# A Szegő theory for rational functions

*A. Bultheel*
*P. González-Vera*
*E. Hendriksen*
*O. Njåstad*

*Report TW 131, May 1990, Revised November 1991*

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

# A Szegő theory for rational functions

*A. Bultheel*
*P. González-Vera*
*E. Hendriksen*
*O. Njåstad*

*Report TW 131, May 1990, Revised November 1991*

Department of Computer Science, K.U.Leuven

**Abstract**

We shall consider nested spaces $\mathcal{L}_n$, $n = 0, 1, 2, \ldots$ 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 real 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. We give generalizations of most of the polynomial results : recurrence relations, continued fractions, reproducing kernels and Christoffel-Darboux relations, approximating properties and relation with the algorithm of Pick-Nevanlinna, quadrature, etc.

**Keywords :** Rational interpolation, Orthogonal functions, Szegő theory, Pick-Nevanlinna interpolation, reproducing kernels.
**MSC :** Primary : 42C05, Secondary : 30DF50, 41A20, 41A55.

# A Szegő theory for rational functions

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

May 1990, Revised November 1991

**Abstract**

We shall consider nested spaces $\mathcal{L}_n$, $n = 0, 1, 2, \ldots$ 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 real 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. We give generalizations of most of the polynomial results : recurrence relations, continued fractions, reproducing kernels and Christoffel-Darboux relations, approximating properties and relation with the algorithm of Pick-Nevanlinna, quadrature, etc.

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

# Contents

# 1 Introduction

The particularly rich and fascinating theory of polynomials orthogonal on the unit circle needs no advertising. These polynomials are named after Szegő since his pioneering work on them. His book on orthogonal polynomials [78] 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. The Szegő polynomials were studied by several authors. For example they play an important role in books by Geronimus [39], Freud [36], Grenander and Szegő [42] and many more modern books on orthogonal polynomials.

Szegő's interest in these polynomials was inspired by the investigation of the eigenvalue distribution of Toeplitz forms, an even older subject which was related to coefficient problems as initiated by Carathéodory [19, 20] and Carathéodory and Fejér [21] and further discussed by F. Riesz [70, 71], Gronwall [43], Schur [75, 76, 77], Hamel [44] and many others. The papers by Schur contained a continued fraction like algorithm to actually check if the given coefficients (moments) correspond to a bounded analytic function. The algorithm produces some coefficients (Schur coefficients) that turned out later to be 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 [65, 66, 67]. Nevanlinna was not aware of Pick's work when he developed the same theory in a long memoir in 1919 [59]. See also his later work [60, 61, 62]. Nevanlinna also gave an algorithm which directly generalized the algorithm given by Schur.

Since then, these problems and a myriad of generalizations played an important role in several books, like in Akhiezer [6], Kreĭn and Nudel'man [52], Walsh [79] and more recently in Donoghue [29], Garnett [37], Rosenblum and Rovnyak [72] etc.

Some of the more recent interest in this subject was stimulated by the the work of Adamyan, Arov and Kreĭn [2, 3] and most of all by their fundamental papers [4, 5]. We should also mention Sarason's paper [74] which had great influence on some developments made in later publications. These results relate the theory to operator theoretic methods for Hankel and Toeplitz operators.

Besides this, there is also a long history where the same theory is approached from several application fields. Grenander and Szegő themselves discussed the application in the theory of probability and statistics [42]. But you find also the applications in the prediction theory of stationary stochastic processes in work by Kolmogorov [51] and Wiener [80]. Some benchmark papers on this topic are collected in [50]. The book by Wiener contained a reprint from Levinson's celebrated paper [54], which is in fact a reformulation of the Szegő recursions. Other engineering applications are network theory (see e.g. Belevitch [10] and Youla and Saito [81]), spectral estimation (see Papoulis [64] for an excellent survey), maximum entropy analysis as formulated by Burg [18] (see the survey paper [53]), transmission lines and scattering theory as studied by Arov, Redheffer [69] and Dewilde and Dym [24, 26], digital filtering (see the survey of Kailath [49]), speech processing (see [56] or the tutorial paper by Makhoul [55]), etc.

It is from these engineering applications that emerged also methods for inverting and factorizing Toeplitz or related matrices (see [45]) and people are now even using these ideas for designing systolic arrays for the solution of a number of linear algebra problems [11]. The linear algebra literature in this connection has a complete history of its own, which we shall not mention here. Most of it was devoted to Toeplitz and Hankel matrices or related matrices which appear in relation with a theory of Schur-Szegő. It is however only recently that people are starting to look at matrices that are related to interpolation problems.

We could go on like this and probably be never complete in summing up all the application fields and this is without ever touching all the related generalizations of this theory that were obtained recently or the analog theory that has been developed for the complex half-plane instead of the unit circle or the continuous analog of Wiener-Hopf factorization. We just stop here by referring to a survey paper on the applications of Pick-Nevanlinna theory by Delsarte, Genin and Kamp [22].

In all this theory and application papers, the approach of the Nevanlinna-Pick theory from the point of view of the orthogonal functions has not been fully put forward. We try to give in this paper an approach to the theory which is an immediate generalization of the theory of Szegő 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.

Note that at any point in the discussion we can replace all the interpolation points $\alpha_k$ 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. Similarly, section 3 contains some properties of reproducing kernels. These reproducing kernels, or evaluating kernels, played an important role in the Szegő theory of orthogonal polynomials, they will be even more important in our development. Section 4 gives some generalities on positive real functions, i.e., analytic functions in the unit disk with positive real part, also known as Carathéodory functions since they appeared in the Carathéodory-Fejér problem. We give also the relation with the bounded analytic functions or Schur functions because it was Schur who used these functions in his algorithm to solve the Carathéodory-Fejér problem. A last important tool in our analysis are the J-contractive matrices studied by Potapov. They are introduced in section 5. It is then time to be more specific and we then introduce the fundamental spaces of rational functions which will generalize the spaces of polynomials of finite degree as they feature in the Szegő theory. They are defined in section 6 and some useful calculation techniques in these spaces are discussed in section 7. The latter computational properties will be used on almost every page to follow. Rather than starting with the orthogonal functions themselves, it will turn out to be easier to start with the reproducing kernels. These kernels feature in the solutions of some extremal problems in the rational function spaces (section 8). So their appearance is somehow natural when discussing the solution of the Szegő extremal problem (see section 22). In connection with this problem and other convergence results, it is important to know whether these rational functions are a complete system in $L_p$ or $H_p$. Some of these completeness problems are discussed in section 9. The relations between the kernels and the orthogonal functions are the Christoffel-Darboux relations. These are derived in section 10. We then obtain a recurrence relation for the reproducing kernels (section 11) and we also give a normalized version (section 12). The latter has the advantage that the recursion can be described with a J-unitary matrix. It will become clear in section 13, when we give the recurrence for the orthogonal functions, that they are somewhat less simple to handle, since they can not be simply described in terms of a J-unitary recursion. It is however possible to get some recurrence that generalizes the Szegő relations and one can also define functions of the second kind (section 14). Like in Szegő's theory, they appear as another independent solution of the recurrence for the orthogonal functions, exactly like the Szegő polynomials of the second kind. These second kind functions appear in the expression for the positive real interpolant of the positive real function that can be related to the measure defining the orthogonality. Before we discuss these interpolating properties, we give first some generalities on the relation that exist between the equality of the inner product in the rational function spaces for different measures and corresponding interpolation properties for the associated positive real functions (section 15) and the relation between orthogonal functions and quadrature (section 16). We then turn to the interpolation properties of the kernels (section 17) and the orthogonal functions (section 19) and the relation with the algorithm of Pick-Nevanlinna (section 18). In section 20 we give some continued fractions that can be obtained from the previously given recurrences and also the three-term recurrence that emerges from a contraction of these continued fractions. Next, we prove some Favard type theorems which state that if there is a recurrence relation then there is a measure for which you have orthogonality for the rational functions, or for which the kernels are reproducing. These are proved in section 21. Finally, section 22 gives the generalization of the Szegő problem, as it can be solved as a limiting approximation process in our rational function spaces. We could formulate it as finding the projection of $z^{-1}$ onto the space $H_2(\mu)$, that is the space of polynomials closed in the $L_2(\mu)$ metric. A crucial fact will then be to find out when the space $H_2(\mu)$ is not only spanned by the polynomials, which is the original Szegő approach, but it will also be spanned by the rational functions under certain conditions. All this is related to convergence results of the previous theory. Some further convergence results are discussed in section 23.

## 2 Setting up the scenery

We shall be concerned with complex function theory on the unit circle. The complex number field is denoted by $\mathbf{C}$. We use the following notations for the unit circle, the open unit disc and its complement :

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

The upper bar denotes complex conjugation when appropriate or closures when it concerns sets, e.g., $\overline{\mathbf{D}} = \mathbf{D} \cup \mathbf{T}$ is the closed unit disc and $\overline{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ is the Riemann sphere. The real axis is denoted as $\mathbf{R}$. Real and imaginary parts of a complex number $z$ are indicated by $\Re z$ and $\Im z$ respectively: $z = \Re z + \Im z$, $\mathbf{i}$ is reserved for the unit on the imaginary axis and the open right half plane is denoted as

$$\mathbf{H} = \{z : \Re z > 0\}.$$

By $\Pi_n$ we mean the set of polynomials of degree at most $n$. The set of complex functions holomorphic on $X$ are denoted by $H(X)$.

Let $\mu$ be a positive measure on $\mathbf{T}$, whose support in an infinite set. It is characterized by a distribution function $\int d\mu$ which has an infinite number of points of increase. If $u = e^{\mathbf{i}\alpha} \in \mathbf{T}$ is a point of discontinuity of the distribution function, then $\mu(\{u\})$ is the *concentrated mass* at $u$. The metric spaces $L_p(\mu)$, $0 < p \leq \infty$ 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

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

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

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

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 $N$ contains all spaces $H_p$ for $0 < p \leq \infty$. It can be characterized by

$$f \in N \Longleftrightarrow f = g/h; \qquad g, h \in H_\infty$$

which is a theorem by F. and R. Nevanlinna [32, p.16]. It is known that each function $f \in N$ has a nontangential limit to $\mathbf{T}$ a.e. and $\log|f| \in L_1$, unless $f \equiv 0$ [32, 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

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

Note that on $\mathbf{T}$, $f_*(z)$ is just $\overline{f(z)}$.

The Hardy and Nevanlinna classes of analytic functions in $\mathbf{E}$ are indicated by a prime, e.g.,

$$H'_p = \{f : f_* \in H_p\} \text{ and } N' = \{f : f_* \in N\}$$

Let the Lebesgue decomposition of $\mu$ be $\mu = \mu_a + \mu_s$ with $\mu_a$ satisfying

$$d\mu_a = \omega d\lambda \tag{2.1}$$

the *absolutely continuous part* : $\mu_a \ll \lambda$. The function $\mu' = d\mu_a/d\lambda = \omega \in L_1$ is a weight function. The remaining $\mu_s$ is the *singular part* w.r.t. $\lambda$ : $\mu_s \perp \lambda$.

Define the *moments* as the Fourier coefficients

$$c_k = \int e^{-\mathbf{i}k\theta} d\mu(\theta) \qquad , \qquad k \in \mathbf{Z}. \tag{2.2}$$

Clearly $c_{-k} = \bar{c}_k$ and $|c_k| \leq c_0$ for a real measure $\mu$. Computations will simplify substantially if we suppose the measure to be normalized. This means that we divide out $c_0$ which is always possible since it is not zero and we shall thus set $c_0 = 1$ from now on, which is no restriction of the generality. In other words, we work with the normalized measure which satisfies $\int d\mu = 1$.

With the positive measure $\mu$, we can associate a *positive real function* (that is a function analytic in $\mathbf{D}$ whose real part is positive there) :

$$\Omega_\mu(z) = \mathbf{i}c + \int \frac{e^{\mathbf{i}\theta} + z}{e^{\mathbf{i}\theta} - z} d\mu(\theta) \quad , \quad c \in \mathbf{R}, \quad z \in \mathbf{D}. \tag{2.3}$$

This function is analytic in $\mathbf{D}$ and belongs to $H_p$ for all $p < 1$ ([32, p.34]) and hence it has a nontangential limit a.e. The constant $\mathbf{i}c$ is the imaginary part of $\Omega_\mu(0)$. The integral representation is called the *Riesz-Herglotz* representation. The relation between $\mu$ and $\Omega_\mu$ is one-to-one except for the constant $c$. Since $\mu$ is uniquely defined by $\Omega_\mu$, we shall refer to it as the Riesz-Herglotz measure for $\Omega_\mu$.

The kernel in (2.3) shall be denoted as $D(t, z)$.

$$D(t, z) = \frac{t + z}{t - z}, \quad z \in \mathbf{D}, t \in \mathbf{T}$$

The *Poisson kernel* is defined as

$$P(t, z) = \Re\, D(t, z) = \frac{1 - |z|^2}{|t - z|^2}, \quad t \in \mathbf{T}.$$

This relation could be generalized for $t$ off $\mathbf{T}$ as

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

with the substar conjugate for $t$. The latter reduces to the previous definition for $t \in \mathbf{T}$. This kernel appears in the *Poisson-Stieltjes integral*

$$\Re\, \Omega_\mu(z) = \int \Re\, \frac{t + z}{t - z} d\mu(t) = \int \frac{1 - |z|^2}{|t - z|^2} d\mu(t) \quad , \quad z \in \mathbf{D} \tag{2.4}$$

which also has a radial limit given by

$$\lim_{r \to 1^-} \Re\, \Omega_\mu(re^{\mathbf{i}\theta}) = \mu'(e^{\mathbf{i}\theta}) \quad \text{a.e.} \tag{2.5}$$

where $\mu'$ is the symmetric derivative of $\mu$, i.e.,

$$\mu'(e^{\mathbf{i}\theta}) = \lim_{h \to 0} \frac{\mu((\theta - h, \theta + h))}{2h}.$$

See [32, p.4]. The relation (2.4) shows that indeed $\Re\, \Omega_\mu(z) > 0$ in $\mathbf{D}$, since the integrand at the right hand side is positive on $\mathbf{T}$. This explains why $\Omega_\mu$ is called a positive real function. Note that while $\Omega_\mu(z)$ depends on the measure $\mu$, including its singular part, the radial limit function of $\Re\, \Omega_\mu$ depends on its absolutely continuous part $\mu'$ only.

If $\Omega_\mu \in H_1$, the analysis simplifies considerably, since then $\mu$ should be absolutely continuous, the Fourier coefficients of $\mu$ are equal to the Taylor coefficients of $\Omega_\mu$, since indeed writing

$$D(t, z) = \frac{t + z}{t - z} = 1 + 2 \sum_{k=1}^{\infty} z^k t^{-k}, \quad z \in \mathbf{D}$$

gives

$$\Omega_\mu(z) = \mathbf{i}c + c_0 + 2\sum_{k=1}^{\infty} c_k z^k \tag{2.6}$$

which converges uniformely in $\mathbf{D}$. Any positive real function $\Omega$ of $H_1$ with $\Omega(0) > 0$ can be characterized by

$$\Omega(z) = \int D(t,z)\Re\,\Omega(t)d\lambda(t). \tag{2.7}$$

Note that the converse is not true : the measure $\mu$ can be absolutely continuous without $\Omega_\mu$ being an $H_1$ function.

The class of positive real functions is denoted by $\mathcal{P}$.

$$\mathcal{P} = \{f \in H(\mathbf{D}) : \Re\, f(z) > 0, \quad z \in \mathbf{D}\}.$$

More on class $\mathcal{P}$ functions and their relation to class $\mathcal{B}$ functions will be given in a later section.

Another useful observation we can make here is that

$$g(z) = \int D(t,z)f(t)d\mu(t), \qquad f \in L_1(\mu)$$

represents an analytic function in $\mathbf{D}$, while, by a symmetry argument, $h$ defined by

$$\overline{h(z)} = -\int \frac{1+t\overline{z}}{1-t\overline{z}} f(t)d\mu(t), \qquad f \in L_1(\mu)$$

represents an analytic function in $\mathbf{E}$.

If $\log\omega \in L_1$, with $\omega = \mu'$, then we can define

$$\sigma(z) = c\ \exp\{\frac{1}{2}\int D(t,z)\log\omega(t)d\lambda(t)\} \quad , z \in \mathbf{D}, \quad |c| = 1. \tag{2.8}$$

which we shall call the *spectral factor* of $\omega$. It is an *outer* function in $H_2$. Outer implies that $\sigma$ as well as $1/\sigma$ are both in $H_2$. See e.g. [73]. Since it is in $H_2$, it has a radial limit which satisfies

$$\omega(e^{\mathbf{i}\theta}) = |\sigma(e^{\mathbf{i}\theta})|^2 \quad \text{a.e.} \tag{2.9}$$

Note also that we have

$$|\sigma(z)|^2 = \exp\{\int P(t,z)\log\omega(t)d\lambda(t)\} \quad , z \in \mathbf{D}.$$

The inequality

$$(1-|z|^2)|\sigma(z)|^2 \le \int \mu' d\lambda$$

holds. (See [42, p.25].)

As you can see from its definition, the spectral factor $\sigma$ does not depend on the singular part of the measure, but is completely defined in terms of the absolutely continuous part. Recall that $d\mu_s = d\mu - \mu' d\lambda = d\mu - d\mu_a$. From the Szegő theory of orthogonal polynomials, we know that $1/\sigma$ vanishes $d\mu_s$ a.e. if $\log\mu' \in L_1$ as was shown in Freud's book [36, 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*. Szegő's theory has been extended beyond this condition if $\mu' > 0$ a.e. on $\mathbf{T}$ [58].

An *inner* function $U$ in $H_p$ is a function with

$$|U(e^{\mathbf{i}\theta})| = 1 \quad \text{a.e.}$$

Such an inner function has the general form

$$U(z) = cB(z)S(z)$$

with $|c| = 1$,

$$B(z) = z^n \prod_{j\geq 1} \frac{\overline{\alpha}_j}{|\alpha_j|} \frac{\alpha_j - z}{1 - \overline{\alpha}_j z}$$

$0 \neq \alpha_j \in \mathbf{D}$ for all $j$, $n \geq 0$ is a *Blaschke product* and $S(z)$ is a singular factor which has the form

$$S(z) = \exp\{-\int D(t,z) d\nu(\theta)\}$$

where $\nu$ is a finite positive singular measure on $\mathbf{T}$.

Since on $\mathbf{T}$ we have for any inner function : $|U|^2 = 1$ or $UU_* = 1$, we can write $U = 1/U_*$ on $\mathbf{T}$. Because $U \in H_p$, we can extend it analytically to $\mathbf{D}$ and because $U_* \in H_p'$, we can extend this analytically to $\mathbf{E}$. In this way, $U$ has an analytic extension to the whole complex Riemann sphere, where we have to exclude the poles $1/\overline{\alpha}_j$, $j = 1, 2, \ldots$ of course as well as the points of $\mathbf{T}$ which are in the support of $\nu$. One says that inner functions allow a *pseudo-meromorphic extension* across the unit circle to the complete Riemann sphere [31]. The nontangential limits from outside or inside the unit circle coincide. See also [37, p.75 ff]. Douglas, Shapiro and Shields [31] showed that a general function $f \in H_2$ has a pseudo-meromorphic extension across $\mathbf{T}$ if there exists an inner function $U \in H_2$ such that on $\mathbf{T}$ we have $\overline{U}f \in H_2'$ or, equivalently, 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 $\overline{\mathbf{C}}$.

Suppose that the spectral factor $\sigma$ has such a pseudo-meromorphic extension, then the relations

$$\omega = \sigma\sigma_* = \frac{1}{2}[\Omega + \Omega_*] \tag{2.10}$$

are valid on $\mathbf{T}$ but these relations can now be extended to $\overline{\mathbf{C}}$.

We now conclude this section by observing that

$$\langle f, g\rangle_\mu = \int f g_* d\mu = \langle g_*, f_*\rangle_\mu = \langle f g_*, 1\rangle_\mu = \langle Uf, Ug\rangle_\mu$$

if $U$ is an inner function.

# 3 Reproducing kernel spaces

In this section we recall some definitions and properties of reproducing kernel spaces.

**Definition 3.1 (Reproducing kernel)** *Let $H$ be a Hilbert space of functions defined on $X$ with inner product $\langle \cdot, \cdot \rangle$. Then we call $k_w(z) = k(z, w)$ a* reproducing kernel *if*

1. *$k_w(z) \in H$ for all $w \in X$*
2. *$\langle f, k_w \rangle = f(w)$ for all $w \in X$ and $f \in H$.*

It is a well known property [57] that if the Hilbert space is separable and $\{\phi_k\}_{k \in \Gamma}$ is an orthonormal basis, then the unique reproducing kernel is given by

$$k(z, w) = \sum_{k \in \Gamma} \phi_k(z) \overline{\phi_k(w)}$$

These reproducing kernels can also be used to find best approximants in subspaces as the following property shows.

**Theorem 3.1** *Let $H$ be a separable Hilbert space and $K$ a closed subspace with reproducing kernel $k_w(z) = k(z, w)$. Then the best approximant (w.r.t. the norm $\|\cdot\| = \langle \cdot, \cdot \rangle^{1/2}$) of $f \in H$ from $K$ is given by*

$$h(w) = \langle f, k_w \rangle.$$

*This $h$ is the orthogonal projection of $f$ onto $K$.*

**Proof.** Suppose $\{\phi_k : k \in \Gamma'\}$ is an orthonormal basis for $K$. Extend this with $\{\phi_k : k \in \Gamma''\}$ such that $\{\phi_k : k \in \Gamma = \Gamma' \cup \Gamma''\}$ is an orthonormal basis for $H$. Then the kernel of $K$ is given by $k_w = \sum_{k \in \Gamma'} \phi_k \overline{\phi_k(w)}$. Any element $f \in H$ can be expanded as

$$f = \sum_{k \in \Gamma} a_k \phi_k \qquad \text{with} \qquad a_k = \langle f, \phi_k \rangle.$$

The best approximant from $K$ is given by

$$h = \sum_{k \in \Gamma'} a_k \phi_k$$

while

$$\langle f, k_w \rangle = \sum_{k \in \Gamma'} \langle f, \phi_k \rangle \phi_k(w) = \sum_{k \in \Gamma'} a_k \phi_k(w) = h(w).$$

This proves the theorem. □

With these kernels, it is also possible to solve a number of classical extremal problems in Hilbert spaces. We find in [57, p.44] the following theorem.

**Theorem 3.2** *Let $H$ be a Hilbert space with reproducing kernel $k(z, w)$. Then all the solutions of the following problem*

$$P^1(a, w): \qquad \sup \{|f(w)|^2 : \|f\| = a, \quad w \in X\}$$

*are given by*

$$f = \eta\, a\, k(z, w)[k(w, w)]^{-1/2} \quad , \quad |\eta| = 1$$

*and these are all the solutions. The supremum is*

$$|a|^2 k(w, w).$$

*The problem*

$$P^2(a,w): \qquad \inf\,\{\|f\|^2 : f(w) = a, \quad w \in X\}$$

*reaches an infimum for*

$$f = a\,k(z,w)[k(w,w)]^{-1}$$

*and this solution is unique. The minimum reached is*

$$|a|^2[k(w,w)]^{-1}.$$

**Proof.** This theorem was given in [57] for $a = 1$, but the introduction of $a$ is trivial. □

The problems $P^1(a,w)$ and $P^2(a,w)$ are related to dual extremal problems as can be found in [30, p.133] in a much more general context of Banach spaces.

Problem $P^2(a,w)$ can be read as the problem of finding the orthogonal projection in $H$ of 0 onto the space $V = \{f \in H : f(w) = a\}$.

## 4 The classes $\mathcal{P}$ and $\mathcal{B}$

The class of *positive real* functions was already partially discussed in section 2. It is also known as the class of *Carathéodory* functions. We shall denote this class as (recall that $H(\mathbf{D})$ denotes analytic functions in $\mathbf{D}$ and $\mathbf{H}$ is the (open) right half plane)

$$\mathcal{P} = \{f \in H(\mathbf{D}) : f(\mathbf{D}) \subset \mathbf{H}\}. \tag{4.1}$$

The class of *bounded analytic* functions or *Schur* functions is defined as

$$\mathcal{B} = \{f \in H(\mathbf{D}) : f(\mathbf{D}) \subset \mathbf{D}\}. \tag{4.2}$$

Since they can be regarded as the unit ball in $H(\mathbf{D})$, we have chosen the notation $\mathcal{B}$ for it.

We shall also use the following notations for slightly larger classes :

$$\overline{\mathcal{P}} = \{f \in H(\mathbf{D}) : f(\mathbf{D}) \subset \overline{\mathbf{H}} = \mathbf{H} \cup \mathbf{iR}\} \tag{4.3}$$

$$\overline{\mathcal{B}} = \{f \in H(\mathbf{D}) : f(\mathbf{D}) \subset \overline{\mathbf{D}} = \mathbf{D} \cup \mathbf{T}\}. \tag{4.4}$$

The classical *Schwarz'* lemma for functions in $\mathcal{B}$ reads

**Lemma 4.1 (Schwarz' lemma)** *Suppose $f \in \mathcal{B}$ and $f(0) = 0$. Then*

$$|f'(0)| \leq 1 \quad \text{and} \quad |f(z)| \leq |z|, \quad z \in \mathbf{D}. \tag{4.5}$$

*Equality holds if and only if $f(z) = c\,z$ with $|c| = 1$.*

**Proof.** This is a classical theorem and we are not going to prove it here. See e.g., [17, p.191] or [37, p.1]. □

A *Möbius transform* is a conformal map of the unit circle/disc onto itself. It has the general form

$$\zeta : z \mapsto \frac{az + b}{\overline{b}z + \overline{a}}, \quad |b| < |a| \tag{4.6}$$

or equivalently

$$\zeta_\alpha : z \mapsto \eta \frac{z - \alpha}{1 - \overline{\alpha}z}, \quad |\alpha| < 1, |\eta| = 1. \tag{4.7}$$

Note that $\zeta_\alpha$ is the most general conformal map of this type which transforms $\alpha$ into the origin. The unit circle $\mathbf{T}$ is transformed into itself. The inverse transformation is given by

$$\zeta_\alpha^{-1} : w \mapsto \frac{w/\eta + \alpha}{1 + \overline{\alpha}w/\eta}. \tag{4.8}$$

Clearly $\zeta_\alpha$ is a function from the class $\mathcal{B}$.

The expression

$$\rho(z, w) = |\zeta_w(z)| = \left|\frac{z - w}{1 - \overline{w}z}\right|; \quad z, w \in \mathbf{D}$$

is invariant under Möbius transformations. It is called the *pseudohyperbolic distance* and it forms a metric in $\mathbf{D}$.

A form of the Schwarz lemma which is invariant w.r.t. this distance can now be formulated :

**Theorem 4.2** *Let $f \in \mathcal{B}$ and $z$ and $w \in \mathbf{D}$. Then*

$$\rho(f(z), f(w)) = \left|\frac{f(z) - f(w)}{1 - \overline{f(w)}f(z)}\right| \leq \left|\frac{z - w}{1 - \overline{w}z}\right| = \rho(z, w), \quad z \neq w \tag{4.9}$$

*and*

$$\frac{|f'(z)|}{1 - |f(x)|^2} \leq \frac{1}{1 - |z|^2}, \quad z \in \mathbf{D}. \tag{4.10}$$

*Equality holds if and only if $f$ is a Möbius transformation.*

**Proof.** Also this result is classical. See e.g., [17, p.192] or [37, p.2]. □

**Notes:**

1. Inequality (4.9) says that $f \in \mathcal{B}$ is Lipschitz continuous w.r.t. the pseudohyperbolic distance.
2. The second form (4.10) is the limiting case of the first one (4.9) for $z \to w$.

The following property forms the basis of the Pick-Nevanlinna algorithm.

**Theorem 4.3** *Let $\zeta_\alpha$ be a Möbius transform as defined in (4.7).*

1. *Let $f \in \mathcal{B}$ and $\alpha \in \mathbf{D}$. Then $\zeta_\alpha(f) \in \mathcal{B}$. More precisely:*

$$\zeta_\alpha(\mathcal{B}) = \mathcal{B}. \tag{4.11}$$

2. *If $f \in \mathcal{B}$ and $\alpha \in \mathbf{D}$, then*

$$\zeta_{f(\alpha)}(f)/\zeta_\alpha \in \mathcal{B}. \tag{4.12}$$

3. *If $f \in \mathcal{B}$ and $f(\alpha) = 0$ for some $\alpha \in \mathbf{D}$, then $f/\zeta_\alpha \in \mathcal{B}$.*

**Proof.**

1. Since $\zeta_\alpha \in \mathcal{B}$ and the composition of functions in $\mathcal{B}$ is also in $\mathcal{B}$, we find that $\zeta_\alpha(\mathcal{B}) \subset \mathcal{B}$. Hence $\mathcal{B} \subset \zeta_\alpha^{-1}(\mathcal{B})$. But since $\zeta_\alpha^{-1} = \zeta_{-\alpha}$ (take $\eta = 1$ in (4.7), without loss of generality) we also have $\zeta_{-\alpha}(\mathcal{B}) \subset \mathcal{B}$. Thus

$$\mathcal{B} \subset \zeta_\alpha^{-1}(\mathcal{B}) = \zeta_{-\alpha}(\mathcal{B}) \subset \mathcal{B},$$

so that equality holds.

2. This is a rewriting of the invariant form of Schwarz' lemma.
3. This is a special case of 2 because $f(\alpha) = 0$.

□

The link with class $\mathcal{P}$ functions can be made as follows. The *Cayley transform*

$$c : z \mapsto \frac{1-z}{1+z} \tag{4.13}$$

is a one-to-one map of $\mathbf{D}$ onto $\mathbf{H}$ and of $\mathbf{T}$ onto $\mathbf{i}\overline{\mathbf{R}}$ ($-1$ is mapped onto $\infty$). The following result is now simple to see.

**Theorem 4.4** *The following relations between class $\mathcal{P}$ and class $\mathcal{B}$ exist.*

1. *The Cayley transform c is a one-to-one map of $\mathcal{P}$ onto $\mathcal{B}$. I.e.,*

$$c(\mathcal{B}) = \mathcal{P} \text{ and } c(\mathcal{P}) = \mathcal{B} \tag{4.14}$$

2. *For the extended classes, define $\overline{\mathcal{B}}' = \overline{\mathcal{B}} \backslash \{-1\}$, i.e., we exclude the constant function $f \equiv -1$. Then*

$$c(\overline{\mathcal{B}}') = \overline{\mathcal{P}} \text{ and } c(\overline{\mathcal{P}}) = \overline{\mathcal{B}}'. \tag{4.15}$$

**Proof.**

1. If $f \in \mathcal{B}$, then $|f| < 1$ in $\mathbf{D}$ so that $|1+f| > 0$ in $\mathbf{D}$. Hence $c(f) \in H(\mathbf{D})$ and conversely, if $f \in \mathcal{P}$, then $1+f$ has strictly positive real part in $\mathbf{D}$ and therefore $1+f$ does not vanish. Thus again, $f \in H(\mathbf{D})$. The rest follows from the one-to-one map given by the Cayley transform.
2. Here we have to exclude $f(z) \equiv -1$ because then the transform would fail.

This concludes our proof. □

We conclude this section with a property of inner functions from $\mathcal{B}$.

**Theorem 4.5** *Let $f \in \mathcal{B}$ be an inner function. Then a.e. the following relations hold*

$$|f(z)| < 1 \text{ in } \mathbf{D}, \quad |f(z)| = 1 \text{ on } \mathbf{T}, \quad |f(z)| > 1 \text{ in } \mathbf{E}. \tag{4.16}$$

**Proof.** This is a simple consequence of the definitions. (2) is from the definition of an inner function, (1) follows from the maximum modulus principle and because of the technique of pseudomeromophic extension, it easily follows that (3) holds. □

## 5 J-unitary and J-contractive matrix functions

We shall consider $2 \times 2$ matrices $\theta$ 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

$$\theta^H \, J \, \theta = J \qquad \text{on } \mathbf{T} \tag{5.1}$$

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 (5.1) as

$$\theta_* \, J \, \theta = J \text{ on } \mathbf{T}. \tag{5.2}$$

As we did for inner functions in $\mathcal{B}$, we can define a pseudomeromorphic extension for such a $\theta$-matrix. Indeed, it follows from (5.2) that $|\det \, \theta| = 1$ a.e. on $\mathbf{T}$. Hence, $\theta$ is invertible on $\mathbf{T}$ a.e. and therefore also in $\mathbf{D}$ a.e. From the relation

$$\theta_* = J \, \theta^{-1} \, J$$

which holds a.e. on the unit circle $\mathbf{T}$, we can extend the right hand side to $\mathbf{D}$ and hence we define also $\theta_*(z) = [\theta(1/\overline{z})]^H$ for $z \in \mathbf{D}$, which is equivalent with defining $\theta(y) = [\theta_*(1/\overline{y})]^H$ for $y \in \mathbf{E}$. Thus $\theta$ is defined on the sphere $\overline{\mathbf{C}}$.

