ON A SPECIAL LAURENT-HERMITE INTERPOLATION PROBLEM

Adhemar Bultheel

We present a recursive algorithm for the construction of a rational approximation for a given Laurent series which in a certain sense interpolates at the zeros of its numerator. If certain symmetry conditions are satisfied, the algorithms of Nevanlinna-Pick and Schur are found as special cases. We give also an interpolation of the algorithm as a coupled recursion for reproducing kernels of indefinite inner product spaces defined with the aid of the given Laurent series. In the symmetric case, the approximation can be given a least squares interpretation. The interpolation points can then be chosen in an optimal way. A numerical example of the latter problem is given.

1. Problem formulation

The Laurent-Hermite interpolation problem as we define it here is an attempt to generalize the concept of Laurent-Padé approximation to a related interpolation problem.

We introduce the definition gradually and start with the non-Laurent problem :

Suppose $F(x)$ is an analytic function in a region of the complex plane (suppose it is the unit disc) and let $\alpha_0, \alpha_1, \alpha_2, \ldots$ be a sequence of points (some of them may coincide) in this region. The (rational) Hermite interpolation problem consists in finding two polynomials $P^{m,n}$ and $Q^{m,n}$ of degrees bounded by m and n respectively such that $R^{m,n} = P^{m,n}/Q^{m,n}$ interpolates F in $\alpha_0, \alpha_1, \ldots \alpha_N$ with $N \geqslant m+n$.

Thus

$$[F - R^{m,n}](x) = g(x) \prod_0^N (x-\alpha_i)$$

with g analytic in the disc. If all points α_i are different, this is the Cauchy problem. If all points $\alpha_i = 0$, this is the Padé problem.

For the Laurent-Padé problem $F(x)$ need not be analytic everywhere inside the unit disc. Suppose it has a Laurent series expansion around $x = 0$ valid in a region containing the unit circle $|x| = 1$. Then one tries to find trigonometric polynomials $P^{m,n}$ and $Q^{m,n}$ of degree not exceeding m and n respectively such that $R^{m,n} = P^{m,n}/Q^{m,n}$ has a Laurent series expansion in the neighborhood of $|x| = 1$ being such that $F - R^{m,n}$ contains no terms in x^i for $|i| \leqslant N$.

This problem has received some attention [11,6]. It is known to be related to the Schur algorithm [1,16] and thus to two-point Padé approximation [17]. This means that one constructs a rational function approximating two power series simultaniously, viz. the "analytic part" of F and the "coanalytic part" of F. One part is approximated at the origin, the other at infinity.

The Laurent-Hermite interpolation problem is a direct generalization of this. We interpolate the "analytic part" at $0 = \alpha_0, \alpha_1, \alpha_2, \ldots$, a sequence of points inside the unit disc and the "coanalytic part" at points $\infty = 1/\bar{\beta}_0, 1/\bar{\beta}_1, 1/\bar{\beta}_2, \ldots$ outside the unit circle (upper bar denotes complex conjugation) so that we have in the end

$$F(x) - R^{m,n}(x) = \Delta_1^{m,n}(x) \prod_0^N (x-\alpha_i) + \Delta_2^{m,n}(1/x) \prod_0^N (\frac{1}{x} - \bar{\beta}_i) \tag{1}$$

with $\Delta_1^{m,n}(x)$ and $\Delta_2^{m,n}(1/x)$ as functions of x, analytic inside the unit disc, everywhere, except for poles of $R^{m,n}$.

In this paper we are not going to consider this general case but an interesting special version of it. Suppose we know the zeros of F, or at least, can estimate them. If they are $\alpha_1, \alpha_2, \ldots$ and $1/\bar{\beta}_1, 1/\bar{\beta}_2, \ldots$ with $|\alpha_i|, |\beta_i| < 1$, then it is wise to propose an approximant $R^{m,n}$ of the form

$$R^{m,n}(x) = \frac{\prod_1^m (x-\alpha_i)}{\psi_n(x)} \cdot \frac{\prod_1^m (1/x-\bar{\beta}_i)}{\phi_{n^\star}(1/x)} \tag{2}$$

with $\phi_{n^\star}$ and ψ_n polynomials of degree at most n. If the estimates of α_i and β_i are exact and if F(x) is of the form (2), then we shall formulate and algorithm that gives $R^{m,n}(x) = F(x)$ after max(m,n) steps. If the estimates α_i and β_i are not exact, then our procedure still provides an approximation in a restricted Laurent-Hermite sense. Of course we can not go as far as in (1) because we fix the numerator beforehand, thus reducing the degrees of freedom. However all remaining parameters in the denominator are used to satisfy (1) as far as possible. More precisely we require ψ_n and $\phi_{n^\star}$ to be such that

$$F(x) - \frac{\prod_1^m (x-\alpha_i)(1/x-\bar{\beta}_i)}{\psi_n(x)\phi_{n^\star}(1/x)} = \Delta_1^{m,n}(x) \prod_0^m (x-\alpha_i)/\psi_n(x) + \Delta_2^{m,n}(1/x) \prod_0^m (1/x-\bar{\beta}_i)/\phi_{n^\star}(1/x)$$

with as before $\alpha_0 = \beta_0 = 0$ and $\Delta_1^{m,n}(x)$ and $\Delta_2^{m,n}(1/x)$ as functions of x are analytic inside the unit disc.

First remark that if $m < n$, then we can always add $\alpha_{m+1} = \alpha_{m+2} = \ldots = \alpha_n = \beta_{m+1} = \beta_{m+2} = \ldots = \beta_n = 0$ because this does not change the numerator. The case $m > n$ gives more trouble. We don't like to consider it at this moment and suppose that in this case m-n zeros are simply devided out of $F(x)$ so that we restrict ourselves from now on to the case $m = n$.

