N generators $1, \ldots, N$ with lexicographic order $\prec$. The empty word is the identity element, the length of the word σ is denoted by $|\sigma|$, and the length of the empty word is 0. If $\sigma = i_1 \cdots i_k$, then we write e_σ instead of $e_{i_1} \otimes \cdots \otimes e_{i_k}$, so that any element of $\mathcal{T}_N$ can be uniquely written in the form $x = \sum_{\sigma\in\mathbb{F}_N^+} c_\sigma e_\sigma$, where only finitely many of the complex numbers c_σ are different from 0.

Another construction of $\mathcal{T}_N$ can be obtained as follows. Let S be a unital semigroup and denote by $F_0(S)$ the set of functions $\phi : S \to \mathbb{C}$ with the property that $\phi(s) \neq 0$ for only finitely many values of s. This set has a natural vector space structure and $B_S = \{\delta_s : s \in S\}$ is a vector basis for $F_0(S)$, where δ_s is the Kronecker symbol associated to $s \in S$. Also, $F_0(S)$ is a unital associative algebra with respect to the product

$$\phi * \psi = \left(\sum_{s\in S} \phi(s)\delta_s\right) * \left(\sum_{t\in S} \psi(t)\delta_t\right) = \sum_{s,t\in S} \phi(s)\psi(t)\delta_{st}.$$

It is readily seen that $F_0(\mathbb{F}_N^+)$ is isomorphic to $\mathcal{T}_N$. Since each element ϕ in $F_0(\mathbb{F}_N^+)$ can be uniquely written as a (finite) sum $\phi = \sum_{\sigma\in\mathbb{F}_N^+} c_\sigma\delta_\sigma$, the isomorphism is the linear extension Φ_1 of the mapping $\delta_\sigma \to e_\sigma$ $(\sigma \in \mathbb{F}_N^+)$.

Another copy of the tensor algebra is given by the algebra $\mathcal{P}_N^0$ of polynomials in N non-commuting indeterminates $X_1, \ldots, X_N$ with complex coefficients. Each element $P \in \mathcal{P}_N^0$ can be uniquely written in the form $P = \sum_{\sigma\in\mathbb{F}_N^+} c_\sigma X_\sigma$ with $c_\sigma \neq 0$ for finitely many σ's and $X_\sigma = X_{i_1} \cdots X_{i_k}$ where $\sigma = i_1 \cdots i_k \in \mathbb{F}_N^+$. The linear extension Φ_2 of the mapping $\delta_\sigma \to X_\sigma$ $(\sigma \in \mathbb{F}_N^+)$ gives an isomorphism of $\mathcal{T}_N$ with $\mathcal{P}_N^0$.

Yet another copy of $\mathcal{T}_N$ inside the algebra of lower triangular operators allowed for the connection with displacement structure established in [7]. Thus let $\mathcal{E}$ be a Hilbert space and define $\mathcal{E}_0 = \mathcal{E}$ and, for $k \geq 1$,

$$\mathcal{E}_k = \underbrace{\mathcal{E}_{k-1} \oplus \cdots \oplus \mathcal{E}_{k-1}}_{N\, terms} = \mathcal{E}_{k-1}^{\oplus N}. \tag{2.1}$$

For $\mathcal{E} = \mathbb{C}$ we have that $\mathbb{C}_k$ can be identified with $(\mathbb{C}^N)^{\otimes k}$ and $\mathcal{T}_N$ is isomorphic to the algebra $\mathcal{L}_N^0$ of lower triangular operators $T = [T_{ij}] \in \mathcal{L}(\oplus_{k \geq 0} \mathbb{C}_k)$ with the property

$$T_{ij} = \underbrace{T_{i-1,j-1} \oplus \cdots \oplus T_{i-1,j-1}}_{N\, terms} = T_{i-1,j-1}^{\oplus N} \tag{2.2}$$

for $i, j \geq 1$ with $i \leq j$ and $T_{j0} = 0$ for all sufficiently large $j's$. The isomorphism is given by the map Φ_3 defined as follows: Let $x = (x_0, x_1, \ldots) \in \mathcal{T}_N$ ($x_p \in (\mathbb{C}^N)^{\otimes p}$ is the pth homogeneous component of x). Then $x_p = \sum_{|\sigma|=p} c_\sigma e_\sigma$ and, for $j \geq 0$, T_{j0} denotes the column matrix $[c_\sigma]_{|\sigma|=j}^T$, where "$T$" denotes the matrix transpose. Then $T_{j0} = 0$ for all sufficiently large j's and we can define $T \in \mathcal{L}(\oplus_{k \geq 0} \mathbb{C}_k)$ by using (2.2). Finally, set $\Phi_3(x) = T$.

2.2 Displacement structure. We can now describe the displacement structure of the tensor algebra. We write this connection for $\mathcal{L}_N^0$. Then it can be easily translated into any other realization of the tensor algebra. Let $F_k = [T_{ij}^k] \in \mathcal{L}(\oplus_{k \geq 0} \mathbb{C}_k)$ ($k = 1, \ldots, N$) be isometries defined by the formulae $T_{ij} = 0$ for $i \neq j+1$ and $T_{i+1,i}$ is a block-column matrix consisting of N blocks of dimension dim $\mathbb{C}_i$, all of them zero except for the kth block which is the identity on $\mathbb{C}_k$. We have the following result noticed in [7].

Theorem 2.1. *Let $T \in \mathcal{L}_N^0$ and define $A = I - TT^*$. Then*

$$A - \sum_{k=1}^{N} F_k A F_k^* = G J_{11} G^* \tag{2.3}$$

where

$$G = \begin{bmatrix} 1 & T_{00} \\ 0 & T_{01} \\ \vdots & \vdots \end{bmatrix} \qquad \text{and} \qquad J_{11} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The model $\mathcal{L}_N^0$ of the tensor algebra is also useful in order to extend this algebra to some topological tensor algebras (see [12]). Here we consider only the norm topology and denote by $\mathcal{L}_N$ the algebra of all lower triangular operators $T = [T_{ij}] \in \mathcal{L}(\oplus_{k \geq 0} \mathbb{C}_k)$ satisfying (2.2).

2.3 Multiscale processes. Multiscale processes are stochastic processes indexed by nodes on a tree. They became quite popular lately (see [1, 2]) and have potential to model the self-similarity of fractional Brownian motion leading to iterative algorithms in computer vision, remote sensing, etc. Here we restrict our attention to the case of the Cayley tree, in which each node has N branches. The vertices of the Cayley tree are indexed by $\mathbb{F}_N^+$.

Let $(X, \mathcal{F}, P)$ be a probability space and let $\{v_\sigma\}_{\sigma \in \mathbb{F}_N^+} \subset L^2(P)$ be a family of random variables. Its covariance kernel is

$$K(\sigma, \tau) = \int_X \overline{v}_\sigma v_\tau \, dP$$

and assume that the process is stationary in the sense (considered earlier, e.g. [10]) that

$$K(\tau\sigma, \tau\sigma') = K(\sigma, \sigma') \qquad (\tau, \sigma, \sigma' \in \mathbb{F}_N^+) \tag{2.4}$$
$$K(\sigma, \tau) = 0 \ \text{ if there is no } \alpha \in \mathbb{F}_N^+ \text{ such that } \sigma = \alpha\tau \text{ or } \tau = \alpha\sigma. \tag{2.5}$$

Conversely, by the invariant Kolmogorov decomposition theorem (see, e.g., [15: Chapter II]) there exists an isometric representation u of $\mathbb{F}_N^+$ on a Hilbert space $\mathcal{K}$ and a mapping $v : \mathbb{F}_N^+ \to \mathcal{K}$ such that

- $K(\sigma, \tau) = \langle v(\tau), v(\sigma) \rangle_{\mathcal{K}}$ and $u(\tau)v(\sigma) = v(\tau\sigma)$ for all $\sigma, \tau \in \mathbb{F}_N^+$
- the set $\{v(\sigma) : \sigma \in \mathbb{F}_N^+\}$ is total in $\mathcal{K}$
- $u(1), \ldots, u(N)$ are isometries with orthogonal ranges.

