ORTHOGONAL MATRIX FUNCTIONS RELATED TO THE MULTIVARIABLE NEVANLINNA-PICK PROBLEM

by A. BULTHEEL

Report TW 49, May 1980
Applied Mathematics and Programming Division
KATHOLIEKE UNIVERSITEIT LEUVEN (BELGIUM)

Paper presented at the Belgian Mathematical Society Conference,
22-23 May 1980, Louvain-la-Neuve

ORTHOGONAL MATRIX FUNCTIONS RELATED TO THE MULTIVARIABLE NEVANLINNA-PICK PROBLEM

A. BULTHEEL, Katholieke Universiteit Leuven

ABSTRACT

The relation between the Wiener-Masani theory of multivariable prediction and the multivariable Szegö theory of orthogonal polynomials recently studied by Delsarte et al. and Youla et al. is well understood. The classical Levinson and Schur algorithms are based on this theory and produce AR (autoregressive) filters by "incoming" resp. "outgoing" recursions. The Nevanlinna-Pick (NP) algorithm generalizes the Schur algorithm to produce ARMA filters. In this note, the corresponding incoming algorithm is studied which gives recursively reproducing kernels and orthogonal functions in a Hilbert module, thus building a resolution of submodules, constructed on the transmission zeros of the filter.

1. INTRODUCTION

A survey of results that were obtained in the author's doctoral thesis will be given, without going into the details of the proofs. This is the multivariable extension of a similar note [1] treating the scalar case. Here, an extension of Szegö's matrix polynomials, orthogonal on the unit circle [2,3] is given. We will be concerned with a Hilbert module $L^2(\omega)$ [4] of matrix valued functions, with complex variable.

Because of the non commutativity of the matrix product, most of the definitions and theorems will have a right (R-) and a left (L-) version. In most cases we will only give one of both and leave the other formulation to the devotion of the reader. So, in our Hilbert module we shall have an R-inner product, defined by

$$(f,g)_{\omega R} = \frac{1}{2\pi}\int_0^{2\pi} \tilde{f}\, d\omega\, g \qquad (1.1)$$

Here $\tilde{}$ denotes complex conjugate transpose and ω is some non negative (hermitian) measure on the Borel sets of the unit circle T. Some care has to be taken in defining (1.1) (see [5]) in order for $L^2(\omega)$ to be complete under the trace norm

$\|f\|_{tR} = \text{trace}(f,f)_{\omega R}^{1/2}$. In $L^2(\omega)$ we can speak of R-orthogonality, R-projection, etc.. The parahermitian conjugate (with respect to T) for $f \in L^2(\omega)$ is defined by $f_\star(z) = \tilde{\bar{f}}(1/\bar{z})$.

A very powerful condition imposed on ω is the Szegö condition (SC) : [6]

$$\int_0^{2\pi} \log \det \omega' \, d\theta < \infty.$$

where ω' is the Lebesgue derivative of ω. It implies e.g. that the Schmidt scheme can be used to orthogonalize a given basis for $L^2(\omega)$. It also implies the existence of R (and L) spectral factors of ω', i.e. $\omega' = \sigma_\star^R \sigma^R$ on T and $\sigma^R \in H^2$ is outer [6,7]. (H^2 are the matrix functions, analytic in the unit disc and square integrable on T w.r.t. the Lebesgue measure).

2. THE INTERPOLATION SPACES L_n

We shall define in this section a nested set of submodules associated with complex numbers $\alpha_1, \alpha_2, \ldots$, all in the unit disc D. First of all define for $A_n = \{\alpha_k\}_0^n \subset D$ the Blaschke factors

$$\zeta_k(z) = \frac{\alpha_k - z}{1 - \bar{\alpha}_k z} \frac{|\alpha_k|}{\alpha_k} \quad \text{with} \quad \frac{|\alpha_k|}{\alpha_k} = 1 \quad \text{if} \quad \alpha_k = 0.$$

From this, partial Blaschke products $U_n(z) = \zeta_1(z)\zeta_2(z) \ldots \zeta_n(z)$, $n \geq 1$ and $U_0(z)=1$ can be constructed and finally we define $L_n = \text{span}\{U_k \, I_N\}_0^n$ where linear combinations have to be taken over K, the set of constant N×N matrices and I_N is the identity of K. Remark that if $\alpha_k = 0$, $\forall k$, L_n will be the set of matrix polynomials of degree at most n. Suppose

$$A_n = \{\underbrace{\beta_1, \ldots, \beta_1}_{\nu_1}, \underbrace{\beta_2, \ldots, \beta_2}_{\nu_2}, \ldots, \underbrace{\beta_m, \ldots, \beta_m}_{\nu_m}\}$$

with $\nu_1 \geq 0$, $\nu_i \geq 1$ for $i > 1$, $\Sigma_1^m \nu_i = n$ and $\beta_1 = 0$, then $\{W_k\}_0^n = \{w_k I_N\}_0^n$ will be a basis for L_n if we take $\{w_k\}_0^n = \{1, z, \ldots, z^{\nu_1}, (1-\bar{\beta}_2 z)^{-1}, \ldots (1-\bar{\beta}_2 z)^{-\nu_2}, \ldots, (1-\bar{\beta}_m z)^{-1} \ldots (1-\bar{\beta}_m z)^{-\nu_m}\}$.

