Centrum voor Wiskunde en Informatica

REPORTRAPPORT

On orthonormal polynomials and kernel polynomials associated with a matrix spectral distribution function

K.O. Dzhaparidze, R.H.P. Janssen

Department of Operations Research, Statistics, and System Theory

Report BS-R9418 May 1994

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.
SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

On Orthonormal Polynomials and Kernel Polynomials Associated with a Matrix Spectral Distribution Function

Kacha Dzhaparidze and René H.P. Janssen

CWI

P.O. Box 94079, 1009 GB Amsterdam, The Netherlands

and

Delft University of Technology

Mekelweg 4, 2628 CD Delft, The Netherlands

Abstract

In this report we deal with problems of constructing orthonormal polynomials and kernel polynomials associated with a matrix spectral distribution function F, defined on $(-\pi,\pi]$. In section 2 the recurrence relations between the orthonormal polynomials are presented, as well as the Christoffel-Darboux formula for the kernel polynomials. In section 3 new results are derived concerning the spectral distribution function F dominated by another spectral distribution function F_0 in the sense that

$$dF(\lambda) = r(e^{i\lambda}\mathrm{I})\, dF_0(\lambda)\, r(e^{i\lambda}\mathrm{I})^*$$

where r is some matrix polynomial. It is shown how to construct the orthonormal polynomials and kernel polynomials associated with F, given the polynomial r and the orthonormal polynomials and kernel polynomials associated with F_0. In section 4 the results are applied to multiple time series analysis.

AMS Subject Classification (1991): 62M15.
Keywords and Phrases: matrix spectral distribution function, orthonormal polynomials, kernel polynomials, Christoffel-Darboux formula, time series analysis.

Report BS-R9418
ISSN 0924-0659
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

1 Introduction

In [5], chapter 2, recurrence relations are presented between polynomials in a single variable $z \in \mathbb{C}$, which are orthonormal on the unit circle $z = e^{i\lambda}$ $(-\pi < \lambda \leq \pi)$ with respect to a scalar spectral distribution function F. These recurrence relations are derived from the well-known Christoffel-Darboux formula for the corresponding kernel polynomials. For the extension to the case of a matrix spectral distribution function F, with associated matrix polynomials in a variable $z \in \mathbb{C}$, see e.g. [4], where further references can be found. Some results from [4] are presented in section 2, for instance recurrence relations and the Christoffel-Darboux formula for matrix polynomials in the matrix variable Z (not necessarily $\mathrm{Z} = z\mathrm{I}$ as in [4]), which are applied in sections 3 and 4.

Unlike section 2 where the matrix spectral distribution function F is arbitrary, we treat in section 3 matrix spectral distribution functions F of a special form: we assume that F is dominated by another matrix spectral distribution function F_0 so that (33) holds with some matrix polynomial r, called the transfer polynomial. This choice of the form (33) for the spectral distribution function is motivated by applications in time series analysis. We show how to construct the orthonormal polynomials and kernel polynomials associated with F, given the transfer polynomial r and the orthonormal polynomials and kernel polynomials associated with F_0. This construction is based on formula (40), which can be viewed as an extension of the Christoffel-Darboux formula. The results of section 3 are partly known: see [3] for similar results concerning a scalar valued spectral distribution function.

In section 4 special types of transfer polynomials r are treated. In case the matrix polynomial r possesses zeros (unlike the scalar case, matrix polynomials need not have zeros), the relations of section 3 take a particular useful form. See subsections 4.4 - 4.6 for other special cases. In the concluding subsection 4.7 we discuss briefly applications to time series analysis in the spirit of [2].

2 Orthonormal and Kernel Polynomials

2.1 Polynomials

In this subsection we will introduce polynomials in the matrix variable $\mathrm{Z} \in \mathcal{M}_d(\mathbb{C})$, where for fixed $d \in \mathbb{N}^+$, $\mathcal{M}_d(\mathbb{C})$ is the set of all $d \times d$ complex valued matrices. On the complex vector space of these polynomials, an inner product $< \cdot ; \cdot >$ will be defined with respect to a matrix spectral distribution function $F : (-\pi, \pi] \to \mathcal{M}_d(\mathbb{C})$, i.e. a function F with the following three properties (cf. [1] or [6]):

1. the increments, $F(\lambda_2) - F(\lambda_1)$, $\lambda_2 \geq \lambda_1$, are Hermitian non-negative definite.

2. $\lambda \mapsto F(\lambda)$ is continuous from the right.

3. $\lim_{\lambda \downarrow -\pi} F(\lambda)$ exists.

Fix a $d \in \mathbb{N}^+$. For $n \in \mathbb{N}$, the monomial $e_n : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ is then the mapping such that

$$\mathrm{Z} \longmapsto \mathrm{Z}^n,$$

with the convention $\mathrm{Z}^0 = \mathrm{I}$. Note that the product of the monomials e_n and e_m is the monomial e_{n+m}.

Using throughout the notation $\mathcal{M}_{k \times l}$ for the set of all $(k+1) \times (l+1)$ block matrices with blocks from $\mathcal{M}_d(\mathbb{C})$ (for some $k, l \in \mathbb{N}$), we also define the mapping $\mathbf{e}_n : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_{n \times 0}$ with components $e_j, j = 0, \ldots, n$.

Once monomials are introduced, polynomials can be defined: a polynomial p is a mapping $\mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ defined by

$$p = \sum_{k=0}^{n} \mathrm{P}_k e_k = \mathsf{P}_n \mathbf{e}_n \tag{1}$$

for some $n \in \mathbb{N}$, where $\mathsf{P}_n = (\mathrm{P}_0, \ldots, \mathrm{P}_n) \in \mathcal{M}_{0 \times n}$. It is said that the polynomial p of form (1) has degree n if $\mathrm{P}_n \neq 0$. The matrix P_n is then called the leading coefficient of p.

Let $\mathcal{P}$ denote the set of polynomials. Note that $p \cdot e$ belongs to $\mathcal{P}$ for all polynomials p and all monomials e, whereas the product of two arbitrary polynomials is in general not a polynomial anymore (for more details see subsection 3.1).

Obviously $(\mathcal{P}, \mathcal{M}_d(\mathbb{C}))$ is a vector space, provided scalar multiplication is understood as matrix multiplication on the left. On this vector space, an inner product is defined. A mapping $< \cdot ; \cdot >: \mathcal{P} \times \mathcal{P} \to \mathcal{M}_d(\mathbb{C})$ is an inner product if it possesses the following four properties:

1. $< \mathrm{A}p + \mathrm{B}q ; r > = \mathrm{A} < p ; r > + \mathrm{B} < q ; r >$ for all $p, q, r \in \mathcal{P}$ and $\mathrm{A}, \mathrm{B} \in \mathcal{M}_d(\mathbb{C})$.

2. $< p ; q > = < q ; p >^*$ for all $p, q \in \mathcal{P}$.

3. $< p ; p >$ is Hermitian non-negative definite for all $p \in \mathcal{P}$.

4. $< p; p >= 0 \Rightarrow p = 0$ for all $p \in \mathcal{P}$.

In this paper we *assume* that the mapping $< \cdot ; \cdot >$ given by

$$< p; q >:= \int_{(-\pi,\pi]} p(e^{i\lambda}\mathrm{I})\, dF(\lambda)\, q(e^{i\lambda}\mathrm{I})^* \tag{2}$$

is an inner product, where $F : (-\pi, \pi] \to \mathcal{M}_d(\mathbb{C})$ is a matrix spectral distribution function. The mapping defined by (2) indeed satisfies properties 1-3 of an inner product for an arbitrary F. So, by assuming that (2) defines an inner product, we restricted the class of spectral functions F, in order to satisfy also property 4. Note that (2) is an inner product if and only if the spectral function F yields positive definite Hermitian block Toeplitz matrices

$$\mathrm{H}_n := \int_{(-\pi,\pi]} \mathbf{e}_n(e^{i\lambda}\mathrm{I})\, dF(\lambda)\, \mathbf{e}_n(e^{i\lambda}\mathrm{I})^* \tag{3}$$

for all $n \in \mathbb{N}$. This is easily verified by rewriting the inner product (2) of the polynomials $p = \mathsf{P}_n\mathbf{e}_n$ and $q = \mathsf{Q}_n\mathbf{e}_n$ in the form

$$< p; q >= \mathsf{P}_n\mathrm{H}_n\mathsf{Q}_n^*, \tag{4}$$

for any $\mathsf{P}_n, \mathsf{Q}_n \in \mathcal{M}_{0\times n}$.

By assumption the matrices H_n, associated with a given spectral function F by (3), are invertible for all $n \in \mathbb{N}$ and so all expressions of the form $\mathbf{X}^*\mathrm{H}_n^{-1}\mathbf{Y}$ are well defined for any $\mathbf{X} \in \mathcal{M}_{n\times k}$ and $\mathbf{Y} \in \mathcal{M}_{n\times l}$ with $k, l \in \mathbb{N}$. Throughout this paper a number of expressions of this type are treated. It seems therefore useful to make use of the following abridged notation: for fixed $n \in \mathbb{N}$

$$\mathbf{X}^*\mathrm{H}_n^{-1}\mathbf{Y} = (\mathbf{X}; \mathbf{Y}) \tag{5}$$

where $\mathbf{X} \in \mathcal{M}_{n\times k}$ and $\mathbf{Y} \in \mathcal{M}_{n\times l}$ with $k, l \in \mathbb{N}$. The special vector $\mathbf{X} \in \mathcal{M}_{n\times 0}$ with components $\mathrm{X}_j = 0$ for $j \neq m$ and $\mathrm{X}_m = \mathrm{I}$ for fixed $m \in \{0, \ldots, n\}$, called the m^{th} unit vector, is denoted by $\mathbf{U}_n^m$.

Since for every $p \in \mathcal{P}$ the matrix $< p, p >$ is Hermitian non-negative definite, we have the unique decomposition $< p, p >= \mathrm{LL}^*$ with $\mathrm{L} \in \mathcal{L}_d(\mathbb{C})$, where $\mathcal{L}_d(\mathbb{C})$ is the set of all $d \times d$ complex lower triangular matrices with non-negative real diagonal elements. This matrix L will be denoted by $< p, p >^{\frac{1}{2}}$. Using this notation, the norm $\|\cdot\| : \mathcal{P} \to \mathcal{L}_d(\mathbb{C})$ is then the mapping

$$p \longmapsto < p, p >^{\frac{1}{2}} .$$

2.2 Orthonormal Polynomials and their Reciprocals

Let $\mathcal{L}_d^+(\mathbb{C})$ be the set of all $d \times d$ complex lower triangular matrices with positive real diagonal elements.

Theorem 1 *There exists a unique system of polynomials* $\{\phi_n\}_{n\in\mathbb{N}}$ *such that*

1. ϕ_n *is a polynomial of degree n.*

2. the leading coefficient of ϕ_n*, denoted by* $\Phi_{n,n}$*, belongs to* $\mathcal{L}_d^+(\mathbb{C})$.

3. the polynomials $\{\phi_n\}_{n\in\mathbb{N}}$ *are orthonormal, i.e.*

$$< \phi_n; \phi_m >= \delta_{n-m}\mathrm{I} = \begin{cases} \mathrm{I} & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

Proof: First such a system will be constructed explicitly by using the Gram-Schmidt orthogonalization procedure: let $\psi_0 = e_0$ and $\phi_0 = ||\psi_0||^{-1}\psi_0$. Given $\phi_0, \ldots, \phi_n$, let ψ_{n+1} be defined by

$$\psi_{n+1} := e_{n+1} - \sum_{k=0}^{n} < e_{n+1}; \phi_k > \phi_k.$$

The matrix $< \psi_{n+1}; \psi_{n+1} >$ is Hermitian positive definite, so $||\psi_{n+1}||^{-1}$ exists. Define now $\phi_{n+1} := ||\psi_{n+1}||^{-1}\psi_{n+1}$. Then the polynomials $\{\phi_n\}_{n\in\mathbb{N}}$ satisfy the three conditions of this theorem.

It remains to prove uniqueness. Suppose that there is another system $\{\hat{\phi}_n\}_{n\in\mathbb{N}}$ that satisfies the three conditions of this theorem. Let $\mathrm{C} := \hat{\Phi}_{n,n}\Phi_{n,n}^{-1}$. Then $\mathrm{C}\phi_n - \hat{\phi}_n$ is a polynomial of degree less than n, hence $< \mathrm{C}\phi_n - \hat{\phi}_n; \mathrm{C}\phi_n - \hat{\phi}_n >= 0$, which yields $\mathrm{C}\phi_n - \hat{\phi}_n = 0$. So $\Phi_{n,n}^{-1}\phi_n = \hat{\Phi}_{n,n}^{-1}\hat{\phi}_n$ and it remains to verify the identity $\Phi_{n,n} = \hat{\Phi}_{n,n}$. By taking in the last expression the inner product first with ϕ_n and then with $\hat{\phi}_n$, we get

$$< \phi_n; \hat{\phi}_n >= \Phi_{n,n}\hat{\Phi}_{n,n}^{-1} = (\hat{\Phi}_{n,n}\Phi_{n,n}^{-1})^*.$$

Thus $\Phi_{n,n}^*\Phi_{n,n} = \hat{\Phi}_{n,n}^*\hat{\Phi}_{n,n}$, i.e. $\Phi_{n,n} = \hat{\Phi}_{n,n}$ since both $\Phi_{n,n}$ and $\hat{\Phi}_{n,n}$ belong to $\mathcal{L}_d^+(\mathbb{C})$. Hence $\phi_n = \hat{\phi}_n$. □

For $n \in \mathbb{N}$ let (cf. (5))

$$\xi_n = (\mathbf{U}_n^n; \mathbf{e}_n), \tag{6}$$

where $\mathbf{U}_n^n \in \mathcal{M}_{n\times 0}$ is the n^{th} unit vector. Then $\phi_n = ||\xi_n||^{-1}\xi_n$, as is easily verified by calculating $< e_k; \xi_n >$ for $0 \leq k \leq n$.

The polynomial *reciprocal* to ϕ_n, denoted by ϕ_n^ρ, is uniquely defined by the following four properties

1. ϕ_n^ρ belongs to the span of $e_0, \ldots, e_n$.

2. the norm of ϕ_n^ρ equals I.

3. ϕ_n^ρ is orthogonal to $e_1, \ldots, e_n$.

4. the coefficient of e_0 in ϕ_n^ρ belongs to $\mathcal{L}_d^+(\mathbb{C})$.