We shall call the matrix functions satisfying

$$\theta_* \, J \, \theta = J \text{ a.e. in } \overline{\mathbf{C}}$$

*J-unitary matrices* and denote the set of these matrices as

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

We have the following properties for J-unitary matrices:

**Theorem 5.1** *For elements of $\mathbf{T}_J$ the following relations hold :*

1. $\theta_1, \theta_2 \in \mathbf{T}_J \Rightarrow \theta_1\theta_2 \in \mathbf{T}_J$
2. $\theta \in \mathbf{T}_J \Rightarrow |\det \, \theta| = 1$
3. $\theta \in \mathbf{T}_J \Rightarrow \theta^{-1} = J\theta_* J$
4. $\theta \in \mathbf{T}_J \Rightarrow \theta \, J\theta_* = J$
5. *If* $\theta = [\theta_{ij}] \in \mathbf{T}_J$ *then*
   (a) $\theta_{11*}\theta_{11} - \theta_{21*}\theta_{21} = \theta_{22*}\theta_{22} - \theta_{12*}\theta_{12} = 1$
   (b) $\theta_{11*}\theta_{12} - \theta_{21*}\theta_{22} = \theta_{11*}\theta_{21} - \theta_{12*}\theta_{22} = 0$
   (c) $\theta_{12*}\theta_{12} - \theta_{21*}\theta_{21} = \theta_{11*}\theta_{11} - \theta_{22*}\theta_{22} = 0$
   (d) $(\theta_{11} + \theta_{12})_*^{-1}(\theta_{11} - \theta_{12})_* = (\theta_{22} + \theta_{21})^{-1}(\theta_{22} - \theta_{21})$

6. *Let* $\theta = [\theta_{ij}] \in \mathbf{T}_J$ *and set* $a = \theta_{11} - \theta_{12}$*;* $b = \theta_{11} + \theta_{12}$*;* $c = \theta_{22} - \theta_{21}$*; and* $d = \theta_{22} + \theta_{21}$*. Then*

$$\frac{1}{2}\left[\frac{a}{b} + \frac{a_*}{b_*}\right] = \frac{1}{bb_*} = \frac{1}{2}\left[\frac{c}{d} + \frac{c_*}{d_*}\right] = \frac{1}{dd_*}$$

$$ab_* + a_*b = cd_* + c_*d = 2.$$

**Proof.** Parts 1-3 are trivial to check. Part 4 follows from $\theta_* J\theta = J$ so that $J\theta_* = \theta^{-1}J$ and by multiplying with $\theta$, we get $\theta J\theta_* = J$. Part 5 is just an explicitation of $\theta_* J\theta = J = \theta J\theta_*$. The results of part 5 then give an easy proof for part 6. □

An important example of a constant J-unitary matrix is

$$U_\rho = (1 - |\rho|^2)^{-1/2} \begin{bmatrix} 1 & -\rho \\ -\overline{\rho} & 1 \end{bmatrix} \quad , \quad \rho \in \mathbf{D}. \tag{5.3}$$

In fact, this example turns out to be almost the most general constant J-unitary matrix.

**Theorem 5.2** *The most general constant* $\theta \in \mathbf{T}_J$ *is given by*

$$\begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix} U_\rho \tag{5.4}$$

*with* $|\eta_i| = 1, i = 1, 2$ *and* $U_\rho$*,* $\rho \in \mathbf{D}$ *as given in (5.3).*

**Proof.** This is a matter of simple algebra. You can make use of the properties given in Theorem 5.1. □

A simple nonconstant matrix from the class $\mathbf{T}_J$ is given by the *Blaschke-Potapov factor* with a zero in $\alpha \in \mathbf{D}$.

$$B_\alpha = \begin{bmatrix} \zeta_\alpha & 0 \\ 0 & 1 \end{bmatrix} = \zeta_\alpha \oplus 1 \quad ; \quad \zeta_\alpha(z) = \frac{z - \alpha}{1 - \overline{\alpha} z}, \quad \alpha \in \mathbf{D}. \tag{5.5}$$

The J-unitary matrices that are also *J-contractive* in $\mathbf{D}$ form an important class we shall often need in this paper. J-contractive in $\mathbf{D}$ means

$$\theta^H J\theta \leq J \text{ a.e. in } \mathbf{D}.$$

By the inequality, we mean that $J - \theta^H J\theta \geq 0$, i.e., this is positive semi definite. The class of J-unitary, strictly J-contractive matrices is denoted by

$$\mathbf{D}_J = \{\theta \in \mathbf{T}_J; \theta^H J\theta < J \text{ a.e. in } \mathbf{D}\} \tag{5.6}$$

and its closure

$$\overline{\mathbf{D}}_J = \{\theta \in \mathbf{T}_J; \theta^H J\theta \leq J \text{ a.e. in } \mathbf{D}\}. \tag{5.7}$$

For these matrices a number of additional properties can be proved. The following theorem is due to Dewilde and Dym [24, p.448].

**Theorem 5.3** *For* $\theta = [\theta_{ij}] \in \overline{\mathbf{D}}_J$ *the following holds.*

1. $\theta^H \in \overline{\mathbf{D}}_J$.
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.*

**Proof.** Part 1 was shown in Potapov [68, p.171].

Part 2 follows from part 1 and the definition of $\mathbf{D}_J$ which imply that for $z \in \mathbf{D}$ it holds that a.e. $\theta(z)^{-1}J\theta(z)^{-H} \geq J$. Now make use of $\theta(z)^{-1} = J\theta(1/\overline{z})^H J$ to get

$$\theta(w)^H J\theta(w) \geq J \text{ with } w = 1/\overline{z} \in \mathbf{E}. \tag{5.8}$$

Using part 1, we also have

$$\theta(1/\overline{z})J\theta(1/\overline{z})^H \geq J, \quad z \in \mathbf{D}.$$

Then it follows from writing out the (1,1) element that

$$|\theta_{11*}|^2 - |\theta_{12*}|^2 \geq 1 \text{ a.e. in } \mathbf{D}. \tag{5.9}$$

Therefore, $(\theta_{11*}+\theta_{12*})$ is not zero and its inverse is analytic in $\mathbf{D}$.

Computing the real part of the expression of part 4, we get, using (5.9)

$$\Re\left(\frac{\theta_{11*}-\theta_{12*}}{\theta_{11*}+\theta_{12*}}\right) = \frac{|\theta_{11*}|^2 - |\theta_{12*}|^2}{|\theta_{11*}+\theta_{12*}|^2} \geq \frac{1}{|\theta_{11*}+\theta_{12*}|^2} \geq 0.$$

From this part 4 follows. Since the left hand side in the previous expression is a harmonic majorant in $\mathbf{D}$ for the analytic function $|\theta_{11*}+\theta_{12*}|^{-2}$, it follows from [32, Theorem 2.12, p.28] that part 3 is true.

Part 5 and 6 follow from the (2,2) element in much the same way as 3 and 4 followed from the (1,1) element in (5.8).

To prove the last part, note that we had for $z \in \mathbf{D}$ that $\theta_* J\theta_*^H \geq J$. So we get in $\mathbf{D}$ :

$$[1 \;\; 1]\theta_* J\theta_*^H[1 \;\; 1]^T \geq [1 \;\; 1]J[1 \;\; 1]^T = 0$$

with equality on $\mathbf{T}$. Working out gives

$$1 - a\overline{a} \geq 0$$

where $a = (\theta_{11*}+\theta_{12*})^{-1}(\theta_{21*}+\theta_{22*})$ and with equality on $\mathbf{T}$. This identifies $a$ as an inner function. □

The following theorem describes a simple matrix from the class $\mathbf{D}_J$.

**Theorem 5.4** *The most general first degree matrix in* $\mathbf{D}_J$ *with a zero at* $z = \alpha \in \mathbf{D}$ *is given by*

$$\begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix} U_\rho \, B_\alpha \, U_\gamma$$

*with* $\eta_1, \eta_2 \in \mathbf{T}$*,* $U_\rho$ *and* $U_\gamma$ *constant J-unitary matrices as defined in (5.3) for* $\rho$ *and* $\gamma \in \mathbf{D}$ *and* $B_\alpha$ *is a Blaschke-Potapov factor as in (5.5).*

**Proof.** This is a classical result that can be found e.g. in Potapov [68, p.187-188]. □

## 6 The spaces $\mathcal{L}_n$

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

$$\zeta_i(z) = \frac{\overline{\alpha}_i}{|\alpha_i|}\frac{\alpha_i - z}{1-\overline{\alpha}_i z} \text{ if } \alpha_i \neq 0 \quad \text{or} \quad \zeta_i(z) = z \text{ if } \alpha_i = 0. \tag{6.1}$$

In what follows we shall use only the first notation and suppose by convention that $\overline{\alpha}_i/|\alpha_i|$ is equal to $-1$ for $\alpha_i = 0$.

Next we define finite Blaschke products recursively as

$$B_0 = 1 \text{ and } B_n = B_{n-1}\zeta_n \text{ for } n \geq 1. \tag{6.2}$$

We then consider the spaces

$$\mathcal{L}_n = \text{span}\{B_k : k = 0, 1, \ldots, n\}. \tag{6.3}$$

They will often be considered as subspaces of $L_2(\mu)$ but from time to time we shall also consider them as subspaces of $L_2(\lambda)$ or some other space. Note that if all $\alpha_i = 0$, then $\mathcal{L}_n$ is just the space of polynomials of degree at most $n$. Thus in that case $\mathcal{L}_n = \Pi_n$. There are of course many equivalent ways to describe the spaces $\mathcal{L}_n$. One of them is to say that $\mathcal{L}_n$ is a space of rational functions with prescribed poles $1/\overline{\alpha}_i$, $i = 1, \ldots, n$ which are all in $\mathbf{E}$.

$$\mathcal{L}_n = \{f = \frac{p(z)}{\prod_{i=1}^n (1-\overline{\alpha}_i z)}; p \in \Pi_n\}.$$

The spaces $\mathcal{L}_n$ depend upon the numbers in

$$A_n = \{\alpha_i : \alpha_i \in \mathbf{D}, i = 1, \ldots, n\}.$$

By $A_{n*}$ we shall denote the set

$$A_{n*} = \{1/\overline{\alpha}_i : \alpha_i \in A_n\}.$$

Some of the $\alpha_i$ can be repeated a number of times. So we could rearrange them and make the repetition explicit by setting

$$A_n = \{\underbrace{\beta_1, \ldots, \beta_1}_{\nu_1}, \underbrace{\beta_2, \ldots, \beta_2}_{\nu_2}, \underbrace{\beta_m, \ldots, \beta_m}_{\nu_m}\}. \tag{6.4}$$

We fix $\beta_1$ to be 0, but it may happen that it does not appear. Therefore we set $\nu_1 \geq 0$. All the other $\nu_i, i = 2, \ldots, m$ are positive integers and $\sum_1^m \nu_i = n$.

The basis $\{B_k : k = 0, \ldots, n\}$ is not the only possible choice to span $\mathcal{L}_n$ of course. With $A_n$ as described in (6.4), we can use as a possible basis

$$\begin{aligned}\{w_k : k = 0, \ldots, n\} \quad = \quad & \{1, z, \ldots, z^{\nu_1}, (1-\overline{\beta}_2 z)^{-1}, \ldots, (1-\overline{\beta}_2 z)^{-\nu_2}, \ldots, \\ & (1-\overline{\beta}_m z)^{-1}, \ldots, (1-\overline{\beta}_m z)^{-\nu_m}\}.\end{aligned} \tag{6.5}$$

The advantage of working with the basis $\{B_k : k = 0, \ldots, n\}$ is that repetition of points and distinction between $\alpha_i = 0$ or $\alpha_i \neq 0$ need no special notation as in some other choices like e.g., (6.5).

Here is yet another way to characterize the spaces $\mathcal{L}_n$. Define

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

with $B_n$ the finite Blaschke product associated with $\alpha_1, \ldots, \alpha_n$. Clearly $\mathcal{M}_n$ is a shift invariant subspace of $H_2$ since

$$f \in \mathcal{M}_n \Rightarrow Sf \in \mathcal{M}_n$$

where $Sf(z) = zf(z)$ represents the shift operator. As a matter of fact, the famous *Beurling theorem* says that any shift invariant subspace can be characterized as $UH_2$ where $U$ is an inner function.

The sequence $\{\mathcal{M}_n : n = 0, 1, \ldots\}$ contains shrinking subspaces, i.e., $\mathcal{M}_{n+1} \subset \mathcal{M}_n \subset \cdots \subset \mathcal{M}_0 = zH_2$. If we define

$$\mathcal{L}_n = H_2 \ominus \mathcal{M}_n = \mathcal{M}_n^\perp = \{f \in H_2, \langle f, g\rangle = 0 \text{ for all } g \in \mathcal{M}_n\}, \tag{6.6}$$

then the sequence $\{\mathcal{L}_n : n = 0, 1, \ldots\}$ is a growing sequence : $\mathcal{L}_{n+1} \supset \mathcal{L}_n \supset \cdots \supset \mathcal{L}_0 = \mathbf{C}$. The choice of the notation $\mathcal{L}_n$ in the previous definition may seem confusing at the moment, since we reserved this notation for the spaces defined in (6.3). Our next Theorem will show that the spaces of (6.3) and (6.6) are actually the same. We shall do this by proving that a basis for $\mathcal{L}_n$ is given by $\{B_k : k = 0, \ldots, n\}$.

**Theorem 6.1** *Define the spaces $\mathcal{M}_n = zB_n H_2$ and $\mathcal{L}_n = H_2 \ominus \mathcal{M}_n = \mathcal{M}_n^\perp$ where $B_n$ is a finite Blaschke product of degree $n$. Then*

$$\mathcal{L}_n = \operatorname{span}\{B_k : k = 0, \ldots, n\}$$

**Proof.** The result is obvious for $n = 0$. We shall then prove that $B_n \in \mathcal{L}_n - \mathcal{L}_{n-1}$ which implies that the Blaschke products form indeed a basis. First we show that $B_n \in \mathcal{L}_n$. Choose some $f = zB_n g \in zB_n H_2$ ($g \in H_2$). Then

$$\langle f, B_n\rangle = \int zB_n g B_{n*} d\lambda = \int zg d\lambda = 0$$

since $B_{n*} = 1/B_n$ and $g$ has vanishing negative Fourier coefficients. Hence $B_n \perp zB_nH_2$ and therefore $B_n \in \mathcal{L}_n$. On the other hand $B_n \notin \mathcal{L}_{n-1}$ since for $f \in \mathcal{M}_{n-1}$ :

$$\langle f, B_n\rangle = \int zB_{n-1} g\, B_{n*} d\lambda = \int zg/\zeta_n d\lambda$$

with $1/\zeta_n = \alpha_n/|\alpha_n| \cdot (1 - \overline{\alpha}_n z)/(\alpha_n - z)$, which gives by Cauchy's formula

$$\langle f, B_n\rangle = -g(\alpha_n)\frac{\alpha_n}{|\alpha_n|}\alpha_n(1 - |\alpha_n|^2)$$

which is not zero for all $g \in H_2$. Hence, $B_n$ is not orthogonal to $\mathcal{M}_{n-1}$, thus not in $\mathcal{L}_{n-1}$. $\square$

The previous theorem shows that, we can identify $\mathcal{L}_n$ as defined in (6.6) with the originally introduced space $\mathcal{L}_n$ of (6.3).

$$\mathcal{L}_n = H_2 \ominus zB_nH_2 = \mathcal{M}_n^\perp = \operatorname{span}\{B_k : k = 0, \ldots, n\}.$$

In the special case where all $\alpha_i = 0$, the spaces $\mathcal{L}_n$ reduce to the spaces $\Pi_n$ of polynomials. It is well known that in that case the Gram matrix of the basis $Z_n^T = [1\ z\ z^2\ \ldots\ z^n]$ in $L_2(\mu)$ is given by

$$G_n = \langle Z_n, Z_n^T\rangle_\mu = [\langle z^i, z^j\rangle_\mu] = [c_{j-i}]$$

which is a positive definite Toeplitz matrix containing the moments of $\mu$. If all the $\alpha_i$ are distinct, then the basis $w_k$ which we mentioned previously reduces to $W_n^T = [w_0\ w_1 \ldots w_n] = [1,\ (1 - \overline{\alpha}_1 z)^{-1} \ldots (1 - \overline{\alpha}_n z)^{-1}]$. Set by convention $\alpha_0 = 0$ and denote $\Omega_\mu$ (see (2.3)) as $\Omega$, which is supposed to be normalized such that $\Omega(0) = 1$. This gives the Gram matrix

$$G_n = \langle W_n, W_n^T\rangle_\mu = \frac{1}{2}\left[\frac{\Omega(\alpha_j) + \overline{\Omega(\alpha_i)}}{1 - \alpha_j\overline{\alpha}_i}\right].$$

This is a so called *Pick matrix* named after G. Pick who used the positive definiteness of this matrix as a criterion to characterize the solvability of the Pick-Nevanlinna interpolation problem. In the more general case where some of the $\alpha_k$ do coincide, the Gram matrix looks more complicated and involves derivatives of $\Omega$ evaluated at the $\alpha_i$. To see this, we give a technical lemma first.

**Lemma 6.2** *Let $D_w(z) = D(z,w) = (z+w)(z-w)^{-1}$ be the reproducing kernel for $H_2$, then*

$$\partial_w^k D_w(z) = 2(k!)z(z-w)^{-(k+1)}, \quad k \geq 1$$

*where $\partial_w^k$ denotes the $k$-th derivative with respect to $w$. We also have*

$$[\partial_w^k D_w(z)]_* = 2(k!)z^k(1-\overline{w}z)^{-(k+1)}, \quad k \geq 1,$$

*where the substar transform is with respect to $z$. Furthermore, if*

$$\Omega(w) = \Im\,\Omega(0) + \int D_w(z)d\mu(z),$$

*then*

$$\Omega^{(k)}(w) = \partial_w^k \Omega(w) = \int \partial_w^k D_w(z)d\mu(z) = 2(k!)\int \frac{z}{(z-w)^{k+1}}d\mu(z), \quad k \geq 1.$$

*Note also that*

$$\overline{\Omega^{(k)}(w)} = \int [\partial_w^k D_w(z)]_* d\mu(z).$$

**Proof.** This is a matter of simple algebra and we leave this to the reader. □

With this lemma, we can now prove the following theorem.

**Theorem 6.3** *If we choose the basis (6.5) for the space $\mathcal{L}_n$, then the Gram matrix*

$$G_n = [\langle w_i, w_j \rangle_\mu]$$

*will only depend upon*

$$\Omega^{(t)}(\beta_s) \quad t = 0,1,\ldots,\nu_s - 1,\ s = 1,2,\ldots,m$$

*The superscript $^{(t)}$ denotes the $t$-th derivative.*

**Proof.** One possible form of the elements in $G_n$ involves an integral like

$$\int \frac{1}{(1-\overline{\alpha}z)^k}\frac{z^l}{(z-\beta)^l}d\mu(z). \tag{6.7}$$

It should be clear from the previous lemma that the function to be integrated can be written as a linear combination of $\partial_w^i D_w(z)$ evaluated at $w = \beta$ for $i = 0,\ldots,l$ and $[\partial_w^j D_w(z)]_*$ evaluated at $w = \alpha$ for $j = 0,\ldots,k$. Thus (6.7) can be written as a linear combination of $\Omega^{(i)}(\beta)$ and $\overline{\Omega^{(j)}(\alpha)}$ for $i = 0,\ldots,l$ and $j = 0,\ldots,k$.

The other possibilities for $w_i$ and $w_j$ can be treated similarly □

We shall call a matrix of the form $G_n$ as in the previous theorem a *generalized Pick matrix.* The Gram matrix for any other basis, e.g., the basis $\{B_k : k = 0,\ldots,n\}$ is of course equivalent with a generalized Pick matrix. Turning this argument around, we can show that every generalized Pick matrix is equivalent with a Toeplitz matrix. This is explicitly done in the next theorem which is due to Delsarte, Genin and Kamp [23].

**Theorem 6.4** *A generalized Pick matrix is equivalent with a Toeplitz matrix.*

**Proof.** We just note that also

$$\{\ell_k^{(n)} : k = 0,\ldots,n\} = \{\frac{z^k}{\pi_n(z)}; k = 0,\ldots,n\}, \quad \pi_n(z) = \prod_{i=1}^n (1-\overline{\alpha}_i z)$$

is a basis for $\mathcal{L}_n$. The Gram matrix for this basis is obviously a Toeplitz matrix since its entries depend only on the difference of the indices (and $n$) :

$$\langle \ell_i^{(n)}, \ell_j^{(n)} \rangle_\mu = \int \frac{z^{i-j}}{\pi_{n*}\pi_n}d\mu = \langle z^i, z^j \rangle_\pi$$

with $d\pi = (\pi_{n*}\pi_n)^{-1}d\mu$.

It is not difficult to find the transformation matrix explicitly, at least not in the case where all the $\alpha_i$ are different. Suppose we express the basis functions $w_i$ in terms of $\ell_j^{(n)}$ as

$$w_i = \sum_{j=0}^{n} \ell_j^{(n)} a_{ij}, \quad i = 0, \ldots, n.$$

which we can also express as

$$w_i(z)\pi_n(z) = \sum_{j=0}^{n} z^j a_{ij}, \quad i = 0, \ldots, n$$

Since this is a relation between polynomials of degree at most $n$, we can take the transform $p^*(z) = z^n\overline{p(1/\overline{z})}$ to get

$$\pi_{ni}(z) = \sum_{j=0}^{n} z^{n-j}\overline{a}_{ij}, \quad i = 0, \ldots, n \tag{6.8}$$

where $\pi_{ni} = (w_i\pi_n)^*$ is a polynomial of degree at most $n$. So we get again a polynomial relation. We can now find the coefficients $a_{ij}$ by requiring that the polynomial on the right interpolates the polynomial on the left in the points $\{\alpha_i : i = 0, \ldots, n\}$ where we have supposed that $\alpha_0 = 0$. Interpolation has to be understood in Hermite sense. This means that, if a point $\alpha_i$ appears $\nu_i$ times, then one interpolates the first $(\nu_i - 1)$ derivatives. In case of all different $\alpha_i$, we get ordinary interpolation. The result then is

$$\Pi = VA^H$$

where $A = [a_{ij}]$, $V$ is a Vandermonde matrix $V = [\alpha_i^{n-j}]$ and $\Pi$ is a diagonal matrix

$$\Pi = \mathrm{diag}[\pi_{ni}] \text{ with } \pi_{ni} = \prod_{j\neq i, j=0}^{n} (\alpha_i - \alpha_j).$$

In the case of confluent points, the matrices $\Pi$ and $V$ can still be found, only they are much more complicated to write down. If we set

$$W^T = [w_0, \ldots, w_n] \text{ and } L^T = [\ell_0^{(n)}, \ldots, \ell_n^{(n)}]$$

with $W = AL$, then the Pick matrix is given by

$$P_n = \langle W, W^T\rangle_\mu$$

and the Toeplitz matrix

$$T_n = \langle L, L^T\rangle_\mu$$

will be related by

$$P_n = A\langle L, L^T\rangle_\mu A^H = AT_nA^H = \Pi^H V^{-H} T_n V^{-1}\Pi.$$

This concludes the proof of our theorem. □

From the proof of the previous theorem, it follows that another interesting basis is

$$\{n_k : k = 0, \ldots, n\} = \left\{\frac{1}{\pi_n}, \frac{z}{\pi_n}, \frac{z(z-\alpha_1)}{\pi_n}, \ldots, \frac{z(z-\alpha_1)\cdots(z-\alpha_{n-1})}{\pi_n}\right\}$$

where as always $\pi_n = \prod_{i=1}^{n}(1-\overline{\alpha}_i z)$. The coefficient $a_{ij}$ from the previous theorem is then a $j$-th order divided difference of the function $\pi_n w_i$ at the points $1/\overline{\alpha}_i : i = 0, \ldots, j$. Again, here we can have confluency. See the book of Donoghue [29].

Some other bases can be found in the literature like e.g.,

$$\left\{1, \frac{\gamma_1}{1-\overline{\alpha}_1 z}, B_1 \frac{\gamma_2}{1-\overline{\alpha}_2 z}, \ldots, B_{n-1} \frac{\gamma_n}{1-\overline{\alpha}_n z}\right\} \tag{6.9}$$

with $\gamma_k = (1-|\alpha_k|^2)^{1/2}$. It is an orthonormal basis for $\mathcal{L}_n$ w.r.t. the Lebesgue measure if all the points $\alpha_i, i = 1, \ldots$ are different and nonzero [7, p.138].

If the points $\alpha_i$ are renamed like in (6.5) as $\beta_j$, then the basis

$$\left\{\frac{1}{1-\overline{\beta}_i z}\left(\frac{z-\beta_i}{1-\overline{\beta}_i z}\right)^k : k = 0, \ldots, \nu_i - 1; i = 1, \ldots, m\right\}$$

has also attracted a lot of attention. See e.g. [46, p.149 ff] where expressions for the elements of the corresponding Gram matrix (w.r.t. Lebesgue measure) can be found. See also [72, p.27]. In [27] it is called the Malmquist basis.

Let us elaborate a bit further on this orthonormal basis in $H_2$. Let us define as in (6.9)

$$v_k = B_{k-1} \frac{\gamma_k}{1-\overline{\alpha}_k z}; \quad k = 1, \ldots, n. \tag{6.10}$$

Then this forms, together with $v_0 = 1$, an orthonormal basis in $H_2$ for $\mathcal{L}_n$ if all the $\alpha_i$ are mutually different and nonzero. However, it holds in general, also if some points coincide or are equal to zero that

$$\mathcal{L}_n = \operatorname{span}\{1, (z-w)v_1, (z-w)v_2, \ldots, (z-w)v_n)\}$$

for any $w$. The interesting thing about this basis is that we can now write

$$\mathcal{L}_n(w) = \operatorname{span}\{(z-w)v_1, \ldots, (z-w)v_n)\} = \{f \in \mathcal{L}_n : f(w) = 0\}.$$

Define

$$\begin{aligned} \mathcal{K}_n(w) &= \operatorname{span}\{v_1, \ldots, v_n\} = \{(z-w)^{-1} f : f \in \mathcal{L}_n(w)\} = \\ \mathcal{L}_n(0)^* &= \{f \in \mathcal{L}_n : f^* \in \mathcal{L}_n(0)\} = \\ &= \{p/\pi_n : p \in \Pi_{n-1}, \pi_n = \prod_{i=1}^n (1-\overline{\alpha}_i z)\} \end{aligned}$$

which is of course independent of $w : \mathcal{K}_n(w) = \mathcal{K}_n(0)$. Define

$$z^{-1}\mathcal{M}_n = \mathcal{M}_n(0) = B_n H_2,$$

then we can prove the following property (see also Walsh [79, p. 225]).

**Theorem 6.5** *With the spaces as defined above we have that*

$$\mathcal{K}_n(0) = \{z^{-1} f : f \in \mathcal{L}_n, f(0) = 0\} = \{f \in \mathcal{L}_n : f^*(0) = 0\} = \mathcal{L}_n(0)^*$$

*is the orthogonal complement (w.r.t. the Lebesgue measure) of $\mathcal{M}_n(0) = B_n H_2$. Thus*

$$\mathcal{K}_n(0) = H_2 \ominus B_n H_2 = \mathcal{M}_n(0)^\perp.$$

*Similarly, it also holds that*

$$\mathcal{L}_n = H_2 \ominus z B_n H_2 = H_2 \ominus \mathcal{M}_n = \mathcal{M}_n^\perp.$$

**Proof.** Take a function from $B_nH_2$ which is of the form $B_nf$ with $f \in H_2$ and a function from $\mathcal{K}_n(0)$ which has the form $z^{-1}p_n(z)/\pi_n(z)$ where $p_n(z) \in \Pi_n(0)$, a polynomial of degree at most $n$ which vanishes at the origin and as before $\pi_n = \prod_1^n(1-\overline{\alpha}_i z)$. The inner product of these functions equals

$$\int B_n(z)f(z)\frac{zp_{n*}(z)}{\pi_{n*}(z)}d\lambda = \eta_n \int f(z)\frac{zp_n^*(z)}{\pi_n(z)}d\lambda, \qquad \eta_n = \prod_{i=1}^n(-\overline{\alpha}_i/|\alpha_i|)$$

where $p_n^*(z) = \overline{p_n(1/\overline{z})}$ and $B_n = \eta_n\pi_n^*/\pi_n$. Since $fp_n^*/\pi_n$ is analytic in $\mathbf{D}$, this integral is zero.

The other orthogonality relation follows similarly. An element from $\mathcal{L}_n$ can be written as $c+zk_n$ with $c \in \mathbf{C}$ and $k_n \in \mathcal{K}_n(0)$. We should prove that $\langle zf, c+zk_n\rangle = 0$ for all $f \in B_nH_2 = \mathcal{M}_n(0)$, $c \in CC$ and $k_n \in KK_n(0)$. The inner product equals $\langle zf, c\rangle + \langle zf, zk_n\rangle$. The second term vanishes by the first part of this theorem. The first term equals $\overline{c}\int zf(z)d\lambda(z) = 0$ since $f$ is analytic. Hence the theorem is proved. □

The previous result says for example that a function in $H_2$ is orthogonal to $\mathcal{L}_n$ if and only if it vanishes in the point set $A_n^0 = \{0, \alpha_1, \ldots, \alpha_n\}$, thus the difference between a function $f \in H_2$ and its orthogonal projection onto $\mathcal{L}_n$ should vanish in $A_n^0$. In other words, the orthogonal projection of $f \in H_2$ onto $\mathcal{L}_n$ should interpolate $f$ in the points $A_n^0$. We shall come back to this property in Section 9.

The next result says that we can also find a basis for $\mathcal{L}_n$ from its reproducing kernel.

**Theorem 6.6** *Let $k_n(z,w)$ be a reproducing kernel for $\mathcal{L}_n$. Then for a set $\{\xi_0, \ldots, \xi_n\}$ of distinct points in $\mathbf{D}$, the functions $\{k_n(z,\xi_j)\}$, $j = 0, \ldots, n$ form a basis for $\mathcal{L}_n$.*

**Proof.** Certainly, the functions are all in $\mathcal{L}_n$. They are also linearly independent. Suppose they were not, then there is a nonzero vector $A = [a_0, \ldots, a_n]^T \in \mathbf{C}^n$ such that for all $z$

$$\sum_{j=0}^n k_n(z,\xi_j)a_j = 0$$

Take $z = \xi_i$ for $i = 0, \ldots, n$, then this would mean that the following system

$$\sum_{j=0}^n k_n(\xi_i,\xi_j)a_j = 0, \quad i = 0, \ldots, n$$

has a non trivial solution. However, since the matrix

$$[k_n(\xi_i,\xi_j)] = [\langle k_n(\cdot,\xi_j), k_n(\cdot,\xi_i)\rangle_\mu],$$

this matrix is positive definite the system has only the zero solution. Thus the functions $\{k_n(z,\xi_j), j = 0, \ldots, n\}$ form indeed a basis for $\mathcal{L}_n$. In fact, a positive definite (Gram) matrix is basically equivalent to a reproducing kernel. See [8, p. 344] or [29, chap. 10]. □

# 7 Calculus in $\mathcal{L}_n$

Recall that we already defined the substar transform $f_*(z) = \overline{f(1/\overline{z})}$. Now if $f \in \mathcal{L}_n$, we shall also define the superstar transform as $f^*(z) = B_n(z) f_*(z)$ where $B_n$ is the finite Blaschke product with zeros from $A_n = \{\alpha_i : i = 1, \ldots, n\}$. Note that if all $\alpha_i$ are zero, $\mathcal{L}_n = \Pi_n$ and then it is natural to define the superstar transform for polynomials as $p^*(z) = z^n p_*(z)$ if $p \in \Pi_n$. Thus

$$\left(\sum_{i=0}^{n} a_k z^k\right)^* = \sum_{i=0}^{n} \overline{a}_i z^{n-i} = \overline{a}_n + \overline{a}_{n-1} z + \cdots + \overline{a}_0 z^n.$$

For functions in $\mathcal{L}_n$, this generalizes to

$$\left(\sum_{i=0}^{n} a_i B_i(z)\right)^* = \sum_{i=0}^{n} \overline{a}_i B_{n\setminus i}(z) = \overline{a}_n + \overline{a}_{n-1}\zeta_n + \cdots + \overline{a}_0 B_n$$

where

$$B_{n\setminus i}(z) = \frac{B_n(z)}{B_i(z)} = \prod_{j=i+1}^{n} \zeta_j(z), \qquad 0 \le i \le n.$$

Note also that if we define $\pi_n(z) = \prod_{i=1}^{n}(1 - \overline{\alpha}_i z)$, then we can write $B_n(z)$ as

$$B_n(z) = \frac{\pi_n^*(z)}{\pi_n(z)} \eta_n \text{ with } \eta_n = \prod_{i=1}^{n} \left(\frac{-\overline{\alpha}_i}{|\alpha_i|}\right) \in \mathbf{T}.$$

Since in the polynomial case, $\overline{p^*(0)}$ is the *leading coefficient* of the polynomial $p$, we shall generalize this concept and call $\overline{f_n^*(\alpha_n)}$ the leading coefficient of $f_n \in \mathcal{L}_n$, i.e.

$$f_n = \overline{f_n^*(\alpha_n)} B_n + a_{n-1} B_{n-1} + \cdots + a_0$$

If $f_n^*(\alpha_n) = 1$, we shall say that it is *monic*. The following technical properties can be trivially verified but they are very useful if you want to do computations in $\mathcal{L}_n$.