This class of multiscale processes would be suitable to model branching processes without "past". If a "past" should be attached to a process as above, we could try to consider processes indexed by the nodes of the tree associated to the free group on N generators $1, \ldots, N$. As mentioned in Introduction, we plan to look at this matter in a sequel of this paper. Here we focus on processes with covariance kernel satisfying (2.4) - (2.5). It was shown in [5] that such a kernel has displacement structure. Also, it is clear that for all $j, k \geq 1$

$$[K(\sigma, \tau)]_{|\sigma|=j, |\tau|=k} = \big(K(\sigma', \tau')]_{|\sigma'|=j-1, |\tau'|=k-1}\big)^{\oplus N} \tag{2.6}$$

so that the kernel is determined by the elements $s_\sigma = K(\emptyset, \sigma)$ $(\sigma \in \mathbb{F}_N^+)$.

By [3: Theorem 1.5.3] each positive definite kernel K on $\mathbb{F}_N^+$ is uniquely determined by a family of contractions $\{\gamma_{\sigma,\tau} : \sigma, \tau \in \mathbb{F}_N^+, \sigma \preceq \tau\}$ such that $\gamma_{\sigma,\sigma} = 0$ $(\sigma \in \mathbb{F}_N^+)$ and otherwise $\gamma_{\sigma,\tau} \in \mathcal{L}(\mathcal{D}_{\gamma_{\sigma+1,\tau}}, \mathcal{D}_{\gamma^*_{\sigma,\tau-1}})$ (for a contraction T between two Hilbert spaces $D_T = (I - T^*T)^{1/2}$ denotes the defect operator of T and $\mathcal{D}_T$ is the defect space of T defined as the closure of the range of D_T – note that in our case $\gamma_{\sigma,\tau}$ are just complex numbers and the condition $\gamma_{\sigma,\tau} \in \mathcal{L}(\mathcal{D}_{\gamma_{\sigma+1,\tau}}, \mathcal{D}_{\gamma^*_{\sigma,\tau-1}})$ for $\sigma \prec \tau$ encodes the fact that $|\gamma_{\sigma+1,\tau}| = 1$ or $|\gamma_{\sigma,\tau-1}| = 1$ implies $\gamma_{\sigma,\tau} = 0$; also, $\tau - 1$ denotes the predecessor of τ with respect to the lexicographic order $\prec$ on $\mathbb{F}_N^+$, while $\sigma + 1$ denotes the successor of σ). In addition, the positive definite kernel K satisfies (2.4) - (2.5) if and only if

$$\gamma_{\tau\sigma,\tau\sigma'} = \gamma_{\sigma,\sigma'} \quad (\tau, \sigma, \sigma' \in \mathbb{F}_N^+) \tag{2.7}$$
$$\gamma_{\sigma,\tau} = 0 \ \text{ if there is no } \alpha \in \mathbb{F}_N^+ \text{ such that } \sigma = \alpha\tau \text{ or } \tau = \alpha\sigma. \tag{2.8}$$

We define $\gamma_\sigma = \gamma_{\emptyset,\sigma}$ $(\sigma \in \mathbb{F}_N^+)$ and we notice that $\{\gamma_{\sigma,\tau} : \sigma, \tau \in \mathbb{F}_N^+, \sigma \preceq \tau\}$ is uniquely determined by $\{\gamma_\sigma\}_{\sigma \in \mathbb{F}_N^+}$ by the formula

$$[\gamma_{\sigma,\tau}]_{|\sigma|=j, |\tau|=k} = \big([\gamma_{\sigma',\tau'}]_{|\sigma'|=j-1, |\tau'|=k-1}\big)^{\oplus N} \qquad (j, k \geq 1). \tag{2.9}$$

3. Szegö polynomials

We introduce polynomials in several non-commuting variables orthogonal with respect to a positive definite kernel K satisfying (2.4) - (2.5). We extend some elements of the Szegö theory to this setting.

The kernel K being given, we can introduce an inner product on $F_0(\mathbb{F}_N^+)$ in the usual manner by

$$\langle \phi, \psi \rangle_K = \sum_{\sigma,\tau \in \mathbb{F}_N^+} K(\sigma,\tau)\phi(\tau)\overline{\psi(\sigma)}. \tag{3.1}$$

By factoring out the subspace $\mathcal{N}_K = \{\phi \in F_0(\mathbb{F}_N^+) : \langle \phi, \phi \rangle_K = 0\}$ and completing with respect to the norm induced by (3.1) we obtain a Hilbert space denoted $\mathcal{H}_K$. A similar structure can be introduced on $\mathcal{P}_N^0$. Let $P = \sum_{\sigma \in \mathbb{F}_N^+} c_\sigma X_\sigma$ and $Q = \sum_{\sigma \in \mathbb{F}_N^+} d_\sigma X_\sigma$ be elements in $\mathcal{P}_N^0$. Then define

$$\langle P, Q \rangle_K = \sum_{\sigma,\tau \in \mathbb{F}_N^+} K(\sigma,\tau) c_\tau \overline{d}_\sigma. \tag{3.2}$$

By factoring out the subspace $\mathcal{M}_K = \{P \in \mathcal{P}_N^0 : \langle P, P \rangle_K = 0\}$ and completing with respect to the norm induced by (3.2) we obtain a Hilbert space denoted $L^2(K)$. One can check that the map Φ_2 defined by $\delta_\sigma \to X_\sigma$ $(\sigma \in \mathbb{F}_N^+)$ extends to a unitary operator from $\mathcal{H}_K$ onto $L^2(K)$.

From now on we assume that for any $\alpha \in \mathbb{F}_N^+$ the matrix $[K(\sigma,\tau)]_{\sigma,\tau \preceq \alpha}$ is invertible. This implies that $\mathcal{M}_K = 0$ and $\mathcal{P}_N^0$ can be viewed as a subspace of $L^2(K)$. Also, for any $\alpha \in \mathbb{F}_N^+$, $\{X_\sigma\}_{\sigma \preceq \alpha}$ is a linearly independent family in $L^2(K)$. Then the Gram-Schmidt procedure gives a family $\{\varphi_\sigma\}_{\sigma \in \mathbb{F}_N^+}$ of elements in $\mathcal{P}_N^0$ such that

$$\varphi_\sigma = \sum_{\tau \preceq \sigma} a_{\sigma,\tau} X_\tau \quad (a_{\sigma,\sigma} > 0) \tag{3.3}$$

$$\langle \varphi_\sigma, \varphi_\tau \rangle_K = 0 \quad (\emptyset \preceq \sigma \prec \tau). \tag{3.4}$$

An explicit formula for the orthogonal polynomials φ_σ can be obtained in the same manner as in the classical (one variable) case. Define for $\sigma \in \mathbb{F}_N^+$

$$D_\sigma = \det\left[K(\sigma',\tau')\right]_{\sigma',\tau' \preceq \sigma} \tag{3.5}$$

and let $\{\gamma_\sigma\}_{\sigma \in \mathbb{F}_N^+}$ be the parameters associated to K as described in Subsection 2.3. Note that since all the matrices $[K(\sigma,\tau)]_{\sigma,\tau \preceq \alpha}$ $(\alpha \in \mathbb{F}_N^+)$ are assumed to be invertible, it follows that $|\gamma_\sigma| < 1$ for all $\sigma \in \mathbb{F}_N^+$.