So if

$$\xi_n^\rho = (\mathrm{U}_n^0; \mathrm{e}_n),$$

where $\mathrm{U}_n^0 \in \mathcal{M}_{n\times 0}$ is the 0^{th} unit vector, then $\phi_n^\rho = \|\xi_n^\rho\|^{-1}\xi_n^\rho$, as is easily verified by calculating $< e_k; \xi_n^\rho >$ for $0 \le k \le n$.

Note that for every spectral distribution function F, the mapping $\overline{F} : (-\pi, \pi] \to \mathcal{M}_d(\mathbb{C})$ is also a spectral distribution function. The orthonormal polynomials associated with F and $\overline{F}$ and their reciprocals are related as follows:

Theorem 2 *Let $\{\phi_n^{\overline{F}}\}_{n\in\mathbb{N}}$ be the orthonormal polynomials associated with $\overline{F}$ and let $\{\phi_n^{F\rho}\}_{n\in\mathbb{N}}$ be the polynomials reciprocal to the orthonormal polynomials associated with F.*

If $\phi_n^{\overline{F}} = \sum_{k=0}^{n} \Phi_{n,k} e_k$, then $\phi_n^{F\rho} = \sum_{k=0}^{n} \overline{\Phi}_{n,n-k} e_k$.

The term 'reciprocal' is justified by the following considerations (cf. [5]): if the matrix spectral distribution function F has real increments, i.e. $dF = d\overline{F}$ (which is always true for $d = 1$), then $\phi_n^\rho(\mathrm{Z}) = \overline{\phi}_n(\mathrm{Z}^{-1})\mathrm{Z}^n$. Note that the condition $dF = d\overline{F}$ is equivalent to $dF = dF^{\mathrm{T}}$, since $dF = dF^*$ by definition.

<u>Proof:</u> Let $< \cdot;\cdot >_F$ and $< \cdot;\cdot >_{\overline{F}}$ denote the inner product with respect to F and $\overline{F}$ respectively. Then for $\phi_n^{F\rho} = \sum_{k=0}^{n} \overline{\Phi}_{n,n-k} e_k$ we have

$$< e_m; \phi_n^{F\rho} >_F = [< \phi_n^{\overline{F}}, e_{n-m} >_{\overline{F}}]^{\mathrm{T}} \tag{7}$$

which is verified as follows. By (3) we get $[\mathrm{H}_n^{\overline{F}}]^{\mathrm{T}} = \overline{\mathrm{H}_n^{\overline{F}}} = \mathrm{J}_n \mathrm{H}_n^F \mathrm{J}_n$ where $\mathrm{J}_n \in \mathcal{M}_{n\times n}$ is such that for $0 \le k, l \le n$

$$\mathrm{U}_n^{k*} \mathrm{J}_n \mathrm{U}_n^l = \delta_{k+l-n} \mathrm{I}.$$

This identity and (4) yield (7):

$$\begin{aligned}
< e_m; \phi_n^{F\rho} >_F &= \mathrm{U}_n^{m*} \mathrm{H}_n^F (\overline{\Phi}_{n,n}, \ldots, \overline{\Phi}_{n,0})^* = \mathrm{U}_n^{m*} \mathrm{J}_n \mathrm{J}_n \mathrm{H}_n^F \mathrm{J}_n \mathrm{J}_n (\Phi_{n,n}, \ldots, \Phi_{n,0})^{\mathrm{T}} \\
&= \mathrm{U}_n^{n-m*} [\mathrm{H}_n^{\overline{F}}]^{\mathrm{T}} (\Phi_{n,0}, \ldots, \Phi_{n,n})^{\mathrm{T}} = [(\Phi_{n,0}, \ldots, \Phi_{n,n}) \mathrm{H}_n^{\overline{F}} \mathrm{U}_n^{n-m}]^{\mathrm{T}} \\
&= [< \phi_n^{\overline{F}}, e_{n-m} >_{\overline{F}}]^{\mathrm{T}}.
\end{aligned}$$

We deduce now the desired statement from (7). First note that the coefficient of e_0 in $\phi_n^{F\rho}$ equals $\overline{\Phi}_{n,n}$ and therefore belongs to $\mathcal{L}_d^+(\mathbb{C})$. By the definition of reciprocals it suffices to show that for $\phi_n^{F\rho} = \sum_{k=0}^{n} \overline{\Phi}_{n,n-k} e_k$ we have $< e_m; \phi_n^{F\rho} >_F = 0$ with $1 \le m \le n$ and $< \phi_n^{F\rho}; \phi_n^{F\rho} >_F = \mathrm{I}$. Both identities follow from (7). The former identity is obvious and the latter is obtained as follows:

$$\begin{aligned}
< \phi_n^{F\rho}; \phi_n^{F\rho} >_F &= \overline{\Phi}_{n,n} < e_0; \phi_n^{F\rho} >_F = \overline{\Phi}_{n,n} [< \phi_n^{\overline{F}}; e_n >_{\overline{F}}]^{\mathrm{T}} \\
&= [< \phi_n^{\overline{F}}; \Phi_{n,n} e_n >_{\overline{F}}]^{\mathrm{T}} = [< \phi_n^{\overline{F}}; \phi_n^{\overline{F}} >_{\overline{F}}]^{\mathrm{T}} = \mathrm{I}.
\end{aligned}$$

□

We conclude this subsection with two simple examples of orthonormal polynomials and their reciprocals associated with certain matrix spectral distribution functions F.

Example 1: Let the increments of $F : (-\pi, \pi] \to \mathcal{M}_d(\mathbb{C})$ be given by

$$dF(\lambda) = \frac{1}{2\pi}\Sigma\, d\lambda, \tag{8}$$

where the matrix Σ is Hermitian positive definite. So there exists a matrix $\Sigma^{\frac{1}{2}} \in \mathcal{L}_d^{+}(\mathbb{C})$ such that $\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}*} = \Sigma$, whose inverse $\Sigma^{-\frac{1}{2}}$ belongs to $\mathcal{L}_d^{+}(\mathbb{C})$ as well. The function F in this example is associated with 'White Noise'. Then for all $n \in \mathbb{N}$

$$\phi_n = \Sigma^{-\frac{1}{2}} e_n \text{ and } \phi_n^{\rho} = \Sigma^{-\frac{1}{2}} e_0.$$

Example 2: Suppose

$$dF(\lambda) = \beta(e^{i\lambda}\mathrm{I})^{-1}\, dF_0(\lambda)\, \beta(e^{i\lambda}\mathrm{I})^{-1*}, \tag{9}$$

where β is a polynomial of degree $\eta \in \mathbb{N}^{+}$ with a leading coefficient equal to I, while F_0 is such as in example 1. Moreover all roots of the equation $|\beta(z\mathrm{I})| = 0$ are lying inside the unit circle, where $|\cdot|$ denotes the determinant. This function F is associated with 'autoregression of η^{th} order'. Then the first $\eta + 1$ polynomials ϕ_n and ϕ_n^{ρ}, $0 \le n \le \eta$, are constructed according to the recurrence procedure described in subsection 2.4 and for all $n \ge \eta$

$$\phi_n = \Sigma^{-\frac{1}{2}}\, \beta\, e_{n-\eta} \text{ and } \phi_n^{\rho} = \phi_\eta^{\rho} \tag{10}$$

(see subsections 2.4 and 4.4 for more details).

2.3 Christoffel-Darboux Formula for Kernel Polynomials

In this subsection kernel polynomials are introduced. We will give a short proof of the well-known Christoffel-Darboux formula.

Let $n \in \mathbb{N}$. For fixed $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$, define the kernel polynomial $s_n^{\mathrm{A}} : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ by

$$s_n^{\mathrm{A}} := \sum_{k=0}^{n} \phi_k(\mathrm{A})^* \phi_k. \tag{11}$$

Sometimes we will write $s_n(\mathrm{A}, \cdot)$ instead of s_n^{A}.

Theorem 3 *Let $n \in \mathbb{N}$. Then for a fixed $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$*

$$< p; s_n^{\mathrm{A}} >= p(\mathrm{A}) \tag{12}$$

for all polynomials p of degree $\le n$. The kernel polynomial s_n^{A} is uniquely determined by the reproducing property (12).

Proof: A polynomial p of degree $\leq n$ can always be written in the form $p = \sum_{k=0}^{n} C_k\phi_k$, with some matrices $C_k \in \mathcal{M}_d(\mathbb{C})$. Then

$$< p; s_n^A >= \sum_{k=0}^{n} C_k < \phi_k; s_n^A >= \sum_{k=0}^{n} C_k\phi_k(A) = p(A).$$

Suppose that there exists another polynomial $\hat{s}_n^A$ with the property (12). Then

$$< \phi_k; s_n^A - \hat{s}_n^A >= 0$$

for all $0 \leq k \leq n$, which yields $s_n^A = \hat{s}_n^A$. □

Note that with the notation (5) we also have the representation

$$s_n^A = (\mathbf{e}_n(A); \mathbf{e}_n). \tag{13}$$

This can be proved by checking property (12).

In the next theorem the Chistoffel-Darboux formula is derived (cf. [5]).

Theorem 4 *Let $n \in \mathbb{N}$. For a fixed $A \in \mathcal{M}_d(\mathbb{C})$*

$$s_n(A, \cdot) - e_1(A)^* s_n(A, \cdot)e_1(\cdot) = \phi_{n+1}^\rho(A)^*\phi_{n+1}^\rho(\cdot) - \phi_{n+1}(A)^*\phi_{n+1}(\cdot). \tag{14}$$

Proof: First note that

$$s_{n+1}(A, \cdot) = e_1(A)^* \sum_{k=0}^{n} \phi_k(A)^*\phi_k(\cdot)e_1(\cdot) + \phi_{n+1}^\rho(A)^*\phi_{n+1}^\rho(\cdot) \tag{15}$$

since $\{\phi_0 e_1, \ldots, \phi_n e_1, \phi_{n+1}^\rho\}$ is an orthonormal basis of span$\{e_0, \ldots, e_{n+1}\}$. This can be verified by using the definition of ϕ_{n+1}^ρ and the fact that $< \phi_k e_1; \phi_l e_1 >=< \phi_k; \phi_l >$.

Now (11) gives

$$s_{n+1}(A, \cdot) = s_n(A, \cdot) + \phi_{n+1}(A)^*\phi_{n+1}(\cdot)$$

and (11) and (15) give

$$s_{n+1}(A, \cdot) = e_1(A)^* s_n(A, \cdot)e_1(\cdot) + \phi_{n+1}^\rho(A)^*\phi_{n+1}^\rho(\cdot).$$

By combining these two formulas for $s_{n+1}(A, \cdot)$ we get (14). □

Corollary 1 *Fix $n \in \mathbb{N}$ and let $\Phi_{n+1,0}^\rho = \phi_{n+1}^\rho(0)$. Then*

$$s_{n+1}^0 = (\Phi_{n+1,0}^\rho)^*\phi_{n+1}^\rho \tag{16}$$

and

$$s_{n+1}(0,0) = (\Phi_{n+1,0}^\rho)^*\Phi_{n+1,0}^\rho. \tag{17}$$

Proof: The identity (16) is obtained by evaluating (15) at $A = 0$, and (17) is a special case of (16). □

2.4 Recurrence Relations

In this subsection recurrence relations between orthonormal polynomials and their reciprocals are derived. These relations yield an algorithm for constructing the polynomials $\{\phi_n\}_{n\in\mathbb{N}}$ recurrently.

Theorem 5 *For each $n \in \mathbb{N}$ the orthonormal polynomials ϕ_n and ϕ_n^ρ given by*

$$\phi_n = \sum_{k=0}^{n} \Phi_{n,k} e_k \text{ and } \phi_n^\rho = \sum_{k=0}^{n} \Phi_{n,k}^\rho e_k \tag{18}$$

satisfy the following identity

$$\Phi_{n+1,n+1}^{-1} \phi_{n+1} = \Phi_{n,n}^{-1} \phi_n e_1 + C_n \phi_n^\rho \tag{19}$$

with

$$C_n = - < e_{n+1}; \phi_n^\rho > \tag{20}$$

so that

$$\Phi_{n+1,n+1}^{-1} (\Phi_{n+1,n+1}^{-1})^* = \Phi_{n,n}^{-1} (\Phi_{n,n}^{-1})^* - C_n C_n^*. \tag{21}$$

Proof: The polynomials $\phi_n e_1$ and ϕ_n^ρ are linearly independent. Moreover these polynomials, as well as the polynomial ϕ_{n+1}, are orthogonal to $e_1, \ldots, e_n$. Hence $\phi_{n+1} \in \text{span}\{\phi_n e_1; \phi_n^\rho\}$, i.e. $\phi_{n+1} = A\phi_n e_1 + B\phi_n^\rho$ for some $A, B \in \mathcal{M}_d(\mathbb{C})$. By comparing the coefficients of e_0 and e_{n+1} in the last expression, we get relation (19), provided

$$C_n = \Phi_{n+1,n+1}^{-1} \Phi_{n+1,0} (\Phi_{n,0}^\rho)^{-1}. \tag{22}$$

Thus, to complete the proof of (19), we have to verify (cf. (20) and (22))

$$\Phi_{n+1,n+1}^{-1} \Phi_{n+1,0} (\Phi_{n,0}^\rho)^{-1} = - < e_{n+1}; \phi_n^\rho > . \tag{23}$$

Let $\psi_n := \Phi_{n+1,n+1}^{-1} \phi_{n+1} - e_{n+1} - \Phi_{n+1,n+1}^{-1} \Phi_{n+1,0} e_0$. Then we have $< \psi_n; \phi_n^\rho >= 0$, since $\psi_n \in \text{span}\{e_1, \ldots, e_n\}$. This implies (23), due to $< \phi_{n+1}; \phi_n^\rho >= 0$ and

$$< e_0; \phi_n^\rho >= (\Phi_{n,0}^\rho)^{-1} < \phi_n^\rho; \phi_n^\rho >= (\Phi_{n,0}^\rho)^{-1}.$$

By taking the inner product with e_{n+1} in (19) and using the relation (20) we get

$$\Phi_{n+1,n+1}^{-1} < \phi_{n+1}; e_{n+1} >= \Phi_{n,n}^{-1} < \phi_n e_1; e_{n+1} > -C_n C_n^*,$$

which yields (21). Indeed $< \phi_n e_1; e_{n+1} >=< \phi_n e_1; \Phi_{n,n}^{-1} \phi_n e_1 >= (\Phi_{n,n}^{-1})^*$ since

$$< \phi_n e_1; \phi_n e_1 >=< \phi_n; \phi_n >= I.$$

□

The recurrence relations established in theorem 5 allow for the construction of the orthonormal polynomials and their reciprocals according to the following

Algorithm: Suppose ϕ_n, ϕ_n^ρ and s_n^0 are given. In order to determine $\phi_{n+1}, \phi_{n+1}^\rho$ and s_{n+1}^0, carry out the following calculations:

1. calculate C_n by using (20)
2. calculate $\Phi_{n+1,n+1} \in \mathcal{L}_d^+(\mathbb{C})$ by using (21)
3. determine ϕ_{n+1} by using (19)
4. determine $s_{n+1}^0 = s_n^0 + \Phi_{n+1,0}^* \phi_{n+1}$
5. calculate $\Phi_{n+1,0}^\rho \in \mathcal{L}_d^+(\mathbb{C})$ by using (17)
6. determine ϕ_{n+1}^ρ by using (16).