**Theorem 7.1** *In $\mathcal{L}_n$, the following relations hold.*

1. *If $f \in \mathcal{L}_n$, then*
   - (a) $(f_*)_* = (f^*)^* = f$
   - (b) $(f_*)^* = f B_n$
   - (c) $(f^*)_* = f B_{n*} = f/B_n$
2. *For the finite Blaschke products we have*
   - (a) $B_n^* = 1$
   - (b) $B_{n*} = 1/B_n$
3. *Define $\pi_n(z) = \prod_{i=1}^{n}(1 - \overline{\alpha}_i z)$, and let $f \in \mathcal{L}_n$ be given by $f = p_n/\pi_n$, with $p_n$ a polynomial. Then*
   - (a) $B_n = \eta_n \pi_n^*/\pi_n$ *with* $\eta_n = \prod_{i=1}^{n}(-\overline{\alpha}_i/|\alpha_i|)$
   - (b) $f_* = p_{n*}/\pi_{n*} = p_n^*/\pi_n^*$
   - (c) $f^* = \eta_n p_n^*/\pi_n$.
4. *If $f, g \in \mathcal{L}_n$, then*
   $\langle f, g\rangle_\mu = \langle f_*, g_*\rangle_\mu = \langle g^*, f^*\rangle_\mu$

5. *If $\phi_n \in \mathcal{L}_n$ and $\langle \phi_n, \mathcal{L}_{n-1}\rangle_\mu = 0$, i.e. $\phi_n \perp \mathcal{L}_{n-1}$, then $\langle \phi_n^*, \zeta_n \mathcal{L}_{n-1}\rangle_\mu = 0$, i.e. $\phi_n^* \perp \zeta_n \mathcal{L}_{n-1}$ or in the notation of the last section $\phi_n^* \perp \mathcal{L}_n(\alpha_n)$ where $\mathcal{L}_n(\alpha_n) = \{f \in \mathcal{L}_n : f(\alpha_n) = 0\}$.*

6. *If $f \in \mathcal{L}_n$ has a zero/pole $\beta$ in $\mathbf{D}$ ($\mathbf{E}$,$\mathbf{T}$), then $f_* \in \mathcal{L}_{n*}$ has a zero/pole $1/\overline{\beta}$ in $\mathbf{E}$ ($\mathbf{D}$,$\mathbf{T}$) and $f^* \in \mathcal{L}_n$ has the same zeros as $f_*$ and the same poles as $f^*$. The latter are the elements from $A_{n*} = \{1/\overline{\alpha}_1, \ldots, 1/\overline{\alpha}_n\}$.*

**Proof.** These results are so simple to verify that we do not give the proof explicitly. □

Let us now introduce the orthonormal basis $\{\phi_k\}$ for the space $\mathcal{L}_n$ with respect to the inner product $\langle\cdot,\cdot\rangle_\mu$. This means

$$\mathcal{L}_n = \operatorname{span}\{\phi_k : k = 0, \ldots, n\}$$

while $\langle \phi_i, \phi_j\rangle_\mu = \delta_{ij}$ We can always choose the leading coefficient $\kappa_n = \overline{\phi_n^*(\alpha_n)}$ of $\phi_n$ to be real and positive. We shall use from now on the notation $\kappa_n$ to denote this coefficient. The functions $\Phi_n = \kappa_n^{-1}\phi_n$ are then the monic orthogonal basis functions.

From a previous section we know that the reproducing kernel for $\mathcal{L}_n$ is given by

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

We can obtain the following determinant expressions.

**Theorem 7.2** *Let $E_n^T = [e_0, \ldots, e_n]$ be a vector of basis functions for $\mathcal{L}_n$ and let $G_n$ denote the Gram matrix w.r.t. $\mu$:*

$$G_n = \langle E_n, E_n^T\rangle_\mu = [\langle e_i, e_j\rangle_\mu].$$

*Then the reproducing kernel $k_n(z,w)$ for $\mathcal{L}_n$ is expressed as*

$$k_n(z,w) = \frac{-1}{\det G_n} \det \begin{bmatrix} G_n & E_n(z) \\ E_n(w)^H & 0 \end{bmatrix}.$$

**Proof.** Suppose that by a Gram-Schmidt orthogonalization procedure, the basis $E_n$ is transformed into an orthonormal basis $F_n = [\phi_0, \ldots, \phi_n]^T$ and that

$$\phi_k = \sum_{i=0}^{k} l_{ki} e_i.$$

then the vector $F_n$ can be written as $F_n = L_n E_n$ with $L_n$ the lower triangular matrix containing the $l_{ki}$ coefficients. Now express that these $\phi_k$ form an orthonormal set. Then you get

$$I = \langle F_n, F_n^T\rangle_\mu = L_n \langle E_n, E_n^T\rangle_\mu L_n^H = L_n G_n L_n^H.$$

Hence $G_n^{-1} = L_n^H L_n$. Therefore

$$k_n(z,w) = F_n(w)^H F_n(z) = E_n(w)^H L_n^H L_n E_n(z) = E_n(w)^H G_n^{-1} E_n(z).$$

The last line gives the result by a standard argument for determinants. □

From this fact the following simple but basic relations can be derived.

**Theorem 7.3** *Let $k_n(z,w)$ be the reproducing kernel for $\mathcal{L}_n$ which is based on the points $A_n = \{\alpha_i : i = 1, \ldots, n\}$ and $\{\phi_k\}$ a set of orthonormal basis functions with leading coefficients $\kappa_k > 0$. Then the following relations hold.*

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. $k_n(\alpha_n, \alpha_n) = \kappa_n^2$.

**Proof.**

1. If $\{e_0, \ldots, e_n\}$ is a basis for $\mathcal{L}_n$ with Gram matrix $G_n$, then $\{B_n e_{k*} : k = 0, \ldots, n\}$ is also a basis with Gram matrix $G_n^T$. Indeed,

$$\langle B_n e_{i*}, B_n e_{j*} \rangle_\mu = \langle e_j, e_i \rangle_\mu.$$

Let us denote this basis in a vector form, which, according to our earlier convention of substar transform for matrices, should be denoted as

$$B_n(z)E_{n*}(z) = B_n(z)[e_{0*}(z), \ldots, e_{n*}(z)] \text{ if } E_n^T(z) = [e_0(z), \ldots, e_n(z)].$$

Hence, when applying the determinant formula of the previous theorem for this new basis, we get

$$\begin{aligned} k_n(z, w) &= \frac{-1}{\det G_n} \det \left[ \begin{array}{cc} G_n^T & B_n(z)E_{n*}(z)^T \\ \hline \overline{B_n(w)}\, \overline{E_{n*}(w)} & 0 \end{array} \right] \\ &= \frac{-B_n(z)\overline{B_n(w)}}{\det G_n} \det \left[ \begin{array}{cc} G_n & E_{n*}(w)^H \\ E_{n*}(z) & 0 \end{array} \right] \\ &= B_n(z)\overline{B_n(w)} k_n(1/\overline{w}, 1/\overline{z}) \end{aligned}$$

2. The first part gives that

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

where $B_{n\backslash k} = B_n/B_k$. For $w \to \alpha_n$, we see that $B_{n\backslash k}(\alpha_n) = \delta_{kn}$, so that

$$k_n(z, \alpha_n) = \phi_n^*(z)\overline{\phi_n^*(\alpha_n)} = \phi_n^*(z)\kappa_n.$$

3. If you also let $z$ tend to $\alpha_n$, then you directly get the third result.

□

We can now also obtain determinant expressions for $\kappa_n$ and $\phi_n$.

**Theorem 7.4** *Consider the basis of finite Blaschke products $\{B_k(z) : k = 0, \ldots, n\}$ for the space $\mathcal{L}_n$ and denote the elements from the corresponding Gram matrix as*

$$\mu_{ij} = \langle B_i, B_j \rangle_\mu$$

*and set $G_n = [\mu_{ij}]$. The orthonormal basis functions are denoted as $\phi_k$ and their leading coefficient as $\kappa_k$. Then the following equalities hold*

1. $\kappa_n^2 = \det G_{n-1} / \det G_n$,

2. $\phi_n(z) = (\det G_{n-1} \det G_n)^{-1/2} \det \begin{bmatrix} \mu_{00} & \cdots & \mu_{0n} \\ \vdots & & \vdots \\ \mu_{n-1,0} & \cdots & \mu_{n-1,n} \\ B_0(z) & \cdots & B_n(z) \end{bmatrix}$.

**Proof.** We know from the previous theorem that

$$k_n(z,\alpha_n) = \frac{-B_n(z)}{\det G_n} \det \begin{bmatrix} & & \gamma_0 \\ & G_n & \vdots \\ & & \gamma_n \\ B_{0*}(z) \cdots B_{n*}(z) & & 0 \end{bmatrix}$$

with $\gamma_k = B_{n\backslash k}(\alpha_n) = \delta_{kn}$. Hence we get

$$k_n(z,\alpha_n) = \kappa_n \phi_n^*(z) = \frac{B_n(z)}{\det G_n} \det \begin{bmatrix} \mu_{00} & \cdots & \mu_{0n} \\ \vdots & & \vdots \\ \mu_{n-1,0} & \cdots & \mu_{n-1,n} \\ B_{0*}(z) & \cdots & B_{n*}(z) \end{bmatrix}$$

which then gives first

$$\begin{aligned} k_n(\alpha_n,\alpha_n) = \kappa_n^2 &= \frac{1}{\det G_n} \det \begin{bmatrix} \mu_{00} & \cdots & \mu_{0n} \\ \vdots & & \vdots \\ \mu_{n-1,0} & \cdots & \mu_{n-1,n} \\ \gamma_0 & \cdots & \gamma_n \end{bmatrix} \\ &= \frac{\det G_{n-1}}{\det G_n} \end{aligned}$$

because $\gamma_k = \delta_{kn}$. So that also

$$\phi_n(z) = \kappa_n^{-1} k_n(z,\alpha_n)^* = (\det G_{n-1} \det G_n)^{-1/2} \det \begin{bmatrix} \mu_{00} & \cdots & \mu_{0n} \\ \vdots & & \vdots \\ \mu_{n-1,0} & \cdots & \mu_{n-1,n} \\ B_0 & \cdots & B_n \end{bmatrix}$$

which concludes the proof. $\square$

## 8 Extremal problems in $\mathcal{L}_n$

In this section we shall review some of the extremal problems that can be solved with reproducing kernels in the special case of $\mathcal{L}_n$. From now on we shall use the notation $\|\cdot\|_\mu$ to mean $\langle\cdot,\cdot\rangle_\mu^{1/2}$. From the problem $P^1(a,w)$ which we considered in section 3, we can now derive

**Theorem 8.1** *All the solutions of the following optimization problem*

$$P_n^1(1,\alpha_n): \qquad \sup\{|f(0)|^2 : \|f\|_\mu = 1, f\in\mathcal{L}_n\}$$

*are given by*

$$f = \eta\phi_n^* \quad , \quad |\eta| = 1$$

*where $\phi_n$ is the $n$-th orthonormal basis function of $\mathcal{L}_n$ with leading coefficient $\kappa_n = \phi_n^*(\alpha_n)$. The maximum is equal to $\kappa_n^2$.*

**Proof.** This follows immediately from the Theorem 3.2 and the properties given in Theorem 7.3. □

Also for the second optimization problem of Theorem 3.2 we formulate a special case.

**Theorem 8.2** *The optimization problem*

$$P_n^2(1,\alpha_n): \qquad \inf\{\|f\|_\mu^2 : f(\alpha_n) = 1, f\in\mathcal{L}_n\}$$

*has a unique solution given by*

$$f = \Phi_n^* = \kappa_n^{-1}\phi_n^*$$

*where $\Phi_n$ is the $n$-th monic orthogonal basis function in $\mathcal{L}_n$, $\phi_n$ the orthonormal one with $\phi_n^*(\alpha_n) = \kappa_n > 0$. The minimal value is $\kappa_n^{-2}$.*

**Proof.** Also this follows from Theorem 3.2 and the properties in Theorem 7.3. □

Since in $\mathcal{L}_n$ it holds that

$$\|f\|_\mu = \|f^*\|_\mu$$

we also have solved the following problem.

**Corollary 8.3** *The unique solution of the problem*

$$\inf\{\|f\|_\mu^2 : f\in\mathcal{L}_n^M\}$$

*where $\mathcal{L}_n^M$ denotes all monic elements of $\mathcal{L}_n$ is given by the $n$-th monic orthogonal basis function $\Phi_n = \kappa_n^{-1}\phi_n$ of $\mathcal{L}_n$ and the minimum is $\kappa_n^{-2}$, with $\kappa_n > 0$ the leading coefficient of the orthonormal one.*

Recall the definition of $\mathcal{L}_n(w)$, given in section 6

$$\mathcal{L}_n(w) = \{(z-w)f : f\in\mathcal{K}_n(w)\} = \{f : f\in\mathcal{L}_n : f(w) = 0\}.$$

The problem of finding the orthogonal projection of 1 onto $\mathcal{L}_n(w)$ is related to a classical Szegő problem.

**Theorem 8.4** *Define the following problem in $\mathcal{L}_n(w)$ which is defined as above with $\mathcal{L}_n$ the rational function space based on the point set $A_n = \{\alpha_1,\ldots,\alpha_n\}$.*

$$P_n^3(w): \qquad \inf\{\|1-f\|_\mu^2 : f\in\mathcal{L}_n(w)\}.$$

*Then $P_n^3(w)$ has the unique solution*

$$f = 1 - \frac{k_n(z,w)}{k_n(w,w)}$$

*where $k_n(z,w)$ is the reproducing kernel of $\mathcal{L}_n$. The minimum is given by $[k_n(w,w)]^{-1}$.*

**Proof.** This problem can be reduced to problem $P_n^2(1,w)$ by noting that

$$\mathcal{L}_n(w) = \{f = 1 - g : g \in \mathcal{L}_n, g(w) = 1\}$$

and thus is

$$\inf\{\|1-f\|_\mu^2 : f \in \mathcal{L}_n(w)\} = \inf\{\|g\|_\mu^2 : g \in \mathcal{L}_n, g(w) = 1\}, \qquad f = 1 - g.$$

From this the result follows easily. □

This theorem has an easy corollary.

**Corollary 8.5** *If $k_n(z,w)$ is the reproducing kernel for $\mathcal{L}_n$, then $k_n(w,w)$ is nondecreasing with $n$ if $w \in \mathbf{D}$.*

**Proof.** Since the $\mathcal{K}_n(w)$ are nested as $\mathcal{K}_n(w) \subset \mathcal{K}_{n+1}(w)$, the minimum $[k_n(w,w)]^{-1}$ can not increase with $n$. □

Since there are so many optimization problems whose solutions can be expressed in terms of the reproducing kernel, one can ask whether there is an optimization problem that has this kernel for its solution. It turns out that it gives an approximation to the spectral factor that can be related to the measure $\mu$ as explained in section 2. More precisely, it approximates some function $s_w(z)$, depending on $w$ which is related to the spectral factor. This function is defined in the next Theorem. It is known as the *Szegő kernel* associated with $\mu$. See [42, p.51-52].

**Theorem 8.6** *Let $\sigma$ be the spectral factor, related to the measure $\mu$ as explained in section 2. That is we suppose $\sigma_*\sigma = \mu'$. Define for $w \in \mathbf{D}$ the function $s(z,w) = s_w(z)$ with*

$$s_w(z) = [\sigma(z)(1-\overline{w}z)\overline{\sigma(w)}]^{-1}, \qquad w \in \mathbf{D}.$$

*Define the problem*

$$S_n(w): \qquad \inf\{\|f - s_w\|_\mu^2 : f \in \mathcal{L}_n\}.$$

*Then $S_n(w)$ has a solution $f(z) = k_n(z,w)$, where $k_n(z,w)$ is the reproducing kernel for $(\mathcal{L}_n,\mu)$ and the minimum is given by*

$$[(1-|w|^2)|\sigma(w)|^2]^{-1} - k_n(w,w).$$

**Proof.** First note that $s_w \in H_2$ since $1/\sigma \in H_2$ and $w \in \mathbf{D}$ and $\sigma(w) \neq 0$. Now you can also reduce this problem $S_n(w)$ to the problem $P_n^2(1,w)$ by observing that

$$\|f - s_w\|_\mu^2 = \|f\|_\mu^2 + \|s_w\|_\mu^2 - 2\Re\,\langle f, s_w\rangle_\mu.$$

The last term can be obtained by the following computation

$$\begin{aligned}
\langle f, s_w\rangle_\mu &= \int f(z)\frac{1}{\sigma_*(z)\sigma(w)}\frac{z}{z-w}d\mu(z)\\
&= \int f(z)\frac{\sigma(z)}{\sigma(w)}\frac{z}{z-w}d\lambda(z)\\
&= f(w).
\end{aligned}$$

The second equality follows from the fact that $1/\sigma = 0$ $d\mu_s$-a.e. where $d\mu_s = d\mu - \mu' d\lambda$. On the other hand, using again the fact that $1/\sigma$ vanishes $d\mu_s$-a.e., we get

$$\begin{aligned}
\|s_w\|_\mu^2 &= \frac{1}{|\sigma(w)|^2}\int \frac{d\mu(z)}{\sigma(z)\sigma_*(z)|1-\overline{w}z|^2}\\
&= \frac{1}{|\sigma(w)|^2}\int \frac{d\lambda(z)}{|1-\overline{w}z|^2}\\
&= \frac{1}{|\sigma(w)|^2}\frac{1}{1-|w|^2}
\end{aligned}$$

It is now clear that we have to minimize

$$\inf_a \{ \inf_{f\in\mathcal{L}_n, f(w)=a} \|f\|_\mu^2 - 2\Re\, a\} = \inf_a \{\frac{|a|^2}{k_n(w,w)} - 2\Re\, a\}$$

The solution of the latter problem is easily seen to be found for $a = k_n(w,w)$, so that we get the solution of the Theorem. □

This Theorem can also be reformulated as follows.

**Corollary 8.7** *Let the measure $\mu$ of the previous theorem satisfy $d\mu = P(\cdot, w)d\nu$ with $P(z,w)$ the Poisson kernel. Denote by $\sigma_\mu$ and $\sigma_\nu$ the spectral factors of $\mu$ and $\nu$ respectively. Then the problem*

$$\inf\{\|f - [\sigma_\nu \overline{\sigma_\nu(w)}]^{-1}\|_\mu^2 : f \in \mathcal{L}_n\}$$

*reaches a minimum $|\sigma_\nu(w)|^{-1} - k_n(w,w)$ for $f(z) = k_n(z,w)$ where $k_n$ is the reproducing kernel for $\mathcal{L}_n$ w.r.t. $d\mu$.*

**Proof.** Note that the outer spectral factors are related by

$$\sigma_\mu = \sigma_\nu \frac{\sqrt{1-|w|^2}}{1-\overline{w}z}.$$

Fill this into the expression for $s_w$ of the previous Theorem to find that it is equal to $[\sigma_\nu(z)\overline{\sigma_\nu(w)}]^{-1}$. The result then follows easily. □

# 9 Density in $L_p$ and $H_p$

As with the powers of $z$, it is possible to prove that the basis functions of finite Blaschke products and their inverses are complete in the space $L_p$ if and only if $\sum(1-|\alpha_k|) \to \infty$. In analogy with the powers of $z$ we can define the finite Blaschke products $B_n$ for $n = 0, 1, \ldots$ as before and we set by definition $B_{-n} = B_{n*} = 1/B_n$ for $n = 1, 2, \ldots$ Hence, the $\{B_n\}_{n\in\mathbf{Z}}$ span the spaces $\mathcal{R}_n = \mathcal{L}_n + \mathcal{L}_{n*}$, $n = 0, 1, \ldots$ The completeness of the functions $\{B_n\}_{n\in\mathbf{Z}}$ is the same as the density of $\mathcal{R}_\infty$ in $L_p$. If $\mathcal{R}$ denotes the closure in $L_p$ of $\mathcal{R}_\infty$, then if $\mathcal{R}_\infty$ is dense in $L_p$, $\mathcal{R}$ should coincide with $L_p$.

In order to prove this density property, we start with a Lemma that can be found in Achieser [1, p. 243].

**Lemma 9.1** *Let $z_1, z_2, \ldots, z_n$ be some fixed points in $\mathbf{C}$. We have the following optimization result in $L_p$, $0 < p \leq \infty$*

$$\min_q \left\{ \|f\|_p : f(z) = \frac{q(z)}{(z-z_1)\cdots(z-z_n)}, q \in \Pi_N^M, N \geq n \right\} = \prod_{k=1}^{n} \frac{1}{|z_k|_+} \tag{9.1}$$

*where $q$ ranges over all monic polynomials of degree at most $N$ ($q \in \Pi_N^M$) and $|z|_+ = \max\{|z|, 1\}$. The unique solution is obtained for $q = Q$ with $Q$ as in (9.2) below.*

**Proof.** Suppose for simplicity that $z_1, \ldots, z_d$ are all in $\mathbf{E}$ while $z_{d+1}, \ldots, z_n$ are all in $\overline{\mathbf{D}}$. It is clear that the right hand side value can be reached. It is obtained for $q = Q$ with

$$Q(z) = \frac{z^{N-n}(z\overline{z}_1 - 1)\cdots(z\overline{z}_d - 1)(z - z_{d+1})\cdots(z - z_n)}{\overline{z}_1 \cdots \overline{z}_d} \tag{9.2}$$

and this is the unique solution. Thus we have to show that for any other monic polynomial $P \in \Pi_N^M$

$$\left\| \frac{P(z)}{\pi(z)} \right\|_p > \left\| \frac{Q(z)}{\pi(z)} \right\|_p \tag{9.3}$$

where $\pi(z) = (z - z_1)\cdots(z - z_n)$. Since $\|f\|_\infty \geq \|f\|_p$, it is sufficient to prove that (9.3) is true for $0 < p < \infty$. Suppose that $q_1, \ldots, q_j$ are all the zeros of $Q$ in the closed unit disk. Thus $Q(z) = (z - q_1)\cdots(z - q_j)R(z)$ with $|R(0)| > 1$. The function

$$g(z) = \frac{(1 - \overline{q}_1 z)\cdots(1 - \overline{q}_j z)R(z)}{(z - z_1)\cdots(z - z_d)(1 - \overline{z}_{d+1}z)\cdots(1 - \overline{z}_n z)} \tag{9.4}$$

is analytic in $\mathbf{D}$ and therefore

$$\int \left|\frac{Q}{\pi}\right|^p d\lambda = \int |g|^p d\lambda \geq \left| \int g(z)^p d\lambda \right| = |g(0)|^p = \frac{|R(0)|^p}{|z_1 \cdots z_d|^p} \geq \frac{1}{|z_1 \cdots z_d|^p}$$

with inequality if $P \neq Q$. $\square$

With the previous Lemma, we can now prove the following completeness theorem. It is a slight generalization of a result in Achieser [1, p. 244] where the poles were supposed to be simple. On the other hand Achieser is more general since poles need not appear in reflection pairs $(\alpha_k, 1/\overline{\alpha}_k)$ as we suppose here.

**Theorem 9.2** *For given $\alpha_1, \alpha_2, \ldots$ all in $\mathbf{D}$, let the finite Blaschke products $B_n$ be defined as before for $n \geq 0$ and $B_{-n} = 1/B_n$ for $n = 1, 2, \ldots$ Then the system $\{B_n\}_{n\in\mathbf{Z}}$ is complete in any $L_p$ space ($1 \leq p \leq \infty$) if and only if $\sum(1 - |\alpha_k|) \to \infty$.*

**Proof.** First note that the case where infinitely many $\alpha_k$ are equal to zero can immediately be discarted since in that case the system contains the trigonometric polynomials and these are known to be complete while the sum certainly diverges. So we suppose that only finitely many (say $q$) of the $\alpha_k$ that are zero. Without loss of generality we suppose $\alpha_1 = \cdots = \alpha_q = 0$. In this case the system contains the trigonometric polynomials of degree at most $q$. Suppose we set

$$\begin{aligned}\mathcal{R}_n &= \text{span}\{B_k : k = -n, \ldots, n\} = \mathcal{L}_n + \mathcal{L}_{n*} \\ &= \left\{ \frac{z^{-n} p_{2n}(z)}{D(z)} : p_{2n} \in \Pi_{2n}, D(z) = z^q \prod_{k=q+1}^{n} [(z - 1/\overline{\alpha}_k)(z - \alpha_k)] \right\}.\end{aligned}$$

It is sufficient to prove that the divergence of the sum implies completeness in $L_\infty$ and conversely that completeness in $L_1$ implies divergence of the sum.

*Completeness in $L_1$ ⇒ divergence of the sum.* By the previous Lemma we have for $p = 1$ that

$$\begin{aligned}\inf_{f \in \mathcal{R}_n} \|z^{q+1} - f(z)\|_1 &= \inf_{P \in \Pi_{2n}} \left\| z^{q+1} - \frac{P(z)}{D(z)} \right\|_1 \\ &= \inf_{Q \in \Pi_{2n+1}^M} \left\| \frac{Q(z)}{D(z)} \right\|_1 = \prod_{k=q+1}^{n} |\alpha_k|.\end{aligned}$$

Since the system is supposed to be complete in $L_1$, the previous expression should go to zero. Whence $\sum(1 - |\alpha_k|) \to \infty$.

*Divergence of the sum ⇒ completeness in $L_\infty$.* We already know that all the powers $z^k$, $k = -q, \ldots, q$ are in the system, so we should prove that all $z^k$ for $|k| > q$ can be approximated arbitrarily close in $L_\infty$ by elements from $\mathcal{R}_n$ for $n$ sufficiently large. This follows from the following observations

$$\begin{aligned}&\inf_{f \in \mathcal{R}_n, a_i} \|z^{m+q} + a_1 z^{m+q-1} + \cdots + a_{m-1} z^{q+1} + f(z)\|_\infty \\ &\quad = \inf_{q \in \Pi_{2n+m-1}} \left\| \frac{z^{2n+m} + q(z)}{D(z)} \right\|_\infty = \prod_{k=q+1}^{n} |\alpha_k|\end{aligned}$$

for $m = 1, 2, \ldots$ and

$$\begin{aligned}&\inf_{f \in \mathcal{R}_n, a_i} \|z^{-m-q} + a_1 z^{-m-q+1} + \cdots + a_{m-1} z^{-q-1} + f(z)\|_\infty \\ &\quad = \inf_{f_* \in \mathcal{R}_n, a_i} \|z^{m+q} + \overline{a}_1 z^{m+q-1} + \cdots + \overline{a}_{m-1} z^{q+1} + f_*(z)\|_\infty = \prod_{k=q+1}^{n} |\alpha_k|\end{aligned}$$

for $m = 1, 2, \ldots$ Since the sum diverges to infinity, we must have that the right hand side $\prod_{q+1}^{n} |\alpha_k| \to 0$. By an induction argument, it then follows that $z^{\pm(q+m)}$ for $m = 1, 2, \ldots$ can be approximated arbitrary close by the system $\{B_n\}$. This means that it is complete in $L_\infty$. □

While the previous Theorem was concerned with the density of $\mathcal{R}_\infty$ in $L_p$, we shall now study the density of $\mathcal{L}_\infty$ in $H_2$. It can already be expected from 6.5 which characterized $\mathcal{L}_n$ as $H_2 \ominus z B_n H_2$. If $\sum(1 - |\alpha_k|)$ diverges, the Blaschke product goes to zero in $\mathbf{D}$, which suggests that $\mathcal{L}_n$ becomes $H_2$. In fact the previous characterization links interpolation with least squares approximation as we can find in the next Theorem which can be found in the book of Walsh [79, p. 224].

**Theorem 9.3** *Let $f \in H_2$ be given. Then the following least squares approximation problem*

$$\inf_{f_n} \{\|f - f_n\|_2 : f_n \in \mathcal{L}_n\}$$

*has a unique solution which is the function $f_n \in \mathcal{L}_n$ which interpolates $f$ in the point set $A_n^0 = \{0, \alpha_1, \ldots, \alpha_n\}$.*

**Proof.** The details can be found in Walsh's book, but it may be sufficient to observe that the interpolation error will vanish in the point set $A_n^0$ and will therefore be orthogonal to $\mathcal{L}_n$ (Theorem 6.5). By the orthogonality principle of least squares approximation, this identifies the interpolant as the least squares approximant. □

In fact Walsh observes that we may replace in the previous Theorem $f \in H_2$ by $f_1 \in L_2$. The best least squares approximant is the interpolant for its Cauchy integral

$$f(z) = \int \frac{f_1(t)}{t-z} t d\lambda(t)$$

in the point set $A_n^0$. See Walsh [79, p. 225].

Completely analogous is the following result which uses the other part of Theorem 6.5.

**Theorem 9.4** *Let $f \in H_2$ be given and recall the definition $\mathcal{K}_n(0) = \{f_n \in \mathcal{L}_n : f_n^*(0) = 0\}$. The unique solution of*

$$\inf_{f_n} \{\|f - f_n\|_2 : f_n \in \mathcal{K}_n(0)\}$$

*is the function $f_n \in \mathcal{K}_n(0)$ which interpolates $f$ in the point set $A_n = \{\alpha_1, \ldots, \alpha_n\}$.*

Let us now return to our density problem. The idea that the interpolation error, which now turns out to be also related to the $L_2$ approximation error, is proportional to a Blaschke product, forms the backbone of the following convergence result which can be found again in Walsh [79, p. 305-306].

**Theorem 9.5** *Let $f \in H_2$ and suppose that $\sum(1-|\alpha_k|)$ diverges. Let $f_n \in \mathcal{L}_n$ be the function which interpolates $f$ in the point set $A_n^0 = \{0, \alpha_1, \ldots, \alpha_n\}$. (The zero is in this theorem not essential. It could be replaced by any other $\alpha_0 \in \mathbf{D}$.) Then $f_n$ converges uniformly on compact subsets of $\mathbf{D}$ to $f(z)$.*

*If $f \in H_2$ is also analytic on $\mathbf{T}$, then the convergence is also uniform on $\mathbf{T}$.*

**Proof.** The details of the proof can be found in Walsh's book. The basic idea is to use the error formula

$$f(z) - f_n(z) = \int B_n(z) B_{n*}(t) \frac{f(t)t}{t-z} d\lambda(t).$$

For $z \in \mathbf{D}$, use the fact that $B_n(z)$ converges to zero while the modulus of $B_{n*}(t)$ is 1.

For $z \in \mathbf{T}$, take the integral over a circle slightly larger than the unit circle. Then $|B_n(z)| = 1$ while $(|t| > 1)$

$$B_{n*}(t) = \prod \frac{\alpha_i}{|\alpha_i|} \frac{1 - \overline{\alpha}_i t}{\alpha_i - t} = \prod \frac{\alpha_i}{|\alpha_i|} \frac{\overline{\alpha}_i - 1/t}{1 - \alpha_i/t} \to 0.$$

Whence the Theorem follows. □

The second part of the Theorem implies that the set $\{B_n\}_{n \geq 0}$ is complete in $H_2$. The result is stronger. It says that if the sum diverges, it is possible for an arbitrary polynomial $p$ and any $\epsilon > 0$, to find $n$ sufficiently large such that there is an $f_n \in \mathcal{L}_n$ with $|p - f_n| < \epsilon$ uniformly on $\overline{\mathbf{D}}$.

For a more general positive measure $\mu$, we can also prove the density of $\mathcal{L}_\infty$ in $H_2(\mu)$. Since every element from some $\mathcal{L}_n$ is meromorphic in $\mathbf{C}$ and analytic in $\overline{\mathbf{D}}$ and since $L_2(\mu)$ is a complete metric space [32, p.69], all $\mathcal{L}_n \subseteq H_2(\mu)$. Hence $\mathcal{L}_\infty \subseteq \mathcal{L} \subseteq H_2(\mu)$ where $\mathcal{L}$ denotes now the $L_2(\mu)$ closure of $\mathcal{L}_\infty$. It was shown in [24, Appendix] for an absolutely continuous measure $d\mu = \mu' d\lambda$, that we have equality in the previous inclusion if and only if $\sum(1 - |\alpha_i|)$ diverges. This happens to be the condition for the infinite Blaschke product to go to zero.