Theorem 3.1.

(1) $\varphi_\emptyset = 1$ *and, for* $\emptyset \prec \sigma$,

$$\varphi_\sigma = \frac{1}{\sqrt{D_{\sigma-1} D_\sigma}} \det \begin{bmatrix} [K(\sigma',\tau')]_{\sigma' \prec \sigma; \tau' \preceq \sigma} \\ 1 \quad X_1 \quad \cdots \quad X_\sigma \end{bmatrix}. \tag{3.6}$$

(2) *For* $\emptyset \prec \sigma = i_1 \cdots i_k$,

$$\varphi_\sigma = \frac{1}{\prod_{1\leq j\leq k}(1-|\gamma_{i_j\ldots i_k}|^2)^{1/2}}\big(X_\sigma + \textit{ lower order terms}\big).$$

Proof. The proof is similar to the classical one. Thus we deduce from orthogonality condition (3.4) that $\langle \varphi_\sigma, X_{\tau'}\rangle_K = 0$ for $\emptyset \preceq \tau' \prec \sigma$, which implies that $\sum_{\tau\preceq\sigma} a_{\sigma,\tau}K(\tau',\tau) = 0$ for $\emptyset \preceq \tau' \prec \sigma$. Using the Cramer rules for the system

$$\left.\begin{array}{c}\displaystyle\sum_{\tau\preceq\sigma} a_{\sigma,\tau}K(\tau',\tau) = 0 \quad (\emptyset \preceq \tau' \prec \sigma)\\ \displaystyle\sum_{\tau\preceq\sigma} a_{\sigma,\tau}X_\tau = \varphi_\sigma\end{array}\right\}$$

with unknowns $a_{\sigma,\tau}$, we deduce

$$a_{\sigma,\sigma} = \frac{\varphi_\sigma D_{\sigma-1}}{\det\begin{bmatrix}[K(\sigma',\tau')]_{\sigma'\prec\sigma;\tau'\preceq\sigma}\\ 1 \quad X_1 \quad \cdots \quad X_\sigma\end{bmatrix}}\,.$$

Therefore,

$$\varphi_\sigma = \frac{a_{\sigma,\sigma}}{D_{\sigma-1}}\det\begin{bmatrix}[K(\sigma',\tau')]_{\sigma'\prec\sigma;\tau'\preceq\sigma}\\ 1 \; X_1 \; \ldots \; X_\sigma\end{bmatrix}.$$

We now compute $a_{\sigma,\sigma}$ and D_σ in terms of the parameters $\{\gamma_\sigma\}_{\sigma\in\mathbb{F}_N^+}$ of K. First we notice that

$$\left\langle \det\begin{bmatrix}[K(\sigma',\tau')]_{\sigma'\prec\sigma;\tau'\preceq\sigma}\\ 1 \quad X_1 \quad \cdots \quad X_\sigma\end{bmatrix},\ X_\sigma\right\rangle_K = D_\sigma$$

and since $X_\sigma = \frac{1}{a_{\sigma,\sigma}}\varphi_\sigma + \sum_{\tau\prec\sigma} c_\tau X_\tau$ we deduce

$$D_\sigma = \left\langle \frac{D_{\sigma-1}}{a_{\sigma,\sigma}}\,\varphi_\sigma, \frac{1}{a_{\sigma,\sigma}}\,\varphi_\sigma + \sum_{\tau\prec\sigma} c_\tau X_\tau\right\rangle_K = \frac{D_{\sigma-1}}{a^2_{\sigma,\sigma}}$$

so that $\frac{1}{a^2_{\sigma,\sigma}} = \frac{D_\sigma}{D_{\sigma-1}}$ which gives (3.6).

In order to compute D_σ in terms of $\{\gamma_\sigma\}_{\sigma\in\mathbb{F}_N^+}$ we use [3: Theorem 1.5.10] and the special structure of D_σ. Thus

$$D_\sigma = \prod_{\emptyset\prec\sigma',\tau'\preceq\sigma}(1-|\gamma_{\sigma',\tau'}|^2)$$

and for $\emptyset \prec \sigma = i_1 \ldots i_k$ we deduce

$$\frac{1}{a^2_{\sigma,\sigma}} = \frac{D_\sigma}{D_{\sigma-1}} = \prod_{1\leq j\leq k}(1-|\gamma_{i_j\ldots i_k}|^2).$$

Then

$$\begin{aligned}\varphi_\sigma &= a_{\sigma,\sigma}X_\sigma + \sum_{\tau\prec\sigma} a_{\sigma,\sigma}c_\tau X_\tau\\ &= \frac{1}{\prod_{1\leq j\leq k}(1-|\gamma_{i_j\ldots i_k}|^2)^{1/2}}\,(X_\sigma + \text{ lower order terms})\end{aligned}$$

which gives (3.7) ∎

We illustrate this result for $N = 2$. From now on it is convenient to use the notation $d_\sigma = (1 - |\gamma_\sigma|^2)^{1/2}$ $(\sigma \in \mathbb{F}_N^+ - \{\emptyset\})$.

Example. Let $N = 2$ and assume the positive kernel K satisfies the conditions in Theorem 3.1. We have $D_\emptyset = 1$ and the next three determinants are

$$D_1 = \det \begin{bmatrix} 1 & s_1 \\ \overline{s}_1 & 1 \end{bmatrix} = d_1^2$$

$$D_2 = \det \begin{bmatrix} 1 & s_1 & s_2 \\ \overline{s}_1 & 1 & 0 \\ \overline{s}_2 & 0 & 1 \end{bmatrix} = d_1^2 d_2^2$$

$$D_{11} = \det \begin{bmatrix} 1 & s_1 & s_2 & s_{11} \\ \overline{s}_1 & 1 & 0 & s_1 \\ \overline{s}_2 & 0 & 1 & 0 \\ \overline{s}_{11} & \overline{s}_1 & 0 & 1 \end{bmatrix} = d_1^4 d_2^2 d_{11}^2.$$

Using Theorem 3.1 we can easily calculate the first four orthogonal polynomials of K. Thus, $\varphi_\emptyset = 1$ and then

$$\varphi_1 = \frac{1}{d_1} \det \begin{bmatrix} 1 & s_1 \\ 1 & X_1 \end{bmatrix} = -\frac{\gamma_1}{d_1} + \frac{1}{d_1} X_1$$

$$\varphi_2 = \frac{1}{d_1^2 d_2} \det \begin{bmatrix} 1 & s_1 & s_2 \\ \overline{s}_1 & 1 & 0 \\ 1 & X_1 & X_2 \end{bmatrix} = -\frac{\gamma_2}{d_1 d_2} + \frac{\overline{\gamma}_1 \gamma_2}{d_1 d_2} X_1 + \frac{1}{d_2} X_2$$

where we used the fact that $s_2 = d_1 \gamma_2$. Then, after some calculations,

$$\begin{aligned} \varphi_{11} &= \frac{1}{d_1^3 d_2^2 d_{11}} \det \begin{bmatrix} 1 & s_1 & s_2 & s_{11} \\ \overline{s}_1 & 1 & 0 & s_1 \\ \overline{s}_2 & 0 & 1 & 0 \\ 1 & X_1 & X_2 & X_1^2 \end{bmatrix} \\ &= -\frac{\gamma_{11}}{d_1 d_2 d_{11}} + \Big(-\frac{\gamma_1}{d_1 d_{11}} + \frac{\gamma_{11} \overline{\gamma}_1}{d_1 d_2 d_{11}} \Big) X_1 + \frac{\gamma_{11} \overline{\gamma}_2}{d_2 d_{11}} X_2 + \frac{1}{d_{11} d_1} X_1^2. \end{aligned}$$

We establish now that the orthogonal polynomials introduced above satisfy equations similar to the classical Szegö difference equations.