In theorem 5 ϕ_{n+1} is expressed in terms of ϕ_n and ϕ_n^ρ. It will be shown now that also ϕ_{n+1}^ρ can be expressed in terms of ϕ_n and ϕ_n^ρ.

Theorem 6 *For each $n \in \mathbb{N}$ the orthonormal polynomials ϕ_n and ϕ_n^ρ given by (18) satisfy the following identity*

$$(\Phi_{n+1,0}^\rho)^{-1}\phi_{n+1}^\rho = (\Phi_{n,0}^\rho)^{-1}\phi_n^\rho + D_n \phi_n e_1 \tag{24}$$

with

$$D_n = - < e_0; \phi_n e_1 > \tag{25}$$

so that

$$(\Phi_{n+1,0}^\rho)^{-1}(\Phi_{n+1,0}^\rho)^{-1*} = (\Phi_{n,0}^\rho)^{-1}(\Phi_{n,0}^\rho)^{-1*} - D_n D_n^*. \tag{26}$$

Proof: By using the same arguments as in theorem 5, we get $\phi_{n+1}^\rho = A\phi_n e_1 + B\phi_n^\rho$ for some $A, B \in \mathcal{M}_d(\mathbb{C})$. By comparing the coefficients of e_0 and e_{n+1} in the last expression, we get relation (24), provided

$$D_n = (\Phi_{n+1,0}^\rho)^{-1}\Phi_{n+1,n+1}^\rho \Phi_{n,n}^{-1}.$$

The proof of (25) is now based on

$$- < e_0; \phi_n e_1 >= (\Phi_{n+1,0}^\rho)^{-1}\Phi_{n+1,n+1}^\rho \Phi_{n,n}^{-1}.$$

To prove the last relation, use the same arguments as in theorem 5 when proving (23), making use of $\psi_n := (\Phi_{n+1,0}^\rho)^{-1}\phi_{n+1}^\rho - e_0 - (\Phi_{n+1,0}^\rho)^{-1}\Phi_{n+1,n+1}^\rho e_{n+1}$ this time and taking into consideration that $< \psi_n; \phi_n e_1 >= 0$.

By taking the inner product with e_0 in (24) we get the desired relation (26). □

Example: Consider the autoregressive case of example 2 in subsection 2.2. Since for all $n \geq \eta$ the orthonormal polynomials ϕ_n are given by (10), we have $\phi_n e_1 = \phi_{n+1}$. Thus $D_n = 0$ for all $n \geq \eta$ by definition (25). Due to (24) and (26) we get $\phi_n^\rho = \phi_\eta^\rho$ for all $n \geq \eta$ (cf. (10)).

2.5 Determinants

In this subsection we derive the commonly used expression for the determinant of the matrix $s_{n+1}(0,0)$ as the quotient of the determinants $|\mathrm{H}_n|$ and $|\mathrm{H}_{n+1}|$. We use here the representation (alternative to (17))

$$s_{n+1}(0,0)^{-1} = \mathrm{H}_0 - (\mathrm{B}_n;\mathrm{B}_n) \tag{27}$$

where

$$\mathrm{B}_n = \int_{(-\pi,\pi]} e^{i\lambda}\,\mathrm{e}_n(e^{i\lambda}\mathrm{I})\,dF(\lambda)\,.$$

In order to verify (27), note first that by (3) and the definition of the vector B_n, the block Toeplitz matrix H_{n+1} may be presented in the form

$$\mathrm{H}_{n+1} = \begin{pmatrix} \mathrm{H}_0 & \mathrm{B}_n^* \\ \mathrm{B}_n & \mathrm{H}_n \end{pmatrix} = \begin{pmatrix} \mathrm{I} & 0 \\ \mathrm{D} & \mathrm{I} \end{pmatrix}^* \begin{pmatrix} \mathrm{C} & 0 \\ 0 & \mathrm{H}_n \end{pmatrix} \begin{pmatrix} \mathrm{I} & 0 \\ \mathrm{D} & \mathrm{I} \end{pmatrix}$$

with $\mathrm{D} = \mathrm{H}_n^{-1}\mathrm{B}_n$ and $\mathrm{C} = \mathrm{H}_0 - \mathrm{B}_n^*\mathrm{H}_n^{-1}\mathrm{B}_n$. Therefore

$$\mathrm{H}_{n+1}^{-1} = \begin{pmatrix} \mathrm{I} & 0 \\ -\mathrm{D} & \mathrm{I} \end{pmatrix} \begin{pmatrix} \mathrm{C}^{-1} & 0 \\ 0 & \mathrm{H}_n^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{I} & 0 \\ -\mathrm{D} & \mathrm{I} \end{pmatrix}^* . \tag{28}$$

By (5), (13) and (28) we get (27), since

$$\mathrm{C}^{-1} = \mathrm{U}_{n+1}^{n+1*}\mathrm{H}_{n+1}^{-1}\mathrm{U}_{n+1}^{n+1} = s_{n+1}(0,0). \tag{29}$$

Theorem 7 *For all $n \in \mathbb{N}$*

$$|s_{n+1}(0,0)| = \frac{|\mathrm{H}_n|}{|\mathrm{H}_{n+1}|}. \tag{30}$$

<u>Proof:</u> By (28) and (29) we get (30):

$$|\mathrm{H}_{n+1}|^{-1} = |\mathrm{C}^{-1}|\,|\mathrm{H}_n|^{-1} = |s_{n+1}(0,0)|\,|\mathrm{H}_n|^{-1}.$$

□

Since $s_0(\mathrm{A},\mathrm{Z}) \equiv \mathrm{H}_0^{-1}$ by (5) and (13), we get by applying (17) and (30) that

$$\frac{1}{|\mathrm{H}_n|} = \prod_{k=0}^{n} |s_k(0,0)| = \prod_{k=0}^{n} \overline{|\Phi_{k,0}^{\rho}|}|\Phi_{k,0}^{\rho}| \tag{31}$$

for all $n \in \mathbb{N}$. In case $dF = \overline{dF}$ we also have $\Phi_{k,0}^{\rho} = \overline{\Phi}_{k,k}$, in view of theorem 2. Therefore, in this particular case, (31) reduces to

$$\frac{1}{|\mathrm{H}_n|} = \prod_{k=0}^{n} |\Phi_{k,k}|\overline{|\Phi_{k,k}|}. \tag{32}$$

3 Construction of Polynomials

In contrast to section 2, we focus here our attention to matrix spectral distribution functions F of a special form: we assume that F is related to another matrix spectral distribution function F_0 by

$$dF(\lambda) = r(e^{i\lambda}\mathrm{I})\, dF_0(\lambda)\, r(e^{i\lambda}\mathrm{I})^*, \tag{33}$$

where r is a polynomial with a leading coefficient equal to I. Spectral functions F of the form (33) are often encountered in time series analysis, for it is associated with moving average models in which the observed time series with the spectral distribution F is modelled as a linear transformation of some basic time series with the spectral distribution F_0 (not necessarily White Noise), and the transformation is characterized by a polynomial transfer function r (see e.g. [7], [8] or [9]). We will show how to construct the orthonormal polynomials and kernel polynomials associated with F, given the transfer polynomial r and the orthonormal polynomials and kernel polynomials associated with F_0.

3.1 Product of Polynomials

As was mentioned in subsection 2.1, the product of a polynomial and a monomial is a polynomial, but obviously the product of two arbitrary polynomials is not a polynomial since matrices do not in general commute. It is necessary therefore to modify this notion in the following natural way:

Let p and q be polynomials. If $p = \sum_{k=0}^{n} \mathrm{P}_k e_k$, for some matrices $\mathrm{P}_k \in \mathcal{M}_d(\mathbb{C})$, then the *product* of the polynomials p and q, denoted by $p \lhd q$, is the polynomial

$$p \lhd q = \sum_{k=0}^{n} \mathrm{P}_k\, q\, e_k.$$

In general $p \lhd q(\mathrm{Z}) \neq p(\mathrm{Z})q(\mathrm{Z})$ (we write $p \lhd q(\mathrm{Z})$ instead of $p(\mathrm{Z}) \lhd q(\mathrm{Z})$, for simplicity), unless $q(\mathrm{Z})$ and Z are commutative. For instance, $p \lhd q(z\mathrm{I}) = p(z\mathrm{I})q(z\mathrm{I})$ for $z \in \mathbb{C}$. Observe also that the product $p \lhd q$ evaluated at $\mathrm{Z} \in \mathcal{M}_d^q(\mathbb{C})$, where for a fixed $q \in \mathcal{P}$ the subset $\mathcal{M}_d^q(\mathbb{C})$ of $\mathcal{M}_d(\mathbb{C})$ consists of all matrices Z such that $q(\mathrm{Z})$ is invertible, satisfies

$$p \lhd q(\mathrm{Z}) = p(\mathrm{Z}_q)\, q(\mathrm{Z}) \tag{34}$$

with $\mathrm{Z}_q := q(\mathrm{Z})\,\mathrm{Z}\, q(\mathrm{Z})^{-1}$. This is easily verified since $\mathrm{Z}_q^n = q(\mathrm{Z})\mathrm{Z}^n q(\mathrm{Z})^{-1}$ for $n \in \mathbb{N}$.

The binary operation $\lhd$ is closed, i.e. $p \lhd q \in \mathcal{P}$ for all $p, q \in \mathcal{P}$. Moreover $[\mathcal{P}, \lhd]$ is a monoid, since

(i) the operation $\lhd$ is associative: $p \lhd (q \lhd r) = (p \lhd q) \lhd r$ for all $p, q, r \in \mathcal{P}$, and

(ii) there is an identity, namely e_0.

Property (i) is verified as follows: if $p = \sum_{k=0}^{n} \mathrm{P}_k e_k$ and $q = \sum_{l=0}^{m} \mathrm{Q}_l e_l$ then the polynomial $p \triangleleft q = \sum_{k=0}^{n} \sum_{l=0}^{m} \mathrm{P}_k \mathrm{Q}_l e_{l+k}$ and therefore

$$p \triangleleft (q \triangleleft r) = \sum_{k=0}^{n} \mathrm{P}_k (q \triangleleft r) e_k = \sum_{k=0}^{n} \mathrm{P}_k (\sum_{l=0}^{m} \mathrm{Q}_l r e_l) e_k = \sum_{k=0}^{n} \sum_{l=0}^{m} \mathrm{P}_k \mathrm{Q}_l r e_{l+k} = (p \triangleleft q) \triangleleft r.$$

The algebraic system $[\mathcal{P}, +, \triangleleft]$ is a ring since

(i) $[\mathcal{P}, +]$ is an abelian group.

(ii) $[\mathcal{P}, \triangleleft]$ is a semigroup.

(iii) the operation $\triangleleft$ is distributive over the operation $+$.

Let $\mathcal{M}_d^-(\mathbb{C})$ denote the set of all invertible $d \times d$ complex valued matrices. The ring $[\mathcal{P}, +, \triangleleft]$ has the following additional property that if $p, r \in \mathcal{P}$ and the leading coefficient of r belongs to $\mathcal{M}_d^-(\mathbb{C})$, then

$$p \triangleleft r = 0 \;\Rightarrow\; p = 0. \tag{35}$$

This property is verified as follows: suppose $p \neq 0$, so $p = \sum_{k=0}^{n} \mathrm{P}_k e_k$ with $\mathrm{P}_n \neq 0$. Let $r = \sum_{l=0}^{m} \mathrm{R}_l e_l$ with $\mathrm{R}_m \in \mathcal{M}_d^-(\mathbb{C})$. Then by assumption $p \triangleleft r = \sum_{k=0}^{n} \sum_{l=0}^{m} \mathrm{P}_k \mathrm{R}_l e_{l+k} = 0$. Hence the coefficient of e_{m+n} equals zero, i.e. $\mathrm{P}_n \mathrm{R}_m = 0$. Since $\mathrm{R}_m \in \mathcal{M}_d^-(\mathbb{C})$ this yields $\mathrm{P}_n = 0$, which contradicts the assumption $\mathrm{P}_n \neq 0$. Thus $p = 0$ and (35) is proved.

Due to distributivity, we have the following important consequence of (35): if $p, q, r \in \mathcal{P}$ and the leading coefficient of r belongs to $\mathcal{M}_d^-(\mathbb{C})$, then

$$p \triangleleft r = q \triangleleft r \;\Rightarrow\; p = q. \tag{36}$$

Along with the operation $\triangleleft$, we need in subsection 3.2 the operation $\triangleright$ dual to $\triangleleft$, in the sense that

$$p^* \triangleright q^* = (p \triangleleft q)^*$$

for all $p, q \in \mathcal{P}$.

3.2 Extension of the Christoffel-Darboux Formula

From now on we consider the case of two matrix spectral distribution functions F and F_0 related by (33) where

$$r = \sum_{m=0}^{\eta} \mathrm{R}_m e_m \tag{37}$$

is a polynomial of degree $\eta \in \mathbb{N}$ with $\mathrm{R}_\eta = \mathrm{I}$.

The extension of the Christoffel-Darboux formula, asserted in theorem 9, relates kernel polynomials associated with F and kernel polynomials associated with F_0. It is therefore necessary to index these polynomials (as well as orthonormal polynomials, inner products, etc.) by F and F_0 respectively.