We have the following property.

**Theorem 9.6** *Let $\mu$ be a positive measure on $\mathbf{T}$. We have equality of the spaces $(\mathcal{L}, \mu) = H_2(\mu) \equiv (\Pi, \mu)$ if $\sum_1^n (1 - |\alpha_i|)$ diverges for $n \to \infty$. Conversely, let $\log \mu' \in L_1$, then the density of $\mathcal{L}_\infty$ in $L_2(\mu)$ implies the divergence of the sum $\sum_1^n (1 - |\alpha_i|)$ for $n \to \infty$.*

**Proof.** The first part goes as follows. Since we already know that $\mathcal{L} \subseteq H_2(\mu)$, we have only to show that there is no function in $H_2(\mu)$ which is orthogonal to $\mathcal{L}$ except for the zero function. Suppose $f$ is such a function in $H_2(\mu)$ and $\langle f, B_k \rangle_\mu = 0$ for all $k \geq 0$. Thus $\int f(t)/B_k(t) d\mu(t) = 0$ for all $k \geq 0$. This implies that

$$F(\alpha_k) = \int \frac{tf(t)d\mu(t)}{t - \alpha_k} = 0 \text{ for all } k = 0, 1, 2, \ldots$$

where the function

$$F(z) = \int \frac{tf(t)d\mu(t)}{t - z}$$

is analytic inside $\mathbf{D}$ as the Cauchy-Stieltjes integral of the measure $d\nu(t) = f(t)d\mu(t)$ and belongs to $H_p$ for any $p < 1$ [32, Theorem 3.5, p.39]. Suppose that there are infinitely many different $\alpha_i$. Otherwise one needs a slight adaptation to find that the repeated $\alpha_k$ are multiple zeros of $F$. Then we can use the following argument. Since any function in the Nevanlinna class $N \supset H_p$ which has zeros at the points $z = 0, \alpha_1, \alpha_2, \ldots$ and $\sum(1 - |\alpha_i|) = \infty$, vanishes identically in $\mathbf{D}$ [73, Corollary p.335], we have $F(z) \equiv 0$ in $\mathbf{D}$. Now this is equivalent with

$$\int f(t)t^{-k} d\mu(t) = 0 \text{ for all } k = 0, 1, 2, \ldots$$

(compare with [32, Theorem 3.7, p.40]). Thus $f \in H_2(\mu)$, while at the same time, the last relations imply that $f \perp_\mu H_2(\mu)$. Thus $f = 0$ $\mu$-a.e.

For the second part, we can use a similar construct as in [24]. If $\sum(1 - |\alpha_i|) < \infty$, then it is known that $B_n(z)$ converges to a Blaschke product $B(z)$ which is an inner function, i.e., bounded by 1 in $\mathbf{D}$, while its radial limit is equal to 1 a.e. It has zeros in $z = \alpha_1, \alpha_2, \ldots$ Let us take $f(z) = zB(z)\sigma(z)/\sigma_*(z)$, with $\sigma(z)$ the outer spectal factor of $\mu$. Since $|f(t)| = 1$ a.e. on $\mathbf{T}$ we have $f \in L_2(\mu)$. Let $g$ be the orthogonal projection of $f$ onto $H_2$ and define $h = g/\sigma$. Clearly $h \in H_2(\mu)$ because $h\sigma = g \in H_2$ (see [36, Theorem 3.4, p. 215]). On the other hand, it is orthogonal to $\mathcal{L}_\infty$, since the orthogonality $h \perp_\mu \mathcal{L}_\infty$ does not depend on the singular part $\mu_s$ of the measure $\mu$ because

$$\left| \int \frac{g(t)}{\sigma(t)B_k(t)} d\mu_s(t) \right|^2 = |\langle h, B_k \rangle_{\mu_s}|^2 \leq \left\| \frac{g(t)}{B_k(t)} \right\|_{\mu_s}^2 \left\| \frac{1}{\sigma(t)} \right\|_{\mu_s}^2 = \int \left| \frac{g(t)}{B_k(t)} \right|^2 d\mu_s \cdot \int \left| \frac{1}{\sigma(t)} \right|^2 d\mu_s.$$

The last factor is zero because $1/\sigma$ vanishes $d\mu_s$-a.e., and the other factor is finite because $|g/B_k| = 1$ a.e. on $\mathbf{T}$. Hence $\langle h, B_k \rangle_{\mu_s} = 0$ for all $k = 0, 1, \ldots$ Thus the situation is exactly as in the case of an absolutely continuous measure. □

Since we know from the classical Szegő theory that the polynomials are dense in $L_2(\mu)$ if and only if $\log \mu' \notin L_1$, we can combine this with the previous Theorem to get the following Corollary.

**Corollary 9.7** *The space $\mathcal{L}_\infty$ is dense in $L_2(\mu)$ if $\sum(1 - |\alpha_i|)$ diverges and $\log \mu' \notin L_1$.*

**Proof.** The divergence of the sum implies the density of $\mathcal{L}_\infty$ in $H_2(\mu)$, the fact that $\log \mu' \notin L_1$ implies that $H_2(\mu)$ is dense in $L_2(\mu)$. □

The technique used in the first part of the Theorem 9.6 can be used to prove also that $\mathcal{R}_\infty$ is dense in $L_2(\mu)$. So it is possible to show that the following is true.

**Corollary 9.8** *Suppose $\sum_1^n (1 - |\alpha_i|)$ diverges for $n \to \infty$ and that $\mu$ is a positive measure on $\mathbf{T}$. Then $\mathcal{R}_\infty = \operatorname{span}\{B_n\}_{n \in \mathbf{Z}}$ is dense in $L_2(\mu)$.*

**Proof.** Exactly like in the Theorem, we have that $\mathcal{R}_\infty \subseteq L_2(\mu)$ and we have to show that if $f \in L_2(\mu)$ and orthogonal to $\mathcal{R}_\infty$, then it is zero. It follows from the Theorem that if $\int f(t)B_k(t)d\mu(t) = 0$ for all $k \geq 0$, then $\int f(t)t^{-k}d\mu(t) = 0$ for all $k = 0, 1, \ldots$ Similarly, if $\int f(t)B_k(t)d\mu(t) = 0$ for $k = 0, 1, 2, \ldots$, then, using the same assumption of infinitely many different $\alpha_k$, that

$$G(\alpha_k) = \int \frac{f(t)}{1 - t\overline{\alpha}_k} d\mu(t) = \int \frac{tf_*(t)}{t - \alpha_k} d\mu(t) = 0 \text{ for all } k = 0, 1, \ldots$$

where now

$$G(z) = \int \frac{t f_*(t) d\mu(t)}{t - z}$$

is again analytic in $\mathbf{D}$. We then derive in a similar way that also

$$\int f_*(t) t^{-k} d\mu(t) = \int f(t) t^k d\mu(t) = 0 \text{ for all } k = 0, 1, \ldots$$

Thus $f$ is orthogonal to all the elements in $\{t^n\}_{n \in \mathbf{Z}}$, which is complete in $L_2(\mu)$. Thus $f$ is at the same time in $L_2(\mu)$ and orthogonal to it. Hence it is zero. in $\square$

Another kind of density result relates to the representation of positive functions in $L_1$. It is e.g. well known that $f \in L_1$, $f \geq 0$ a.e. on $\mathbf{T}$ and $\log f \in L_1$ if and only if $f = |g|^2$ with $g \in H_2$. In fact, any positive trigonometric polynomial can be written as the square modulus of a polynomial of the same degree. As we know, we can take $g$ to be the spectral factor, which may be chosen an outer function. You find this result in any standard work, e.g., in Grenander and Szegő [42, p. 23-26]. The density of the trigonometric polynomials then implies that any positive function from $L_1$ with integrable logarithm can be approximated arbitrary well by a positive trigonometric polynomial, hence by the square modulus of an outer polynomial can be generalized as follows.

**Theorem 9.9** *Let $\sum(1 - |\alpha_i|)$ diverge and $f \in L_1$. Suppose $f \geq 0$ a.e. on $\mathbf{T}$ and $\log f \in L_1$. Then for every $\epsilon > 0$, there is some $f_n \in \mathcal{L}_n$ for $n$ sufficiently large such that $\|f - |f_n|^2\|_1 < \epsilon$.*

**Proof.** By the above mentioned property, we may replace $f$ by $|g|^2$ with $g \in H_2$ and outer. Thus we have to prove that there exists an $f_n$ such that $\||g|^2 - |f_n|^2\|_1 < \epsilon$. By a property which can be found in the book by Rudin [73, p. 78], we may use for $p = 2$

$$\||g|^p - |h|^p\|_1 \leq 2pR^{p-1}\|g - h\|_p, \text{ with } \max\{\|g\|_p, \|h\|_p\} \leq R.$$

Thus

$$\||g|^2 - |f_n|^2\|_1 \leq 4R \left[\int |g - f_n|^2 d\lambda\right]^{1/2}.$$

We can always find an interpolant $f_n$ which makes $|g - f_n|$ arbitrary small and hence also $|f_n| < |g| + \epsilon_1$ with $\epsilon_1 > 0$ arbitrary small. Thus because $g \in H_2$, also $f_n$ will be in $H_2$, so that $R$ is bounded, while on the other hand $\|g - f_n\|_2$ can be made as small as we want. This proves the Theorem. $\square$

# 10 Christoffel-Darboux relations

We now prove the Christoffel-Darboux relations. We start with some technical lemmas.

**Lemma 10.1** *Let $f \in \mathcal{L}_n$.*

1. *If $g$ and $h$ are defined by the relations $f(z) - f(w) = (z-w)g(z) = \frac{z-w}{1-\overline{\alpha}_n z}h(z)$ then*
   (a) *$p_1(z)g(z) \in \mathcal{L}_n$ for all $p_1 \in \Pi_1$, an arbitrary polynomial of degree at most 1. Especially $g(z) \in \mathcal{L}_n$.*
   (b) *$h \in \mathcal{L}_{n-1}$.*
2. *If $f(w) = 0$, then $\frac{1-\overline{\alpha}_n z}{z-w} f(z) \in \mathcal{L}_{n-1}$.*

**Proof.** Clearly $g(z)$ can be written as $p_{n-1}(z)/\pi_n(z)$ with $\pi_n(z) = \prod_{k=1}^n (1 - \overline{\alpha}_k z)$ and $p_{n-1}(z) \in \Pi_{n-1}$. This implies (a). Furthermore, $h(z) = (1 - \overline{\alpha}_n z)g(z)$, which gives $h(z) = p_{n-1}(z)/\pi_{n-1} \in \mathcal{L}_{n-1}$ and this is (b).

The second result is a special case of (1b) for $f(w) = 0$. □

**Lemma 10.2** *Let $\phi_k$ denote as before the orthonormal basis functions for $\mathcal{L}_n$ and $\zeta_k$ the Blaschke factor based on $\alpha_k$. As functions of $z$, with $w$ some parameter, we have*

$$l_n^0(z,w) = \frac{\phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(z)\overline{\phi_{n+1}(w)}}{1 - \zeta_{n+1}(z)\overline{\zeta_{n+1}(w)}} \in \mathcal{L}_n$$

*and for $k = 1, \ldots, n$*

$$l_n^k(z,w) = \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - \zeta_k(z)\overline{\zeta_k(w)}\phi_n(z)\overline{\phi_n(w)}}{1 - \zeta_k(z)\overline{\zeta_k(w)}} \in \mathcal{L}_n.$$

**Proof.** An easy computation gives

$$1 - \zeta_k(z)\overline{\zeta_k(w)} = 1 - \frac{(\alpha_k - z)(\overline{\alpha}_k - \overline{w})}{(1-\overline{\alpha}_k z)(1 - \alpha_k \overline{w})} = \frac{(1-|\alpha_k|^2)(1 - z\overline{w})}{(1-\overline{\alpha}_k z)(1-\alpha_k\overline{w})}.$$

According to part (2) of the previous lemma, we only have to prove that the numerator of $l_n^0$ is zero for $z = 1/\overline{w}$. Call this numerator $N(z,w)$. Thus we have to prove that $N(1/\overline{w}, w) = 0$. Now,

$$\begin{aligned}
N(1/\overline{w}, w) &= \phi_{n+1}^*(1/\overline{w})\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(1/\overline{w})\overline{\phi_{n+1}(w)} \\
&= B_{n+1}(1/\overline{w})\overline{\phi_{n+1}(w)}\overline{B_{n+1}(w)}\phi_{n+1}(1/\overline{w}) - \phi_{n+1}(1/\overline{w})\overline{\phi_{n+1}(w)} \\
&= 0
\end{aligned}$$

which proves the first part.

For the second part, we have to prove, according to part (1a) of the previous lemma, as in part one of this theorem that the numerator of $l_n^k(z,w)$ is zero for $z = 1/\overline{w}$. Let us call this numerator again $N(z,w)$. Then

$$\begin{aligned}
N(1/\overline{w}, w) &= \phi_n^*(1/\overline{w})\overline{\phi_n^*(w)} - \phi_n(1/\overline{w})\overline{\phi_n(w)} \\
&= B_n(1/\overline{w})\overline{\phi_n(w)}\overline{B_n(w)}\phi_n(1/\overline{w}) - \phi_n(1/\overline{w})\overline{\phi_n(w)} \\
&= 0
\end{aligned}$$

and this proves the second part. □

Now we can prove the Christoffel-Darboux relations.

**Theorem 10.3 (Christoffel-Darboux relations)** *The following relations hold between reproducing kernel and orthonormal basis functions of $\mathcal{L}_n$.*

$$k_n(z,w) = \frac{\phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(z)\overline{\phi_{n+1}(w)}}{1-\zeta_{n+1}(z)\overline{\zeta_{n+1}(w)}} \quad (= l_n^0(z,w)) \tag{10.1}$$

*and*

$$k_n(z,w) = \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - \zeta_n(z)\overline{\zeta_n(w)}\phi_n(z)\overline{\phi_n(w)}}{1-\zeta_n(z)\overline{\zeta_n(w)}} \quad (= l_n^n(z,w)). \tag{10.2}$$

**Proof.** The proof we give is completely analogous to the proof that Szegő gave for these relations in the polynomial case. See [78, p.293]. We have shown in the previous lemma that the right hand sides are elements from $\mathcal{L}_n$. We only have to show that they reproduce. So choose some $f \in \mathcal{L}_n$. Then

$$\langle f, l_n^0(\cdot,w)\rangle_\mu = f(w)\langle 1, l_n^0(\cdot,w)\rangle_\mu + \langle f - f(w), l_n^0(\cdot,w)\rangle_\mu. \tag{10.3}$$

By Lemma 10.1a, we find that if $f(z) - f(w) = (z-w)g(z)$, then $g \in \mathcal{L}_{n-1}$ and $p_1 g \in \mathcal{L}_n$ for all $p_1 \in \Pi_1$. Thus if we call $N(z,w)$ the numerator of $l_n^0$, we get

$$\begin{aligned} \langle f - f(w), l_n^0\rangle_\mu &= \frac{1-\overline{\alpha}_{n+1}w}{1-|\alpha_{n+1}|^2}\langle (z-w)g(z), \frac{1-\overline{\alpha}_{n+1}z}{1-z\overline{w}}N(z,w)\rangle_\mu \\ &= \frac{1-\overline{\alpha}_{n+1}w}{1-|\alpha_{n+1}|^2}\langle (z-\alpha_{n+1})g(z), N(z,w)\rangle_\mu. \end{aligned}$$

Because $(z-\alpha_{n+1})g(z) \in \mathcal{L}_n$, the inner product in the right hand side gives

$$\begin{aligned} \langle (z-\alpha_{n+1})g(z), N(z,w)\rangle_\mu &= \langle (z-\alpha_{n+1})g(z), \phi_{n+1}^*(z)\rangle_\mu \phi_{n+1}^*(w) \\ &= \langle \zeta_{n+1}(z)h(z), \phi_{n+1}^*(z)\rangle_\mu \phi_{n+1}^*(w) \end{aligned}$$

with $h(z) = -|\alpha_{n+1}|/\overline{\alpha}_{n+1}\ (1-\overline{\alpha}_{n+1}z)g(z) \in \mathcal{L}_n$. This is zero because of Theorem 7.1(4). It remains to be shown that $\langle 1, l_n^0(\cdot,w)\rangle_\mu = \eta(w) = 1$. Apply (10.3) to $f(\cdot) = l_n^0(\cdot,z)$. Then we get

$$\langle l_n^0(\cdot,z), l_n^0(\cdot,w)\rangle_\mu = l_n^0(w,z)\eta(w).$$

If we interchange $z$ and $w$, we get

$$\langle l_n^0(\cdot,w), l_n^0(\cdot,z)\rangle_\mu = l_n^0(z,w)\eta(z).$$

Since the left hand sides are each others conjugate and because also $l_n^0(z,w)$ is the complex conjugate of $l_n^0(w,z)$, we find $\eta(z) = \overline{\eta(w)}$. Thus $\eta$ is a constant and this constant is 1 because for $z = w = \alpha_{n+1}$, we get

$$\begin{aligned} k_n(\alpha_{n+1},\alpha_{n+1}) &= \eta l_n^0(\alpha_{n+1},\alpha_{n+1}) = \eta(\kappa_{n+1}^2 - |\phi_{n+1}(\alpha_{n+1})|^2) \\ &= k_{n+1}(\alpha_{n+1},\alpha_{n+1}) - |\phi_{n+1}(\alpha_{n+1})|^2 \\ &= \kappa_{n+1}^2 - |\phi_{n+1}(\alpha_{n+1})|^2. \end{aligned}$$

So that $\eta = 1$.

Analogously, we get when $N(z,w)$ is now the numerator of $l_n^n$

$$\langle f(z) - f(w), l_n^n(z,w)\rangle_\mu = c\frac{1-\overline{\alpha}_n w}{1-|\alpha_n|^2}\langle \zeta_n(z)h(z), N(z,w)\rangle_\mu$$

with $h \in \mathcal{L}_{n-1}$ by Lemma 10.1. The inner product of the right hand side is zero because $\phi_n^* \perp \zeta_n h \in \zeta_n\mathcal{L}_{n-1}$ and $\zeta_n\phi_n \perp \zeta_n h$ because $h \in \mathcal{L}_{n-1}$. The rest follows exactly as in the proof of the previous formula. □

From these relations, we find a useful corollary.

**Corollary 10.4** *For the orthonormal functions of $\mathcal{L}_n$, it holds that for all $n \geq 0$*

1. $\phi_n^*(z) \neq 0$ *for* $z \in \mathbf{D}$ *and* $\phi_n(z) \neq 0$ *for* $z \in \mathbf{E}$.
2. $|\phi_{n+1}(z)/\phi_{n+1}^*(z)| < 1$ *for* $z \in \mathbf{D}$ *and* $|\phi_{n+1}(z)/\phi_{n+1}^*(z)| > 1$ *for* $z \in \mathbf{E}$.

**Proof.** From the first Christoffel-Darboux relation (10.1), we get for $w = z$

$$(1 - |\zeta_{n+1}(z)|^2)k_n(z,z) = |\phi_{n+1}^*(z)|^2 - |\phi_{n+1}(z)|^2.$$

Because $|\zeta_{n+1}(z)| < 1$ for $z \in \mathbf{D}$, and $k_n(z,z) > 0$, we get

$$|\phi_{n+1}^*(z)|^2 > |\phi_{n+1}(z)|^2 \geq 0, \text{ for } z \in \mathbf{D}.$$

Hence $\phi_{n+1}^*(z) \neq 0$ for $z \in \mathbf{D}$ and thus

$$|\phi_{n+1}(z)/\phi_{n+1}^*(z)| < 1 \text{ for } z \in \mathbf{D}.$$

The proof for $\mathbf{E}$ is analogous. □

## 11 Recurrence relations for the kernels

The reproducing kernels for the spaces $\mathcal{L}_n$ satisfy some recursions which can be found from the Christoffel-Darboux relations. We shall derive them below.

**Theorem 11.1** *Let $k_n(z,w)$ be the reproducing kernel for $\mathcal{L}_n$. Then*

$$\begin{bmatrix} k^*_{n+1}(z,w) \\ k_{n+1}(z,w) \end{bmatrix} = t_{n+1}(z,w) \begin{bmatrix} k^*_n(z,w) \\ k_n(z,w) \end{bmatrix} \tag{11.1}$$

*where the superstar conjugation is with respect to $z$. The matrix $t_{n+1}$ is given by*

$$t_{n+1}(z,w) = c \begin{bmatrix} 1 & \overline{\rho}_{n+1} \\ \rho_{n+1} & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n+1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \overline{\gamma}_{n+1} \\ \gamma_{n+1} & 1 \end{bmatrix}$$

*with*

$$\begin{aligned} c &= (1-|\rho_{n+1}|^2)^{-1} \\ \rho_{n+1} &= \rho_{n+1}(w) = \overline{\phi_{n+1}(w)}/\phi^*_{n+1}(w) \\ \gamma_{n+1} &= \gamma_{n+1}(w) = -\zeta_{n+1}(w)\rho_{n+1}(w) \end{aligned}$$

*and $\phi_{n+1}$ is the $(n+1)$-st orthonormal basis function.*

**Proof.** Obviously,

$$k_{n+1}(z,w) = k_n(z,w) + \phi_{n+1}(z)\overline{\phi_{n+1}(w)}. \tag{11.2}$$

From the Christoffel-Darboux relation (10.1), we get

$$\phi_{n+1}(z)\overline{\phi_{n+1}(w)} = \phi^*_{n+1}(z)\overline{\phi^*_{n+1}(w)} - (1-\zeta_{n+1}(z)\overline{\zeta_{n+1}(w)})k_n(z,w). \tag{11.3}$$

Substitute this into (11.2) to get

$$k_{n+1}(z,w) = \zeta_{n+1}(z)\overline{\zeta_{n+1}(w)}k_n(z,w) + \phi^*_{n+1}(z)\overline{\phi^*_{n+1}(w)}. \tag{11.4}$$

Now, take the superstar conjugate with respect to $z$ of (11.2).

$$\phi^*_{n+1}(z)\overline{\phi^*_{n+1}(w)}\ \overline{\rho_{n+1}(w)} = k^*_{n+1}(z,w) - \zeta_{n+1}(z)k^*_n(z,w). \tag{11.5}$$

Substitute this into (11.4) to get

$$k_{n+1}(z,w)\overline{\rho}_{n+1} = \zeta_{n+1}(z)\overline{\zeta_{n+1}(w)}k_n(z,w)\overline{\rho}_{n+1} + k^*_{n+1}(z,w) - \zeta_{n+1}(z)k^*_n(z,w). \tag{11.6}$$

The superstar conjugate of (11.6) is

$$k^*_{n+1}(z,w)\rho_{n+1} = \zeta_{n+1}(w)k^*_n(z,w)\rho_{n+1} + k_{n+1}(z,w) - k_n(z,w). \tag{11.7}$$

From (11.6) and (11.7), the result follows. □

Note that by Corollary 10.4, $|\rho_n(w)| < 1$ for $w \in \mathbf{D}$ and also $\gamma_n(w) = -\zeta_n(w)\rho_n(w)$ is in $\mathbf{D}$ for all $w \in \mathbf{D}$. The following Corollary generalizes a well known result from the polynomial case. See e.g. [42, p.40].

**Corollary 11.2** *The following statements hold.*

1. *Because of the normalization of the measure $\mu$, which was $\int d\mu = 1$, it follows that $\phi_0 = \kappa_0 = 1$, so that also $k_0(z,w) = 1$ and consequently $s_0(z) = k^*_0(z,w)/k_0(z,w) = 1$. For $n \geq 1$, define $s_n(z,w) = k^*_n(z,w)/k_n(z,w)$. Then it holds that $s_n(z,w) \in \mathbf{D}$ for all $z,w \in \mathbf{D}$. In other words, for $n \geq 1$ and $w \in \mathbf{D}$ fixed, $s_n(z,w) \in \mathcal{B}$.*

2. *For some fixed $w \in \mathbf{D}(\mathbf{E},\mathbf{T})$, the reproducing kernel $k_n(z,w)$ has its zeros in $\mathbf{E}(\mathbf{D},\mathbf{T})$.*

3. *The zeros of the orthonormal basis functions $\phi_n$ are all in $\mathbf{D}$.*

**Proof.**

1. From the recurrence relation given in the previous theorem, we can deduce that the function $s_{n+1}(z,w)$ is obtained from $s_n(z,w)$ by a transformation which we denote as $\tau$, i.e.,

   $$s_{n+1}(z,w) = \tau(s_n(z,w))$$

   where $\tau$ is a succession of three transformations $\tau_1 \circ \tau_2 \circ \tau_3$ with

   $$\begin{aligned} \tau_3 &: \ t \longmapsto \frac{t+\overline{\gamma}_{n+1}}{1+\gamma_{n+1}t} \\ \tau_2 &: \ t \longmapsto \zeta_{n+1}(z)t \\ \tau_1 &: \ t \longmapsto \frac{t+\overline{\rho}_{n+1}}{1+\rho_{n+1}t}. \end{aligned}$$

   In other words, in the notation of section 4, $\tau_3 = \zeta_{-\overline{\gamma}_{n+1}}$ and $\tau_1 = \zeta_{-\overline{\rho}_{n+1}}$. Because $|\gamma_{n+1}| < 1$ and $|\rho_{n+1}| < 1$ for $w \in \mathbf{D}$, it follows that all these transformations are contractions of $\mathbf{D}$, thus all the $s_n$ are Schur functions as it follows from Theorem 4.3.

2. For this part a transcription of the proof given in Szegő [78, p.292] can be made. Szegő proved this result in the polynomial case for an absolutely continuous measure $d\mu = \mu' d\lambda$. This absolute continuity however is not essential and his proof can be easily translated to a direct proof for the rational case. If we start from the fact that the theorem is true for the polynomial case, then we can make the following observation. If

   $$k_n(z,w) = p_n(z,w)/\pi_n(z), \quad p_n(z,w) \in \Pi_n$$

   where $\pi_n(z) = \prod_{k=1}^n (1-\overline{\alpha}_k z)$, then one easily sees that $p_n(z,w)$ is a (polynomial) reproducing kernel for $(\Pi_n, \mu_n)$ with $d\mu_n = |\pi_n(z)|^{-2} d\mu$. Hence, the zeros of $k_n(z,w)$, which are the zeros of $p_n(z,w)$ are exactly as was stated.

   Another, direct, proof can be given, using the J-unitary recursions as given in the next section.

3. If we set $w = \alpha_n$ in the previous result and use $k_n(z,\alpha_n) = \kappa_n \phi_n^*(z)$, we find that $\phi_n^*$ has all its zeros (and poles) in $\mathbf{E}$. Consequently, $\phi_n$ has all its zeros in $\mathbf{D}$ (and poles is $\mathbf{E}$). Note that this sharpens the result of Corollary 10.4(1).

$\square$

The recursion for the kernels can easily be inverted to give $k_{n-1}$ and $k_{n-1}^*$ in terms of $k_n$ and $k_n^*$. Moreover, the recursion coefficients $\rho_n$ and $\gamma_n$ are completely defined in terms of $k_n$, since indeed $\rho_n = \overline{k_n(w,\alpha_n)^*/k_n(w,\alpha_n)}$, so that the kernel $k_{n-1}$ is uniquely defined for given $k_n$. By induction, all the previous kernels $k_j$ will also be fixed by $k_n$. Thus to check if in $\mathcal{L}_n$ all the kernels with respect to two different measures $\mu$ and $\nu$ are the same, it is sufficient to check that the $k_n$ are the same.

# 12 Normalized recursions for the kernels

The recursions for the reproducing kernels involved matrices that were almost J-unitary matrices. Since J-unitary matrices had a lot of interesting properties, we want to normalize these recursions. It turns out that the recursion is given by a J-unitary matrix if we consider normalized kernels, which we shall denote with a capital :

$$K_n(z,w) = k_n(z,w)[k_n(w,w)]^{-1/2}. \tag{12.1}$$

This can be easily inverted to give $k_n(z,w) = K_n(w,w)K_n(z,w)$. Note that $K_n(z,\alpha_n) = \phi_n^*(z)$, the $n$-th orthonormal basis function. The next theorem gives the normalized recursion.

**Theorem 12.1** *The normalized kernels $K_n(z,w)$ defined in (12.1) satisfy the recursion*

$$\begin{bmatrix} K_n^*(z,w) \\ K_n(z,w) \end{bmatrix} = \theta_n(z,w) \begin{bmatrix} K_{n-1}^*(z,w) \\ K_{n-1}(z,w) \end{bmatrix} \tag{12.2}$$

*where the superstar conjugation is with respect to $z$. The matrix $\theta_n$ is given by*

$$\theta_n(z,w) = c \begin{bmatrix} 1 & \overline{\rho}_n \\ \rho_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix} d \begin{bmatrix} 1 & \overline{\gamma}_n \\ \gamma_n & 1 \end{bmatrix}$$

*with*

$$\begin{aligned} c &= (1-|\rho_n|^2)^{-1/2} \text{ and } d = (1-|\gamma_n|^2)^{-1/2} \\ \rho_n &= \rho_n(w) = \overline{\phi_n(w)}/\phi_n^*(w) \\ \gamma_n &= \gamma_n(w) = -\zeta_n(w)\rho_n(w) \end{aligned}$$

*and $\phi_n$ is the $n$-th orthonormal basis function for $\mathcal{L}_n$.*

**Proof.** Note that in the notation of section 5, the matrix $\theta_n$ can also be written as $\theta_n = U_{-\overline{\rho}_n} B_{\alpha_n} U_{-\overline{\gamma}_n}$ with the exception of the factor $\overline{\alpha}_n/|\alpha_n|$ in $\zeta_n$.

From the normalization, it follows that the $\theta_n$ matrix of this theorem and the $t_n$ matrix of Theorem 11.1 are related by

$$t_n(z,w) = \left[\frac{k_n(w,w)}{k_{n-1}(w,w)}\right]^{1/2} \theta_n(z,w).$$

From Theorem 11.1 with $z = w$ and $\gamma_n = -\zeta_n(w)\rho_n$, we get

$$k_n(w,w) = (1-|\zeta_n(w)\rho_n|^2)(1-|\rho_n|^2)^{-1} k_{n-1}(w,w). \tag{12.3}$$

This gives the normalized recursion. □

**Corollary 12.2** *If $\rho_k(w) = \overline{\phi_k(w)}/\phi_k^*(w)$ and $\gamma_k(w) = -\zeta_k(w)\rho_k(w)$ as in the previous Theorem, we have the following expression for $k_n(w,w)$.*

$$k_n(w,w) = \prod_{k=1}^{n} \frac{(1-|\gamma_k(w)|^2)}{(1-|\rho_k(w)|^2)}. \tag{12.4}$$

**Proof.** This is a direct consequence of the formula (12.3) which is repeatedly applied and the fact that $k_0(w,w) = 1$. □

Now define

$$\Theta_n = \theta_n\theta_{n-1}\cdots\theta_1, \quad n \geq 1. \tag{12.5}$$

It is not difficult to show that each $\Theta_k = [\Theta_k^{ij}]$, has the property that

$$\Theta_k^{11} = (\Theta_k^{22})^* \text{ and } \Theta_k^{12} = (\Theta_k^{21})^* \quad \text{(superstar with respect to } z\text{)}. \tag{12.6}$$

Notice that $k_0 = K_0 = 1$. Let us define

$$\theta_0 = \begin{bmatrix} K_0^* & L_0^* \\ K_0 & -L_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{12.7}$$

with $L_0 = K_0 = 1$. Then we immediately get

$$\Theta_n \theta_0 = \begin{bmatrix} K_n^*(z,w) & L_n^*(z,w) \\ K_n(z,w) & -L_n(z,w) \end{bmatrix}. \tag{12.8}$$

The first column is obtained as a consequence of the recurrence for the normalized kernels. The elements in the second column satisfy the same recurrence but with different initial conditions. Notice that $L_n(z,w) \in \mathcal{L}_n$ for every $w$. That you get $L_n^*(z,w)$ on top of $-L_n(z,w)$ is a consequence of ((12.6). The elements of $\Theta_n$ can be expressed in terms of these $K_n$ and $L_n$ by multiplying with $\theta_0^{-1}$.

$$\Theta_n = \frac{1}{2} \begin{bmatrix} K_n^* + L_n^* & K_n^* - L_n^* \\ K_n - L_n & K_n + L_n \end{bmatrix}. \tag{12.9}$$

From property (6) of Theorem 5.1, we now get the next Theorem.

