# Numerical quadrature and orthogonal rational functions

A. Bultheel*

**Abstract**

The purpose of these lecture notes is to give an introduction to the theory of orthogonal rational functions (ORF) on the unit circle and on (part of) the real line. In particular we shall concentrate on their use in the construction of quadrature formulas which include computational aspects and related interpolation problems.

There are however many aspects that are intimately related to these ideas that we do not explore in depth like continued fractions, Nevanlinna-Pick and Carathéodory interpolation, moment problems, and many other aspects of what is generally known as Schur analysis.

## Contents

**1 Quadrature formulas** **4**
- 1.1 The general setting . . . 4
- 1.2 Interpolatory qf . . . 4
- 1.3 Euclidean spaces . . . 5

**2 The polynomial case** **5**
- 2.1 Orthogonal polynomials on $\mathbb{R}$ and Gauss-qf . . . 5
- 2.2 Orthogonal polynomials on $\mathbb{T}$ and Szegő-qf . . . 6

**3 The rational case** **7**
- 3.1 The unit circle . . . 7
- 3.2 The real line . . . 7

**4 Orthogonal rational functions on the unit circle** **9**
- 4.1 Preliminaries . . . 9
- 4.2 The fundamental spaces . . . 11
- 4.3 Reproducing kernels . . . 12
- 4.4 Recurrence relations . . . 14
- 4.5 Quadrature . . . 18
- 4.6 Interpolation . . . 19
- 4.7 Interpolation and quadrature using the kernels . . . 21
- 4.8 Interpolatory product rules . . . 22

*Department of Computer Science, KU Leuven, Belgium. This text is based on work that appeared in many papers with coauthors: Maria-José Cantero, Ruyman Cruz Barroso, Leila Daruis, Karl Deckers, Pablo González-Vera, Erik Hendriksen, Andreas Lasarow, Olav Njåstad, Francisco Perdomo-Pío, Joris Van Deun, Patrick Van gucht. This is dedicated to the memory of Pablo González Vera.

**5 Orthogonal rational functions on the real line** **23**
5.1 Preliminaries . . . 23
5.2 The fundamental spaces . . . 25
5.3 Recurrence relation . . . 25
5.4 Quadrature . . . 29
5.5 Quasi-orthogonal functions . . . 30
5.6 Interpolation . . . 31
5.7 Quadrature with kernels . . . 33
5.8 The half-line and the interval . . . 33
5.9 More general situations . . . 34

**6 Computational aspects** **35**
6.1 The real line . . . 35
6.2 The unit circle . . . 36

**7 Rational Chebyshev quadrature** **40**
7.1 Rational Gauss-Chebyshev quadrature . . . 40
7.2 Rational Szegő quadrature with Chebyshev weights . . . 42
7.3 Rational modifications of a measure on $\mathbb{T}$ . . . 43
7.4 Near best rational approximation . . . 44

**8 Quadrature with preassigned nodes** **46**
8.1 Rational Szegő qf with preassigned nodes . . . 46
8.1.1 Rational Szegő-Radau qf . . . 46
8.1.2 Rational Szegő-Lobatto qf . . . 46
8.1.3 Rational Szegő qf with more than two preassigned nodes . . . 48
8.2 Rational Gauss qf on $[-1, 1]$ with preassigned nodes . . . 49
8.2.1 Rational Gauss-Radau qf . . . 49
8.2.2 Qf on $\mathbb{T}$ and $I$ with prescribed nodes . . . 50

**9 Error bound** **51**
9.1 An estimate on $\mathbb{R}$ . . . 51
9.2 An estimate on $\mathbb{T}$ . . . 51
9.3 A general estimate . . . 52

**10 Some applications** **54**
10.1 Rational Krylov methods . . . 54
10.2 Identification and model reduction . . . 55

**11 What we did not discuss** **58**

**Appendices** **61**

**A Moment problems (polynomial case)** **61**
A.1 The real line and its subsets . . . 61
A.2 The unit circle . . . 62

**B Orthogonal polynomials (a crash course)** **63**
B.1 The orthogonal polynomials . . . . . 63
B.1.1 Recurrence . . . . . 63
B.1.2 Related polynomials . . . . . 64
B.2 Reproducing kernels . . . . . 65
B.3 Hardy, Lebesgue and other spaces . . . . . 65

**C Density and the proof of Favard's theorem** **67**
C.1 Density for $\mathbb{T}$ . . . . . 67
C.2 Density for $\mathbb{R}$ . . . . . 67
C.3 Proof of Favard theorem on $\mathbb{T}$ . . . . . 68
C.4 Proof of Favard theorem on $\mathbb{R}$ . . . . . 68

**D Asymptotics for $\mathbb{T}$** **69**
D.1 Orthogonal polynomials w.r.t. varying measures . . . . . 69
D.2 Szegő's condition and asymptotics . . . . . 70
D.3 Szegő's problem . . . . . 73

# 1 Quadrature formulas

In these notes we shall be particularly interested in the generalizing Gauss-type quadrature formulas (the real line) and Szegő quadrature formulas (the unit circle) from a polynomial setting to one with rational functions. This will obviously involve also the study of orthogonal polynomials and their generalization to orthogonal rational functions and several related topics. This study will profit from an interplay of linear algebra and (complex) analysis.

## 1.1 The general setting

We shall consider a real finite nonnegative measure $\mu$ and integrals of the form $I_\mu\{f\} = \int f(z)\mathrm{d}\mu(z)$ for functions $f$ belonging to some space $\mathcal{L}$. The support of the measure is either supposed to belong to the extended real line $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ or the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$.

The general idea is to consider a sequence of finite dimensional subspaces of $\mathcal{L}$:

$$\mathcal{L}_0 \subset \mathcal{L}_1 \subset \mathcal{L}_2 \subset \cdots \subset \mathcal{L}_n \subset \cdots \subset \mathcal{L}_\infty = \bigcup_{k=0}^{\infty} \mathcal{L}_k \subset \mathcal{L} = \mathrm{cl}(\mathcal{L}_\infty).$$

We shall assume that $\mathcal{L}_n$ has dimension $n+1$ ($\mathcal{L}_{-1} = \varnothing$). A classical example is $\mathcal{L}_n = \mathcal{P}_n$, the space of polynomials of degree at most $n$.

We shall discuss quadrature formulas of the form

$$I_n\{f\} = \sum_{k=1}^{n} w_{nk} f(\xi_{nk}).$$

The $\{\xi_{nk}\}_{k=1}^n$ are called the *nodes* and $\{w_{nk}\}_{k=1}^n$ are called the *weights* of the quadrature formula (qf). Introducing the discrete measure $\mu_n$ with $n$ masses $\{w_{nk}\}_{k=1}^n$ in the nodes $\{\xi_{nk}\}_{k=1}^n$, i.e., $\mu_n = \sum_{k=1}^n w_{nk}\delta_{nk}$, we can denote the qf also as

$$I_n\{f\} = I_{\mu_n}\{f\} = \int f(z)\mathrm{d}\mu_n(z).$$

To be as efficient as possible, one wants to choose the nodes and weights to be such that the qf is exact in the largest possible subspace $\mathcal{L}_{d(n)}$.

The largest possible $d(n)$ is called the *degree of exactness* (DoE) of the qf $I_n$ and the largest possible subspace $\mathcal{L}_{d(n)}$ for which the qf integrates exactly is the *(maximal) domain of validity*. For example for $n$-point Gaussian qf, the (polynomial or algebraic) DoE is $2n-1$ and the maximal domain of validity is $\mathcal{P}_{2n-1}$.

Since we may assume that the integrand $f$ is known in the support of $\mu$, it is natural to require that $\{\xi_{nk}\}_{k=1}^n \subset \mathrm{supp}\,(\mu)$. For stability reasons it is also required that the weights of the qf are positive.

## 1.2 Interpolatory qf

If we fix the nodes in advance (for example equidistant in a finite interval of integration), then only the weights are left as linear parameters in the qf. Since the linear space $\mathcal{L}_n$ has dimension $n+1$, one may try to define the weights such that $I_\mu\{f\} = I_n\{f\}$ for all

$f \in \mathcal{L}_n$. This is obtained by *interpolatory qf*. That means that we choose in $\mathcal{L}_n$ a Lagrange-type basis of functions $\{\ell_{n1}, \ell_{n2}, \ldots, \ell_{nn}\}$ defined by the conditions $\ell_{nk}(\xi_{nl}) = \delta_{kl}$. Then generically $f \in \mathcal{L}_{n-1}$ is defined by $n$ interpolation conditions, and so any $f \in \mathcal{L}_{n-1}$ can be written as $f(z) = \sum_{k=1}^{n} \ell_{nk}(z) f(\xi_{nk})$ and the interpolating qf is obtained as the integral of the interpolant:

$$I_\mu\{f\} = \sum_{k=1}^{n} \int \ell_{nk}(z) \mathrm{d}\mu(z) f(\xi_{nk}) = \sum_{k=1}^{n} w_{nk} f(\xi_{nk}) = I_n\{f\}, \quad w_{nk} = \int \ell_{nk}(z) \mathrm{d}\mu(z).$$

Interpolatory qf with $n+1$ nodes are guaranteed to be exact in $\mathcal{L}_n$ and this is usually maximal if no extra conditions are imposed on the nodes. It may be slightly larger for example the $n$-point Newton-Cotes formulas with $n$ even are exact in $\mathcal{L}_n$.

### 1.3 Euclidean spaces

For the choice of nodes that are optimal with respect to the degree of exactness, one has to rely on an inner product and corresponding orthogonal basis functions. If the spaces $\mathcal{L}$ allows the definition of an inner product $\langle f, g \rangle_\mu = \int \overline{f(z)} g(z) \mathrm{d}\mu(z)$, then there is a substantial difference between the case $\mathbb{R}$ and the case $\mathbb{T}$.

Indeed, on $\mathbb{R}$, $z$ will vary over real values and hence, defining $f_{\bar{*}}(z) = \overline{f(\overline{z})}$, which is just $\overline{f(z)}$ for $z \in \mathbb{R}$, the inner product on $\mathcal{L}$ involves integrands in $\mathcal{L}_{\bar{*}} \cdot \mathcal{L}$ where $\mathcal{L}_{\bar{*}} = \{f : f_{\bar{*}} \in \mathcal{L}\}$. The sub-$\bar{*}$ involution is defined everywhere in $\mathbb{C}$ and coincides on $\mathbb{R}$ with just taking the complex conjugate. In most cases also $f$ will be a real function (i.e., $\mathcal{L}_{\bar{*}} = \mathcal{L}$) and then the inner product is a real one. This will be assumed most of the time in these notes.

On the other hand, for $z \in \mathbb{T}$, $\overline{f(z)} = \overline{f}(1/z) = \overline{f(1/\overline{z})}$. This transformation we shall denote as $f_*(z) = \overline{f(1/\overline{z})}$. Again it is defined everywhere in $\mathbb{C}$, but on $\mathbb{T}$ it coincides with just taking the complex conjugate. Thus $\langle f, g \rangle_\mu = \int f_*(z) g(z) \mathrm{d}\mu(z)$. Note that $f_*$ need not belong to $\mathcal{L}_n$ if $f \in \mathcal{L}_n$. For example if $\mathcal{L}_n = \mathcal{P}_n$, then $f$ is a polynomial in $z$ while $f_*$ is a polynomial in $1/z$. Thus the inner product on $\mathbb{T}$ requires integrals of functions in $\mathcal{L}_* \cdot \mathcal{L}$ with $L_* = \{f : f_* \in \mathcal{L}\}$.

Now assume that the polynomials are dense in $\mathcal{L}$. Then on $\mathbb{R}$ it is sufficient to consider moments $m_k = \int z^k \mathrm{d}\mu(z)$, $k = 0, 1, \ldots$ to define an inner product on $\mathcal{L}$ because a product of polynomials is again a polynomial. However, on $\mathbb{T}$, in inner product on $\mathcal{L}$ can only be defined if we know all the moments $m_k = \int z^k \mathrm{d}\mu(z)$, $z \in \mathbb{Z}$, because we need to integrate Laurent polynomials containing positive and negative powers of $z$.

We shall next briefly recall some details about the polynomial case, i.e., when $\mathcal{L}_n = \mathcal{P}_n$ in both cases: $\mathbb{R}$ and $\mathbb{T}$, which we shall subsequently generalize it to the rational case, which is the main subject of these notes.

## 2 The polynomial case

In this section we consider the classical cases where $\mathcal{L}_n = \mathcal{P}_n$. We start with the case $\mathbb{R}$ and treat $\mathbb{T}$ next.

### 2.1 Orthogonal polynomials on $\mathbb{R}$ and Gauss-qf

Consider the situation where $\mathrm{supp}\,(\mu)$ is a (possibly infinite) interval $I \subseteq \mathbb{R}$ and take $\mathcal{L}_n = \mathcal{P}_n$. By a particular choice of the nodes $\{\xi_{nk}\}$ in the interpolatory qf $I_n$ we may obtain a

much higher degree of exactness. We have the inner product $\langle f,g\rangle_\mu = \int f(x)g(x)\mathrm{d}\mu(x)$ and we consider an orthogonal polynomial sequence (OPS) $\{\phi_0,\phi_1,\ldots\}$ with $\phi_n \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$ and $\langle x^k,\phi_n(x)\rangle_\mu = 0$ for $k=0,1,\ldots,n-1$. We assume for definiteness that the leading coefficients of the $\phi_n$ are positive. I.e., $\phi_n(x) = \kappa_n x^n + \kappa_n' x^{n-1} + \cdots$ with $\kappa_n > 0$. It is well known that the zeros of $\phi_n$ are simple and in the support of the measure $\mu$. Let us denote them as $\{\xi_{nk}\}_{k=1}^n$ and use them to construct an interpolating qf. Take $f \in \mathcal{L}_m$, $m \geq n$, and write it as $f = q\phi_n + r$ where the quotient $q \in \mathcal{L}_{m-n}$ and the rest $r \in \mathcal{L}_{n-1}$. Clearly

$$I_\mu\{f\} = \int q(x)\phi_n(x)\mathrm{d}\mu(x) + I_\mu\{r\} = \langle q,\phi_n\rangle_\mu + I_\mu\{r\}.$$

Note that $\phi_n$ is orthogonal to $q \in \mathcal{L}_{m-n}$ if $m-n \leq n-1$ or $m \leq 2n-1$. This means that $I_\mu\{f\} = I_\mu\{r\}$ if $f \in \mathcal{L}_{2n-1}$ and since $I_\mu\{r\} = I_n\{r\}$ because $r \in \mathcal{L}_{n-1}$. On the other hand,

$$I_n\{f\} = \sum_{k=1}^n w_{nk} q(\xi_{nk})\phi_n(\xi_{nk}) + I_n\{r\} = I_n\{r\}$$

because $\phi_n(\xi_{nk}) = 0$. Thus $I_\mu\{f\} = I_\mu\{r\} = I_n\{r\} = I_n\{f\}$, which means that the qf is exact for all polynomials of degree at most $2n-1$. One can prove that $2n-1$ is the maximal possible degree of exactness for these *Gauss qf*. Moreover the weights are positive since indeed $0 < I_\mu\{\ell_{nj}^2\} = I_n\{\ell_{nj}^2\} = \sum_{k=1}^n w_{nk}\ell_{nj}^2(\xi_{nk}) = w_{nj}$ because $\ell_{nj}^2 \in \mathcal{L}_{2n-2} \subset \mathcal{L}_{2n-1}$ and $\ell_{nj}(\xi_{nk}) = \delta_{kj}$

## 2.2 Orthogonal polynomials on $\mathbb{T}$ and Szegő-qf

As mentioned above, the situation for the circle is different and we need integrals of Laurent polynomials to define an inner product. Therefore we define $\mathcal{L}_n = \mathcal{P}_n$ and $\mathcal{L}_{n*} = \{f : f_* \in \mathcal{L}_n\}$, i.e., the polynomials in $1/z$ of degree at most $n$. We need to consider integrals of functions in spaces like $\mathcal{L}_{p*}\cdot\mathcal{L}_q = \{f = gh : g \in \mathcal{L}_{p*},\ h \in \mathcal{L}_q\} = \{\sum_{k=-p}^q a_k z^k : a_k \in \mathbb{C}\}$. The inner product is then $\langle f,g\rangle_\mu = \int f_*(z)g(z)\mathrm{d}\mu(z) = \int \overline{f(z)}g(z)\mathrm{d}\mu(z)$. The OPS $\{\phi_0,\phi_1,\ldots\}$ is satisfied as before $\phi_n \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$ and $\langle z^k,\phi_n(z)\rangle_\mu = 0$ for $k=0,1,\ldots,n-1$. Note that the inner product has some particular properties like $\langle f,g\rangle_\mu = \langle z^k f(z), z^k g(z)\rangle_\mu$ and $\langle f,g\rangle_\mu = \langle g_*,f_*\rangle_\mu$.

Now that we have provided an inner product, there is another complication, before we can construct a qf similar to the Gauss situation. Namely, the zeros of $\phi_n$ are not on $\mathbb{T}$ but they are known to be inside the open disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. So these can not be used as nodes. There is a way out by considering *para-orthogonal polynomials*. They are defined by

$$Q_n(z) = \phi_n(z) + \tau\phi_n^*(z), \quad \phi_n^*(z) = z^n\phi_{n*}(z), \quad \tau \in \mathbb{T}.$$

Note that $\phi_n^*$ and hence also $Q_n \in \mathcal{L}_n$. It is known that their zeros $\{\xi_{nk}\}_{k=1}^n$ are simple and are on $\mathbb{T}$. The para-orthogonal polynomials are almost orthogonal in the sense that $\langle z^k,Q_n(z)\rangle_\mu = 0$ for $k=1,2,\ldots,n-1$ while $\langle z^n,Q_n(z)\rangle_\mu\,\langle 1,Q_n(z)\rangle_\mu \neq 0$. Indeed, using the properties of the inner product one may show that $\langle z^k,\phi_n^*(z)\rangle_\mu = 0$ for $k=1,2,\ldots,n$ and since $\phi_n$ is orthogonal to $\mathcal{L}_{n-1}$ it follows that $Q_n$ is orthogonal to $\mathrm{span}\{z^k : k=1,2,\ldots,n-1\} = \mathcal{L}_{n-1} \cap \mathcal{L}_n(0)$ where $\mathcal{L}_n(0) = \{f \in \mathcal{L}_n : f(0) = 0\}$.

Now we can see what happens if we take the zeros of $Q_n$ as the nodes of an interpolatory qf. Consider the space $\mathcal{L}_{n-1,n-1} = \mathcal{L}_{(n-1)*}\cdot\mathcal{L}_{n-1}$ which we shall denote as $\mathcal{R}_{n-1}$. Take

$f \in \mathcal{R}_{n-1}$. Let $R$ be the polynomial in $\mathcal{L}_{n-1}$ that interpolates $f$ in the zeros of $\phi_n$. Then clearly $f = R + E$ where the interpolation error $E$ should vanish in the zeros of $Q_n$, which means that $E$ can be written as $E(z) = Q_n(z)S_*(z)$ with $S \in \mathcal{L}_{n-1}$. Thus

$$I_\mu\{f\} = I_\mu\{R\} + I_\mu\{Q_n S_*\} = I_\mu\{R\} + \langle S, Q_n \rangle_\mu = I_\mu\{R\}.$$

Because the qf is of interpolatory type in the zeros of $Q_n$, hence exact in $\mathcal{L}_{n-1}$, we have $I_\mu\{R\} = I_n\{R\}$ while $I_n\{Q_n S_*\} = 0$. Hence $I_\mu\{f\} = I_\mu\{R\} = I_n\{R\} = I_n\{f\}$. Again one may prove that the weights of the qf are positive and that the domain of validity $\mathcal{R}_{n-1}$ is the largest possible one, i.e., a qf of this type can not be exact in $\mathcal{L}_{(n-1)*} \cdot \mathcal{L}_n$ or in $\mathcal{L}_{n*} \cdot \mathcal{L}_{n-1}$. This is a *Szegő qf*. Note that the nodes, and hence also the weights of a Szegő-qf depends on the choice of the parameter $\tau \in \mathbb{T}$. This gives an extra degree of freedom that can be used to place one of the nodes at a particular place on $\mathbb{T}$.