Let $\mathcal{P}^r$ be the linear space of the polynomials $p \triangleleft r$ with $p \in \mathcal{P}$:

$$\mathcal{P}^r := \{p \triangleleft r | p \in \mathcal{P}\}.$$

For all polynomials $p, q \in \mathcal{P}$ we have

$$< p \triangleleft r; q \triangleleft r >_{F_0} = < p; q >_F . \tag{38}$$

Indeed, since $p \triangleleft r$ evaluated at $e^{i\lambda}\mathrm{I}$ coincides with $p\, r$ at $e^{i\lambda}\mathrm{I}$, we get

$$\begin{aligned} < p \triangleleft r; q \triangleleft r >_{F_0} &= \int_{(-\pi,\pi]} p(e^{i\lambda}\mathrm{I})r(e^{i\lambda}\mathrm{I})\, dF_0(\lambda)\, r(e^{i\lambda}\mathrm{I})^* q(e^{i\lambda}\mathrm{I})^* \\ &= \int_{(-\pi,\pi]} p(e^{i\lambda}\mathrm{I})\, dF(\lambda)\, q(e^{i\lambda}\mathrm{I})^* = < p; q >_F . \end{aligned}$$

We will now give an orthonormal basis (with respect to $< \cdot ; \cdot >_{F_0}$) of $\mathcal{P}^r$:

Theorem 8 *Let $\{\phi_n^F\}_{n \in \mathbb{N}}$ be the system of orthonormal polynomials associated with F. The system of polynomials $\{\phi_n^F \triangleleft r\}_{n \in \mathbb{N}}$ is the unique orthonormal basis of $\mathcal{P}^r$ such that*

1. *$\phi_n^F \triangleleft r$ is a polynomial of degree $n + \eta$, belonging to $\mathcal{P}^r$.*
2. *the leading coefficient of $\phi_n^F \triangleleft r$ belongs to $\mathcal{L}_d^+(\mathbb{C})$.*
3. *the polynomials $\{\phi_n^F \triangleleft r\}_{n \in \mathbb{N}}$ are orthonormal with respect to $< \cdot ; \cdot >_{F_0}$.*

Proof: In view of (38) the properties 1-3 of the polynomials $\phi_n^F \triangleleft r$ are directly verified. Uniqueness of the system $\{\phi_n^F \triangleleft r\}_{n \in \mathbb{N}}$ can be proved as follows. Suppose that there is another system $\{\hat{\phi}_n^F \triangleleft r\}_{n \in \mathbb{N}}$ that satisfies the three conditions of this theorem. Then by (38) $\{\hat{\phi}_n^F\}_{n \in \mathbb{N}}$ is a system of orthonormal polynomials satisfying the conditions of theorem 1. Thus $\phi_n = \hat{\phi}_n$ for all $n \in \mathbb{N}$, that is the system $\{\phi_n^F \triangleleft r\}_{n \in \mathbb{N}}$ is unique. □

For all $n \in \mathbb{N}$, denote by $\mathcal{P}_n$ the linear space of the polynomials of degree $\leq n$. We will need the following unique decomposition of the space $\mathcal{P}_{n+\eta}$:

$$\mathcal{P}_{n+\eta} = \mathcal{P}^r_{n+\eta} \oplus \mathcal{P}^{r\perp}_{n+\eta} \tag{39}$$

with

$$\mathcal{P}^r_{n+\eta} := \{p \triangleleft r | p \in \mathcal{P}_n\} \text{ and } \mathcal{P}^{r\perp}_{n+\eta} := \{q | q \in \mathcal{P}_{n+\eta} \; and \; < p; q >_{F_0} = 0 \; for \; all \; p \in \mathcal{P}^r_{n+\eta}\}.$$

By theorem 8, the system of polynomials $\{\phi_k^F \triangleleft r : 0 \leq k \leq n\}$ is an orthonormal basis of $\mathcal{P}^r_{n+\eta}$. To construct an orthonormal basis of $\mathcal{P}^{r\perp}_{n+\eta}$ we need the following polynomials:

For a fixed $n \in \mathbb{N}$ and a fixed $k \in \{0,\ldots,\eta-1\}$ let

$$\zeta_{k,n}^{F_0} = \sum_{l=0}^{n} \mathrm{B}_l\,(e_l \triangleleft r) + \sum_{m=0}^{\eta-1-k} \mathrm{C}_{\eta-1-m}\, e_{\eta-1-m}$$

with $\mathrm{B}_l, \mathrm{C}_{\eta-1-m} \in \mathcal{M}_d(\mathbb{C})$, be the unique polynomial such that

1. $\zeta_{k,n}^{F_0}$ belongs to the span of $e_0 \triangleleft r, \ldots, e_n \triangleleft r, e_k, \ldots, e_{\eta-1}$.
2. $\zeta_{k,n}^{F_0}$ is orthogonal (with respect to F_0) to $e_0 \triangleleft r, \ldots, e_n \triangleleft r, e_{k+1}, \ldots, e_{\eta-1}$.
3. the norm (with respect to F_0) of $\zeta_{k,n}^{F_0}$ equals I.
4. the coefficient C_k of $\zeta_{k,n}^{F_0}$ belongs to $\mathcal{L}_d^{+}(\mathbb{C})$.

Thus, the system of polynomials $\{\zeta_{k,n}^{F_0} : 0 \leq k \leq \eta - 1\}$ is an orthonormal basis of $\mathcal{P}_{n+\eta}^{r\perp}$. (the explicit expression for the polynomials $\zeta_{k,n}^{F_0}$ will be given in the remark at the end of subsection 3.3). Due to this fact we can establish the following relationship between kernel polynomials associated with F and kernel polynomials associated with F_0.

Theorem 9 *For each* $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ *and each* $n \in \mathbb{N}$

$$r(\mathrm{A})^* \triangleright s_n^F(\mathrm{A},\cdot) \triangleleft r(\cdot) = s_n^{F_0}(\mathrm{A},\cdot) - \sum_{k=0}^{\eta-1} \left[\zeta_{k,n}^{F_0}(\mathrm{A})^* \zeta_{k,n}^{F_0}(\cdot) - \phi_{n+k+1}^{F_0}(\mathrm{A})^* \phi_{n+k+1}^{F_0}(\cdot)\right]. \qquad (40)$$

By the definition in subsection 3.1 of the operation $\triangleleft$ and its dual $\triangleright$, it is easily verified that if Z and $r(\mathrm{Z})$ are commutative, then the left-hand side of (40) evaluated at Z reduces to $r(\mathrm{A})^* s_n^F(\mathrm{A},\mathrm{Z}) r(\mathrm{Z})$ and if $\mathrm{A},\mathrm{Z} \in \mathcal{M}_d^r(\mathbb{C})$, then it reduces to $r(\mathrm{A})^* s_n^F(\mathrm{A}_r,\mathrm{Z}_r) r(\mathrm{Z})$.

The relationship (40) can be viewed as an extension of the Christoffel-Darboux formula. Indeed, with the choice $r = e_1$ (so $\eta = 1$) we get $dF = dF_0$. Then $\zeta_{0,n}^{F_0} = \phi_{n+1}^{F_0 \rho}$ and formula (40) reduces to the Christoffel-Darboux formula (14).

<u>Proof:</u> Let $n \in \mathbb{N}$ and $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ be fixed. Define the kernel polynomial $s_{n+\eta}^r(\mathrm{A},\cdot) : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ by

$$s_{n+\eta}^r(\mathrm{A},\cdot) := \sum_{k=0}^{n} [\phi_k^F \triangleleft r(\mathrm{A})]^*\, \phi_k^F \triangleleft r(\cdot). \qquad (41)$$

The polynomial $s_{n+\eta}^r(\mathrm{A},\cdot)$ is the only polynomial in $\mathcal{P}_{n+\eta}^r$ with the following reproducing property:

$$< p \triangleleft r(\cdot); s_{n+\eta}^r(\mathrm{A},\cdot) >_{F_0} = p \triangleleft r(\mathrm{A}) \qquad (42)$$

for all polynomials $p \in \mathcal{P}_n$. Due to (11), the expression (41) reduces to

$$s_{n+\eta}^r(\mathrm{A},\cdot) = r(\mathrm{A})^* \triangleright s_n^F(\mathrm{A},\cdot) \triangleleft r(\cdot). \qquad (43)$$

The kernel polynomial

$$s^{r\perp}_{n+\eta}(\mathrm{A},\cdot)=\sum_{k=0}^{\eta-1}\zeta^{F_0}_{k,n}(\mathrm{A})^*\zeta^{F_0}_{k,n}(\cdot) \tag{44}$$

belongs to $\mathcal{P}^{r\perp}_{n+\eta}$. It is the only polynomial in $\mathcal{P}^{r\perp}_{n+\eta}$ with the following reproducing property:

$$< q(\cdot); s^{r\perp}_{n+\eta}(\mathrm{A},\cdot) >_{F_0}= q(\mathrm{A})$$

for all $q\in\mathcal{P}^{r\perp}_{n+\eta}$.

We now show that the unique kernel polynomial $s^{F_0}_{n+\eta}(\mathrm{A},\cdot)$ in $\mathcal{P}_{n+\eta}$ can be decomposed as

$$s^{F_0}_{n+\eta}(\mathrm{A},\cdot)=s^{r}_{n+\eta}(\mathrm{A},\cdot)+s^{r\perp}_{n+\eta}(\mathrm{A},\cdot) \tag{45}$$

with $s^{r}_{n+\eta}(\mathrm{A},\cdot)\in\mathcal{P}^{r}_{n+\eta}$ and $s^{r\perp}_{n+\eta}(\mathrm{A},\cdot)\in\mathcal{P}^{r\perp}_{n+\eta}$ and this decomposition is unique. Obviously it suffices to verify that the sum on the right-hand side of (45) possesses the reproducing property in $\mathcal{P}_{n+\eta}$. Indeed, for all $p\in\mathcal{P}_{n+\eta}$

$$< p(\cdot); s^{r}_{n+\eta}(\mathrm{A},\cdot)+s^{r\perp}_{n+\eta}(\mathrm{A},\cdot) >_{F_0}= p(\mathrm{A}),$$

in view of the decomposition (39) and the reproducing property of $s^{r}_{n+\eta}(\mathrm{A},\cdot)$ and $s^{r\perp}_{n+\eta}(\mathrm{A},\cdot)$ in $\mathcal{P}^{r}_{n+\eta}$ and $\mathcal{P}^{r\perp}_{n+\eta}$, respectively.

Now formulas (43), (44) and (45) yield

$$s^{F_0}_{n+\eta}(\mathrm{A},\cdot)=r(\mathrm{A})^*\triangleright s^{F}_{n}(\mathrm{A},\cdot)\triangleleft r(\cdot)+\sum_{k=0}^{\eta-1}\zeta^{F_0}_{k,n}(\mathrm{A})^*\zeta^{F_0}_{k,n}(\cdot).$$

Obviously, we also have (cf. (11))

$$s^{F_0}_{n+\eta}(\mathrm{A},\cdot)=s^{F_0}_{n}(\mathrm{A},\cdot)+\sum_{k=0}^{\eta-1}\phi^{F_0}_{n+k+1}(\mathrm{A})^*\phi^{F_0}_{n+k+1}(\cdot).$$

By combining these two formulas for $s^{F_0}_{n+\eta}(\mathrm{A},\cdot)$, we get (40). □

3.3 Construction of Kernel Polynomials

In this subsection we establish yet another relationship between the kernel polynomials associated with F and F_0, respectively. Consequently useful expressions are derived for constructing the kernel s^F_n and the inverse of the block Toeplitz matrix H^F_n.

Fix $n\in\mathbb{N}$. With the polynomial r given by (37) associate the matrix $\mathrm{C}^*_{n,\eta}\in\mathcal{M}_{n\times n+\eta}$ such that for $k=0,\ldots,n$ and $l=0,\ldots,n+\eta$

$$\mathrm{U}^{k*}_{n}\mathrm{C}^*_{n,\eta}\mathrm{U}^{l}_{n+\eta}=\begin{cases}\mathrm{R}_{l-k} & \text{if } 0\le l-k\le\eta\\ 0 & \text{otherwise}\end{cases}$$

where $\mathrm{R}_0, \ldots, \mathrm{R}_{\eta-1}, \mathrm{R}_\eta = \mathrm{I}$ are the coefficients of the polynomial r. It is simply verified that

$$\mathrm{H}_n^F = \mathrm{C}_{n,\eta}^* \mathrm{H}_{n+\eta}^{F_0} \mathrm{C}_{n,\eta} \tag{46}$$

due to (3), (33) and

$$\mathrm{C}_{n,\eta}^* \mathbf{e}_{n+\eta} = \mathbf{e}_n \triangleleft r, \tag{47}$$

where $\mathbf{e}_n \triangleleft r \in \mathcal{M}_{n\times 0}$ is the vector with components $e_j \triangleleft r$, $j = 0, \ldots, n$. By the last relation the elements in $\mathcal{P}_{n+\eta}^r$ are characterized as the linear combinations of the polynomials $\mathrm{U}_n^{m*} \mathrm{C}_{n,\eta}^* \mathbf{e}_{n+\eta}$ with $m \in \{0, \ldots, n\}$. Indeed, if $p = \mathsf{P}_n \mathbf{e}_n$ belongs to $\mathcal{P}_n$ ($\mathsf{P}_n \in \mathcal{M}_{0\times n}$), then by (47) the elements in $\mathcal{P}_{n+\eta}^r$ are given by

$$p \triangleleft r = \mathsf{P}_n \mathbf{e}_n \triangleleft r = \mathsf{P}_n \mathrm{C}_{n,\eta}^* \mathbf{e}_{n+\eta}. \tag{48}$$

We will now characterize the elements in $\mathcal{P}_{n+\eta}^{r\perp}$. To this end, we first determine the null space of $\mathrm{C}_{n,\eta}^*$. Let $\mathrm{C}_{n,\eta}^* = (\mathcal{C}_{n,1}^*, \mathcal{C}_{n,2}^*)$, where $\mathcal{C}_{n,1}^*$ and $\mathcal{C}_{n,2}^*$ are elements of $\mathcal{M}_{n\times\eta-1}$ and $\mathcal{M}_{n\times n}$ respectively. Note that $\mathcal{C}_{n,2}^*$ is invertible. Let $\mathrm{V}_{n,\eta}^* = (\mathcal{V}_{n,1}^*, \mathcal{V}_{n,2}^*)$, where $\mathcal{V}_{n,1}^*$ and $\mathcal{V}_{n,2}^*$ are elements of $\mathcal{M}_{\eta-1\times\eta-1}$ and $\mathcal{M}_{\eta-1\times n}$ respectively, such that

(i) $\mathcal{V}_{n,1}^*$ is arbitrary but invertible and

(ii) the columns of $\mathrm{V}_{n,\eta}$ span the null space of $\mathrm{C}_{n,\eta}^*$, i.e. $\mathrm{C}_{n,\eta}^* \mathrm{V}_{n,\eta} = 0$.

