

The computation of orthogonal rational functions and their interpolating properties

Adhemar Bultheel ^{*†}
Pablo González-Vera [‡]
Erik Hendriksen [§]
Olav Njåstad [¶]

May 1990

Abstract

We shall consider nested spaces \mathcal{L}_n , $n = 0, 1, 2, \dots$ of rational functions with n prescribed poles outside the unit disk of the complex plane. We study orthogonal basis functions of these spaces for a general positive measure on the unit circle. In the special case where all poles are placed at infinity, $\mathcal{L}_n = \Pi_n$, the polynomials of degree at most n . Thus the present paper is a study of orthogonal rational functions, which generalize the orthogonal Szegő polynomials. In this paper we shall concentrate on the functions of the second kind which are natural generalizations of the corresponding polynomials.

Keywords : Rational interpolation, Orthogonal functions, Szegő theory, Pick-Nevanlinna interpolation.

AMS(MOS) Classification : Primary : 42C05, Secondary : 30DF50, 41A20, 41A55

^{*}Department of Computer Science, K.U.Leuven, Belgium

[†]The work of the first author is partially supported by a research grant from the Belgian National Fund for Scientific Research

[‡]Department of Mathematical Analysis, University of La Laguna, Tenerife, Spain

[§]Department of Mathematics, University of Amsterdam, The Netherlands

[¶]Department of Mathematics, University of Trondheim-NTH, Norway

1 Introduction

The rich and fascinating theory of polynomials orthogonal on the unit circle is well known. These polynomials are named after Szegő since his pioneering work on them. His book on orthogonal polynomials [57] was first published in 1939 but the ideas were already published in several papers in the twenties. It is also in Szegő's book that the notion of a reproducing kernel is clearly introduced.

Szegő's interest in these polynomials was inspired by coefficient problems that were investigated by among others Carathéodory, Fejér [11, 12, 13], F. Riesz [50, 51], and Schur [54, 55, 56]. The Schur algorithm and the Schur coefficients were named after the continued fraction like algorithm designed by Schur to actually check if the given coefficients (moments) correspond to a bounded analytic function. It was observed later that these Schur coefficients were exactly the complex conjugates of the coefficients that appeared in the recurrence relation for the orthogonal polynomials as formulated by Szegő.

It was Pick who first considered an interpolation problem as a generalization of the coefficient problems of Carathéodory [46, 47, 48]. Independently, Nevanlinna developed the same theory in a long memoir in 1919 [43]. The Pick-Nevanlinna algorithm is a direct generalization of the Schur algorithm.

Since then, these problems played an important role in several books, like in Akhiezer [1], Kreĭn and Nudel'man [36], Walsh [58] and more recently in Donoghue [21], Garnett [27], Rosenblum and Rovnyak [52] etc.

Among the many applications of the theory, we should mention the prediction theory of stationary stochastic processes first discussed by Kolmogorov [35] and Wiener [59]. The book by Wiener contained a reprint from Levinson's celebrated paper [38], which is in fact a reformulation of the Szegő recursions. See also the classic book of Grenander en Szegő [32]. Other engineering applications are network theory (see e.g. Belevitch [3] and Youla and Saito [60]), spectral estimation (see Papoulis [45] for an excellent survey), maximum entropy analysis as formulated by Burg [10] (see the survey paper [37]), transmission lines and scattering theory as studied by Arov, Redheffer [49], Dewilde, Viera, Kailath [19] and Dewilde and Dym [17, 18], digital filtering (see the survey of Kailath [33]), speech processing (see [42] or the tutorial paper by Makhoul [41]), etc. See also the survey of Genin [28] for more applications. All of these applications can be described by the same mathematical basics.

In all these papers on theory and applications, the approach of the Pick-Nevanlinna theory from the point of view of the orthogonal functions has not always been fully put forward, although many interesting results have been obtained implicitly or explicitly. Unfortunately, except for the papers of Djrbashian (see the survey paper [20]), and more recently, [5, 7, 9, 24, 44] most of these results are buried in publications mainly written for an engineering audience. To the best of our knowledge, in none of these publications, not even in the applied papers, there has been given a direct rational generalization of the so called polynomials of the second kind, as they are known in the classical theory. It is the subject of this paper to describe these functions of the second kind. They satisfy the same recurrence relation as the orthogonal functions, but with opposite sign for the recurrence coefficients which generalize the Schur-Szegő parameters. We shall describe the recurrence, give two algorithms to compute the recursion coefficients and discuss the interpolation properties related to these functions.

We try to give in this paper an approach to the theory which is an immediate generalization of the Szegő theory for orthogonal polynomials. This theory is supposed to be related to the interpolation theory of Pick and Nevanlinna like the Szegő theory was related to the Schur and Carathéodory-Fejér coefficient problems. In fact, this paper is roughly a generalization to the rational case of the first chapter of Geronimus book [29].

Note that at any point in the discussion we can replace all the interpolation points α_k by zero and recover at any moment the corresponding result of the polynomial case. In this respect it is a natural generalization of the Szegő theory.

The outline of the paper is as follows. In section 2, we start with some general definitions

and properties from complex analysis that we shall need in the sequel of the paper. Section 3 introduces the fundamental spaces of rational functions. In section 4 we derive the Szegő type recurrence relation for the orthogonal rational functions and in the next section, the so called functions of the second kind are introduced and it is shown how they play a role in a rational interpolation problem of Pick-Nevanlinna type. The J-unitary and J-contractive matrices, defined in section 6 can be brought into the recurrence and they are important to derive properties for the interpolants like being positive real functions (i.e., Carathéodory functions). The interpolating properties that were derived before inspired the derivation in section 7 of an algorithm of the Pick-Nevanlinna type to find the recursion coefficients. In the last section, we prove a convergence result when the Blaschke product based on the interpolation points diverges.

2 Complex analysis and notations

We shall be concerned with complex function theory on the unit circle. We start with the introduction of some notation.

The complex number field is denoted by \mathbf{C} , while \mathbf{R} is the real axis. We use the following notations for the unit circle, the open unit disc and the exterior of the unit circle:

$$\mathbf{T} = \{z : |z| = 1\}, \quad \mathbf{D} = \{z : |z| < 1\}, \quad \mathbf{E} = \{z : |z| > 1\}.$$

The unit on the imaginary axis is denoted by i .

Let μ be a positive measure on \mathbf{T} , whose support is an infinite set. It is characterized by a distribution function $\int d\mu$ which has an infinite number of points of increase. The function classes $L_p(\mu)$, $0 < p \leq \infty$ are well known. The normalized Lebesgue measure is denoted by $\lambda : d\lambda = (2\pi)^{-1}d\theta$. If $\mu = \lambda$, we just write L_p instead of $L_p(\lambda)$. The inner product in $L_2(\mu)$ is denoted by

$$\langle f, g \rangle_\mu = \int f(e^{i\theta}) \overline{g(e^{i\theta})} d\mu(\theta)$$

The integration will always be over the unit circle in one form or another and we shall take the freedom to write the previous integral in different forms

$$\langle f, g \rangle_\mu = \int f \bar{g} d\mu = \int f(e^{i\theta}) \overline{g(e^{i\theta})} d\mu(\theta) = \int f(z) \overline{g(z)} d\mu(z).$$

The Hardy spaces of L_p functions analytic in \mathbf{D} are denoted by H_p . They are Banach spaces for $1 \leq p \leq \infty$.

The Nevanlinna class N is defined by

$$f \in N \iff f = g/h; \quad g, h \in H_\infty.$$

It contains all spaces H_p for $0 < p \leq \infty$ [23, p.16]. It is known that each function $f \in N$, not vanishing identically, has a nontangential limit to \mathbf{T} a.e. and $\log |f| \in L_1$ [23, p.17].

The operation of taking the complex conjugate on the unit circle is extended to the whole complex plane \mathbf{C} by the involution operation

$$f_*(z) = \overline{f(1/\bar{z})}.$$

Note that on \mathbf{T} , $f_*(z)$ is just $\overline{f(z)}$. Hence, it is obvious that $\langle f, g \rangle_\mu = \langle g_*, f_* \rangle_\mu$. One can also see immediately that if f is rational, then f_* has a pole/zero in $\alpha \in \overline{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ if and only if f has a pole/zero in $1/\bar{\alpha}$.

If μ is absolutely continuous w.r.t. the Lebesgue measure λ , then $d\mu = d\mu_a = \mu' d\lambda$, for some weight function $\mu' \in L_1$, but in general, there is also a singular part μ_s : $\mu = \mu_a + \mu_s$.

Define the *moments* as the Fourier coefficients

$$c_k = \int e^{-ik\theta} d\mu(\theta) \quad , \quad k \in \mathbf{Z}. \quad (2.1)$$

Clearly $c_{-k} = \bar{c}_k$ and $|c_k| \leq c_0$ for a real measure μ . Without loss of generality, we suppose the measure to be normalized: $c_0 = \int d\mu = 1$.

With the positive measure μ , we can associate

$$\Omega_\mu(z) = ic + \int \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(\theta) \quad , \quad c \in \mathbf{R}, \quad z \in \mathbf{D}. \quad (2.2)$$

It belongs to H_p for all $p < 1$ ([23, p.34]).

The kernel in (2.2) shall be denoted as $D(t, z)$.

$$D(t, z) = \frac{t + z}{t - z}$$

with “real part”

$$P(t, z) = \frac{1}{2}[D(t, z) + D(t, z)_*]$$

(substar conjugate for t) that reduces to the *Poisson kernel* $(1 - |z|^2)/(|t - z|^2)$ for $t \in \mathbf{T}$. Thus the real part of Ω_μ

$$\Re \Omega_\mu(z) = \int \Re \frac{t + z}{t - z} d\mu(t) = \int \frac{1 - |z|^2}{|t - z|^2} d\mu(t) \quad (2.3)$$

is positive in \mathbf{D} . Therefore we call Ω_μ a positive real function. The radial limit gives

$$\lim_{r \rightarrow 1^-} \Re \Omega_\mu(re^{i\theta}) = \mu'(e^{i\theta}) \quad \text{a.e.} \quad (2.4)$$

See [23, p.4]. The class of positive real functions, analytic in \mathbf{D} is denoted by \mathcal{P} . This class is often referred to as the Carathéodory class. Every positive real function has a *Riesz-Herglotz representation* of the form (2.2). A Cayley transform maps the right half plane into the unit circle, so that the Cayley transform of the class \mathcal{P} is the class \mathcal{B} of analytic functions in \mathbf{D} , which are bounded by 1 in modulus: the so called Schur class.