# 3 The rational case

Our next step is to generalize the Gauss and Szegő qf which are based on polynomial subspaces, to more general rational subspaces. That is subspaces of rational functions with prescribed poles. The polynomial situation considered above is a special case in the sense that a polynomial is a rational function for which we have placed all the poles at infinity. In that sense several of the properties we mentioned above will be special cases of properties proved for the rational case.

The distinction that was made between $\mathbb{R}$ and $\mathbb{T}$ in the polynomial case will be even more prominent in the rational case.

## 3.1 The unit circle

Let us first consider the unit circle. The poles for polynomials are at infinity. That is as far as possible from the support of the measure on $\mathbb{T}$. So an obvious generalization is to have a sequence of poles in the neighborhood of $\infty$, in any case away from the boundary $\mathbb{T}$. In fact we could allow also poles inside the open unit disk, but then again staying away from $\mathbb{T}$. Most of the theory was developed for poles in the points $\{1/\overline{\alpha}_k\}$ where $\{\alpha_k\}$ is a sequence of points in the open unit disk. Thus rational functions whose denominators are of the form $\prod_k (1 - \overline{\alpha}_k z)$. However, one might also allow such a sequence in the Riemann sphere $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ as long as one stays away from $\mathbb{T}$ and on condition that the functions will remain meaningful which needs the condition $\overline{\alpha}_k \alpha_l \neq 1$ for all $k, l$ (implying that no $\alpha_k \in \mathbb{T}$). In the trigonometric moment problem one investigates the existence and characterization of a positive measure $\mu$ supported on $\mathbb{T}$ such that $\int z^{-k} \mathrm{d}\mu(z) = m_k$ for a given sequence of trigonometric moments $m_k$, $k \in \mathbb{Z}$. Taking poles in the neighborhood of $\infty$ is connected to a rational or multipoint generalization of the trigonometric moment problem. It also relates to Nevanlinna-Pick or other interpolation problems in the complex plane.

## 3.2 The real line

For the case $\mathbb{R}$ the situation is even more complicated. The poles at infinity are on the extended real line $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, and since rational functions can be finite at $\infty$, infinity may well be in the support of the measure. If the integral is on a half line, then $\infty$ is a

boundary point, and if the integral is over a finite interval such as e.g. $[0, 1]$, then of course $\infty$ is again far away from the support. So there are many possible situations to consider, which also correspond to many different classical moment problems. So one may wonder what it means to take poles "in the neighborhood of $\infty$" in this case. Should they be on $\mathbb{R}$ or can they be complex? If the support of $\mu$ is all of $\mathbb{R}$, then one could take real poles, but one should be careful to assume that the integrals of functions with poles in the region of integration do exist. This is related to the *Hamburger moment problem* in the polynomial case and to the *strong Hamburger moment problem* if there are also poles at 0, i.e., when one considers Laurent polynomials. In the rational case this means (real) poles "near infinity" and poles "near the origin".
If the integral is over the positive real line $\mathbb{R}_+$, then we could take real poles outside $\mathbb{R}_+$ by taking them on the negative real axis. Some of them are "in the neighborhood of $-\infty$" and others are "in the neighborhood of 0". This is related to the *Stieltjes* and the *strong Stieltjes moment problem*.
If the integral is over $[0, 1]$, then we could take the poles either real but outside $[0, 1]$ or even anywhere in the complex plane, excluding $[0, 1]$. This is related to the *Hausdorff moment problem*.

If we choose poles not at infinity but in a neighborhood of $\infty$, we obtain again multipoint or rational generalizations of the polynomial problems just mentioned. It also corresponds to (complex) (Hermite) interpolation in multiple points for some particular functions derived from the measure. For some notes on these moment problems, see the appendix.

# 4 Orthogonal rational functions on the unit circle

We give in this chapter, an introduction to the theory of ORF for a measure that is supported on the unit circle of the complex plane. All the results of this part (and more) can be found in the monograph [28].

## 4.1 Preliminaries

We denote

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

For any function $f$ we introduce the *para-hermitian conjugate* $f_*$, defined by $f_*(z) = \overline{f(1/\overline{z})}$. Note that for $t \in \mathbb{T}$, $f_*(t) = \overline{f(t)}$.

Let $\mu$ be a probability measure on $\mathbb{T}$ with infinite support and $L_2(\mu)$ the Hilbert space with inner product $\langle f, g\rangle_\mu = \int f(t)\overline{g(t)}\mathrm{d}\mu(t)$. If $\mathrm{d}\mu = \mathrm{d}\lambda$ is the normalized Lebesgue measure $\mathrm{d}\lambda = \mathrm{d}t/(2\pi)$ we drop $\mu$ from the notation. The trigonometric moments are denoted by $c_k = \int t^{-k}\mathrm{d}\mu(t)$, $k \in \mathbb{Z}$.

Introducing the Riesz-Herglotz kernel

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

we can associate with $\mu$ its Riesz-Herglotz transform

$$\Omega_\mu(z) = \mathbf{i}c + \int D(t, z)\mathrm{d}\mu(t)$$

with $\Omega_\mu(0) = 1 + \mathbf{i}c$. This $\Omega_\mu$ belongs to the class $\mathcal{C}$ of Carathéodory functions

$$\mathcal{C} = \{f \in H(\mathbb{D}) : \mathrm{Re}\, f(\mathbb{D}) > 0\}$$

where $H(\mathbb{D})$ denotes the functions analytic in $\mathbb{D}$.

**Exercise 4.1** Prove that

$$\mathrm{Re}\,\Omega_\mu(z) = \int P(t, z)\mathrm{d}\mu(t)$$

with Poisson kernel

$$P(t, z) = \frac{1}{2}[D(t, z) + \overline{D(1/\overline{t}, z)}] = \frac{t(1 - |z|^2)}{(t - z)(1 - \overline{z}t)}, \tag{4.1}$$

which can be simplified for $t \in \mathbb{T}$ as follows

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

Moreover with $c_k = \int t^{-k}\mathrm{d}\mu(t)$

$$\Omega_\mu(z) = c_0 + 2\sum_{k=1}^{\infty} c_k z^k, \quad (c_0 = 1)$$

and the nontangential limit of $\mathrm{Re}\,\Omega_\mu(z)$ when $|z| \to 1$ is a.e. $(\mathrm{d}\lambda)$ equal to $\mu'$ which is the density function of the absolute continuous part of $\mu$. $\diamond$

Let $H_2(\mathbb{D})$ be the classical Hardy space of functions analytic in $\mathbb{D}$ that have a boundary value on $\mathbb{T}$ that is square integrable. Every function $f \in H_2(\mathbb{D})$ has a canonical inner-outer factorization. This means that it can be written as $f = UF$ with $U$ inner and $F$ outer. An inner function belongs to the class of bounded analytic functions

$$\mathcal{B}(\mathbb{D}) = \{f \in H(\mathbb{D}) : f(\mathbb{D}) \subset \mathbb{D}\}$$

and $|f(t)| = 1$ a.e. on $\mathbb{T}$. A *Blaschke product* is an example of an inner function. It is defined as $B(z) = \prod_n \zeta_n(z)$ with

$$\zeta_n(z) = z_n \frac{z - \alpha_n}{1 - \overline{\alpha}_n z}, \quad z_n = -\frac{|\alpha_n|}{\alpha_n}, \quad \alpha_n \in \mathbb{D} \setminus \{0\}$$

and $z_n = 1$ if $\alpha_n = 0$. It is well known that a Blaschke product converges (to a finite function not identically zero) if and only if

$$\sum_n (1 - |\alpha_n|) < \infty$$

which means that the $|\alpha_n|$ should approach 1 fast enough.

Any inner function is of the form

$$U(z) = e^{i\gamma} B(z) S(z), \quad \gamma \in \mathbb{R}, \quad S(z) = \exp\left\{-\int D(t,z) \mathrm{d}\nu(t)\right\}$$

with $\nu$ a bounded positive singular ($\nu' = 0$ a.e.) measure and $B$ a Blaschke product that catches all the zeros of $U$.
An outer function in $H_2(\mathbb{D})$ is a function of the form

$$F(z) = e^{i\gamma} \exp\left\{\int D(t,z) \log \psi(t) \mathrm{d}\lambda(t)\right\}, \quad \gamma \in \mathbb{R}$$

where $\log \psi \in L_1$ and $\psi \in L_2$.

If $\log \mu' \in L_1$ (Szegő's condition), then we may define the spectral factor of $\mu$

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

It is an outer function, uniquely defined up to a constant $c \in \mathbb{T}$. The constant $c$ can be fixed by requiring for example that $\sigma(0) > 0$. This $\sigma(z)$ has a nontangential limit to the boundary $\mathbb{T}$ that satisfies $|\sigma(t)|^2 = \mu'(t)$, a.e., $t \in \mathbb{T}$.

**Exercise 4.2** Prove that

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

$\diamond$

More about the material in this section can be found in many textbooks a.o. [78, 111, 103, 85, 139, 166].

## 4.2 The fundamental spaces

We select a sequence $\{\alpha_k\}_{k=0}^{\infty} \subset \mathbb{D}$ with $\alpha_0 = 0$ and define the partial Blaschke products[1]

$$B_0 = 1, \quad B_n(z) = B_{n-1}(z)\zeta_n(z) = \prod_{k=1}^{n} z_k \frac{z - \alpha_k}{1 - \overline{\alpha}_k z}, \quad n \geq 1.$$

The functions $\{B_0, B_1, \ldots, B_n\}$ span the space

$$\mathcal{L}_n = \left\{ \frac{p_n}{\pi_n} : \pi_n(z) = \prod_{k=1}^{n}(1 - \overline{\alpha}_k z), \quad p_n \in \mathcal{P}_n \right\} \tag{4.2}$$

where $\mathcal{P}_n$ is the space of polynomials of degree at most $n$.

If we set all $\alpha_k = 0$, then $\mathcal{L}_n = \mathcal{P}_n$. In that case the Gram matrix of $\mathcal{L}_n$ for the standard basis is a *Toeplitz matrix* with entries $\langle z^i, z^j \rangle_\mu = c_{j-i}$ where $i, j = 0, \ldots, n$. If all the $\alpha_k$ are mutually distinct, then the Gram matrix for the basis $\{1, 1/(1 - \overline{\alpha}_1 z), \ldots, 1/(1 - \overline{\alpha}_n z)\}$ has entries

$$\left\langle \frac{1}{1 - \overline{\alpha}_i z}, \frac{1}{1 - \overline{\alpha}_j z} \right\rangle_\mu = \frac{1}{2(1 - \overline{\alpha}_i \alpha_j)} \int \left[ \overline{D(t, \alpha_i)} + D(t, \alpha_j) \right] \mathrm{d}\mu(t) = \frac{1}{2} \frac{\overline{\Omega_\mu(\alpha_i)} + \Omega_\mu(\alpha_j)}{1 - \overline{\alpha}_j \alpha_i}.$$

Such a matrix is called a *Pick matrix*.

**Exercise 4.3** If $\partial_w^k$ denotes the $k$th derivative with respect to the variable $w$, prove that

$$\Omega_\mu^{(k)}(w) = \partial_w^k \Omega_\mu(w) = \int \partial_w^k D(t, w) \mathrm{d}\mu(t) = 2(k!) \int \frac{t \mathrm{d}\mu(t)}{(t - w)^{k+1}}$$

and

$$\overline{\Omega_\mu^{(k)}(w)} = \int \overline{\partial_w^k D(t, w)} \mathrm{d}\mu(t).$$

$\diamond$

In the more general situation where we want to construct a Gram matrix for the space $\mathcal{L}_n$ for fixed $n$ and where some of the of the $\alpha_k$ are repeated, then we can rearrange them so that equal $\alpha_k$ are grouped. We can then take the basis consisting of functions of the form $1/(1 - \overline{\alpha}_k z), \ldots, 1/(1 - \overline{\alpha}_k z)^{\nu_k}$ if $\alpha_k$ is repeated $\nu_k$ times, and similarly for the other $\alpha_i$. Technically, it is rather difficult to write a general expression of an entry in the Gram matrix for such a basis, but with the help of the previous exercise, one can show the following.

**Exercise 4.4** Suppose that there are $m+1$ different $\alpha_0, \ldots, \alpha_m$ (recall $\alpha_0 = 0$) which appear with multiplicity $\nu_0, \ldots, \nu_m$ respectively with $\sum_{i=0}^{m} \nu_i = n + 1$, then the Gram matrix for the basis of $\mathcal{L}_n$ that we just mentioned will only depend on $\Omega_\mu^{(k)}(\alpha_i)$, $k = 0, \ldots, \nu_i - 1$, $i = 0, \ldots, m$. We could call this a generalized Pick matrix. $\diamond$

Whatever the basis is that we choose for $\mathcal{L}_n$, it will always be similar to a Toeplitz matrix.

[1]The numbers $z_k \in \mathbb{T}$ are defined as $-\overline{\alpha}_k/|\alpha_k|$ if $\alpha_k \neq 0$ and they are equal to 1 otherwise. They are only needed if we want to let $n \to \infty$ and look at the convergence of the (infinite) Blaschke product. For computations with finite $n$, we may remove it, i.e., just set $z_k = 1$ for all $k$.

**Theorem 4.1** *If $W = [w_0, \ldots, w_n]^T$ is a basis for $\mathcal{L}_n$, with Gram matrix $G_n(W) = \langle W, W \rangle_\mu = [\langle w_k, w_l \rangle_\mu]_{k,l=0}^n$, then there is a Toeplitz matrix $T$ and an invertible matrix $V$ such that $VG_n(W)V^H = T$.*

**Proof.** Note that we may always choose the basis $\{\ell_k(z) = z^k/\pi_n(z) : k = 0, \ldots, n\}$ with $\pi_n(z)$ as in (4.2). Note that the Gram matrix for this basis is Toeplitz, whose entries are trigonometric moments for the measure $\mathrm{d}\mu(t)/|\pi_n(t)|^2$. Since every basis transformation is represented by an invertible matrix, the theorem follows. □

## 4.3 Reproducing kernels

Suppose $\{e_0, \ldots, e_n\}$ is a basis for $\mathcal{L}_n$, orthonormal with respect to $\mu$. Then $k_n(z, w) = \sum_{k=0}^n e_k(z)\overline{e_k(w)}$ is a reproducing kernel for $\mathcal{L}_n$. I.e., $\langle f(\cdot), k_n(\cdot, w) \rangle_\mu = f(w)$ for all $f \in \mathcal{L}_n$ and $w \in \mathbb{D}$. If $f \in \mathcal{L}_m$ with $m > n$, then $\langle f(\cdot), k_n(\cdot, w) \rangle_\mu$ gives the orthogonal projection of $f$ onto $\mathcal{L}_n$. Both observations follow immediately by writing $f$ as a linear combination of the $e_k$ and using the orthonormality. Note that the reproducing kernel is unique. The previous expression does not depend on a particular choice of the orthonormal basis. For example, it is immediately seen that if $\{e_k\}$ is an orthonormal basis for $\mathcal{L}_n$, then $\{B_n e_{k*}\}$ is also an orthonormal basis because $B_n e_{k*} \in \mathcal{L}_n$ for $k = 0, \ldots, n$ and

$$\langle B_n e_{k*}, B_n e_{l*} \rangle_\mu = \int |B_n(t)|^2 \overline{e_k(t)} e_l(t) \mathrm{d}\mu(t) = \langle e_l, e_k \rangle = \delta_{k,l}.$$

Therefore, by using $B_n(z)e_{k*}(z) = B_n(z)\overline{e_k(1/\overline{z})}$ as orthonormal basis:

$$k_n(z, w) = B_n(z)\overline{B_n(w)} \sum_{k=0}^n \overline{e_k(1/\overline{z})} e_k(1/\overline{w}) = B_n(z)\overline{B_n(w)} k_n(1/\overline{w}, 1/\overline{z}).$$

By using a basis transformation, one may express $k_n(z, w)$ in terms of *any* basis for $\mathcal{L}_n$ as follows.

**Theorem 4.2** *If $E_n(z) = [e_0(z), \ldots, e_n(z)]^T$ represents any basis for $\mathcal{L}_n$, then the reproducing kernel is given by*

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

*where $G_n = \langle E_n, E_n \rangle_\mu$ is the Gram matrix of $E_n$. The superscript $^H$ refers to the complex conjugate (Hermitian) transpose.*

**Proof.** Let $F_n = VE_n$ with $V$ invertible be a column of orthonormal basis functions. Then $I = \langle F_n, F_n \rangle_\mu = V \langle E_n, E_n \rangle_\mu V^H = VG_nV^H$, so that $G_n^{-1} = V^H V$ and thus $k_n(z, w) = F_n(w)^H F_n(z) = E_n(w)^H V^H V E_n(z) = E_n(w)^H G_n^{-1} E_n(z)$, which proves the theorem. □

From now on, we shall use the notation $\phi_k$ for the orthonormal basis for $\mathcal{L}_n$ that is ordered such that $\phi_0 = 1$ and $\phi_k \in \mathcal{L}_k \setminus \mathcal{L}_{k-1}$ for $k = 1, 2, \ldots$. Using the partial Blaschke products $B_k$, which also form a basis, we can write $\phi_n = a_{n,n}B_n + a_{n,n-1}B_{n-1} + \cdots + a_{n,1}B_1 + a_{n,0}B_0$. We shall denote the leading coefficient $a_{n,n}$ (with respect to the basis $B_k$) by $\kappa_n$. Since $\phi_n$

is uniquely defined up to a constant of modulus 1, we can fix $\phi_n$ uniquely, for example by assuming that $\kappa_n > 0$, which we shall do throughout this lecture, unless stated otherwise. Note that the trailing coefficient $a_{n,0}$ (with respect to the basis $B_k$) is given by $\phi_n(\alpha_1)$. To derive a similar notation for the leading coefficient, we need a generalization of the *reciprocal* function. We shall denote it by a superscript $^*$. In general we set for any function $f \in \mathcal{L}_n$: $f^*(z) = B_n(z)f_*(z)$. Note that the superstar notation is ambiguous since it depends on the $n$ that is considered. So a notation like $f^{[n]}$ instead of $f^*$ would be more appropriate. However, in order not to overload the notation, if not explicitly mentioned, it should be clear from the context which $n$ is intended. Note that with this notation we have $\phi_n^*(\alpha_n) = \kappa_n > 0$. We can now immediately write down as a consequence of the previous theorem:

**Corollary 4.3** *If the $\phi_n = \kappa_n B_n + \cdots$ with $\kappa_n > 0$ are the orthonormal basis functions for $\mathcal{L}_n$ as introduced above and $k_n(z, w)$ is the reproducing kernel, then $k_n(z, \alpha_n) = \kappa_n \phi_n^*(z)$ and $k_n(\alpha_n, \alpha_n) = \kappa_n^2$.*

**Proof.** Exercise. □

The following Christoffel-Darboux (CD) relations hold:

**Theorem 4.4** *With the notation just introduced we have*

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

*If $z = w = t \in \mathbb{T}$, then one may pass to the limit which will introduce derivatives:*

$$P(t, \alpha_{n+1})k_n(t,t) = t[\phi_{n+1}'(t)\overline{\phi_{n+1}(t)} - \phi_{n+1}^{*'}(t)\overline{\phi_{n+1}^*(t)}],$$

*with $P(t, z)$ the Poisson kernel (4.1).*

**Proof.** Because the numerators and denominators of the expressions in the right hand side vanish for $z = 1/\overline{w}$, this zero cancels out and the right hand sides are in $\mathcal{L}_n$ as a function of $z$. Using $k_n(z,w) = B_n(z)\overline{B_n(w)}\sum_{k=0}^n \phi_{k*}(z)\overline{\phi_{k*}(w)}$ (which follows from Theorem 4.2) we may write

$$\frac{k_{n+1}(z,w)}{B_{n+1}(z)\overline{B_{n+1}(w)}} - \frac{k_n(z,w)}{B_n(z)\overline{B_n(w)}} = \phi_{(n+1)*}(z)\overline{\phi_{(n+1)*}(w)}.$$

Multiplying by the denominator gives

$$[k_n(z,w) + \phi_{n+1}(z)\overline{\phi_{n+1}(w)}] - \zeta_{n+1}(z)\overline{\zeta_{n+1}(w)}k_n(z,w) = \phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)}$$

which gives the first formula when $z$ and $w$ do not coincide on $\mathbb{T}$. The other formula is obtained when in the previous formula we replace $n$ by $n-1$ so that

$$k_n(z,w) + \zeta_n(z)\overline{\zeta_n(w)}[\phi_n(z)\overline{\phi_n(w)} - k_n(z,w)] = \phi_n^*(z)\overline{\phi_n^*(w)}$$

from which the proof follows. The confluent case is left as an exercise. □

It can be seen that this relation does not depend on the fact that all $\alpha_k \in \mathbb{D}$.