The Gram matrix $P_n = [(W_i,W_j)_{\omega R}]_{i,j=0}^{n}$ only depends on $(d/dz)^r(\Omega(z))_{z=\beta_s}$, $s = 1,\ldots,m$; $r = 0,1,\ldots,\nu_s-1$ where

$$\Omega(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\omega(\theta). \tag{2.1}$$

P_n is called a block Pick matrix.

If the Schmidt scheme is applied, we obtain a set of R-orthogonal functions $\{\phi_k^R\}_0^n$. For $f \in L_n$ we define its reciprocal as $f^\star(z) = f_\star(z)U_n(z)$ and its leading coefficient as $f^\star(\alpha_n)$. The symbol κ_n^R is reserved for the leading coefficient of ϕ_n^R, $n \geqslant 0$. Remark that ϕ_n^R (and also κ_n^R) is only defined up to a constant right unitary factor.

3. R-KERNELS AND OPTIMIZATION PROBLEMS

An R-(reproducing) kernel $k_n^R(x,y)$ for L_n is a matrix function of two complex variables such that

(1) $k_n^R(x,y) \in L_n$ for constant x

(2) $(k_n^R(x,.),f(.))_{\omega R} = f(x)$, $\forall f \in L_n$.

An excellent book on reproducing kernels is [8], from which the following properties are taken in a matrix adaptation.

Theorem 3.1

(1) $k_n^R(x,y) = \tilde{\tilde{k}}_n^R(y,x)$

(2) $k_n^R(x,y) = \Sigma_0^n \phi_k^R(y)\tilde{\tilde{\phi}}_k^R(x)$ where $\{\phi_k^R\}_0^n$ is the orthonormal basis for L_n.

(3) The R-projection of $h \in L^2(\omega)$ onto L_n is given by $(h/L_n)_{\omega R} = (k_n^R(x,.),h(.))_{\omega R}$.

(4) $0 < k_i^R(x,x) \leqslant k_j^R(x,x)$ for $0 \leqslant i \leqslant j$

Here for f and $g \in K$ and hermitian, $f > (\geqslant)\ g$ means that f-g is positive (semi) definite. It is known that in the scalar case, the reproducing kernel solves some optimization problems. Let us first introduce the following convention :

For Φ, a map from W, the set of p×q matrices into K, we say that min $\Phi(f)$ has a solution F in $V \subset W$ and a minimum $m \in K$ if $m = \Phi(F) \leqslant \Phi(f)$, $\forall f \in V$ and this is denoted as $\delta[\min_{f\in V\subset W} \Phi(f)] = \{F,m\}$. If the solution is unique, we replace δ by $\delta!$.

For a complex number α, define the sets $V_1(\alpha) = \{f \in L_n \,|\, f(\alpha) = I_N\}$ and $V_2^R = \{f \in L_n \,|\, (f,f)_{\omega R} = I_N\}$ and the related optimization problems

$$[P_1^R(\alpha)]_n : \min_{f \in V_1(\alpha)} (f,f)_{\omega R} \quad \text{and} \quad [P_2^R(\alpha)]_n : \max_{f \in V_2^R} f(\alpha)\tilde{\tilde{f}}(\alpha)$$

then, using the matrix version of the Cauchy-Schwarz inequality, the following solutions are found as in [8]

Theorem 3.2

$\delta![P_1^R(\alpha)]_n = \{k_n^R(\alpha,y)k_n^R(\alpha,\alpha)^{-1}, k_n^R(\alpha,\alpha)^{-1}\}$ (this is the R-projection of 0 onto $V_1(\alpha)$) and

$\delta[P_2^R(\alpha)]_n = \{k_n^R(\alpha,y)k_n^R(\alpha,\alpha)^{-L/2}, k_n^R(\alpha,\alpha)\}$

where $f^{-L/2}$ is an arbitrary left square root of $f^{-1} \in K$, i.e. any $N \times N$ matrix $f^{-L/2}$ such that $f^{-L/2}\,\tilde{\tilde{f}}^{-L/2} = f^{-1}$. It is only defined up to a right unitary factor, so that the above solution exhausts the whole solution set.