Theorem 3.2. *The orthogonal polynomials satisfy the recurrences* $\varphi_\emptyset = \varphi_\emptyset^\sharp = 1$ *and, for* $k \in \{1, \ldots, N\}$ *and* $\sigma \in \mathbb{F}_N^+$,

$$\varphi_{k\sigma} = \frac{1}{d_{k\sigma}} \big(X_k \varphi_\sigma - \gamma_{k\sigma} \varphi_{k\sigma-1}^\sharp \big) \tag{3.8}$$

$$\varphi_{k\sigma}^\sharp = \frac{1}{d_{k\sigma}} \big(-\overline{\gamma}_{k\sigma} X_k \varphi_\sigma + \varphi_{k\sigma-1}^\sharp \big). \tag{3.9}$$

Proof. We deduce this result from similar formulae obtained for an arbitrary positive definite kernel. In this way we can show the meaning of the polynomials $\varphi_\sigma^\sharp$ $(\sigma \in \mathbb{F}_N^+)$. Let $[t_{i,j}]_{i,j \geq 1}$ be a positive definite kernel on $\mathbb{N}$ and assume that each matrix $A^{(i,j)} = [t_{k,l}]_{1 \leq i \leq k, l \leq j}$ is invertible. Also, assume $t_{k,k} = 1$ for all $k \geq 1$. Let

$F_{i,j}$ be the upper Cholesky factor of $A^{(i,j)}$, so that $F_{i,j}$ is an upper triangular matrix with positive diagonal and $A^{(i,j)} = F_{i,j}^* F_{i,j}$. A dual, lower Cholesky factor is obtained as follows: define the symmetry of appropriate dimension

$$\mathcal{J} = \begin{bmatrix} 0 & 0 & \dots & 0 & I \\ 0 & 0 & & I & 0 \\ \vdots & & \ddots & & \\ 0 & I & & & 0 \\ I & 0 & & 0 & 0 \end{bmatrix}$$

and then let $\tilde{F}_{i,j}$ denote the upper Cholesky factor of $B^{(i,j)} = \mathcal{J} A^{(i,j)} \mathcal{J}$. If $G_{i,j} = \mathcal{J} \tilde{F}_{i,j} \mathcal{J}$, then

$$A^{(i,j)} = \mathcal{J} B^{(i,j)} \mathcal{J} = \mathcal{J} F_{i,j}^* F_{i,j} \mathcal{J} = G_{i,j}^* G_{i,j}$$

and $G_{i,j}$ is a lower triangular matrix with positive diagonal, called the lower Cholesky factor of $A^{(i,j)}$. Let $P_{i,j}$ be the last column of $F_{i,j}^{-1}$ and let $P_{i,j}^\sharp$ be the first column of $G_{i,j}^{-1}$, that is $P_{i,j} = F_{i,j}^{-1} E$ and $P_{i,j}^\sharp = G_{i,j}^{-1} \mathcal{J} E$ where $E = [0 \ \cdots \ 0 \ I]^T$. Let $\{r_{i,j}\}_{1<i\leq j}$ be the parameters associated to $[t_{i,j}]_{i,j\geq 1}$ by [3: Theorem 1.5.3] and let $\rho_{i,j} = (1 - |r_{i,j}|^2)^{1/2}$. We have

$$P_{1,n} = \frac{1}{d_{1,n}} \begin{bmatrix} 0 \\ P_{2,n} \end{bmatrix} - \frac{r_{1,n}}{d_{1,n}} \begin{bmatrix} P_{1,n-1}^\sharp \\ 0 \end{bmatrix} \tag{3.10}$$

$$P_{1,n}^\sharp = -\frac{\overline{r}_{1,n}}{d_{1,n}} \begin{bmatrix} 0 \\ P_{2,n} \end{bmatrix} + \frac{1}{d_{1,n}} \begin{bmatrix} P_{1,n-1}^\sharp \\ 0 \end{bmatrix}. \tag{3.11}$$

These formulae are presumable known to the experts. For the sake of completeness we give a proof here based on results and notation from [3]. First we introduce for $i < j$ the elements

$$L_i^{(j)} = L\big(\{r_{i,k}\}_{k=i+1}^j\big) = \Big[r_{i,i+1} \ \rho_{i,i+1} r_{i,i+1} \ \cdots \ \rho_{i,i+1} \ \cdots \ \rho_{i,j-1} r_{i,j} \Big] \tag{3.12}$$

$$C_j^{(i)} = \begin{bmatrix} r_{j-1,j} \\ \vdots \\ r_{i+1,j}\rho_{i+2,j} \cdots \rho_{j-1,j} \\ r_{i,j}\rho_{i+1,j} \cdots \rho_{j-1,j} \end{bmatrix}$$

$$K_i^{(j)} = \begin{bmatrix} \overline{r}_{i,i+1}\rho_{i,i+2} \cdots \rho_{i,j} \\ \vdots \\ \overline{r}_{i,j-1}\rho_{i,j} \\ \overline{r}_{i,j} \end{bmatrix} = \begin{bmatrix} K_i^{(j-1)}\rho_{i,j} \\ \overline{r}_{i,j} \end{bmatrix}.$$

Also, we define inductively $D_i^{(i+1)} = \rho_{i,i+1}$ and

$$D_i^{(j)} = D\big(\{r_{i,k}\}_{k=i+1}^j\big) = \begin{bmatrix} D_i^{(j-1)} & -K_i^{(j-1)} r_{i,j} \\ 0 & \rho_{i,j} \end{bmatrix}. \tag{3.13}$$

We also need to review the factorization of unitary matrices. This is an extension of Euler's description of $SO(3)$. First we define

$$R_{j-i}(r_{i,k}) = I_{k-1-i} \oplus \begin{bmatrix} r_{i,k} & \rho_{i,k} \\ \rho_{i,k} & -\bar{r}_{i,k} \end{bmatrix} \oplus I_{j-k-1}$$

where I_{k-1-i} is the identity matrix of size $k-1-i$. Then

$$\begin{aligned} R_{i,j} &= R_{j-i}(r_{i,i+1}) \cdots R_{j-i}(r_{i,j}) \\ U_{i,j} &= R_{i,j}(U_{i+1,j} \oplus 1). \end{aligned}$$