Note that if we set $\phi_k = p_k/\pi_k$ with $\pi_k(z) = \prod_{i=0}^k (1 - \overline{\alpha}_i z)$, then the $p_k$ do not form an orthogonal polynomial sequence w.r.t. a positive measure, so that the theory of ORF is not quite the same as the theory of orthogonal polynomials w.r.t. a varying measure. With the CD formula, setting $z = w$, it is not difficult to derive the following property.

**Theorem 4.5** *For all $n \geq 0$, $\phi_n^*(z) \neq 0$ for $z \in \mathbb{D}$ (hence, $\phi_n(z) \neq 0$ for $z \in \mathbb{E}$) and $|\phi_{n+1}(z)/\phi_{n+1}^*(z)| < 1$, $(= 1, > 1)$ for $z \in \mathbb{D}$, $(\mathbb{T}, \mathbb{E})$.*

**Proof.** Exercise. □

## 4.4 Recurrence relations

The kernels satisfy a fundamental recurrence relation

**Theorem 4.6** *The kernels satisfy (superstar w.r.t. $z$)*

$$\begin{bmatrix} k_n^*(z,w) \\ k_n(z,w) \end{bmatrix} = t_n(z,w) \begin{bmatrix} k_{n-1}^*(z,w) \\ k_{n-1}(z,w) \end{bmatrix}$$

*with*

$$t_n(z,w) = c_n \begin{bmatrix} 1 & \overline{\rho}_n \\ \rho_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \overline{\gamma_n} \\ \gamma_n & 1 \end{bmatrix}$$

*where*

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

**Proof.** The CD relation implies

$$k_n(z,w) = \phi_n^*(z)\overline{\phi_n^*(w)} + \zeta_n(z)\overline{\zeta_n(w)}k_{n-1}(z,w).$$

Multiply this with $\overline{\rho_n(w)}$ and substitute (superstar w.r.t. $z$)

$$\phi_n^*(z)\overline{\phi_n^*(w)\rho_n(w)} = k_n^*(z,w) - \zeta_n(z)k_{n-1}^*(z,w)$$

to get

$$k_n(z,w)\overline{\rho_n(w)} = \zeta_n(z)\overline{\zeta_n(w)}k_{n-1}(z,w)\overline{\rho_n(w)} + k_n^*(z,w) - \zeta_n(z)k_{n-1}^*(z,w).$$

Take the superstar conjugate of this relation and solve for $k_n^*(z,w)$ and $k_n(z,w)$ and the result follows. □

The CD relation also implies a recurrence relation for the $\phi_n$.

**Theorem 4.7** *The orthonormal functions satisfy*

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = \tilde{t}_n(z) \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix}$$

*where the matrix $\tilde{t}_n(z)$ can be written in different ways:*

$$\begin{aligned} \tilde{t}_n(z) &= \frac{\kappa_n}{\kappa_{n-1}} \begin{bmatrix} 1 & 0 \\ 0 & z_n \end{bmatrix} \begin{bmatrix} \varepsilon_n & \delta_n \\ \overline{\delta}_n & \overline{\varepsilon}_n \end{bmatrix} \begin{bmatrix} \frac{z-\alpha_{n-1}}{1-\overline{\alpha}_n z} & 0 \\ 0 & \frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z} \end{bmatrix} \\ &= \frac{\kappa_n}{\kappa_{n-1}} \frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z} \begin{bmatrix} \overline{z}_{n-1}\varepsilon_n & 0 \\ 0 & z_n\overline{\varepsilon}_n \end{bmatrix} \begin{bmatrix} 1 & \overline{\varsigma}_n \\ \varsigma_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \\ &= c_n \frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z} \begin{bmatrix} \eta_n^1 & 0 \\ 0 & \eta_n^2 \end{bmatrix} \begin{bmatrix} 1 & \overline{\varsigma}_n \\ \varsigma_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix}, \quad \eta_n^1, \eta_n^2 \in \mathbb{T}, \end{aligned}$$

*with*

$$c_n = \sqrt{\frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2}}\frac{1}{\sqrt{1-|\varsigma_n|^2}}$$

$$\varepsilon_n = z_n\frac{1-\overline{\alpha}_{n-1}\alpha_n}{1-|\alpha_{n-1}|^2}\,\frac{\overline{\phi_n^*(\alpha_{n-1})}}{\kappa_n}, \quad \delta_n = \frac{1-\overline{\alpha}_n\alpha_{n-1}}{1-|\alpha_{n-1}|^2}\,\frac{\phi_n(\alpha_{n-1})}{\kappa_n},$$

$$\varsigma_n = \overline{z}_{n-1}\frac{\overline{\delta}_n}{\overline{\varepsilon}_n} = \eta_n^1\frac{\overline{\phi_n(\alpha_{n-1})}}{\phi_n^*(\alpha_{n-1})} \in \mathbb{D},$$

$$\eta_n^1 = z_n\overline{z}_{n-1}\frac{1-\overline{\alpha}_{n-1}\alpha_n}{1-\overline{\alpha}_n\alpha_{n-1}} \in \mathbb{T}, \quad \eta_n^2 = \overline{z}_{n-1}z_n\overline{\eta}_n^1 = \frac{1-\overline{\alpha}_n\alpha_{n-1}}{|1-\overline{\alpha}_n\alpha_{n-1}|}\frac{\phi_n^*(\alpha_{n-1})}{|\phi_n^*(\alpha_{n-1})|} \in \mathbb{T}.$$

**Proof.**(sketch) From the superstar conjugate (w.r.t. $z$) of the CD relation we find

$$\frac{\phi_n^*(z)\phi_n(\alpha_{n-1}) - \phi_n(z)\phi_n^*(\alpha_{n-1})}{\zeta_n(\alpha_{n-1}) - \zeta_n(z)} = \phi_{n-1}(z)\kappa_{n-1}.$$

The superstar conjugate of this relation is

$$\frac{\phi_n^*(z)\overline{\phi_n^*(\alpha_{n-1})} - \phi_n(z)\overline{\phi_n(\alpha_{n-1})}}{1-\zeta_n(z)\overline{\zeta_n(\alpha_{n-1})}} = \phi_{n-1}^*(z)\kappa_{n-1}.$$

Eliminate $\phi_n^*(z)$ between these two relations and the first form of the recurrence for $\phi_n(z)$ is obtained. The second one follows immediately from this because the CD relations imply $(|\phi_n^*(z)|^2 - |\phi_n(z)|^2)/(1-|\zeta_n(z)|^2) > 0$ so that $\phi_n^*(\alpha_{n-1}) \neq 0$ and $\varsigma_n \in \mathbb{D}$. For the third one, recall that $\kappa_n = \phi_n^*(\alpha_n)$, and use again the CD relation with $z = w = \alpha_n$ to obtain the appropriate factor in front. □

It may be easier to catch the idea of the formula by rewriting the first relation as

$$\phi_n(z) = \tilde{c}_n\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z}[\zeta_{n-1}(z)\phi_{n-1}(z) + \overline{\varsigma}_n\phi_{n-1}^*(z)], \quad \tilde{c}_n = c_n\eta_n^1$$

or, disregarding the normalizations, also

$$\phi_n(z) = \frac{z-\alpha_{n-1}}{1-\overline{\alpha}_n z}\phi_{n-1}(z) + \overline{\varsigma}_n\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z}\phi_{n-1}^*(z).$$

The other relation is just the superstar conjugate of this one.

Note that if all the $\alpha_k$ are zero, then we recover the polynomial case. The recurrence relation is just the Szegő recurrence and the $\varsigma_n$ being then equal to $\overline{\phi_n(0)}/\overline{\phi_n^*(0)}$, are the Szegő parameters (and are sometimes called Schur or reflection coefficients or partial correlation (PARCOR) coefficients, but currently people start using the name of *Verblunsky coefficients*).

Since this is derived from the CD relation, the first recurrence does not depend on the fact that the $\alpha_k$ are in $\mathbb{D}$. It is not so difficult to generalize the whole theory to the case where $\alpha_k$'s are in $\mathbb{D} \cup \mathbb{E}$ (the $\alpha_k \in \mathbb{T}$ require a different treatment). However, we can not allow $\overline{\alpha}_{n-1}\alpha_n = 1$ because the $\varsigma_n$ is not defined. Denoting $\phi_n = p_n/\pi_n$ with $p_n \in \mathcal{P}_n$, then we call the system *degenerate* if $p_n(\overline{\alpha}_{n-1}) = 0$. If $p_n(\alpha_{n-1}) = 0$, then we call the system *exceptional*. If all the $\alpha_k$ are in $\mathbb{D}$ or all in $\mathbb{E}$, then the system is non-degenerate and $\varsigma_n \in \mathbb{D}$.

The vector $[\phi_n, \phi_n^*]$ is a solution of the recurrence with initial condition $[\phi_0, \phi_0^*] = [1, 1]$. This is not the only solution. With the initial condition $[1, -1]$ we get another, independent

solution. It will be formulated in terms of some functions $\psi_n \in \mathcal{L}_n$ which we shall call the ORF *of the second kind*. We introduce them as follows.

$$\psi_0 = 1, \quad \psi_n(z) = \int D(t,z)[\phi_n(t) - \phi_n(z)]\mathrm{d}\mu(t), \quad n \geq 1.$$

**Exercise 4.5** Let $f$ be such that $f_* \in \mathcal{L}_{n-1}$, then prove by orthogonality that

$$\int D(t,z)[f(t)/f(z) - 1]\phi_n(t)\mathrm{d}\mu(t) = 0.$$

From this derive that for such an $f$

$$\psi_n(z) = \int D(t,z)[\phi_n(t)f(t)/f(z) - \phi_n(z)]\mathrm{d}\mu(t), \quad n \geq 1. \tag{4.3}$$

Similarly derive that for an $f \neq 0$ such that $f \in \mathcal{L}_n$ and $f(\alpha_n) = 0$ it holds that

$$\psi_n^*(z) = \int D(t,z)\left[\phi_n^*(t)f(t)/f(z) - \phi_n^*(z)\right]\mathrm{d}\mu(t), \quad n \geq 1.$$

$\diamond$

We can now prove that $(\psi_n, -\psi_n^*)$ is another solution of the recurrence relation for $(\phi_n, \phi_n^*)$, which corresponds to the initial condition $(1, -1)$. In other words $(\psi_n, \psi_n^*)$ satisfies the same recurrence relation as $(\phi_n, \phi_n^*)$, but with $\varsigma_n$ replaced by $-\varsigma_n$.

**Theorem 4.8** *With the notation of theorem 4.7 we have*[2]

$$\begin{bmatrix} \phi_n(z) & \psi_n(z) \\ \phi_n^*(z) & -\psi_n^*(z) \end{bmatrix} = \tilde{t}_n(z) \begin{bmatrix} \phi_{n-1}(z) & \psi_{n-1}(z) \\ \phi_{n-1}^*(z) & -\psi_{n-1}^*(z) \end{bmatrix}. \tag{4.4}$$

**Proof.** We only have to prove the relation for the $\psi_n$. From the definition and the previous exercise we have for $n > 1$

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

Multiply from the left by $\tilde{t}_n(z)$ and the right hand side becomes

$$-\Omega_\mu(z) \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} + \int D(t,z)Q(t,z) \begin{bmatrix} \phi_n(t) \\ \phi_n^*(t) \end{bmatrix} \mathrm{d}\mu(t)$$

with

$$Q(t,z) = \frac{(z - \alpha_{n-1})(1 - \overline{\alpha}_n t)}{(1 - \overline{\alpha}_n z)(t - \alpha_{n-1})}.$$

Using a technique like in the previous exercise, it can be shown that the integral equals

$$\int D(t,z) \begin{bmatrix} \phi_n(t) \\ \frac{\varsigma_n(z)}{\varsigma_n(t)}\phi_n^*(t) \end{bmatrix} \mathrm{d}\mu(t)$$

[2]clearly $[\psi_n, -\psi_n^*]$ satisfies the same recurrence as $[\phi_n, \phi_n^*]$. This is equivalent with saying that $[\psi_n, \psi_n^*]$ satisfies the same recurrence as $[\phi_n, \phi_n^*]$ but with $-\varsigma_k$ instead of $\varsigma_k$, i.e., with Szegő parameters of opposite sign.

so that in the right hand side we get again the same expression as in the starting relation but with $n-1$ replaced by $n$. This proves the theorem for $n > 1$. Checking the theorem for $n = 1$ is left as an exercise. □

Far reaching generalizations of the CD relations can be obtained for any couple of solutions $(x_n, x_n^+)$ and $(y_n, y_n^+)$ of this recurrence. We give them without proof. See [28, Thm. 4.3.1].

**Theorem 4.9** *Given two solutions $(x_n, x_n^+)$ and $(y_n, y_n^+)$ of this recurrence, we can define*

$$F_n(z,w) = x_n^+(z)y_n(w) - x_n(z)y_n^+(w),$$

*and then we have a Liouville-Ostrogradskii formula*

$$\frac{F_n(z,w)}{1-\zeta_n(z)/\zeta_n(w)} - B_n(w)\frac{F_0(z,w)}{1-\zeta_0(z)/\zeta_0(w)} = -\sum_{k=0}^{n-1} x_k(z)y_k^+(w)\frac{B_n(w)}{B_k(w)}$$

*and with the definition*

$$G_n(z,w) = x_n^+(z)\overline{y_n^+(w)} - x_n(z)\overline{y_n(w)},$$

*we have a Green formula*

$$\frac{G_n(z,w)}{1-\zeta_n(z)\overline{\zeta_n(w)}} - \frac{G_0(z,w)}{1-\zeta_0(z)\overline{\zeta_0(w)}} = \sum_{k=0}^{n-1} x_k(z)\overline{y_k(w)}.$$

Choosing $(x_n, x_n^+)$ and/or $(y_n, y_n^+)$ equal to $(\phi_n, \phi_n^*)$ or $(\psi_n, \psi_n^*)$ gives several identities which we leave for the reader to discover. We just give one example:

$$\psi_n(z)\phi_n^*(z) + \psi_n^*(z)\phi_n(z) = 2B_n(z)P(z,\alpha_n) \tag{4.5}$$

with $P(z,w)$ the Poisson kernel (4.1). It is obtained by taking $z = w$ and $(x, x^+) = (\phi, \phi^*)$ and $(y, y^+) = (\psi, -\psi^*)$ in the Liouville-Ostrogradskii formula. It can also be obtained by taking determinants in (4.4). Therefore we refer to it as the *determinant formula*.

Note that the CD relation was crucial in our development so far. From this, we derived the recurrence relation for the kernels, and for the ORF which gave rise to the introduction of the second kind functions.

This can be inverted: if we have a recurrence relation for $\phi_n$ as given above, then they will be ORF with respect to some positive measure on $\mathbb{T}$. This is a Favard type theorem which reads as follows:

**Theorem 4.10 (Favard)** *Given a sequence $\{\alpha_0 = 0, \alpha_1, \alpha_2, \ldots\} \subset \mathbb{D}$ and suppose that with initial condition $\phi_0 = 1$, the $\phi_n$ are generated by a recurrence relation of the 3rd form given in Theorem 4.7 with all $\varsigma_n \in \mathbb{D}$, and the unimodular constants such that $\phi_n^*(\alpha_n) > 0$, then these $\phi_n$ will form a sequence of ORF with respect to a probability measure on $\mathbb{T}$.*

We formulate the theorem here as a motivation for the introduction of quadrature formulas and will give its proof later in section C.

Because the functions of the second kind satisfy a recurrence relation of this form (the $\varsigma_n$ have to be replaced by $-\varsigma_n$), it follows by the Favard theorem that they also are a system of ORF, with respect to a positive measure on $\mathbb{T}$. Thus we can attribute to $\psi_n$ the same

properties as we can attribute to the $\phi_n$ (location of the zeros, CD relations, etc.). If $\Omega_\mu \in \mathcal{C}$ is the Riesz-Herglotz transform of $\mu$, then clearly $1/\Omega_\mu$ is also in $\mathcal{C}$. So it can be written as the Riesz-Herglotz transform of some measure $\nu$. It can be shown that if the $\phi_n$ are orthogonal w.r.t. $\mu$, then the associated functions of the second kind $\psi_n$, will be orthogonal w.r.t. the associated $\nu$.

To obtain a constructive proof for this theorem is closely related to moment problems for these ORF and techniques used in convergence and asymptotic analysis etc. The idea is to construct a measure that generates an inner product that entails orthogonality in $\mathcal{L}_n$ and then let $n$ tend to infinity. The first part is related to numerical quadrature and interpolation which we shall consider in the next section.

## 4.5 Quadrature

One way to obtain a quadrature formula for the integral $I_\mu\{f\} = \int f(t)\mathrm{d}\mu(t)$ is to interpolate $f$ by a function $f_n$ in a number of points $\{\xi_{ni}\}_{i=1}^n \subset \mathbb{T}$, and to approximate $I_\mu\{f\}$ by $I_n\{f\}$ with $I_n\{f\} = I_\mu\{f_n\}$.

**Exercise 4.6** Consider the function $f_n \in \mathcal{L}_{p,q} = \{fg : f \in \mathcal{L}_{p*}, g \in \mathcal{L}_q\}$, $p, q \geq 0$, $p + q = n - 1$, where $\mathcal{L}_{p*} = \{f : f_* \in \mathcal{L}_p\}$. Show that this $f_n$ shall interpolate $f$ in the points $\{\xi_{ni}\}_{i=1}^n \subset \mathbb{T}$ if

$$f_n(t) = \sum_{k=1}^n L_{n,k}(t) f(\xi_{nk}), \quad L_{n,k}(t) = \ell_{n,k}(t) \left( \prod_{i=1}^q \frac{1 - \overline{\alpha}_i \xi_{nk}}{1 - \overline{\alpha}_i t} \right) \prod_{i=1}^p \frac{\xi_{nk} - \alpha_i}{t - \alpha_i},$$

where $\ell_{n,k}$ are the classical Lagrange polynomials, i.e., $\ell_{n,k}(\xi_{nj}) = \delta_{k,j}$. Note that the $L_{n,k} \in \mathcal{L}_{p,q}$ are rational generalizations for which also $L_{n,k}(\xi_{nj}) = \delta_{k,j}$. $\diamond$

Thus the quadrature formula is $I_n\{f\} = \sum_{k=1}^n \lambda_{n,k} f(\xi_{nk})$, $\lambda_{n,k} = I_\mu\{L_{n,k}\}$. This is called an interpolating quadrature formula, which is obviously exact for all $f \in \mathcal{L}_{p,q}$.