If $\log \mu' \in L_1$, then we can define a spectral factor

$$\sigma(z) = c \exp\left\{\frac{1}{2} \int D(t, z) \log \mu'(t) d\lambda(t)\right\} \quad , \quad z \in \mathbf{D}, \quad |c| = 1. \quad (2.5)$$

It is an *outer* function in H_2 . See e.g. [53]. Furthermore

$$|\sigma(z)|^2 = \exp\left\{\int P(t, z) \log \mu'(t) d\lambda(t)\right\}$$

has a radial limit which satisfies $\mu'(e^{i\theta}) = |\sigma(e^{i\theta})|^2$ a.e.

From the Szegő theory of orthogonal polynomials, we know that $1/\sigma$ vanishes $d\mu_s$ a.e. if $\log \mu' \in L_1$ [26, p. 202].

The condition $\log \mu' \in L_1$ is fundamental in the theory of Szegő for orthogonal polynomials on the unit circle. We shall therefore call it the *Szegő condition*.

Douglas, Shapiro and Shields [22] showed that a general function $f \in H_2$ has a pseudo-meromorphic extension across \mathbf{T} if f can be factored as $f = h_*/U_*$ on \mathbf{T} with $h \in H_2$ and U inner in H_2 . Again, the left hand side has an extension to \mathbf{D} and the right hand side to \mathbf{E} , which defines f in the sphere $\bar{\mathbf{C}}$.

We suppose that the spectral factor σ has such a pseudo-meromorphic extension in the sense of Douglas, Shapiro and Shields [22] (see also [17]), then the relations

$$\mu' = \sigma\sigma_* = \frac{1}{2}[\Omega + \Omega_*] \quad (2.6)$$

which hold on \mathbf{T} can be extended to the whole Riemann sphere $\bar{\mathbf{C}}$.

3 The spaces \mathcal{L}_n

In this section we shall introduce the spaces \mathcal{L}_n which are the fundamental spaces that most sections of this paper will be dealing with. Define for $\alpha_i \in \mathbf{D}$ a *Blaschke factor* as

$$\zeta_i(z) = \frac{\bar{\alpha}_i}{|\alpha_i|} \frac{\alpha_i - z}{1 - \bar{\alpha}_i z} \text{ if } \alpha_i \neq 0 \quad \text{and} \quad \zeta_i(z) = z \text{ if } \alpha_i = 0. \quad (3.1)$$

In what follows we shall use only the first notation and suppose by convention that $\bar{\alpha}_i/|\alpha_i|$ is equal to -1 for $\alpha_i = 0$. Note also the useful property $\zeta_{i*}(z) = 1/\zeta_i(z)$.

These factors appear in finite Blaschke products

$$B_0 = 1 \text{ and } B_n = B_{n-1}\zeta_n \text{ for } n \geq 1 \quad (3.2)$$

which span the fundamental spaces

$$\mathcal{L}_n = \text{span}\{B_k : k = 0, 1, \dots, n\} = \left\{ \frac{p_n}{\pi_n} : \pi_n = \prod_{i=1}^n (1 - \bar{\alpha}_i z); p_n \in \Pi_n \right\} \quad (3.3)$$

where Π_n denotes the space of polynomials of degree at most n . Thus \mathcal{L}_n is a space of rational functions with prescribed poles $1/\bar{\alpha}_i$, $i = 1, \dots, n$ which are all in \mathbf{E} .

Note that \mathcal{L}_n is isometric with Π_n in the following sense

$$\left\langle \frac{p_n}{\pi_n}, \frac{q_n}{\pi_n} \right\rangle_\mu = \langle p_n, q_n \rangle_{\mu_n}, \quad p_n, q_n \in \Pi_n$$

where $d\mu_n = |\pi_n|^{-2} d\mu$. Thus instead of studying the spaces of rational functions with respect to a fixed measure, one could also study spaces of polynomials with respect to a varying measure, an approach which was used by G. Lopéz in [39, 40].

It can be shown [58, p.227] that \mathcal{L}_n is the orthogonal complement in H_2 of the shift invariant subspace

$$\mathcal{M}_n = zB_n H_2$$

which is a consequence of Beurling's theorem [27, p.82].

If $f_n \in \mathcal{L}_n$, then we define its superstar conjugate as

$$f_n^*(z) = B_n(z) \overline{f_n(z)}.$$

Note that this transformation depends on n . It must be clear from the context what n is.

In analogy with the polynomial case where $\alpha_i = 0$, $i = 0, \dots, n$, $B_n(z) = z^n$ and $\overline{p^*(0)}$ the leading coefficient, we shall also call here $\overline{f_n^*(\alpha_n)}$ the *leading coefficient* of $f_n \in \mathcal{L}_n$.

The superstar conjugate is the direct generalization of what is used by Freud [26, p. 194], Geronimus [29, p. 6] and many others in the polynomial case. One has in general the relation

$$f_n(z) = \sum_{k=0}^n a_k B_k(z) \leftrightarrow f_n^*(z) = B_n(z) \sum_{k=0}^n \frac{\bar{a}_k}{B_k(z)}$$

so that the leading coefficient of f_n is $\overline{f_n^*(\alpha_n)} = a_n$.

We can orthonormalize the basis $\{B_k\}_k$ in the space $L_2(\mu)$ with a Gram-Schmidt procedure to find the rational basis functions which will be denoted in the sequel as $\{\phi_k\}_k$. They are uniquely defined if we require that their leading coefficient $\kappa_n = \phi_n^*(\alpha_n)$ is positive.

The function

$$k_n(z, w) = \sum_{k=0}^n \phi_k(z) \overline{\phi_k(w)}$$

is called the reproducing kernel for \mathcal{L}_n since it holds that $f_n(w) = \langle f_n(z), k_n(z, w) \rangle_\mu$ for every $f_n \in \mathcal{L}_n$.

The following properties have been proved e.g., in [8, 9]

Theorem 3.1 Let $k_n(z, w)$ be the reproducing kernel for \mathcal{L}_n which is based on the points $A_n = \{\alpha_i : i = 1, \dots, n\}$ and $\{\phi_k\}$ a set of orthonormal basis functions in $L_2(\mu)$ with leading coefficients $\kappa_k > 0$. Then

1. $k_n(z, w) = B_n(z)\overline{B_n(w)}k_n(1/\overline{w}, 1/\overline{z})$
2. $k_n(z, \alpha_n) = \kappa_n\phi_n^*(z)$
3. $k_n(\alpha_n, \alpha_n) = \kappa_n^2$
4. and the following Christoffel-Darboux relations hold

$$\begin{aligned} k_n(z, w) &= \frac{\phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(z)\overline{\phi_{n+1}(w)}}{1 - \zeta_{n+1}(z)\overline{\zeta_{n+1}(w)}} \\ &= \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - \zeta_n(z)\overline{\zeta_n(w)}\phi_n(z)\overline{\phi_n(w)}}{1 - \zeta_n(z)\overline{\zeta_n(w)}}. \end{aligned}$$

It is an easy consequence of the first Christoffel-Darboux relation that $\phi_{n+1}(z)/\phi_{n+1}^*(z) \in \mathbf{D}$ for $z \in \mathbf{D}$. You need to observe that $k_n(z, z) > 0$ and that $\zeta_{n+1}(z) \in \mathbf{D}$ for all $z \in \mathbf{D}$ to find that $|\phi_{n+1}^*|^2 - |\phi_{n+1}|^2 > 0$ in \mathbf{D} . Similarly it follows that $\phi_{n+1}(z)/\phi_{n+1}^*(z) \in \mathbf{E}$ (or \mathbf{T}) for $z \in \mathbf{E}$ (or \mathbf{T}). We can therefore conclude that $\phi_n^*(z)$ has no zeros in \mathbf{D} .

4 Recursions for the orthogonal functions

Because of the importance of the recurrence relations for the orthogonal polynomials as studied by Szegő, it is a natural question to ask whether it is possible to find such a recurrence relation also for the rational case. In [8], there were some recursions that were derived from the recursions for the reproducing kernels. The trick is that by setting $w = \alpha_n$, the reproducing kernel essentially gives (a superstar of) an orthonormal function (up to a constant). Here, we want to give a direct recursion for the orthonormal functions, which should reduce to the Szegő recursion for the polynomials if all $\alpha_k = 0$. This is done in the following theorem. Note that when all α_k are zero, this reduces to the matrix notation for the Szegő recursion for polynomials.

Theorem 4.1 For the orthonormal basis functions in \mathcal{L}_n , a recursion of the following form exists

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = N_n \frac{1 - \overline{\alpha}_{n-1}z}{1 - \overline{\alpha}_n z} \begin{bmatrix} 1 & \overline{\lambda}_n \\ \lambda_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} \quad (4.1)$$

where the matrix N_n is a constant $e_n > 0$ times a unitary matrix

$$N_n = e_n \begin{bmatrix} \eta_n^1 & 0 \\ 0 & \eta_n^2 \end{bmatrix}$$

with η_n^1 and $\eta_n^2 \in \mathbf{T}$. The constant η_n^1 is chosen such that $\kappa_n = \phi_n^*(\alpha_n) > 0$. The other constant η_n^2 is related to η_n^1 by

$$\eta_n^2 = \overline{\eta_n^1} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \frac{\overline{\alpha}_n}{|\alpha_n|}.$$

The parameter λ_n is given by

$$\lambda_n = \eta \frac{\overline{\phi_n(\alpha_{n-1})}}{\phi_n^*(\alpha_{n-1})} \in \mathbf{D} \text{ with } \eta = \frac{1 - \alpha_n \overline{\alpha}_{n-1}}{1 - \overline{\alpha}_n \alpha_{n-1}} \frac{\overline{\alpha}_n}{|\alpha_n|} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \in \mathbf{T}.$$

The value of e_n follows from the proof below, but a simpler expression is given in Theorem 4.2 below.

The formula (4.1) holds from $n = 1$ on if we define $\alpha_0 = 0$.

Proof. First we prove the existence of constants c_n and d_n such that

$$\frac{1 - \bar{\alpha}_n z}{z - \alpha_{n-1}} \phi_n - d_n \phi_{n-1} - c_n \frac{1 - \bar{\alpha}_{n-1} z}{z - \alpha_{n-1}} \phi_{n-1}^* \in \mathcal{L}_{n-2} \quad (4.2)$$

Let us define as before $\pi_k(z) = \prod_{i=1}^k (1 - \bar{\alpha}_i z)$ and the polynomials p_k are defined by $\phi_k = p_k / \pi_k$. Note that we can rewrite (4.2) as

$$\frac{p_n - d_n(z - \alpha_{n-1})p_{n-1} - c_n(1 - \bar{\alpha}_{n-1}z)p_{n-1}^* \eta_{n-1}}{(z - \alpha_{n-1})\pi_{n-1}(z)} = \frac{N(z)}{D(z)}$$

where we have used

$$\eta_k = \prod_{i=1}^k \frac{-\bar{\alpha}_i}{|\alpha_i|} \in \mathbf{T}.$$

If this has to be in \mathcal{L}_{n-2} , then we should require that $N(\alpha_{n-1}) = N(1/\bar{\alpha}_{n-1}) = 0$ or, which is the same, $N(\alpha_{n-1}) = N^*(\alpha_{n-1}) = 0$. If so, we can drop the factors $(z - \alpha_{n-1})(1 - \bar{\alpha}_{n-1}z)$ from numerator and denominator.