2. Justification

The Laurent-Hermite interpolation problem has recently found a very interesting application in the construction of least squares ARMA (auto-regressive, moving average) filters [2, 3, 4, 7, 8, 9]. In this special case the given Laurent series $F(x) = \sum_{-\infty}^{\infty} f_k x^k$ represents an autocorrelation function and therefore $f_{-k} = \bar{f}_k$ and for all n the Toeplitz matrices constructed on the parameters $\{f_j\}_{j=-n}^{n}$ are positive definite. For the general theory we refer to the literature where it is shown how the classical Wiener-Massani AR-prediction theory [18] can be nicely adapted in this context. To explain the basic idea however we may restrict ourselves to the case that $F(x)$ is rational. $F(x)$ is then factorizable as $F(x) = s(x).\overline{s(1/\bar{x})}$ with $s(x)$ rational having all its poles and zeros inside the unit disc. $s(x)$ is in essence the filter one wants to construct from the covariance data $F(x)$ and the location of the poles and zeros of $s(x)$ are crucial for stability reasons. The example shows also why we do not consider the general Laurent-Hermite problem. Indeed the result of this would deliver an approximant to $F(x)$ that may not be factorizable in general (i.e. not be positive on the unit circle) and thus gives no approximation to $s(x)$ which is the ultimate goal. By restricting ourselves to the special problem described before, we fix the numerator as a product of linear factors with zeros separated as desired and the denominator is found as a product of two polynomials. Thus we directly find the approximant of $s(x)$ because the zeros of the denominator are automatically separated. This results from the properties of $F(x)$ and the algorithm we are going to define in the next section. Just as in the classical case of AR filtering, the approximant can be shown to be optimal in a weighted least squares sense. At the end of the paper we shall give a numerical example illustrating these ideas. The remaining sections give a general framework for these techniques. We give in section 3 the description of an algorithm in the style of Schur [1,16] and Nevanlinna-Pick [1] to construct the approximant recursively. Section 4 defines reproducing kernels for a certain indefinite inner product space and derives a Christoffel-Darboux type formula. In section 5 it is briefly sketched how these reproducing kernels satisfy a recursion that turns out to be exactly the same as the one described in section 3.

3. A Schur-like algorithm

In this section we describe an algorithm that shall give the required approximant recursively. It is a continued fraction type of algorithm which is a generalization of a similar recursive scheme proposed by Shur [1,16].

Given $F(x) = \sum_{-\infty}^{\infty} f_k x^k$, $f_0 = 1$, we split this up into two parts :

$F^+(x) = 1 + 2\sum_{1}^{\infty} f_k x^k$ and $F^-(1/x) = 1 + 2\sum_{1}^{\infty} f_k x^k$, thus $F(x) = \frac{1}{2}[F^-(1/x)+F^+(x)]$.

We call $\frac{1}{2}[1+F^+(x)]$ the analytic part of $F(x)$ and $-\frac{1}{2}[1-F^-(1/x)]$ the coanalytic part of $F(x)$.

We construct the matrix

$$\Delta_0(x) = \frac{1}{2}\begin{bmatrix} 1+F^+(x) & -(1-F^+(x)) \\ -(1-F^-(1/x)) & 1+F^-(1/x) \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} \Delta_{11}^0(x) & \Delta_{12}^0(x) \\ \Delta_{21}^0(1/x) & \Delta_{22}^0(1/x) \end{bmatrix}$$

Now we successively transform $\Delta_0(x)$ into $\Delta_1(x), \Delta_2(x), \ldots,$ such that

$$\Delta_n(x) = \begin{bmatrix} \Delta_{11}^n(x) & \Delta_{12}^n(x) \\ \Delta_{21}^n(x) & \Delta_{22}^n(1/x) \end{bmatrix}$$

satisfies :

$$\Delta_{11}^n(x) = \prod_{1}^{n}(x-\alpha_i).\tilde{\Delta}_{11}^n(x) \quad ; \quad \Delta_{12}^n(x) = \prod_{0}^{n}(x-\alpha_i)\tilde{\Delta}_{12}^n(x)$$
$$\Delta_{21}^n(1/x) = \frac{1}{x}\prod_{1}^{n}(1-\bar{\beta}_i x).\tilde{\Delta}_{21}^n(1/x) \quad ; \quad \Delta_{22}^n(1/x) = \prod_{1}^{n}(1-\bar{\beta}_i x)\tilde{\Delta}_{22}^n(1/x) \qquad (3)$$

with $\tilde{\Delta}_{1j}^n(x)$ and $\tilde{\Delta}_{2j}^n(1/x)$ as functions of x analytic inside the unit disc.

Clearly these requirements are met for n = 0.

The transform from $\Delta_{n-1}(x)$ to $\Delta_n(x)$ is described as follows :

$$\Delta_n(x) = \Delta_{n-1}(x)\theta_n(x)$$

with

$$\theta_n(x) = \begin{bmatrix} 1 & -\gamma_n^\alpha \\ -\overline{\gamma_n^\beta} & 1 \end{bmatrix} \begin{bmatrix} x-\alpha_n & 0 \\ 0 & 1-\bar{\beta}_n x \end{bmatrix} \begin{bmatrix} 1 & -\rho_n^\alpha \\ -\overline{\rho_n^\beta} & 1 \end{bmatrix}$$

$$\rho_n^\alpha = \tilde{\Delta}_{12}^{n-1}(\alpha_n)/\tilde{\Delta}_{11}^{n-1}(\alpha_n) \quad ; \quad \gamma_n^\alpha = \alpha_n\, \rho_n^\alpha$$