□

It is only a slight extension to see that for $\alpha \in \mathbb{C}$ and $\mu \in K$, the solution of the following problem :

$$[P_1^R(\alpha,\mu)]_n : \min_{f \in \mu V_1(\alpha)} (f,f)_{\omega R} \text{ with } \mu V_1(\alpha) = \{f \in L_n \,|\, f(\alpha) = \mu\}$$

is given by

Corollary 3.3

$\delta![P_1^R(\alpha,\mu)]_n = \{k_n^R(\alpha,y)k^R(\alpha,\alpha)^{-1}\ \mu, \tilde{\tilde{\mu}}\ k^R(\alpha,\alpha)^{-1}\ \mu\}$

□

Essential is that a reciprocal problem may be solved in L_n. Suppose $\mu V_1^\star(\alpha) = \{f^\star \,|\, f \in \mu V_1(\alpha)\}$ and associate with this set the optimization problem

$$[P_1^{R\star}(\alpha,\mu)]_n : \min_{f \in \mu V_1^\star(\alpha)} (f,f)_{\omega R}$$

then its solution may be expressed with the L-kernel :

Theorem 3.4

$$\delta![P_1^{R\star}(\alpha,\mu)]_n = \{k_n^{L\star}(x,\alpha)k_n^L(\alpha,\alpha)^{-1} \tilde{\mu},\ \mu\, k_n^L(\alpha,\alpha)^{-1} \tilde{\mu}\}$$

Indication of the proof : use $\min_{f \in \mu V_1(\alpha)}{}^{\star} (f,f)_{\omega L} = \min_{f^\star \in \mu V_1(\alpha)} (f^\star,f^\star)_{\omega R}$

and the L-version of corollary 3.3.

□

Some other projection results may be given. Therefore introduce for $\alpha \in D$:

$L_n(\alpha) = \{f \in L_n \mid f(\alpha) = 0\}$ and $K_n(\alpha) = (z-\alpha)^{-1} L_n(\alpha)$.

Theorem 3.5

(1) The R-projection of I_N onto $K_n(\alpha)$ is given by :

$$\delta![P_3^R(\alpha)]_n = \{I_N - k_n^R(\alpha,y)k_n^R(\alpha,\alpha)^{-1},\ k_n^R(\alpha,\alpha)^{-1}\}$$

with

$$[P_3^R(\alpha)]_n : \min_{f \in K_n(\alpha)} (I_N-f, I_N-f)_{\omega R}$$

(2) The least squares approximation of the R-spectral factor σ^R of ω' in L_n results from the solution of

$$[P_4^R(x)]_n : \min_{f \in L_n} (s_x-f, s_x-f)_{\omega R} \quad \text{where}$$

$$s_x(y) = \sigma^R(y)^{-1} \tilde{\sigma}^R(x)^{-1}/(1-\bar{x}y) \ (\in H^2) \quad \text{for } x \in D,$$

which is given by :

$$\delta![P_4^R(x)]_n = \{k_n^R(x,y),\ \frac{(\sigma^R(x)\sigma^R(x))^{-1}}{1-|x|^2} - k_n^R(x,x)\}$$

proof

(1) From the definition of $K_n(\alpha)$ and $V_1(\alpha)$, it is clear that

$\min_{f \in K_n(\alpha)} (I_N-f, I_N-f)_{\omega R} = \min_{g \in V_1(\alpha)} (g,g)_{\omega R}$, and the latter problem is solved by theorem 3.2.

(2) Here we remark that

$(s_x-f, s_x-f)_{\omega R} = (s_x,s_x)_{\omega R} + (f,f)_{\omega R} - 2\mathrm{Herm}(s_x,f)_{\omega R}$ with $(s_x,f)_{\omega R} = f(x)$ and $(s_x,s_x)_{\omega R} = (\tilde{\sigma}^R(x)\sigma^R(x))^{-1}/(1-|x|^2)$.

But for $0 < \mu \in K$, it is easy to verify that

$\delta![\min_{f \in K} (\tilde{\bar{f}}\, \mu^{-1} f - 2\mathrm{Herm}\, f)] = \{\mu, -\mu\}$, which completes the proof.

□

With the aid of these optimization results, it requires little effort to adapt the proof for the polynomial case [9, p. 290] to the result of lemma 3.6 (1) below, relating L- and R-kernels. This in turn implies the other relations given.

Lemma 3.6

(1) for $\alpha \in \mathbb{C}$: $k_n^R(\alpha,z) = \tilde{U}_n(\alpha) U_n(z) k_n^L(1/\bar{\alpha}, 1/\bar{z})$

(2) $k_n^R(\alpha_n, z) = \phi_n^{L\star}(z) \kappa_n^L$

(3) $\phi_n^R(x)\phi_n^{R\star}(x) = \phi_n^{L\star}(x)\phi_n^L(x)$

□

4. CHRISTOFFEL-DARBOUX RELATIONS AND RECURSIONS

We have now presented all the material to adapt the proof in [9, p. 293] to obtain the following results.