To obtain the largest possible domain in which we get an exact quadrature formula, we have to choose the nodes $\xi_{nk}$ in a particular way. Like in Gaussian quadrature formulas, we could try to choose them as the zeros of $\phi_n$, but that is impossible because the ORF $\phi_n$, just like orthogonal polynomials, have all their zeros inside $\mathbb{D}$, and not on $\mathbb{T}$. However *para-orthogonal rational functions* (PORFs) will have poles on $\mathbb{T}$. First we define $\mathcal{L}_{n-1}(\alpha_n) = \{f \in \mathcal{L}_{n-1} : f(\alpha_n) = 0\}$.

**Definition 4.1** *We say that $Q_n \in \mathcal{L}_n$ is a PORF if it is orthogonal to $\mathcal{L}_{n-1}(\alpha_n)$ and $\langle 1, Q_n \rangle_\mu \neq 0 \neq \langle B_n, Q_n \rangle_\mu$.*

The following result holds.

**Theorem 4.11** *The function $Q_n(z, \tau) = \phi_n(z) + \tau \phi_n^*(z)$, $\tau \in \mathbb{T}$ is a PORF and it has $n$ simple zeros on $\mathbb{T}$.*

**Proof.** By the CD relations we know that $|\phi_n^*|^2 > |\phi_n|^2$ in $\mathbb{D}$ and the opposite inequality holds in $\mathbb{E}$. Thus $Q_n$ can only have zeros in $\mathbb{T}$. These zeros are simple, otherwise, we would have $Q_n(\xi) = Q_n'(\xi) = 0$, which implies that in $\xi$ we have $\phi_n/\phi_n^* = -\tau = -1/\overline{\tau} = \overline{\phi_n^*/\phi_n}$ or $\phi_n' \overline{\phi_n} - (\phi_n^*)' \overline{\phi_n^*} = 0$. But letting $z$ and $w$ approach $\xi \in \mathbb{T}$ in the CD formula leads to $\frac{(1-\overline{\alpha}_n \xi)(\xi - \alpha_n)}{1-|\alpha_n|^2} [\phi_n'(\xi)\overline{\phi_n(\xi)} - (\phi_n^*)'(\xi)\overline{\phi_n^*(\xi)}] = k_{n-1}(\xi, \xi) > 0$, which is a contradiction.

Concerning the para-orthogonality we note that $\langle f, \phi_n \rangle_\mu = 0$ if $f \in \mathcal{L}_{n-1}$ while if $f \in \mathcal{L}_n(\alpha_n)$, then it can be written as $B_n g_*$ with $g \in \mathcal{L}_{n-1}$, and thus $\langle f, \phi_n^* \rangle_\mu = \langle B_n g_*, B_n \phi_{n*} \rangle_\mu = \langle \phi_n, g \rangle_\mu = 0$. On the other hand $\langle 1, Q_n \rangle_\mu = \tau \langle 1, \phi_n^* \rangle_\mu = \tau \langle \phi_n, B_n \rangle_\mu \neq 0$, and similarly for the other inequality. □

Now take an arbitrary function $R \in \mathcal{L}_{n-1,n-1}$, suppose we interpolate it with a function $R_n \in \mathcal{L}_{0,n-1} = \mathcal{L}_{n-1}$, using the zeros $\{\xi_{n1}, \ldots \xi_{nn}\}$ of $Q_n$. Then simply writing it out shows that the interpolation error can be written as

$$E_n(z) = R(z) - R_n(z) = Q_n(z)S(z), \quad S \in \mathcal{L}_{n-1} \cap \mathcal{L}_n(\alpha_n) = \mathcal{L}_{n-1}(\alpha_n).$$

Thus, because of the para-orthogonality of $Q_n$, we get an integration error $I_\mu\{E_n\} = 0$, so that the quadrature formula is exact in $\mathcal{L}_{n-1,n-1}$. This is the highest possible degree of exactness that can be obtained with $n$ nodes in the sense that with $n$ nodes on $\mathbb{T}$, it is impossible to integrate exactly all functions in $\mathcal{L}_{n,n-1}$ or $\mathcal{L}_{n-1,n}$. In this sense the quadrature formula is optimal and we shall refer to it as the *rational (or R-) Szegő quadrature formula.* We denote the (discrete) measure having masses $\lambda_{nk}$ in the points $\xi_{nk}$ by $\mu_n$ so that by definition $I_n\{f\} = I_{\mu_n}\{f\}$.

**Exercise 4.7** To note the relation between reproducing kernels and PORFs, prove that if $w \in \mathbb{T}$, then $c(1 - \zeta_n(z)\overline{\zeta_n(w)})k_{n-1}(z, w)$ is a PORF where $c = -1/\overline{\phi_n(w)}$ and $\overline{\tau} = -\phi_n^*(w)/\phi_n(w) \in \mathbb{T}$.
Clearly $z = w$ is a zero of this PORF. Thus we can fix one of the zeros at $w \in \mathbb{T}$ in this representation.

◇

**Exercise 4.8** The zeros of two consecutive PORFs $Q_n$ and $Q_{n-1}$ are all simple and on $\mathbb{T}$. Show that they interlace. See [35]. ◇

**Exercise 4.9** Prove that a PORF $Q_n(z, \tau)$ can also be written in terms of $\phi_{n-1}$:

$$Q_n(z, \tau) = (1 - \tau\overline{\varsigma}_n)\left[\frac{z - \alpha_{n-1}}{1 - \overline{\alpha}_n z}\phi_{n-1}(z) + \tilde{\tau}\frac{1 - \overline{\alpha}_{n-1}z}{1 - \overline{\alpha}_n z}\phi_{n-1}^*(z)\right]$$

with $\tilde{\tau} = \frac{\tau + \varsigma_n}{1 + \tau\overline{\varsigma}_n} \in \mathbb{T}$. Use the recurrence (disregarding normalization). ◇

## 4.6 Interpolation

We have used interpolation to construct a quadrature formula and found that integration with respect to $\mu$ and $\mu_n$ is the same in $\mathcal{L}_{n-1,n-1}$. This happens if and only if their Riesz-Herglotz transforms take the same values in the points $\{\alpha_0, \ldots, \alpha_n\}$ (taking multiplicity into account) since the respective Gram matrices depend only on their (derivative) values in the poles $\{\alpha_k\}_{k=0}^{n-1}$ (see exercise 4.4). So let us first find out what the Riesz-Herglotz transform of $\mu_n$ is.

**Theorem 4.12** *Set* $\Omega_n(z) = \int D(t, z)\mathrm{d}\mu_n(t) = -\frac{P_n(z)}{Q_n(z)}$. *Then for* $n \geq 1$

$$P_n(z) = \int D(t, z)[Q_n(t)f(t)/f(z) - Q_n(z)]\mathrm{d}\mu(t) = \psi_n(z) - \tau\psi_n^*(z)$$

*for any* $f$ *such that* $f_* \in \mathcal{L}_{n-1}(\alpha_n)$.

**Proof.** If we write the terms explicitly, then it becomes clear that the previous integrand is in $\mathcal{L}_{n-1,n-1}$ for $n \geq 2$, and thus we can replace the integral by the R-Szegő quadrature formula with nodes the zeros of $Q_n$, giving $P_n(z)f(z) = -f(z)Q_n(z)I_n\{D(\cdot,z)\} = -f(z)Q_n(z)\Omega_n(z)$. The case $n = 1$ is left as an exercise. The expression with $\psi_k$ follows from (4.3) and its superstar. □

Because of this result we could call $P_n$ a PORF of the second kind. It is para-orthogonal with respect to the same measure as for which the $\psi_n$ are orthogonal.

The previous theorem has an interesting corollary.

**Corollary 4.13** *The weights of the R-Szegő formula are given by*

$$\lambda_{nk} = \frac{1}{2\xi_{nk}} \frac{P_n(\xi_{nk})}{Q_n'(\xi_{nk})} = \frac{1}{\sum_{i=0}^{n-1} |\phi_i(\xi_{nk})|^2} > 0.$$

**Proof.** Recall $\Omega_n(z) = -P_n(z)/Q_n(z) = I_n\{D(\cdot,z)\} = \sum_{k=1}^{n} \lambda_{nk} D(\xi_{nk}, z)$ and $Q_n(\xi_{nk}) = 0$. Thus after multiplying by $(z - \xi_{nk})$ we have

$$(z - \xi_{nk})\Omega_n(z) = \sum_{j \neq k} \lambda_{nj} D(\xi_{nj}, z)(z - \xi_{nk}) + \lambda_{nk} D(\xi_{nk}, z)(z - \xi_{nk}).$$

Taking the limit for $z \to \xi_{nk}$ gives the first formula. For the expression with the kernel, use the fact that $Q_n(\xi_{nk}) = 0$, thus $\tau = 1/\overline{\tau} = -\phi_n(\xi_{nk})/\phi_n^*(\xi_{nk})$, and the confluent CD formula to get

$$Q_n'(\xi_{nk}) = \frac{P(\xi_{nk}, \alpha_n)}{\xi_{nk}\overline{\phi_n(\xi_{nk})}} k_{n-1}(\xi_{nk}, \xi_{nk})$$

with $P(t,z)$ the Poisson kernel (4.1), while the determinant formula leads to $P_n(\xi_{nk}) = 2P(\xi_{nk}, \alpha_n)/\overline{\phi_n(\xi_{nk})}$, so the corollary is proved. □

We can now derive interpolation properties for $\Omega_n = -P_n/Q_n$:

**Theorem 4.14** *Let $\Omega_n$ and $\Omega_\mu$ be the Riesz-Herglotz transforms of $\mu_n$ and $\mu$ respectively, then $\Omega_\mu(z) - \Omega_n(z) = zB_{n-1}(z)h(z)$, for $n \geq 1$ with $h$ analytic in $\mathbb{D}$.*

**Proof.** This follows because of exercise 4.4. □

Since we assumed $\tau \in \mathbb{T}$, we cannot set $\tau = 0$, but we do have the same type of interpolation with an extra interpolation in $\alpha_n$.

**Theorem 4.15** *For the ORF and the ORF of the second kind we have*

$$\Omega_\mu(z)\phi_n(z) + \psi_n(z) = zB_{n-1}(z)h(z), \quad \Omega_\mu(z)\phi_n^*(z) - \psi_n^*(z) = zB_n(z)g(z)$$

*with $h$ and $g$ analytic in $\mathbb{D}$.*

**Proof.** Since $[\phi_n\Omega_\mu + \psi_n]/B_{n-1} = \int D(t,z)\phi_n(t)/B_{n-1}(t)\mathrm{d}\mu(t)$, we find a function analytic in $\mathbb{D}$ because it is a Cauchy-Stieltjes integral and setting $z = 0$, so that $D(t,0) = 1$, the integral becomes $\langle \phi_n, B_{n-1} \rangle_\mu = 0$. The second relation is similarly proved since now for $z = 0$ we find $\langle \phi_n^*, B_n \rangle_\mu = 0$. □

Note that by the determinant formula, we have for $R_n(z) = \psi_n^*/\phi_n^*$ that for $t \in \mathbb{T}$: $w_n(t) = \operatorname{Re} R_n(t) = \frac{1}{2}[R_n(t) + R_{n*}(t)] = P(t,\alpha_n)/|\phi_n(t)|^2$. Because of the extra interpolation in $\alpha_n$, one might expect that the absolute continuous measure $w_n(t)\mathrm{d}\lambda(t)$ gives the same integrals in $\mathcal{L}_{n,n}$, and it does indeed.

**Theorem 4.16** *The inner product in $\mathcal{L}_n$ is the same for the measure $\mu$ and for the absolute continuous measure with weight $w_n = P(\cdot, \alpha_n)/|\phi_n|^2$.*

**Proof.** Since $\int |\phi_n(t)|^2 w_n(t) \mathrm{d}\lambda(t) = \int P(t, \alpha_n) \mathrm{d}\lambda(t) = 1 = \|\phi_n\|_\mu^2$, the norm is maintained. Moreover

$$\langle \phi_n, \phi_k \rangle_{w_n} = \int P(t, \alpha_n) \frac{\phi_{k*}(t)}{\phi_{n*}(t)} \mathrm{d}\lambda(t) = \int P(t, \alpha_n) \left\{ \frac{\phi_k^*(t) B_n(t)/B_k(t)}{\phi_n^*(t)} \right\} \mathrm{d}\lambda(t) = 0,$$

which follows because the factor in curly brackets is analytic in $\mathbb{D} \cup \mathbb{T}$, so that we may apply Poisson's formula to find zero because $B_n(z)/B_k(z)$ is zero in $\alpha_n$. Thus $\phi_n$ has norm 1 and is orthogonal to $\mathcal{L}_{n-1}$ for the weight $w_n$, and because the recurrence relation then defines all the previous $\phi_k$ uniquely (provided they have the proper normalization of positive leading coefficient), the theorem follows. □

We note that the para-orthogonality conditions alone do not define the functions $Q_n$ completely. The PORFs that we proposed are also $\overline{\tau}$-invariant, which means that $Q_n^* = \overline{\tau} Q_n$. The latter is essential to guarantee that they have $n$ simple zeros on $\mathbb{T}$. It can be shown that this invariance property and the para-orthogonality completely defines the $Q_n$ up to the parameter $\tau \in \mathbb{T}$ and a normalizing constant factor.

## 4.7 Interpolation and quadrature using the kernels

After we formulated the recurrence relation for the kernels in Theorem 4.4, we gave the recurrence for the ORF in Theorem 4.7, but from there on, the kernels were neglected. However, what has been developed for the ORF, can be repeated and generalized to the kernels. Since the methodology is completely analogous, we leave the proofs in this section as a major exercise.

First of all, we note that the kernels are produced by their recurrence relation when we give the initial conditions $(k_0^*, k_0) = (1, 1)$. As in the case of the ORF, we obtain another independent solution when we start with $(l_0^*, -l_0) = (1, -1)$. The resulting kernels could be called kernels of the second kind. They satisfy

$$\begin{bmatrix} l_n^*(z,w) \\ -l_n(z,w) \end{bmatrix} = t_n(z,w) \begin{bmatrix} l_{n-1}^*(z,w) \\ -l_{n-1}(z,w) \end{bmatrix}$$

with $t_n(z, w)$ as in Theorem 4.4. By our remark after Theorem 4.10, it should be clear that $l_n(z, w) = \sum_{k=0}^n \psi_k(z)\overline{\psi_k(w)}$ is a reproducing kernel for the space $\mathcal{L}_n$ w.r.t. the same measure as for which the $\psi_k$ are the ORF. The interpolation properties are a bit harder to obtain because they require a parameterized (in $w$) Riesz-Herglotz transform:

$$\Omega_\mu(z,w) = \int \frac{D(t,z)}{P(t,w)} \mathrm{d}\mu(t) + c, \quad c = \frac{\overline{w} \int t \mathrm{d}\mu(t) - w \int t^{-1} \mathrm{d}\mu(t)}{1 - |w|^2} \in \mathbf{i}\mathbb{R}, \quad (z,w) \in \mathbb{D} \times \mathbb{D},$$

$D(t, z)$ the Dirichlet and $P(t, w)$ the Poisson kernel. The $c$ is chosen to make $\Omega_\mu(w, w)$ real (and hence it equals 1). Then it is possible to formulate an interpolation property saying that $l_n(z, w) - k_n(z, w)\Omega_\mu(z, w)$ vanishes for all $z \in \{w, \alpha_1, \ldots, \alpha_n\}$. Thus the special role of $\alpha_0 = 0$ has been removed and is played by an arbitrary $w \in \mathbb{D}$. Since $k_n(z, w)$ does not vanish for any $z \in \mathbb{D} \cup \mathbb{T}$ if $w \in \mathbb{D}$, we may consider $\Omega_n(z, w) = l_n(z, w)/k_n(z, w)$, which is the parameterized Riesz-Herglotz transform of the absolutely continuous measure $P(t, w)k_n(w, w)/|k_n(t, w)|^2 \mathrm{d}\lambda(t)$. The inner product in $\mathcal{L}_n$ w.r.t. the latter will not depend on $w$ and will give the same results as the inner product w.r.t. $\mathrm{d}\mu(t)$.

## 4.8 Interpolatory product rules

To approximate integrals that express an inner product in $\mathcal{L}_n$, i.e., $\langle f, g\rangle_\mu = \int f_*(t)g(t)\mathrm{d}\mu(t)$ (recall that for $t \in \mathbb{T}$, $f_*(t) = \overline{f(1/\overline{t})} = \overline{f(t)}$), one may consider the basis $\{B_0, B_1, \ldots\}$ for the $g$-factor and the basis $\{B_0, B_{1*}, B_{2*}, \ldots\} = \{B_0, 1/B_1, 1/B_2, \ldots\}$ for the $f$-factor. So by orthogonalizing the basis $\{B_0, B_1, 1/B_1, B_2, 1/B_2, \ldots\}$ we get another set of orthogonal rational functions spanning the successive rational subspaces $\mathcal{L}_{0,0} \subset \mathcal{L}_{0,1} \subset \mathcal{L}_{1,1} \subset \mathcal{L}_{1,2} \subset \cdots$ where

$$\mathcal{L}_{m,n} = \mathcal{L}_{m*} + \mathcal{L}_n = \left\{ \frac{p(z)}{\pi_m^*(z)\pi_n(z)} : p \in \mathcal{P}_{m+n} \right\}.$$

In this case we have poles $\{\alpha_1, \alpha_2, \ldots\}$ and $\{1/\overline{\alpha}_1, 1/\overline{\alpha}_2, \ldots\}$ as well. The first ones are in $\mathbb{D}$, the second ones are in $\mathbb{E}$. Thus our general theory needs a slight adaptation. However, when just doubling the poles like $\{\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4, \tilde{\alpha}_5, \ldots\} = \{\alpha_1, \alpha_1, \alpha_2, \alpha_2, \alpha_3, \ldots\}$ and constructing the corresponding ORF (we denote them as $\tilde{\phi}_n$), then

$$\mathcal{L}_{n-1,n} \perp \phi_{2n}^+ = B_{n*}\tilde{\phi}_{2n}^* \in \mathcal{L}_{n,n} \text{ while } \mathcal{L}_{n,n} \perp \phi_{2n+1}^+ = B_{n*}\tilde{\phi}_{2n+1} \in \mathcal{L}_{n,n+1}.$$

A similar derivation can be made for the nesting $\mathcal{L}_{0,0} \subset \mathcal{L}_{1,0} \subset \mathcal{L}_{1,1} \subset \mathcal{L}_{2,1} \subset \cdots$ and then we get rational functions $\phi_n^-$ satisfying

$$\mathcal{L}_{n,n-1} \perp \phi_{2n}^- = B_{n*}\tilde{\phi}_{2n} \in \mathcal{L}_{n,n} \text{ while } \mathcal{L}_{n,n} \perp \phi_{2n+1}^- = B_{(n+1)*}\tilde{\phi}_{2n+1}^* \in \mathcal{L}_{n+1,n}.$$

The recurrence relations become more involved. There is some entanglement of the recurrences for $\phi_n$ and $\hat{\phi}_n$, but they can be separated, resulting in "three-term recurrence relations" of the form

$$\phi_n^\pm(z) = a_n^\pm(z)\phi_{n-1}^\pm(z) + b_n^\pm(z)\phi_{n-2}^\pm(z)$$

The $a_n^\pm$ and $b_n^\pm$ are all rational function of the first degree in $z$.

Using the general theory for the "tilde-sequence", one may take the zeros of the PORF $\tilde{\phi}_N + \tau\tilde{\phi}_N^*$, $\tau \in \mathbb{T}$ (which are simple and on $\mathbb{T}$) and construct the corresponding interpolatory qf. If $N = 2n+1$, then this formula will be exact in $\mathcal{L}_{n,n}$. Thus there is no extra gain: a $(2n+1)$-point qf gives exactness in a space of dimension $2n+1$.

For the details we refer to [20, 25, 26, 29].