$$\overline{\rho_n^\beta} = \tilde{\Delta}_{21}^{n-1}(\bar{\beta}_n)/\tilde{\Delta}_{22}^{n-1}(\bar{\beta}_n) \quad ; \quad \gamma_n^\beta = \beta_n\, \rho_n^\beta$$

if α_n and $\beta_n \neq 0$. For $\alpha_n = 0$ or $\beta_n = 0$ we would have $\rho_n^\alpha = \rho_n^\beta = 0$. In that case we have to replace the formulas for ρ_n^α and ρ_n^β by

$$\rho_n^\alpha = \frac{d}{dx}\Gamma_n^\alpha(x)\bigg|_{x=0} \quad , \; \gamma_n^\alpha = 0, \quad \Gamma_n^\alpha(x) = \tilde{\Delta}_{12}^{n-1}(x)/\tilde{\Delta}_{11}^{n-1}(x)$$

and

$$\overline{\rho_n^\beta} = \frac{d}{dx}\Gamma_n^\beta(x)\bigg|_{x=0} \quad , \; \gamma_n^\beta = 0, \quad \Gamma_n^\beta(x) = \tilde{\Delta}_{21}^{n-1}(x)/\tilde{\Delta}_{22}^{n-1}(x)$$

It can be verified that $\rho_n^{\alpha/\beta}$ and $\gamma_n^{\alpha/\beta}$ are chosen such that if $\Delta_{n-1}(x)$ satisfies (3) with n replaced by (n-1), then also $\Delta_n(x)$ will satisfy (3).

Now consider the matrix

$$\Theta_n(x) = \theta_1(x)\theta_2(x) \;\ldots\; \theta_n(x).$$

It is clear from its construction that the elements of $\Theta_n(x)$ are polynomials in x of degree n. By taking linear combinations of elements in $\Theta_n(x)$ we define the polynomials $f_n^\star, g_n, \psi_n$ and $\phi_n^\star$ by the relation

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \Theta_n(x) = \begin{bmatrix} \phi_n^\star & \psi_n \\ f_n^\star & -g_n \end{bmatrix}$$

or equivalently

$$\Theta_n(x) = \frac{1}{2}\begin{bmatrix} \phi_n^\star + f_n^\star & \psi_n - g_n \\ \phi_n^\star - f_n^\star & \psi_n + g_n \end{bmatrix} .$$

Multiplying out the product $\Delta_0(x)\Theta_n(x) = \Delta_n(x)$ where $\Delta_n(x)$ has the properties (3) gives for the (1,2) element :

$$-g_n(x) + F^+(x)\psi_n(x) = x \prod_1^n (x-\alpha_i)\tilde{\Delta}_{12}(x)$$

and for the (2,1) element :

$$-f_n^\star(x) + F^-(1/x)\phi_n^\star(x) = \frac{1}{x}\prod_1^n (1-\bar{\beta}_i x)\tilde{\Delta}_{21}(1/x)$$

or if we divide by x^n and set $f_n^\star(x)/x^n = f_{n\star}(1/x)$ and $\psi_n^\star(x)/x^n = \psi_{n\star}(1/x)$:

$$-f_{n\star}(1/x) + F^-(1/x)\phi_{n\star}(1/x) = \frac{1}{x}\prod_1^n (\frac{1}{x} - \bar{\beta}_i)\tilde{\Delta}_{21}(1/x).$$

Now

$$\det \Theta_n(x) = \prod_1^n \det \theta_i(x)$$

which gives

$$\phi_n^\star(x)g_n(x) + f_n^\star(x)\psi_n(x) = \prod_1^n (x-\alpha_i)(1-\bar{\beta}_i x)(1-\gamma_i^\alpha\overline{\gamma_i^\beta})(1-\rho_i^\alpha\overline{\rho_i^\beta})$$

$$= 2C_n \prod_1^n (x-\alpha_i)(1-\bar{\beta}_i x)$$

where C_n is a constant.

If $\psi_n(x)$ and $\phi_{n\star}(1/x)$ have no zeros on the unit circle, then in a certain region containing this circle we may define

$$F_n^+(x) = g_n(x)/\psi_n(x) \quad \text{and} \quad F_n^-(1/x) = f_{n\star}(1/x)/\phi_{n\star}(1/x)$$

and

$$F_n(x) = \frac{1}{2}[F_n^-(1/x) + F_n^+(x)] = C_n \frac{\prod_1^n (x-\alpha_i)(\frac{1}{x} - \bar{\beta}_i)}{\psi_n(x)\phi_{n\star}(1/x)} .$$

We have

$$F(x) - F_n(x) = \prod_0^n (x-\alpha_i)\tilde{\Delta}_{12}(x)/\psi_n(x) + \prod_0^n (\frac{1}{x} - \bar{\beta}_i)\tilde{\Delta}_{21}(\frac{1}{x})/\phi_{n\star}(\frac{1}{x})$$

as required.

This completes the description of the algorithm.

4. Indefinite inner product spaces, reproducing kernels and Christoffel-Darboux relations

Like the Schur algorithm is related to the Szegö polynomials orthogonal on the unit circle, the algorithm of the previous is related to orthogonal (rational) functions and reproducing kernels for some indefinite inner product space which we shall introduce in this section.