Theorem 4.1

(1) For all $n \geqslant 1$ we have

$$k_n^R(x,y) = \frac{\phi_{n+1}^L{}^{\star}(y)\tilde{\phi}_{n+1}^{L}{}^{\star}(x) - \phi_{n+1}^R(y)\tilde{\phi}_{n+1}^R(x)}{1 - \zeta_{n+1}(y)\bar{\zeta}_{n+1}(x)}$$

$$= \frac{\phi_n^{L\star}(y)\tilde{\phi}_n^{L\star}(x) - \zeta_n(y)\bar{\zeta}_n(x)\phi_n^R(y)\tilde{\phi}_n^R(x)}{1 - \zeta_n(y)\bar{\zeta}_n(x)}$$

which are Christoffel-Darboux relations.

(2)
$$\begin{bmatrix} I_N & -\tilde{r}_{n+1}^R(x) r_{n+1}^L(x) & 0 \\ 0 & I_N - r_{n+1}^L(x)\tilde{r}_{n+1}^R(x) \end{bmatrix} \begin{bmatrix} k_{n+1}^{R\star}(x,y) \\ k_{n+1}^L(y,x) \end{bmatrix} =$$

$$\begin{bmatrix} I_N & \tilde{r}_{n+1}^R(x) \\ r_{n+1}^L(x) & I_N \end{bmatrix} \begin{bmatrix} \zeta_{n+1}(y) I_N & 0 \\ 0 & I_N \end{bmatrix} \begin{bmatrix} I_N & -\bar{\zeta}_{n+1}(x)\tilde{r}_{n+1}^R(x) \\ -\zeta_{n+1}(x) r_{n+1}^L(x) & I_N \end{bmatrix} \begin{bmatrix} k_n^{R\star}(x,y) \\ k_n^L(y,x) \end{bmatrix}$$

with $r^R_{n+1} = [\phi^{R\star}_{n+1}(x)]^{-1}\ \tilde{\phi}^R_{n+1}(x)$; $r^L_{n+1} = \tilde{\phi}^L_{n+1}(x)[\phi^{L\star}_{n+1}(x)]^{-1}$

□

The next step is to normalize these recursions to obtain more symmetry. r^L_n and r^R_n are normalized to become generalizations of the reflection coefficients that give information on the stability (zerolocation) of the basisfunctions. To this end we introduce first some definitions and further notation.

5. CONTRACTIONS AND UNITARY MATRICES

For a square matrix G, we define the following sets of square matrices : $\mathcal{T}_G(\mathcal{D}_G,\bar{\mathcal{D}}_G) = \{\theta \mid \theta\ G\ \tilde{\tilde{\theta}} = (<,\leq)G\}$, which are called G-unitary (strictly contractive, contractive).

Two possibilities for G will be considered :

(1) G = I, and in that case we drop the indicator G

(2) $G = J \stackrel{\text{def}}{=} \begin{bmatrix} -I_N & 0 \\ 0 & -I_N \end{bmatrix}$.

For function matrices involved in this paper one may replace $\tilde{\tilde{\theta}}$ by $\theta_\star$ and then we write $\star$ - $\mathcal{T}_G$ instead of $\mathcal{T}_G$ etc. .

The most general form of $\theta \in \mathcal{T}_J$ is

$$\begin{bmatrix} (I_N-\rho\tilde{\tilde{\rho}})^{-1/2} & 0 \\ 0 & u_3(I_N-\tilde{\tilde{\rho}}\rho)^{-1/2} \end{bmatrix} \begin{bmatrix} I_N & \rho \\ \rho & I_N \end{bmatrix} \begin{bmatrix} u_1 & 0 \\ 0 & u_2 \end{bmatrix} \tag{5.1}$$

with $u_i \in \mathcal{T}$, i = 1,2,3, $\rho \in \mathcal{D}$ and 1/2 referring to the hermitian square root. For $u_1 = u_3 = I_N$ and $u_2 = -I_N$ we will refer to (5.1) as θ_ρ. Another example :

$$b_\alpha(z) = \begin{bmatrix} \frac{\alpha-z}{1-\bar{\alpha}z}\ \frac{|\alpha|}{\alpha} I_N & 0 \\ 0 & I_N \end{bmatrix} \in \bar{\mathcal{D}}_J(\mathcal{D}_J,\mathcal{T}_J) \iff \alpha \in \bar{D}(D,T) \tag{5.2}$$

or in extended sense $b_\alpha(z) \in \star$ - $\mathcal{T}_J$. On J-unitary matrices see e.g. [10].

With $\theta_{2N\times 2N} \in \bar{\mathcal{D}}_J$, we associate the homographic transform of $x \in K$: $B^R_\theta(x) = (\theta_{11}x+\theta_{12})(\theta_{21}x+\theta_{22})^{-1}$, which is contractive in the sense that $B^R_\theta(\bar{\mathcal{D}}) \subset (=)\ \bar{\mathcal{D}}$ if $\theta \in \bar{\mathcal{D}}_J(\mathcal{T}_J)$. These transforms are the working instrument of the algorithms of Schur

and Nevanlinna-Pick [11-13] to study the Schur class [14] of matrix functions, i.e. the set $\bar{S} = \{\Gamma(z) \mid \Gamma(z) \in \bar{\mathcal{D}}$ and analytic $\forall z \in D\}$. Roughly speaking, the Cayley transform $c(f) = (I_N - f)(I_N + f)^{-1}$ makes the study of Schur type functions identical to the study of the class $\bar{C} = \{\Omega(z) \mid \Omega(z)$ is analytic and $\mathrm{Herm}\Omega(z) \geqslant 0 \ \forall z \in D\}$ which is called Carathéodory's class [14].