The first condition $N(\alpha_{n-1}) = 0$ gives (p_{n-1}^* is the reverse polynomial in the Freud sense: $p_{n-1}^*(z) = z^{n-1} \overline{p_{n-1}(1/\bar{z})} = z^{n-1} p_{n-1*}(z)$)

$$c_n = \frac{\bar{\eta}_{n-1}}{1 - |\alpha_{n-1}|^2} \frac{p_n(\alpha_{n-1})}{p_{n-1}^*(\alpha_{n-1})}.$$

The second condition $N(1/\bar{\alpha}_{n-1}) = 0$ defines d_n :

$$\bar{d}_n = \frac{1}{1 - |\alpha_{n-1}|^2} \frac{p_n^*(\alpha_{n-1})}{p_{n-1}^*(\alpha_{n-1})}.$$

Note that $p_{n-1}^*(\alpha_{n-1}) = \phi_{n-1}^*(\alpha_{n-1})\pi_{n-1}(\alpha_{n-1})\bar{\eta}_{n-1} = \kappa_{n-1}\pi_{n-1}(\alpha_{n-1})\bar{\eta}_{n-1} \neq 0$. We can therefore also write

$$c_n = \frac{1 - \bar{\alpha}_n \alpha_{n-1}}{1 - |\alpha_{n-1}|^2} \frac{\phi_n(\alpha_{n-1})}{\kappa_{n-1}}$$

and

$$\bar{d}_n = -\frac{1 - \bar{\alpha}_n \alpha_{n-1}}{1 - |\alpha_{n-1}|^2} \frac{\alpha_n}{|\alpha_n|} \frac{\phi_n^*(\alpha_{n-1})}{\kappa_{n-1}}. \quad (4.3)$$

Thus we have proved that with the previous choices of c_n and d_n , the expression in (4.2) is in \mathcal{L}_{n-2} . However, at the same time it is orthogonal to \mathcal{L}_{n-2} . To check this, we note that for every $k \leq n-2$, ϕ_k is orthogonal to the first term in (4.2) because

$$\left\langle \frac{1 - \bar{\alpha}_n z}{z - \alpha_{n-1}} \phi_n, \phi_k \right\rangle_\mu = \left\langle \phi_n, \frac{z - \alpha_n}{1 - \bar{\alpha}_{n-1} z} \phi_k \right\rangle_\mu$$

and this is zero because the right factor is in \mathcal{L}_{n-1} . ϕ_k is trivially orthogonal to the second term in (4.2). Finally, it is also orthogonal to the third term since

$$\left\langle \frac{1 - \bar{\alpha}_{n-1} z}{z - \alpha_{n-1}} \phi_{n-1}^*, \phi_k \right\rangle_\mu = \left\langle \phi_{n-1}^*, \frac{z - \alpha_{n-1}}{1 - \bar{\alpha}_{n-1} z} \phi_k \right\rangle_\mu$$

and this is zero because ϕ_j^* is orthogonal to any function from \mathcal{L}_j which vanishes at $z = \alpha_j$. We may thus conclude that the expression in (4.2) is zero. Hence

$$\begin{aligned} \phi_n &= d_n \frac{z - \alpha_{n-1}}{1 - \bar{\alpha}_n z} \phi_{n-1} + c_n \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \phi_{n-1}^* \\ &= d_n \frac{-\alpha_{n-1}}{|\alpha_{n-1}|} \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} [\zeta_{n-1}(z) \phi_{n-1} + \bar{\lambda}_n \phi_{n-1}^*] \end{aligned} \quad (4.4)$$

with

$$\lambda_n = -\frac{\bar{c}_n}{\bar{d}_n} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \quad (4.5)$$

Note that we can write λ_n as

$$\lambda_n = \eta \frac{\overline{\phi_n(\alpha_{n-1})}}{\phi_n^*(\alpha_{n-1})} \text{ with } \eta = \frac{1 - \alpha_n \bar{\alpha}_{n-1}}{1 - \bar{\alpha}_n \alpha_{n-1}} \frac{\bar{\alpha}_n}{|\alpha_n|} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \in \mathbf{T}$$

We then know from the remark at the end of the previous section that $\lambda_n \in \mathbf{D}$ as a consequence of the Christoffel-Darboux relation. Taking the superstar conjugate, we can find the recurrence as claimed. One can choose $e_n = |d_n| > 0$. The values of η_n^1 and η_n^2 can readily be computed to be

$$\eta_n^1 = -\frac{d_n}{|d_n|} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \text{ and } \eta_n^2 = -\frac{\bar{d}_n}{|d_n|} \frac{\bar{\alpha}_n}{|\alpha_n|}$$

which are indeed on \mathbf{T} .

It remains to check the initial conditions of the recurrence, i.e., for $n = 1$. Now, since $\phi_0 = \phi_0^* = 1$, we can always put

$$\phi_1(z) = e_1 \eta_1^1 \frac{1}{1 - \bar{\alpha}_1 z} [z \phi_0 + \bar{\lambda}_1 \phi_0^*]$$

where $e_1 \in \mathbf{R}$ and $\eta_1^1 \in \mathbf{T}$. Hence the constants η_1^1 and λ_1 should satisfy

$$\phi_1(0) = e_1 \eta_1^1 \bar{\lambda}_1 \text{ and } \phi_1^*(0) = -e_1 \bar{\eta}_1^1 \bar{\alpha}_1 / |\alpha_1|.$$

These can be solved for η_1^1 and λ_1 , and as you can easily check, the result corresponds to the general formula if you take $\alpha_0 = 0$ and use $\alpha_0/|\alpha_0| = -1$.

This gives the first of the two coupled recursions of (4.1). The other recurrence is found by taking the superstar conjugate of the first one. They are equivalent to each other. \square

The previous expressions for the recursion coefficients λ_n are not very practical, since they use function values of ϕ_n and ϕ_n^* to compute these. The following theorem gives, at least in principle, more reasonable expressions.

Theorem 4.2 *The recursion coefficient λ_n from the previous theorem can also be expressed as*

$$\lambda_n = \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \frac{\langle \phi_k, \frac{z - \alpha_{n-1}}{1 - \bar{\alpha}_n z} \phi_{n-1} \rangle_\mu}{\langle \phi_k, \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \phi_{n-1}^* \rangle_\mu}, \quad k = 0, 1, \dots, n-1$$

and the value of $e_n > 0$ can be obtained as the positive square root of

$$e_n^2 = \frac{1 - |\alpha_n|^2}{1 - |\alpha_{n-1}|^2} \frac{1}{1 - |\lambda_n|^2}.$$

Proof. Use the relation (4.4) for ϕ_n and express that it is orthogonal to ϕ_k . Then you get.

$$\langle \phi_k, \frac{z - \alpha_{n-1}}{1 - \bar{\alpha}_n z} \phi_{n-1} \rangle_\mu \bar{d}_n + \langle \phi_k, \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \phi_{n-1}^* \rangle_\mu \bar{c}_n = 0$$

Use then the defining relation of (4.5) and the expression for the ratio of \bar{c}_n/\bar{d}_n that you can get from the previous relation. Then you will find the expression for λ_n .

To find the expression for e_n^2 , we should prove that

$$e_n^2 (1 - |\lambda_n|^2) = \frac{1 - |\alpha_n|^2}{1 - |\alpha_{n-1}|^2}. \quad (4.6)$$

Fill in $e_n^2 = |d_n|^2$ with d_n given by (4.3) and the expression for λ_n to find

$$\begin{aligned} e_n^2(1 - |\lambda_n|^2) &= \frac{|1 - \bar{\alpha}_n \alpha_{n-1}|^2}{(1 - |\alpha_{n-1}|^2)^2} \frac{|\phi_n^*(\alpha_{n-1})|^2}{|\phi_{n-1}^*(\alpha_{n-1})|^2} \left(1 - \frac{|\phi_n(\alpha_{n-1})|^2}{|\phi_n^*(\alpha_{n-1})|^2}\right) \\ &= \frac{|1 - \bar{\alpha}_n \alpha_{n-1}|^2}{(1 - |\alpha_{n-1}|^2)^2} \frac{1}{|\phi_{n-1}^*(\alpha_{n-1})|^2} (|\phi_n^*(\alpha_{n-1})|^2 - |\phi_n(\alpha_{n-1})|^2) \\ &= \frac{|1 - \bar{\alpha}_n \alpha_{n-1}|^2}{(1 - |\alpha_{n-1}|^2)^2} (1 - |\zeta_n(\alpha_{n-1})|^2) \end{aligned}$$

where for the third line we used the Christoffel-Darboux relation. It is just a matter of writing $\zeta_n(\alpha_{n-1})$ explicitly and simplification to find that you get indeed the right hand side of (4.6). \square

Note that from a numerical point of view, the most interesting choice is probably to take $k = 0$ in the previous expression for λ_n , thus to use $\phi_0 = 1$ instead of an arbitrary ϕ_k with $0 \leq k < n$. However, further research should decide on the numerical performance of the computation of the λ_n by the formula of Theorem 4.2 and compare it with the stability of the Nevanlinna-Pick algorithm or the algorithm implicit in Theorem 7.1 to obtain these parameters. For a numerical stability analysis in the polynomial case, see e.g. [14, 6].

The presence of the η_n^1 and η_n^2 are a bit cumbersome to deal with in certain circumstances. They are needed because of our choice of the orthonormal functions to satisfy $\phi_n^*(\alpha_n) = \kappa_n > 0$. It is possible to get rid of the η 's by rotating the orthonormal functions. That is, we multiply them by some number $\epsilon_n \in \mathbf{T}$. This number can be chosen to avoid the rotations needed in the recurrence (4.1). Therefore we define

$$\epsilon_0 = 1 \text{ and } \epsilon_n = -\epsilon_{n-1} \frac{\bar{d}_n}{|d_n|} \frac{\bar{\alpha}_n}{|\alpha_n|} \text{ for } n \geq 1. \quad (4.7)$$

where d_n is as in (4.1) and use this as a rotation for ϕ_n . The rotated orthonormal functions, which are still orthonormal, will be denoted by $\Phi_n = \epsilon_n \phi_n$. These basis functions now satisfy a recurrence relation as given in the next theorem.