**Theorem 12.3** *Let $w \in \mathbf{D}$ be a given number and let $K_n$ and $L_n$ be as defined above by the normalized recurrence matrix. Then*

1. $\frac{1}{2}\left[\frac{L_n^*}{K_n^*} + \frac{L_n}{K_n}\right] = \frac{B_n}{K_n^* K_n} = \frac{1}{K_{n*} K_n} = \frac{1}{2}\left[\frac{L_{n*}}{K_{n*}} + \frac{L_n}{K_n}\right]$
2. $L_n / K_n \in \mathcal{P}$
3. $1/K_n \in H_2$.

**Proof.** Take $a = \theta_{11} - \theta_{12} = L_n^*$, $b = \theta_{11} + \theta_{12} = K_n^*$, and use $(f^*)_* = B_n^{-1} f_*$ from Theorem 7.1. Theorems 5.1 and 5.3 then lead to the result. □

The following is some useful observation.

**Theorem 12.4** *If $K_n^*$ and $K_n$ are generated from (12.2) with some coefficients $\rho_n$ and $\gamma_n$, then $L_n^*$ and $L_n$ as defined in (12.8) are generated by exactly the same relation (12.2) where you have to replace $\rho_n$ by $-\rho_n$ and $\gamma_n$ by $-\gamma_n$.*

**Proof.** We can remove the minus sign in the defining relation (12.8) by writing it as

$$\begin{bmatrix} L_n^* \\ L_n \end{bmatrix} = J\theta_n J \begin{bmatrix} L_{n-1}^* \\ L_{n-1} \end{bmatrix}$$

with

$$J\theta_n J = c \begin{bmatrix} 1 & -\overline{\rho}_n \\ -\rho_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n & 0 \\ 0 & 1 \end{bmatrix} d \begin{bmatrix} 1 & -\overline{\gamma}_n \\ -\gamma_n & 1 \end{bmatrix}$$

where the coefficients have the same meaning as in Theorem 12.1. □

# 13 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. These turn out to be a bit more complicated. In view of the J-unitary recursions which we saw in the previous section for the reproducing kernels, the recursions for the orthonormal basis seems not to be so nice.

**Theorem 13.1** *For the orthonormal basis functions in $\mathcal{L}_n$, a recursion of the following form exists*

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = N_n \frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z} \begin{bmatrix} 1 & \overline{\lambda}_n \\ \lambda_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} \tag{13.1}$$

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

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

*The parameter $\lambda_n$ is given by*

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

*and $\rho_n(w)$ as defined in Theorem 11.1.*

*The formula (13.1)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

$$\frac{1-\overline{\alpha}_n z}{z-\alpha_{n-1}} \phi_n - d_n \phi_{n-1} - c_n \frac{1-\overline{\alpha}_{n-1}z}{z-\alpha_{n-1}} \phi_{n-1}^* \in \mathcal{L}_{n-2} \tag{13.2}$$

Let us define as before $\pi_k(z) = \prod_{i=1}^k (1-\overline{\alpha}_i z)$ and the polynomials $p_k$ are defined by $\phi_k = p_k/\pi_k$. Note that we can use Theorem 7.1 to rewrite this as

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

where we have used $\eta_k \in \mathbf{T}$ as defined in Theorem 7.1. If this has to be in $\mathcal{L}_{n-2}$, then we should require that $N(\alpha_{n-1}) = N(1/\overline{\alpha}_{n-1}) = 0$ or, which is the same, $N(\alpha_{n-1}) = N^*(\alpha_{n-1}) = 0$. The first condition gives

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

The second condition defines $d_n$.

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

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

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

and

$$\overline{d}_n = -\frac{1-\overline{\alpha}_n\alpha_{n-1}}{1-|\alpha_{n-1}|^2}\frac{\alpha_n}{|\alpha_n|}\frac{\phi_n^*(\alpha_{n-1})}{\kappa_{n-1}}. \tag{13.3}$$

Thus we have proved that with the previous choices of $c_n$ and $d_n$, the expression in (13.2) 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 $k \leq n-2$, $\phi_k$ is orthogonal to the first term in (13.2) because

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

and this is zero because the right factor is in $\mathcal{L}_{n-1}$. $\phi_k$ is trivially orthogonal to the second term in (13.2). Finally, it is also orthogonal to the third term since

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

and this is zero by Theorem 7.1. We may thus conclude that the expression in (13.2) is zero. Hence

$$\begin{aligned} \phi_n &= d_n\frac{z-\alpha_{n-1}}{1-\overline{\alpha}_n z}\phi_{n-1} + c_n\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z}\phi_{n-1}^* \\ &= d_n\frac{-\alpha_{n-1}}{|\alpha_{n-1}|}\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z}[\zeta_{n-1}(z)\phi_{n-1} + \overline{\lambda}_n\phi_{n-1}^*] \end{aligned} \tag{13.4}$$

with

$$\lambda_n = -\frac{\overline{c}_n}{\overline{d}_n}\frac{\alpha_{n-1}}{|\alpha_{n-1}|} \tag{13.5}$$

Note that we can write $\lambda_n$ as

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

and $\rho_n(w)$ as defined in Theorem 11.1. We then know from Corollary 10.4 that $\rho_n \in \mathbf{D}$ and thus also $\lambda_n \in \mathbf{D}$. Recall that $\overline{\alpha}_k/|\alpha_k| = -1$ if $\alpha_k = 0$. Taking the superstar conjugate, we can find the recurrence as claimed. One can choose e.g., $e_n = |d_n| \in \mathbf{R}$. The values of $\eta_n^1$ and $\eta_n^2$ can readily be computed to be

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

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

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

where $e_1 \in \mathbf{R}$ and $\eta_1^1 \in \mathbf{T}$. Hence the constants $\eta_1^1$ and $\lambda_1$ should satisfy

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

These can be solved for $\eta_1^1$ and $\lambda_1$, 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 (13.1). The other recurrence is found by taking the superstar conjugate of the first one. They are equivalent to each other. □

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

**Theorem 13.2** *The recursion coefficient $\lambda_n$ from the previous Theorem can also be expressed as*

$$\lambda_n = \frac{\alpha_{n-1}}{|\alpha_{n-1}|} \frac{\langle \phi_{n-1}, \frac{z-\alpha_{n-1}}{1-\overline{\alpha}_n z} \phi_{n-1} \rangle_\mu}{\langle \phi_{n-1}, \frac{1-\overline{\alpha}_{n-1} z}{1-\overline{\alpha}_n z} \phi^*_{n-1} \rangle_\mu},$$

*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 (13.4) for $\phi_n$ and express that it is orthogonal to $\phi_{n-1}$. Then you get.

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

Use then the defining relation of (13.5) and the expression for the ratio of $\overline{c}_n/\overline{d}_n$ that you can get from the previous relation. Then you will find the expression for $\lambda_n$.

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

$$e_n^2(1-|\lambda_n|^2) = \frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2}. \tag{13.6}$$

Fill in $e_n^2 = |d_n|^2$ with $d_n$ given by (13.3) and the expression for $\lambda_n$ to find

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

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 (13.6). □

The presence of the $\eta_n^1$ and $\eta_n^2$ 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 $\eta$'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 (13.1). Therefore we define

$$\epsilon_0 = 1 \text{ and } \epsilon_n = -\epsilon_{n-1} \frac{\overline{d}_n}{|d_n|} \frac{\overline{\alpha}_n}{|\alpha_n|} \text{ for } n \geq 1. \tag{13.7}$$

where $d_n$ is as in (13.1) and use this as a rotation for $\phi_n$. 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 13.3** *Let $\phi_n$ be the orthonormal functions satisfying the recurrence relation of Theorem 13.1 and denote by $\Phi_n$ the rotated orthonormal functions $\Phi_n = \epsilon_n \phi_n$ as introduced above. Then these satisfy the recurrence relation*

$$\begin{bmatrix} \Phi_n(z) \\ \Phi^*_n(z) \end{bmatrix} = e_n \frac{1-\overline{\alpha}_{n-1} z}{1-\overline{\alpha}_n z} \begin{bmatrix} 1 & \overline{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix} \begin{bmatrix} Z_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Phi_{n-1}(z) \\ \Phi^*_{n-1}(z) \end{bmatrix} \tag{13.8}$$

*where*

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

*and*

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

**Proof.** You can start with the recurrence (13.1) and express the $\phi_n$ in terms of the $\Phi_n$, which results in the relation

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

with the matrix $M_n$ defined by

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

Use in this matrix the definitions of $\epsilon_n$, of $\eta_n^i$ and $\Lambda_n$ and some algebra will give the result. □

In the next sections, we still go on developing the results for the $\phi_n$ in the first place, but virtually the same results hold true for the rotated functions $\Phi_n$. Occasionally we shall state the result for $\Phi_n$ in a remark. The rotated functions are however important for the interpolation algorithm to be given later in Section 19.

It is also possible to get relations between successive orthogonal functions from the recursions for the (normalized) kernels $K_n(z,w)$ in the previous section. We give the result in a slightly more general form

**Theorem 13.4** *Let the J-unitary contractive matrices $\Theta_n(z,w)$ be as defined in (12.5). Then*

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = \Theta_n(z,\alpha_n)\Theta_{n-k}^{-1}(z,\alpha_{n-k}) \begin{bmatrix} \phi_{n-k}(z) \\ \phi_{n-k}^*(z) \end{bmatrix}. \tag{13.9}$$

**Proof.** We recall that $K_n(z,\alpha_n) = \phi_n^*(z)$. Hence it follows that

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = \Theta_n(z,\alpha_n) \begin{bmatrix} \phi_0(z) \\ \phi_0^*(z) \end{bmatrix}.$$

Because $\Theta_n$ is J-unitary and therefore invertible (see Theorem 5.1) the result easily follows. □

Recall the definition of the $L_n(z,w)$ from (12.9) and suppose we set by definition $L_n(z,\alpha_n) = \chi_n^*(z) \in \mathcal{L}_n$. As a special case of (12.9) we thus get

$$\Theta_n(z,\alpha_n) = \begin{bmatrix} \phi_n + \chi_n & \phi_n - \chi_n \\ \phi_n^* - \chi_n^* & \phi_n^* + \chi_n^* \end{bmatrix}. \tag{13.10}$$

We can now formulate a special case of Theorem 12.3.

**Theorem 13.5** *Let $\phi_n$ and $\chi_n$ be as defined above. Then*

1. $\frac{1}{2}\left[\frac{\chi_n}{\phi_n} + \frac{\chi_n^*}{\phi_n^*}\right] = \frac{B_n}{\phi_n\phi_n^*} = \frac{1}{\phi_n\phi_{n*}}$
2. $\chi_n^*/\phi_n^* = \chi_{n*}/\phi_{n*} \in \mathcal{P}$
3. $1/\phi_n^*$ *and hence also* $1/\phi_{n*} \in H_2$
4. $\phi_n/\phi_n^* \in \mathcal{B}$.

**Proof.** Use Theorems 12.3, 11.2 and some properties from section 7. □

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

**Theorem 13.6** *Given the orthonormal function $\phi_n$ with $\phi_n^*(\alpha_n) = \kappa_n > 0$, all the previous orthonormal functions $\phi_k, k < n$ are uniquely defined if they are similarly normalized by $\phi_k^*(\alpha_k) = \kappa_k > 0$. They can be found with the recursions*

$$\begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} = \frac{1}{1-|\lambda_n|^2} \frac{1-\overline{\alpha}_n z}{1-\overline{\alpha}_{n-1} z} \begin{bmatrix} 1/\zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\overline{\lambda}_n \\ -\lambda_n & 1 \end{bmatrix} N_n^{-1} \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} \quad (13.11)$$

*with all the quantities appearing in this formula as in Theorem 13.1.*

**Proof.** The formula (13.11) is evidently the inverse of the recurrence formula (13.1). Since the coefficients $\lambda_n$ and the matrix $N_n$ is completely defined in terms of $\phi_n$, the $\phi_{n-1}$ is uniquely defined. By induction, all the previous $\phi_k$ are uniquely defined. □

In fact this is a simple consequence of the note given at the end of section 11. The kernels are uniquely defined in terms of the last one. The orthonormal functions will also be unique if they have the normalization mentioned.

## 14 Functions of the second kind

In this section we shall define some functions $\psi_k$ 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 $\phi_n$. 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 $\Omega_\mu$. Here are some equivalent definitions for $z \in \mathbf{D}$

$$\begin{aligned}\psi_n(z) &= \int [\frac{2t}{t-z}\phi_n(t) - \frac{t+z}{t-z}\phi_n(z)]d\mu(t) \qquad (14.1)\\ &= \int D(t,z)[\phi_n(t)-\phi_n(z)]d\mu(t) + \int \phi_n(t)d\mu(t)\\ &= \begin{cases} 1 & , \text{ if } n=0 \\ \int D(t,z)[\phi_n(t)-\phi_n(z)]d\mu(t) & , \text{ if } n \geq 1. \end{cases} \qquad (14.2)\end{aligned}$$

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.

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

**Lemma 14.1** *The functions $\psi_n$ 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 (14.2) has the form

$$[\phi_n(t)-\phi_n(z)](t+z)/(t-z).$$

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

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

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 14.2** *To define the functions of the second kind for $n > 0$, we may replace (14.2) by*

$$\frac{\psi_n(z)}{B_k(z)} = \int D(t,z)[\frac{\phi_n(t)}{B_k(t)} - \frac{\phi_n(z)}{B_k(z)}]d\mu(t) = \int [\frac{2t}{t-z}\frac{\phi_n(t)}{B_k(t)} - \frac{t+z}{t-z}\frac{\phi_n(z)}{B_k(z)}]d\mu(t) \qquad (14.3)$$

*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

$$\int \frac{t+z}{t-z}[1-\frac{B_k(z)}{B_k(t)}]\phi_n(t)d\mu(t) = 0 \text{ or } \int \frac{t}{t-z}[1-\frac{B_k(z)}{B_k(t)}]\phi_n(t)d\mu(t) = 0$$

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 $\psi_n^*$.

**Lemma 14.3** *The superstar conjugate of the functions of the second kind are given by*

$$\psi_n^*(z) = \int [\frac{t+z}{t-z}\phi_n^*(z) - \frac{2z}{t-z}\frac{B_{n\setminus k}(z)}{B_{n\setminus k}(t)}\phi_n^*(t)]d\mu(t) \tag{14.4}$$

*for any* $0 \leq k < n$. *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 previous expression implies that $\psi_0^* = 1$, since we get

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

The last term equals $\langle \phi_n^*, B_{n\setminus k}\rangle_\mu B_{n\setminus k}(z) = \delta_{n0}$ because of the orthogonality properties of $\phi_n^*$. Since the first term vanishes for $n = 0$, the Lemma is true for $n = 0$. So, suppose that $n > 0$. The relation (14.4) then follows immediately from (14.3) by taking the superstar conjugate. This proves the Lemma. □

Note that like in (14.3), we can give an equivalent form of (14.4) as follows.

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

where, as we know, the last term is $\delta_{0n}$. 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 14.4** *Let* $\Omega = \Omega_\mu$ *be the positive real function with Riesz-Herglotz measure* $\mu$. *Then for the functions of the second kind, it holds that*

$$\frac{\phi_n\Omega + \psi_n}{B_{n-1}} = \begin{cases} \Omega + 1 \in H(\mathbf{D}) & , \quad n = 0 \\ g \in H(\mathbf{D}) \text{ and } g(0) = 0 & , \quad n > 0. \end{cases} \tag{14.5}$$

*For their superstar conjugates, we find*

$$\frac{\phi_n^*\Omega - \psi_n^*}{B_n} = \begin{cases} \Omega - 1 \in H(\mathbf{D}), \Omega(0) - 1 = 0 & , \quad n = 0 \\ h \in H(\mathbf{D}) \text{ and } h(0) = 0 & , \quad n > 0. \end{cases} \tag{14.6}$$

**Proof.** For $n = 0$, the relation (14.5) is obvious knowing that $\phi_0 = \psi_0 = 1$.

Use (14.3) for $k = n-1$ and $n > 0$ to write the left hand side of (14.5) as

$$2\int \frac{t}{t-z}\frac{\phi_n(t)}{B_{n-1}(t)}d\mu(t). \tag{14.7}$$

For $z = 0$, the integral equals

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

This implies that $\phi_n/B_{n-1} \in L_1(\mu)$. Therefore (14.7) is analytic in $\mathbf{D}$ as a Cauchy-Stieltjes integral.

For the relation (14.6), one can similarly check the case $n = 0$ and for $n > 0$, use (14.4) with $k = 0$ to see that the left hand side equals

$$\int D(t,z)\frac{\phi_n^*(t)}{B_n(t)}d\mu(t) \tag{14.8}$$

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 (14.8) as

$$\begin{aligned}
\int D(t,z)\frac{\phi_n^*(t)}{B_n(t)}d\mu(t) &= \int D(t,z)\frac{\phi_n^*(t)}{B_n(t)}d\mu(t) + \int \frac{\phi_n^*(t)}{B_n(t)}d\mu(t) \\
&= \int [D(t,z)+1]\frac{\phi_n^*(t)}{B_n(t)}d\mu(t) \\
&= \int \frac{2t}{t-z}\frac{\phi_n^*(t)}{B_n(t)}d\mu(t)
\end{aligned}$$

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

This proves the theorem. □

Like in the polynomial case, these functions of the second kind satisfy the same recurrence relations as the orthogonal functions. We prove the following theorem.

**Theorem 14.5** *For the functions of the second kind a recursion of the following form exists*

$$\begin{bmatrix} \psi_n(z) \\ -\psi_n^*(z) \end{bmatrix} = N_n \frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z} \begin{bmatrix} 1 & \overline{\lambda}_n \\ \lambda_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \psi_{n-1}(z) \\ -\psi_{n-1}^*(z) \end{bmatrix} \tag{14.9}$$

*where the recurrence matrix is exactly as in Theorem 13.1.*

**Proof.** As in the case of Theorem 13.1, it is sufficient to prove only one of the two associated recursions. The other one follows by duality. We shall prove the second one. First note that by our previous lemma's we can write for $n > 1$

$$\begin{bmatrix} \psi_{n-1}(z) \\ -\psi_{n-1}^*(z) \end{bmatrix} = -\Omega(z)\begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} + \int \frac{2z}{t-z}\begin{bmatrix} \phi_{n-1}(t) \\ \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)}\phi_{n-1}^*(t) \end{bmatrix} d\mu(t).$$

Multiply from the left with

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

then the right hand side becomes

$$-\Omega(z)\phi_n^*(z) + 2ze_n\eta_n^2\int\frac{f(t,z)}{t-z}d\mu(t) \tag{14.10}$$

with

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

Using the recursion for $\phi_n^*$, we thus get that (14.10) can be replaced by

$$-\Omega(z)\phi_n^*(z) + 2z\int\frac{1}{t-z}\frac{\alpha_{n-1}-z}{\alpha_{n-1}-t}\frac{1-\overline{\alpha}_n t}{1-\overline{\alpha}_n z}\phi_n^*(t)d\mu(t).$$

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

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

This can indeed be done, since the difference equals

$$\begin{aligned}
&2z\int\frac{1}{t-z}\left[\frac{\alpha_{n-1}-z}{\alpha_{n-1}-t} - \frac{\alpha_n-z}{\alpha_n-t}\right]\frac{1-\overline{\alpha}_n t}{1-\overline{\alpha}_n z}\phi_n^*(t)d\mu(t) \\
&\quad = 2z\int f_*(t)\phi_n^*(t)d\mu(t) = 2z\langle\phi_n^*, f\rangle_\mu
\end{aligned}$$

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 $\psi_1(z) = e_1\eta_1^1(z - \overline{\lambda}_1)/(1 - \overline{\alpha}_1 z)$. From the definition, we get

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

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

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

Now we use the expression we get from Theorem 13.1 for $\phi_1$ in terms of $\lambda_1$ and express the orthogonality relation $\langle \phi_1, 1\rangle_\mu = 0$ to find

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

Fill this into the last expression and you find

$$\psi_1(z) = e_1\eta_1^1 \frac{z-\overline{\lambda}_1}{1-\overline{\alpha}_1 z}(1+\overline{\alpha}_1\overline{\lambda}_1)\int \frac{1}{1-\overline{\alpha}_1 t}d\mu(t). \tag{14.11}$$

We have to find an expression for the remaining integral. Therefore we use again the expression for $\lambda_1$ from Theorem 13.2 to get

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

From this relation we finally get

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

Now the recursion for $\psi_1$ is proved and this concludes the proof of the Theorem. □

The interpolation properties cause the following Theorem to be true.

**Theorem 14.6** *Let $\phi_n$ be the orthonormal functions of $\mathcal{L}_n$ with respect to the measure $\mu$. Define the absolutely continuous measure $\mu_n$ 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 $\mu_n$ and $\mu$ is the same : $\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_{\mu_n}$.*

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

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 $\phi_n^*$ 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 $\phi_n$ is a function of norm 1 and orthogonal to $\mathcal{L}_{n-1}$ both with respect to $\mu_n$ and with respect to $\mu$. By Theorem 13.6 this $\phi_n$ will uniquely define all the previous $\phi_k$, provided they are normalized properly with $\phi_k^*(\alpha_k) = \kappa_k > 0$. Thus the orthonormal system in $\mathcal{L}_n$ for $\mu_n$ and for $\mu$ 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 $\phi_k$, it also holds that $\langle f, g \rangle_\mu = \langle f, g \rangle_{\mu_n}$ for every $f$ and $g \in \mathcal{L}_n$. □

The previous Theorem was proved for orthogonal polynomials e.g. in [36, p. 199].

We shall now derive some determinant formula and some other properties of these functions like we did for the kernels and for the orthogonal functions at the end of the previous section. Therefore we need a J-unitary matrix. This is obtained in the next Lemma.

**Lemma 14.7** *Let $t_n$ denote the recursion matrix*

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

*with all the parameters as defined in Theorems 13.1 and 14.5. Set $T_n = t_n t_{n-1} \cdots t_1$ (recall $\alpha_0 = 0$). Then*

$$T_n = \frac{1}{2} \begin{bmatrix} \phi_n + \psi_n & \phi_n - \psi_n \\ \phi_n^* - \psi_n^* & \phi_n^* + \psi_n^* \end{bmatrix}. \tag{14.12}$$

*There exists a positive constant $c_n$ such that*

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

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

**Proof.** The first relation follows easily from

$$\begin{bmatrix} \psi_n & \phi_n \\ -\psi_n^* & \phi_n^* \end{bmatrix} = T_n \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{14.13}$$

by inverting the right most matrix.

Now note that $t_k$ can be written as

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

with

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

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

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

**Theorem 14.8** *With the notation introduced in the previous Lemma, we have :*

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

2. $\frac{1}{2} \left[ \frac{\psi_n}{\phi_n} + \frac{\psi_n^*}{\phi_n^*} \right] = \frac{P(z, \alpha_n)}{\phi_n \phi_{n*}}$ *with $P(z, w)$ the Poisson kernel.*

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

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

**Proof.** The first determinant relation follows by taking the determinant of (14.13), giving

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

The second relation is a direct consequence of part 1.

That $\psi_n^*/\phi_n^* \in \mathcal{P}$ is because $\Theta_n$ is J-contractive and the factor $c_n/(1-\overline{\alpha}_n z)$ relating $\Theta_n$ 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. □

As with the functions $\phi_n$, we could rotate the functions of the second kind to give $\Psi_n = \epsilon_n\psi_n$ where the $\epsilon_n$ are as defined in (13.7). For these rotated $\Psi_n$ a recurrence like in Theorem 13.3 exists. Most of the properties of $\psi_n$ are transferred to $\Psi_n$.

## 15 Measures and interpolation

In this section, we want to show how interpolation properties of positive real functions is practically equivalent with the equality of inner products in some $\mathcal{L}_n$-spaces. In Section 14 we already saw that $\Omega_n = \psi_n^*/\phi_n^*$ was in $\mathcal{P}$ and interpolated $\Omega_\mu \in \mathcal{P}$ in the points of $A_n^0 = \{0, \alpha_1, \ldots, \alpha_n\}$. In the same section, it was shown that $d\mu_n = P(\cdot, \alpha_n)/|\phi_n|^2 d\lambda = \Re\, \Omega_n d\lambda$ and $d\mu$ are two measures defining the same inner product in $\mathcal{L}_n$. That this is not a coincidence will be shown in this section.

We shall therefore first prove a simple Lemma.

**Lemma 15.1** *Let $\mu$ be a normalized positive measure on $\mathbf{T}$ and let the positive real function $\Omega = \Omega_\mu$ be associated with it by (2.3) with $c = 0$. This means that $\Omega_\mu(0) > 0$. Define also the positive real function $\Omega_\mu(z, w)$ with $w \in \mathbf{D}$ some parameter, by*

$$\Omega_\mu(z, w) = \int \frac{D(t, z)}{P(t, w)} d\mu(t) + c$$

*with $c = c_w[\overline{w}c_{-1} - wc_1] \in \mathbf{iR}$, $c_w = 1/(1 - |w|^2)$, $P(t, w)$ being the Poisson kernel and $c_k = \overline{c}_{-k}$ the moments of $\mu$. Then the relation between $\Omega_\mu(z)$ and $\Omega_\mu(z, w)$ is given by*

$$\Omega_\mu(z, w) = c_w[(z - w)(z^{-1} - \overline{w})\Omega_\mu(z) + (z^{-1}w - \overline{w}z)c_0] = \frac{\Omega_\mu(z)}{P(z, w)} + c_w(z^{-1}w - \overline{w}z)c_0.$$

*Note that $c_0 = \Omega_\mu(0) = \Omega_\mu(w, w)$.*

**Proof.** First, note the simple relations

$$\begin{aligned} D(t, z)t &= (t + z) + zD(t, z) \\ D(t, z)t^{-1} &= -(t^{-1} + z^{-1}) + z^{-1}D(t, z). \end{aligned}$$

This can be used to give the following relations.

$$\begin{aligned} \Omega_\mu(z, w) &= c_w \int D(t, z)(t - w)(t^{-1} - \overline{w})d\mu(t) + c \\ &= c_w(z - w)(\frac{1}{z} - \overline{w}) \int D(t, z)d\mu(t) + c_w(\frac{w}{z} - \overline{w}z) \int d\mu + c_w \int (\frac{w}{t} - \overline{w}t)d\mu(t) + c \\ &= c_w(z - w)(z^{-1} - \overline{w})\Omega_\mu(z) + c_w(z^{-1}w - \overline{w}z)c_0 \end{aligned}$$

and this proves the Lemma. □

Note that the choice of $c$ in the previous Lemma is such that $\Omega_\mu(w, w) = \Omega_\mu(0) = 1 \in \mathbf{R}$. You can also see easily that $\Omega_\mu(z, 0) = \Omega_\mu(z)$.

Now we get to an interpolation result. Part of it confirms what was shown in Theorem 6.4.

**Theorem 15.2** *Let $\mu$ be a normalized positive measure on $\mathbf{T}$ and associate with it as in Lemma 15.1 the positive real functions $\Omega_\mu(z)$ and $\Omega_\mu(z, w)$, for some arbitrary $w \in \mathbf{D}$. Let $\nu$ be another normalized positive measure on $\mathbf{T}$ and associate with it in a similar way the positive real functions $\Omega_\nu(z)$ and $\Omega_\nu(z, w)$. Suppose that on $\mathcal{L}_n$ the inner product w.r.t. $\mu$ and $\nu$ is the same : $\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_\nu$. Then, with $\pi_n^* = \prod_{i=1}^n (z - \alpha_i)$,*

$$\frac{\Omega_\mu(z, w) - \Omega_\nu(z, w)}{\pi_n^*(z)} = g_w(z) \in H(\mathbf{D}) \text{ and } g_w(w) = 0. \tag{15.1}$$

*This means that $\Omega_\nu(\cdot, w)$ is an Hermite interpolant for $\Omega_\mu(\cdot, w)$ in the point set $A_n^w = \{w, \alpha_1, \ldots, \alpha_n\}$, multiplicities counted.*

**Proof.** By Lemma 15.1 we know that (15.1) holds if and only if $(\Omega_\mu(z) - \Omega_\nu(z))/\pi_n^*(z)$ is analytic in $\mathbf{D}$ and vanishes at the origin. Thus we shall prove the latter instead of (15.1).

Consider

$$\left[\frac{1}{\pi_n^*(z)} - \frac{1}{\pi_n^*(t)}\right] D(t,z) = \frac{1}{\pi_n^*(t)} \left[\frac{\pi_n^*(t) - \pi_n^*(z)}{\pi_n^*(z)} \frac{t+z}{t-z}\right] = \frac{p^*(t)}{\pi_n^*(t)}$$

Because $\pi_n^*(t) - \pi_n^*(z)$ is zero for $t = z$, we may conclude that the factor $p^*$ in the right hand side is a polynomial of degree $n$ in $t$. Since

$$\frac{p^*}{\pi_n^*} = \frac{p_*}{\pi_{n*}} = f_*(t) \text{ with } f = \frac{p}{\pi_n} \in \mathcal{L}_n$$

we can see that

$$0 = \langle 1, f\rangle_{\mu-\nu} = \frac{1}{\pi_n^*(z)} \int D(t,z) d(\mu-\nu)(t) - \int \frac{D(t,z)}{\pi_n^*(t)} d(\mu-\nu)(t).$$

Thus, by evaluating the first term of the right hand side

$$\frac{\Omega_\mu(z) - \Omega_\nu(z)}{\pi_n^*(z)} = \int \frac{D(t,z)}{\pi_n^*(t)} d(\mu-\nu)(t).$$

By construction, the right hand side is analytic in $\mathbf{D}$. Furthermore, $\Omega_\mu(0) - \Omega_\nu(0) = \int d(\mu-\nu) = 0$, so that (15.1) is now proved. □

There is also a kind of inverse for Theorem 15.2.

**Theorem 15.3** *Suppose that for some positive real functions $\Omega_\mu$ and $\Omega_\nu$,*

$$\frac{\Omega_\mu - \Omega_\nu}{\pi_n^*} = g \in H(\mathbf{D}) \textit{ and } g(w) = 0 \textit{ for some } w \in \mathbf{D}. \tag{15.2}$$

*The measures $\mu$ and $\nu$ are the Riesz-Herglotz measures for $\Omega_\mu$ and $\Omega_\nu$ respectively. Then the measures*

$$d\mu_w(z) = P(z,w) d\mu(z) \textit{ and } d\nu_w(z) = P(z,w) d\nu(z), \quad z \in \mathbf{T}$$

*with $P(z,w)$, the Poisson kernel, define on $\mathcal{L}_n$ the same inner products : $\langle\cdot,\cdot\rangle_{\mu_w} = \langle\cdot,\cdot\rangle_{\nu_w}$.*

**Proof.** We can again use Lemma 15.1 to find that $\Omega_\mu$ is an Hermite interpolant of $\Omega_\nu$ in the point set $A_n^w$ if and only if $\Omega_{\mu_w}$ is an Hermite interpolant of $\Omega_{\nu_w}$ in the point set $A_n^0$. Thus the problem reduces to the following one. If $\Omega_{\mu_w}$ is an Hermite interpolant in $A_n^0$ for $\Omega_{\nu_w}$, prove that on $\mathcal{L}_n$, $\langle\cdot,\cdot\rangle_{\mu_w} = \langle\cdot,\cdot\rangle_{\nu_w}$. It follows from Theorem 6.3 that we can choose a basis in $\mathcal{L}_n$ which will make the Gram matrix only depend on function values and derivatives of $\Omega_{\mu_w}$ and $\Omega_{\nu_w}$ respectively in points from the set $A_n^0$. By the mutual interpolation property these are the same. Thus the Gram matrices are the same and therefore also the inner products. □

This Theorem has also a Corollary which is worth formulating.