By property (ii) we have $\mathcal{C}_{n,1}^* \mathcal{V}_{n,1} + \mathcal{C}_{n,2}^* \mathcal{V}_{n,2} = 0$. So

$$\mathcal{V}_{n,2} = -(\mathcal{C}_{n,2}^*)^{-1} \mathcal{C}_{n,1}^* \mathcal{V}_{n,1} = -(\mathcal{C}_{n,1} \mathcal{C}_{n,2}^{-1})^* \mathcal{V}_{n,1}. \tag{49}$$

We will show now that the elements in $\mathcal{P}_{n+\eta}^{r\perp}$ are characterized as the linear combinations of the polynomials $\mathrm{U}_{\eta-1}^{m*}(\mathrm{V}_{n,\eta}; \mathbf{e}_{n+\eta})^{F_0}$ with $m \in \{0, \ldots, \eta-1\}$, where we again use the notation (5). Indeed, fix $\mathsf{Q}_{\eta-1} \in \mathcal{M}_{0\times\eta-1}$ and consider the polynomial

$$\mathsf{Q}_{\eta-1}(\mathrm{V}_{n,\eta}; \mathbf{e}_{n+\eta})^{F_0}. \tag{50}$$

It suffices to show that the inner product (with respect to F_0) of this polynomial with an element in $\mathcal{P}_{n+\eta}^r$, say $p \triangleleft r \in \mathcal{P}_{n+\eta}^r$, equals 0. This is directly verified by using (4) and (48):

$$< p \triangleleft r; q >_{F_0} = \mathsf{P}_n \mathrm{C}_{n,\eta}^* \mathrm{H}_{n+\eta}^{F_0} (\mathrm{H}_{n+\eta}^{F_0})^{-1} \mathrm{V}_{n,\eta} \mathsf{Q}_{\eta-1}^* = 0.$$

We are now in a position to establish the relationship between the kernel polynomials associated with F and F_0, respectively.

Theorem 10 *Let* $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ *be fixed. Then*

$$\begin{aligned} r(\mathrm{A})^* \triangleright s_n^F(\mathrm{A}, \cdot) \triangleleft r(\cdot) &= (\mathbf{e}_{n+\eta}(\mathrm{A}); \mathbf{e}_{n+\eta}(\cdot))^{F_0} - \\ &\quad (\mathbf{e}_{n+\eta}(\mathrm{A}); \mathrm{V}_{n,\eta})^{F_0} [(\mathrm{V}_{n,\eta}; \mathrm{V}_{n,\eta})^{F_0}]^{-1} (\mathrm{V}_{n,\eta}; \mathbf{e}_{n+\eta}(\cdot))^{F_0}. \end{aligned} \tag{51}$$

Recall that $(\mathrm{e}_{n+\eta}(\mathrm{A}); \mathrm{e}_{n+\eta}(\cdot))^{F_0} = s^{F_0}_{n+\eta}(\mathrm{A}, \cdot)$; cf. (13).

Proof: Let $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ be fixed. As was already noted, the first term in the right-hand side of (51) is the kernel polynomial $s^{F_0}_{n+\eta}(\mathrm{A}, \cdot)$. By (43) the left-hand side of (51) is the kernel polynomial $s^{r}_{n+\eta}(\mathrm{A}, \cdot)$. By (45) it suffices thus to proof

$$s^{r\perp}_{n+\eta}(\mathrm{A}, \cdot) = (\mathrm{e}_{n+\eta}(\mathrm{A}); \mathrm{V}_{n,\eta})^{F_0}[(\mathrm{V}_{n,\eta}; \mathrm{V}_{n,\eta})^{F_0}]^{-1}(\mathrm{V}_{n,\eta}; \mathrm{e}_{n+\eta}(\cdot))^{F_0}. \tag{52}$$

Since by (50) each polynomial in $\mathcal{P}^{r\perp}_{n+\eta}$ has the form $\mathrm{Q}_{\eta-1}(\mathrm{V}_{n,\eta}; \mathrm{e}_{n+\eta})^{F_0}$ with some $\mathrm{Q}_{\eta-1} \in \mathcal{M}_{0\times\eta-1}$, it is easily verified that the kernel polynomial given by (52) belongs to $\mathcal{P}^{r\perp}_{n+\eta}$ and has the reproducing property in $\mathcal{P}^{r\perp}_{n+\eta}$. □

Corollary 2 *Let* $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ *be fixed. Then*

$$r(\mathrm{A})^* \triangleright s^F_n(\mathrm{A}, \cdot) \triangleleft r(\cdot) = \mathrm{e}_{n+\eta}(\mathrm{A})^* \,\Pi_{n+\eta}\, \mathrm{e}_{n+\eta}(\cdot), \tag{53}$$

where

$$\Pi_{n+\eta} = [\mathrm{H}^{F_0}_{n+\eta}]^{-1} - [\mathrm{H}^{F_0}_{n+\eta}]^{-1}\mathrm{V}_{n,\eta}(\mathrm{V}^*_{n,\eta}[\mathrm{H}^{F_0}_{n+\eta}]^{-1}\mathrm{V}_{n,\eta})^{-1}\mathrm{V}^*_{n,\eta}[\mathrm{H}^{F_0}_{n+\eta}]^{-1}. \tag{54}$$

Proof: By using (5) and (13), formula (53) is a direct consequence of (51). □

It is shown in the next theorem how the inverse of the Toeplitz matrix H^F_n is calculated, given the inverse of the Toeplitz matrix $\mathrm{H}^{F_0}_{n+\eta}$ and the polynomial r.

Theorem 11 *For each* $n \in \mathbb{N}$

$$[\mathrm{H}^F_n]^{-1} = \mathrm{T}^*_{n,\eta}\Pi_{n+\eta}\mathrm{T}_{n,\eta} \tag{55}$$

with the matrix $\mathrm{T}_{n,\eta} \in \mathcal{M}_{n+\eta,n}$ *given by* $\mathrm{T}_{n,\eta} = (0, \mathcal{C}^{-1}_{n,2})^*$.

Proof: It suffices to prove that

$$\mathrm{e}_{n+\eta} = \mathrm{T}_{n,\eta}\mathrm{e}_n \triangleleft r + \mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1}\mathrm{e}_{\eta-1}. \tag{56}$$

Indeed, since $\Pi_{n+\eta}\mathrm{V}_{n,\eta} = 0$ the desired formula (55) follows from (53) and (56). In order to prove (56), note that (47) yields

$$\mathrm{e}_n \triangleleft r = \mathcal{C}^*_{n,1}\mathrm{e}_{\eta-1} + \mathcal{C}^*_{n,2}\mathrm{e}_n\, \mathrm{e}_\eta.$$

Hence by using (49) we get

$$\mathrm{e}_n\, \mathrm{e}_\eta = [\mathcal{C}^*_{n,2}]^{-1}\mathrm{e}_n \triangleleft r + \mathcal{V}_{n,2}\mathcal{V}^{-1}_{n,1}\mathrm{e}_{\eta-1}$$

which easily reduces to (56). □

Formula (55) yields the following explicit expression for the kernel polynomial $s^F_n(\mathrm{A}, \cdot)$ (cf. (51)).

Corollary 3 *For each $n \in \mathbb{N}$*

$$s_n^F(\mathrm{A},\cdot) = [\mathrm{T}_{n,\eta}\mathbf{e}_n(\mathrm{A})]^*\Pi_{n+\eta}\mathrm{T}_{n,\eta}\mathbf{e}_n(\cdot). \qquad (57)$$

Proof: By using (13) and (55) we get (57). □

Remark: If we choose for $\mathcal{V}_{n,1}$ the identity matrix, then the polynomials $\zeta_{k,n}^{F_0}$, $k = 0,\dots,\eta-1$ used in (40) coincide with the components of the vector

$$[(\mathrm{V}_{n,\eta};\mathrm{V}_{n,\eta})^{F_0}]^{-\frac{1}{2}}(\mathrm{V}_{n,\eta};\mathbf{e}_{n+\eta}(\cdot))^{F_0} \in \mathcal{M}_{\eta-1\times 0}.$$

Indeed, this can be checked by using (56).

3.4 Construction of Orthonormal Polynomials

The derivative of a polynomial $p = \sum_{k=0}^{n} \mathrm{P}_k e_k$ is defined as usual by

$$p^{(1)} := \sum_{k=1}^{n} k\mathrm{P}_k e_{k-1}.$$

The N^{th} derivative of a polynomial p is defined by $p^{(N)} := [p^{(N-1)}]^{(1)}$, with the convention $p = p^{(0)}$. Along with $p^{(N)}(\mathrm{Z})$, we also use the common notation $(\frac{d^N}{d\mathrm{Z}^N})p(\mathrm{Z})$ for the derivative $p^{(N)}$ evaluated at Z.

The relationship (51) between the kernel polynomials associated with F and F_0, respectively, provides for a similar relationship between the orthonormal polynomials ϕ_n^F and $\phi_n^{F_0}$: see the next theorem in which the polynomial $\xi_n^F = (\mathbf{U}_n^n;\mathbf{e}_n(\cdot))^F$ is related to the polynomial $\xi_{n+\eta}^{F_0} = (\mathbf{U}_{n+\eta}^{n+\eta};\mathbf{e}_{n+\eta}(\cdot))^{F_0}$ (cf. (6); recall that $\phi_n = \|\xi_n\|^{-1}\xi_n$).

Theorem 12 *For all $n \in \mathbb{N}$*

$$\begin{aligned}(\mathbf{U}_n^n;\mathbf{e}_n(\cdot))^F \lhd r(\cdot) \;=\;& (\mathbf{U}_{n+\eta}^{n+\eta};\mathbf{e}_{n+\eta}(\cdot))^{F_0} - \\ & (\mathbf{U}_{n+\eta}^{n+\eta};\mathrm{V}_{n,\eta})^{F_0}[(\mathrm{V}_{n,\eta};\mathrm{V}_{n,\eta})^{F_0}]^{-1}(\mathrm{V}_{n,\eta};\mathbf{e}_{n+\eta}(\cdot))^{F_0}. \end{aligned} \qquad (58)$$

Proof: We get the desired relation (58), by taking the $(n+\eta)^{th}$ derivative with respect to A^* of both sides in (51). The result on the right-hand side is obvious. The calculation on the left-hand side is equally simple, provided the following consequence of (47) is taken into consideration

$$s_n^F(\mathrm{A},\cdot) \lhd r(\cdot) = \mathbf{e}_n(\mathrm{A})^*(\mathrm{H}_n^F)^{-1}\mathrm{C}_{n,\eta}^*\mathbf{e}_{n+\eta}(\cdot).$$

Indeed, we then only need to verify that

$$\frac{1}{(n+\eta)!}[\mathrm{C}_{n,\eta}^*\mathbf{e}_{n+\eta}]^{(n+\eta)} = \mathrm{C}_{n,\eta}^*\mathbf{U}_{n+\eta}^{n+\eta} = \mathbf{U}_n^n$$

(recall the definition of the N^{th} derivative $p^{(N)}$ of a polynomial p at the end of subsection 3.1). □

Formula (58) yields the following explicit expression for the polynomials $\xi_n^F = (\mathbf{U}_n^n;\mathbf{e}_n(\cdot))^F$

Corollary 4 *For each $n \in \mathbb{N}$*

$$(\mathrm{U}_n^n;\mathrm{e}_n(\cdot))^F = \mathrm{U}_{n+\eta}^{n+\eta *}\Pi_{n+\eta}\mathrm{T}_{n,\eta}\mathrm{e}_n(\cdot). \tag{59}$$

Proof: By (54), (56) and (58) we get

$$(\mathrm{U}_n^n;\mathrm{e}_n(\cdot))^F \triangleleft r(\cdot) = \mathrm{U}_{n+\eta}^{n+\eta *}\Pi_{n+\eta}\mathrm{e}_{n+\eta}(\cdot) = \mathrm{U}_{n+\eta}^{n+\eta *}\Pi_{n+\eta}\mathrm{T}_{n,\eta}\mathrm{e}_n(\cdot) \triangleleft r(\cdot).$$

According to (36) this implies the desired relation (59). □

3.5 Determinants

The determinant of H_n^F can be calculated according to the following relationship.