Theorem 4.3 *Let ϕ_n be the orthonormal functions satisfying the recurrence relation of Theorem 4.1 and denote by Φ_n the rotated orthonormal functions $\Phi_n = \epsilon_n \phi_n$ as introduced above. Then these satisfy the recurrence relation*

$$\begin{bmatrix} \Phi_n(z) \\ \Phi_n^*(z) \end{bmatrix} = e_n \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \begin{bmatrix} 1 & \bar{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix} \begin{bmatrix} Z_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Phi_{n-1}(z) \\ \Phi_{n-1}^*(z) \end{bmatrix} \quad (4.8)$$

where

$$\begin{aligned} \Lambda_n &= \bar{\epsilon}_{n-1}^2 \frac{\alpha_n}{|\alpha_n|} \frac{\bar{\alpha}_{n-1}}{|\alpha_{n-1}|} \lambda_n \\ &= \bar{\epsilon}_{n-1}^2 \frac{1 - \alpha_n \bar{\alpha}_{n-1}}{1 - \bar{\alpha}_n \alpha_{n-1}} \frac{\overline{\Phi_n(\alpha_{n-1})}}{\Phi_n^*(\alpha_{n-1})} \end{aligned}$$

and

$$Z_{n-1} = \frac{\bar{\alpha}_n}{|\alpha_n|} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \zeta_{n-1}.$$

Proof. You can start with the recurrence (4.1) and express the ϕ_n in terms of the Φ_n , which results in the relation

$$\begin{bmatrix} \Phi_n(z) \\ \Phi_n^*(z) \end{bmatrix} = e_n \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} M_n \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Phi_{n-1}(z) \\ \Phi_{n-1}^*(z) \end{bmatrix}$$

with the matrix M_n defined by

$$M_n = \begin{bmatrix} \epsilon_n \eta_n^1 \bar{\epsilon}_{n-1} & \epsilon_n \eta_n^1 \epsilon_{n-1} \bar{\lambda}_n \\ \bar{\epsilon}_n \eta_n^2 \bar{\epsilon}_{n-1} \lambda_n & \bar{\epsilon}_n \eta_n^2 \epsilon_{n-1} \end{bmatrix}.$$

Use in this matrix the definitions of ϵ_n , of η_n^i and Λ_n and some algebra will give the result. \square

In the next sections, we still go on developing the results for the ϕ_n in the first place, but virtually the same results hold true for the rotated functions Φ_n . Occasionally we shall state the result for Φ_n in a remark. The rotated functions are however important for the interpolation algorithm to be given later in section 7.

It will be useful to write an inverse form of the recursion formulas as in the next theorem.

Theorem 4.4 *Given the orthonormal function ϕ_n with $\phi_n^*(\alpha_n) = \kappa_n > 0$, all the previous orthonormal functions $\phi_k, k < n$ are uniquely defined if they are normalized by $\phi_k^*(\alpha_k) = \kappa_k > 0$. They can be found with the recursions*

$$\begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} = \frac{1}{1 - |\lambda_n|^2} \frac{1 - \bar{\alpha}_n z}{1 - \bar{\alpha}_{n-1} z} \begin{bmatrix} 1/\zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\bar{\lambda}_n \\ -\lambda_n & 1 \end{bmatrix} N_n^{-1} \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} \quad (4.9)$$

with all the quantities appearing in this formula as in Theorem 4.1.

Proof. The formula (4.9) is evidently the inverse of the recurrence formula (4.1). Since the coefficients λ_n and the matrix N_n are completely defined in terms of ϕ_n , the function ϕ_{n-1} is uniquely defined. By induction, all the previous ϕ_k are uniquely defined. \square

Note that Theorem 4.4 reduces in the polynomial case, i.e. when all $\alpha_k = 0$, to the backward recursion of the Szegő polynomials.

5 Functions of the second kind and interpolation

In this section we shall define some functions ψ_k which are the rational analogues of the polynomials of the second kind which appear in the Szegő theory. We shall call them *functions of the second kind*. They are defined first in terms of the orthogonal functions ϕ_n . We then show that they satisfy the same recurrence relation as the orthogonal functions and that they can be used to get rational approximants for the positive real function Ω_μ . To give the definitions of these functions of the second kind, we recall that $D(t, z) = (t + z)(t - z)^{-1}$, and introduce also the kernel

$$E(t, z) = D(t, z) + 1 = \frac{2t}{t - z}.$$

Note that when $t \in \mathbf{T}$, then taking substar with respect to z gives $D(t, z)_* = -D(t, z) = D(z, t)$ and thus $E(z, t) = E(t, z)_* = 1 - D(t, z)$. Here are some equivalent definitions for $z \in \mathbf{D}$

$$\psi_n(z) = \int [E(t, z)\phi_n(t) - \frac{t+z}{t-z}\phi_n(z)]d\mu(t) \quad (5.1)$$

$$\begin{aligned} &= \int D(t, z)[\phi_n(t) - \phi_n(z)]d\mu(t) + \int \phi_n(t)d\mu(t) \\ &= \begin{cases} 1 & , \text{ if } n = 0 \\ \int D(t, z)[\phi_n(t) - \phi_n(z)]d\mu(t) & , \text{ if } n \geq 1. \end{cases} \quad (5.2) \end{aligned}$$

The last equality follows from the fact that $\langle 1, \phi_n \rangle_\mu = \delta_{0n}$. These definitions are for $z \in \mathbf{D}$, but, as we show below, these functions are rational and can therefore be defined in the whole complex plane. Formula (5.2) can be compared with the formulas for the polynomials of the second kind in (1.13) of Geronimus book [29, p. 10].

We shall first show that these are functions from \mathcal{L}_n .

Lemma 5.1 *The functions ψ_n of the second kind belong to \mathcal{L}_n .*

Proof. This is trivially true for $n = 0$. For $n \geq 1$, note that the integrand in (5.2) has the form

$$[\phi_n(t) - \phi_n(z)]D(t, z).$$

The term in square brackets vanishes for $t = z$, so that the integral can be written as

$$\psi_n(z) = \int \frac{(t-z) \sum_{k=0}^n a_k(t) z^k}{(t-z)\pi_n(z)} d\mu(t) = \frac{\sum_{k=0}^n [\int a_k(t) d\mu(t)] z^k}{\pi_n(z)}$$

and this is clearly an element in \mathcal{L}_n . □

We can obtain more general expressions for these functions of the second kind as shown below.

Lemma 5.2 *To define the functions of the second kind for $n > 0$, we may replace (5.2) by*

$$\frac{\psi_n(z)}{B_k(z)} = \int D(t, z) \left[\frac{\phi_n(t)}{B_k(t)} - \frac{\phi_n(z)}{B_k(z)} \right] d\mu(t) = \int \left[E(t, z) \frac{\phi_n(t)}{B_k(t)} - D(t, z) \frac{\phi_n(z)}{B_k(z)} \right] d\mu(t) \quad (5.3)$$

for any $0 \leq k < n$. The second formula holds also for $n = 0$, if you then take $B_k = 1$.

Proof. We only consider the case $n > 0$. To prove the first or the second formula, we only have to check that

$$\int D(t, z) \left[1 - \frac{B_k(z)}{B_k(t)} \right] \phi_n(t) d\mu(t) = 0 \quad \text{or} \quad \int E(t, z) \left[1 - \frac{B_k(z)}{B_k(t)} \right] \phi_n(t) d\mu(t) = 0$$

depending on the case. The proof is the same for both of them. Since the term in square brackets vanishes for $z = t$, it follows that we can write the integral as

$$\int \frac{p(t)}{\pi_k^*(t)} \phi_n(t) d\mu(t)$$

with p a polynomial of degree at most k . The latter is of the form $\langle \phi_n, f \rangle_\mu$ with $f \in \mathcal{L}_k$. Since $k < n$ and $\phi_n \perp \mathcal{L}_{n-1}$, this is zero. □

We show next an expression for ψ_n^* .

Lemma 5.3 *The superstar conjugate of the functions of the second kind satisfy*

$$\begin{aligned} \frac{\psi_n^*(z)}{B_{n \setminus k}(z)} &= \int D(z, t) \left[\frac{\phi_n^*(t)}{B_{n \setminus k}(t)} - \frac{\phi_n^*(z)}{B_{n \setminus k}(z)} \right] d\mu(t) \\ &= \int \left[E(z, t) \frac{\phi_n^*(t)}{B_{n \setminus k}(t)} - D(z, t) \frac{\phi_n^*(z)}{B_{n \setminus k}(z)} \right] d\mu(t) \end{aligned} \quad (5.4)$$

for any $0 \leq k < n$. The second relation also holds for $n = 0$. As before, we set $B_{n \setminus k} = B_n/B_k$ for $n > 0$ and it equals 1 for $n = 0$.

Proof. Note that the second expression implies that $\psi_0^* = 1$, since we get

$$\psi_0^*(z) = \int [E(z, t) - D(z, t)] d\mu(t) = \int d\mu(t) = 1$$

which proves the second formula for $n = 0$. So, suppose that $n > 0$. Then the relations follow immediately from (5.3) in Lemma 5.2 by taking the superstar conjugate. This proves the lemma. □

Note that like in (5.3), we can give an equivalent form of (5.4) as follows.

$$-\frac{\psi_n^*(z)}{B_{n \setminus k}(z)} = \int D(t, z) \left[\frac{\phi_n^*(t)}{B_{n \setminus k}(t)} - \frac{\phi_n^*(z)}{B_{n \setminus k}(z)} \right] d\mu(t) - \int \frac{\phi_n^*(t)}{B_{n \setminus k}(t)} d\mu(t)$$

where, as we know, the last term is δ_{0n} . For $k = 0$, this takes the even simpler form

$$-\psi_{n*}(z) = \int D(t, z)[\phi_{n*}(t) - \phi_{n*}(z)]d\mu(t) - \delta_{0n}.$$

Since by definition

$$\int D(t, z)\phi_n(z)d\mu(t) = \phi_n(z) \int D(t, z)d\mu(t) = \phi_n(z)\Omega(z),$$

we can derive the following interpolation properties.

Theorem 5.4 *Let $\Omega = \Omega_\mu$ be the positive real function with Riesz-Herglotz measure μ , normalized by $\Omega(0) = 1$. Then for the functions of the second kind, it holds that ($B_{-1} = 1$)*

$$\frac{\phi_n\Omega + \psi_n}{B_{n-1}} = \begin{cases} \Omega + 1 \in H(\mathbf{D}) & , \quad n = 0 \\ g \in H(\mathbf{D}) \text{ and } g(0) = 0 & , \quad n > 0. \end{cases} \quad (5.5)$$

For their superstar conjugates, we find

$$\frac{\phi_n^*\Omega - \psi_n^*}{B_n} = \begin{cases} \Omega - 1 \in H(\mathbf{D}), \Omega(0) - 1 = 0 & , \quad n = 0 \\ h \in H(\mathbf{D}) \text{ and } h(0) = 0 & , \quad n > 0. \end{cases} \quad (5.6)$$

Proof. For $n = 0$, the relation (5.5) is obvious knowing that $\phi_0 = \psi_0 = 1$.