# 5 Orthogonal rational functions on the real line

As we have mentioned above, there are several possibilities, depending on whether the support of the measure is the whole real line, the positive real line or a finite interval. We shall start with the first case: the support is the whole real line. Then there are still two possibilities: either the poles for the rational functions are all off the extended real line (in which case the usual polynomials are not included) or they are all on the extended real line. In the first case, the theory is basically a transformation of the unit circle case, which we do not treat in detail here. So we shall concentrate on the case of poles on the real line. Obviously one can make the transform to the unit circle, giving an analogous theory of ORF on $\mathbb{T}$ whose poles are all on $\mathbb{T}$ as well. We shall not go into that situation. Later we shall discuss the case where the support of the measure is a real interval, possibly half infinite. In that case it is possible of course to take real poles outside the support, or even take complex poles anywhere in $\mathbb{C}$, excluding the support. The analogous situation of a measure supported on a subinterval of $\mathbb{T}$ has not been explored in great detail. Also the case of a measure supported on several nonoverlapping intervals will not be discussed here.

## 5.1 Preliminaries

Instead of Lebesgue and Hardy spaces of the circle and the disk, we need Lebesgue and Hardy spaces for the real line and the upper half plane. Which requires some adaptations. We denote

$$\mathbb{U} = \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}, \quad \mathbb{R} = \{z \in \mathbb{C} : \operatorname{Im}(z) = 0\}, \quad \mathbb{L} = \{z \in \mathbb{C} : \operatorname{Im}(z) < 0\}.$$

The Hardy space for $\mathbb{U}$ is defined as the class of functions $f$ analytic in $\mathbb{U}$ and such that

$$\sup_{y>0} \int_{-\infty}^{\infty} |f(x + \mathbf{i}y)|^p \mathrm{d}x = \|f\|_p^p < \infty, \quad 0 < p < \infty$$

and such that $\|f\|_\infty = \sup_{z\in\mathbb{U}} |f(z)| < \infty$ for $p = \infty$. The *normalized Lebesgue measure* on $\mathbb{R}$ is $\mathrm{d}\lambda(x) = \frac{\mathrm{d}x}{\pi}$. The substar conjugate will also get a new definition: $f_{\bar{*}}(z) = \overline{f(\overline{z})}$, which is the complex conjugate $\overline{f(z)}$ for $z \in \mathbb{R}$. Thus just as for the circle it extends the definition of complex conjugate from the real line to the whole complex space.

The Carathéodory functions are defined as $\mathcal{C}(\mathbb{U}) = \{f \in H(\mathbb{U}) : f(\mathbb{U}) \subset \mathbb{U}\}$ and the Schur class as $\mathcal{B}(\mathbb{U}) = \{f \in H(\mathbb{U}) : f(\mathbb{U}) \subset \mathbb{D}\}$.[3] The *Nevanlinna kernel* is replacing the Riesz-Herglotz kernel. It is defined as

$$D(t, z) = \frac{1}{\mathbf{i}} \frac{1 + tz}{t - z}.$$

and the *Poisson kernel* is $P(t, z) = \frac{\operatorname{Im} z}{|t-z|^2}$. Note however that

$$\frac{1}{2}[D(t, z) + \overline{D(\bar{t}, z)}] = P(t, z)(1 + t^2) = \frac{1}{2\mathbf{i}} \frac{z - \overline{z}}{(t - z)(t - \overline{z})}(1 + t^2).$$

[3] In systems theory, the frequency axis is the imaginary axis, in which case the variables are of the form $\mathbf{i}\omega$ with $\omega \in \mathbb{R}$. Also the $\mathbf{i}$ in the denominator of $D(t, z)$ is removed, so that our Carathéodory functions will be replaced by *positive real* functions, mapping the right-half plane to itself.

All $\mathcal{C}(\mathbb{U})$ functions have the following Nevanlinna representation

$$\Omega_{\mathring{\mu}}(z) = \mathbf{i}c + \int D(t,z)\frac{\mathrm{d}\mu(t)}{1+t^2} = \mathbf{i}c + \int D(t,z)\mathrm{d}\mathring{\mu}(t), \quad t \in \mathbb{R}, \quad z \in \mathbb{U}.$$

Since we shall need this factor $(1+t^2)^{-1}$ a lot, we introduced the special notation $\mathrm{d}\mathring{\mu}(t) = \frac{\mathrm{d}\mu(t)}{1+t^2}$ and we shall use a similar notation for other measures as well.

We assume that $\mu$ is supported on the *extended* real line. A mass point at $\infty$ would contribute a term of the form $-\mathbf{i}bz$ with $b > 0$. If $|y\Omega(\mathbf{i}y)|$ is bounded for $y > 1$, there will be no mass point at $\infty$ and one has the much simpler *Hamburger representation* using a *Cauchy kernel*

$$\Omega_\mu(z) = \frac{2}{\mathbf{i}}\int \frac{\mathrm{d}\mu(t)}{t-z}.$$

The real part is given by

$$\operatorname{Re}\Omega_{\mathring{\mu}}(z) = by + \int \operatorname{Re} D(t,z)\mathrm{d}\mathring{\mu}(t) = by + \int P(t,z)\mathrm{d}\mu(t) > 0, \quad z = x + \mathbf{i}y.$$

$\Omega_{\mathring{\mu}}$ is a *positive real function.*

By Fatou's theorem,

$$\lim_{y\downarrow 0} \operatorname{Re}\ \Omega_{\mathring{\mu}}(x+\mathbf{i}y) = \mu'(x).$$

Just like in the case of the disk, any function $f \in H_2(\mathbb{U})$ has an *inner-outer* factorization $f = UF$ with $U$ inner and $F$ outer. The definitions of inner and outer functions are like for the disk with $\mathbb{D}$ replaced by $\mathbb{U}$ and $\mathbb{T}$ replaced by $\mathbb{R}$. An example of an inner function is a *Blaschke product* with zeros $\{\beta_i\}_{i=1}^n \subset \mathbb{U}$: $B_n(z) = \prod_n \zeta_n(z)$ with

$$\zeta_n(z) = z_n\frac{z-\beta_n}{z+\beta_n}, \quad z_n = \begin{cases} \frac{|1+\beta_n^2|}{1+\beta_n^2}, & \beta_n \in \mathbb{U}\setminus\{\mathbf{i}\} \\ 1, & \beta_n = \mathbf{i}. \end{cases}$$

The Blaschke product converges if and only if

$$\sum \frac{\operatorname{Im}\beta_n}{1+|\beta_n|^2} < \infty.$$

If the $\beta_n$ do not approach $\infty$ then the condition is equivalent with $\sum_n \operatorname{Im}\beta_n < \infty$.

The *spectral factor* of $\mu'$ is an example of an outer function. It is defined as

$$\sigma(z) = c\exp\left\{\frac{1}{2}\int D(t,z)\log\mu'(x)\mathrm{d}\mathring{\lambda}(x)\right\}, \quad z \in \mathbb{U}, \quad c \in \mathbb{T}.$$

It is defined if *Szegő's condition* is satisfied, i.d., if

$$\int \log\mu'(x)\mathrm{d}\mathring{\lambda}(x) > -\infty.$$

If it exists, then $\sigma$ is unique up to the factor $c \in \mathbb{T}$. The boundary function satisfies $|\sigma(x)|^2 = \mu'(x)$ a.e. $x \in \mathbb{R}$.

## 5.2 The fundamental spaces

We select now a sequence $\{\alpha_n\}_{n=1}^{\infty} \subset \hat{\mathbb{R}} \setminus \{0\}$ (we set $\alpha_0 = \infty$) and define the products

$$b_0 = 1, \quad b_n(z) = \prod_{k=1}^{n} Z_k(z), \quad n = 1, 2, \ldots$$

with

$$Z_k(z) = \frac{z}{1 - z/\alpha_k}, \quad k = 1, 2, \ldots$$

Note that $Z_{k*}(z) = Z_k(z)$, hence also $b_{n*}(z) = b_n(z)$ for any $z \in \mathbb{C}$.

Obviously, if all $\alpha_k = \infty$, then $b_n(z) = z^n$ is the standard polynomial basis. The functions $\{b_k : k = 0, 1, \ldots, n\}$ span the space

$$\mathcal{L}_n = \left\{ \frac{p_n(z)}{\pi_n(z)} : \pi_n(z) = \prod_{k=1}^{n} (1 - z/\alpha_k),\ p_n \in \mathcal{P}_n \right\} = \left\{ \frac{q_n(z)}{\pi_n^*(z)} : \pi_n^*(z) = \prod_{k=1}^{n} (z - \alpha_k),\ q_n \in \mathcal{P}_n \right\}$$

with $\mathcal{P}_n$ the space of polynomials of degree at most $n$. Note $p_n(z) = q_n(z) \prod_{k=1}^{n} (-\alpha_k)$. Sometimes it is more convenient to work with the second form, but it is less convenient when dealing with some $\alpha_k = \infty$. We shall say that a polynomial $p \in \mathcal{P}_n$ has a zero or order $r$ at infinity when its degree is $n - r$.

## 5.3 Recurrence relation

Again, we can take any basis that spans the successive nested sequence of subspaces $\mathcal{L}_n$ and apply a Gram-Schmidt algorithm to it to obtain a sequence of orthogonal rational functions. To make the sequence unique, we shall assume that the leading coefficients (with respect to the basis $b_n$) of the ORF are positive. Thus we assume that the $n$th ORF is of the form

$$\phi_n(z) = \kappa_n b_n(z) + \kappa_n' b_{n-1}(z) + \cdots + \phi_n(0), \quad \kappa_n > 0.$$

Note that the coefficients in this expansion are all real and hence that $\phi_{n*} = \phi_n$.

We say that $\phi_n$ is *singular* if $\phi_n = p_n/\pi_n$ and $p_n(\alpha_{n-1}) = 0$, otherwise it is said to be *regular*.

**Theorem 5.1** *The following recurrence relation for the ORF holds if and only if $\phi_n$ and $\phi_{n-1}$ are regular*

$$\begin{aligned} \frac{\phi_n(z)}{Z_n(z)} &= \left( A_n + \frac{B_n}{Z_{n-2}(z)} \right) \phi_{n-1}(z) + C_n \frac{\phi_{n-2}(z)}{Z_{n-2}(z)} \\ &= (E_n Z_{n-1}(z) + B_n) \frac{\phi_{n-1}(z)}{Z_{n-1}(z)} + C_n \frac{\phi_{n-2}(z)}{Z_{n-2}(z)} \end{aligned}$$

*where $C_n = -E_n/E_{n-1} \neq 0$ and $E_n = A_n + B_n/Z_{n-2}(\alpha_{n-1}) = A_n - B_n/Z_{n-1}(\alpha_{n-2}) \neq 0$.*

**Proof.** First note that it is always possible to choose $A_n$ such that[4]

$$W_n(z) = \frac{Z_{n-2}(z)}{Z_n(z)} \phi_n(z) - A_n Z_{n-2}(z) \phi_{n-1}(z) \in \mathcal{L}_{n-1}.$$

[4]This is where we need $\phi_{n-1}$ to be regular and $\alpha_{n-2} \neq 0$ because we have to choose $A_n = p_n(\alpha_{n-2})/[\alpha_{n-2} p_{n-1}(\alpha_{n-2})]$.

Furthermore it is easily seen that $W_n \perp \mathcal{L}_{n-3}$ so that it can be written as $W_n(z) = B_n\phi_{n-1}(z) + C_n\phi_{n-2}(z)$ and the first recurrence follows.
For the second form, one should realize that $D_n = 1/Z_{n-2}(z) - 1/Z_{n-1}(z)$ is a constant that does not depend on $z$, and hence can be written as $D_n = 1/Z_{n-2}(\alpha_{n-1}) = -1/Z_{n-1}(\alpha_{n-2})$.
The fact that $E_n \neq 0$ follows by taking the numerator of the recurrence for $z = \alpha_{n-1}$,[5] while $C_n \neq 0$ follows if you write down that $b_{n-1}/Z_n \in \mathcal{L}_{n-1}$ is orthogonal to $\phi_n$. Namely $0 = \left\langle Z_n(A_n + \frac{B_n}{Z_{n-2}})\phi_{n-1}, \frac{b_{n-1}}{Z_n} \right\rangle + C_n \left\langle \frac{Z_n}{Z_{n-2}}\phi_{n-2}, \frac{b_{n-1}}{Z_n} \right\rangle$ or $0 = \left\langle (A_n + \frac{B_n}{Z_{n-2}})b_{n-1}, \phi_{n-1} \right\rangle + C_n \left\langle \frac{b_{n-1}}{Z_{n-2}}, \phi_{n-2} \right\rangle$ with the first product nonzero and hence also $C_n$.
The 'only if' part is simple. □

It can be seen that $E_n = 0$ if and only if $\phi_n$ is singular and that

$$E_n = \frac{1}{\kappa_{n-1}} \left[ \kappa_n + \frac{\kappa'_n}{Z_n(\alpha_{n-1}))} \right] = \frac{1}{\kappa_{n-1}} \left[ \kappa_n + \kappa'_n D_{n+1} \right]. \tag{5.1}$$

It is somewhat harder to prove that $E_n = -C_n E_{n-1}$, $n \geq 2$.
If the whole sequence $\phi_n$ is nonsingular, then one may start the recurrence with $\phi_0 = 1$ and $\phi_{-1} = 0$ if it is assumed that $\int \mathrm{d}\mathring{\mu}(t) = 1$, $\alpha_0 = \infty$ so that $Z_0(z) = z$.

**Exercise 5.1** Prove the claims in the footnotes of the proof of Theorem 5.1: If $\phi_{n-1}$ regular, then $A_n$ exists and if $\phi_n$ regular, then $E_n \neq 0$. Note that $A_n$ can still exist if $\phi_{n-1}$ is singular as long as $\lim_{z\to\alpha_{n-2}} p_n(z)/[zp_{n-1}(z)]$ exists. It will be shown later in Thm. 5.10 that $p_{n-1}$ can only have simple zeros, hence

$$\lim_{z\to\alpha_{n-2}} \frac{p_n(z)}{zp_{n-1}(z)} = \frac{p'_n(\alpha_{n-2})}{\alpha_{n-2}p'_{n-1}(\alpha_{n-2})}$$

exists.
Prove the formula (5.1) for $n \geq 2$ if the recurrence holds.
Also try to prove $E_n = -C_n E_{n-1}$, $n \geq 2$. Hint: use the recurrence in $\langle b_{n-1}/Z_n, \phi_n \rangle = 0$ and $\langle b_i, \phi_i \rangle = 1/\kappa_i$ together with (5.1).
The latter implies an alternative proof of $C_n \neq 0$ if $\phi_{n-1}$ and $\phi_n$ are regular because then $E_n E_{n-1} \neq 0$.

◇

**Question 5.1** What happens with the recurrence relation when regularity of $\phi_n$ and/or $\phi_{n-1}$ is given up? ◯

**Exercise 5.2** Derive the recurrence relation for orthogonal polynomials (orthogonalization of $1, z, z^2 \ldots$)

$$\phi_n(z) = (A_n z + B_n)\phi_{n-1}(z) + C_n\phi_{n-2}(z), \quad A_n \neq 0, C_n \neq 0$$

and for Laurent polynomials (orthogonalization of $1, , z^{-1}, z, z^{-2}, z^2, \ldots$)

$$\begin{aligned} \phi_{2k}(z) &= (A_{2k}z + B_{2k})\phi_{2k-1}(z) + C_{2k}\phi_{2k-2}(z), && k = 1, 2, \ldots \\ \phi_{2k+1}(z) &= (A_{2k+1}/z + B_{2k+1})\phi_{2k}(z) + C_{2k+1}\phi_{2k-1}(z), && k = 1, 2, \ldots \end{aligned}$$

with $A_n C_n \neq 0$, as a special case of the previous recurrence. ◇

[5] Here we need $\phi_n$ regular and $\alpha_{n-1} \neq 0$ because $p_n(\alpha_{n-1}) = E_n p_{n-1}(\alpha_{n-1})\alpha_{n-1}$.

For the Laurent polynomials, one may even generalize further and introduce the successive powers of $z$ and $z^{-1}$ switching arbitrarily between introducing the next positive or the next negative power. See e.g., [77].

This recurrence relation is second order, so there is another independent solution which is called the sequence of functions *of the second kind*. They are given by[6]

$$\psi_{-1}(z) = \mathbf{i}(1+z^2), \quad \psi_0(z) = \mathbf{i}z, \quad \psi_n(z) = \int D(x,z)[\phi_n(x) - \phi_n(z)]\mathrm{d}\mathring{\mu}(x), \quad n \geq 1.$$

Note that these $\psi_k$ do *not* belong to $\mathcal{L}_k$. One could e.g., generate $\sigma_k$ from the recurrence relation for $k \geq 1$ starting with initial conditions $\sigma_{-1} = 1$, $\sigma_0 = 0$. In that case $\sigma_k \in \mathcal{L}_n(\infty)$. This means that they are in $\mathcal{L}_k$ but the numerator is of degree $n-1$ at most.

The recursion can be written as

$$\begin{bmatrix} \frac{\phi_n(z)}{Z_n(z)} & \frac{\psi_n(z)}{Z_n(z)} \\ \frac{\phi_{n-1}(z)}{Z_{n-1}(z)} & \frac{\psi_{n-1}(z)}{Z_{n-1}(z)} \end{bmatrix} = \begin{bmatrix} E_n Z_{n-1}(z) + B_n & C_n \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\phi_{n-1}(z)}{Z_{n-1}(z)} & \frac{\psi_{n-1}(z)}{Z_{n-1}(z)} \\ \frac{\phi_{n-2}(z)}{Z_{n-2}(z)} & \frac{\psi_{n-2}(z)}{Z_{n-2}(z)} \end{bmatrix}$$

Taking determinants recursively, using $E_n = -C_n E_{n-1}$ and the initial conditions leads to the *determinant formula*

**Theorem 5.2** *For a regular sequence of ORFs it holds that*

$$\frac{\phi_n(z)}{Z_n(z)}\frac{\psi_{n-1}(z)}{Z_{n-1}(z)} - \frac{\psi_n(z)}{Z_n(z)}\frac{\phi_{n-1}(z)}{Z_{n-1}(z)} = \mathbf{i}\frac{1+z^2}{z^2}E_n$$

This is a special case of the *Liouville-Ostrogradskii* formula which can be proved independently of the regularity condition [28, Thm. 11.2.3].

**Theorem 5.3** *For the ORFs $\phi_k$ and the second kind functions $\psi_k$ it holds that for $n \geq 1$*

$$\frac{\phi_n(z)}{Z_n(z)}\frac{\psi_{n-1}(w)}{Z_{n-1}(w)} - \frac{\psi_n(w)}{Z_n(w)}\frac{\phi_{n-1}(z)}{Z_{n-1}(z)} = \left[\sum_{k=1}^{n-1}\phi_k(z)\psi_k(w) - D(z,w)\right] H(z,w)E_n$$

*where $H(z,w) = \frac{1}{Z_{n-1}(z)} - \frac{1}{Z_{n-1}(w)} = \frac{1}{z} - \frac{1}{w} = \frac{w-z}{zw}$.*

As before we have a *reproducing kernel* for $\mathcal{L}_n$ given by $k_n(x,y) = \sum_{k=0}^{n}\phi_k(z)\phi_k(y)$. It is again related to the *Christoffel-Darboux relation* that now takes the following form.

**Theorem 5.4** *The Christoffel-Darboux relation is*

$$\frac{\phi_n(y)\phi_{n-1}(x)}{Z_n(y)Z_{n-1}(x)} - \frac{\phi_n(x)\phi_{n-1}(y)}{Z_n(x)Z_{n-1}(y)} = H(x,y)E_n\sum_{k=0}^{n-1}\phi_k(x)\phi_k(y)$$

*with $H(x,y)$ as above. Or*

$$\frac{f_n(x)f_{n-1}(y) - f_n(y)f_{n-1}(x)}{x-y} = E_n\sum_{k=0}^{n-1}\phi_k(x)\phi_k(y) = E_n k_{n-1}(x,y)$$

*where $f_n(x) = \phi_n(x)(1 - x/\alpha_n)$.*
*The confluent form ($y \to x$) is*

$$f_n'(x)f_{n-1}(x) - f_n(x)f_{n-1}'(x) = E_n k_{n-1}(x,x).$$

[6] If the coefficients of $\phi_n$ are real, then the coefficients of $\psi_n$ will be purely imaginary because of the factor $\mathbf{i}$ in $D(x,z)$.