Consider the space F of functions having a Laurent series expansion in a region containing the unit circle. We define two subspaces of rational functions of a certain degree n with given poles outside the unit disc :

$$\mathcal{L}_n^\alpha = \{ \frac{p_n(x)}{\prod_1^n (1-\bar{\alpha}_j x)} \quad , p_n \text{ polynomial of degree } n, |\alpha_j| < 1 \}$$

$$\mathcal{L}_n^\beta = \{ \frac{q_n(x)}{\prod_1^n (1-\bar{\beta}_j x)} \quad , q_n \text{ polynomial of degree } n, |\beta_j| < 1 \}$$

For fixed $F(x) = \sum_{-\infty}^{\infty} f_k x^k$, $f_0 = 1$, we define a linear functional C such that

$$C(x^k) = f_{-k} \qquad k = \dots,-1,0,1,2,\dots$$

and we associate with it the following indefinite inner product over F

$$\langle f,g \rangle = C(f\, g_\star)$$

where by definition $h_\star(x) = \overline{h(1/\bar{x})}$.

We define also transformations from $\mathcal{L}_n^\alpha$ into $\mathcal{L}_n^\beta$ and conversely by

$$f_n \in \mathcal{L}_n^\alpha \to f_n^{\alpha\beta} \overset{\text{def}}{=} U_n^{\alpha\beta}(x) f_{n\star}(x) \in \mathcal{L}_n^\beta$$

$$g_n \in \mathcal{L}_n^\beta \to g_n^{\beta\alpha} \overset{\text{def}}{=} U_n^{\beta\alpha}(x) g_{n\star}(x) \in \mathcal{L}_n^\alpha$$

where

$$U_n^{\alpha\beta}(x) = \prod_{i=1}^n \frac{x-\alpha_i}{1-\bar{\beta}_i x} \quad \text{and} \quad U_n^{\beta\alpha}(x) = \prod_{i=1}^n \frac{x-\beta_i}{1-\bar{\alpha}_i x}$$

With these transformations we have the following relations as can be easily verified

PROPERTY 4.1 For $f \in \mathcal{L}_n^{\alpha}$ and $g \in \mathcal{L}_n^{\beta}$:

$$<f,g> = <g_\star, f_\star> = <g^{\beta\alpha}, f^{\alpha\beta}>$$

The recursive algorithm of the previous section is closely related to the construction of reproducing kernels [14] for $\mathcal{L}_n^{\alpha}$ and $\mathcal{L}_n^{\beta}$, $n = 0,1,\ldots$. These kernels are defined in the following way :

The kernels $k_n(x,y)$ and $\ell_n(x,y)$ are in $\mathcal{L}_n^{\beta}$ resp. $\mathcal{L}_n^{\alpha}$ as a function of x for fixed y and are such that

$$<f(.),k_n(.,y)> = f(y) \qquad \forall f \in \mathcal{L}_n^{\alpha}$$

$$<\ell(.,y),g(.)> = \overline{g(y)} \qquad \forall g \in \mathcal{L}_n^{\beta}.$$

As in the classical case, under certain non-degeneracy conditions for the product space, (which we suppose to be satisfied wherever needed) these kernels are uniquely defined and can be expressed as a combination of a biorthogonal basis for $\mathcal{L}_n^{\alpha}$ and $\mathcal{L}_n^{\beta}$.

Indeed, suppose $\{\lambda_j\}_0^n$ and $\{\mu_j\}_0^n$ form a basis for $\mathcal{L}_n^{\alpha}$ and $\mathcal{L}_n^{\beta}$ respectively and are biorthonormal in the sense that $<\lambda_i,\mu_j> = \delta_{ij}$. It is easily verified that

$$k_n(x,y) = \sum_0^n \overline{\lambda_j(y)}\ \mu_j(x) \quad \text{and} \quad \ell_n(y,x) = \sum_0^n \overline{\mu_j(x)}\ \lambda_j(y)$$

LEMMA 4.2

$$\ell_n(y,x) = \overline{U_n^{\beta\alpha}(x)}\ U_n^{\beta\alpha}(y)\ \overline{k_n(1/\overline{y},1/\overline{x})}$$

PROOF

By definition is $<f(x),k_n(x,y)> = f(y)$, $\forall f \in \mathcal{L}_n^{\alpha}$. With property 4.1 then also

$$<U_n^{\beta\alpha}(x)\ \overline{k_n(1/\overline{x},y)},\ U^{\alpha\beta}(x)\ \overline{f(1/\overline{x})}> = f(y).$$

Take the $^{\alpha\beta}$ transform with respect to y then

$$<\overline{U_n^{\alpha\beta}(y)}\ U_n^{\beta\alpha}(x)\ \overline{k_n(1/\overline{x},1/\overline{y})},\ f^{\alpha\beta}(x)> = \overline{f^{\alpha\beta}(y)}.$$

If $f \in \mathcal{L}_n^{\alpha}$ is arbitrary then $f^{\alpha\beta} \in \mathcal{L}_n^{\beta}$ is arbitrary so that the result follows by definition of $\ell_n(x,y)$.

Using the biorthonormal basis functions, the following corollary is simple to prove

COROLLARY 4.3

$$\ell_n(y,\alpha_n) = \kappa_n^{\lambda}\,\mu_n^{\beta\alpha}(y) \quad , \quad \ell_n(\beta_n,x) = \overline{\kappa_n^{\mu}}\,\overline{\kappa_n^{\alpha\beta}}(x)$$

and

$$\ell_n(\beta_n,\alpha_n) = \kappa_n^{\lambda}\,\overline{\kappa_n^{\mu}}$$

with

$$\kappa_n^{\lambda} = \overline{\lambda_n^{\alpha\beta}(\alpha_n)} \quad \text{and} \quad \kappa_n^{\mu} = \overline{\mu_n^{\beta\alpha}(\beta_n)}$$

The following theorem gives the analogue of the Christoffel-Darboux relation for orthogonal polynomials, now generalized to a relation for the biorthonormal systems.