Use (5.3) for $k = n - 1$ and $n > 0$ to write the left hand side of (5.5) as

$$\int E(t, z) \frac{\phi_n(t)}{B_{n-1}(t)} d\mu(t) \quad (5.7)$$

where as before $E(t, z) = D(t, z) + 1$. For $z = 0$, the integral equals

$$2 \int \frac{\phi_n}{B_{n-1}} d\mu = 2\langle \phi_n, B_{n-1} \rangle_\mu = 0.$$

This implies that $\phi_n/B_{n-1} \in L_1(\mu)$. Therefore (5.7) is analytic in \mathbf{D} as a Cauchy-Stieltjes integral.

For the relation (5.6), one can similarly check the case $n = 0$ and for $n > 0$, use (5.4) with $k = 0$ to see that the left hand side equals

$$\int D(t, z) \frac{\phi_n^*(t)}{B_n(t)} d\mu(t) \quad (5.8)$$

which for $z = 0$ equals $\langle \phi_n^*, B_n \rangle_\mu$ and this is zero because $\phi_n^* \perp \zeta_n \mathcal{L}_{n-1}$. Thus $\phi_n^*/B_n \in L_1(\mu)$. We can also rewrite (5.8) as

$$\begin{aligned} \int D(t, z) \frac{\phi_n^*(t)}{B_n(t)} d\mu(t) &= \int D(t, z) \frac{\phi_n^*(t)}{B_n(t)} d\mu(t) + \int \frac{\phi_n^*(t)}{B_n(t)} d\mu(t) \\ &= \int [D(t, z) + 1] \frac{\phi_n^*(t)}{B_n(t)} d\mu(t) \\ &= \int E(t, z) \frac{\phi_n^*(t)}{B_n(t)} d\mu(t) \end{aligned}$$

and this is again analytic for z in \mathbf{D} as a Cauchy-Stieltjes integral.

This proves the theorem. □

The previous theorem should be compared with the corresponding properties in the polynomial case as given e.g. in (1.19) of Geronimus book [29, p. 11].

Like in the polynomial case, these functions of the second kind satisfy the same recurrence relations as the orthogonal functions but with opposite sign for the parameters λ_k . Taking this sign outside the transition matrix of the recursion, gives formula (5.9) as shown in the following theorem.

Theorem 5.5 For the functions of the second kind a recursion of the following form exists

$$\begin{bmatrix} \psi_n(z) \\ -\psi_n^*(z) \end{bmatrix} = N_n \frac{1 - \bar{\alpha}_{n-1}z}{1 - \bar{\alpha}_n z} \begin{bmatrix} 1 & \bar{\lambda}_n \\ \lambda_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \psi_{n-1}(z) \\ -\psi_{n-1}^*(z) \end{bmatrix} \quad (5.9)$$

where the recurrence matrix is exactly as in theorem 4.1.

Proof. As in the case of theorem 4.1, it is sufficient to prove only one of the two associated recursions. The other one follows by applying the superstar operation. We shall prove the second one. First note that by our previous lemma's we can write for $n > 1$

$$\begin{bmatrix} \psi_{n-1}(z) \\ -\psi_{n-1}^*(z) \end{bmatrix} = -\Omega(z) \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} + \int D(t, z) \begin{bmatrix} \phi_{n-1}(t) \\ \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} \phi_{n-1}^*(t) \end{bmatrix} d\mu(t).$$

Multiply from the left with

$$e_n \eta_n^2 \frac{1 - \bar{\alpha}_{n-1}z}{1 - \bar{\alpha}_n z} [\lambda_n \zeta_{n-1}(z) \quad 1]$$

then the right hand side becomes

$$-\Omega(z) \phi_n^*(z) + e_n \eta_n^2 \int D(t, z) f(t, z) d\mu(t) \quad (5.10)$$

with

$$\begin{aligned} f(t, z) &= \frac{1 - \bar{\alpha}_{n-1}z}{1 - \bar{\alpha}_n z} [\lambda_n \zeta_{n-1}(z) \phi_{n-1}(t) + \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} \phi_{n-1}^*(t)] \\ &= \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} \frac{1 - \bar{\alpha}_{n-1}z}{1 - \bar{\alpha}_n z} [\lambda_n \zeta_{n-1}(t) \phi_{n-1}(t) + \phi_{n-1}^*(t)]. \end{aligned}$$

Using the recursion for ϕ_n^* , we thus get that (5.10) can be replaced by

$$-\Omega(z) \phi_n^*(z) + \int D(t, z) \frac{\alpha_{n-1} - z}{\alpha_{n-1} - t} \frac{1 - \bar{\alpha}_n t}{1 - \bar{\alpha}_n z} \phi_n^*(t) d\mu(t).$$

This will equal $-\psi_n^*(z)$ if we may replace the latter integral by

$$\int D(t, z) \frac{\alpha_n - z}{\alpha_n - t} \frac{1 - \bar{\alpha}_n t}{1 - \bar{\alpha}_n z} \phi_n^*(t) d\mu(t).$$

This can indeed be done, since the difference equals

$$\begin{aligned} &\int D(t, z) \left[\frac{\alpha_{n-1} - z}{\alpha_{n-1} - t} - \frac{\alpha_n - z}{\alpha_n - t} \right] \frac{1 - \bar{\alpha}_n t}{1 - \bar{\alpha}_n z} \phi_n^*(t) d\mu(t) \\ &= \int f_*(t) \phi_n^*(t) d\mu(t) = \langle \phi_n^*, f \rangle_\mu \end{aligned}$$

with $f \in \zeta_n \mathcal{L}_{n-1}$. This gives zero because of the orthogonality. This proves the theorem for $n > 1$.

For $n = 1$, we have to show that $\psi_1(z) = e_1 \eta_1^1 (z - \bar{\lambda}_1) / (1 - \bar{\alpha}_1 z)$. From the definition, we get

$$\psi_1(z) = \int D(t, z) [\phi_1(t) - \phi_1(z)] d\mu(t).$$

Now we replace ϕ_1 by its expression from the recurrence relation which is $\phi_1(z) = e_1 \eta_1^1 (z + \bar{\lambda}_1) / (1 - \bar{\alpha}_1 z)$. After some computations, this results in

$$\begin{aligned} \psi_1(z) &= \frac{e_1 \eta_1^1}{1 - \bar{\alpha}_1 z} \int \frac{t + z}{t - z} \left[\frac{(t + \bar{\lambda}_1)(1 - \bar{\alpha}_1 z)}{1 - \bar{\alpha}_1 t} - (z + \bar{\lambda}_1) \right] d\mu(t) \\ &= \frac{e_1 \eta_1^1 (1 + \bar{\alpha}_1 \bar{\lambda}_1)}{1 - \bar{\alpha}_1 z} \int \frac{t + z}{1 - \bar{\alpha}_1 t} d\mu(t). \end{aligned}$$

Now we use the expression we get from theorem 4.1 for ϕ_1 in terms of λ_1 and use the orthogonality relation $\langle \phi_1, 1 \rangle_\mu = 0$ to find

$$\int \frac{t}{1 - \bar{\alpha}_1 t} d\mu(t) = -\bar{\lambda}_1 \int \frac{1}{1 - \bar{\alpha}_1 t} d\mu(t).$$

Fill this into the last expression and you find

$$\psi_1(z) = e_1 \eta_1^1 \frac{z - \bar{\lambda}_1}{1 - \bar{\alpha}_1 z} (1 + \bar{\alpha}_1 \bar{\lambda}_1) \int \frac{1}{1 - \bar{\alpha}_1 t} d\mu(t). \quad (5.11)$$

We have to find an expression for the remaining integral. Therefore we use again the expression for λ_1 from theorem 4.2 to get

$$\begin{aligned} \int \frac{d\mu(t)}{1 - \bar{\alpha}_1 t} &= 1 + \bar{\alpha}_1 \int \frac{td\mu(t)}{1 - \bar{\alpha}_1 t} \\ &= 1 - \bar{\alpha}_1 \bar{\lambda}_1 \int \frac{d\mu(t)}{1 - \bar{\alpha}_1 t}. \end{aligned}$$

From this relation we finally get

$$(1 + \bar{\alpha}_1 \bar{\lambda}_1) \int \frac{d\mu(t)}{1 - \bar{\alpha}_1 t} = 1.$$

Now the recursion for ψ_1 is proved and this concludes the proof of the theorem. \square

The interpolation properties cause the following theorem to be true.

Theorem 5.6 *Let ϕ_n be the orthonormal functions of \mathcal{L}_n with respect to the measure μ . Define the absolutely continuous measure μ_n by $d\mu_n(t) = P(t, \alpha_n) |\phi_n(t)|^{-2} d\lambda(t)$ where P is the Poisson kernel. Then on \mathcal{L}_n , the inner product with respect to μ_n and μ is the same : $\langle \cdot, \cdot \rangle_\mu = \langle \cdot, \cdot \rangle_{\mu_n}$.*

Proof. We prove first that the norm of ϕ_n is the same. $\|\phi_n\|_{\mu_n}^2 = \int P(t, \alpha_n) |\phi_n|^2 / |\phi_n|^2 d\lambda = 1 = \|\phi_n\|_\mu^2$.

Next we show that $\langle \phi_n, \phi_k \rangle_\mu$ and $\langle \phi_n, \phi_k \rangle_{\mu_n}$ is the same for $k < n$. They are both zero.

$$\begin{aligned} \langle \phi_n, \phi_k \rangle_{\mu_n} &= \int \frac{\phi_{k^*}(t)}{\phi_{n^*}(t)} P(t, \alpha_n) d\lambda(t) \\ &= \int \frac{\phi_k^*(t) B_{n \setminus k}(t)}{\phi_n^*(t)} P(t, \alpha_n) d\lambda(t) \end{aligned}$$

Since ϕ_n^* has its zeros in \mathbf{E} , we know that $B_{n \setminus k} \phi_k^* / \phi_n^*$ is analytic in the closed unit disk and then we may apply Poisson's formula which gives zero because $B_{n \setminus k}(\alpha_n) = 0$. Of course also $\langle \phi_n, \phi_k \rangle_\mu = 0$. Hence ϕ_n is a function of norm 1 and orthogonal to \mathcal{L}_{n-1} both with respect to μ_n and with respect to μ . By theorem 4.4 this ϕ_n will uniquely define all the previous ϕ_k , provided they are normalized properly with $\phi_k^*(\alpha_k) = \kappa_k > 0$. Thus the orthonormal system in \mathcal{L}_n for μ_n and for μ is the same : $\langle \phi_k, \phi_i \rangle_\mu = \langle \phi_k, \phi_i \rangle_{\mu_n} = \delta_{ki}$. Since every element from \mathcal{L}_n can be expressed as a linear combination of the ϕ_k , it also holds that $\langle f, g \rangle_\mu = \langle f, g \rangle_{\mu_n}$ for every f and $g \in \mathcal{L}_n$. \square

The previous theorem was proved for orthogonal polynomials e.g. in [26, p. 198–199].