A simple proof mimics the proof in the polynomial case given in the appendix. Recall that the recurrence depends on the ORF being regular. However, it is possible to prove the CD relation without using the recurrence. Singularity is something that does not occur in the polynomial case. Neither will it occur in situations where the (real) poles are in some part of the real line outside the support of the measure because that is where the zeros are located which rules out singularity of the ORF. For example one might choose the poles on the negative real axis if the measure is supported on the positive real line (Stieltjes situation).

Since the proof of the CD relation depends on the recurrence, one can prove a much more general result for any sequence of solutions of the recurrence relation. The CD relation is then just a special case. Since $\{\phi_k\}$ and $\{\psi_k\}$ are two independent solutions of the recurrence, we can write any solution as $\chi_k(z;s) = \psi_k(z) + s\phi_k(z)$ with $s \in \mathbb{C}$ arbitrary. If also $z, w, t \in \mathbb{C}$, then one may prove the so-called *Green's formula*

**Theorem 5.5 (Green's formula)** *If $\chi_n(z;s)$ and $\chi_n(w;t)$ are complex solutions of the recurrence relation, then*

$$\frac{\overline{\chi_n(w;t)}\chi_{n-1}(z;s)}{\overline{Z_n(w)}Z_{n-1}(z)} - \frac{\chi_n(z;s)\overline{\chi_{n-1}(w;t)}}{Z_n(z)\overline{Z_{n-1}(w)}} = H(z,\overline{w})E_n\left[\sum_{k=1}^{n-1}\chi_k(z;s)\overline{\chi_k(w;t)} + X\right]$$

*with $X = s\bar{t} - 1 + D(z,\overline{w})(\bar{t} - s)$.*

Of course if the variables $z$ and $w$ and the parameters $s$ and $t$ are real, then all the complex conjugates can be removed using $\overline{\psi_k(x)} = -\psi_k(x)$.

As special cases, one finds several determinantal formulas as corollaries:

**Corollary 5.6**

$$\frac{\overline{\psi_n(w)}\psi_{n-1}(z)}{\overline{Z_n(w)}Z_{n-1}(z)} - \frac{\overline{\psi_{n-1}(w)}\psi_n(z)}{\overline{Z_{n-1}(w)}Z_n(z)} = H(z,\overline{w})E_n\left[\sum_{k=0}^{n-1}\psi_k(z)\overline{\psi_k(w)} + 1\right]$$

*and*

$$\frac{\overline{\psi_n(w)}\phi_{n-1}(z)}{\overline{Z_n(w)}Z_{n-1}(z)} - \frac{\overline{\psi_{n-1}(w)}\phi_n(z)}{\overline{Z_{n-1}(w)}Z_n(z)} = H(z,\overline{w})E_n\left[\sum_{k=0}^{n-1}\phi_k(z)\overline{\psi_k(w)} + D(z,\overline{w})\right]$$

Orthogonality of the subsequent rational functions, the three-term recurrence relation and the CD relation are all more or less equivalent under some technical conditions like initial conditions, regularity etc. There is just like in the polynomial case a *Favard type theorem* which says:

**Theorem 5.7 (Favard)** *If a (finite) sequence of rational functions satisfy a recurrence relation of the form given above with $E_nC_n \neq 0$, then they will form a sequence of ORF with respect to some bounded positive measure $\mu$.*

For example if the $\phi_n$ are orthogonal w.r.t. a measure $\mu$ whose Nevanlinna transform is $\Omega_{\mathring{\mu}}(z) = \mathbf{i}c + \int D(t,z)\mathrm{d}\mathring{\mu}(t)$, then the $\psi_n$ will be orthogonal w.r.t. a measure whose Nevanlinna transform is $1/\Omega_{\mathring{\mu}}$.

## 5.4 Quadrature

We may again construct interpolating quadrature formulas by selecting a set of points $\{\xi_{nk}\}_{k=1}^n \subset \mathbb{R}$ and construct a rational function $f_n \in \mathcal{L}_{n-1}$ in these points that interpolates the integrand $f$ in these points, then clearly $I_{\mathring{\mu}}(f) = \int f(x)\mathrm{d}\mathring{\mu}(x)$ will be equal to the quadrature $I_n\{f\} = I_{\mathring{\mu}}\{f_n\}$, i.e., the interpolating quadrature formula will be exact in $\mathcal{L}_{n-1}$.

The fundamental Lagrange interpolating basis functions are

$$L_{ni}(x) = \prod_{k=1}^{n-1} \frac{\xi_{ni} - \alpha_k}{x - \alpha_k} \prod_{k\neq i}^{n} \frac{x - \xi_{nk}}{\xi_{ni} - \xi_{nk}} = \frac{\pi^*_{n-1}(\xi_{ni})}{\pi^*_{n-1}(x)} \frac{q_n(x)}{(x - \xi_{ni})q_n'(\xi_{ni})} \in \mathcal{L}_{n-1}, \quad i = 1, 2, \ldots, n$$

where $\pi^*_{n-1}(x) = \prod_{k=1}^{n-1}(x - \alpha_k)$ and the node polynomial is $q_n(x) = \prod_{k=1}^{n}(x - \xi_{nk})$, so that they satisfy $L_{ni}(\xi_{nk}) = \delta_{ik}$. Thus

$$I_{\mathring{\mu}}\{f\} = \int f(x)\mathrm{d}\mathring{\mu}(x) \approx I_n\{f\} = I_{\mathring{\mu}}\{f_n\} = \sum_{k=1}^{n} \lambda_{nk} f(\xi_{nk}), \quad \lambda_{nk} = I_{\mathring{\mu}}\{L_{nk}\}$$

and the $\approx$ will be an exact equality for all $f \in \mathcal{L}_{n-1}$.

The challenge is to choose the quadrature nodes $\{\xi_{nk}\}$ such that the degree of exactness is as high as possible. So consider the spaces $\mathcal{L}_{n-1,N} = \mathcal{L}_{n-1} \cdot \mathcal{L}_N$. The error of the quadrature formula for $f \in \mathcal{L}_{n-1,N}$ will be

$$I_{\mathring{\mu}}\{f - f_n\} = \left\langle q_n, \frac{f - f_n}{q_n} \right\rangle_{\mathring{\mu}} = \int \frac{q_n(x)}{\pi_{n-1}(z)} \frac{p_{N-1}(x)}{\pi_N(x)} \mathrm{d}\mathring{\mu}(x) = \int \frac{q_n(x)}{\pi_n(z)} \frac{p_{N-1}(x)(1 - \frac{x}{\alpha_n})}{\pi_N(x)} \mathrm{d}\mathring{\mu}(x).$$

Since $\frac{p_{N-1}(x)(1-\frac{x}{\alpha_n})}{\pi_N(x)} \in \mathcal{L}_N(\alpha_n) = \{f = \frac{p_N}{\pi_N} \in \mathcal{L}_n : p_N(\alpha_n) = 0\}$, it means that the quadrature formula will be exact if $Q_n = \frac{q_n}{\pi_n} \perp \mathcal{L}_N(\alpha_n)$. In other words, to obtain exactness in $\mathcal{L}_{n,N}$, we should choose as nodes of our quadrature formula the zeros of a function $\frac{q_n}{\pi_n} \in \mathcal{L}_n$ that is orthogonal to $\mathcal{L}_N(\alpha_n)$ which implies that the latter should be a subspace of $\mathcal{L}_n$, i.e., $N \leq n$. Thus if we take the zeros of the ORF $\phi_n \perp \mathcal{L}_n(\alpha_n) = \mathcal{L}_{n-1}$, the quadrature formula will be exact in $\mathcal{L}_{n-1,n}$. All this holds under the assumption that the zeros of $\phi_n$ are all in $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ and simple.

More generally, if $R_n \in \mathcal{L}_n$ has $n$ simple zeros in $\hat{\mathbb{R}}$ and let it be orthogonal to $\mathcal{L}_{n-m}(\alpha_n)$, $0 \leq m \leq n-1$, then the interpolatory qf whose nodes are the zeros of $R_n$ will be exact in $\mathcal{L}_{n-1,n-m}$.

Note that in the polynomial case where all $\alpha_k = \infty$, then $\mathcal{L}_{n-1} \cdot \mathcal{L}_n = \mathcal{P}_{2n-1}$ and we get the classical Gauss quadrature formulas.

We may conclude:

**Theorem 5.8** *If $\phi_n$ has $n$ simple zeros in $\hat{\mathbb{R}}$, and we take these as the nodes of the interpolatory quadrature formula, then the corresponding quadrature is exact in $\mathcal{L}_{n-1,n} = \mathcal{L}_{n-1} \cdot \mathcal{L}_n$ and all the weights are positive.*

**Proof.** We only need to prove that the weights are positive. Therefore consider the function $f(z) = L_{ni}(z)[L_{ni}(z) - 1] \in \mathcal{L}_{n-1,n-1} \subset \mathcal{L}_{n-1,n}$. Since $L_{ni}(\xi_{nk}) = \delta_{ik}$, it is clear that $f(z)$ vanishes at all points $\{\xi_{nk}\}_{k=1}^n$. Because the quadrature is exact in $\mathcal{L}_{n-1,n}$, $I_{\mathring{\mu}}\{f\} = I_n\{f\} = 0$. Thus the weights of the qf are given by $\lambda_{ni} = I_{\mathring{\mu}}\{L_{ni}\} = I_{\mathring{\mu}}\{L_{ni}^2\} > 0$. □

## 5.5 Quasi-orthogonal functions

In the case of the unit circle, we needed the PORFs because these provided zeros on $\mathbb{T}$ whereas the ORF themselves had zeros inside $\mathbb{D}$ and were therefore not convenient to derive rational Szegő qf.
In the present case of the real line, the ORF do provide zeros on $\mathbb{R}$. Thus if they are simple, they can be used as quadrature nodes. However, there is some analog of the PORFs for the real line. These are called *quasi-orthogonal functions* (QORFs), and they are defined as follows. If the $\phi_n$ are the usual ORF, then the QORFs are given by

$$Q_n(z,\tau) = Q_n(z) = \phi_n(z) + \tau \frac{Z_n(z)}{Z_{n-1}(z)}\phi_{n-1}(z) = \phi_n(z) + \tau \frac{z-\alpha_{n-1}}{z-\alpha_n}\phi_{n-1}(z) \in \mathcal{L}_n.$$

$\tau \in \hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ is an arbitrary parameter. We shall first check the orthogonality properties of $Q_n$. The first term is orthogonal to $\mathcal{L}_{n-1}$ and the second one is orthogonal to all functions $f$ such that $\left\langle \frac{1-x/\alpha_{n-1}}{1-x/\alpha_n}\phi_{n-1}(x), f(x) \right\rangle_{\mathring{\mu}} = 0$ or $\frac{1-x/\alpha_{n-1}}{1-x/\alpha_n} f(x) \in \mathcal{L}_{n-2}$ or $f \in \frac{1-x/\alpha_n}{1-x/\alpha_{n-1}}\mathcal{L}_{n-2} = \mathcal{L}_{n-1}(\alpha_n)$. Therefore $Q_n(z)$ is orthogonal to $\mathcal{L}_{n-1} \cap \mathcal{L}_{n-1}(\alpha_n) = \mathcal{L}_{n-1}(\alpha_n)$.

So we have as a result:

**Theorem 5.9** *If $Q_n(z)$ has $n$ simple zeros in $\hat{\mathbb{R}}$, then these zeros can be used as nodes in an interpolating quadrature formula. This quadrature formula will be exact in $\mathcal{L}_{n-1,n-1}$ and the weights are positive.*

**Proof.** This is an immediate consequence of our results in the previous section. One may take $m = 1$. □

We now investigate this quadrature in a bit more detail and derive some alternative expressions for the quadrature weights. With some appropriate interpretation, the results will also apply to the case $\tau = 0$ discussed in the previous section.

Since everything is real, the numerator polynomial of $Q_n(x)$ will have real coefficients, all complex poles will appear as complex conjugate pairs (multiplicity included). We shall then prove that $Q_n(z)$ has $n$ simple zeros on $\hat{\mathbb{R}}$.

**Theorem 5.10** *Let the QORFs be given by $Q_n(x) = \frac{p_n(x)}{\pi_n^*(x)}$. Then $p_n(z)$ has $n$ simple zeros on $\hat{\mathbb{R}}$.*

**Proof.** Suppose that $x_1, \ldots, x_l$ are all the (real) zeros of $p_n(x)$ with odd multiplicities $m_i+1$, $i = 1, \ldots, l$. All the other zeros are taken in pairs $(\xi_j, \overline{\xi}_j)$, so that

$$p_n(x) = c\prod_{i=1}^{l}(x-x_i)^{2m_i+1}\prod_{j=1}^{k}(x-\xi_j)^2, \quad n = 2M+l+2k, \quad c \neq 0$$

$M = m_1 + \cdots + m_l$. We have to prove that $M = k = 0$, or equivalently since these are nonnegative that $M + k = 0$. The proof goes by contradiction. Suppose $M + k \geq 1$ and define

$$T(x) = \frac{(x-x_1)\cdots(x-x_l)(x-\alpha_n)}{(x-\alpha_1)\cdots(x-\alpha_{n-1})} \in \mathcal{L}_{n-1}.$$

We show that $\langle T, Q_n\rangle_{\mathring{\mu}} = \langle T, \phi_n\rangle_{\mathring{\mu}} + \tau \left\langle T, \frac{x-\alpha_{n-1}}{x-\alpha_n}\phi_{n-1}(x)\right\rangle_{\mathring{\mu}} = 0$. Clearly the first term $\langle T, \phi_n\rangle = 0$. Furthermore define $S(x) = \frac{x-\alpha_{n-1}}{x-\alpha_n}T(x) \in \mathcal{L}_{n-2}$ then $\langle S, \phi_{n-1}\rangle_{\mathring{\mu}} = 0$ or

$$\langle S, \phi_{n-1}\rangle_{\mathring{\mu}} = \left\langle T(x), \frac{x-\alpha_{n-1}}{x-\alpha_n}\phi_{n-1}(x)\right\rangle_{\mathring{\mu}} = 0.$$

However $\langle T, Q_n\rangle_{\mathring{\mu}} = \int |W(x)|^2 \mathrm{d}\mathring{\mu}(x) = c\,\langle W, W\rangle_{\mathring{\mu}} > 0$ because

$$W(x) = \frac{\prod_{i=1}^{l}(x-x_i)^{m_i+1}\prod_{j=1}^{k}(x-\xi_j)}{\pi^*_{n-1}(x)} \neq 0.$$

□

This result applies for all $\tau \in \hat{\mathbb{R}}$, thus also for $\tau = 0$ when $Q_n(x) = \phi_n(x)$. The numerators of $\phi_n$ have simple real zeros. The problem is however that it can not be assured that none of these zeros belongs to the set of poles $A_n = \{\alpha_i\}_{i=1}^n$, i.e., that the zeros of the numerator are also zeros of the rational functions.

We shall skip the proof, but it can be shown that if $\phi_n$ is regular (i.e., its numerator does not vanish at $\alpha_{n-1}$), then the numerators of $\phi_n$ and $\phi_{n-1}$ cannot have a common zero, and therefore the numerator of $Q_n(z)$ will only vanish at a given point $w \in \hat{\mathbb{R}}$ for exactly one value of $\tau \in \hat{\mathbb{R}}$. Consequently, there are at most $n$ values of $\tau$ for which a zero of the numerator of $Q_n(z)$ will coincide with one of the poles $\{\alpha_i\}_{i=1}^n$, which means that if $\phi_n$ is regular, then, except for at most $n$ values of $\tau$, the QORF $Q_n(z,\tau)$ has $n$ simple zeros on $\hat{\mathbb{R}}$. If $\tau$ is not one of these $n$ exceptional values we shall call it *regular* and the corresponding $Q_n(z)$ will be called *regular* too. As a particular case, we can say that if $\phi_n$ is regular and if $\tau = 0$ is a regular value, then $\phi_n$ will have $n$ simple zeros in $\hat{\mathbb{R}}$.

**Exercise 5.3** Find a relation between QORFs and reproducing kernels in analogy with the relation between QORFs and reproducing kernels for the circle. ◇

## 5.6 Interpolation

Because the poles can be in the support of the measure, we have to be a bit careful and we therefore assume that $\int |b_n(x)|^2 \mathrm{d}\mathring{\mu}(x) < \infty$ for all $n$. This implies for example that there is no mass in any of the poles $\mathcal{A}_n = \{\alpha_k\}_{k=1}^n$ or at $\infty$ and $\Omega_\mu^{(k)}$ will be continuous at $\alpha \in A_n$ for $k = 0, 1, \ldots \alpha^\# - 1$ where $\alpha^\#$ is the multiplicity of $\alpha$ in $\mathcal{A}_n$.

To define all the products $\langle b_k, b_l\rangle_{\mathring{\mu}}$, $k, l = 0, \ldots, n$ and hence the inner product in $\mathcal{L}_n$, we need more. All these are completely defined by the values or $\Omega_{\mathring{\mu}}^{(k)}(\alpha)$ with $\alpha \in \mathcal{A}_n \cup \{\mathbf{i}\}$ and $k = 0, \ldots, 2\alpha^\# - 1$.

The qf based on the quasi-orthogonal nodes is exact in $\mathcal{L}_{n-1}\cdot\mathcal{L}_{n-1}$. Denote by $\mathring{\mu}_n$ the discrete measure defined by the quadrature formula and set $\Omega_n = \Omega_{\mathring{\mu}_n}$. Then exactness implies that the values of $\Omega_{\mathring{\mu}}^{(k)}$ and of $\Omega_n^{(k)}$ at the poles in $A_n$ are the same. Thus that $\Omega_n$ is an Hermite interpolant for $\Omega_{\mathring{\mu}}$ in the points $\{\mathbf{i}, -\mathbf{i}, \alpha_1, \alpha_1, \alpha_2, \alpha_2, \ldots, \alpha_{n-1}, \alpha_{n-1}\}$ if the nodes are the zeros of $Q_n$. The interpolation in the point $\mathbf{i}$ is a consequence of $D(x, \mathbf{i}) = 1$ and thus $\int \mathrm{d}\mathring{\mu} = 1 = \Omega_{\mathring{\mu}}(\mathbf{i}) = \overline{\Omega_{\mathring{\mu}}(-\mathbf{i})}$. If $\tau = 0$ and thus $Q_n = \phi_n$ and, as will be shown below, $\Omega_n = -\psi_n/\phi_n$, then $\alpha_n$ is an extra interpolation point.

An explicit expression for $\Omega_n$ exists. Therefore we first introduce the QORFs of the second kind as[7]

$$P_n(z,\tau) = P_n(z) = \psi_n(z) + \tau\frac{Z_n(z)}{Z_{n-1}(z)}\psi_{n-1}(z) = \psi_n(z) + \tau\frac{z-\alpha_{n-1}}{z-\alpha_n}\psi_{n-1}(z).$$

It can be shown that

$$\frac{Z_{n-1}(z)}{Z_n(z)}P_n(z) = \int D(x,z)\left[\frac{Z_{n-1}(x)}{Z_n(x)}Q_n(x) - \frac{Z_{n-1}(z)}{Z_n(z)}Q_n(z)\right]\mathrm{d}\mathring{\mu}(x).$$

Because the integrand is in $\mathcal{L}_{n-1}$, our qf whose nodes $\xi_{nk}$ are zeros of $Q_n$ will integrate exactly, this is equal to $-[\sum_k D(\xi_{nk},z)\lambda_{nk}]\frac{Z_{n-1}(z)}{Z_n(z)}Q_n(z)$ and the sum is $\Omega_n(z) = \int D(x,z)\mathrm{d}\mathring{\mu}_n(x)$ with $\mathring{\mu}_n$ the discrete measure representing the qf. Thus we have shown that $\Omega_n(z) = -\frac{P_n(z)}{Q_n(z)}$ (with $\Omega_n(z) = -\frac{\psi_n(z)}{\phi_n(z)}$ if $\tau = 0$).