It is a result of Riesz-Herglotz that $\bar{C}$ type functions can be represented in a form like (2.1).

It is computationally easier to introduce R-homogeneous coordinates for matrices [15]. If Δ is $2N \times N$, then it is clear that $x = d^R(\Delta) \overset{\text{def}}{=} \Delta_1 \Delta_2^{-1} \in \bar{S}$ if we take $A \in \bar{A}^R = \{\Delta \mid \Delta$ is analytic in D and $d^R(\Delta) \in \bar{S}\}$. We call this the set of R-acceptable matrices [16]. The contraction property of the homographic transforms then translates into

$$\theta \bar{A}^R \subset (=) \bar{A}^R \quad \text{if} \quad \theta \in \bar{\mathcal{D}}_J(T_J).$$

6. J-UNITARY RECURSIONS AND THE NP ALGORITHM

We shall give in this section a normalized version of the previously given recursion and also outline the NP algorithm in homogeneous coordinates which will both give J-unitary recursions. In the next section we shall investigate the structure of these J-unitary matrices.

Theorem 6.1

Define the normalized kernels : $\overset{\circ}{k}{}^L_n(x,y) = (k^L_n(x,x)^{L/2})^{-1} k^L_n(x,y)$ and $\overset{\circ}{k}{}^R_n(x,y) = k^R_n(x,y)(k^R_n(x,x)^{R/2})^{-1}$ where L/2 and R/2 indicate L- resp. R-square roots, then it is possible to choose them such that the following recursion exists :

$$\begin{bmatrix} \overset{\circ}{k}{}^R_{n+1}(x,y) \cdot \\ \\ \overset{\circ}{k}{}^L_{n+1}(y,x) \end{bmatrix} = \theta_n \begin{bmatrix} \overset{\circ}{k}{}^R_n(x,y) \\ \\ \overset{\circ}{k}{}^L_n(y,x) \end{bmatrix} = \underbrace{\theta_n \, \theta_{n-1} \cdots \theta_0}_{\Theta_{n+1}} \begin{bmatrix} I_N \\ \\ I_N \end{bmatrix}$$

with $\theta_n = \theta_{-\tilde{\rho}_n} b_{\alpha_{n+1}}(y) \, \theta_{-\bar{\zeta}_{n+1}(x)\tilde{\rho}_n} \in \star - T_J$

where θ_γ and b_α were defined in section 5, and

$$-\rho_n = (k_n^L(x,x)^{L/2})^{-1} \tilde{\phi}_{n+1}^L(x)(\phi_{n+1}^{L\star}(x))^{-1} k_n^R(x,x)^{L/2}$$

$$= (k_n^L(x,x)^{L/2})^{\simeq} \phi_{n+1}^R(x)(\tilde{\phi}_{n+1}^{R\star}(x))^{-1}((k_n^R(x,x)^{L/2})^{-1})^{\simeq} \in \mathcal{D}$$

□

The proof involves elaborate, but straightforward computations. If $\alpha_1 = \alpha_2 = \ldots = 0$ and if one sets $x = 0$ in the previous recursion, then it follows from lemma 3.6 (2) that theorem 6.1 gives the well known recursion for the Szegö orthogonal polynomials.

This recursion is also related to the NP problem which has a long history, but the multivariable case has only been approached recently [16-18]. The NP algorithm may be used e.g. to check if a given function matrix Γ_0 is of Schur type. Omitting details, one can say that the algorithm constructs recursively the set of functions $\{\Gamma_k\}_1^\infty$ related to a chosen set $\{\alpha_k\}_1^\infty$ by successive homographic transforms described in section 5, and extracting a Blaschke factor ζ_k after each transform. If all the $\gamma_k = \Gamma_k(\alpha_{k+1})$ are in $\mathcal{D}$, then Γ_0 is a Schur function. Meanwhile, some rational approximation problems are solved as will be explained in a next section. In the special case that $\alpha_k = 0$, $\forall k$, the NP algorithm coincides with the Schur algorithm [19]. A homogeneous version has a general step that can be formulated as follows : If $d^L(\Delta_n) \overset{\text{def}}{=} \Delta_{2n}^{-1}\Delta_{1n} = \Gamma_n$, then Δ_n is an L-acceptable matrix and the homographic transform followed by a ζ_{n+1} extraction and a normalization at a fixed point $x \in D$ given by $\Delta_{n+1} = \Delta_n \theta_n^{-1}$ with $\theta_n^{-1} = \theta_{-\tilde{\gamma}_n} b_{\alpha_{n+1}}^{-1} \theta_{-\tilde{\rho}_n}$ where $\gamma_n = \Gamma_n(\alpha_{n+1})$ and $\rho_n = g_n(x)$ with $g_n = d^L(\Delta_n \theta_{-\tilde{\gamma}_n} b_{\alpha_{n+1}}^{-1})$.