**Corollary 15.4** *Let $\mu$ be a given measure to which we associate the positive real function $\Omega_\mu(z,w)$ as in Lemma 15.1. Suppose $\Omega_n$ is a positive real function which is an Hermite interpolant for $\Omega_\mu(z,w)$ in the point set $A_n^w$. Let $\mu_n$ be the Riesz-Herglotz measure corresponding with $\Omega_n$. Define the measure $\mu_n^w$ by $d\mu_n^w(z) = P(z,w) d\mu_n(z)$ with $P(z,w)$ the Poisson kernel. Then on $\mathcal{L}_n$, the inner product $\langle\cdot,\cdot\rangle_{\mu_n^w}$ is independent of $w$.*

**Proof.** You can apply the previous Theorem with $\Omega_n(z)$ in the role of $\Omega_\mu(z,w)$ and $\Omega_\mu(z,w)$ in the role of $\Omega_\nu(z,w)$. Hence we have to substitute for the corresponding measures $d\mu(z)$ and $d\nu(z)$ respectively $d\mu(z)/P(z,w)$ and $\mu_n(z)$. You will find that on $\mathcal{L}_n$ the inner products w.r.t. $\mu_n^w$ and $\mu$ are the same : $\langle\cdot,\cdot\rangle_{\mu_n^w} = \langle\cdot,\cdot\rangle_\mu$. This implies that $\langle\cdot,\cdot\rangle_{\mu_n^w}$ is independent of $w$. □

## 16 Quadrature on the unit circle

We now consider the functions

$$f_n(z) = f_n(z, w) = \phi_n(z) + w\phi_n^*(z) \in \mathcal{L}_n \qquad (16.1)$$

with $\phi_n$ the orthonormal functions of $\mathcal{L}_n$. In accordance with the polynomial costume, we can call these quasi-orthogonal functions. See e.g., in Akhiezer [1, p.10]. The function $f_n$ is orthogonal to $\mathcal{L}_{n-1} \cap \zeta_n \mathcal{L}_{n-1}$ since $\langle \phi_n + w\phi_n^*, g\rangle_\mu = \langle \phi_n, g\rangle_\mu + w\langle \phi_n^*, g\rangle_\mu = 0$ for any $g \in \mathcal{L}_{n-1} \cap \zeta_n \mathcal{L}_{n-1}$. Note that it holds that

$$\mathcal{L}_{n-1} \cap \zeta_n \mathcal{L}_{n-1} = \{g \in \mathcal{L}_n : g(\alpha_n) = g^*(\alpha_n) = 0\}.$$

(see Theorem 7.1). The zeros of these quasi-orthogonal functions have the following property.

**Theorem 16.1** *Let $w \in \mathbf{T}$ be given and define the quasi-orthogonal functions $f_n$ by (16.1). Then all the zeros of $f_n$ are on $\mathbf{T}$ and they are simple.*

**Proof.** Since $\phi_n^*$ doesn't vanish in $\mathbf{D}$, the ratio $\phi_n/\phi_n^*$ is well defined in $\mathbf{D}$ and it equals $p_n/p_n^*$ if the polynomial $p_n \in \Pi_n$ is defined by $\phi_n = p_n/\pi_n$ with as always $\pi_n(z) = \prod_{i=0}^n (1 - \overline{\alpha}_i z)$. Since $b_n = p_n/p_n^*$ is a finite Blaschke product, with all its poles in $\mathbf{E}$, we have $|b_n(z)| < 1$ in $\mathbf{D}$. Let $\alpha$ be a zero of $f_n$ in $\mathbf{D}$, then $f_n(\alpha) = 0$ and this implies $p_n(\alpha)/p_n^*(\alpha) = b_n(\alpha) = -w$. Since $w \in \mathbf{T}$ and $|b_n(\alpha)| < 1$, we get a contradiction. Hence, $f_n$ has no zeros in $\mathbf{D}$. Because on the other hand, $f_n^* = \overline{w} f_n$, it follows that if $\alpha$ is a zero of $f_n$, then also $f_n(1/\overline{\alpha}) = \overline{B_{n*}(\alpha)} w \overline{f_n(\alpha)} = 0$. Thus zeros appear in pairs $(\alpha, 1/\overline{\alpha})$. Because we showed that there are no zeros in $\mathbf{D}$, there can also be no zeros in $\mathbf{E}$ by this duality property. Thus we may conclude that all the zeros are on $\mathbf{T}$.

Now we prove that they are simple zeros. Suppose there are only $s \leq n-2$ zeros with an odd multiplicity. Call them $\xi_1, \ldots, \xi_s$, possibly repeated if they are multiple. They are all in $\mathbf{T}$. Now we note that $(z - \xi_i)^2 = (z - \xi_i)(z - 1/\overline{\xi}_i) = c_i z (z - \xi_i)(z - \xi_i)_*$ with $c_i = -\xi_i$. Hence, $f_n = N/\pi_n$ with $N$ of the form

$$N(z) = c \prod_{i=1}^{s} (z - \xi_i) z^t \prod_{i=s+1}^{s+t} (z - \xi_i)(z - \xi_i)_* \text{ with } n - s = 2t, \quad t \geq 1$$

for some constant $c$. Consider now the function $T(z) = M(z)/\pi_n(z)$ with $M(z)$ of the form $c \prod_{i=1}^s (z - \xi_i) z^{t-1} (z - \alpha_n)$. Clearly $T \in \mathcal{L}_{n-1}$, hence it is orthogonal to $\phi_n$, but also $T(z) \in (z - \alpha_n)\mathcal{L}_{n-1} = (z - \alpha_n)\,\mathrm{span}\{B_k : k = 1, \ldots, n-1\}$. This means that it is of the form $(z - \alpha_n) p_{n-1}(z)/\pi_{n-1}$ with $p_{n-1} \in \Pi_{n-2}$. Thus it is also in $\zeta_n \mathcal{L}_{n-1}$ and therefore $\langle T, \phi_n^*\rangle_\mu = 0$. Consequently, $T$ is orthogonal to $f_n$. On the other hand, if we write explicitly $\langle f_n, T\rangle_\mu$, we get

$$|c|^2 \int \frac{\prod_{i=1}^s |z - \xi_i|^2}{|\pi_n(z)|^2} \prod_{i=s+1}^{s+t} |z - \xi_i|^2 d\mu = \|S\|_\mu^2 > 0$$

since $S \not\equiv 0$. This is a contradiction so that $s \geq n-1$. This means that all the zeros of $f_n$ should be simple and on $\mathbf{T}$. □

Now we consider the space of rational functions

$$\mathcal{R}_n = \mathcal{L}_n + \mathcal{L}_{n*}$$

where $\mathcal{L}_{n*} = \{f : f_* \in \mathcal{L}_n\}$. The space $\mathcal{R}_n$ can also be characterized as

$$\mathcal{R}_n = \{\frac{t_n(z)}{\pi_n(z)\pi_{n*}(z)} : t_n \in \Pi_n + \Pi_{n*}\}.$$

$\Pi_n + \Pi_{n*}$ is the set of trigonometric polynomials $\sum_{k=-n}^n a_k z^k$, and $\pi_n(z) = \prod_{k=0}^n (1 - \overline{\alpha}_k z)$. Let $\ell_{ni}(z) = \prod_{k \neq i} (z - \xi_k)/(\xi_i - \xi_k)$ denote the Lagrange polynomial for the interpolation points

$\xi_1, \xi_2, \ldots, \xi_n$, so that $\sum_{k=1}^n R(\xi_k)\ell_{nk}(z)$ is the interpolating polynomial of degree $n$ for $R$, which we take from $\mathcal{R}_{n-1}$. Let $L_{ni}$ be defined by

$$L_{ni}(z) = \ell_{ni}\frac{\pi_{n-1}(\xi_i)}{\pi_{n-1}(z)}$$

and $\xi_k, k = 1, \ldots, n$ the zeros of $f_n = \phi_n + w\phi_n^*$, which are all on $\mathbf{T}$. Note that $L_{nk}(\xi_i) = \delta_{ki}$, so that $\sum_1^n R(\xi_k)L_{nk}(z)$ is an interpolant for $R(z)$ from $\mathcal{L}_{n-1}$ in the points $\xi_1, \ldots, \xi_n$. Consider

$$\begin{aligned} E(z) &= R(z) - \sum_{k=1}^n R(\xi_k)L_{nk}(z) \\ &= \frac{t_{n-1}(z) - \sum_{k=1}^n R(\xi_k)\ell_{nk}(z)\pi_{n-1}(\xi_k)\pi_{n-1\,*}(z)}{\pi_{n-1}(z)\pi_{n-1\,*}(z)}. \end{aligned}$$

From the interpolating property, we find

$$E(z) = \frac{\prod_{k=1}^n (z - \xi_k)q_{n-2}(z)z^{-(n-1)}}{\pi_{n-1}(z)\pi_{n-1\,*}(z)}, \quad q_{n-2} \in \Pi_{n-2}.$$

Now, because $\xi_i, i = 1, \ldots, n$ are the zeros of $f_n(z)$, we can write this also as

$$E(z) = f_n(z)\frac{(1 - \overline{\alpha}_n z)q_{n-2}(z)z^{-(n-1)}}{\pi_{n-1\,*}(z)}.$$

The second factor can be written as $S_*$ with $S$ defined by

$$S(z) = \frac{(z - \alpha_n)q_{n-2}^*(z)}{\pi_{n-1}(z)}.$$

Observe that $S \in \mathcal{L}_{n-1}$ and also $S \in \zeta_n\mathcal{L}_{n-1}$, so that $S$ is orthogonal to $\phi_n$ and to $\phi_n^*$, hence orthogonal to $f_n$. Thus $\langle f_n, S\rangle_\mu = \int E d\mu = 0$. In other words, if $R \in \mathcal{R}_{n-1}$ and $\xi_k, k = 1, \ldots, n$ are the zeros of $f_n = \phi_n + w\phi_n^*$, then

$$\int R d\mu = \sum_{i=1}^n R(\xi_i)\int L_{ni}d\mu = \sum_{i=1}^n R(\xi_i)\lambda_{ni},$$

which is the Gauss quadrature on the unit circle.

Since $L_{ni}L_{ni*} \in \mathcal{R}_{n-1}$ and also $L_{ni}L_{ni*} - L_{ni} \in \mathcal{R}_{n-1}$, we get

$$\int (L_{ni}L_{ni*} - L_{ni})d\mu = \sum \lambda_{ni}0 = 0.$$

Thus

$$\lambda_{ni} = \int L_{ni}d\mu = \int |L_{ni}|^2 d\mu > 0.$$

Thus we have proved the following Theorem.

**Theorem 16.2** *Let $\{\phi_k\}$ be an orthonormal system for $\mathcal{L}_n$, $w \in \mathbf{T}$ and $f_n = \phi_n + w\phi_n^*$. Then $f_n$ has $n$ simple zeros on $\mathbf{T}$ : $\xi_1, \ldots, \xi_n$. Let $\ell_{ni}$ denote the Lagrange polynomials for the interpolation points $\xi_1, \ldots, \xi_n$. Then the Gauss quadrature formula*

$$\sum_{i=1}^n R(\xi_i)\lambda_{ni} \textit{ with } \lambda_{ni} = \int \frac{\ell_{ni}(z)}{\pi_{n-1}(z)}\pi_{n-1}(\xi_i)d\mu > 0$$

*is exact for all $R \in \mathcal{R}_{n-1} = \mathcal{L}_{n-1} + \mathcal{L}_{n-1\,*}$.*

# 17 Interpolation properties for the kernels

In this section we will show how to find a rational interpolant, or if you prefer, a *multipoint Padé type approximant* for the positive real function $\Omega = \Omega_\mu$ that was related to the measure $\mu$ by the Poisson-Stieltjes integral (2.3).

We first prove the following Theorem which can be found in [25, p.48] or [26, p.654].

**Theorem 17.1** *Let $K_n(z,w)$ be the normalized kernel of $(\mathcal{L}_n,\mu)$ and define the absolutely continuous measure on the unit circle $d\mu_n(z) = |K_n(z,0)|^{-2}d\lambda(z)$, $z \in \mathbf{T}$. Then in the space $\mathcal{L}_n$ we have equality of the inner products $\langle\cdot,\cdot\rangle_{\mu_n} = \langle\cdot,\cdot\rangle_\mu$*

**Proof.** We only have to show that for all $f \in \mathcal{L}_n$ and arbitrary $w \in \mathbf{D}$

$$\langle f, K_n(\cdot,w)\rangle_{\mu_n} = \langle f, K_n(\cdot,w)\rangle_\mu = f(w)/K_n(w,w) \tag{17.1}$$

because the $K_n(\cdot,w_i)$ with $w_i, i=0,\ldots,n$ some set of distinct points in $\mathbf{D}$ form a basis for $\mathcal{L}_n$.

We first note that, if we define $\pi_n(z) = \prod_{i=1}^n(1-\overline{\alpha}_i z)$, and let $\{p_k\}$ be a system of orthonormal polynomials with respect to the measure $d\pi = |\pi_n(z)|^{-2}d\mu$, then

$$k_\pi(z,w) = \sum_{k=0}^n p_k(z)\overline{p_k(w)}$$

is a reproducing kernel for the space $(\Pi_n,\pi)$ of polynomials where the inner product is with respect to the measure $\pi$. Thus for any polynomial $q \in \Pi_n$, it holds that

$$\langle q, k_\pi(\cdot,w)\rangle_\pi = q(w).$$

Hence, dividing by $\pi_n(w)$, we get

$$\langle \frac{q(\cdot)}{\pi_n(\cdot)}, \frac{k_\pi(\cdot,w)}{\pi_n(\cdot)\overline{\pi_n(w)}}\rangle_\mu = \frac{q(w)}{\pi_n(w)}$$

which means that

$$\frac{k_\pi(z,w)}{\pi_n(z)\overline{\pi_n(w)}} = k_n(z,w)$$

is the reproducing kernel for $(\mathcal{L}_n,\mu)$.

On the other hand $K_n(z,0) = p_n^*(z)/\pi_n(z)$ since indeed, by the Christoffel-Darboux relation for polynomials (see e.g., [42, p.41] or derive it from our Theorem 10.3 by setting all $\alpha_i = 0$)

$$k_\pi(z,w) = \frac{p_n^*(z)\overline{p_n^*(w)} - z\overline{w}p_n(z)\overline{p_n(w)}}{1-z\overline{w}}$$

so that $k_\pi(z,0) = p_n^*(z)\overline{p_n^*(0)}$. If we now suppose the polynomials $p_n$ to be chosen such that $p_n(0) > 0$, then it easily follows that $K_n(z,0) = p_n^*(z)/\pi_n(z)$.

With these tools, we can now see that (17.1) will be proved if we can show that

$$\langle p(\cdot)/\pi_n(\cdot), k_n(\cdot,w)\rangle_{\mu_n}\pi_n(w) = p(w)$$

for arbitrary polynomial $p \in \Pi_n$. The left hand side can be written explicitly as

$$\int \frac{|\pi_n(z)|^2}{|p_n^*(z)|^2}\frac{p(z)}{|\pi_n(z)|^2}\left[\frac{z\overline{p_n^*(z)}p_n^*(w) - w\overline{p_n(z)}p_n(w)}{z-w}\right]d\lambda(z)$$

$$\begin{aligned}
&= p_n^*(w)\int \frac{p(z)}{|p_n^*(z)|^2}\frac{z\overline{p_n^*(z)}}{z-w}d\lambda(z) - wp_n(w)\int\frac{p(z)}{|p_n^*(z)|^2}\frac{\overline{p_n(z)}}{z-w}d\lambda(z) \\
&= p_n^*(w)\int\frac{p(z)}{p_n^*(z)}\frac{z}{z-w}d\lambda(z) - wp_n(w)\int\frac{p(z)}{p_n(z)}\frac{1}{z-w}d\lambda(z) \\
&= p_n^*(w)\frac{p(w)}{p_n^*(w)} - wp(w)0 = p(w).
\end{aligned}$$

The last line follows from the Cauchy formula and the fact that $p_n^*(z) \neq 0$ in $\overline{\mathbf{D}}$ while $p_n(z) \neq 0$ in $\overline{\mathbf{E}}$. $\square$

Note that for the polynomial case, i.e., when all $\alpha_i = 0$, then $d\mu_n = |\phi_n^*(z)|^{-2} d\lambda = |\phi_n(z)|^{-2} d\lambda$ and a theorem in this style can be found in [36, p.198].

From the previous result and the Theorems from Section 15 we can now find the interpolation property.

**Theorem 17.2** *Let us define*

$$\Omega(z) = \int D(t,z) d\mu(t) \text{ and } \Omega_n(z) = \int D(t,z) d\mu_n(t), \qquad D(t,z) = (t+z)/(t-z)$$

*with $d\mu_n(z) = |K_n(z,0)|^{-2} d\lambda(z)$ as defined in Theorem 17.1. Then*

$$\Omega_n(z) = \frac{L_n(z,0)}{K_n(z,0)} \tag{17.2}$$

*with $K_n(t,z)$ the normalized kernel and $L_n(t,z)$ the associated function as defined in (12.8). Furthermore*

$$\mu_n' = \omega_n = \frac{1}{2}[\Omega_n + \Omega_{n*}] = \frac{1}{K_n(z,0) K_{n*}(z,0)} \tag{17.3}$$

*is positive on $\mathbf{T}$ and the spectral factor of $d\mu_n$ is given by*

$$\sigma_n(z) = 1/K_n(z,0). \tag{17.4}$$

*Moreover, the function $g$ defined by*

$$\frac{\Omega - \Omega_n}{B_n} = g \tag{17.5}$$

*is analytic in $\mathbf{D}$ and $g(0) = 0$.*

**Proof.** By Theorem 12.3, we get for $t \in \mathbf{T}$,

$$\frac{1}{|K_n(t,0)|^2} = \Re \left[\frac{L_n(t,0)}{K_n(t,0)}\right] = \Re\, \Omega_n(t)$$

while $\Omega_n \in \mathcal{P}$. Thus (17.2) follows by (2.7). Since $1/K_n(z,0)$ is outer in $H_2$, also (17.4) follows and (17.5) is a consequence of the previous Theorem. $\square$

The interpolation result of the last Theorem states that $\Omega_n$ is a *partial multipoint Padé approximant* of $\Omega$ in the points of $A_n^0 = \{0\} \cup A_n = \{0, \alpha_1, \ldots, \alpha_n\}$. It is only a *partial* interpolant since it is of degree type $(n/n)$ while only $n+1$ interpolation conditions are satisfied. Because $\Omega - \Omega_n = z B_n g$ with $g$ analytic in $\mathbf{D}$, we find by taking the substar conjugate that $\Omega_* - \Omega_{n*} = z^{-1} B_{n*} g_*$ with $g_*$ analytic in $\mathbf{E}$. Summing up gives for $\omega_n = \mu_n' = [K_n(z,0) K_{n*}(z,0)]^{-1}$

$$\mu' - \mu_n' = z B_n g + z^{-1} B_{n*} g_*$$

This generalizes the notion of Laurent-Padé approximant [14], since in the case where all $\alpha_i = 0$, $\omega_n$ is the inverse of a Laurent polynomial of degree $n$, which fits the expansion of $\mu'$ from $-n$ till $+n$. In the present case, $\omega_n$ takes the form

$$\omega_n = \frac{1}{K_n(z,0) K_{n*}(z,0)} = \frac{\pi_n(z) \pi_{n*}(z)}{p_n(z) p_{n*}(z)}$$

where $p_n(z) = K_n(z,0) \pi_n(z) \in \Pi_n$ is a polynomial. Thus $\omega_n$ is the ratio of two Laurent polynomials of degree $n$ and fits only $2n+2$ interpolation conditions. It is a *partial* Laurent-Padé approximant since by fixing the interpolation points $\alpha_i$, one fixes the zeros and hence the numerator of the approximant.

# 18 The interpolation algorithm of Pick-Nevanlinna

In this section we describe the algorithm of Pick-Nevanlinna for interpolation of class $\mathcal{P}$ functions or equivalently class $\mathcal{B}$ functions. The recursions can be described by J-unitary matrices. This approach gives an alternative way of computing the recurrence coefficients $\rho_k$ and $\gamma_k$, exactly like the duality of the two algorithms considered in [14].

Suppose we start with some function $S_0 \in \mathcal{B}$ which is zero at $w \in \mathbf{D}$. Since it depends on the parameter $w$, we write it as a function in $z$ but include the dependence on $w$ explicitly when appropriate. We now transform this $S_0$ into some other $S_1 \in \mathcal{B}$ in three steps. Let $\alpha_1$ in $\mathbf{D}$ be given. Then

$$S_1(z,w) = \tau_{31} \circ \tau_{21} \circ \tau_{11}(S_0(z,w)) = \tau_1(S_0(z,w))$$

where

$$\begin{aligned}
\tau_{11} \quad &: \quad S_0 \mapsto S_1' = \frac{S_0 - \gamma_1}{1 - \overline{\gamma}_1 S_0}, \quad \gamma_1 = \gamma_1(w) = S_0(\alpha_1, w) \\
\tau_{21} \quad &: \quad S_1' \mapsto S_1'' = S_1'/\zeta_1, \quad \zeta_1(z) = \frac{\overline{\alpha}_1}{|\alpha_1|}\frac{\alpha_1 - z}{1 - \overline{\alpha}_1 z} \\
\tau_{31} \quad &: \quad S_1'' \mapsto S_1 = \frac{S_1'' - \rho_1}{1 - \overline{\rho}_1 S_1''}, \quad \rho_1 = \rho_1(w) = S_1''(w,w)
\end{aligned}$$

Clearly, $S_1$ is again a function in $\mathcal{B}$ and it will be zero at $w$. This follows from Theorem 4.3. What was done in the previous transformations is the following. First, $S_0$ is transformed into $S_1'$ to make it zero in $\alpha_1$. This zero can be taken out by dividing with $\zeta_1$. The last step will normalize $S_1$ by making it zero at $w$ just like $S_0$ was. We are now in a position like the one we started with and we can repeat the same procedure with some point $\alpha_2$ from $\mathbf{D}$ to produce $S_2$. Note that this procedure can go on indefinitely as long as the $\alpha_i$ is in $\mathbf{D}$ since both $\gamma_i$ and $\rho_i$ will be in $\mathbf{D}$ as evaluations of functions in $\mathcal{B}$. Note the following relation between $\gamma_k$ and $\rho_k$. Since $S_{k-1}(w,w) = 0$, it follows that

$$S_k'(w,w) = \frac{S_{k-1}(w,w) - \gamma_k}{1 - \overline{\gamma}_k S_{k-1}(w,w)} = -\gamma_k$$

while on the other hand

$$S_k'(w,w) = \zeta_k(w) S_k''(w,w) = \zeta_k(w)\rho_k$$

so that for all $k > 0 : \gamma_k = -\zeta_k(w)\rho_k$.

We can also invert the previous procedure. Suppose the coefficients $\rho_k$ and $\gamma_k$ for $k = 1, \ldots, n$ are produced by the previous algorithm starting from some $S_0$. Now choose some $\Gamma_0 \in \mathcal{B}$ such that $\Gamma_0(w) = 0$. Then generate the sequence $\Gamma_{k+1} = \tau_{n-k}^{-1}(\Gamma_k)$ for $k = 0, \ldots, n-1$. It turns out that $\Gamma_n$ shall interpolate $S_0$ in the points $A_n^w = \{w, \alpha_1, \ldots, \alpha_n\}$. Indeed, denote $n-k$ as $j$, then it holds in general that, $\Gamma_j$ will interpolate $S_k$ in the points $\{w, \alpha_n, \ldots, \alpha_{k+1}\}$. We show this for the case where all $\alpha_k$ are different. If some of them are confluent, the proof becomes messy by technicalities. For that case, we refer to the homogeneous formulation to be given later in this section. If the interpolation points are all different, then the result follows easily by induction. For $j = 0$, we have interpolation at $w$ only since both $\Gamma_0$ and $S_n$ are zero in that point. The induction step can be proved by noting that $\Gamma_{j+1} = \tau_k^{-1}(\Gamma_j)$ and $S_{k-1} = \tau_k^{-1}(S_k)$ so that the interpolation from the previous step in inherited. There is one extra interpolation condition satisfied, viz., in the point $\alpha_k$ since $\Gamma_{j+1}(\alpha_k) = \gamma_k = S_{k-1}(\alpha_k, w)$.

We have now (almost) proved the following Theorem.

**Theorem 18.1** *Let $S_0(z,w) \in \mathcal{B}$ be a Schur function which is zero at $z = w$. Construct iteratively for a sequence of points $\{\alpha_k : k > 0\} \subset \mathbf{D}$ the functions $S_k$ by $S_k(z,w) = \tau_k(S_{k-1}(z,w))$ where*

$\tau_k = \tau_{1k} \circ \tau_{2k} \circ \tau_{3k}$,

$$\tau_{1k} \quad : \quad S_{k-1} \mapsto S'_k = \frac{S_{k-1} - \gamma_k}{1 - \overline{\gamma}_k S_{k-1}}, \quad \gamma_k = \gamma_k(w) = S_{k-1}(\alpha_k, w) \tag{18.1}$$

$$\tau_{2k} \quad : \quad S'_k \mapsto S''_k = S'_k/\zeta_k, \quad \zeta_k(z) = \frac{\overline{\alpha}_k}{|\alpha_k|} \frac{\alpha_k - z}{1 - \overline{\alpha}_k z} \tag{18.2}$$

$$\tau_{3k} \quad : \quad S''_k \mapsto S_k = \frac{S''_k - \rho_k}{1 - \overline{\rho}_k S''_k}, \quad \rho_k = \rho_k(w) = S''_k(w, w) \tag{18.3}$$

*Then all the $S_k$ are in $\mathcal{B}$ and $S_k(w,w) = 0$. All the $\gamma_k(w)$ and $\rho_k(w)$ are in $\mathbf{D}$ and $\gamma_k(w) = -\zeta_k(w)\rho_k(w)$.*

*Conversely, by choosing an arbitrary $\Gamma_0(z,w) \in \mathcal{B}$ which vanishes for $z = w$, we can construct, using the previous $\gamma_k$ and $\rho_k$ the functions $\Gamma_k$ by $\Gamma_{k+1} = \tau_{n-k}^{-1}(\Gamma_k)$. All these $\Gamma_k$ are in $\mathcal{B}$ and $\Gamma_k$ will interpolate $S_{n-k}$ in the points $\{w, \alpha_n, \ldots, \alpha_{n-k+1}\}$. Specifically*

$$\frac{\Gamma_n - S_0}{B_n(z)} = h(z), \quad h \in H(\mathbf{D}) \text{ and } h(w) = 0.$$

The previous Theorem gives algorithms to check whether a certain function is an element from $\mathcal{B}$. It was shown in the literature ([1, 79] and others) that $S_0$ is in $\mathcal{B}$ if and only if one of the following two cases occur. Either $\gamma_k$ and $\rho_k$ are all in $\mathbf{D}$ for all $k$ or they are all in $\mathbf{D}$ for all $k \le n-1$ and $|\gamma_n| = 1$ and all the remaining $\gamma$'s and $\rho$'s are zero. In the latter case, $S_0$ was a finite Blaschke product of degree $n$.

We shall now give an equivalent homogeneous formulation of the previous algorithm. Suppose that a Schur function $S \in \mathcal{B}$ is described as the ratio of two functions $S = \Delta_1/\Delta_2$ which are both holomorphic in $\mathbf{D}$ and $\Delta_2$ zero-free in $\mathbf{D}$. If $S_k(w) = 0$, then of course $\Delta_1(w) = 0$ too. We place these two functions in a vector

$$\Delta = [\Delta_1 \quad \Delta_2]$$

which can be considered as a set of homogeneous coordinates for $S$. Following Dewilde-Dym [24] we shall call this set of $\Delta$-matrices *admissible* and denote it as

$$\mathcal{A} = \{\Delta = [\Delta_1 \ \ \Delta_2] : \Delta_1, \Delta_2 \in H(\mathbf{D}), \Delta_2(z) \neq 0, z \in \mathbf{D}, \Delta_1/\Delta_2 \in \mathcal{B}\} \tag{18.4}$$

We use $\overline{\mathcal{A}}$ if $\mathcal{B}$ is replaced by $\overline{\mathcal{B}}$ in the previous definition. Note that $\Delta_1/\Delta_2 \in \mathcal{B}$ can also be written as $\Delta^H J \Delta < 1$.

We can now describe the Pick-Nevanlinna algorithm of the previous Theorem in terms of J-unitary matrix multiplications on admissible matrix functions. Let $\Delta_n = [\Delta_{n1} \quad \Delta_{n2}]$ be the admissible matrix containing the homogeneous coordinates for $S_n$. Then the inverse transform $S_{n-1} = \tau_n^{-1}(S_n)$ can be written as

$$\Delta_{n-1} = \Delta_n \theta_n$$

with the matrix $\theta_n$ given by

$$\theta_n(z,w) = c \begin{bmatrix} 1 & \overline{\rho}_n \\ \rho_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix} d \begin{bmatrix} 1 & \overline{\gamma}_n \\ \gamma_n & 1 \end{bmatrix}$$

with

$$\begin{aligned} c &= (1 - |\rho_n|^2)^{-1/2} \text{ and } d = (1 - |\gamma_n|^2)^{-1/2} \\ \gamma_n &= \gamma_n(w) = \Delta_{n-1,1}(\alpha_n, w)/\Delta_{n-1,2}(\alpha_n, w) \\ \rho_n &= \rho_n(w) = \begin{cases} \gamma_n(w)/\zeta_n(w) & \text{if } w \neq \alpha_n \\ \partial_z \left(\Delta_{n-1,1}(z,w)/\Delta_{n-1,2}(z,w)\right)\big|_{z=w} & \text{if } w = \alpha_n \end{cases} \end{aligned}$$

Let us define the J-unitary matrix $\Theta_n$ as $\Theta_n = \theta_n \cdots \theta_1$ for $n \ge 1$. Because this matrix is formally the same as the $\Theta_n$ matrix of section 12, there must exist some $K_n$ and $L_n$, both functions in

$\mathcal{L}_n$, parametrized in $w$, such that a relation like (12.9) holds. The interpolation property given in Theorem 18.1 implies that if we choose $\Delta_n = [0 \;\; 1] \in \mathcal{A}$, then $\Delta_n \Theta_n = [K_n - L_n \;\; K_n + L_n] \in \mathcal{A}$ has the property that $(K_n - L_n)/(K_n + L_n)$ interpolates $S_0$ at the points of $A_n^w$ if they are all different. We can now easily give the proof for confluent points too. If we define $\Theta_n^*$ as $B_n \Theta_{n*}$ we get the form

$$\Theta_n^* = \frac{1}{2} \begin{bmatrix} K_n + L_n & K_n^* - L_n^* \\ K_n - L_n & K_n^* + L_n^* \end{bmatrix}.$$

Because we know that for a J-unitary matrix $\Theta_n^{-1} = J\Theta_{n*}J = B_{n*}J\Theta_n^*J$, we get

$$\Delta_0 J \Theta_n^* J = B_n \Delta_n.$$

Furthermore, since $S_0 = \Delta_{01}/\Delta_{02} \in \mathcal{B}$ and vanishes at $z = w$, the positive real function $\Omega(z,w) = (1 - S_0(z,w))/(1 + S_0(z,w)) \in \mathcal{P}$ will be 1 for $z = w$ : $\Omega(w,w) = 1$. We can write $\Delta_0 = [1 - \Omega(z,w) \;\; 1 + \Omega(z,w)]$. If we multiply this from the right with $J\Theta_n^*J$, then we get

$$\frac{1}{2}[1 - \Omega \;\; 1 + \Omega] \begin{bmatrix} K_n + L_n & -K_n^* + L_n^* \\ -K_n + L_n & K_n^* + L_n^* \end{bmatrix} = B_n [\Delta_{n1} \;\; \Delta_{n2}].$$

Thus

$$[L_n - K_n \Omega \;\; L_n^* \; + K_n^* \Omega] = B_n [\Delta_{n1} \;\; \Delta_{n2}].$$

The first of these relations shows that

$$L_n(z,w) - K_n(z,w)\Omega(z,w) = B_n(z)\Delta_{n1}(z,w) \text{ with } \Delta_{n1} \in H(\mathbf{D}) \text{ and } \Delta_{n1}(w,w) = 0.$$

Because $K_n$ does not vanish in $\mathbf{D}$ and $L_n/K_n \in \mathcal{P}$ by a property of J-unitary matrices, we can also say that the positive real function $\Omega$ is approximated by the positive real function $\Omega_n = L_n/K_n$ such that

$$\frac{\Omega - \Omega_n}{B_n} = h \in H(\mathbf{D}) \text{ with } h(w) = 0.$$

Taking a Cayley transform results in the interpolation property of the Schur function $S_0$ by the Schur function $\Gamma_n = (K_n - L_n)/(K_n + L_n) \in \mathcal{B}$:

$$\frac{S_0 - \Gamma_n}{B_n} = g \in H(\mathbf{D}) \text{ with } g(w) = 0.$$

Suppose $\mu$ is some positive measure of $\mathbf{T}$ and that we associate with it the positive real function $\Omega_\mu(z,w)$ like in Lemma 15.1. Suppose we start the Pick-Nevanlinna algorithm as described above with $\Omega = \Omega_\mu$. Since we just showed that then $\Omega_n$ will interpolate this starting $\Omega_\mu$ at the point set $A_n^w$, it follows from Theorem 15.3 that the inner product on $\mathcal{L}_n$ is the same for the measure $\mu$, and the measure $\nu$ defined by $d\nu(t) = [P(t,w)/|K_n(t,w)|^2]d\lambda(t)$. Thus, as far as the inner product in $\mathcal{L}_n$ is concerned, $\nu$ is not dependent on $w$ (see Corollary 15.4).

$$\begin{aligned} \int f(t)g_*(t)d\mu(t) &= \int f(t)g_*(t)d\nu(t) \\ &= \int f(t)g_*(t)\frac{P(t,w)}{|K_n(t,w)|^2}d\lambda(t) \\ &= \int f(t)g_*(t)\frac{P(t,0)}{|K_n(t,0)|^2}d\lambda(t). \end{aligned}$$

This has the important consequence that $K_n(z,w)$ as generated by the Pick-Nevanlinna algorithm applied to the $\Omega_\mu$, is the normalized reproducing kernel for $\mathcal{L}_n$ w.r.t. the measure $\mu$. Note that the real part of $\Omega_\mu$ has a radial limit satisfying a.e. $\Re\,\Omega_\mu(z,w) = \mu'(z)/P(z,w)$, $z \in \mathbf{T}$, $w \in \mathbf{D}$. To see that up to normalization, $K_n$ reproduces every $f \in \mathcal{L}_n$ note that

$$\begin{aligned}
\langle f, K_n(\cdot, z)\rangle_\mu &= \langle f, K_n(\cdot, z)\rangle_\nu \\
&= \int f(t) K_{n*}(t,z) \frac{P(t,z)}{K_{n*}(t,z)K_n(t,z)} d\lambda(t) \\
&= \int \frac{f(t)}{K_n(t,z)} P(t,z) d\lambda(t) = \frac{f(z)}{K_n(z,z)}
\end{aligned}$$

because $f/K_n$ is analytic in $\mathbf{D}$ so that the Poisson formula holds.