It turns out that any unitary matrix can be written as a matrix of the form of $U_{i,j}$. The main idea for the proof of (3.10) is to use the identity

$$U_{i,j}\mathcal{J}G_{i,j} = F_{i,j} \tag{3.14}$$

which follows from [3: Relations (1.6.10), (6.3.8) and (6.3.9)]. Thus, we notice that (3.14) implies $P^{\sharp}_{i,j} = F_{i,j}U_{i,j}E$ which is more tractable than the original definition of $P^{\sharp}_{i,j}$. This is seen from the following calculations. Using [3: Formula (1.5.7)], the above definition of $D_1^{(n)}$ and the notation $D_1^{-(n)} = (D_1^{(n)})^{-1}$ we obtain

$$\begin{aligned}
P_{1,n} &= \begin{bmatrix} 1 & -L_1^{(n)} D_1^{-(n)} \\ 0 & F_{2,n}^{-1} D_1^{-(n)} \end{bmatrix} E \\
&= \begin{bmatrix} -L_1^{(n)} D_1^{-(n)} E \\ F_{2,n}^{-1} D_1^{-(n)} E \end{bmatrix} \\
&= \begin{bmatrix} -L_1^{(n)} \begin{bmatrix} D_1^{-(n-1)} & \frac{r_{1,n}}{\rho_{1,n}} D_1^{-(n-1)} K_1^{(n-1)} \\ 0 & \frac{1}{\rho_{1,n}} \end{bmatrix} E \\ F_{2,n}^{-1} \begin{bmatrix} D_1^{-(n-1)} & \frac{r_{1,n}}{\rho_{1,n}} D_1^{-(n-1)} K_1^{(n-1)} \\ 0 & \frac{1}{\rho_{1,n}} \end{bmatrix} E \end{bmatrix} \\
&= \begin{bmatrix} -L_1^{(n)} \begin{bmatrix} \frac{r_{1,n}}{\rho_{1,n}} D_1^{-(n-1)} K_1^{(n-1)} \\ \frac{1}{\rho_{1,n}} \end{bmatrix} \\ F_{2,n}^{-1} \begin{bmatrix} \frac{r_{1,n}}{\rho_{1,n}} D_1^{-(n-1)} K_1^{(n-1)} \\ \frac{1}{\rho_{1,n}} \end{bmatrix} \end{bmatrix} \\
&= \frac{1}{\rho_{1,n}} \begin{bmatrix} 0 \\ F_{2,n}^{-1} E \end{bmatrix} + \begin{bmatrix} -L_1^{(n)} \begin{bmatrix} \frac{r_{1,n}}{\rho_{1,n}} D_1^{-(n-1)} K_1^{(n-1)} \\ \frac{1}{\rho_{1,n}} \end{bmatrix} \\ F_{2,n}^{-1} \begin{bmatrix} \frac{r_{1,n}}{\rho_{1,n}} D_1^{-(n-1)^{-1}} K_1^{(n-1)} \\ 0 \end{bmatrix} \end{bmatrix} \\
&= \frac{1}{\rho_{1,n}} \begin{bmatrix} 0 \\ P_{2,n} \end{bmatrix} + \frac{r_{1,n}}{\rho_{1,n}} \begin{bmatrix} -L_1^{(n-1)} D_1^{-(n-1)} K_1^{(n-1)} - \rho_{1,2} \cdots \rho_{1,n-1} \\ F_{2,n}^{-1} \begin{bmatrix} D_1^{-(n-1)} K_1^{(n-1)} \\ 0 \end{bmatrix} \end{bmatrix}.
\end{aligned}$$

The proof of [3: Formula (1.6.15)] gives

$$L_1^{(n-1)} D_1^{-(n-1)} K_1^{(n-1)} + \rho_{1,2} \cdots \rho_{1,n-1} = \frac{1}{\rho_{1,2} \cdots \rho_{1,n-1}}$$

and using [3: Formula (1.5.6)] we deduce

$$F_{2,n}^{-1} \begin{bmatrix} D_1^{-(n-1)} K_1^{(n-1)} \\ 0 \end{bmatrix} = \begin{bmatrix} F_{2,n-1}^{-1} D_1^{-(n-1)} K_1^{(n-1)} \\ 0 \end{bmatrix}.$$

Therefore

$$P_{1,n} = \frac{1}{\rho_{1,n}} \begin{bmatrix} 0 \\ P_{2,n} \end{bmatrix} - \frac{r_{1,n}}{\rho_{1,n}} \begin{bmatrix} \frac{1}{\rho_{1,2}\cdots\rho_{1,n-1}} \\ -F_{2,n-1}^{-1} D_1^{-(n-1)} K_1^{(n-1)} \\ 0 \end{bmatrix}.$$

It remains to show that

$$P_{1,n-1}^{\sharp} = \begin{bmatrix} \frac{1}{\rho_{1,2}\cdots\rho_{1,n-1}} \\ -F_{2,n-1}^{-1} D_1^{-(n-1)} K_1^{(n-1)} \end{bmatrix}.$$

To that end we notice that using [3: Formula (1.5.8)], the definition of $U_{1,n-1}$, the fact that $R_{1,n-1}$ is a unitary matrix, and the notation $L_1^{*(n-1)} = (L_1^{(n-1)})^*$ we obtain

$$\begin{aligned} P_{1,n-1}^{\sharp} &= \begin{bmatrix} 0 & \frac{1}{\rho_{1,2}\cdots\rho_{1,n-1}} \\ F_{2,n-1}^{-1} & -\frac{1}{\rho_{1,2}\cdots\rho_{1,n-1}} F_{2,n-1}^{-1} L_1^{*(n-1)} \end{bmatrix} R_{1,n-1}^* R_{1,n-1} \begin{bmatrix} U_{2,n-1} & 0 \\ 0 & 1 \end{bmatrix} E \\ &= \begin{bmatrix} \frac{1}{\rho_{1,2}\cdots\rho_{1,n-1}} \\ -\frac{1}{\rho_{1,2}\cdots\rho_{1,n-1}} F_{2,n-1}^{-1} L_1^{*(n-1)}) \end{bmatrix}. \end{aligned}$$

It follows that all we have to show is the equality

$$F_{2,n-1}^{-1} D_1^{-(n-1)} K_1^{(n-1)} = \frac{1}{\rho_{1,2} \cdots \rho_{1,n-1}} F_{2,n-1}^{-1} L_1^{*(n-1)}.$$

Now this is a simple consequence of the formula $T^* D_{T^*} = D_T T^*$ for the contraction $T = L_1^{(n-1)}$. Formula (3.11) can be proved in a similar manner.

We rewrite (3.10) - (3.11) for a positive definite kernel K satisfying (2.4) - (2.5). We notice that $P_{1,n-1}^{\sharp}$ is replaced by $\varphi_{k\sigma-1}^{\sharp}$ and then we have to show that $P_{2,n}$ can be expressed in terms of φ_σ. This follows by taking into account relations (2.7) - (2.8) and using systematically (3.10). We can omit the details ∎

The previous recurrence equations look quite similar to the classical Szegö recursions. This type of recurrence equations was also found in [4] in connection with some derivations on $\mathcal{L}_N$. It turns out that these derivations are related to those considered in [11] and later studied in [9].

We also notice a graded form of the recurences (3.8) - (3.9). It is convenient to introduce, using (3.12), for $n \geq 1$ the notation

$$g_n = L(\{\gamma_\sigma\}_{|\sigma|=n}).$$

It was explained in [5] that g_n are the parameters associated to the kernel K in [16]. We also use (3.13) in order to introduce the notation

$$H_n = D(\{\gamma_\sigma\}_{|\sigma|=n}) \qquad (n \geq 1).$$

Let $\sigma(n)$ be the largest word (with respect to the lexicographic order) of lenght n, that is $\sigma(n) = \underbrace{N \cdots N}_{N\,terms}$.