Remark that the θ_n are again $\star$ - J matrices. For more details one should consult the literature.

7. THE STRUCTURE OF THE Θ MATRICES

The θ matrices appearing in the J-unitary recursion of theorem 6.1 and in the NP algorithm one of a similar form. In this section we give some properties related to their structure.

Define $\Theta_0^L = I_{2N}$ and for $n \geqslant 1$ $\Theta_n^L = \theta_{n-1} \ldots \theta_1 \theta_0$ with θ_k of the form

$\theta_{-\tilde{\rho}_k} b_{\alpha_{k+1}} \theta_{-\tilde{\gamma}_k}$ where as above $b_{\alpha_{k+1}}$ depends on y and ρ_k and γ_k depend on x. We consider x fixed and the blocks of Θ_n^L are seen as function matrices in y. Using this notation one can check easily that we have

Theorem 7.1

Θ_n^L may be written in the form

$$\frac{1}{2}\begin{bmatrix} (h_n+k_n)^\star & (h_n-k_n)^\star \\ (f_n-g_n) & (f_n+g_n) \end{bmatrix}$$

with f_n, g_n, h_n and k_n all in L_n. Moreover f_n, h_n, $f_n + g_n$ and $h_n + k_n$ have no zeros in D. By the choice $\gamma_k = \zeta_{k+1}(x)\rho_k$ (Denjoy normalization) is $\Theta_n^L(x,x)$ made upper block triangular and consequently $f_n(x,x) = g_n(x,x)$, $h_n(x,x) = k_n(x,x)$.

□

With the aid of f_n and g_n we may construct a matrix valued measure on T as follows : Let $\Omega_n(y) = f_n^{-1}(x,y)g_n(x,y)$ for x fixed, then $\Omega_n(x) = I_n$ and $\Omega_n(y) \in \bar{C}$. Now $\omega_n' = \frac{1}{2}[\Omega_n+\Omega_{n\star}] = f_n^{-1} f_{n\star}^{-1}$ which is a consequence of the $\star$ - T_J structure, so that ω_n with Lebesgue derivative ω_n' is a positive (hermitian) measure on T. The choice of normalization in $x \in D$ is arbitrary, and it is always possible to bring x to the origin by a simple bilinear transform which doesn't change the properties of the function matrices which are in order. Therefore we suppose from now on that $x = 0$ and write $f_n(y)$ for $f_n(0,y)$ etc..

8. SECOND KIND FUNCTIONS

The concept of Szegõ polynomials of the second kind may be generalized to the interpolation case as follows. For given $\Omega \in \bar{C}$ and $f_n \in L_n$, define $f_n^\#$ as the R-function of the second kind of f_n (w.r.t. Ω) as the R projection of $f_n\Omega$ on L_n in L^2 (L^2 without argument refers to the Lebesgue measure). Remark that $F_n = f_n\Omega$ is analytic in $\bar{D}$. It is a result of Walsh [20, p. 225] that the L_2-R-projection of $f \in H^2$ onto L_n is given by the unique $f_n \in L_n$ which interpolates f in the set $\{0,\alpha_1,\ldots,\alpha_n\} \subset D$. Using this result and the $\star$ - T_J property of the NP θ matrix, we obtain without

great difficulties

Theorem 8.1

Let Θ_n^L be constructed by the NP algorithm, starting from $\Gamma_0 = d^L([I-\Omega\ I+\Omega])$ with Ω, a Carathéodory function such that $\Omega(0) = I_N$, and f_n and g_n defined as in theorem 7.1 (with x = 0), then $g_n = f_n^{\#}$.

□

Remark that $\Omega_n = f_n^{-1} f_n^{\#}$ is a rational matrix function, solving the linearised interpolation condition $f_n^{\#} - f_n\Omega = 2U_n \Delta_{1n}$, with Δ_{1n} analytic in D and vanishing at the origin. Constructing ω_n, a positive measure on T from Ω_n, as done at the end of section 7, we obtain

Corollary 8.2

With the above notation we have

(1) $\Omega_n(z) = \Omega(z)$ for $z \in \{0,\alpha_1,\ldots,\alpha_n\}$ (with adaptation for confluent points).

(2) $(f,g)_{\omega_n R} = (f,g)_{\omega R}$, $\forall f,g \in L_n$

(3) $\omega' - \omega_n' = -2[U_n f_n^{-1} \Delta_{1n} + U_{n\star} \Delta_{1n\star} f_{n\star}^{-1}]$

(primes denote Lebesgue derivatives).