## 19 Interpolation algorithm for the orthonormal functions

We shall try in this section to give an algorithm in the style of the Pick-Nevanlinna algorithm, which will , based on the idea of successive interpolation generate the recursion for the orthonormal functions $\phi_n$ and the functions of the second kind $\psi_n$. As a matter of fact, it is difficult to do this for these functions because of the rotating factors $\eta_n^1$ and $\eta_n^2$ 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 $\Phi_n$ and $\Psi_n$ satisfied a recurrence that got rid of these $\eta$'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 $\epsilon_n$ as defined by (13.7). Define $R_{n1}$ and $R_{n2}$ through

$$\begin{bmatrix} B_{n-1}R_{n1}(z) \\ B_nR_{n2}(z) \end{bmatrix} = \begin{bmatrix} \Phi_n(z) \\ \Phi_n^*(z) \end{bmatrix}\Omega(z) + \begin{bmatrix} \Psi_n(z) \\ -\Psi_n^*(z) \end{bmatrix}. \tag{19.1}$$

Where $\Omega = \Omega_\mu$ is the positive real function associated with the measure $\mu$ 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 (14.5) and (14.6) 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 (19.1) of the remainder functions satisfies the recurrence for the rotated functions as in Theorem 13.3. Hence, also the left hand side shall satisfy

$$\begin{bmatrix} B_{n-1}R_{n1}(z) \\ B_nR_{n2}(z) \end{bmatrix} = e_n\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_nz}\begin{bmatrix} 1 & \overline{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix}\begin{bmatrix} Z_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} B_{n-2}R_{n-1,1}(z) \\ B_{n-1}R_{n-1,2}(z) \end{bmatrix}. \tag{19.2}$$

This can be rewritten as given in the next Theorem.

**Theorem 19.1** *The remainder functions as defined above satisfy the following recursion*

$$(1-\overline{\alpha}_nz)\begin{bmatrix} R_{n1}(z) \\ R_{n2}(z) \end{bmatrix} = e_n\begin{bmatrix} 1 & 0 \\ 0 & 1/\zeta_n(z) \end{bmatrix}\begin{bmatrix} 1 & \overline{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix}(1-\overline{\alpha}_{n-1}z)\begin{bmatrix} \eta_{n-1}R_{n-1,1}(z) \\ R_{n-1,2}(z) \end{bmatrix} \tag{19.3}$$

*with* $e_n > 0$ *and*

$$\Lambda_n = -\overline{\eta}_{n-1}\lim_{z\to\alpha_n}\frac{R_{n-1,2}(z)}{R_{n-1,1}(z)}, \quad \eta_{n-1} = \frac{\alpha_n}{|\alpha_n|}\frac{\overline{\alpha}_{n-1}}{|\alpha_{n-1}|} \text{ and } e_n^2 = \frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2}\frac{1}{1-|\Lambda_n|^2}. \tag{19.4}$$

*The $\Lambda_n$ in the previous expression are the same as the $\Lambda_n$ of Theorem 13.3.*

*We can make the recursion even simpler and avoid the explicit use of the $\eta_{n-1}$ by introducing*

$$r_{n1}(z) = \frac{\overline{\alpha}_n}{|\alpha_n|}R_{n1}(z) \text{ and } r_{n2}(z) = R_{n2}(z). \tag{19.5}$$

*With this notation, the recursion (19.3) becomes*

$$(1-\overline{\alpha}_nz)\begin{bmatrix} r_{n1}(z) \\ r_{n2}(z) \end{bmatrix} = e_n\begin{bmatrix} 1 & 0 \\ 0 & 1/\zeta_n(z) \end{bmatrix}\begin{bmatrix} 1 & \overline{L}_n \\ L_n & 1 \end{bmatrix}(1-\overline{\alpha}_{n-1}z)\begin{bmatrix} r_{n-1,1}(z) \\ r_{n-1,2}(z) \end{bmatrix} \tag{19.6}$$

*with*

$$L_n = \frac{\alpha_n}{|\alpha_n|}\Lambda_n = -\lim_{z\to\alpha_n}\frac{r_{n-1,2}(z)}{r_{n-1,2}(z)}, \qquad e_n = \left[\frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2}\frac{1}{1-|L_n|^2}\right]^{1/2}. \tag{19.7}$$

**Proof.** We shall only prove (19.4), because (19.7) is a direct consequence. We can start from the relation (19.2) and use $\zeta_{n-1} = \eta_{n-1} Z_{n-1}$ to get

$$B_{n-1}(z) \begin{bmatrix} R_{n1}(z) \\ \zeta_n R_{n2}(z) \end{bmatrix} = e_n \frac{1-\overline{\alpha}_{n-1} z}{1-\overline{\alpha}_n z} \begin{bmatrix} 1 & \overline{\Lambda}_n \\ \Lambda_n & 1 \end{bmatrix} B_{n-1}(z) \begin{bmatrix} \eta_{n-1} R_{n-1,1}(z) \\ R_{n-1,2}(z) \end{bmatrix} \tag{19.8}$$

which now easily gives (19.3). To find the expression for $\Lambda_n$, you can use the last line of (19.8) 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 $\Lambda_n$ follows. The expression for $e_n$ was shown in Theorem 13.2. □

The previous Theorem has the following consequence.

**Corollary 19.2** *Define the function $\Gamma_n(z)$ in terms of the remainder functions by*

$$\Gamma_n(z) = \frac{\alpha_n}{|\alpha_n|} \frac{R_{n2}(z)}{R_{n1}(z)} = \frac{r_{n2}(z)}{r_{n1}(z)}. \tag{19.9}$$

*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*

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

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

**Proof.** This follows immediately from the previous Theorem. All the $\Gamma_k$ are in $\mathcal{B}$ because $\Gamma_0$ is, while the Möbius transforms are done with $L_k \in \mathbf{D}$. Moreover the division by $\zeta_n$ respects the analyticity because the function between brackets was made zero in $z = \alpha_n$ by the choice of $L_n$. □

# 20 Continued fractions and three term recurrence

We shall now give some continued fractions which can be given in association with the recurrence relations we have introduced so far. Therefore we borrow a result from [14, p. 19-21].

**Lemma 20.1** *Let us define*

$$t_k = \begin{bmatrix} c_k & d_k \\ a_k & b_k \end{bmatrix}, k \geq 0 \text{ and } T_n = \begin{bmatrix} C_n & D_n \\ A_n & B_n \end{bmatrix}$$

*with $T_0 = t_0$ and $T_n = t_n T_{n-1}$ for $n > 0$. Suppose $d_0 c_k \neq 0$ for all $k > 0$. Furthermore define $R_n = FB_n - A_n$ and $S_n = FD_n - C_n$. Then the formal continued fraction expansion*

$$\begin{aligned} F &= \frac{c_0}{d_0} + \frac{(a_0 d_0 - b_0 c_0) d_0^{-2} \rfloor}{\lceil b_0 d_0^{-1}} + \frac{c_1 \rfloor}{\lceil d_1} + \frac{(b_1 c_1 - a_1 d_1) c_1^{-1} \rfloor}{\lceil a_1 c_1^{-1}} + \cdots + \\ & \quad \frac{c_n \rfloor}{\lceil d_n} + \frac{(b_n c_n - a_n d_n) c_n^{-1} \rfloor}{\lceil a_n c_n^{-1}} + \begin{cases} \dfrac{c_{n+1} \rfloor}{\lceil d_{n+1}} - \dfrac{S_{n+1} \rfloor}{\lceil R_n} \\ -\dfrac{R_n \rfloor}{\lceil S_n} \end{cases} \end{aligned} \tag{20.1}$$

*holds and the successive convergents are*

$$\frac{C_0}{D_0}, \frac{A_0}{B_0}, \frac{C_1}{D_1}, \frac{A_1}{B_1}, \ldots, \frac{C_n}{D_n}, \frac{A_n}{B_n}, \ldots$$

**Proof.** This is a matter of simple algebra. See e.g. [14, Property 2.9 and Property 2.5]. □

This can now be applied to the situation of a recursion like the one for the orthonormal functions. Note that $t_n$ there looks like

$$t_n = \begin{bmatrix} c_k & d_k \\ a_k & b_k \end{bmatrix} = e_k \frac{1 - \overline{\alpha}_{k-1} z}{1 - \overline{\alpha}_k z} \begin{bmatrix} \eta_k^1 \zeta_{k-1} & \eta_k^1 \overline{\lambda}_k \\ \eta_k^2 \lambda_k \zeta_{k-1} & \eta_k^2 \end{bmatrix}.$$

A general term therefore is of the form

$$\frac{c_k \rfloor}{\lceil d_k} + \frac{(b_k c_k - a_k d_k) c_k^{-1} \rfloor}{\lceil a_k c_k^{-1}} = \frac{\zeta_{k-1} \rfloor}{\lceil \overline{\lambda}_k} + \frac{\overline{\eta}_k^1 \eta_k^2 (1 - |\lambda_k|^2) \rfloor}{\lceil \overline{\eta}_k^1 \eta_k^2 \lambda_k}.$$

We can now try several initial conditions and get different convergents. For example, to get the convergents

$$\frac{\psi_0}{\phi_0}, -\frac{\psi_0^*}{\phi_0^*}, \ldots, \frac{\psi_n}{\phi_n}, -\frac{\psi_n^*}{\phi_n^*}, \ldots$$

you need the initial conditions

$$\begin{bmatrix} c_0 & d_0 \\ a_0 & b_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

If you choose $F$ of Lemma 20.1 to be $-\Omega$, then the remainders $R_n$ and $S_n$ can be recovered from (19.1) which can be written as

$$\begin{bmatrix} -\overline{\epsilon}_n B_{n-1} R_{n1}(z) \\ -\epsilon_n B_n R_{n2}(z) \end{bmatrix} = \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} (-\Omega(z)) - \begin{bmatrix} \psi_n(z) \\ -\psi_n^*(z) \end{bmatrix}.$$

The first row gives $S_n$, the second gives $R_n$.

All this put together gives the expansion

$$
\begin{aligned}
-\Omega \;&=\; 1 - \frac{2|}{|1} + \sum_{k=1}^{n}\left(\frac{\zeta_{k-1}|}{|\overline{\lambda}_k} + \frac{\overline{\eta}_k^1\eta_k^2(1-|\lambda_k|^2)|}{|\overline{\eta}_k^1\eta_k^2\lambda_k}\right) + \\
&\left\{
\begin{array}{l}
\dfrac{\zeta_n|}{|\overline{\lambda}_{n+1}} - \dfrac{\overline{\epsilon}_{n+1}(1-\overline{\alpha}_{n+1}z)R_{n+1,1}|}{|e_{n+1}\epsilon_n(1-\overline{\alpha}_n z)\eta_{n+1}^1 R_{n,2}} \\[2ex]
-\dfrac{\epsilon_n\zeta_n R_{n2}|}{|\overline{\epsilon}_n R_{n1}}
\end{array}
\right.
\end{aligned}
\tag{20.2}
$$

From this, you can derive that

$$
\frac{\phi_n - \psi_n}{\phi_n + \psi_n}, \frac{\phi_n^* - \psi_n^*}{\phi_n^* + \psi_n^*} \text{ for } n \geq 1
$$

are the successive convergents of the previous continued fraction without the two initial terms i.e.,

$$
\sum_{k\geq 1}\left(\frac{\zeta_{k-1}|}{|\overline{\lambda}_k} + \frac{\overline{\eta}_k^1\eta_k^2(1-|\lambda_k|^2)|}{|\overline{\eta}_k^1\eta_k^2\lambda_k}\right).
$$

By interchanging rows and columns in the matrices of the recurrence, we get the same convergents in the other order. For example, the convergents of the continued fraction expansion

$$
\begin{aligned}
\Omega \;&=\; 1 - \frac{2|}{|1} + \sum_{k=1}^{n}\left(\frac{1|}{|\zeta_{k-1}\lambda_k} + \frac{\overline{\eta}_k^2\eta_k^1(1-|\lambda_k|^2)\zeta_{k-1}|}{|\overline{\eta}_k^2\eta_k^1\overline{\lambda}_k}\right) + \\
&\left\{
\begin{array}{l}
\dfrac{1|}{|\zeta_n\lambda_{n+1}} - \dfrac{\epsilon_{n+1}B_{n+1}(1-\overline{\alpha}_{n+1}z)R_{n+1,2}|}{|e_{n+1}\eta_{n+1}^2\overline{\epsilon}_n B_{n-1}(1-\overline{\alpha}_n z)R_{n1}} \\[2ex]
-\dfrac{\overline{\epsilon}_n B_{n-1}R_{n1}|}{|\epsilon_n B_n R_{n2}}
\end{array}
\right.
\end{aligned}
\tag{20.3}
$$

are now

$$
\frac{\psi_0^*}{\phi_0^*}, -\frac{\psi_0}{\phi_0}, \ldots, \frac{\psi_n^*}{\phi_n^*}, -\frac{\psi_n}{\phi_n}, \ldots
$$

Taking contractions of these continued fractions will give you the even or odd parts and these give genuine three-term recurrence relations. By definition, the even contraction of the continued fraction

$$
\frac{a_0}{b_0} + \frac{a_1|}{|b_1} + \frac{a_2|}{|b_2} + \cdots \tag{20.4}
$$

is the continued fraction whose convergents are the even convergents of the contracted one. We have the following expression for an even contraction.

**Lemma 20.2** *The even contraction of the continued fraction (20.4) is given by*

$$
\frac{a_0}{b_0} + \frac{a_1 b_2|}{|a_2 + b_1 b_2} + \sum_{k\geq 1}\frac{-a_{2k}b_{2k}^{-1}a_{2k+1}b_{2k+2}|}{|a_{2k+2} + (b_{2k+1} + b_{2k}^{-1}a_{2k+1})b_{2k+2}}
$$

**Proof.** This can be found e.g., in [14, Property 2.8]) □

For example the even convergents of (20.2) i.e., $\psi_k/\phi_k$ for $k = 0, 1, \ldots$ are the successive convergents of the following continued fraction expansion

$$
-\Omega = 1 - \frac{2\overline{\lambda}_1|}{|z + \overline{\lambda}_1} + \sum_{k\geq 1}\frac{-\zeta_{k-1}\eta_k^2\overline{\eta}_k^1(1-|\lambda_k|^2)\dfrac{\overline{\lambda}_{k+1}}{\overline{\lambda}_k}|}{|\zeta_k + \eta_k^2\overline{\eta}_k^1\dfrac{\overline{\lambda}_{k+1}}{\overline{\lambda}_k}}
$$

if all the $\lambda_k$ are different from 0 and 1. From which it follows that the orthonormal functions as well as the functions of the second kind satisfy the three term recurrence relation

$$\overline{\lambda}_k f_{k+1}(z) = (\zeta_k(z)\overline{\lambda}_k + \eta_k^2\overline{\eta}_k^1\overline{\lambda}_{k+1})f_k(z) - (\zeta_{k-1}(z)\eta_k^2\overline{\eta}_k^1(1-|\lambda_k|^2)\overline{\lambda}_{k+1})f_{k-1}(z). \tag{20.5}$$

With the initial conditions $f_0 = 1$, $f_1 = a_1\eta_1^1(z+\overline{\lambda}_1)/(1-\overline{\alpha}_1 z)$ you generate $f_k = \phi_k$ and with initial conditions $f_0 = 1$, $f_1 = a_1\eta_1^1(z-\overline{\lambda}_1)/(1-\overline{\alpha}_1 z)$ you generate $f_k = \psi_k$. For the even contraction of (20.3), we find

$$\Omega = 1 - \frac{2z\lambda_1 \rfloor}{\lceil 1 + z\lambda_1} + \sum_{k\geq 1} \frac{-\zeta_k\eta_k^1\overline{\eta}_k^2(1-|\lambda_k|^2)\dfrac{\lambda_{k+1}}{\lambda_k} \rfloor}{\lceil 1 + \zeta_k\eta_k^1\overline{\eta}_k^2\dfrac{\lambda_{k+1}}{\lambda_k}}$$

with convergents $\psi_k^*/\phi_k^*$ for $k = 0, 1, \ldots$

We can also use the rotated functions and thus avoid the $\eta$'s. The reader can translate the results easily for himself. The rule is to multiply $\phi_n$ and $\psi_n$ and $R_{n1}$ by $\epsilon_n$. For $\phi_n$ and $\psi_n$ this means replace them by $\Phi_n$ and $\Psi_n$ etc. In short, here are the translation rules:

$$\phi \to \Phi, \quad \psi \to \Psi, \quad \lambda \to \Lambda, \quad \zeta \to Z, \quad \eta \to 1, \quad \epsilon \to 1$$

As an example, the translation of (20.3) reads like

$$\begin{aligned}\Omega \;=\; & 1 - \frac{2 \rfloor}{\lceil 1} + \sum_{k=1}^{n}\left(\frac{1 \rfloor}{\lceil Z_{k-1}\Lambda_k} + \frac{(1-|\Lambda_k|^2)Z_{k-1} \rfloor}{\lceil \overline{\Lambda}_k}\right) + \\ & \left\{ \begin{array}{l} \dfrac{1 \rfloor}{\lceil Z_n\Lambda_{n+1}} - \dfrac{B_{n+1}(1-\overline{\alpha}_{n+1}z)R_{n+1,2} \rfloor}{\lceil e_{n+1}B_{n-1}(1-\overline{\alpha}_n z)R_{n1}} \\ -\dfrac{B_{n-1}R_{n1} \rfloor}{\lceil B_n R_{n2}} \end{array}\right.\end{aligned}$$

The convergents are

$$\frac{\Psi_0^*}{\Phi_0^*}, -\frac{\Psi_0}{\Phi_0}, \ldots, \frac{\Psi_n^*}{\Phi_n^*}, -\frac{\Psi_n}{\Phi_n}, \ldots$$

The three term recurrence for the rotated functions is

$$\overline{\Lambda}_k F_{k+1}(z) = (Z_k(z)\overline{\Lambda}_k + \overline{\Lambda}_{k+1})F_k(z) - (Z_{k-1}(z)(1-|\Lambda_k|^2)\overline{\Lambda}_{k+1})F_{k-1}(z). \tag{20.6}$$

It will be clear from our discussion in this section that it is also possible to get continued fractions whose convergents are the ratios of kernels $L_k(z,w)/K_k(z,w)$ etc. Likewise we can obtain three term recursions for these kernels. We leave this to the reader, it is a matter of simple algebra.

## 21 Favard Theorems

In sections 12 and 13, we have seen how the kernels, as well as the orthogonal functions satisfied certain recurrence relations which generalize the Szegő recurrence relations. It is thus true that all rational functions which are orthogonal (or reproducing kernels) with respect to a certain positive measure on the unit circle will satisfy such a recurrence. The converse of this theorem is known as a Favard Theorem, named after Favard's paper [35]. Such a Favard Theorem states that if functions satisfy recurrence relations as we have given in Section 13, then they will give orthogonal rational functions with respect to some positive measure on the unit circle; if we have a recurrence like in Section 12, then it will be reproducing kernels for some positive measure. A simple proof for the Szegő polynomials was recently given in [34]. There, not only the existence of the measure is proved, but the measure is actually constructed. We shall follow in this section a similar approach for the rational case. Related results in a somewhat different setting were obtained in [47, 16].

We shall give a sequence of lemmas which will eventually lead to the proof of the Favard Theorem. We start with the orthogonal functions.

**Lemma 21.1** *Suppose we are given two sequences of numbers $\alpha_k \in \mathbf{D}$ and $\lambda_k \in \mathbf{D}$ for k=1,2,... and set $\alpha_0 = 0$. Define the numbers $e_k > 0$ by their squares*

$$e_k^2 = \frac{1-|\alpha_k|^2}{1-|\alpha_{k-1}|^2}\frac{1}{1-|\lambda_k|^2} \text{ for } k = 1, 2, \ldots \tag{21.1}$$

*Finally define the functions $\phi_k$ by*

$$\phi_0 = 1, \quad \phi_k = e_k \eta_k^1 \frac{1-\overline{\alpha}_{k-1}z}{1-\overline{\alpha}_k z}[\zeta_{k-1}\phi_{k-1} + \overline{\lambda}_k \phi_{k-1}^*], \quad k = 1, 2, \ldots \tag{21.2}$$

*where the numbers $\eta_k^1 \in \mathbf{T}$ are chosen such that $\phi^*(\alpha_k) > 0$.*

*Then the functions $\phi_k^*$ satisfy the following recurrence*

$$\phi_0^* = 1, \quad \phi_k^* = e_k \eta_k^2 \frac{1-\overline{\alpha}_{k-1}z}{1-\overline{\alpha}_k z}[\zeta_{k-1}\lambda_k\phi_{k-1} + \phi_{k-1}^*], \quad k = 1, 2, \ldots \tag{21.3}$$

*with*

$$\eta_k^2 = \overline{\eta}_k^1 \frac{\alpha_{k-1}}{|\alpha_{k-1}|}\frac{\overline{\alpha}_k}{|\alpha_k|}. \tag{21.4}$$

*Moreover, $1/\phi_n^* \in H_2$.*

**Proof.** We can leave this proof to the reader because the first part comes down to simple calculus and the second part follows immediately from the J-unitarity of the recurrence that can be obtained by coupling the recurrences (21.2) and (21.3) in one matrix relation. The result follows as in Theorem 13.5. □

**Lemma 21.2** *Under the conditions of Lemma 21.1, it holds that*

$$\lambda_k = \eta\frac{\overline{\phi_k(\alpha_{k-1})}}{\phi_k^*(\alpha_{k-1})} \text{ with } \eta = \frac{\overline{\alpha}_k}{|\alpha_k|}\frac{\alpha_{k-1}}{|\alpha_{k-1}|}\frac{1-\alpha_k\overline{\alpha}_{k-1}}{1-\overline{\alpha}_k\alpha_{k-1}} \in \mathbf{T}. \tag{21.5}$$

**Proof.** From the recurrences for $\phi_k$ and $\phi_k^*$, we find

$$\overline{\phi_k(\alpha_{k-1})} = e_k\overline{\eta}_k^1\frac{1-|\alpha_{k-1}|^2}{1-\alpha_k\overline{\alpha}_{k-1}}[0 + \lambda_k\phi_{k-1}^*(\alpha_{k-1})] \tag{21.6}$$

and

$$\phi_k^*(\alpha_{k-1}) = e_k\eta_k^2\frac{1-|\alpha_{k-1}|^2}{1-\overline{\alpha}_k\alpha_{k-1}}[0 + \phi_{k-1}^*(\alpha_{k-1})]. \tag{21.7}$$

Dividing (21.6) by (21.7) gives

$$\frac{\overline{\phi_k(\alpha_{k-1})}}{\phi_k(\alpha_{k-1})} = \frac{\overline{\eta}_k^1}{\eta_k^2}\frac{1-\overline{\alpha}_k\alpha_{k-1}}{1-\alpha_k\overline{\alpha}_{k-1}}\lambda_k = \frac{\overline{\alpha}_{k-1}}{|\alpha_{k-1}|}\frac{\alpha_k}{|\alpha_k|}\frac{1-\overline{\alpha}_k\alpha_{k-1}}{1-\alpha_k\overline{\alpha}_{k-1}}\lambda_k = \overline{\eta}\lambda_k.$$

The result now follows. □

**Lemma 21.3** *Define the positive measure $\mu_n$ on* $\mathbf{T}$ *by*

$$d\mu_n(z) = \frac{P(z,\alpha_n)}{|\phi_n(z)|^2}d\lambda(z) \tag{21.8}$$

*with $P(z,w)$ the Poisson kernel. Then the functions constructed in Lemma 21.1 satisfy the orthogonality relations*

$$\langle \phi_k, \phi_\ell\rangle_{\mu_n} = \delta_{k\ell} \textit{ for } 0 \le k,\ell \le n. \tag{21.9}$$

**Proof.** We shall first prove that $\phi_n$ is orthonormal with respect to all its predecessors :

$$\langle \phi_n, \phi_m\rangle_{\mu_n} = \delta_{nm} \text{ for } 0 \le m \le n. \tag{21.10}$$

This is shown as follows

$$\begin{aligned}
\langle \phi_n, \phi_m\rangle_{\mu_n} &= \int \frac{\phi_n(z)\phi_{m*}(z)}{\phi_n(z)\phi_{n*}(z)}P(z,\alpha_n)d\lambda(z)\\
&= \int \frac{\phi_{m*}(z)}{\phi_{n*}(z)}P(z,\alpha_n)d\lambda(z)\\
&= \int B_{n\backslash m}(z)\frac{\phi_m^*(z)}{\phi_n^*(z)}P(z,\alpha_n)d\lambda(z)\\
&= B_{n\backslash m}(z)\frac{\phi_m^*(z)}{\phi_n^*(z)}\bigg|_{z=\alpha_n} = \delta_{nm}.
\end{aligned}$$

The general orthogonality follows from Theorem 13.6 which says that all the previous orthogonal functions are defined in terms of $\phi_n$, orthonormal to $\mathcal{L}_{n-1}$ by the inverse recurrence, and we have just proved by the previous Lemma 21.2 that the recurrence in the Lemma 21.1 is the same as the one from Theorem 13.1. □

We are now ready to prove the following Favard-type Theorem.

**Theorem 21.4** *There exists a unique Borel measure on* $\mathbf{T}$ *for which the $\phi_n$ as constructed in Lemma 21.1 are the orthonormal functions.*

**Proof.** Define

$$\mu_n(t) = \int_0^t \frac{P(e^{\mathbf{i}\theta},\alpha_n)}{|\phi_n(e^{\mathbf{i}\theta})|^2}d\lambda(\theta) = \int_0^t d\mu_n(\theta). \tag{21.11}$$

These are all increasing functions and uniformly bounded ($\int d\mu_n = 1$). Hence, there exists a subsequence such that

$$\lim_{k\to\infty}\mu_{n_k}(\theta) = \mu(\theta)$$

and

$$\lim_{k\to\infty}\int f(e^{\mathbf{i}\theta})d\mu_{n_k}(\theta) = \int f d\mu$$

for all functions $f$ continuous on $\mathbf{T}$. Thus the $\{\phi_n\}$ are an orthonormal system with respect to this measure $\mu$.

The uniqueness follows from the representation of bounded linear functionals (F. Riesz). □

We can now repeat a similar thing for the reproducing kernels.

**Lemma 21.5** *Let $0 = \alpha_0, \alpha_1, \alpha_2, \ldots$ be complex numbers in $\mathbf{D}$. Let $\rho, k = 1, 2, \ldots$ be given functions in $\mathcal{B}$ (i.e., $|\rho_k(w)| < 1$ for $w \in \mathbf{D}$). Define $\gamma_k(w) = -\zeta_k(w)\rho_k(w)$ for $k = 1, 2, \ldots$ (Hence $\gamma_k \in \mathcal{B}$.) Let the functions $K_n(z, w)$ be generated by (12.2), i.e.,*

$$\begin{bmatrix} K_n^*(z,w) \\ K_n(z,w) \end{bmatrix} = \theta_n(z,w) \begin{bmatrix} K_{n-1}^*(z,w) \\ K_{n-1}(z,w) \end{bmatrix}; \begin{bmatrix} K_0^*(z,w) \\ K_0(z,w) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{21.12}$$

*with the matrix $\theta_n$ given by*

$$\theta_n(z,w) = \frac{1}{\sqrt{1-|\rho_n|^2}} \begin{bmatrix} 1 & \overline{\rho}_n \\ \rho_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{1-|\gamma_n|^2}} \begin{bmatrix} 1 & \overline{\gamma}_n \\ \gamma_n & 1 \end{bmatrix}. \tag{21.13}$$

*Then the following properties hold.*

$$1/K_n(z,w) \in H_2 \text{ for } w \in \mathbf{D} \text{ fixed.} \tag{21.14}$$

$$K_n(w,w) = K_{n-1}(w,w)\sqrt{\frac{1-|\gamma_n(w)|^2}{1-|\rho_n(w)|^2}} = \prod_{k=1}^{n} \left( \frac{1-|\gamma_n(w)|^2}{1-|\rho_n(w)|^2} \right)^{1/2} > 0 \tag{21.15}$$

$$\rho_n(w) = \frac{K_{n-1}^*(\alpha_n, w)}{K_{n-1}(\alpha_n, w)}. \tag{21.16}$$

**Proof.** Again, (21.14) follows from Theorem 12.3 as for the orthogonal functions because the recurrence is J-unitary by definition.

The proof of (21.15) is pure calculus, just replace $z$ by $w$ and play a bit with the formulas.

Equality (21.16) is also immediate, since

$$\frac{K_n^*(\alpha_n, w)}{K_n(\alpha_n, w)} = \frac{\overline{\rho}_n[K_{n-1}(\alpha_n, w) + \gamma_n(w)K_{n-1}^*(\alpha_n, w)]}{[K_{n-1}(\alpha_n, w) + \gamma_n(w)K_{n-1}^*(\alpha_n, w)]} = \overline{\rho}_n$$

□

**Lemma 21.6** *Let the $K_n$ be generated as in the previous Lemma 21.5 and define*

$$k_m(z,w) = K_m(z,w)K_m(w,w) \text{ for } k = 0, 1, \ldots \tag{21.17}$$

*Then $k_m(z,w)$ is the reproducing kernel for $\mathcal{L}_m$ $(0 \le m \le n)$ with respect to the measure $\mu_n$ defined by*

$$d\mu_n(z) = \frac{P(z,w)}{|K_n(z,w)|^2} d\lambda(z) \tag{21.18}$$

*where $P(z,w)$ is the Poisson kernel*

**Proof.** We first prove that $k_n$ reproduces functions from $\mathcal{L}_n$. This is easily seen as follows. It holds for all $f \in \mathcal{L}_n$ that

$$\begin{aligned} \langle f, k_n(\cdot, w)\rangle_{\mu_n} &= \int f(z) k_{n*}(z,w) \frac{P(z,w)d\lambda(z)}{K_n(z,w)K_{n*}(z,w)} \\ &= K_n(w,w) \int f(z) \frac{k_{n*}(z,w)}{k_{n*}(z,w)} \frac{P(z,w)}{K_n(z,w)} d\lambda(z) \\ &= f(w) \frac{K_n(w,w)}{K_n(w,w)} = f(w). \end{aligned}$$

We show next the induction step which starts from the fact that $k_m$ reproduces all $f \in \mathcal{L}_m$ with respect to the measure $\mu_n$ and we shall prove that $k_{m-1}$ will then be reproducing with respect to the same measure for all functions in $\mathcal{L}_{m-1}$. That is we have to prove

$$\langle f, k_m(\cdot, w)\rangle_{\mu_n} = f(w), \forall f \in \mathcal{L}_m \Rightarrow \langle f, k_{m-1}(\cdot, w)\rangle_{\mu_n} = f(w), \forall f \in \mathcal{L}_{m-1} \tag{21.19}$$

We first derive the following expression from the recurrence which can be obtained with some patient calculations.

$$k_m(z,w)[\zeta_m(z)\overline{\gamma}_m+\overline{\rho}_m]=[\zeta_m(z)(1-|\gamma_m|^2)]k^*_{m-1}(z,w)+[\zeta_m(z)\overline{\gamma}_m\rho_m+1]k^*_m(z,w)$$

which becomes after dividing by $B_m$

$$\frac{k_m(z,w)}{B_m(z)}[\zeta_m(z)\overline{\gamma}_m+\overline{\rho}_m]=(1-|\gamma_m|^2)k_{m-1*}(z,w)+[\zeta_m(z)\overline{\gamma}_m\rho_m+1]k_{m*}(z,w).$$

Now we use this in the right hand side of (21.19) to get

$$\begin{aligned}\langle f,k_{m-1}(\cdot,w)\rangle_{\mu_n} &= \frac{-1}{1-|\gamma_m|^2}\langle f,[k_{m*}(\zeta_m\overline{\gamma}_m\rho_m+1)]_*\rangle_{\mu_n}+\frac{1}{1-|\gamma_m|^2}\langle f,[\frac{k_m}{B_m}(\zeta_m\overline{\gamma}_m+\overline{\rho}_m]_*\rangle_{\mu_n}\\ &= \frac{-1}{1-|\gamma_m|^2}\langle f(\zeta_m\overline{\gamma}_m\rho_m+1),k_m\rangle_{\mu_n}+\frac{1}{1-|\gamma_m|^2}\langle f(\zeta_m\overline{\gamma}_m+\overline{\rho}_m,k^*_m\rangle_{\mu_n}.\end{aligned}$$

The first term can be simplified because $f(\zeta_m\overline{\gamma}_m\rho_m+1)\in\mathcal{L}_m$ and $k_m$ reproduces. Therefore the inner product in the first term equals

$$f(w)(\zeta_m(w)\overline{\gamma}_m(w)\rho_m(w)+1)=-(1-|\gamma_m(w)|^2).$$

The inner product in the second term equals

$$\langle f(\zeta_m\overline{\gamma}_m+\overline{\rho}_m,k^*_m\rangle_{\mu_n}=\langle k_m,f^*(\gamma_m+\zeta_m\rho_m)\rangle_{\mu_n}.$$

The latter has a factor $f^*(\gamma_m+\zeta_m\rho_m)$ which is again in $\mathcal{L}_m$ and we can use again the reproducing property of $k_m$ to find that that inner product equals the complex conjugate of

$$f^*(w)(\gamma_m(w)+\zeta_m(w)\rho_m(w))$$

which is zero by the definition of $\gamma_m$. Thus, after filling in the last two results for the inner products, we get

$$\langle f,k_{m-1}(\cdot,w)\rangle_{\mu_n}=f(w)+0=f(w)$$

which proves the induction step. □

We can now prove the Favard Theorem for the kernels.