THEOREM 4.4

$$k_n(x,y) = \frac{\overline{\mu_{n+1}^{\beta\alpha}(y)}\;\lambda_{n+1}^{\alpha\beta}(x) - \overline{\lambda_{n+1}(y)}\;\mu_{n+1}(x)}{1-\left(\overline{\dfrac{y-\beta_{n+1}}{1-\bar{\alpha}_{n+1}y}}\right)\left(\dfrac{x-\alpha_{n+1}}{1-\bar{\beta}_{n+1}x}\right)}$$

PROOF

Call the right hand side $R(x,y)$, then we have to prove that $<f(x),R(x,y)> = f(y)$, $\forall f \in \mathcal{L}_n^{\alpha}$. Now $<f(x),R(x,y)> = f(y)<1,R(x,y)> + <(x-y)h(x),R(x,y)>$ where $f(x) - f(y) \overset{\text{def}}{=} (x-y)h(x) \in \mathcal{L}_n^{\alpha}$.

We first show that the second term vanishes. Indeed, working on the denominator of $R(x,y)$ we find with some algebra that

$$<(x-y)h(y),R(x,y)> = c <h_1(x)(1-\bar{\alpha}y),\overline{\mu_{n+1}^{\beta\alpha}(y)}\;\lambda_{n+1}^{\alpha\beta}(x) - \overline{\lambda_{n+1}(y)}\;\mu_{n+1}(x)>$$

with $h_1(x) = (x-\beta_{n+1})h(x) \in \mathcal{L}_n^{\alpha}$. Using property 4.1 and the orthogonality of μ_{n+1} and λ_{n+1} on $\mathcal{L}_n^{\alpha}$ and $\mathcal{L}_n^{\beta}$ respectively, it follows that this is zero.

We thus may conclude that $<f(x),R(x,y)> = f(y)\eta(y)$ with $\eta(y) = <1,R(.,y)>$.

Similarly, we can show that $<R(x,y),f(y)> = \eta'(x)\overline{f(x)}$ where $\eta'(x) = \overline{<R(x,.),1>}$.

Using both results on $<R(x,.),R(.,y)>$ we find that it equals $\eta(y)\overline{R(x,y)}$ and also $\eta'(x)\overline{R(x,y)}$ so that we must have $\eta'(x) = \eta(y) = \text{constant} = \eta$.

Take $x = \alpha_{n+1}$ and $y = \beta_{n+1}$, then we have shown that

$$\eta\; k_n(\alpha_{n+1},\beta_{n+1}) = \kappa_{n+1}^{\mu}\;\overline{\kappa_{n+1}^{\lambda}} - \overline{\lambda_{n+1}(\beta_{n+1})}\;\mu_{n+1}(\alpha_{n+1})$$

The first term in the right hand side equals $k_{n+1}(\alpha_{n+1},\beta_{n+1})$ because of corollary 4.3. So the difference in the right hand side is $k_n(\alpha_{n+1},\beta_{n+1})$ and thus is $\eta = 1$.

5. Recurrence relations

In this section we give a recursion for the reproducing kernels defined in the previous section and briefly indicate how it is related to the recursion in section 3. The proof of the correspondence between both recursions requires more knowledge about the indefinit inner product space which we shall not give in detail. More on this will be published later (see also [2,7,9]).

The reproducing kernels satisfy the recursions given in the following

THEOREM 5.1

$$(1-\rho^{\alpha}_{n+1}(y)\overline{\rho^{\beta}_{n+1}(y)})[\ell^{\alpha\beta}_{n+1}(x,y)k_{n+1}(x,y)] = [\ell^{\alpha\beta}_{n}(x,y)k_{n}(x,y)].$$

$$\begin{bmatrix} 1 & \left(\dfrac{y-\alpha_{n+1}}{1-\bar{\beta}_{n+1}y}\right)\rho^{\alpha}_{n+1}(y) \\ \left(\overline{\dfrac{y-\beta_{n+1}}{1-\bar{\alpha}_{n+1}y}}\right)\overline{\rho^{\beta}_{n+1}(y)} & 1 \end{bmatrix}\begin{bmatrix} \dfrac{x-\alpha_{n+1}}{1-\bar{\beta}_{n+1}x} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -\rho^{\alpha}_{n+1}(y) \\ -\overline{\rho^{\beta}_{n+1}(y)} & 1 \end{bmatrix}$$

with $\rho^{\alpha}_{n+1}(y) = -\mu_{n+1}(y)/\overline{\mu^{\beta\alpha}_{n+1}(y)}$

$$\rho^{\beta}_{n+1}(y) = -\lambda_{n+1}(y)/\overline{\lambda^{\alpha\beta}_{n+1}(y)}$$

PROOF

Obviously

$$k_{n+1}(x,y) - k_n(x,y) = \overline{\lambda_{n+1}(y)}\ \mu_{n+1}(x). \tag{4}$$

$\overline{\lambda_{n+1}(y)}\ \mu_{n+1}(x)$ is herein replaced by an expression that can be found from the Christoffel-Darboux formula. We obtain

$$k_{n+1}(x,y) = \left(\overline{\frac{y-\beta_{n+1}}{1-\bar{\alpha}_{n+1}y}}\right)\left(\frac{x-\alpha_{n+1}}{1-\bar{\beta}_{n+1}x}\right)k_n(x,y) + \overline{\mu^{\beta\alpha}_{n+1}(y)}\;\lambda^{\alpha\beta}_{n+1}(x) \;. \tag{5}$$

The ${}^{\alpha\beta}$ transform of a relation like (4) written down for $\ell_{n+1}(x,y)$ gives

$$\ell^{\alpha\beta}_{n+1}(x,y) = \left(\frac{x-\alpha_{n+1}}{1-\bar{\beta}_{n+1}x}\right)\ell^{\alpha\beta}_n(x,y) + \lambda^{\alpha\beta}_{n+1}(x)\;\mu_{n+1}(y) \;. \tag{6}$$