6 J-unitary, J-contractive matrices

We shall now derive some determinant formula and some other properties of the functions we introduced. Therefore we need a J-unitary matrix which is J-contractive in \mathbf{D} .

We give the following definitions. Let θ be a 2×2 matrix with entries that are functions in the Nevanlinna class N : $\theta = [\theta_{ij}] \in N^{2 \times 2}$. We consider such matrices that are unitary with respect to the indefinite metric

$$J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = 1 \oplus -1.$$

We mean that they satisfy

$$\theta^H J \theta = J \quad \text{on } \mathbf{T} \quad (6.1)$$

where the superscript H denotes complex conjugate transpose. If we define the substar conjugate for matrices as the elementwise substar conjugate of the transposed matrix :

$$\begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}_* = \begin{bmatrix} \theta_{11*} & \theta_{21*} \\ \theta_{12*} & \theta_{22*} \end{bmatrix},$$

then we can write (6.1) as

$$\theta_* J \theta = J \quad (6.2)$$

which is valid a.e. in the whole Riemann sphere $\overline{\mathbf{C}}$ since these θ matrices have a pseudomero-morphic extension across the unit circle. Note that the set of J-unitary matrices is closed under multiplication. We denote the class of J-unitary matrices by

$$\mathbf{T}_J = \{\theta \in N^{2 \times 2} : \theta_* J \theta = J \text{ a.e. in } \overline{\mathbf{C}}\}$$

The class of J-contractive matrices is denoted as

$$\mathbf{D}_J = \{\theta \in N^{2 \times 2} : \theta^H J \theta \leq J \text{ a.e. in } \mathbf{D}\}$$

By the inequality, we mean that $J - \theta^H J \theta$, is positive semi definite. Also the set of J-contractive matrices is closed under multiplication.

The class of J-unitary, J-contractive matrices, which is denoted by $\mathbf{B}_J = \mathbf{T}_J \cap \mathbf{D}_J$, could be called the class of (matrix valued) J-inner functions as in [24].

For these J-inner matrices a number of properties can be proved. The following theorem is due to Dewilde and Dym [17, p.448].

Theorem 6.1 For $\theta = [\theta_{ij}] \in \mathbf{B}_J$ the following holds.

1. $\theta^H \in \mathbf{B}_J$.
2. $\theta^H J \theta \geq J$ a.e. in \mathbf{E} .
3. $(\theta_{11} + \theta_{12})_*^{-1} \in H_2$.
4. $(\theta_{11} + \theta_{12})_*^{-1}(\theta_{11} - \theta_{12})_* \in \mathcal{P}$.
5. $(\theta_{22} + \theta_{21})^{-1} \in H_2$.
6. $(\theta_{22} + \theta_{21})^{-1}(\theta_{22} - \theta_{21}) \in \mathcal{P}$.
7. $(\theta_{11} + \theta_{12})_*^{-1}(\theta_{21} - \theta_{22})_*$ is inner.

The matrix having the properties discussed above will be the Θ_n matrix that is obtained in the next lemma.

Lemma 6.2 Let t_n denote the recursion matrix

$$t_n = N_n \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \begin{bmatrix} 1 & \bar{\lambda}_n \\ \lambda_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix}$$

with all the parameters as defined in theorems 4.1 and 5.5. Set $T_n = t_n t_{n-1} \cdots t_1$ (recall $\alpha_0 = 0$). Then

$$T_n = \frac{1}{2} \begin{bmatrix} \phi_n + \psi_n & \phi_n - \psi_n \\ \phi_n^* - \psi_n^* & \phi_n^* + \psi_n^* \end{bmatrix}. \quad (6.3)$$

There exists a positive constant c_n such that

$$\Theta_n = \frac{1 - \bar{\alpha}_n z}{c_n} T_n$$

is a J -unitary matrix which is J -contractive in \mathbf{D} .

Proof. The first relation follows easily from

$$\begin{bmatrix} \psi_n & \phi_n \\ -\psi_n^* & \phi_n^* \end{bmatrix} = T_n \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (6.4)$$

by solving for T_n .

Now note that t_k can be written as

$$|d_k| (1 - |\lambda_k|^2)^{1/2} \frac{1 - \bar{\alpha}_{k-1} z}{1 - \bar{\alpha}_k z} \theta_k$$

with

$$\theta_k = \begin{bmatrix} \eta_k^1 & 0 \\ 0 & \eta_k^2 \end{bmatrix} (1 - |\lambda_k|^2)^{-1/2} \begin{bmatrix} 1 & \bar{\lambda}_k \\ \lambda_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_{k-1} & 0 \\ 0 & 1 \end{bmatrix},$$

a J -unitary matrix, which is also J -contractive in \mathbf{D} since $|\lambda_k| < 1$. Multiply this out to find $\Theta_n = \theta_n \theta_{n-1} \cdots \theta_1$ and $c_n = \prod_{k=1}^n |d_k| (1 - |\lambda_k|^2)^{1/2}$. \square

Recall that the Poisson kernel is $P(z, w) = (1 - |w|^2)/|z - w|^2$, $z \in \mathbf{T}$. We shall generalize this for values $z \notin \mathbf{T}$ and we shall mean then

$$P(z, w) = \frac{1 - |w|^2}{(z - w)(z - w)_*} = \frac{z(1 - |w|^2)}{(z - w)(1 - \bar{w}z)}$$

and call this the generalized Poisson kernel.

With the previous result, we can now prove the following theorem.

Theorem 6.3 *With the notation introduced in the previous lemma, we have :*

1. the determinant formula $\frac{1}{2}[\psi_n \phi_{n*} + \psi_{n*} \phi_n] = P(z, \alpha_n)$ whence

$$\frac{1}{2} \begin{bmatrix} \psi_n & \psi_n^* \\ \phi_n & \phi_n^* \end{bmatrix} = \frac{P(z, \alpha_n)}{\phi_n \phi_{n*}}$$

with $P(z, w)$ the (generalized) Poisson kernel.

2. $\psi_n^*/\phi_n^* = \psi_{n*}/\phi_{n*} \in \mathcal{P}$. The Riesz-Herglotz measure for this positive real function is the one given in:

$$\frac{\psi_n^*(z)}{\phi_n^*(z)} = \int D(t, z) d\mu_n(t) \text{ with } d\mu_n(t) = \frac{P(t, \alpha_n)}{|\phi_n^*|^2} d\lambda(t).$$

Proof. The first determinant relation follows by taking the determinant of (6.4), giving

$$\begin{aligned} \frac{1}{2} [\psi_n \phi_n^* + \psi_n^* \phi_n] &= \frac{c_n^2}{(1 - \bar{\alpha}_n z)^2} \det \Theta_n \\ &= \frac{z}{(1 - \bar{\alpha}_n z)^2} B_{n-1}(z) \prod_{k=1}^n \frac{1 - |\alpha_k|^2}{1 - |\alpha_{k-1}|^2} \frac{\alpha_{k-1}}{|\alpha_{k-1}|} \frac{\bar{\alpha}_k}{|\alpha_k|} \\ &= \frac{z}{(1 - \bar{\alpha}_n z)(z - \alpha_n)} B_n(z) (1 - |\alpha_n|^2) \\ &= B_n(z) P(z, \alpha_n) \end{aligned}$$

That $\psi_n^*/\phi_n^* \in \mathcal{P}$ follows from theorem 6.1 because Θ_n is J-inner and the factor $c_n/(1 - \bar{\alpha}_n z)$ relating Θ_n and T_n drops out of the ratio.

If $\Omega_n = \psi_n^*/\phi_n^* \in \mathcal{P}$, then $\Re\Omega_n = P(z, \alpha_n)/|\phi_n|^2$, by the first part. Hence the Riesz-Herglotz representation has the form

$$\Omega_n(z) = \int D(t, z) \Re\Omega_n(t) d\lambda(t).$$

The theorem is completely proved. \square

Note that the previous theorem contains generalizations of formulas (1.17) and (1.18) of Geronimus book [29, p. 11].

As with the functions ϕ_n , we could rotate the functions of the second kind to give $\Psi_n = \epsilon_n \psi_n$ where the ϵ_n are as defined in (4.7). For these rotated Ψ_n a recurrence like in theorem 4.3 exists. Most of the properties of ψ_n are transferred to Ψ_n .

7 Interpolation algorithm for the orthonormal functions

In this section we give an algorithm in the style of the Pick-Nevalinna algorithm, which, based on the idea of successive interpolation, will generate the recursion for the orthonormal functions ϕ_n and the functions of the second kind ψ_n . As a matter of fact, it is difficult to do this for these functions because of the rotating factors η_n^1 and η_n^2 in the recurrence relation. These rotations depend on the angle that $\phi_n(\alpha_{n-1})$ forms with the real axis and this is difficult to find without evaluating $\phi_n(\alpha_{n-1})$. However, the rotated functions Φ_n and Ψ_n satisfied a recurrence that got rid of these η 's and it will be possible to find an interpolation algorithm for these rotated functions. That is what we shall currently do.

Recall that $\Phi_n = \epsilon_n \phi_n$ and $\Psi_n = \epsilon_n \psi_n$ with ϵ_n as defined by (4.7). Define R_{n1} and R_{n2} through

$$\begin{bmatrix} B_{n-1}R_{n1}(z) \\ B_n R_{n2}(z) \end{bmatrix} = \begin{bmatrix} \Phi_n(z) \\ \Phi_n^*(z) \end{bmatrix} \Omega(z) + \begin{bmatrix} \Psi_n(z) \\ -\Psi_n^*(z) \end{bmatrix}. \quad (7.1)$$

Where $\Omega = \Omega_\mu$ is the positive real function associated with the measure μ for which the orthogonality holds. Note that the functions in the left hand side are in fact rotated versions of the functions g and h as defined in (5.5) and (5.6) respectively. These are indeed the remainders in the linearized interpolation properties of the rotated functions. We shall call the functions R_{n1} and R_{n2} the remainder functions. The factor B_{-1} has to be understood as 1 and thus $R_{01} = \Omega + 1$ while $R_{02} = \Omega - 1$. For $n > 0$, both R_{n1} and R_{n2} are zero in the origin.