We are now ready to give alternative expressions for the weights of the qf.

**Theorem 5.11** *The weights of the interpolating qf whose nodes are the zeros $\{\xi_k\}$ of $Q_n(z)$ are given by*

$$\lambda_k = \lambda_k(\tau) = \frac{\mathbf{i}}{1+\xi_k^2}\frac{P_n(\xi_k)}{Q_n'(\xi_k)} = \frac{1}{\sum_{j=0}^{n-1}|\phi_j(\xi_k)|^2} > 0, \quad k = 1,\ldots,n.$$

**Proof.** The first formula is proved just like in the case of the circle, except that $D(z,w)$ has now a different form, leading to the different factor $\mathbf{i}/(1+\xi_k^2)$.
For the second formula, set $f_n = \phi_n/Z_n$, so that $Q_n(\xi_k) = 0$ implies $\tau = -f_n(\xi_k)/f_{n-1}(\xi_k)$. Then the CD relation gives

$$\frac{f_{n-1}(\xi_k)}{Z_n(x)}[-Q_n(x)] = \frac{f_{n-1}(\xi_k)}{Z_n(x)}[Q_n(\xi_k) - Q_n(x)] = E_n k_{n-1}(x,\xi_k)\frac{\xi_k - x}{\xi_k x}.$$

Multiply by $Z_n(x)Z_{n-1}(\xi_k)/(\xi_k - x)$ and let $x \to \xi_k$ to get

$$-\phi_{n-1}(\xi_k)Q_n'(\xi_k) = \frac{E_n k_{n-1}(\xi_k,\xi_k)}{(1-\xi_k/\alpha_n)(1-\xi_k/\alpha_{n-1})}.$$

Similarly, using the (complex conjugate of) second formula of Corollary 5.6, we can get

$$\frac{f_{n-1}(\xi_k)}{Z_n(x)}P_n(x) = E_n\left[\sum_{j=0}^{n-1}\psi_j(x)\phi_j(\xi_k) - D(\xi_k,x)\right]\frac{\xi_k - x}{\xi_k x}.$$

Multiply by $Z_n(x)Z_{n-1}(\xi_k)$ and let $x \to \xi_k$ to get

$$\phi_{n-1}(\xi_k)P_n(\xi_k) = E_n\frac{1+\xi_k^2}{\mathbf{i}(1-\xi_k/\alpha_n)(1-\xi_k/\alpha_{n-1})}.$$

A division then gives the result. □

[7]Again, it the coefficients of $\phi_k$'s are real, the coefficients of $\phi_k$'s and hence of $P_n$ are purely imaginary.

## 5.7 Quadrature with kernels

Like in the circle case, we can define an absolutely continuous measure

$$\mathrm{d}\mathring{\mu}_n(x) = \frac{\mathrm{d}\mathring{\lambda}(x)}{|K_n(x,0)|^2}, \quad \mathrm{d}\mathring{\lambda}(x) = \frac{\mathrm{d}x}{\pi(1+x^2)}, \quad K_n(x,\mathbf{i}) = \frac{k_n(x,\mathbf{i})}{\sqrt{k_n(\mathbf{i},\mathbf{i})}}$$

and it can be shown that the inner product in $\mathcal{L}_n$ with respect to $\mathring{\mu}$ and with respect to $\mathring{\mu}_n$ is the same. However, since we do not have a recurrence relation for the reproducing kernels $k_n$, the Carathéodory function for $\mathring{\mu}_n$ is yet unknown so that the corresponding interpolation properties can not be formulated.

**Question 5.2** Can a Carathéodory function be found in this case of the kernels. How much can be generalized when the second variable in the kernel is not $\mathbf{i}$ but a more general $w \in \mathbb{C}$? (see also [28].) ◯

## 5.8 The half-line and the interval

When the support of the measure is only the positive real half-line, then it makes sense to choose poles $\alpha_k$ that are on the negative real line. The limiting case where the poles are alternating between the boundaries 0 and $-\infty$ is a limiting case that is related to the *(strong) Stieltjes moment problem*.

The generalization is that some poles are chosen negative, but close to 0 and another set is negative but close to $-\infty$. The so called *balanced situation* occurs when the successive poles alternate between these two sets. If the two sets are separated and the measure has no mass at $+\infty$, the we are in the *separate balanced Stieltjes (SBS) situation*.

The general formulas still hold for SBS, but there will be no problem about zeros of the numerator of $\phi_n$ (all simple and on $(0,+\infty)$) canceling against poles (all on $(-\infty,0)$) since they are on separate half lines. We may also define *associate functions* $\sigma_n(z) = \int_0^\infty \frac{\phi_n(x)-\phi_n(z)}{x-z}\mathrm{d}\mu(x)$. With appropriate initial conditions, they are a solution of the same recurrence relation as the $\phi_n$ (set $\sigma_{-1} = -1$ and $\sigma_0 = 0$). Note that these are simplified versions of the functions of the second kind $\psi_k$ that we introduced before. This is made possible because there is no mass point at $+\infty$. Note also that the $\mathbf{i}$ factor is left out so that the Nevanlinna transform should also be replaced by the *Stieltjes transform* $\Omega(z) = \int_0^\infty \frac{\mathrm{d}\mu(x)}{x-z}$. The recurrence can be written as

$$\begin{bmatrix} \phi_n & \sigma_n \\ \phi_{n-1} & \sigma_{n-1} \end{bmatrix} = \begin{bmatrix} b_n & a_n \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi_{n-1} & \sigma_{n-1} \\ \phi_{n-2} & \sigma_{n-2} \end{bmatrix}, \quad \begin{bmatrix} \phi_0 & \sigma_0 \\ \phi_{-1} & \sigma_{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We have set $a_n = C_n Z_n/Z_{n-2}$ and $b_n = (E_n Z_{n-1} + B_n)Z_n/Z_{n-1}$.

This is the standard recurrence relation for numerator and denominator of a continued fraction and the convergents are $\frac{\sigma_n}{\phi_n}$. With appropriate sign normalization, it can be proved that

$$\frac{\sigma_n(x)}{\phi_n(x)} - \frac{\sigma_{n-1}(x)}{\phi_{n-1}(x)} = (-1)^n \frac{a_1(x)\cdots a_n(x)}{\phi_n(x)\phi_{n-1}(x)}.$$

Assuming that the poles are alternating between the points near $-\infty$ and 0 like $-\infty < \alpha_{2m} \leq \alpha^- < \alpha^+ \leq \alpha_{2m+1} < 0$, then it can be shown along these lines that on $(\alpha^-,\alpha^+)$, the sequence $\Omega_{2m}(x) = \frac{\sigma_{2m}(x)}{\phi_{2m}(x)}$ is monotonically decreasing and $\Omega_{2m+1}(x) = \frac{\sigma_{2m+1}(x)}{\phi_{2m+1}(x)}$ is monotonically

increasing. Moreover $\Omega_n$ interpolates $\Omega$ in Hermite sense in the points $\{\alpha_0, \alpha_1, \alpha_1, \alpha_2, \alpha_2, \ldots, \alpha_{n-1}, \alpha_{n-1}, \alpha_n\}$.

To approximate the integral $\int_0^\infty f(x)\mathrm{d}\mu(x)$, the rational Gauss quadrature formula whose nodes are the zeros of $\phi_n$ (simple and positive) and weights $1/\sum_{k=0}^{n-1}|\phi_k(x)|^2 > 0$ is exact in $\mathcal{L}_n \cdot \mathcal{L}_{n-1}$.

For more details see [36].

For the interval, say $I = [-1, 1]$, the poles can be chosen anywhere in $\hat{\mathbb{C}} \setminus I$. One way to deal with this situation is to consider the *Joukowski transform* $x = \mathfrak{j}(z) = (z + 1/z)/2$ mapping the open unit disk $\mathbb{D}$ to the cut Riemann sphere $\hat{\mathbb{C}}^I = \hat{\mathbb{C}} \setminus I$. The inverse $\mathfrak{j}^{inv}$ is chosen such that $\hat{\mathbb{C}} \setminus I$ is mapped onto $\mathbb{D}$. If $\mu$ has support on $I$, than $\mathring{\mu} = \mu \circ \mathfrak{j}$ will have support on $\mathbb{T}$. The $\alpha_k \in \hat{\mathbb{C}}^I$ are mapped onto $\beta_k = \mathfrak{j}^{inv}(\alpha_k) \in \mathbb{D}$. In this way one may transform the theory of rational Szegő quadrature on $\mathbb{T}$ into a theory of rational Gauss quadrature on $I$. See e.g. [15], [72]. Of particular interest of course are measures that are symmetric on $\mathbb{T}$, meaning that $\mu(t) = \mu(\bar{t})$. In that case, (assuming the poles are real) also the nodes of the qf (zeros of the PORFs which belong to $\mathbb{T}$) will be real or appear in complex conjugate pairs if $\tau$ is real (i.e., $\pm 1$) and hence the transformed nodes will be real (but outside $I$). However with $\tau = 1$ and $n$ odd, one of the circular nodes will be real (it is $-1$ in this case) which gives a boundary point on $I$ (it is a *Gauss-Radau* formula) and if $\tau = -1$ and $n$ is even, both end points $\pm 1$ will be nodes and we get a *Gauss-Lobatto* formula [68]. More about these will come later.

If one gives up on symmetry and allows arbitrary poles in $\hat{\mathbb{C}}^I$, the situation becomes more complicate as shown in [65].

Another possibility one might exploit is the mapping of $[-1, 1]$ to $[0, \infty]$ by setting $T(x) = \frac{1-x}{1+x} \in [-1, 1]$ for $x \in [0, \infty]$. Relations between the coefficients in the recurrence relations for the orthogonal functions on $[-1, 1]$ and $[0, \infty]$ and the corresponding quadrature formulas can be obtained. See [156].

## 5.9 More general situations

In the case of poles anywhere in the extended complex plane (except 0), a similar theory can be developed, but with extra complications. For example for orthogonality on an interval, the functions can be complex, and thus the inner product does need the complex conjugation $\langle f, g\rangle = \int_{-1}^{1} f(x)\overline{g(\overline{x})}\mathrm{d}\mu(x)$. There are extra conditions needed to guarantee the existence of the recurrence relation. For example, there are two forms of degeneracy when the numerator polynomial of $\phi_n$ vanishes at $\alpha_{n-1}$ or at $\overline{\alpha}_{n-1}$ and some problem may arise if that numerator does not have degree $n$. Although more complicated, most of the results can be generalized.

A 3-term recurrence relation has been proved in [65].

Also rational Gauss qf can be obtained using either QORFs (which are defined to be functions in $\mathcal{L}_n \setminus \mathcal{L}_{n-1}$ and that are orthogonal to $\mathcal{L}_{n-1}(\overline{\alpha}_{n-1})$), or on PORFs (defined as functions in $\mathcal{L}_n \setminus \mathcal{L}_{n-1}$ that can be written in the form $u_n\phi_n(x) + v_n\phi_n^*(x)$ with $u_n, v_n \in \mathbb{C}$ satisfying certain conditions and $\phi_n(x) = \frac{b_n(x)}{\overline{b_n(\overline{x})}}\overline{\phi_n(\overline{x})} = \frac{b_n(x)}{b_{n^*}(x)}\phi_{n^*}(x)$. In both cases, their zeros depend on a complex parameter $\tau$ that can be chosen in such a way that these zeros become real and can thus be used as nodes. They can be most easily obtained by solving a generalized eigenvalue problem that is much like the one considered in the next section. See [69].

# 6 Computational aspects

To compute the nodes and the weights of a quadrature formula, one could e.g., use Gram-Schmidt orthogonalization to construct the $n$th orthogonal function and compute its zeros. This gives you the nodes. Once these are known, the computation of the corresponding weights is relatively simple.

However, this technique is never used. It is now common practice to compute weights and nodes via an eigenvalue problem. So this chapter will deal mostly with linear algebra.

## 6.1 The real line

Suppose for simplicity that $\int \mathrm{d}\mu(t) = 1$. It is well known that if a sequence of orthonormal polynomials $\{\phi_k\}$ is generated by a recurrence relation

$$x\phi_n(x) = a_{n-1}\phi_{n-2}(x) + b_n\phi_{n-1}(x) + a_n\phi_n(x), \quad \phi_0 = 1, \quad \phi_{-1} = 0,$$

then this can be written in a matrix form as follows:

$$x\Phi_n(x) = \Phi_n(x)J_n + [0, 0, \ldots, 0, a_n\phi_n(x)],$$

$$\Phi_n(x) = [\phi_0, \phi_1(x), \ldots, \phi_{n-1}(x)], \quad J_n = \mathrm{tridiag}\begin{pmatrix} a_1, a_2, \ldots, a_{n-1} \\ b_1, b_2, \ldots, b_{n-1}, b_n \\ a_1, a_2, \ldots, a_{n-1} \end{pmatrix}$$

Thus, if $\xi$ is one of the zeros of $\phi_n$, then $\xi\Phi_n(\xi) = \Phi_n(\xi)J_n$, which means that $\xi$ is an eigenvalue of $J_n$ and $\Phi_n(\xi)$ is the corresponding eigenvector. Because the weight in $\xi$ is given by $1/\sum_{i=0}^{n} |\phi_i(\xi)|^2$, the weight is $\|\Phi_n(\xi)\|^2$. In practice, an eigenvector of $J_n$, can be any nonzero multiple of $\Phi_n(\xi)$ but its first component is $\phi_0 = 1$, so that we know how to normalize it.

Writing this down for $n = \infty$ as $x\Phi_\infty(x) = \Phi_\infty(x)J_\infty$, we can say that the infinite Jacobi matrix $J_\infty$ is a representation of the multiplication operator (i.e., multiply with $x$) with respect to the orthogonal polynomial basis $\Phi_\infty$.

For the rational case, we can rewrite the recurrence relation (*if it exists!*) as

$$E_n\phi_{n-1}(x) + B_n\frac{(1 - x/\alpha_{n-1})\phi_{n-1}(x)}{x} = \frac{(1 - x/\alpha_n)\phi_n(x)}{x} + \frac{E_n}{E_{n-1}}\frac{(1 - x/\alpha_{n-2})\phi_{n-2}(x)}{x}.$$

or (recall $\alpha_0 = \infty$)

$$x\Phi_n(x) = \Phi_n(x)\mathrm{diag}\left(1 - \frac{x}{\alpha_0}, \ldots, 1 - \frac{x}{\alpha_{n-1}}\right)J_n + \left[0, 0, \ldots, 0, \left(1 - \frac{x}{\alpha_n}\right)a_n\phi_n(x)\right]$$

$$\Phi_n(x) = [\phi_0, \phi_1(x), \ldots, \phi_{n-1}(x)], \quad J_n = \mathrm{tridiag}\begin{pmatrix} a_1, a_2, \ldots, a_{n-1} \\ b_1, b_2, \ldots, b_{n-1}, b_n \\ a_1, a_2, \ldots, a_{n-1} \end{pmatrix}$$

with $a_k = 1/E_k$ and $b_k = -B_k/E_k$. We define $D_n = \mathrm{diag}(\alpha_0, \ldots, \alpha_{n-1})$, and set $x = \xi$, which is one of the zeros of $\phi_n$, and we suppose that it is nonzero and not in $\{\alpha_0, \ldots, \alpha_{n-1}\}$, then it can be rewritten as

$$\xi\Phi_n(\xi)[I_n + D_n^{-1}J_n] = \Phi_n(\xi)J_n$$

which says that $\xi$ is a generalized eigenvalue of the pencil $(J_n, I_n + D_n^{-1}J_n)$ and $\Phi_n(\xi)$ is the corresponding eigenvector. The quadrature weight at $\xi$ can be written once more as $\|\Phi_n(\xi)\|_2^2$ where it is supposed that $\phi_0$, the first component of $\Phi_n(\xi)$, is 1.

Note that, if we write this again for $n = \infty$, we see that $J_\infty(I_\infty + D_\infty^{-1}J_\infty)^{-1}$ is the representation of the multiplication operator with respect to the rational orthogonal basis $\Phi_\infty$.

An ad hoc interpretation, (which could be turned into a hard argument) is as follows. If $J_\infty$ is the representation of the multiplication operator for the OP basis, then what has been done to move from the polynomial to the rational case is that the standard basis $1, x, x^2, \ldots$ is replaced by a product of factors of the form $\tilde{x} = x(1 - \alpha^{-1}x)^{-1}$ with inverse $x = \tilde{x}(1 + \alpha^{-1}\tilde{x})^{-1}$ where $\alpha \in \{\alpha_1, \alpha_2, \ldots\}$ and this is clearly reflected in the transformation of the representation. We shall work this out a bit further in the next subsection for the case of the circle.

**Question 6.1** This approach is based on the existence of the three-term recurrence relation. This only holds when the system is regular. How should this be done when there are singular $\phi_k$ in the sequence? ◯

**Remark 6.1** In the case of poles anywhere in the complex plane, a similar generalized eigenvalue problem can be considered. There are however extra conditions needed to guarantee the existence of the recurrence relation. Although more complicated, most of the results can be generalized. See for example [69, 152].

## 6.2 The unit circle

The recurrence in this case is not of 3-term type but it is of Szegő type. Let us start again with the polynomial case then the OP satisfy

$$c_n \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = \begin{bmatrix} 1 & a_n \\ \overline{a}_n & 1 \end{bmatrix} \begin{bmatrix} z\phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix}.$$

with $a_n = \phi_n(0)/\phi_n^*(0)$ and $c_n = \sqrt{1 - |a_n|^2}$. Clearly it is possible to express $z\phi_{n-1}(z)$ as a linear combination of $\phi_0, \phi_n, \ldots, \phi_n$. This means that the multiplication operator with respect to the orthogonal basis $\Phi_\infty$ is a full Hessenberg matrix. It is well known that that it is of the form

$$H_\infty = \begin{bmatrix} -a_1 & -c_1a_2 & -c_1c_2a_3 & -c_1c_2c_3a_4 & \cdots \\ c_1 & -\overline{a}_1a_2 & -\overline{a}_1c_2a_3 & -\overline{a}_1c_2c_3a_4 & \cdots \\ 0 & c_2 & -\overline{a}_2a_3 & -\overline{a}_2c_3a_4 & \cdots \\ 0 & 0 & c_3 & -\overline{a}_3a_4 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

Although it is a dense Hessenberg matrix, it has a lot of structure and can be characterized by only the vector $(a_k)$.