Theorem 13 *For all $n \in \mathbb{N}$*

$$|\mathrm{H}_n^F| = |\mathrm{H}_{n+\eta}^{F_0}|\,\frac{|(\mathrm{V}_{n,\eta};\mathrm{V}_{n,\eta})^{F_0}|}{|\mathcal{V}_{n,1}^*||\mathcal{V}_{n,1}|} \tag{60}$$

Proof: Define

$$\mathrm{Q}_n := \begin{pmatrix} -\mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1}, & \mathrm{I}_n \end{pmatrix} = \begin{pmatrix} (\mathcal{C}_{n,1}\mathcal{C}_{n,2}^{-1})^*, & \mathrm{I}_n \end{pmatrix} = (\mathcal{C}_n\mathcal{C}_{n,2}^{-1})^*. \tag{61}$$

Then by (46)

$$\mathrm{Q}_n\mathrm{H}_{n+\eta}^{F_0}\mathrm{Q}_n^* = (\mathcal{C}_{n,2}^{-1})^*\mathrm{H}_n^F\mathcal{C}_{n,2}^{-1},$$

so that

$$|\mathrm{H}_n^F| = |\mathrm{Q}_n\mathrm{H}_{n+\eta}^{F_0}\mathrm{Q}_n^*| \tag{62}$$

by $|\mathcal{C}_{n,2}| = 1$. It will be now proved that

$$|\mathrm{Q}_n\mathrm{H}_{n+\eta}^{F_0}\mathrm{Q}_n^*| = |-\mathrm{I}_{\eta-1}|\begin{vmatrix} 0 & (\mathrm{V}_{n,\eta}\mathcal{V}_{n,1}^{-1})^* \\ \mathrm{V}_{n,\eta}\mathcal{V}_{n,1}^{-1} & \mathrm{H}_{n+\eta}^{F_0} \end{vmatrix}. \tag{63}$$

Note first that if $\mathrm{H}_{n+\eta}^{F_0}$ is partitioned as follows

$$\mathrm{H}_{n+\eta}^{F_0} = \begin{pmatrix} \mathrm{H}_{\eta-1}^{F_0} & \mathrm{B}^* \\ \mathrm{B} & \mathrm{H}_n^{F_0} \end{pmatrix},$$

then by definition (61)

$$\begin{aligned} \mathrm{Q}_n\mathrm{H}_{n+\eta}^{F_0}\mathrm{Q}_n^* &= \mathrm{H}_n^{F_0} + \mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1}\mathrm{H}_{\eta-1}^{F_0}(\mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1})^* - \mathrm{B}(\mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1})^* - \mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1}\mathrm{B}^* \\ &= \mathrm{H}_n^{F_0} - \begin{pmatrix} \mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1}, & \mathrm{B} \end{pmatrix}\begin{pmatrix} -\mathrm{H}_{\eta-1}^{F_0} & \mathrm{I}_{\eta-1} \\ \mathrm{I}_{\eta-1} & 0 \end{pmatrix}\begin{pmatrix} (\mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1})^* \\ \mathrm{B}^* \end{pmatrix}. \end{aligned}$$

Hence

$$\mathrm{Q}_n\mathrm{H}_{n+\eta}^{F_0}\mathrm{Q}_n^* = \mathrm{H}_n^{F_0} - \begin{pmatrix} \mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1}, & \mathrm{B} \end{pmatrix}\begin{pmatrix} 0 & \mathrm{I}_{\eta-1} \\ \mathrm{I}_{\eta-1} & \mathrm{H}_{\eta-1}^{F_0} \end{pmatrix}^{-1}\begin{pmatrix} (\mathcal{V}_{n,2}\mathcal{V}_{n,1}^{-1})^* \\ \mathrm{B}^* \end{pmatrix}, \tag{64}$$

since

$$\begin{pmatrix} -\mathrm{H}^{F_0}_{\eta-1} & \mathrm{I}_{\eta-1} \\ \mathrm{I}_{\eta-1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{I}_{\eta-1} \\ \mathrm{I}_{\eta-1} & \mathrm{H}^{F_0}_{\eta-1} \end{pmatrix}^{-1}.$$

We can now apply to (64) the following well-known formula for the determinant of a block matrix

$$\begin{vmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{vmatrix} = |\mathrm{A}||\mathrm{D} - \mathrm{C}\mathrm{A}^{-1}\mathrm{B}| = |\mathrm{D}||\mathrm{A} - \mathrm{B}\mathrm{D}^{-1}\mathrm{C}|. \tag{65}$$

We get (63), since

$$\begin{pmatrix} 0 & \mathrm{I}_{\eta-1} & (\mathcal{V}_{n,2}\mathcal{V}^{-1}_{n,1})^* \\ \mathrm{I}_{\eta-1} & \mathrm{H}^{F_0}_{\eta-1} & \mathrm{B}^* \\ \mathcal{V}_{n,2}\mathcal{V}^{-1}_{n,1} & \mathrm{B} & \mathrm{H}^{F_0}_{n} \end{pmatrix} = \begin{pmatrix} 0 & (\mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1})^* \\ \mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1} & \mathrm{H}^{F_0}_{n+\eta} \end{pmatrix}$$

and

$$\begin{vmatrix} 0 & \mathrm{I}_{\eta-1} \\ \mathrm{I}_{\eta-1} & \mathrm{H}^{F_0}_{\eta-1} \end{vmatrix} = |-\mathrm{I}_{\eta-1}| = |-\mathrm{I}_{\eta-1}|^{-1}.$$

Thus by (62) and (63) we have

$$|\mathrm{H}^{F}_{n}| = |-\mathrm{I}_{\eta-1}| \begin{vmatrix} 0 & (\mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1})^* \\ \mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1} & \mathrm{H}^{F_0}_{n+\eta} \end{vmatrix}.$$

To complete the proof, apply (65) to the right-hand side of the last relation. We get (60), since

$$\begin{aligned} |\mathrm{H}^{F}_{n}| &= |-\mathrm{I}_{\eta-1}||\mathrm{H}^{F_0}_{n+\eta}||-(\mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1})^*[\mathrm{H}^{F_0}_{n+\eta}]^{-1}\mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1}| \\ &= |\mathrm{H}^{F_0}_{n+\eta}||(\mathcal{V}^{-1}_{n,1})^*(\mathrm{V}_{n,\eta})^*[\mathrm{H}^{F_0}_{n+\eta}]^{-1}\mathrm{V}_{n,\eta}\mathcal{V}^{-1}_{n,1}|. \end{aligned}$$

□

4 Applications

In this section the results of section 3, concerning spectral distribution functions F of form (33) with arbitrary r and F_0, are specified in a number of special situations of practical importance. In subsections 4.1 - 4.3 F_0 is kept arbitrary, while r is specified as follows. First it is assumed that the transfer polynomial r possesses zeros which completely determine r, in the sense that the number of zeros coincide with the order η of r and besides the associated Vandermonde block matrix is invertible (see subsections 4.1 and 4.2; for notational convenience, we first deal with the case of simple zeros, and then with the case of multiple zeros). As usual, a $\mathrm{W} \in \mathcal{M}_d(\mathbb{C})$ is called a zero of r with multiplicity $m \in \mathbb{N}^+$ if $r(\mathrm{W}) = \ldots = r^{(m-1)}(\mathrm{W}) = 0$ and $r^{(m)}(\mathrm{W}) \neq 0$. Note that this assumption restricts the class of transfer polynomials only in case $d > 1$, for in the last case a matrix valued polynomial r need not have zeros. This particular situation deserves special attention, since here the results of section 3 take a useful form (see theorems 14 and 15).

Next, in subsection 4.3, we assume that the transformation characterized by r (of time series with the spectral distribution function F_0 to time series with the spectral distribution function F) is a result of η successive transformations characterized by transfer polynomials of degree 1, say $e_1 - \Theta_0 e_0, \ldots, e_1 - \Theta_{\eta-1} e_0$ with parameters $\Theta_j \in \mathcal{M}_d(\mathbb{C})$. In this case r is the product of these polynomials (in the sense of subsection 3.1; see (67)), and the parameters $\Theta_0, \ldots, \Theta_{\eta-1}$ are in general not zeros of r, unless $d = 1$ or $d > 1$ but $\eta = 1$. It will be shown that here the explicit expression (57) for $s_n^F(\mathrm{A}, \cdot)$ takes a useful form.

In subsections 4.4 - 4.7 we deal with spectral distribution functions which arise in some special time series models: in subsections 4.4, 4.5 and 4.6 we treat autoregression, first order moving average and ARMA(1,1) models respectively. Finally we discuss in subsection 4.7 how to apply the results of the present section to determine the likelihood of a Gaussian zero mean time series, working out the special case of an one dimensional ARMA(1,1) model.

4.1 Transfer Polynomial with Simple Zeros

Suppose that the polynomial r possesses simple zeros $\mathrm{W}_0, \ldots, \mathrm{W}_{\eta-1}$, and the Vandermonde block matrix $\mathbf{V}_\eta \in \mathcal{M}_{\eta-1\times\eta-1}$ associated with these simple zeros is invertible. This matrix is defined as usual: for $0 \leq k, l \leq \eta - 1$

$$\mathbf{U}_{\eta-1}^{k*} \mathbf{V}_\eta \mathbf{U}_{\eta-1}^{l} = (\mathrm{W}_l)^k.$$

We say in this case that r is completely described by its zeros $\mathrm{W}_0, \ldots, \mathrm{W}_{\eta-1}$. It will be shown next that under the present circumstances the right-hand side of (51), relating the kernel polynomials associated with F and F_0, can be expressed in terms of the kernel $s_{n+\eta}^{F_0}$. We use the following notations: the vector with components $s_{n+\eta}^{F_0}(\mathrm{W}_k, \cdot)$, $0 \leq k \leq \eta - 1$, is denoted by $s_{n+\eta}^{F_0}(\mathbf{W}, \cdot) \in \mathcal{M}_{\eta-1\times 0}$, i.e.

$$\mathbf{U}_{\eta-1}^{k*} s_{n+\eta}^{F_0}(\mathbf{W}, \cdot) = s_{n+\eta}^{F_0}(\mathrm{W}_k, \cdot).$$

Similarly, we denote by $s^{F_0}_{n+\eta}(\mathbf{W},\mathbf{W})$ the matrix in $\mathcal{M}_{\eta-1\times\eta-1}$ such that for $0\le k,l\le \eta-1$

$$\mathbf{U}^{k*}_{\eta-1}s^{F_0}_{n+\eta}(\mathbf{W},\mathbf{W})\mathbf{U}^{l}_{\eta-1}=s^{F_0}_{n+\eta}(\mathrm{W}_k,\mathrm{W}_l).$$

Theorem 14 *If the transfer polynomial r of degree $\eta\in\mathbb{N}^+$, given by (37), is completely described by the simple zeros $\mathrm{W}_0,\dots,\mathrm{W}_{\eta-1}$, then for each $\mathrm{A}\in\mathcal{M}_d(\mathbb{C})$*

$$r(\mathrm{A})^*\triangleright s^F_n(\mathrm{A},\cdot)\triangleleft r(\cdot)=s^{F_0}_{n+\eta}(\mathrm{A},\cdot)-s^{F_0}_{n+\eta}(\mathbf{W},\mathrm{A})^*s^{F_0}_{n+\eta}(\mathbf{W},\mathbf{W})^{-1}s^{F_0}_{n+\eta}(\mathbf{W},\cdot).\tag{66}$$

Proof: Formula (66) is a direct consequence of (51) where the matrix $\mathrm{V}_{n,\eta}$ is identified with the matrix $\mathbf{V}_{n,\eta}\in\mathcal{M}_{n+\eta-1\times\eta-1}$ which is defined as follows: for $0\le k\le n+\eta-1$ and $0\le l\le\eta-1$

$$\mathbf{U}^{k*}_{n+\eta-1}\mathbf{V}_{n,\eta}\mathbf{U}^{l}_{\eta-1}=(\mathrm{W}_l)^k.$$

Indeed, with this choice for $\mathrm{V}_{n,\eta}$ we have $\mathrm{C}^*_{n,\eta}\mathrm{V}_{n,\eta}=0$, since the W_j are zeros of r and $\mathcal{V}_{n,1}$ coincides with the Vandermonde block matrix $\mathbf{V}_\eta$ which is assumed invertible (recall that $\mathrm{V}^*_{n,\eta}=(\mathcal{V}^*_{n,1},\mathcal{V}^*_{n,2})$).

With $\mathrm{V}_{n,\eta}=\mathbf{V}_{n,\eta}$ we can write the right-hand side of (51) explicitly in terms of the the kernel polynomial $s^{F_0}_{n+\eta}(\mathrm{A},\cdot)$ evaluated at the zeros of r: for $0\le k,l\le\eta-1$ we have

$$\mathbf{U}^{k*}_{\eta-1}(\mathbf{V}_{n,\eta};\mathbf{e}_{n+\eta}(\cdot))^{F_0}=s^{F_0}_{n+\eta}(\mathrm{W}_k,\cdot)$$

and

$$\mathbf{U}^{k*}_{\eta-1}(\mathbf{V}_{n,\eta};\mathbf{V}_{n,\eta})^{F_0}\mathbf{U}^{l}_{\eta-1}=s^{F_0}_{n+\eta}(\mathrm{W}_k,\mathrm{W}_l)$$

which yields (66). □

4.2 Transfer Polynomial with Multiple Zeros

We say that the polynomial r of degree $\eta\in\mathbb{N}^+$ is completely described by its (multiple) zeros if r has zeros $\mathrm{W}_j\in\mathcal{M}_d(\mathbb{C})$ of multiplicity $\eta_j\in\mathbb{N}^+$, $0\le j\le\tau$, where $\eta=\sum_{j=0}^{\tau}\eta_j$, such that the generalized Vandermonde block matrix $\mathbf{V}_\eta=(\mathbf{V}^{(0)}_\eta,\dots,\mathbf{V}^{(\tau)}_\eta)\in\mathcal{M}_{\eta-1\times\eta-1}$ associated with these zeros is invertible. This matrix is defined as usual: for each $0\le j\le\tau$ the matrix $\mathbf{V}^{(j)}_\eta\in\mathcal{M}_{\eta-1\times\eta_j-1}$ is such that for $0\le k\le\eta-1$ and $0\le l\le\eta_j-1$

$$\mathbf{U}^{k*}_{\eta-1}\mathbf{V}^{(j)}_\eta\mathbf{U}^{l}_{\eta_j-1}=\frac{1}{l!}e^{(l)}_k(\mathrm{W}_j)=\begin{cases}\begin{pmatrix}k\\l\end{pmatrix}\mathrm{W}_j^{k-l} & \text{if } k\ge l\\ 0 & \text{if } k<l\end{cases}$$

The multiplicity of the zeros causes usual complications in notations. For each zero W_j, $0\le j\le\tau$, we need to define the vector $s^{F_0}_{n+\eta}(\mathbf{W}_j,\cdot)\in\mathcal{M}_{\eta_j-1,0}$ with the k^{th} component $\frac{1}{k!}\frac{\partial^k}{\partial\mathrm{A}^{*k}}s^{F_0}_{n+\eta}(\mathrm{A},\cdot)\,|_{\mathrm{A}=\mathrm{W}_j}$, $0\le k\le\eta_j-1$, i.e.

$$\mathbf{U}^{k*}_{\eta_j-1}s^{F_0}_{n+\eta}(\mathbf{W}_j,\cdot)=\frac{1}{k!}\frac{\partial^k}{\partial\mathrm{A}^{*k}}s^{F_0}_{n+\eta}(\mathrm{A},\cdot)\,|_{\mathrm{A}=\mathrm{W}_j}.$$

Similarly, for each couple of zeros W_j and W_m, $0 \leq j, m \leq \tau$, we define the matrix $s^{F_0}_{n+\eta}(\mathbf{W}_j, \mathbf{W}_m) \in \mathcal{M}_{\eta_j-1,\eta_m-1}$, such that for $0 \leq k \leq \eta_j - 1$ and $0 \leq l \leq \eta_m - 1$

$$\mathbf{U}^{k*}_{\eta_j-1} s^{F_0}_{n+\eta}(\mathbf{W}_j, \mathbf{W}_m) \mathbf{U}^l_{\eta_m-1} = \frac{1}{k!\, l!} \frac{\partial^k}{\partial \mathrm{A}^{*k}} \frac{\partial^l}{\partial \mathrm{Z}^l} s^{F_0}_{n+\eta}(\mathrm{A}, \mathrm{Z})\,|_{\mathrm{A}=\mathrm{W}_j, \mathrm{Z}=\mathrm{W}_m}.$$