The right hand side in the defining relation (7.1) of the remainder functions satisfies the recurrence for the rotated functions as in theorem 4.3. Hence, also the left hand side shall satisfy

$$\begin{bmatrix} B_{n-1}R_{n1}(z) \\ B_n R_{n2}(z) \end{bmatrix} = e_n \frac{1 - \bar{\alpha}_{n-1}z}{1 - \bar{\alpha}_n z} \begin{bmatrix} 1 & \bar{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix} \begin{bmatrix} Z_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B_{n-2}R_{n-1,1}(z) \\ B_{n-1}R_{n-1,2}(z) \end{bmatrix}. \quad (7.2)$$

This can be rewritten as given in the next theorem.

Theorem 7.1 *The remainder functions as defined above satisfy the following recursion*

$$(1 - \bar{\alpha}_n z) \begin{bmatrix} R_{n1}(z) \\ R_{n2}(z) \end{bmatrix} = e_n \begin{bmatrix} 1 & 0 \\ 0 & 1/\zeta_n(z) \end{bmatrix} \begin{bmatrix} 1 & \bar{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix} (1 - \bar{\alpha}_{n-1}z) \begin{bmatrix} \eta_{n-1}R_{n-1,1}(z) \\ R_{n-1,2}(z) \end{bmatrix} \quad (7.3)$$

with $e_n > 0$ and

$$\Lambda_n = -\bar{\eta}_{n-1} \lim_{z \rightarrow \alpha_n} \frac{R_{n-1,2}(z)}{R_{n-1,1}(z)}, \quad \eta_{n-1} = \frac{\bar{\alpha}_n}{|\alpha_n|} \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \text{ and } e_n^2 = \frac{1 - |\alpha_n|^2}{1 - |\alpha_{n-1}|^2} \frac{1}{1 - |\Lambda_n|^2}. \quad (7.4)$$

The Λ_n in the previous expression are the same as the Λ_n of theorem 4.3.

We can make the recursion even simpler and avoid the explicit use of the η_{n-1} by introducing

$$r_{n1}(z) = \frac{\bar{\alpha}_n}{|\alpha_n|} R_{n1}(z) \text{ and } r_{n2}(z) = R_{n2}(z). \quad (7.5)$$

With this notation, the recursion (7.3) becomes

$$(1 - \bar{\alpha}_n z) \begin{bmatrix} r_{n1}(z) \\ r_{n2}(z) \end{bmatrix} = e_n \begin{bmatrix} 1 & 0 \\ 0 & 1/\zeta_n(z) \end{bmatrix} \begin{bmatrix} 1 & \bar{L}_n \\ L_n & 1 \end{bmatrix} (1 - \bar{\alpha}_{n-1} z) \begin{bmatrix} r_{n-1,1}(z) \\ r_{n-1,2}(z) \end{bmatrix} \quad (7.6)$$

with

$$L_n = \frac{\alpha_n}{|\alpha_n|} \Lambda_n = - \lim_{z \rightarrow \alpha_n} \frac{r_{n-1,2}(z)}{r_{n-1,1}(z)}, \quad e_n = \left[\frac{1 - |\alpha_n|^2}{1 - |\alpha_{n-1}|^2} \frac{1}{1 - |L_n|^2} \right]^{1/2}. \quad (7.7)$$

Proof. We shall only prove (7.3), because (7.6) is a direct consequence. We can start from the relation (7.2) and use $Z_{n-1} = \eta_{n-1} \zeta_{n-1}$ to get

$$B_{n-1}(z) \begin{bmatrix} R_{n1}(z) \\ \zeta_n R_{n2}(z) \end{bmatrix} = e_n \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \begin{bmatrix} 1 & \bar{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix} B_{n-1}(z) \begin{bmatrix} \eta_{n-1} R_{n-1,1}(z) \\ R_{n-1,2}(z) \end{bmatrix} \quad (7.8)$$

which now easily gives (7.3). To find the expression for Λ_n , you can use the last line of (7.8) for $z = \alpha_n$ which gives

$$0 = \Lambda_n \eta_{n-1} R_{n-1,1}(\alpha_n) + R_{n-1,2}(\alpha_n)$$

from which the expression for Λ_n follows. The expression for e_n was shown in theorem 4.2. \square

The previous theorem has the following consequence formulated in terms of functions from the Schur class \mathcal{B} .

Corollary 7.2 Define the function $\Gamma_n(z)$ in terms of the remainder functions by

$$\Gamma_n(z) = \frac{\alpha_n}{|\alpha_n|} \frac{R_{n2}(z)}{R_{n1}(z)} = \frac{r_{n2}(z)}{r_{n1}(z)}. \quad (7.9)$$

Then $\Gamma_0 = (1 - \Omega)/(1 + \Omega)$ and for all $k \geq 0$, $\Gamma_k \in \mathcal{B}$ and $\Gamma_k(0) = 0$ and they are generated by

$$\Gamma_n = \frac{1}{\zeta_n} \left(\frac{L_n + \Gamma_{n-1}}{1 + \bar{L}_n \Gamma_{n-1}} \right)$$

with $L_n = -\Gamma_{n-1}(\alpha_n)$.

Proof. This follows immediately from the previous theorem. All the Γ_k are in \mathcal{B} because Γ_0 is, while the Möbius transforms are done with $L_k \in \mathbf{D}$ and this maps \mathcal{B} into itself. Moreover the division by ζ_n respects the analyticity because the function between brackets was made zero in $z = \alpha_n$ by the choice of L_n . \square

8 Convergence results

This section includes a result about convergence uniformly on compact subsets of the unit disk of the approximants we obtained. It is a well known fact that an infinite Blaschke product $B(z) = B_\infty(z)$ will converge to zero uniformly on compact subsets of the unit disk if $\sum_{k=1}^{\infty} (1 - |\alpha_k|) = \infty$. See e.g., [58, p.281 ff]. This can be used to obtain some other convergence results of the same type.

Theorem 8.1 Let ϕ_n be the orthonormal functions for \mathcal{L}_n and ψ_n the functions of the second kind. Define $\Omega_n = \psi_n^*/\phi_n^* \in \mathcal{P}$ and let the positive real function Ω be associated with the measure μ as in (2.2) ($c = 0$). Then Ω_n converges to Ω uniformly on compact subsets of \mathbf{D} if $\sum_1^\infty (1 - |\alpha_n|) = \infty$.

Proof. First, we note that $\Omega_n = \psi_n^*/\phi_n^* = \Psi_n^*/\Phi_n^*$, where Φ_n and Ψ_n are the rotated functions as in section 7.

Let T_n be the recurrence matrix for the rotated functions, i.e., $T_n = t_n t_{n-1} \dots t_1$ with t_n the elementary recurrence matrices as in (4.8). Then

$$T_n = \frac{1}{2} \begin{bmatrix} \Phi_n + \Psi_n & \Phi_n - \Psi_n \\ \Phi_n^* - \Psi_n^* & \Phi_n^* + \Psi_n^* \end{bmatrix}.$$

Hence $\Phi_n^*[1 - \Omega_n \ 1 + \Omega_n] = 2[0 \ 1]T_n$ and thus

$$\begin{aligned} [1 - \Omega_n \ 1 + \Omega_n] \begin{bmatrix} \Omega + 1 \\ \Omega - 1 \end{bmatrix} &= \frac{2}{\Phi_n^*} [0 \ 1] T_n \begin{bmatrix} \Omega + 1 \\ \Omega - 1 \end{bmatrix} \\ &= \frac{2}{\Phi_n^*} [0 \ 1] \begin{bmatrix} B_{n-1} R_{n1} \\ B_n R_{n2} \end{bmatrix} = 2 \frac{B_n}{\Phi_n^*} R_{n2} \\ &= 2(\Omega - \Omega_n) \end{aligned}$$

Thus

$$\Omega - \Omega_n = B_n R_{n2} / \Phi_n^* \quad \text{in } \mathbf{D}. \quad (8.1)$$

Now define the Schur functions by Cayley transforms of Ω and Ω_n

$$\Gamma = \frac{\Omega - 1}{\Omega + 1} \in \mathcal{B} \quad \text{and} \quad \Gamma_n = \frac{\Omega_n - 1}{\Omega_n + 1} \in \mathcal{B}.$$

Then,

$$\Gamma - \Gamma_n = 2 \frac{\Omega - \Omega_n}{(1 + \Omega)(1 + \Omega_n)},$$

which, in view of (8.1) gives

$$\frac{\Gamma - \Gamma_n}{B_n} = \frac{R_{n2}}{\Phi_n^*(1 + \Omega)(1 + \Omega_n)} \in H(\mathbf{D}).$$

On the unit circle, we now get $|B_n| = 1$ a.e. and $|\Gamma| \leq 1$ and $|\Gamma_n| \leq 1$, so that

$$\left| \frac{\Gamma - \Gamma_n}{B_n} \right| \leq 2 \quad \text{on } \mathbf{T}.$$

The maximum modulus theorem then gives

$$|\Gamma - \Gamma_n| \leq 2|B_n| \quad \text{in } \mathbf{D}.$$

The right hand side, and hence also the left hand side converges to zero uniformly on compact subsets of \mathbf{D} if $\sum_1^\infty (1 - |\alpha_n|) = \infty$. With inverse Cayley transforms we now find that

$$\Omega - \Omega_n = \frac{1 + \Gamma}{1 - \Gamma} - \frac{1 + \Gamma_n}{1 - \Gamma_n} = 2 \frac{\Gamma - \Gamma_n}{(1 - \Gamma)(1 - \Gamma_n)}$$

which will converge exactly as $\Gamma - \Gamma_n$ does. □

9 Conclusion

We have discussed some results about orthogonal rational functions related to the Nevanlinna-Pick interpolation algorithm exactly like the Szegő theory of orthogonal polynomials is related to the Schur algorithm. We preassumed that the Szegő condition $\log \mu' \in L_1$ was satisfied. Some ideas about the theory without this condition can be found in [18].

Essentially by using a Cayley transformation, it is possible to translate these results to a set of orthogonal rational function with respect to a measure on the real axis and where the poles are for example in the lower half plane.

We did not discuss the possibility to implement so called superfast algorithms as for the polynomial case like e.g., in [4, 2], nor did we give parallel implementations like discussed in several papers in the proceedings [31] or “split” versions similar to the split Levinson and Schur recursions as introduced by Genin [15, 16]. Also a vast literature concerning the generalization to matrix or operator valued functions should be mentioned here, e.g., [52, 25, 5].