□

9. CORRESPONDENCE BETWEEN NP ALGORITHM AND KERNEL RECURSION

We are now ready to give an identification of the θ matrix generated by the kernel recursion and the θ matrix of the NP algorithm. Although they have the same structure, their parameters were generated in a different way, and it is not a priori clear that they represent the same matrices, as will be explained in this section.

Using corollary 8.2 (2), it is possible to prove that f_n is L-orthogonal in $L^2(\omega)$ to $K_n(0)$, (section 3), since we may replace ω by ω_n. This proves that the L-orthogonal projection of I_N on $K_n(0)$ is given by $(I_N/K_n(0))_{\omega L} = I_N - f_n^{-1}(0) f_n(z)$. The same approximation problem was solved in theorem 3.5 in an R-version. Thus we conclude that

Theorem 9.1

With the above notations : $k_n^L(z,0) = \tilde{f}_n(0) f_n(z)$.

□

Remark that $f_n(0)$ is a R-square root of $k_n^L(0,0)$, so that $k_n^L(z,0) = f_n(z)$ modulo a left unitary factor.

Using theorem 6.1 and section 7, it follows that we may define $\ell_n^{\circ L/2}$ by

$$\Theta_n = \frac{1}{2}\begin{bmatrix} (k_n^{\circ R}(x,y)+\ell_n^{\circ R}(x,y))^\star & (k_n^{\circ R}(x,y)-\ell_n^{\circ R}(x,y))^\star \\ (k_n^{\circ L}(y,x)-\ell_n^{\circ L}(y,x)) & (k_n^{\circ L}(y,x)+\ell_n^{\circ L}(y,x)) \end{bmatrix}$$

where Θ_n is the one from theorem 6.1. Because $(k_{n\star}^{\circ L}(z,0)k_n^{\circ L}(z,0))^{-1} = (f_{n\star}(z)f_n(z))^{-1}$ $= \omega_n'$ by the previous theorem, one can also identify the corresponding Ω_n, given by $(k_n^{\circ L})^{-1}\ \ell_n^{\circ L}$ and $f_n^{-1}\ g_n$ respectively. Thus we have

Theorem 9.2

With the notations above : $\ell_n^{\circ L}(z,0) = g_n(z)$ modulo a constant left unitary factor and thus $(k_n^{\circ L}(z,0))^{\#} = \ell_n^{\circ L}(z,0)$.

□

It can be proved that $\ell_n^L(y,x)$ is an L-reproducing kernel for $L^2(w)$ where w is a non-negative measure of T, which has for its associated Carathéodory function $W = \Omega^{-1}$.

At this point we have identified the lower block row of the Θ_n matrix from the normalized recursion and the corresponding row of the Θ_n^L matrix emanating from the NP algorithm. The other row may then also be identified because of the $\star$ - T_J structure of the matrices.

10. CONVERGENCE

If $L_+^2(\omega)$ are the $L^2(\omega)$ matrix functions with vanishing "negative" fouriercoefficients, then, as in the polynomial case, it is crucial to know if the set $\{U_k I_N\}_0^\infty$ is a complete basis in $L_+^2(\omega)$, i.e. if $L = \text{span}\{U_k I_N\}_0^\infty$ (where closure must be taken in the trace norm) includes all of $L_+^2(\omega)$. In the scalar case, it has been shown [21] that for absolute continuous ω, uniform convergence of U_n to zero on compact results of D (which is equivalent to the divergence of the Blaschke product U_∞ or the divergence of $\Sigma(1-|\alpha_u|)$ [20]) is sufficient for the completeness of $\{U_k I_N\}_0^\infty$. Remember also that the Szegö condition (SC) was supposed to hold.

$L^2(\omega)$ convergence of $k_n^R(x,y)$ to the Szegö kernel [9,p. 303] and of $k_n^R(0,y)$ to the inverse of the R-spectral factor is readily obtained because of the projection properties proved on L_n. Remark also that this type of convergence implies under SC that we have uniform convergence on compact subsets of the unit disc.

11. CONCLUSION

It is a well known fact that the Szegö orthonormal polynomials solve the autoregressive (AR) least squares linear prediction problem of discrete time, second order, N-dimensional, stationary stochastic processes [22].
However, for the ARMA filtering problem (autoregressive-moving average) rational functions are involved and the problem becomes nonlinear. Thus it can only be solved in an iterative way. However by fixing the transmission zeros $\{\alpha_k\}$, we obtain a linear problem again. So, given a spectral distribution function ω of the process, it is possible to obtain estimates of the spectral factor in a least squares sense : either by building up the reproducing kernels with the incoming algorithm, extending the Levinson filter, or by extracting continued fraction-like information from the data. These informations are the generalized reflection coefficients ρ_k which should be in $\mathcal{D}$ to ensure filter stability. They are sufficient to realize this extracting or outgoing algorithm in a ladder- or wave digital filter, [16]. Even if only a vague idea about the location of the transmission zeros may be obtained from the spectral or other information, this could result a better estimate by ARMA filtering than by AR filtering. If one wants to obtain the optimal choice of $\{\alpha_k\}$, then $k_n^R(0,0)$ has to be maximized w.r.t. these parameters. $k_n^R(0,0)$ is recursively defined by $k_{n+1}^R(0,0) = k_n^R(0,0)(I_N-\rho_n\tilde{\rho}_n)^{-1}(I_N-\tilde{\rho}_n\rho_n|\zeta_{n+1}|^2)$. $\{\alpha_k\}_1^{n+1}$ however have to be chosen simultaneously because introducing only one free α_n with each n would yield only a suboptimal solution.