Theorem 15 *If the transfer polynomial r of degree $\eta \in \mathbb{N}^+$, given by (37), is completely described by its multiple zeros $\mathrm{W}_0, \ldots, \mathrm{W}_\tau$, then for each $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ the kernel polynomials $s^F_n(\mathrm{A}, \cdot)$ and $s^{F_0}_{n+\eta}(\mathrm{A}, \cdot)$ are related by (66) where*

$$s^{F_0}_{n+\eta}(\mathbf{W}, \cdot) = \begin{pmatrix} s^{F_0}_{n+\eta}(\mathbf{W}_0, \cdot) \\ \vdots \\ s^{F_0}_{n+\eta}(\mathbf{W}_\tau, \cdot) \end{pmatrix}$$

and

$$s^{F_0}_{n+\eta}(\mathbf{W}, \mathbf{W}) = \begin{pmatrix} s^{F_0}_{n+\eta}(\mathbf{W}_0, \mathbf{W}_0) & \cdots & s^{F_0}_{n+\eta}(\mathbf{W}_0, \mathbf{W}_\tau) \\ \cdots & \cdots & \cdots \\ s^{F_0}_{n+\eta}(\mathbf{W}_\tau, \mathbf{W}_0) & \cdots & s^{F_0}_{n+\eta}(\mathbf{W}_\tau, \mathbf{W}_\tau) \end{pmatrix}.$$

Proof: Use the same arguments as in the course of proving theorem 14, based again on (51) with $\mathrm{V}_{n,\eta} = \mathbf{V}_{n,\eta} = (\mathbf{V}^{(0)}_{n,\eta}, \ldots, \mathbf{V}^{(\tau)}_{n,\eta}) \in \mathcal{M}_{n+\eta-1\times\eta-1}$. Here $\mathbf{V}^{(j)}_{n,\eta} \in \mathcal{M}_{n+\eta-1\times\eta_j-1}$, $0 \leq j \leq \tau$, is such that for $0 \leq k \leq n + \eta - 1$ and $0 \leq l \leq \eta_j - 1$

$$\mathbf{U}^{k*}_{\eta-1} \mathbf{V}^{(j)}_{n,\eta} \mathbf{U}^l_{\eta_j-1} = \frac{1}{l!} e^{(l)}_k(\mathrm{W}_j) = \begin{cases} \binom{k}{l} \mathrm{W}_j^{k-l} & \text{if } k \geq l \\ 0 & \text{if } k < l. \end{cases}$$

□

4.3 Product Transfer Polynomial

In this subsection we assume that the polynomial r can be decomposed as follows

$$r = (e_1 - \Theta_{\eta-1} e_0) \triangleleft \ldots \triangleleft (e_1 - \Theta_0 e_0) \tag{67}$$

with parameters Θ_j, $0 \leq j \leq \eta - 1$, belonging to $\mathcal{M}_d(\mathbb{C})$. Surely, in case $d = 1$ the parameters Θ_j are zeros of r, but in general we only know that Θ_0 is a zero of r since for any $p, q \in \mathcal{P}$ we have $q(\Theta) = 0 \Rightarrow (p \triangleleft q)(\Theta) = 0$.

Theorem 16 *Let the transfer polynomial r be given by (67). Then for each $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ and $n \in \mathbb{N}$*

$$s^F_n(\mathrm{A}, \cdot) = \hat{\mathbf{e}}_{n+\eta}(\mathrm{A})^* \Pi_{n+\eta} \hat{\mathbf{e}}_{n+\eta}(\cdot), \tag{68}$$

where $\Pi_{n+\eta}$ is given by (54), while $\hat{\mathbf{e}}_{n+\eta} \in \mathcal{M}_{n+\eta,0}$ is the vector with the first η components equal to 0 and the last $n+1$ components given by $(0 \leq k_0 \leq n)$

$$\mathbf{U}^{k_0+\eta*}_{n+\eta} \hat{\mathbf{e}}_{n+\eta}(\cdot) = \sum_{k_0 \geq \ldots \geq k_\eta \geq 0} \Theta_0^{k_0-k_1} \cdots \Theta_{\eta-1}^{k_{\eta-1}-k_\eta} e_{k_\eta}(\cdot). \tag{69}$$

Proof: The matrix $C^*_{n,2} \in \mathcal{M}_{n \times n}$ associated with r of the form (67) as in subsection 3.3, can also be decomposed:

$$C^*_{n,2} = C(\Theta_{\eta-1}) \cdots C(\Theta_0) \tag{70}$$

where for each $\Theta \in \mathcal{M}_d(\mathbb{C})$ the matrix $C(\Theta) \in \mathcal{M}_{n \times n}$ is such that for $0 \le k, l \le n$

$$U_n^{k*} C(\Theta) U_n^l = \begin{cases} I & \text{if } k = l \\ -\Theta & \text{if } k = l+1 \\ 0 & \text{otherwise.} \end{cases} \tag{71}$$

Since the inverse of the matrix $C(\Theta)$ is such that for $0 \le k, l \le n$

$$U_n^{k*} C(\Theta)^{-1} U_n^l = \begin{cases} \Theta^{k-l} & \text{if } k \ge l \\ 0 & \text{otherwise} \end{cases} \tag{72}$$

we can easily determine the inverse of $C^*_{n,2}$ and verify that for $0 \le k_0 \le n$

$$U_n^{k_0 *} (C^*_{n,2})^{-1} e_n(\cdot) = \sum_{k_0 \ge \dots \ge k_\eta \ge 0} \Theta_0^{k_0 - k_1} \cdots \Theta_{\eta-1}^{k_{\eta-1} - k_\eta} e_{k_\eta}(\cdot).$$

By using theorem 11 and (57) we then get the desired relation (69). □

4.4 Autoregression

Let F be the matrix distribution function of an autoregressive process of η^{th} order whose increments are given by

$$dF(\lambda) = \beta(e^{i\lambda} I)^{-1} \, dF_0(\lambda) \, \beta(e^{i\lambda} I)^{-1*}, \tag{73}$$

where F_0 is the same as in example 1 of subsection 2.2, and β is a polynomial of degree $\eta \in \mathbb{N}^+$ such that it has a leading coefficient that equals I and all roots of the equation

$$|\beta(zI)| = 0 \tag{74}$$

are lying inside the unit circle.

Theorem 17 *The orthonormal polynomials ϕ_n^F associated with the matrix spectral distribution function F of the form (73), are given by*

$$\phi_n = \Sigma^{-\frac{1}{2}} e_{n-\eta} \triangleleft \beta = \Sigma^{-\frac{1}{2}} \beta \, e_{n-\eta} \tag{75}$$

for each $n \ge \eta$.

Proof: For each $0 \le m \le n-1$ we have by using (2), (8) and (73)

$$< e_m; \Sigma^{-\frac{1}{2}} \beta \, e_{n-\eta} >_F = \frac{1}{2\pi} \int_{(-\pi,\pi]} e^{i(m-n)\lambda} [e^{-i\eta\lambda} \beta(e^{i\lambda} I)]^{-1} \, d\lambda \, \Sigma^{\frac{1}{2}}.$$

Thus

$$\overline{< e_m; \Sigma^{-\frac{1}{2}}\, \beta\, e_{n-\eta} >_F} = \frac{1}{i} \int\limits_{|z|=1} z^{n-1-m}[z^\eta \overline{\beta}(z^{-1}\mathrm{I})]^{-1}\, dz\, \overline{\Sigma^{\frac{1}{2}}}.$$

This last integral equals zero, since by (74) the integrand in the last integral is regular for $|z| \leq 1$. □

Corollary 5 *For each* $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ *and each* $n \in \mathbb{N}$

$$s^F_{n+\eta}(\mathrm{A},\cdot) = s^F_{\eta-1}(\mathrm{A},\cdot) + \sum_{k=0}^{n}[\beta(\mathrm{A})e_k(\mathrm{A})]^*\,\Sigma^{-1}\,\beta(\cdot)e_k(\cdot) \tag{76}$$

with $s^F_{\eta-1}(\mathrm{A},\cdot)$ *satisfying*

$$s^F_{\eta-1}(\mathrm{A},\cdot) - e_1(\mathrm{A})^* s^F_{\eta-1}(\mathrm{A},\cdot)e_1(\cdot) = \phi^{F\rho}_\eta(\mathrm{A})^*\phi^{F\rho}_\eta(\cdot) - \phi^F_\eta(\mathrm{A})^*\phi^F_\eta(\cdot). \tag{77}$$

Proof: By using (11) and (75) we get the desired result (76). Formula (77) follows from the Christoffel-Darboux formula (14). □

Note that according to (76), the difference $d_n(\mathrm{A},\cdot) := s^F_{n+\eta}(\mathrm{A},\cdot) - s^F_{\eta-1}(\mathrm{A},\cdot)$ satisfies

$$d_n(\mathrm{A},\cdot) - e_1(\mathrm{A})^* d_n(\mathrm{A},\cdot)e_1(\cdot) = \beta(\mathrm{A})^*\,\Sigma^{-1}\,\beta(\cdot) - [\beta(\mathrm{A})e_{n+1}(\mathrm{A})]^*\,\Sigma^{-1}\,\beta(\cdot)e_{n+1}(\cdot). \tag{78}$$

We can determine this difference by solving (78). For the determination of $s^F_{n+\eta}(\mathrm{A},\cdot)$ we need not only $d_n(\mathrm{A},\cdot)$ but also $s^F_{\eta-1}(\mathrm{A},\cdot)$. In the following example we will give $s^F_{\eta-1}(\mathrm{A},\cdot)$ in case $\eta = 1$.

Example: Suppose $\beta = e_1 - \Gamma e_0$. Then by (76) the kernel polynomial $s^F_{n+1}(\mathrm{A},\cdot)$ evaluated at Z equals

$$s^F_{n+1}(\mathrm{A},\mathrm{Z}) = s^F_0(\mathrm{A},\mathrm{Z}) + \sum_{k=0}^{n}[\beta(\mathrm{A})\mathrm{A}^k]^*\,\Sigma^{-1}\,\beta(\mathrm{Z})\mathrm{Z}^k,$$

where $s^F_0(\mathrm{A},\mathrm{Z})$ is a constant, namely $s^F_0(\mathrm{A},\mathrm{Z})^{-1} = \int\limits_{(-\pi,\pi]} dF(\lambda) =< e_0; e_0 >_F$, which is determined by solving the following equation

$$< e_0; e_0 >_F = \Gamma < e_0; e_0 >_F \Gamma^* + \Sigma. \tag{79}$$

4.5 First Order Moving Average

In this subsection we deal with a matrix distribution function F, whose increments are given by (33), where F_0 is arbitrary (so F_0 is not necessarily the spectral distribution function of White Noise) while r is specified as follows: $r = e_1 - \Theta e_0$ with a parameter $\Theta \in \mathcal{M}_d(\mathbb{C})$. Then we have the following

Theorem 18 *For each* $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ *and* $n \in \mathbb{N}$

$$s_n^F(\mathrm{A},\cdot) = \sum_{k=0}^{n} \psi_k^{F_0}(\mathrm{A})^* \psi_k^{F_0}(\cdot) - \sum_{k=0}^{n}\sum_{l=0}^{n} \psi_k^{F_0}(\mathrm{A})^* \phi_{k+1}^{F_0}(\Theta)[s_{n+1}^{F_0}(\Theta,\Theta)]^{-1} \phi_{l+1}^{F_0}(\Theta)^* \psi_l^{F_0}(\cdot), \quad (80)$$

where the polynomials $\{\psi_k^{F_0}\}_{k\in\mathbb{N}}$ *are defined by*

$$\phi_{k+1}^{F_0} - \phi_{k+1}^{F_0}(\Theta)\, e_0 = \psi_k^{F_0} \triangleleft r. \quad (81)$$

Note that by lemma 1 given at the end of this subsection, the polynomials $\psi_k^{F_0}$ exist and are unique.

<u>Proof:</u> In the present special case of F, formula (66) reduces to

$$r(\mathrm{A})^* \triangleright s_n^F(\mathrm{A},\cdot) \triangleleft r(\cdot) = s_{n+1}^{F_0}(\mathrm{A},\cdot) - s_{n+1}^{F_0}(\mathrm{A},\Theta)[s_{n+1}^{F_0}(\Theta,\Theta)]^{-1} s_{n+1}^{F_0}(\Theta,\cdot). \quad (82)$$

Rewrite the right-hand side of this relation as follows

$$[s_{n+1}^{F_0}(\mathrm{A},\cdot) - s_{n+1}^{F_0}(\mathrm{A},\Theta)] - [s_{n+1}^{F_0}(\Theta,\cdot) - s_{n+1}^{F_0}(\Theta,\Theta)] -$$
$$[s_{n+1}^{F_0}(\mathrm{A},\Theta) - s_{n+1}^{F_0}(\Theta,\Theta)][s_{n+1}^{F_0}(\Theta,\Theta)]^{-1}[s_{n+1}^{F_0}(\Theta,\cdot) - s_{n+1}^{F_0}(\Theta,\Theta)],$$

in order to verify by using (11) that

$$\sum_{k=0}^{n} [\phi_{k+1}^{F_0}(\mathrm{A}) - \phi_{k+1}^{F_0}(\Theta)]^* [\phi_{k+1}^{F_0}(\cdot) - \phi_{k+1}^{F_0}(\Theta)] -$$
$$\left[\sum_{k=0}^{n} [\phi_{k+1}^{F_0}(\mathrm{A}) - \phi_{k+1}^{F_0}(\Theta)]^* \phi_{k+1}^{F_0}(\Theta)\right] [s_{n+1}^{F_0}(\Theta,\Theta)]^{-1} \left[\sum_{l=0}^{n} \phi_{l+1}^{F_0}(\Theta)^* [\phi_{l+1}^{F_0}(\cdot) - \phi_{l+1}^{F_0}(\Theta)]\right].$$

Due to (81), the last expression reduces to $r(\mathrm{A})^* \triangleright \hat{s}_n^F(\mathrm{A},\cdot) \triangleleft r(\cdot)$, where $\hat{s}_n^F(\mathrm{A},\cdot)$ stands for the right-hand side of (80). Hence

$$r(\mathrm{A})^* \triangleright s_n^F(\mathrm{A},\cdot) \triangleleft r(\cdot) = r(\mathrm{A})^* \triangleright \hat{s}_n^F(\mathrm{A},\cdot) \triangleleft r(\cdot),$$

which means, in view of (35), that

$$r(\mathrm{A})^* \triangleright s_n^F(\mathrm{A},\cdot) = r(\mathrm{A})^* \triangleright \hat{s}_n^F(\mathrm{A},\cdot),$$

or equivalently

$$s_n^F(\cdot,\mathrm{A}) \triangleleft r(\mathrm{A}) = \hat{s}_n^F(\cdot,\mathrm{A}) \triangleleft r(\mathrm{A}).$$

Apply again (35) to conclude that $s_n^F(\cdot,\mathrm{A}) = \hat{s}_n^F(\cdot,\mathrm{A})$. This is the desired assertion (80). □

Lemma 1 *Let* $p \in \mathcal{P}$ *be a polynomial of degree* ≥ 1 *with* $p(\Theta) = 0$ *for some* $\Theta \in \mathcal{M}_d(\mathbb{C})$ *and let* $r = e_1 - \Theta e_0$. *Then there exists a unique polynomial* $q \in \mathcal{P}$ *such that* $p = q \triangleleft r$.