Corollary 3.3. *The Szegö polynomials satisfy the recurrences*

$$[\varphi_\sigma]_{|\sigma|=k} = \left([X_1 \ \ldots \ X_N]\,[\varphi_\sigma]^{\oplus N}_{|\sigma|=k-1} - \varphi^\sharp_{\sigma(k)-1} g_k\right) H_k^{-1} \tag{3.15}$$

$$\varphi^\sharp_{\sigma(k)} = \prod_{|\tau|=k} d_\tau^{-1}\left(-[X_1 \ \ldots \ X_N]\,[\varphi_\sigma]^{\oplus N}_{|\sigma|=k-1} g_k^* + \varphi^\sharp_{\sigma(k)-1}\right). \tag{3.16}$$

for $k \geq 1$.

Proof. Both statements follow by direct calculations from Theorem 3.2 ∎

4. Christoffel-Darboux formula

A first consequence of the Szegö formula in the classical case is the Christoffel-Darboux formula. Here we find a similar formula in several non-commuting variables. To that end we introduce additional notation.

Let $\mathcal{E}$ be a Hilbert space. In this paper $\mathcal{E}$ will always be infinite-dimensional. The N-dimensional unit ball of $\mathcal{E}$ is defined by

$$\mathcal{B}_N(\mathcal{E}) = \left\{Z = (Z_1 \ \cdots \ Z_N) : (Z|Z) < I_{\mathcal{E}}\right\}$$

where for two elements $Z = (Z_1 \ \cdots \ Z_N)$ and $W = (W_1 \ \cdots \ W_N)$ in $\mathcal{L}(\mathcal{E})^N$ we define

$$(Z|W) = \sum_{k=1}^{N} Z_k W_k^*. \tag{4.1}$$

We also need a sort of Szegö kernel for $\mathcal{B}_N(\mathcal{E})$. One suggestion was given in [4] to consider the following construction. For $Z \in \mathcal{B}_N(\mathcal{E})$ define

$$E(Z) = [Z_\sigma]^{\infty}_{|\sigma|=0} \in \mathcal{L}(\oplus_{k\geq 0}\mathcal{E}_k, \mathcal{E}). \tag{4.2}$$

Also, we use the notation $\operatorname{diag}(S)$ to denote the diagonal operator in $\mathcal{L}(\oplus_{k\geq 0}\mathcal{E}_k)$ with diagonal S. A Szegö-type kernel on $\mathcal{B}_N(\mathcal{E})$ is given by the formula

$$K_S(Z, W) = E(Z)E(W)^* \qquad \big(Z, W \in \mathcal{B}_N(\mathcal{E})\big).$$

The next result explains two important properties of K.

Lemma 4.1.

(a) *For $T \in \mathcal{L}(\mathcal{E})$ and $Z, W \in \mathcal{B}_N(\mathcal{E})$, $E(Z)\mathrm{diag}\big(T - \sum_{k=1}^N Z_k T W_k^*\big)E(W)^* = T$.*

(b) *The set $\big\{E(W)^*\mathcal{E} : W \in \mathcal{B}_N(\mathcal{E})\big\}$ is total in $\oplus_{k\geq 0}\mathcal{E}_k$.*

Proof. Statement (a): Using directly the definitions,

$$\begin{aligned}
&E(Z)\mathrm{diag}\left(T - \sum_{k=1}^N Z_k T W_k^*\right)E(W)^* \\
&\quad = T + \sum_{|\sigma|\geq 1} Z_\sigma T W_\sigma^* - \sum_{k=1}^N E(Z)\mathrm{diag}(Z_k T W_k^*)E(W)^* \\
&\quad = T + \sum_{|\sigma|\geq 1} Z_\sigma T W_\sigma^* - \sum_{k=1}^N \sum_{|\sigma|\geq 0} Z_\sigma Z_k T W_k^* W_\sigma^* \\
&\quad = T + \sum_{|\sigma|\geq 1} Z_\sigma T W_\sigma^* - \sum_{|\sigma|\geq 1} Z_\sigma T W_\sigma^* \\
&\quad = T.
\end{aligned}$$

Statement (b): Let $e = \{e_\sigma\}_{\sigma\in\mathbb{F}_N^+}$ be an element of $\oplus_{k\geq 0}\mathcal{E}_k$ orthogonal to the linear span of $\{E(W)^*\mathcal{E} : W \in \mathcal{B}_N(\mathcal{E})\}$. Taking $W = 0$, we deduce $e_\emptyset = 0$. Next, we claim that for each $\sigma \in \mathbb{F}_N^+ - \{\emptyset\}$ there exist $W_l = (W_1^l, \ldots, W_N^l) \in \mathcal{B}_N(\mathcal{E})$ $(l = 1, \ldots, 2|\sigma|)$ such that $\mathrm{range}[W_\sigma^{*1} \;\cdots\; W_\sigma^{*2|\sigma|}] = \mathcal{E}$ and $W_\tau^l = 0$ for all $\tau \neq \sigma$ with $|\tau| \geq |\sigma|$. Once this claim is proved, a simple inductive argument gives $e = 0$, so $\{E(W)^*\mathcal{E} : W \in \mathcal{B}_N(\mathcal{E})\}$ is total in $\oplus_{k\geq 0}\mathcal{E}_k$. Therefore we focus on the proof of the claim.

Let $\{e_{ij}^n\}_{i,j=1}^n$ be the matrix units of the algebra M_n of $n \times n$ matrices. Each e_{ij}^n is an $n \times n$ matrix consisting of 1 in the (i,j)th entry and zeros elsewhere. For a Hilbert space $\mathcal{E}_1$ we define $E_{ij}^n = e_{ij}^n \otimes I_{\mathcal{E}_1}$ and we notice that

$$E_{ij}^n E_{kl}^n = \delta_{jk} E_{il}^n \qquad \text{and} \qquad E_{ji}^{*n} = E_{ij}^n. \tag{4.3}$$

Let $\mathcal{E}$ be infinite-dimensional and $\sigma = i_1 \cdots i_k$, so that $\mathcal{E} = \mathcal{E}_1^{\oplus 2|\sigma|}$ for some Hilbert space $\mathcal{E}_1$. For $s = 1, \ldots, N$ we define

$$\begin{aligned}
J_s &= \big\{l \in \{1, \ldots, k\} : i_{k+1-l} = s\big\} \\
W_s^{*p} &= \tfrac{1}{\sqrt{2}} \textstyle\sum_{r\in J_s} E_{r+p-1,r+p}^{2|\sigma|} \quad (p = 1, \ldots, |\sigma|).
\end{aligned}$$

We show that for each $p \in \{1, \ldots, |\sigma|\}$

$$W_\sigma^{*p} = \tfrac{1}{\sqrt{2^k}} E_{p,k+p}^{2|\sigma|} \tag{4.4}$$

$$W_\tau^p = 0 \quad \text{for } \tau \neq \sigma \text{ with } |\tau| \geq |\sigma|. \tag{4.5}$$

Using (4.3) we deduce

$$\begin{aligned}\sum_{s=1}^{N} W_s^p W_s^{*p} &= \frac{1}{2}\sum_{s=1}^{N}\sum_{r\in J_s} E^{2|\sigma|}_{r+p,r+p-1}E^{2|\sigma|}_{r+p-1,r+p} \\ &= \frac{1}{2}\sum_{s=1}^{N}\sum_{r\in J_s} E^{2|\sigma|}_{r+p,r+p} \\ &= \frac{1}{2}\sum_{r=1}^{k} E^{2|\sigma|}_{r+p,r+p} \\ &< I,\end{aligned}$$

hence $W^p \in \mathcal{B}_N(\mathcal{E})$ for each $p = 1, \ldots, |\sigma|$. For each word $\tau = j_1 \cdots j_k \in \mathbb{F}_N^+ - \{\emptyset\}$ we deduce by induction that

$$W^{*p}_{j_k} \cdots W^{*p}_{j_1} = \tfrac{1}{\sqrt{2^k}} \sum_{r\in A_\tau} E^{2|\sigma|}_{r+p-1,r+p+k-1} \tag{4.6}$$

where $A_\tau = \cap_{p=0}^{k-1}(J_{j_{k-p}} - p) \subset \{1, \ldots, N\}$ and $J_{j_{k-p}} - p = \{l - p : l \in J_{i_{k-p}}\}$.