Extract $\lambda^{\alpha\beta}_{n+1}(x)$ from (6) and substitute in (5), then

$$\rho^{\alpha}_{n+1}(y)k_{n+1}(x,y) = \left(\overline{\frac{y-\beta_{n+1}}{1-\bar{\alpha}_{n+1}y}}\right)\left(\frac{x-\alpha_{n+1}}{1-\bar{\beta}_{n+1}x}\right)k_n(x,y)\;\rho^{\alpha}_{n+1}(y)$$

$$- \left[\ell^{\alpha\beta}_{n+1}(x,y) - \left(\frac{x-\alpha_{n+1}}{1-\bar{\beta}_{n+1}x}\right)\ell^{\alpha\beta}_n(x,y)\right] \tag{7}$$

with $\rho^{\alpha}_{n+1}(y) = -\mu_{n+1}(y)/\overline{\mu^{\beta\alpha}_{n+1}(y)}$.

Repeat the same thing for $\ell_{n+1}(x,y)$ and take the ${}^{\alpha\beta}$ transform. This yields :

$$\overline{\rho^{\beta}_{n+1}(y)}\,\ell^{\alpha\beta}_{n+1}(x,y) = \overline{\rho^{\beta}_{n+1}(y)}\left(\frac{y-\alpha_{n+1}}{1-\bar{\beta}_{n+1}y}\right)\ell^{\alpha\beta}_n(x,y) - k_{n+1}(x,y) + k_n(x,y) \tag{8}$$

with $\rho^{\beta}_{n+1}(y) = -\lambda_{n+1}(y)/\overline{\lambda^{\alpha\beta}_{n+1}(y)}$.

Combining (7) and (8) gives the recursion.

It follows simply from Theorem 5.1 that

COROLLARY 5.2

$$(1-\rho^{\alpha}_{n+1}(0)\overline{\rho^{\beta}_{n+1}(0)})k_{n+1}(0,0) = (1-\alpha_{n+1}\bar{\beta}_{n+1}\rho^{\alpha}_{n+1}(0)\overline{\rho^{\beta}_{n+1}(0)})k_n(0,0)$$

and because $k_0(0,0) = 1$:

$$k_n(0,0) = \prod_1^n (1-\alpha_i\bar{\beta}_i\rho^{\alpha}_i(0)\overline{\rho^{\beta}_i(0)})/(1-\rho^{\alpha}_i(0)\overline{\rho^{\beta}_i(0)}).$$

There is a certain similarity between the $\theta_n(x)$ matrices of section 3 and the recursion for the reproducing kernels given above. To show the correspondence more explicitly, take $y = 0$ and suppose that $\phi^{\star}_n(x)$ and $\psi_n(x)$ are the numerators of $\ell^{\alpha\beta}_n(x,0)c_n$ and $k_n(x,0)c_n$ respectively, with

$$c_n = \prod_{1}^{n} (1-\rho_i^\alpha(0)\overline{\rho_i^\beta(0)})$$

The recursion for these polynomials is then

$$[\phi_n^\star(x) \;\; \psi_n(x)] = [\phi_{n-1}^\star(x) \;\; \psi_{n-1}(x)]\theta_n(x) \tag{9}$$

with $\theta_n(x)$ of exactly the same form as the one defined in section 3. Only the definition of ρ_n^α and ρ_n^β is different.

The explicit identification of both recursions requires a more detailed study of the indefinite inner product space. If this space is not degenerate [5] a whole projection theory can be set up and the proofs are relatively simple. Consult [2,7,9] for similar proofs. If $F(x)$ is such that the inner product space is degenerate, then the algorithms may break down and we have a situation similar to the construction of Padé approximants for a function with non-normal Padé table. However the algorithms may be modified to overcome this situation and proofs go through, although everything becomes considerably more complicated.

That we found in section 3 only the recursion of the reproducing kernels for the special case $y = 0$ is due to the fact that we normalized in the Schur algorithm in each step at $x = 0$, and not in a general point $x = y$. If the algorithm of section 3 would have been adopted as such, then we would have found the complete recursion of theorem 5.1.

Thus both algorithms are completely equivalent. The algorithm of section 3 is called of <u>outgoing</u> type because it starts with the whole information about $F(x)$ in $\Delta_0(x)$ and at every step, some information $(\rho_i^\alpha, \rho_i^\beta)$ is extracted and the algorithm goes on with what "remains" of $F(x)$. This algorithm resembles much the division algorithms constructing continued fractions whose convergents are Padé approximants. The algorithm of this section is called of <u>incoming</u> type because one starts with no information ($\phi_0^\star = \psi_0 = 1$) at all and gradually, as the algorithm comes along, more and more information about $F(x)$ is thrown in. This is accumulated in the polynomials of increasing degree. The matrix $\Theta_n(x)$ defined in section 3 contains also the polynomials $f_n^\star$ and g_n. These are related to reproducing kernels as $\phi_n^\star$ and ψ_n are. It are the kernels for similar rational function spaces, now with an indefinit inner product related to the function $\Phi(x) = \frac{1}{2}[\Phi^-(1/x) + \Phi^+(x)]$ with $\Phi^+(x)$ and $\Phi^-(x)$ the formal inverses of

$F^+(x)$ and $F^-(x)$ respectively.

6. The location of transmission zeros, a numerical example

In this section we give a numerical example which is taken from an important application of the previous theory viz. when $F(x)$ is an autocorrelation function for a discrete time stationary stochastic process. This is only a special case of what is developed in the previous sections but, because of the symmetry in the problem, it simplifies the formulae and shows clearly the applicability of the theory. So e.g. $f_k = \bar{f}_{-k}$ and $F(x)$ is positive real on the unit circle. Also $\alpha_i = \beta_i$ $\forall i$ so that $\mathcal{L}_n^{\alpha} = \mathcal{L}_n^{\beta} = \mathcal{L}_n$.