**Theorem 21.7** *There exists a unique Borel measure on the unit circle for which the $K_n(z,w)$ as defined in Lemma 21.5 are the normalized reproducing kernels of the spaces $\mathcal{L}_n$.*

**Proof.** From the foregoing Lemmas, it follows that the $K_n$ are the normalized reproducing kernels with respect to the measure $\mu_n$ as defined in (21.18). By Theorem 17.1, it follows that in $\mathcal{L}_n$ we have equality of the inner products

$$\langle\cdot,\cdot\rangle_{\mu_n}=\langle\cdot,\cdot\rangle_{\nu_n}$$

with the measure $\nu_n$ defined by

$$d\nu_n(z)=\frac{d\lambda(z)}{|K_n(z,0)|^2}. \tag{21.20}$$

Thus the $K_m$ are also the normalized reproducing kernels for $\mathcal{L}_m$, $0\le m\le n$ with respect to the measure $\nu_n$. We can now repeat the proof of Theorem 21.4 for the orthogonal functions. We leave this to the reader. □

## 22 Generalization of the Szegő problem

In this section and the next one we bring together some convergence results. We start with a resume of the treatment given by Grenander and Szegő, [42, chapter 3] for the polynomial case.

The original Szegő problem was to solve the extremal problem, which in the notation of Section 8 can be denoted as $P^2_\infty(1,0)$. That is : find the solution of

$$P^2_\infty(1,0) \quad : \qquad \inf\{\|f\|^2_\mu : f(0)=1, f\in\Pi_\infty\} = \kappa_\infty^{-2}.$$

$\Pi_\infty$ represents the set of all possible polynomials. Since the previous problem is set in $L_2(\mu)$, we could as well replace $\Pi_\infty$ by $\Pi$, which is the $L_2(\mu)$ closure of the polynomials. In other words, $\Pi$ is just another notation for $H_2(\mu)$. The problem was solved by successively considering $f\in\Pi_n$ for $n=0,1,2,\ldots$. An immediate generalization to the rational case forces us to consider the problem $P^2_n(1,\alpha_n)$, treated in Theorem 8.2 which was to find $\inf\|f\|^2_\mu$ when $f$ is in $\mathcal{L}_n$ and $f(\alpha_n)=1$. The minimum is reached for $f=\kappa_n^{-1}\phi_n^*$ and the minimum is $\kappa_n^{-2}$. You recover the polynomial results by setting all $\alpha_k$ equal to zero. The $\kappa_n$ are then the positive leading coefficients in the orthonormal Szegő polynomials $\phi_n(z)$. If we take this process to the limit when $n\to\infty$, we shall have solved the Szegő problem for a general $f\in L_2(\mu)$ if $\Pi_\infty$ is dense in $L_2(\mu)$. So the intimately related problem is to find conditions for the latter to happen. The answer is that $\Pi_\infty$ is dense in $L_2(\mu)$ if and only if the infimum $\kappa_\infty^{-2}=0$ which happens if and only if $\int\log\mu' d\lambda=-\infty$ [42, p.49-50]. In general, the limiting value of $\kappa_n^{-2}$ as $n\to\infty$ is given by $|\sigma(0)|^2=\exp\{\int\log\mu' d\lambda\}$. This is equal to the geometric mean of $\mu'$. If $\log\mu'\notin L_1$, then this has to be replaced by 0. This was proved e.g., in [42, p.44]. See also [36, p.200 ff]. As a generalization, Grenander and Szegő consider next the problem $P^2_n(1,w)$ with $w\in\mathbf{D}$ again for the polynomial case. We know that the infimum is reached for $f=k_n(z,w)[k_n(w,w)]^{-1}$ and that the minimum is found to be $[k_n(w,w)]^{-1}$. The latter function in known as the *Christoffel function.* For $w=0$ we rediscover the previous result in the case of polynomials. In [42, section 3.2] it is shown that this minimum converges to $(1-|w|^2)|\sigma(w)|^2$. We shall generalize the last approach to the rational case. However in the rational case, it requires putting $w=\alpha_n$ to rediscover the former problem. This is an extra complication since this $w$ is supposed to be in $\mathbf{D}$, while when replacing it by $\alpha_n$, it depends upon $n$ and when $n$ tends to $\infty$ this may approach the unit circle $\mathbf{T}$ or may not converge at all. We shall also have to consider $\mathcal{L}_\infty$ and $\mathcal{L}$. In analogy with the polynomials, the latter is again supposed to denote the closure in $L_2(\mu)$ of $\mathcal{L}_\infty$. We shall have solved the same problem as Szegő did if $\mathcal{L}=\Pi\equiv H_2(\mu)$. We have seen in section 9 that this happens when the sum $\sum(1-|\alpha_i|)$ diverges.

In Dewilde and Dym [24] we find most of the other results of this section proved for an absolutely continuous measure $d\mu=\mu' d\lambda$. We now generalize this to the case of a general measure. We follow closely the development in [24]. In [42, p. 51] we find that for polynomials the limit function $\lim_{n\to\infty}k_n(z,w)$ equals the *Szegő kernel*

$$s_w(z)=s(z,w)=[(1-\overline{w}z)\sigma(z)\overline{\sigma(w)}]^{-1}. \tag{22.1}$$

This is the function appearing in Theorem 8.6. Note that it satisfies $s_w(w)>0$. For the next development, it turns out to be preferable to define a normalized form : $\lim_{n\to\infty}K_n(z,w)$ where $K_n$ is the normalized form of $k_n$. We shall prove that it converges to the following function

$$S_w(z)=S(z,w)=\frac{\sqrt{1-|w|^2}}{1-\overline{w}z}\frac{1}{\sigma(z)} \tag{22.2}$$

if its arbitrary factor of modulus 1 is chosen to satisfy $\sigma(w)>0$. This function is different from the one in Theorem 8.6. In this notation the Szegő kernel can be written as $s_w(z)=S_w(w)S_w(z)$ and conversely $S_w(z)=s_w(z)/\sqrt{s_w(w)}$ as you can easily check. Another interesting observation to make is that $1/S_w(z)$ is readily found to be the outer spectral factor of $\mu_w$ where $d\mu_w(z)=d\mu(z)/P(z,w)$.

With the notation just introduced, Theorem 8.6, which implied that

$$\|k_n(z,w)-s_w(z)\|^2_\mu=s_w(w)-k_n(w,w)$$

which can be rephrased as

$$\|k_n(z,w) - S_w(z)S_w(w)\|_\mu^2 = |S_w(w)|^2 - k_n(w,w). \tag{22.3}$$

We shall denote by $P_n$ the orthogonal projection operator in $L_2(\mu)$ onto $\mathcal{L}_n$. Theorem 8.6 then says that $P_n[s_w(z)] = S_w(w)P_n[S_w(z)] = k_n(z,w)$ and that the squared norm of the error is given by the previous expression. As everywhere in this paper we assume that the Szegő condition $\log \mu' \in L_1$ is satisfied. Now we formulate our first convergence result :

**Theorem 22.1** *Let $k_n(z,w)$ be the kernel for $\mathcal{L}_n$ and $s_w(z)$ the Szegő kernel as defined in (22.1). Let $P_n$ denote the projection operator in $L_2(\mu)$ onto $\mathcal{L}_n$. Then*

$$\|k_n(z,w) - P_\infty s_w(z)\|_\mu^2 \to 0 \text{ as } n \to \infty. \tag{22.4}$$

*If moreover $\sum(1-|\alpha_k|) = \infty$, then $P_\infty s_w = s_w$. In that case of course*

$$\|k_n(z,w) - s_w(z)\|_\mu^2 \to 0 \text{ as } n \to \infty. \tag{22.5}$$

*and we also have*

$$k_n(w,w) \to s_w(w) = [(1-|w|^2)|\sigma(w)|^2]^{-1} \tag{22.6}$$

*uniformly on compact subsets of* $\mathbf{D}$.

**Proof.** By Theorem 8.6 we have $k_n(z,w) = P_n s_w(z)$. Hence

$$\|k_n(\cdot,w) - P_\infty s_w\|_\mu^2 = \|P_n s_w - P_\infty s_w\|_\mu^2$$

and the latter tends to zero by definition.

The relation $P_\infty s_w = s_w$ holds if $s_w \in \mathcal{L}$. Since we assumed that $\sum(1-|\alpha_k|) = \infty$, which implies by the previous Lemma that $\mathcal{L} = H_2(\mu)$, it is sufficient to prove that $s_w \in H_2(\mu)$. It is a well known property that a function $f$ will be in $H_2(\mu)$ if and only if $f\sigma \in H_2$ (see [36, Theorem 3.4, p. 215]). We now have

$$\begin{aligned} \|s_w\sigma\|_2^2 &= \frac{1}{|\sigma(w)|^2(1-|w|^2)} \int P(z,w)d\lambda(z) \\ &= \frac{1}{|\sigma(w)|^2(1-|w|^2)} \le \infty \end{aligned}$$

and this implies $s_w \in H_2(\mu)$ because $s_w$ is analytic in $\mathbf{D}$. Since the squared norm of the error equals $s_w(w) - k_n(w,w)$, which has to go to zero, we get the uniform convergence in compact subsets of $\mathbf{D}$ as stated. □

Note that (22.6) gives in fact a solution of the generalized Szegő problem in the rational case. It says that

$$\inf \|f\|_\mu = (1-|w|^2)|\sigma(w)|^2$$

if the infimum is taken over all $f \in \mathcal{L}$ with $f(w) = 1$.

For the normalized kernel we have the following.

**Lemma 22.2** *Let $S_w$ be the normalized Szegő kernel defined in (22.2) and $K_n(z,w)$ the normalized reproducing kernel for $\mathcal{L}_n$. Then*

$$\|K_n(z,w) - S_w(z)\|_\mu^2 = 2\Re[1 - K_n(w,w)/S_w(w)]. \tag{22.7}$$

**Proof.** We evaluate the norm in the left hand side. From the reproducing property of the kernels we readily find

$$\|k_n(z,w)\|_\mu^2 = k_n(w,w).$$

Since $K_n(z,w) = k_n(z,w)/\sqrt{k_n(w,w)}$, we find that $\|K_n(z,w)\|_\mu^2 = 1$.

The norm of the second term is evaluated as follows

$$\begin{aligned}
\|S_w\|_\mu^2 &= \int \frac{1-|w|^2}{|1-\overline{w}z|^2|\sigma(z)|^2} d\mu(z) \\
&= \int \frac{(1-|w|^2)|\sigma(z)|^2}{|1-\overline{w}z|^2|\sigma(z)|^2} d\lambda(z) \\
&= \int P(z,w) d\lambda(z).
\end{aligned}$$

The second equality is because $1/\sigma$ vanishes $d\mu_s$-a.e. This is also used in the next evaluation to give the cross product term.

$$\begin{aligned}
\Re\langle K_n(z,w), S_w(z)\rangle_\mu &= \sqrt{1-|w|^2}\Re \int \frac{z\overline{\sigma(z)}\sigma(z)}{(z-w)\overline{\sigma(z)}} K_n(z,w) d\lambda(z) \\
&= \sqrt{1-|w|^2}\Re\, \sigma(w) K_n(w,w).
\end{aligned}$$

Hence the assertion is proved since $1/S_w(w) = \sqrt{1-|w|^2}\sigma(w)$. □

**Theorem 22.3** *Let $S_w(z)$ be the normalized Szegő kernel satisfying $S_w(w) > 0$ and $K_n(z,w)$ the normalized reproducing kernel for $\mathcal{L}_n$. Suppose also that $\sum(1-|\alpha_k|) = \infty$, i.e., that $\mathcal{L} = H_2(\mu)$. Then*

$$\|K_n(z,w) - S_w(z)\|_\mu^2 \to 0 \text{ as } n \to \infty \tag{22.8}$$

*and also*

$$\left\|1 - \frac{K_n(z,w)}{S_w(z)}\right\|_{\lambda_w}^2 \to 0 \text{ as } n \to \infty \tag{22.9}$$

*and*

$$\left\|1 - \frac{k_n(z,w)}{s_w(z)}\right\|_{\lambda_w}^2 \to 0 \text{ as } n \to \infty \tag{22.10}$$

*where $d\lambda_w(z) = P(z,w)d\lambda(z)$.*

**Proof.** From the previous Theorem, we easily derive, by taking the square roots that $K_n(w,w) \to S_w(w)$. The first result now follows easily from the previous Lemma.

To prove the second one, note that

$$\left\|1 - \frac{K_n(z,w)}{S_w(z)}\right\|_{\lambda_w}^2 = \left\|1 - \frac{K_n(z,w)}{S_w(z)}\right\|_{\nu}^2$$

with

$$d\nu(z) = \frac{P(z,w)}{|\sigma(z)|^2} d\mu_a(z)$$

and this in turn can be bounded as

$$\begin{aligned}
\left\|1 - \frac{K_n(z,w)}{S_w(z)}\right\|_{\nu}^2 &= \|S_w(z) - K_n(z,w)\|_{\mu_a}^2 \\
&\leq \|S_w(z) - K_n(z,w)\|_{\mu}^2
\end{aligned}$$

Since the latter converges to zero, the assertion (22.9) follows.

The last relation is shown similarly. □

The last result (22.9) is a generalization of Theorem 5.8 of [39].

Two functions equal in $L_2(\mu)$ can only differ in $\mathbf{T}$ on a set of $\mu$-measure 0, a fortiori with $\mu_a$-measure 0 where $\mu_a$ is the absolute continuous part of $\mu$. Because we supposed $\log \mu' \in L_1$, $\mu'$ can only be zero on a set of Lebesgue measure zero. Thus the two functions can only differ on a set of Lebesgue measure 0. In other words convergence in $L_2(\mu)$ implies convergence in $L_2$. Since $H_p$ functions are defined (by the Poisson integral) in terms of the boundary values if $p \geq 1$, it follows that for these functions least squares-$\mu$ convergence implies uniform convergence in compact subsets of the open unit disk.

Hence the previous results imply the following.

**Theorem 22.4** *Suppose that $\sum(1 - |\alpha_k|) = \infty$. Then, with the usual notation, the following convergence results hold uniformly for $z$ in compact subsets of $\mathbf{D}$ as $n$ tends to $\infty$.*

$$k_n(z,w) \to s_w(z) \quad w \in \mathbf{D}, \tag{22.11}$$

$$K_n(z,w) \to S_w(z) \quad w \in \mathbf{D}, \tag{22.12}$$

$$\phi_n(z) \to 0. \tag{22.13}$$

**Proof.** The first two results follow by the observations made just before this Theorem. The third one follows from the convergence of $k_n(w,w) = \sum |\phi_k(w)|^2$ to a finite limit. □

**Theorem 22.5** *Let $\kappa_n$ be the leading coefficient of the orthonormal function $\phi_n$ and $\sigma(z)$ the spectral factor of $\mu$. Suppose $\alpha_n \to 0$ for $n \to \infty$, then we have*

$$\kappa_n \to [\sigma(0)]^{-1} \text{ if } \sigma(0) > 0, \tag{22.14}$$

*and*

$$\phi_n^*(z) \to [\sigma(z)]^{-1} \text{ uniformly on compact subsets of } \mathbf{D}. \tag{22.15}$$

*The last one is equivalent with*

$$\frac{\phi_n(z)}{B_n(z)} \to [\sigma_*(z)]^{-1} \text{ uniformly on compact subsets of } \mathbf{E}. \tag{22.16}$$

**Proof.** All this follows easily from the previous results if we use $K_n(z, \alpha_n) = \phi_n$ and fill in $w = 0$. □

The last two theorems generalize a theorem given by Grenander and Szegő [42, p.51].

# 23 Further convergence results and asymptotic behaviour

This section includes a number of convergence results like uniformly on compacts 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., [79, p.281 ff]. This can be used to obtain some other convergence results of the same type.

**Theorem 23.1** *Let $\phi_n$ be the orthonormal functions for $\mathcal{L}_n$ and $\psi_n$ the functions of the second kind. Define $\Omega_n = \psi_n^*/\phi_n^* \in \mathcal{P}$ and let the positive real function $\Omega$ be associated with the measure $\mu$ as in (2.3) ($c = 0$). Then $\Omega_n$ converges to $\Omega$ 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 $\Phi_n$ and $\Psi_n$ are the rotated functions as in section 19.

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 (13.8). Then

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

hence $\Phi_n^*[1-\Omega_n \quad 1+\Omega_n] = [0 \;\; 1]T_n$ and thus

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

Thus

$$\Omega - \Omega_n = \frac{1}{2}B_n R_{n2}/\Phi_n^* \quad \text{in } \mathbf{D}. \tag{23.1}$$

Now define the Schur functions by Cayley transforms of $\Omega$ and $\Omega_n$

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

Then,

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

which, in view of (23.1) 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 $|\Gamma| \le 1$ and $|\Gamma_n| \le 1$, so that

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

The maximum modulus theorem then gives

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

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

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

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

Practically the same proof can be repeated for the following Theorem.

**Theorem 23.2** *Let $K_n(z,w)$ be the normalized kernels for $\mathcal{L}_n$ and $L_n(z,w)$ the associated kernels. Define $\Omega_n(z,w) = L_n(z,w)/K_n(z,w) \in \mathcal{P}$. Let $\Omega(z,w) = \Omega_\mu(z,w)$ be as defined in Lemma 15.1. Then, if $\sum_1^\infty(1-|\alpha_n|) = \infty$, $\Omega_n$ converges uniformly to $\Omega$ on compact subsets of the unit disk for any $w \in \mathbf{D}$.*

**Proof.** You can use the result of Section 18 to find that

$$\frac{\Omega - \Omega_n}{B_n} = \frac{\Delta_{n1}}{K_n} = g \in H(\mathbf{D}).$$

From then on the proof runs exactly like in the previous Theorem. $\square$

The previous Theorem actually corresponds to Lemma 3.4 in [24].

# 24 Conclusion

We have given an introduction to the theory of orthogonal rational functions related to the Nevanlinna-Pick interpolation algorithm exactly like the Szegő theory of orthogonal polynomials is related to the Schur algorithm. The treatment is only introductory since we have not attempted to cover all possible generalizations of the corresponding Szegő theory. We could have discussed the case where $\log \mu' \notin L_1$. Some ideas on this case can be found in [26]. We also could have included the matrix case like in [23, 12] and many other papers or even the case of operator valued functions [72]. We could have stressed more the multipoint Padé aspect of this theory (see [41, 48, 63]) or discussed a more formal setting where orthogonality is defined with respect to a linear functional defined on a set of rational functions like in the previous papers and also in [13]. Much more results can be obtained for the asymptotics of the polynomials or the recursion coefficients and the convergence of Fourier series in these orthogonal functions. The material of positive real functions in the half plane and the corresponding orthogonal functions on the real line is left here undiscussed. And there are of course the many beautiful applications, with their own terminology and their own problem settings.

So we are well aware of the fact that the present discussion is only an appetizing survey which may hopefully invoke some interest in the field, if that were necessary at all. We think it is a fascinating subject, even more fascinating than the theory of orthogonal polynomials, if that doesn't sound too much as a blasphemy.

## References


[1] N.I. Achieser. *Theory of approximation.* Frederick Ungar Publ. Co., New York, 1956.

[2] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. Infinite Hankel matrices and generalized Carathéodory-Fejér and Riesz problems. *Funct. Anal. Appl.*, 2:1–18, 1968.

[3] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and I. Schur. *Funct. Anal. Appl.*, 2:269–281, 1968.

[4] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. *Math. USSR-Sb.*, 15:31–73, 1971.

[5] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. Infinite Hankel block matrices and related extension problems. *Izv. Akad. nauk. Armjan SSR Ser. Mat.*, 6:87–112, 1971. See also *Am. Math. soc. Transl.*, vol. 111, pages 133-156, 1978.

[6] N.I. Akhiezer. *The classical moment problem.* Oliver and Boyd, Edinburgh, 1969. Originally published Moscow, 1961.

[7] A.C. Allison and N.J. Young. Numerical algorithms for the Nevanlinna-Pick problem. *Numerische Math.*, 42:125–145, 1983.

[8] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[9] G.A. Baker Jr. and P.R. Graves-Morris. *Padé Approximants.* Addison-Wesley, Reading, Massachusetts, 1981. (2 volumes).

[10] V. Belevitch. *Classical network theory.* Holden-Day, San Francisco, 1968.

[11] R.P. Brent and F.T. Luk. A systolic array for the solution of linear-time solution of Toeplitz systems of equations. *Journ. VLSI and Comp. Systems*, 1(1):1–23, 1983.

[12] A. Bultheel. Orthogonal matrix functions related to the multivariable Nevanlinna-Pick problem. *Tijdschr. Belgisch Wisk. Genootschap, ser. B*, 32(2):149–170, 1980.

[13] 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.

[14] A. Bultheel. *Laurent series and their Padé approximations*, volume 27 of *Operator Theory: advances and applications.* Birkhäuser Verlag, Basel, 1987.

[15] 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.

[16] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions similar to Szegő polynomials. In C. Brezinski, L. Gori et al., editor, *IMACS transactions on orthogonal polynomials and their applications*, Basel, 1990. J.C. Baltzer AG. To appear.

[17] R.B. Burckel. *An introduction to classical complex analysis.* Birkhäuser Verlag, Basel, 1971.

[18] J.P. Burg Maximum entropy spectral analysis. In D.G. Childers, editor, *Modern spectral analysis*, pages 34–39, IEEE Press, 1978, IEEE Press. Originally presented at 37th Meet. Soc. Exploration Geophysicists, 1967. 1978,

[19] C. Carathéodory. Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen. *Math. Annal.*, 64:95–115, 1907.

[20] C. Carathéodory. Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. *Rend. Circ. Mat. Palermo*, 32:193–217, 1911.

[21] C. Carathéodory and L. Fejér. Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koefficienten und über den Picard-Landauschen Satz. *Rend. Circ. Mat. Palermo*, 32:218–239, 1911.

[22] P. Delsarte, Y. Genin, and Y. Kamp. On the role of the Nevanlinna-Pick problem in circuit and system theory. *Int. J. Circuit Th. Appl.*, 9:177–187, 1981.

[23] Ph. Delsarte, Y. Genin, and Y. Kamp. The Nevanlinna-Pick problem for matrix-valued functions. *SIAM J. Appl. Math.*, 36:47–61, 1979.

[24] 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.

[25] P. Dewilde and H. Dym. Lossless inverse scattering with rational networks: theory and applications. Technical Report 83-14, Delft University of Technology, Dept. of Electr. Engng, Network Theory Section, December 1982.

[26] P. Dewilde and H. Dym. Lossless inverse scattering, digital filters, and estimation theory. *IEEE Trans. on Information Theory*, IT-30:644–662, 1984.

[27] M.M. Djrbashian. A survey on the theory of orthogonal functions and some open problems. In P. Nevai, editor, *Orthogonal polynomials: Theory and practice*, volume 294 of *NATO-ASI Series C*, pages 135–146, Dordrecht, Boston, London, 1990. Kluwer Academic Publishers.

[28] M.M. Djrbashian. A survey on the theory of orthogonal systems and some open problems. In P. Nevai, editor, *Orthogonal polynomials: Theory and applications*, volume 294 of *Series C: Mathematical and Physical Sciences*, pages 135–146, Boston, 1990. Kluwer Academic Publishers.

[29] W.F. Donoghue Jr. *Monotone matrix functions and analytic continuation.* Springer Verlag, Berlin, 1974.

[30] R.G. Douglas. *Banach algebra techniques in operator theory.* Academic Press, New York, 1972.

[31] 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.

[32] P.L. Duren. *The theory of $H^p$ spaces.* Academic Press, New York, 1970.

[33] 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.

[34] T. Erdélyi, P. Nevai, J. Zhang, and J.S. Geronimo. A simple proof of "Favard's theorem" on the unit circle. preprint, 1989.

[35] J. Favard. Sur les polynomes de Tchebicheff. *Comptes Rendus de l'Académie des Sciences, Paris*, 200:2052–2053, 1935.

[36] G. Freud. *Orthogonal polynomials.* Pergamon Press, Oxford, 1971.

[37] J. B. Garnett. *Bounded analytic functions.* Academic Press, New York, 1981.

[38] Ya. Geronimus. *Polynomials orthogonal on a circle and their applications*, volume 3 of *Trans. Math. Monographs, Ser. 1*, pages 1–78. AMS, 1954.

[39] Ya. Geronimus. *Polynomials orthogonal on a circle and interval.* Pergamon Press, Oxford, 1960.

[40] I. Gohberg, editor. *I. Schur methods in operator theory and signal processing*, volume 18 of *Operator Theory: advances and applications.* Birkhäuser, Basel, 1986.

[41] P. González-Vera and O. Njåstad. Szegő functions and multipoint Padé approximation. *J. Comp. Appl. Math.*, 1990. To appear.

[42] Ya.L. Grenander and G. Szegő. *Toeplitz forms and their applications.* University of California Press, Berkley, 1958.

[43] T.H. Gronwall. On the maximum modulus of an analytic function. *Ann. of Math.*, 16(2):77–81, 1914-15.

[44] G. Hamel. Eine Charakteristische Eigenschaft beschränkter analytischer Funktionen. *Math. Annal.*, 78:257–269, 1918.

[45] G. Heinig and C. Rost. *Algebraic methods for Toeplitz-like matrices and operators.* Birkhäuser Verlag, Basel, 1984.

[46] J.W. Helton. *Orbit structure of the Möbius transformation semi-group acting on $H^\infty$ (broad-band matching)*, volume 3 of *Adv. in Math. Suppl. Stud.*, pages 129–197. Academic Press, New York, 1978.

[47] E. Hendriksen and O. Njåstad. A Favard theorem for rational functions. *J. Math. Anal. Appl.*, 142(2):508–520, 1989.

[48] E. Hendriksen and O. Njåstad. Positive multipoint Padé continued fractions. *Proc. Edinburgh Math. Soc.*, 32:261–269, 1989.

[49] T. Kailath. A view of three decades of linear filtering theory. *IEEE Trans. Inf. Th.*, IT-20:146–181, 1974. Reprinted in [50], 10-45.

[50] 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.

[51] A.N. Kolmogorov. Stationary sequences in Hilbert's space. *Bull. Moscow State Univ.*, 2(6):1–40, 1940. Reprinted in [50], 66-89.

[52] 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.

[53] H.J. Landau. Maximum entropy and the moment problem. *Bull. Am. Math. Soc. (New Series)*, 16(1):47–77, 1987.

[54] N. Levinson. The Wiener rms (root mean square) error criterion in filter design and prediction. *J. Math. Phys.*, 25:261–278, 1947.

[55] J. Makhoul. Linear prediction : a tutorial review. *Proc. IEEE*, 63:561–580, 1975.

[56] J.D. Markel and A.H. Gray Jr. *Linear prediction of speech.* Springer Verlag, New York, 1976.

[57] H. Meschkowski. *Hilbertsche Räume mit Kernfunktion.* Springer Verlag, Berlin, 1962.

[58] P. Nevai. Géza Freud, orthogonal polynomials and Christoffel functions. A case study. *J. Approx. Theory*, 48:3–167, 1986.

[59] R. Nevanlinna. Über beschränkte Funktionen die in gegebenen Punkten vorgeschriebene werte annehmen. *Ann. Acad. Sci. Fenn. Ser. A.*, 13(1):71pp., 1919.

[60] R. Nevanlinna. Asymptotische Entwickelungen beschränkter Funktionen und das Stieltjessche Momentenproblem. *Ann. Acad. Sci. Fenn. Ser. A.*, 18(5):53pp., 1922.

[61] R. Nevanlinna. Kriterien für die Randwerte beschränkter Funktionen. *Math. Z.*, 13:1–9, 1922.

[62] R. Nevanlinna. Über beschränkte analytische Funktionen. *Ann. Acad. Sci. Fenn. Ser. A.*, 32(7):75pp., 1929.

[63] O. Njåstad. Multipoint Padé approximation and orthogonal rational functions. In A. Cuyt, editor, *Nonlinear numerical methods and rational approximation*, pages 258–270. D. Reidel Publ. Comp., 1988.

[64] A. Papoulis. Levinson's algorithm, Wold decomposition, and spectral estimation. *SIAM Review*, 27(3):405–441, 1985.

[65] G. Pick. Über die Beschränkungen analytischen Funktionen Welche durch vorgegebene Funktionswerte bewirkt werden. *Math. Ann.*, 77:7–23, 1916.

[66] G. Pick. Über die Beschränkungen analytischen Funktionen durch vorgegebene Funktionswerte. *Math. Ann.*, 78:270–275, 1918.

[67] G. Pick. Über beschränkte Funktionen mit vorgeschriebenen Wertzuordnungen. *Ann. Acad. Sci. Fenn. Ser. A*, 15(3):17pp., 1920.

[68] V.P. Potapov. *The multiplicative structure of J-contractive matrix functions*, volume 15 of *American Mathematical Society Translations, ser. 2*, pages 131–243. AMS, Providence, Rhode Island, 1960.

[69] R. Redheffer. On the relation of transmission-line theory to scattering and transfer. *J. Math. Phys.*, 41:1–41, 1962.

[70] F. Riesz. Über ein Problem des Herrn Carathéodory. *J. Reine Angew. Math.*, 146:83–87, 1916.

[71] F. Riesz. Über Potenzreihen mit vorgeschriebenen Anfangsgliedern. *Acta Math.*, 42:145–171, 1918.

[72] M. Rosenblum and J. Rovnyak. *Hardy classes and operator theory.* Oxford University Press, New York, 1985.

[73] W. Rudin. *Real and complex analysis.* McGraw-Hill, New York, 2nd edition, 1974.

[74] D. Sarason. Generalized interpolation in $H^\infty$. *Trans. Amer. Math. Soc.*, 127:179–203, 1967.

[75] I. Schur. Über ein Satz von C. Carathéodory. *S.-B. Preuss. Akad. Wiss. (Berlin)*, pages 4–15, 1912.

[76] I. Schur. Über Potenzreihen die im Innern des Einheitskreises Beschränkt sind I. *J. Reine Angew. Math.*, 147:205–232, 1917. See also [40, p.31-59].

[77] I. Schur. Über Potenzreihen die im Innern des Einheitskreises Beschränkt sind II. *J. Reine Angew. Math.*, 148:122–145, 1918. See also [40, p.36-88].

[78] G. Szegő. *Orthogonal polynomials*, volume 33 of *AMS Colloquium publications.* American Mathematical Society, Providence, Rhode Island, 3rd edition, 1967. First edition 1939.

[79] J. L. Walsh. *Interpolation and approximation*, volume 20 of *AMS Colloquium publications*. American Mathematical Society, Providence, Rhode Island, 3rd edition, 1960. First edition 1935.

[80] N. Wiener. *Extrapolation, interpolation and smoothing of time series.* Wiley, New York, 1949.

[81] D.C. Youla and M. Saito. Interpolation with positive-real functions. *J. Franklin Inst.*, 284(2):77–108, 1967.