We show that $A_\sigma = \{1\}$ and $A_\tau = \emptyset$ for $\tau \neq \sigma$. Let $q \in A_\tau$. Therefore, for any $p \in \{0, \ldots, k-1\}$ we must have $q + p \in J_{j_{k-p}}$ or $i_{k+1-q-p} = j_{k-p}$. For $p = k-1$ we deduce $j_1 = i_{2-q}$ and since $2 - q \geq 1$, it follows that $q \leq 1$. Also, $q \geq 1$, therefore the only element that can be in A_τ is $q = 1$, in which case we must have $\tau = \sigma$. Since $l \in J_{i_{k+1-l}}$ for each $l = 1, \ldots, k-1$, we deduce that $A_\sigma = \{1\}$ and $A_\tau = \emptyset$ for $\tau \neq \sigma$. Formula (4.6) implies (4.4). In a similar manner we can construct a family W^p $(p = |\sigma|+1, \ldots, 2|\sigma|)$ such that $W^{*p}_\sigma = \frac{1}{\sqrt{2^k}} E^{2|\sigma|}_{p+k,p}$ and $W^p_\tau = 0$ for $\tau \neq \sigma$ with $|\tau| \geq |\sigma|$. Thus for $s = 1, \ldots, N$ we define

$$\begin{aligned}&K_s = \big\{l \in \{1, \ldots, k\} : i_k = s\big\} \\ &W^{*p}_s = \tfrac{1}{\sqrt{2}}\textstyle\sum_{r\in K_s} E^{2|\sigma|}_{r+p-k,r+p-k-1} \quad (p = |\sigma|+1, \ldots, 2|\sigma|).\end{aligned}$$

Now

$$\big[W^{*1}_\sigma \ \cdots \ W^{*2|\sigma|}_\sigma\big] = \tfrac{1}{\sqrt{2^k}}\big[E^{2|\sigma|}_{1,k+1} \ \cdots \ E^{2|\sigma|}_{k,2k} \ E^{2|\sigma|}_{k+1,1} \ \cdots \ E^{2|\sigma|}_{2k,k}\big]$$

whose range is $\mathcal{E}$. This concludes the proof ∎

We note that the result given by Lemma 4.1/(b) is not true in the case $\mathcal{E}$ finite-dimensional. The meaning of the result is that in the case $\mathcal{E}$ infinite-dimensional E is precisely the Kolmogorov decomposition of the kernel K_S.

We now let a polynomial $P = \sum_{\sigma\in\mathbb{F}_N^+} c_\sigma X_\sigma \in \mathcal{P}_N^0$ take values on $\mathcal{B}_N(\mathcal{E})$ by the formula

$$P(Z) = \sum_{\sigma\in\mathbb{F}_N^+} c_\sigma Z_\sigma \qquad \big(Z \in \mathcal{B}_N(\mathcal{E})\big). \tag{4.7}$$

Define the Cristoffel-Darboux kernel by the formula

$$\begin{aligned}&K_{CD}(Z,W) = \\ &E(Z)\mathrm{diag}\Big(\varphi^\sharp_{\sigma(n)}(Z)\varphi^\sharp_{\sigma(n)}(W)^* - \sum_{|\tau|=n}\varphi_\tau(Z)\varphi_\tau(W)^*\Big)E(W)^*\end{aligned} \tag{4.8}$$

for $Z, W \in \mathcal{B}_N(\mathcal{E})$.

Theorem 4.2. *For any $Z, W \in \mathcal{B}_N(\mathcal{E})$,*

$$K_{CD}(Z,W) = \sum_{0\leq|\tau|<n} \varphi_\tau(Z)\varphi_\tau(W)^*.$$

Proof. From (3.8) - (3.9) we deduce

$$\begin{aligned}&\varphi^\sharp_{k\sigma}(Z)\varphi^\sharp_{k\sigma}(W)^* - \varphi_{k\sigma}(Z)\varphi_{k\sigma}(W)^* \\ &\qquad = \varphi^\sharp_{k\sigma-1}(Z)\varphi^\sharp_{k\sigma-1}(W)^* - Z_k\varphi_{k\sigma}(Z)\varphi_{k\sigma}(W)^*W_k^*\end{aligned}$$

for any $k \in \{1,\dots,N\}$, $\sigma \in \mathbb{F}_N^+$ and $Z, W \in \mathcal{B}_N(\mathcal{E})$. Adding all these relations for $k \in \{1,\dots,N\}$ and $0 \leq |\sigma| \leq n-1$, we deduce

$$\varphi^\sharp_{\sigma(n)}(Z)\varphi^\sharp_{\sigma(n)}(W)^* - \sum_{0\leq|\sigma|\leq n}\varphi_\sigma(Z)\varphi_\sigma(W)^* = \sum_{k=1}^N \sum_{0\leq|\sigma|\leq n-1} Z_k\varphi_\sigma(Z)\varphi_\sigma(W)^*W_k^*.$$

This relation and Lemma 4.1 give

$$\begin{aligned}K_{CD}(Z,W) &= E(Z)\,\mathrm{diag}\Big(\sum_{0\leq|\sigma|<n}\varphi_\sigma(Z)\varphi_\sigma(W)^* \\ &\qquad -\sum_{k=1}^N Z_k\Big(\sum_{0\leq|\sigma|<n}\varphi_\sigma(Z)\varphi_\sigma(W)^*\Big)W_k^*\Big)E(W)^* \\ &= \sum_{0\leq|\tau|<n}\varphi_\tau(Z)\varphi_\tau(W)^*\end{aligned}$$

and the statement is proven ∎

We can show one more application of Lemma 4.1. For a formal power series

$$f = \textstyle\sum_{\sigma\in\mathbb{F}_N^+} c_\sigma X_\sigma$$

in N non-commuting variables $X_1,\dots,X_N$ we denote by T_f the lower triangular infinite matrix associated to f as described in Subsection 2.1. We denote by $\mathcal{S}_N$ the Schur class of those formal power series f with the property that T_f is a contraction in $\mathcal{L}(\oplus_{k\geq 0}\mathbb{C}_k)$. If $\mathcal{E}$ is an infinite-dimensional Hilbert space, then we can define $f(Z)$ for $Z \in \mathcal{B}_N(\mathcal{E})$ as in [4] by the formula

$$f(Z) = E(Z)(T_f \otimes I_{\mathcal{E}})/\mathcal{E}. \tag{4.9}$$

We notice that this definition is consistent with (4.7). We extend a familiar characterization of the Schur class to the setting of this paper.