The general theory for this application has appeared elsewhere [2,3,7,8,9] and it is not repeated here. We mention only the main results.

The most important consequence of the specialisation is that we can now give a least squares optimality interpretation for the approximant, weighted with $F(x)$ itself. Indeed, suppose $F(x)$ has the factorization $s(x).\overline{s(1/\bar{x})}$, then

$$\min_{h_n \in \mathcal{L}_n} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|\frac{1}{s(e^{j\theta})} - h_n(e^{j\theta})\right|^2 F(e^{j\theta})d\theta$$

is attained for $h_n = k_n(x,0)s(0)$ and the minimum is given by $S_n = 1 - k_n(0,0)s(0)^2$. Remark however that the solution is parametrized in the chosen transmission zeros α_i that define $\mathcal{L}_n$. So the previous methods do not only give us an algorithm to find an approximant with given transmission zeros, but we can even improve upon this result if we are willing to optimize the least squares error S_n as a function of $\alpha_0, \alpha_1, \alpha_2, \ldots$. Therefore we have to maximize $k_n(0,0)$, or equivalently to minimize $\prod_1^n (1-|\rho_i|^2)/(1-|\gamma_i|^2)$ (see corollary 5.2 with $\rho_i^{\alpha} = \rho_i^{\beta} = \rho_i$ and $\gamma_i = \alpha_i\rho_i$). Thus we have a tool to choose the optimal location of the interpolation points α_i (= transmission zeros). We shall even find the exact transmission zeros, provided the given $F(x)$ is rational and that the degree of the rational approximant is high enough (at least as high as the degree of the given function). The objective function to be minimized is evaluated after n steps of the recursive algorithm defined in section 3, but, taking the symmetry into account, it becomes even simpler. It is familiar with the interpolation algorithm of Nevanlinna and Pick [1]. So if n is not too high,a function evaluation is relatively cheap.

For a numerical example we took for $F(x)$ the Laurent series valid at $|x|=1$ for the function

$$\frac{(x-\alpha)^2}{(x-\beta_1)^2 (x-\beta_2)(x-\bar{\beta}_2)} \; ; \; \alpha = 0.5; \; \beta_1 = -0.5; \; \beta_2 = 0.5 \exp (j\pi/4)$$

The first coefficients are listed below :

$f_0 = 3.925130$	$f_1 = -3.205387$
$f_2 = 2.104909$	$f_3 = -1.367445$
$f_4 = 0.825844$	$f_5 = -0.478882$
$f_6 = 0.279871$	$f_7 = -0.156157$
$f_8 = 0.087150$	$f_9 = -0.048429$
$f_{10} = 0.026176$	$f_{11} = -0.014318$
$f_{12} = 0.007713$	$f_{13} = -0.004113$
$f_{14} = 0.002214$	$f_{15} = -0.001170$
$f_{16} = 0.000610$	$f_{17} = -0.000328$
$f_{18} = 0.000171$	$f_{19} = -0.000090$

If we estimate the zeros at the origin and at $\alpha = 0.5$ exactly, then we find after four steps of the Schur algorithm the exact approximant with a minimal value of $S_4 = 0.90398$. However numerical experiments show that the location of α at 0.5 is not crucial at all because any other value of α in the interval [0.4,0.8] (if we know α to be real) gives about the same minimum. This illustrates that the approximation is in many cases (even in this simple example) rather insensitive to an exact location of the transmission zeros. If however we look at this procedure as a method to find the exact transmission zeros, then the example illustrates the ill conditioning of this problem. It is comparable with a situation of least squares exponential approximation or any other least squares rational approximation problem. Other, more complicated examples all showed the same characteristic behaviour, the situation becoming worse as the transmission zeros approached the unit circle.
The convergence of this special case has been studied at least theoretically.

The general situation as described in previous sections however is still an open question. In this case we can also have non-normal situations as for the Padé approximation problem. The study of these anomalies is still under investigation and will be reported on later. The theory is however appealing and several potential applications in minimax rational approximation [10], systems theory, coding theory [13], networks [9] and scattering theory [12] are developed.

References

1. Akhiezer, N.I. : The classical moment problem, Oliver and Boyd, Edinburgh-London, 1965.

2. Bultheel, A. : Recursieve rationale benaderingen; Ph.D. Thesis, K.U.Leuven, Sept. 1979.

3. Bultheel, A. , Dewilde, P. : On the optimal location of transmission zeros in least squares ARMA filtering, Report TW 51, K.U.Leuven, June 1980.

4. Bultheel, A., Dewilde, P. : On the relation between Padé approximation and Levinson/Schur recursive methods in M. Kunt - F. de Coulon (eds.), Proceedings EUSIPCO-80, North-Holland, Amsterdam, 1980, 517-523.

5. Bognar, J. : Indefinite inner product spaces, Springer-Verlag, Berlin,1974.

6. Chisholm, J.S.R., Common, A. : Chebyshev Padé approximants, in L. Wuytack (ed.), Padé approximation and its applications, Springer-Verlag, Berlin, (1979), 1-19.

7. Dewilde, P., Dym, H. : Schur recursions, error formulas and convergence of rational estimators for stationary stochastic sequences, to appear in IEEE Trans. Inf. Theory.

8. Dewilde, P. : On the convergence of the generalized Szegö-Levinson least square error algorithm, Tech. Rept. 82, TH Delft, 1979.

9. Dewilde, P., Vieira, A., Kailath, T. : On a generalized Szegö-Levinson realization algorithm for optimal linear predictors based on a network synthesis approach, IEEE Trans. CAS-25, (1978), 663-675.