References

- [1] N.I. Akhiezer. *The classical moment problem*. Oliver and Boyd, Edinburgh, 1969. Originally published Moscow, 1961.
- [2] G.S. Ammar and W.B. Gragg. The generalized Schur algorithm for the superfast solution of Toeplitz systems. In J. Gilewicz, M. Pindor, and W. Siemaszko, editors, *Rational approximation and its applications in mathematics and physics*, volume 1237 of *Lect. Notes Math.*, pages 315–330, Berlin, 1987. Springer-Verlag.
- [3] V. Belevitch. *Classical network theory*. Holden-Day, San Francisco, 1968.
- [4] R.P. Brent, F.G. Gustavson, and D.Y. Yun. Fast solution of Toeplitz systems of equations and computation of Padé approximants. *J. of Algorithms*, 1:259–295, 1980.
- [5] A. Bultheel. Orthogonal matrix functions related to the multivariable Nevanlinna-Pick problem. *Tijdschr. Belgisch Wisk. Genootschap, ser. B*, 32(2):149–170, 1980.
- [6] A. Bultheel. Error analysis of incoming and outgoing schemes for the trigonometric moment problem. In M.G. de Bruin and H. van Rossum, editors, *Padé approximation and its applications, Amsterdam 1980*, volume 888 of *LNM*, pages 100–109, Berlin, 1980 1981. Springer-Verlag.
- [7] A. Bultheel. On a special Laurent-Hermite interpolation problem. In L. Collatz, G. Meinardus and H. Werner, editor, *Numerische methoden der Approximationstheorie 6*, pages 63–79, Basel, 1981. Birkhäuser Verlag. ISNM Vol. 59.
- [8] A. Bultheel and P. Dewilde. Orthogonal functions related to the Nevanlinna-Pick problem. In P. Dewilde, editor, *Mathematical Theory of Networks and Systems*, pages 207–211, North Hollywood, 1979. Western Periodicals. Proceedings MTNS Conference, Delft, The Netherlands.
- [9] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. A Szegő theory for rational functions. Technical Report TW-131, K. U. Leuven, Dept. of Computer Science, May 1990.
- [10] J.P. Burg. Maximum entropy spectral analysis. In D.G. Childers, editor, *Modern spectral analysis*, pages 34–39, New York, 1978. IEEE Press. Originally presented at 37th Meet. Soc. Exploration Geophysicists, 1967.
- [11] C. Carathéodory. Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen. *Math. Annal.*, 64:95–115, 1907.
- [12] C. Carathéodory. Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. *Rend. Circ. Mat. Palermo*, 32:193–217, 1911.
- [13] C. Carathéodory and L. Fejér. Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz. *Rend. Circ. Mat. Palermo*, 32:218–239, 1911.

- [14] G. Cybenko. The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations. *SIAM J. Sci. Stat. Comp.*, 1:303–319, 1980.
- [15] P. Delsarte and Y. Genin. An introduction to the class of split Levinson algorithms. In Golub and Van Dooren [31], pages 112–130.
- [16] P. Delsarte and Y. Genin. On the split approach based algorithms for DSP problems. In Golub and Van Dooren [31], pages 131–147.
- [17] P. Dewilde and H. Dym. Schur recursions, error formulas, and convergence of rational estimators for stationary stochastic sequences. *IEEE Trans. on Information Theory*, IT-27:446–461, 1981.
- [18] P. Dewilde and H. Dym. Lossless inverse scattering, digital filters, and estimation theory. *IEEE Trans. on Information Theory*, IT-30:644–662, 1984.
- [19] P. Dewilde, A. Viera, and T. Kailath. On a generalized Szegő-Levinson realization algorithm for optimal linear predictors based on a network synthesis approach. *IEEE Trans. on Circuit and Systems*, 25:663–675, 1978.
- [20] M.M. Džrbashian. A survey on the theory of orthogonal systems and some open problems. In P. Nevai, editor, *Orthogonal polynomials: Theory and practice*, volume 294 of *Series C: Mathematical and Physical Sciences*, pages 135–146, Boston, 1990. NATO-ASI, Kluwer Academic Publishers.
- [21] W.F. Donoghue Jr. *Monotone matrix functions and analytic continuation*. Springer Verlag, Berlin, 1974.
- [22] R.G. Douglas, H.S. Shapiro, and A.L. Shields. Cyclic vectors and invariant subspaces for the backward shift operator. *Ann. Inst. Fourier*, 20:37–76, 1970.
- [23] P.L. Duren. *The theory of H^p spaces*, volume 38 of *Pure and Applied Mathematics*. Academic Press, New York, 1970.
- [24] H. Dym. *J contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*, volume 71 of *CBMS regional conference series in mathematics*. American Mathematical Society, Providence, Rhode Island, 1989.
- [25] C. Foias and A. Frazho. *The commutant lifting approach to interpolation problems*, volume 44 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1990.
- [26] G. Freud. *Orthogonal polynomials*. Pergamon Press, Oxford, 1971.
- [27] J. B. Garnett. *Bounded analytic functions*. Academic Press, New York, 1981.
- [28] Y. Genin. An introduction to the modern theory of positive functions and some of its applications to signal processing, circuits and system problems. In *Proc. Europ. Conf. Circuit Th. Design.*, pages 195–234, Paris, 1987.
- [29] Ya. Geronimus. *Orthogonal polynomials*. Consultants Bureau, New York, 1961.
- [30] I. Gohberg, editor. *I. Schur methods in operator theory and signal processing*, volume 18 of *Operator Theory: advances and applications*. Birkhäuser, Basel, 1986.
- [31] G. Golub and P. Van Dooren, editors. *Numerical linear algebra, digital signal processing and parallel algorithms*, volume 70 of *NATO-ASI Series, F: Computer and Systems Sciences*. Springer-Verlag, Berlin, 1991.

- [32] Ya.L. Grenander and G. Szegő. *Toeplitz forms and their applications*. University of California Press, Berkley, 1958.
- [33] T. Kailath. A view of three decades of linear filtering theory. *IEEE Trans. Inf. Th.*, IT-20:146–181, 1974. Reprinted in [34], 10-45.
- [34] T. Kailath et al., editor. *Linear least-squares estimation*, volume 17 of *Benchmark papers in electrical engineering and computer science*. Dowden, Hutchinson and Ross, Stroudsburg, Pa., 1977.
- [35] A.N. Kolmogorov. Stationary sequences in Hilbert's space. *Bull. Moscow State Univ.*, 2(6):1–40, 1940. Reprinted in [34], 66-89.
- [36] M.G. Kreĭn and A.A. Nudel'man. *The Markov moment problem and extremal problems*, volume 50 of *Transl. Math. Monographs*. American Mathematical Society, Providence, Rhode Island, 1977.
- [37] H.J. Landau. Maximum entropy and the moment problem. *Bull. Am. Math. Soc. (New Series)*, 16(1):47–77, 1987.
- [38] N. Levinson. The Wiener rms (root mean square) error criterion in filter design and prediction. *J. Math. Phys.*, 25:261–278, 1947.
- [39] G.L. López. On the asymptotics of the ratio of orthogonal polynomials and convergence of multipoint Padé approximants. *Math. USSR Sbornik*, 56:207–219, 1985.
- [40] G.L. López. Asymptotics of polynomials orthogonal with respect to varying measures. *Constructive Approximation*, 5:199–219, 1989.
- [41] J. Makhoul. Linear prediction : a tutorial review. *Proc. IEEE*, 63:561–580, 1975.
- [42] J.D. Markel and A.H. Gray Jr. *Linear prediction of speech*. Springer Verlag, New York, 1976.
- [43] R. Nevanlinna. Über beschränkte Funktionen die in gegebenen Punkten vorgeschriebene werte annehmen. *Ann. Acad. Sci. Fenn. Ser. A.*, 13(1):71pp., 1919.
- [44] O. Njåstad. Multipoint Padé approximation and orthogonal rational functions. In A. Cuyt, editor, *Nonlinear numerical methods and rational approximation*, pages 258–270, Dordrecht, 1988. D. Reidel Publ. Comp.
- [45] A. Papoulis. Levinson's algorithm, Wold decomposition, and spectral estimation. *SIAM Review*, 27(3):405–441, 1985.
- [46] G. Pick. Über die Beschränkungen analytischen Funktionen Welche durch vorgegebene Funktionswerte bewirkt werden. *Math. Ann.*, 77:7–23, 1916.
- [47] G. Pick. Über die Beschränkungen analytischen Funktionen durch vorgegebene Funktionswerte. *Math. Ann.*, 78:270–275, 1918.
- [48] G. Pick. Über beschränkte Funktionen mit vorgeschriebenen Wertzuordnungen. *Ann. Acad. Sci. Fenn. Ser. A*, 15(3):17pp., 1920.
- [49] R. Redheffer. On the relation of transmission-line theory to scattering and transfer. *J. Math. Phys.*, 41:1–41, 1962.
- [50] F. Riesz. Über ein Problem des Herrn Carathéodory. *J. Reine Angew. Math.*, 146:83–87, 1916.
- [51] F. Riesz. Über Potenzreihen mit vorgeschriebenen Anfangsgliedern. *Acta Math.*, 42:145–171, 1918.

- [52] M. Rosenblum and J. Rovnyak. *Hardy classes and operator theory*. Oxford University Press, New York, 1985.
- [53] W. Rudin. *Real and complex analysis*. McGraw-Hill, New York, 2nd edition, 1974.
- [54] I. Schur. Über ein Satz von C. Carathéodory. *S.-B. Preuss. Akad. Wiss. (Berlin)*, pages 4–15, 1912.
- [55] I. Schur. Über Potenzreihen die im Innern des Einheitskreises Beschränkt sind I. *J. Reine Angew. Math.*, 147:205–232, 1917. See also [30, p.31-59].
- [56] I. Schur. Über Potenzreihen die im Innern des Einheitskreises Beschränkt sind II. *J. Reine Angew. Math.*, 148:122–145, 1918. See also [30, p.36-88].
- [57] G. Szegő. *Orthogonal polynomials*, volume 33 of *AMS Colloquium publications*. American Mathematical Society, Providence, Rhode Island, 3rd edition, 1967. First edition 1939.
- [58] J. L. Walsh. *Interpolation and approximation*, volume 20 of *AMS Colloquium publications*. American Mathematical Society, Providence, Rhode Island, 3rd edition, 1960. First edition 1935.
- [59] N. Wiener. *Extrapolation, interpolation and smoothing of time series*. Wiley, New York, 1949.
- [60] D.C. Youla and M. Saito. Interpolation with positive-real functions. *J. Franklin Inst.*, 284(2):77–108, 1967.