Theorem 4.3. *The formal power series f belongs to $\mathcal{S}_N$ if and only if*

$$C_f(Z,W) = E(Z)\mathrm{diag}\big(I - f(Z)f(W)^*\big)E(W)^* \qquad \big(Z, W \in \mathcal{B}_N(\mathcal{E})\big)$$

is a positive definite kernel on $\mathcal{B}_N(\mathcal{E})$.

Proof. Using [4: Lemma 3.1] we deduce that for $Z, W \in \mathcal{B}_N(\mathcal{E})$

$$\begin{aligned} &E(Z)\big(I-(T_f\otimes I_{\mathcal{E}})(T_f\otimes I_{\mathcal{E}})^*\big)E(W)^* \\ &\quad = E(Z)E(W)^* - E(Z)(T_f\otimes I_{\mathcal{E}})(T_f\otimes I_{\mathcal{E}})^*E(W)^* \\ &\quad = E(Z)E(W)^* - E(Z)\mathrm{diag}(f(Z))\mathrm{diag}(f(W)^*)E(W)^* \\ &\quad = E(Z)\mathrm{diag}\big(I-f(Z)f(W)^*\big)E(W)^* \\ &\quad = C_f(Z,W). \end{aligned}$$

This relation implies that if $f \in \mathcal{S}_N$, then C_f is a positive definite kernel on $\mathcal{B}_N(\mathcal{E})$. For the converse implication we have to use in addition Lemma 4.1 ∎

5. Inverse problems

In this brief section we prove a Favard type result for orthogonal polynomials in several non-commuting variables.

Theorem 5.1. *Let $\{\gamma_\sigma\}_{\sigma\in\mathbb{F}_N^+}$ be a family of complex numbers with $\gamma_\emptyset = 0$ and $|\gamma_\sigma| < 1$ for $\sigma \in \mathbb{F}_N^+ - \{\emptyset\}$. Then there exists a unique positive definite kernel K satisfying* (2.4) – (2.5) *such that the polynomials φ_σ $(\sigma \in \mathbb{F}_N^+)$ defined by the recursions $\varphi_\emptyset = \varphi_\emptyset^\sharp = 1$ and for $k \in \{1,\dots,N\}$ and $\sigma \in \mathbb{F}_N^+$ by*

$$\begin{aligned} \varphi_{k\sigma} &= \frac{1}{d_{k\sigma}}\big(X_k\varphi_\sigma - \gamma_{k\sigma}\varphi^s harp_{k\sigma-1}\big) \\ \varphi_{k\sigma}^\sharp &= \frac{1}{d_{k\sigma}}\big(-\overline{\gamma}_{k\sigma}X_k\varphi_\sigma + \varphi_{k\sigma-1}^\sharp\big) \end{aligned}$$

are orthogonal with respect to K.

Proof. Once again we rely on some results that are known for arbitrary positive definite kernels on the set of integers. In this way, the proof is quite straightforward. Let $\{\gamma_{\sigma,\tau} : \sigma,\tau \in \mathbb{F}_N^+ \text{ with } \sigma \preceq \tau\}$ be the family of complex numbers associated to $\{\gamma_\sigma\}_{\sigma\in\mathbb{F}_N^+}$ by (2.9). Let K be the positive definite kernel associated to $\{\gamma_{\sigma,\tau} : \sigma,\tau \in \mathbb{F}_N^+ \text{ with } \sigma \preceq \tau\}$ by [3: Theorem 1.5.3]. By Theorem 3.2, the polynomials φ_σ $(\sigma \in \mathbb{F}_N^+)$ defined by the recurrences $\varphi_\emptyset = \varphi_\emptyset^\sharp = 1$ and for $k \in \{1,\dots,N\}$ and $\sigma \in \mathbb{F}_N^+$ by

$$\begin{aligned} \varphi_{k\sigma} &= \frac{1}{d_{k\sigma}}\big(X_k\varphi_\sigma - \gamma_{k\sigma}\varphi_{k\sigma-1}^\sharp\big) \\ \varphi_{k\sigma}^\sharp &= \frac{1}{d_{k\sigma}}\big(-\overline{\gamma}_{k\sigma}X_k\varphi_\sigma + \varphi_{k\sigma-1}^\sharp\big) \end{aligned}$$

must be the orthogonal polynomials of K ∎

References

[1] Basseville, M., Benveniste, A., Chou, K. C., Golden, S. A., Nikoukhah, R. and A. S. Wilsky: *Modeling and estimation of multiresolution stochastic processes.* IEEE Trans. Info. Theory. 38 (1992), 766 – 784.

[2] Chou, K. C., Wilsky, A. S. and A. Benveniste: *Multiscale recursive estimation, data fusion and regularization.* IEEE Trans. Aut. Control 39 (1994), 464 – 478.

[3] Constantinescu, T.: *Schur Parameters, Factorization and Dilation Problems.* Basel: Birkhäuser Verlag 1996.

[4] Constantinescu, T. and J. L. Johnson: *A note on noncommutative interpolation.* Canad. Math. Bull. (to appear).

[5] Constantinescu, T. and J. L. Johnson: *Tensor algebras and displacement structure.* Part I: *The Schur algorithm.* Z. Anal. Anw. 21 (2002), 3 – 20.

[6] Constantinescu, T. and J. L. Johnson: *Orthogonal partial isometries and their isometric extensions* (in preparation).

[7] Constantinescu, T., Sayed, A. H. and T. Kailath: *Inverse scattering experiments, structured matrix inequalities, and tensor algebra.* Lin. Alg. Appl. 343-344 (2002), 147 – 169.

[8] Dubovoj, V. K., Fritzsche, B. and B. Kirstein: *Matricial Version of the Classical Schur Problem.* Stuttgart: Teubner 1992.

[9] Fox, R. H.: *Free differential calculus*, Parts I and II. Ann. Math. 57 (1953), 547 – 560 and 58 (1954), 196 – 210.

[10] Frazho, A. E.: *On stochastic bilinear systems.* In: Modeling and Applications of Stochastic Processes (ed.: U. B. Desai). Boston: Kluwer Acad. 1988, pp. 215 – 241.

[11] Hausdorff, F.: *Die symbolische Exponentialformel in der Gruppentheorie.* Ber. Sächs. Akad. Wiss. (Math. Phys. Klasse) Leipzig 58 (1906), 19 – 48.

[12] Hofmann, G.: *Topologien auf Tensoralgebren.* Wiss. Z. Univ. Leipzig, Math-Naturw. Reihe 33 (1984), 16 – 24.

[13] Kailath, T., Kung, S. Y. and M. Morf: *Displacement rank of a matrix.* Bull. Amer. Math. Soc. 1 (1979), 769 – 773.

[14] Kailath, T. and A. H. Sayed: *Displacement structure: theory and applications.* SIAM Rev. 37 (1995), 297 – 386.

[15] Parthasarathy, K. R.: *An Introduction to Quantum Stochastic Calculus.* Basel: Birkhäuser Verlag 1992.

[16] Popescu, G.: *Structure and entropy for Toeplitz kernels.* C.R. Acad. Sci. Paris, Sér. 1 Math. 329 (1999), 129 – 134.

[17] Szegö, G.: *Orthogonal Polynomials* (Colloquium Publications: Vol. 23). Providence, Rhode Island (USA): Amer. Math. Soc. 1939.

Received 02.01.2002