<u>Proof:</u> Let $p = \sum_{k=0}^{n} \mathrm{P}_k e_k$ and $q = \sum_{l=0}^{n-1} \mathrm{Q}_l e_l$ with $\mathrm{Q}_l = \sum_{m=0}^{n-1-l} \mathrm{P}_{m+l+1}\Theta^m$. Then it is easily verified that $p = q \triangleleft r$. By property (36) q is unique. □

4.6 ARMA(1,1) Model

In this subsection we consider a matrix distribution function F with increments given by

$$dF(\lambda) = (e^{i\lambda}\mathrm{I} - \Lambda)^{-1}(e^{i\lambda}\mathrm{I} - \Upsilon)\, dF_0(\lambda)\, (e^{i\lambda}\mathrm{I} - \Upsilon)^{*}(e^{i\lambda}\mathrm{I} - \Lambda)^{-1*}, \qquad (83)$$

where F_0 is the spectral distribution function of White Noise and $\Lambda, \Upsilon \in \mathcal{M}_d(\mathbb{C})$. This spectral distribution function F is associated with an ARMA(1,1) process if all eigenvalues of Λ are inside the unit circle. We will assume this throughout this subsection.

Theorem 19 *Suppose* $\Lambda - \Upsilon \in \mathcal{M}_d^{-}(\mathbb{C})$ *and define*

$$\Gamma = (\Lambda - \Upsilon)^{-1}\Lambda(\Lambda - \Upsilon) \text{ and } \Theta = (\Lambda - \Upsilon)^{-1}\Upsilon(\Lambda - \Upsilon). \qquad (84)$$

Then for each $\mathrm{A} \in \mathcal{M}_d(\mathbb{C})$ *and* $n \in \mathbb{N}$ *the kernel polynomial* $s_n^F(\mathrm{A}, \cdot)$ *is given by (80) with*

$$\psi_k^{F_0} = \Sigma^{-\frac{1}{2}} \left[e_k + (\Theta - \Gamma) \sum_{m=0}^{k-1} \Theta^{k-1-m} e_m \right]. \qquad (85)$$

Proof: Due to lemma 2, given at the end of this subsection, the spectral distribution function F considered here satisfies

$$dF(\lambda) = (e^{i\lambda}\mathrm{I} - \Theta)(e^{i\lambda}\mathrm{I} - \Gamma)^{-1}\, dF_0(\lambda)\, (e^{i\lambda}\mathrm{I} - \Gamma)^{-1*}(e^{i\lambda}\mathrm{I} - \Theta)^{*} \qquad (86)$$

with the same F_0 as in (83) and the parameters $\Theta, \Gamma \in \mathcal{M}_d(\mathbb{C})$ given by (84). Since (86) is presented in the form (33) with $r = e_1 - \Theta e_0$ and

$$dF_0(\lambda) = \beta(e^{i\lambda}\mathrm{I})^{-1}[\frac{1}{2\pi}\Sigma\, d\lambda]\beta(e^{i\lambda}\mathrm{I})^{-1*}, \qquad (87)$$

where $\beta = e_1 - \Gamma e_0$, we deal here with the case treated already in subsection 4.5 (with the special F_0 this time, given by (87)). Thus formula (80) holds, and it remains to verify that for the special F_0 with the increments (87), the equation (81) is satisfied with $\psi_k^{F_0}$ given by (85). Since the orthonormal polynomials $\phi_{n+1}^{F_0}$, $n \in \mathbb{N}$, are given by (75), we have

$$\begin{aligned}
\Sigma^{\frac{1}{2}}[\phi_{k+1}^{F_0} - \phi_{k+1}^{F_0}(\Theta)] &= e_{k+1} - \Gamma e_k - [e_{k+1}(\Theta) - \Gamma e_k(\Theta)] \\
&= [e_{k+1} - e_{k+1}(\Theta)] - \Gamma[e_k - e_k(\Theta)] \\
&= \left(\sum_{m=0}^{k} \Theta^{k-m} e_m - \Gamma \sum_{m=0}^{k-1} \Theta^{k-1-m} e_m \right) \lhd r \\
&= \left(e_k + (\Theta - \Gamma) \sum_{m=0}^{k-1} \Theta^{k-1-m} e_m \right) \lhd r.
\end{aligned}$$

□

Lemma 2 *Suppose* $\Lambda - \Upsilon \in \mathcal{M}_d^-(\mathbb{C})$. *Then the equations (84) are equivalent to the following two equations:*

$$\Lambda + \Theta = \Upsilon + \Gamma \text{ and } \Lambda\Theta = \Upsilon\Gamma. \tag{88}$$

Note that the eigenvalues of Γ and Λ coincide, as well as the eigenvalues of Θ and Υ.

Proof: We prove first that (88) implies (84). By (88) $\Lambda\Lambda + \Lambda\Theta = \Lambda\Upsilon + \Lambda\Gamma$. Since $\Lambda\Theta = \Upsilon\Gamma$ we get $\Lambda\Lambda + \Upsilon\Gamma = \Lambda\Upsilon + \Lambda\Gamma$, which is equivalent to $\Lambda(\Lambda - \Upsilon) = (\Lambda - \Upsilon)\Gamma$. This yields the expression (84) for Γ. The expression for Θ is obtained by using $\Lambda + \Theta = \Upsilon + \Gamma$.
We prove now the converse statement. By (84) $\Gamma - \Theta = (\Lambda - \Upsilon)^{-1}(\Lambda - \Upsilon)(\Lambda - \Upsilon) = \Lambda - \Upsilon$, so $\Lambda + \Theta = \Upsilon + \Gamma$. This yields $\Lambda\Upsilon + \Lambda\Gamma - \Lambda\Lambda = \Lambda\Theta$. Moreover the expression (84) for Γ yields $\Lambda\Upsilon + \Lambda\Gamma - \Lambda\Lambda = \Upsilon\Gamma$. Hence $\Lambda\Theta = \Upsilon\Gamma$. □

4.7 ARMA(1,1): The One Dimensional Case

The results of the present paper allow for determining the likelihood of zero mean Gaussian time series $\{X_n\}_{n\in\mathbb{Z}}$ in all particular cases treated in this section. If $d = 1$, for instance, then the log-likelihood $L_n(X_n)$ of the sample $X_0, \dots, X_n$ is (here $\mathbf{X}_n$ is the column vector with components $X_k, 0 \le k \le n$)

$$L_n(\mathbf{X}_n) = -\frac{1}{2}\left[(n+1)\log 2\pi + \log|H_n^F| + (\mathbf{X}_n;\mathbf{X}_n)_F\right], \tag{89}$$

where F is the spectral distribution function of $\{X_n\}_{n\in\mathbb{Z}}$. If, in addition, F is of the form (33), then one can apply formulas (55) and (60) in order to specify the right-hand side of (89). We will demonstrate this with the following simple

Example: Let $d = 1$. Consider the zero mean Gaussian ARMA(1,1) time series $\{X_n\}_{n\in\mathbb{Z}}$ with the spectral density

$$\frac{dF(\lambda)}{d\lambda} = \frac{\sigma^2}{2\pi}\frac{|e^{i\lambda} - \theta|^2}{|e^{i\lambda} - \gamma|^2}$$

(cf. (83) or (86) with $\gamma = \Gamma = \Lambda$, $|\gamma| < 1$, $\theta = \Theta = \Upsilon$, $\gamma \ne \theta$ and $\sigma^2 = \Sigma > 0$). We will show here that

$$L_n(\mathbf{X}_n) = -\frac{1}{2}\left[(n+1)\log(2\pi\sigma^2) - \log\hat{K}_n(\gamma,\theta) + \frac{1}{\sigma^2}\left[\sum_{k=0}^{n}|Y_k|^2 - K_n(\gamma,\theta)\left|\sum_{k=0}^{n}\bar{\theta}^k Y_k\right|^2\right]\right], \tag{90}$$

where

$$Y_k = X_k + (\theta - \gamma)\sum_{m=0}^{k-1}\theta^{k-1-m}X_m = \sum_{m=0}^{k}\theta^{k-m}X_m - \gamma\sum_{m=0}^{k-1}\theta^{k-1-m}X_m, \tag{91}$$

$$K_n(\gamma,\theta) = \frac{|\theta - \gamma|^2}{1 - |\gamma|^2 + |\theta - \gamma|^2 \sum_{k=0}^{n}|\theta|^{2k}} \tag{92}$$

and

$$\hat{\mathrm{K}}_n(\gamma,\theta) = \frac{1-|\gamma|^2}{1-|\gamma|^2+|\theta-\gamma|^2\sum_{k=0}^{n}|\theta|^{2k}}. \tag{93}$$

Obviously it suffices to show that (cf. (89) and (90))

$$\sigma^2(\mathrm{X}_n;\mathrm{X}_n)_F = \sum_{k=0}^{n}|\mathrm{Y}_k|^2 - \mathrm{K}_n(\gamma,\theta)\left|\sum_{k=0}^{n}\overline{\theta}^k\mathrm{Y}_k\right|^2, \tag{94}$$

and

$$|\mathrm{H}_n^F| = \frac{\sigma^{2n+2}}{\hat{\mathrm{K}}_n(\gamma,\theta)}. \tag{95}$$

In view of (13), one can deduce (94) by verifying first that for $\lambda,\mu\in(-\pi,\pi]$

$$s_n^F(e^{i\mu},e^{i\lambda}) = \frac{1}{\sigma^2}\left[\sum_{k=0}^{n}(\sigma\psi_k^{F_0}(e^{i\mu}))^*\sigma\psi_k^{F_0}(e^{i\lambda}) - \mathrm{K}_n(\gamma,\theta)\sum_{k=0}^{n}\sum_{l=0}^{n}(\sigma\psi_k^{F_0}(e^{i\mu}))^*\theta^k\overline{\theta}^l\sigma\psi_l^{F_0}(e^{i\lambda})\right], \tag{96}$$

with

$$\sigma\psi_k^{F_0} = e_k + (\theta-\gamma)\sum_{m=0}^{k-1}\theta^{k-1-m}e_m,$$

and then by substituting $e^{ik\lambda}$ and $e^{-ik\mu}$ in the last expression by X_k and $\overline{\mathrm{X}}_k$ respectively, for each $0\leq k\leq n$. But (96) is obtained by applying theorem 19. Indeed, put $d=1$ and

$$\frac{dF_0(\lambda)}{d\lambda} = \frac{\sigma^2}{2\pi}\frac{1}{|e^{i\lambda}-\gamma|^2}.$$

Then for all $n\in\mathbb{N}$ the corresponding orthonormal polynomials evaluated at θ are

$$\phi_{n+1}^{F_0}(\theta) = \frac{1}{\sigma}(\theta-\gamma)\theta^n$$

and

$$s_{n+1}^{F_0}(\theta,\theta) = \frac{1}{\sigma^2}\left[1-|\gamma|^2+|\theta-\gamma|^2\sum_{k=0}^{n}|\theta|^{2k}\right], \tag{97}$$

since

$$s_{n+1}^{F_0}(\theta,\theta) - s_0^{F_0}(\theta,\theta) = \frac{1}{\sigma^2}|\theta-\gamma|^2\sum_{k=0}^{n}|\theta|^{2k} \text{ and } [s_0^{F_0}(\theta,\theta)]^{-1} = \frac{\sigma^2}{1-|\gamma|^2}. \tag{98}$$

Now we will determine $|\mathrm{H}_n^F|$ by using (60) with $\mathrm{V}_n^* = (1,\theta^*,\ldots,(\theta^*)^{n+1})$. We get

$$|\mathrm{H}_n^F| = |\mathrm{H}_{n+1}^{F_0}|s_{n+1}^{F_0}(\theta,\theta),$$

which yields (95) by (97) and the following consequence of (32) and (98):

$$\frac{1}{|\mathrm{H}_{n+1}^{F_0}|} = s_0^{F_0}(\theta,\theta)\prod_{k=1}^{n+1}|\Phi_{k,k}^{F_0}|\overline{|\Phi_{k,k}^{F_0}|} = \frac{1-|\gamma|^2}{\sigma^2}\prod_{k=1}^{n+1}\frac{1}{\sigma^2} = \frac{1-|\gamma|^2}{\sigma^{2n+4}}.$$

References

[1] P. Brockwell and A. Davis, *Time Series: Theory and Methods.* Springer, 1987.

[2] K. Dzhaparidze, *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.* Springer Series in Statistics, 1985.

[3] K. Dzhaparidze, *On Constructing Kernel Polynomials of a Spectral Function: Application to ARMA Models.* CWI Report BS-R9119, Amsterdam, 1991.

[4] I. Gohberg (ed.), *Orthogonal Matrix-valued Polynomials and Applications.* Birkhaüser, 1988.

[5] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications.* University of California Press, Berkeley, 1958.

[6] E. J. Hannan, *Multiple Time Series.* Wiley, 1970.

[7] E. J. Hannan and M. Deistler, *The Statistical Theory of Linear Systems.* Wiley, 1988.

[8] Y. A. Rozanov, *Stationary Random Processes.* Holden-Day series in time series analysis, 1967.

[9] P. Whittle, *Prediction and Regulation by Linear Least-Square Methods.* Basil Blackwell, 1984.