10. Genin, Y., Kung, S.Y. : A two-variable approach to the model reduction problem with Hankel norm criterion. Subm. for publication IEEE Trans. 1980.

11. Gragg, W.B. : Laurent, Fourier and Chebyshev-Padé tables in [15] (1977), 61-72.

12. Lax, S.D., Phillips, R.S. : Scattering theory, Academic Press, New York, 1967.

13. McEliece, R.J. : The theory of information and coding, Addison-Wesley, Reading, Mass., 1977.

14. Meschkowski, H. : Hilbertsche Räume mit Kernfunktion, Springer-Verlag, Berlin, 1962.

15. Saff, E.B., Varga, R.S. (eds.) : Padé and rational approximation, Academic Press, New York, 1977.

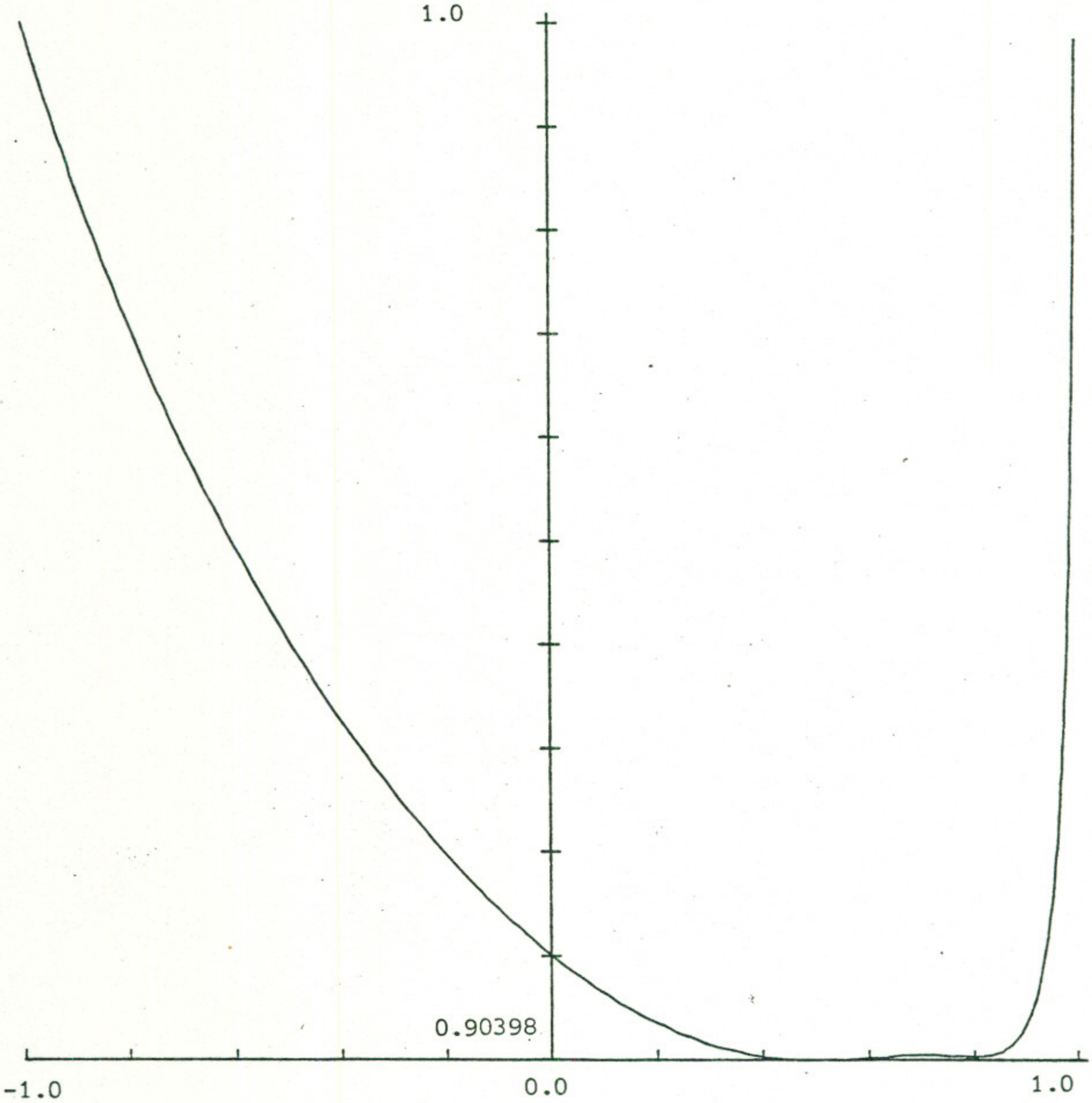
16. Schur, I. : Über Potenzreihen die im Innern des Einheitskreises beschränkt sind, Z. Reine Angew. Math. 147 (1917), 205-232.

17. Thron, W.J. : Two-point Padé tables, T-fractions and sequences of Schur, in [15], (1977), 215-226.

18. Wiener, N., Masani, P. : The prediction theory of multivariate stochastic processes, Acta Mathematica, 98 (1957), 111-150; 99 (1958), 93-139.

Adhemar BULTHEEL
Katholieke Universiteit Leuven
Afdeling Toegepaste Wiskunde en Programmatie
Celestijnenlaan 200 A
B-3030 Heverlee (BELGIUM)

APPENDIX :



Objective function $\prod_{1}^{4}(1-|\rho_i|^2)/(1-|\gamma_i|^2)$ for the numerical example after two extractions at the origin and two extractions in $\alpha \in [-1,1]$.

The minimum 0.90398 is obtained for $\alpha = 0.5$.