12. REFERENCES

[1] A. BULTHEEL, P. DEWILDE, Orthogonal functions related to the Nevalinna-Pick problem, Proceedings MTNS conference, Delft, 1979, P. Dewilde (ed.), Western Periodicals, 207-212.

[2] P. DELSARTE, Y. GENIN, Y. KAMP, Orthogonal polynomial matrices on the unit circle, IEEE Trans. CAS-25 (1978) 149-160.

[3] D.C. YOULA, N.N. KAZANJIAN, Bauer-type factorization of positive matrices and the theory of matrix polynomials, orthogonal on the unit circle, IEEE Trans. CAS-25 (1978) 57-69.

[4] R. DELANGHE, F. BACKX, Hypercomplex function theory and Hilbert modules with reproducing kernel, Proc. London Math. Soc. 37 (1978) 545-576.

[5] M. ROSENBERG, The square integrability of matrix valued functions with respect to a non-negative hermitian measure, Duke Math. J. 31 (1964) 291-298.

[6] H. HELSON, Lectures on invariant subspaces, Ac. Press, N.Y., 1964.

[7] M. ROSENBLUM, J. ROVNYAK, The factorization problem for non-negative valued functions, Bull. Am. Math. Soc. 77 (1979) 287-318.

[8] H. MESCHKOWSKI, Hilbertsche Räume mit Kernfunktion, Springer Verlag, Berlin, 1962.

[9] G. SZEGÖ, Orthogonal polynomials, AMS Colloquium publ. 23, AMS Providence, Rhode Island, 1939.

[10] V.P. POTAPOV, The multiplicative structure of J-unitary matrix functions, AMS translat. ser. 2, 15 (1960) 131-243.

[11] J. SCHUR, Ueber Potenzreihen die im Innern des Einheitskreises beschränkt sind, Z. Reine Angew. Math. 147 (1917) 205-232, 148 (1918) 122-145.

[12] R. NEVANLINNA, Ueber beschränkte analytische Funktionen, Ann. Sci. Fenn. ser. A 32 (1929) 1-75.

[13] G. PICK, Ueber beschränkte Funktionen mit vorgegeben Wertzuordnungen, Ann. Acad. Sci., Fenn. Ser. A 15 (1920).

[14] N.I. AKHIEZER, The classical moment problem. Oliver and Boyd, Edinburgh, 1965.

[15] LOO-KENG HUA, Geometries of symmetric matrices I, II, III, Trans AMS 57 (1945) 441, 61 (1947) 193-255.

[16] P. DEWILDE, A. VIEIRA, T. KAILATH, On a generalized Szegö-Levinson realization algorithm for optimal linear predictors based on a network synthesis approach, IEEE Trans CAS-25 (1978) 663-675.

[17] P. DELSARTE, Y. GENIN, Y. KAMP, The Nevanlinna-Pick problem for matrix valued functions, SIAM J. Appl. Math. 36 (1979) 47-61.

[18] I.P. FEDČINA, A criterion for the solvability of the Nevanlinna-Pick tangent problem, Mat. Issled 7 (1972) 213-227, MR 50 # 4960.

-, A description of the solutions of the Nevanlinna-Pick tangent problem, Akad. Nauk. Armjan SSR Dokl. 60 (1975) 37-42, MR52 # 5974.

-, The tangential Nevanlinna-Pick problem with multiple points, Akad. Nauk. Armjan SSR Dokl. 61 (1975) 214-218, MR 53 # 13576.

[19] P. DELSARTE, Y. GENIN, Y. KAMP, Schur parametrization of positive definite block Toeplitz systems, SIAM J. Math. 36 (1979) 34-46.

[20] J.L. WALSH, Interpolation and approximation by rational functions in the complex domain, AMS Colloquium publ. 20, AMS, Providence, Rhode Island, 1960.

[21] P. DEWILDE, H. DYM, Schur recursions, error formulas and convergence of rational estimators for stationary stochastic sequences, T.H. Delft, rept. 92, 1979.

[22] N. WIENER, P. MASANI, The prediction theory of multivariable stochastic processes, I The regularity condition, Acta Math. 98 (1957) 111-150, II The linear predictor, Acta Math. 99 (1958) 93-138.