It turns out that sparsity is made much more explicit if one does not orthogonalize the basis $1, z, z^2, \ldots$, but extends it to the space spanned by the basis $1, z, z^{-1}, z^2, z^{-2}, \ldots$ and orthogonalize that basis instead. It is also the natural habitat of the reciprocal polynomials $\phi_k^*$ since it is not difficult to show that an orthogonal basis is given by

$$\chi_{2n}(z) = z^{-n}\varphi_{2n}^*(z), \qquad \chi_{2n+1}(z) = z^{-n}\varphi_{2n+1}(z), \qquad n \geq 0,$$

and the representation of the multiplication operator is the famous five-diagonal matrix CMV matrix (see [50, 49, 51, 52], [142], [168]).

$$C_\infty = \begin{pmatrix} -a_1 & -c_1a_2 & c_1c_2 & 0 & 0 & 0 & 0 & \cdots \\ c_1 & -\overline{a}_1a_2 & \overline{a}_1c_2 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -c_2a_3 & -\overline{a}_2a_3 & -c_3a_4 & c_3c_4 & 0 & 0 & \cdots \\ 0 & c_2c_3 & \overline{a}_2c_3 & -\overline{a}_3a_4 & \overline{a}_3c_4 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -c_4a_5 & -\overline{a}_4a_5 & -c_5a_6 & c_5c_6 & \cdots \\ 0 & 0 & 0 & c_4c_5 & \overline{a}_4c_5 & -\overline{a}_5a_6 & \overline{a}_5c_6 & \cdots \\ 0 & 0 & 0 & 0 & 0 & -c_6a_7 & -\overline{a}_6a_7 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{6.1}$$

This five-diagonal representation has a factorization $C_\infty = C_{o\infty}C_{e\infty}$ as a product of $2\times 2$ block-diagonal symmetric unitary matrices

$$C_{o\infty} = \begin{bmatrix} \Theta_1 & & & & \\ & \Theta_3 & & & \\ & & \Theta_5 & & \\ & & & \ddots & \end{bmatrix}, \quad C_{e\infty} = \begin{bmatrix} I_1 & & & & \\ & \Theta_2 & & & \\ & & \Theta_4 & & \\ & & & \ddots & \end{bmatrix} \text{ with } \Theta_k = \begin{bmatrix} -a_k & c_k \\ c_k & \bar{a}_k \end{bmatrix}.$$

Truncation of the infinite matrices leads to

$$z\Phi_n(z) = \Phi_n(z)H_n + [0, 0, \ldots, 0, c_n\phi_n(z)].$$

From the recurrence relation it is not difficult to see that the para-orthogonal polynomial $Q_n(z,\tau) = \phi_n(z) + \tau\phi_n^*(z)$ satisfies $c_nQ_n(z,\tau) = (1+\overline{a}_n\tau)[z\phi_{n-1}(z) + \tilde{\tau}\phi_{n-1}^*(z)]$ with $\tilde{\tau} = \tilde{\zeta}_{a_n}(\tau) = \frac{\tau+a_n}{1+\overline{a}_n\tau} \in \mathbb{T}$. Hence, we obtain the $Q_n(z,\tau)$, up to a constant multiple, by applying the recurrence with parameters $a_1, a_2, \ldots, a_{n-1}$ and then replace $a_n$ by $\tilde{\tau}$. The multiplicative constant is irrelevant because we are mainly interested in the zeros of $Q_n(z,\tau)$ which deliver the nodes of the quadrature. Let us call this modified Hessenberg matrix $H_n^\tau$ then $z\Phi_n(z) = \Phi_n(z)H_n^\tau + [0, 0, \ldots, 0, cQ_n(z,\tau)]$ and if $\xi_{nk}$ is a zero of $Q_n(z,\tau)$, then

$$\xi_{nk}\Phi_n(\xi_{nk}) = \Phi_n(\xi_{nk})H_n^\tau$$

with $\Phi_n(z) = [\phi_0, \phi_1, \ldots, \phi_{n-1}]$ and $H_n^\tau$ is an $n\times n$ matrix, which is a unitary truncation because $\tau \in \mathbb{T}$. The $\xi_{nk}$ are eigenvalues of $H_n^\tau$ and the eigenvectors are $\Phi_n(\xi_{nk})$.

We now return to the rational case. For more details, please consult [162, 10]. If we relax the normalization of the ORF a bit, we may rewrite the recurrence of theorem 4.7 as follows:

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = e_n\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z}\begin{bmatrix} 1 & a_n \\ \overline{a}_n & 1 \end{bmatrix}\begin{bmatrix} \zeta_{n-1}(z)\phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix}.$$

with

$$a_n = \frac{\phi_n(\alpha_{n-1})}{\phi_n^*(\alpha_{n-1})} \quad \text{and} \quad e_n = \sqrt{\frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2}\,\frac{1}{1-|a_n|^2}}.$$

Along the same lines as in the polynomial case, it can be shown that the representation of the multiplication operator in $\mathcal{L}_\infty$ with respect to the ORF basis is given by a transformation of the Hessenberg matrix, namely $\tilde{\zeta}_D(H_\infty)$ where $D = D_\infty$ is the diagonal matrix with entries the $\alpha_k$ and $\tilde{\zeta}_D(H) = \eta_D^{-1}(H+D)(I+D^HH)^{-1}\eta_{D^H}$ where the superscript $^H$ refers to the

adjoint and $\eta_D = \sqrt{I - DD^H}$. Again the transformation of the original basis is clearly reflected here. The polynomial factors $z$ are replaced by factors $\zeta = \frac{z-\alpha}{1-\overline{\alpha}z}$ whose inverse is $z = \tilde{\zeta}_\alpha(\zeta) = \frac{\zeta+\alpha}{1+\overline{\alpha}\zeta}$, $\alpha \in \{\alpha_1, \alpha_2, \ldots\}$.

For the nodes of the quadrature, we need the zeros of PORFs $Q_n(z,\tau) = \phi_n(z) + \tau\phi_n^*(z)$, $\tau \in \mathbb{T}$.

**Exercise 6.1** Use the recurrence relation to show that

$$Q_n(z,\tau) = \phi_n(z) + \tau\phi_n^*(z) = (1+\overline{a}_n\tau)e_n\frac{1-\overline{\alpha}_{n-1}z}{1-\overline{\alpha}_n z}[\zeta_{n-1}(z)\phi_{n-1}(z) + \tilde{\zeta}_{a_n}(\tau)\phi_{n-1}^*(z)]$$

where $\tilde{\zeta}_a(v) = \frac{v+a}{1+\overline{a}v}$ (a scalar version of the $\tilde{\zeta}$ used above). $\diamond$

In concrete terms, this means that we can compute $Q_n(z,\tau)$ by using the recurrence for the ORF with parameters $a_1, a_2, \ldots, a_{n-1}$, but use $\tilde{\tau} = \tilde{\zeta}_{a_n}(\tau) \in \mathbb{T}$ instead of $a_n$. As in the polynomial case we denote the corresponding Hessenberg matrix by $H_n^\tau$. Then the unitary truncation of $\tilde{\zeta}_D(H)$ is given by $V_n^\tau = \tilde{\zeta}_D(H_n^\tau)$. So we have this theorem.

**Theorem 6.1** *The zeros $\xi_{nk}$ of the PORF $Q_n(z,\tau)$ are the eigenvalues of $V_n^\tau = \tilde{\zeta}_D(H_n^\tau)$. Equivalently, the $\xi_{nk}$ are given as generalized eigenvalues of the pencil $(D_n + H_n^\tau, I_n + D_n^H H_n^\tau)$. The corresponding (left) eigenvectors are $\Phi_n(\xi_{nk})\eta_{D_n}^{-1/2}$.*

For the other representation of the multiplication operator, we need another basis in the rational case which is now given by orthogonalizing the following basis for $\mathcal{L}_* \cdot \mathcal{L}$. Let us therefore introduce the even and the odd Blaschke products:

$$B_0^e = B_0^o = 1, \quad B_n^o = \zeta_1\zeta_3\cdots\zeta_{2n-1}, \quad B_n^e = \zeta_2\zeta_4\cdots\zeta_{2n}, n \geq 1.$$

and consider the rational functions

$$\chi_{2n} = B_{n*}^e\phi_{2n}^*, \quad \chi_{2n+1} = B_{n*}^o\phi_{2n+1}^*, \quad n \geq 0,$$

then the subspaces $\mathcal{M} = \operatorname{span}\{\chi_0, \chi_1, \chi_2, \ldots\}$ are

$$\mathcal{M}_{2n} = B_{n*}^e\mathcal{L}_{2n} = \operatorname{span}\{B_{0*}^e, B_1^o, B_{1*}^e, \ldots, B_{(n-1)*}^e, B_n^o\}$$

and

$$\mathcal{M}_{2n+1} = B_{(n+1)*}^e\mathcal{L}_{2n+1} = \operatorname{span}\{B_{0*}^e, B_1^o, B_{1*}^e, \ldots, B_n^o, B_{n*}^e\},$$

i.e., $\mathcal{M}_{2n}, \mathcal{M}_{2n+1}$ are the sets of ORF whose poles are in $\{\hat{\alpha}_1, \alpha_2, \ldots, \alpha_{2n-2}, \hat{\alpha}_{2n-1}\}$ and $\{\hat{\alpha}_1, \alpha_2, \ldots, \hat{\alpha}_{2n-1}, \alpha_{2n}\}$ (repeated according to multiplicity). Moreover these are orthogonal, because the orthogonality of the $\chi_k$ is equivalent to the orthogonality of the $\phi_k$.

Note that $\mathcal{M}_{2n} = \mathcal{R}_{n,n} = \mathcal{L}_{n*} \cdot \mathcal{L}_n = \mathcal{L}_{n*} + \mathcal{L}_n$ and $\mathcal{M}_{2n+1} = \mathcal{R}_{n,n+1} = \mathcal{L}_{n*} \cdot \mathcal{L}_{n+1} = \mathcal{L}_{n*} + \mathcal{L}_{n+1}$ where we have used $\mathcal{L}_{k*} = \{f : f_* \in \mathcal{L}_k\}$. These spaces $\mathcal{R}_{p,q}$ are rational or multipoint generalizations of the Laurent polynomials.

With the basis $(\chi_k)_{k=0}^\infty$ for $\mathcal{L}_* \cdot \mathcal{L}$, which is dense[8] in $L_\mu^2$, we get again a simple matrix representation of the multiplication operator, given as an operator Möbius transform of the $C_\infty$ matrix, namely $\tilde{\zeta}_{D_\infty}(C_\infty)$, the same transformation used before.

With the same arguments used for the Hessenberg representation we can truncate and modify the last Schur parameter $a_n$ into $\tilde{\zeta}_{a_n}(\tau)$, $\tau \in \mathbb{T}$, so that the zeros of the quadrature formula are given as follows.

[8]This basis is dense if $\sum(1 - |\alpha_k|) = \infty$. This coincides with the rational moment problem being determinate.

**Theorem 6.2** *The zeros $\xi_{nk}$ of the PORF $Q_n(z,\tau)$ are the eigenvalues of $U_n^\tau = \tilde{\zeta}_{D_n}(C_n^\tau)$ with (right) eigenvectors $X_n(\xi_{nk})^H = [\chi_0(\xi_{nk}), \chi_1(\xi_{nk}), \ldots, \chi_{n-1}(\xi_{nk}]^H$.*
*Equivalently, these zeros are the generalized eigenvalues of the pencil $(D_n + C_n^\tau, I_n + D_n^H C_n^\tau)$ and the corresponding (left) eigenvectors are given by $X_n(\xi_{nk})\eta_{D_n}^{-1}$.*

Note that with the even-odd factorization of $C$, we can rewrite the previous eigenvalue problem. Because $U_n^\tau$ is unitary, and for any unitary matrix it holds that $U = [U^H]^{-1}$ so that

$$U_n^\tau = \tilde{\zeta}_{D_n}(C_n^\tau) = \eta_{D_n}([C_n^\tau]^H + D_n^H)^{-1}(I_n + [C_n^\tau]^H D_n)\eta_{D_n}^{-1}.$$

This implies that the eigenvalue problem $U_n^\tau E_n = E_n\Lambda_n$ is also

$$\begin{aligned}
\eta_{D_n}((C_n^\tau)^H + D_n^H)^{-1}(I_n + (C_n^\tau)^H D_n)\eta_{D_n}^{-1}E_n &= E_n\Lambda_n \\
((C_n^\tau)^H + D_n^H)^{-1}(I_n + (C_n^\tau)^H D_n)F_n &= F_n\Lambda_n, \quad F_n = \eta_{D_n}^{-1}E_n \\
((C_{on}^\tau)^H + C_{en}^\tau D_n^H)^{-1}(C_{en}^\tau + (C_{on}^\tau)^H D_n)F_n &= F_n\Lambda_n,
\end{aligned}$$

since also $C_{on}^\tau$ and $C_{en}^\tau$ are unitary truncations of the matrices $C_{o\infty}$ and $C_{e\infty}$. This means that we also have the following.

**Theorem 6.3** *The $\xi_{nk}$ of the PORF are given as generalized eigenvalues of the pencil $(C_{en}^\tau + (C_{on}^\tau)^H D_n, (C_{on}^\tau)^H + C_{en}^\tau D_n^H)$. The corresponding (right) eigenvectors are $X_n(\xi_{nk}) = \{[\chi_0(\xi_{nk}), \chi_1(\xi_{nk}), \ldots, \chi_{n-1}(\xi_{nk})]\eta_{D_n}^{-1}\}^H$.*

The previous theorems say how to compute the nodes of the rational Szegő qf. The corresponding weights $\lambda_{nk}$ are given (as in the polynomial case) by the values that we had before $\lambda_{nk} = 1/\sum_{j=0}^{n-1}|\phi_j(\xi_{nk})|^2$ where $\phi_k$ are the orthonormal ORF. This is the inverse of the squared norm of the eigenvectors in the previous theorems. Since these eigenvectors $E_{nk}$ are only defined up to a multiplicative constant, we can choose a normalized eigenvector $F_{nk} = E_{nk}/\|E_{nk}\|$ and for example with the eigenvector $\Phi_n(\xi_{nk})$ we should have $F_{nk} = c\Phi_n(\xi_{nk})$ with $c = \|\Phi_n(\xi_{nk})\|^{-1} = \lambda_{nk}^{1/2}$. Because $\phi_0$ was supposed to be 1 by our normalization of the measure, it follows that $\lambda_{nk}$ is the square of the first element in the normalized eigenvector. For the cases with the ORF $\chi_k$ it suffices to note that $\sum_{j=0}^{n-1}|\chi_j(\xi_{nk})|^2 = \sum_{j=0}^{n-1}|\phi_j(\xi_{nk})|^2$ to come to the same conclusion.

**Question 6.2** In a situation where the measure on $\mathbb{T}$ is symmetric (meaning that $\mathrm{d}\mu(e^{\mathbf{i}\theta}) = \mathrm{d}\mu(e^{-\mathbf{i}\theta})$) and if also the poles are chosen in a symmetric way, (meaning that they are either real on complex conjugate), then it might be possible to divide the complexity of the computations by a factor 2. See for example [13] for the polynomial case. Or possibly complex arithmetic can be replaced by real arithmetic by introducing the complex conjugate pairs in $2 \times 2$ blocks, i.e., working with a Hessenberg or a CMV matrix matrix having $2 \times 2$ blocks. ◯

# 7 Rational Chebyshev quadrature

Unlike in the polynomial case, there are almost no explicit forms known for orthogonal rational functions or for their recurrence coefficients.

One such an exception are the ORF on $\mathbb{T}$ orthogonal with respect to the Lebesgue measure. These are called by Walsh [167] Takenaka-Malmquist functions (after [148] and [120]). The orthonormal functions are given by

$$B_0 = 1, \quad \frac{\sqrt{1-|\alpha_k|^2}}{1-\overline{\alpha}_k z}\tilde{B}_{k-1}(z), \quad k \geq 1, \text{ with } \tilde{B}_k = \zeta_0 B_k = \prod_{j=0}^{k} \zeta_k, \quad \zeta_0(z) = z.$$

## 7.1 Rational Gauss-Chebyshev quadrature

On the real line, such ORF are less well known, except of course for the special case of polynomials. There is however the possibility to transform the previous basis to an orthogonal basis on the interval $I = [-1, 1]$ as we shall currently do.

The idea is to consider ORF on $I$ with poles $A = \{\alpha_1, \alpha_2, \ldots\} \subset \hat{\mathbb{C}}^I$ where $\hat{\mathbb{C}}^I = \hat{\mathbb{C}} \setminus I = (\mathbb{C} \cup \{\infty\}) \setminus [-1, 1]$ denotes the cut sphere. These are obtained by orthogonalizing the usual basis

$$b_0 = 1, \quad b_k = \prod_{j=1}^{k} Z_j, \quad Z_j(x) = \frac{x}{1 - x/\alpha_j}, \quad j = 1, \ldots, k, \quad k \geq 1.$$

The *Joukowski transform* $x = \mathfrak{j}(z) = (z + 1/z)/2$ maps the open unit disk $\mathbb{D}$ onto $\hat{\mathbb{C}}^I$ and $\mathbb{T}$ onto $I$. The inverse mapping is $z = \mathfrak{j}^{inv}(x) = x \pm \sqrt{x^2 - 1}$ where the square root is chosen so that $x \in \hat{\mathbb{C}}^I$ is mapped onto $z \in \mathbb{D}$. A weight $w$ on $I$ is mapped onto a weight $\mathring{w}(\theta) = w(\cos\theta)|\sin\theta|$ on $[-\pi, \pi]$ so that $\int_{-1}^{1} f(x)w(x)\mathrm{d}x = \frac{1}{2}\int_{-\pi}^{\pi} f(\cos\theta)\mathring{w}(\theta)\mathrm{d}\theta$.

If $A = \{\alpha_1, \alpha_2, \ldots\}$ is a set of poles in $\hat{\mathbb{C}}^I$, then these correspond to a set $B = \{\beta_1, \beta_2, \ldots\} \subset \mathbb{D}$ where $\beta_k = \mathfrak{j}^{inv}(\alpha_k)$. With these poles in $B$, we construct the Blaschke products

$$B_0 = 1, \quad B_k = \prod_{j=1}^{k} \zeta_j, \quad \zeta_j(z) = \frac{z - \beta_j}{1 - \overline{\beta}_j z}, \quad j = 1, \ldots, k, \quad k \geq 1.$$

In this section we define $f_{\bar{*}}(z) = \overline{f(\overline{z})}$ and

$$f^{\bar{*}}(z) = \frac{b_n(z)}{b_{n\bar{*}}(z)} f_{\bar{*}}(z) \quad \text{for } f \in \mathcal{L}_n = \operatorname{span}\{b_0, b_1, \ldots, b_n\}.$$

This means that the net result of the operation $f^{\bar{*}}$ is that only the numerator coefficients are replaced by their complex conjugates. It implies that $\mathcal{L}_n^{\bar{*}} = \mathcal{L}_n$.

**Exercise 7.1** Show that $Z_k(\mathfrak{j}(z)) = E\zeta_{k\bar{*}}(z) + F + E/\zeta_k(z)$ where $E$ and $F$ are constants depending on $\beta_k$. ◇

Let us now consider the following weights on $I$.

$$w^{(1)}(x) = (1 - x^2)^{-1/2}, \quad w^{(3)}(x) = \left(\frac{1-x}{1+x}\right)^{1/2}, \; `w^{(2)}(x) = (1 - x^2)^{1/2}$$

The Joukowski transform maps a weight $w$ on $I$ onto a weight $\mathring{w}(\theta) = w(\cos\theta)|\sin\theta|$ on $[-\pi,\pi]$ so that $\int_{-1}^{1} f(x)w(x)\mathrm{d}x = \frac{1}{2}\int_{-\pi}^{\pi} f(\cos\theta)\mathring{w}(\theta)\mathrm{d}\theta$. This transformation of the weight is also sometimes called the *Szegő transformation*.

In the polynomial case, this idea is pretty old and can be found in Szegő's book [146, p. 31] in the case that all $\alpha_k = \infty$ and $w(x) = w(\cos\theta) = |g(e^{\mathbf{i}\theta})|^2$ is a positive polynomial on $I$. Then Szegő gives a relation between the orthogonal polynomials with respect to $w$ and those with respect to $w^{(i)}/w$. These are the so-called *Bernstein-Szegő measures* and the corresponding OPs are the *Bernstein-Szegő polynomials*.

**Exercise 7.2** With $\alpha_k \in \hat{\mathbb{R}}^I$, $\zeta_k(z) = \frac{z-\beta_k}{1-\overline{\beta}_k z}$, $x = \mathfrak{j}(z)$, i.e., $x = \cos\theta = \mathfrak{j}(e^{\mathbf{i}\theta})$, $z = e^{\mathbf{i}\theta}$ and $\alpha_k = \mathfrak{j}(\beta_k)$, prove that $\zeta_k(z) = \exp\left\{2\mathbf{i}\arctan\frac{\sin\theta}{\cos\theta-\beta_k} - \mathbf{i}\theta\right\}$. ◇

Considering the numerators of our rational functions as the polynomials and assimilating the denominators in the weights, we can use Szegő's theory to find explicit expressions for our ORF with respect to the $w^{(i)}$ by passing to the unit circle. We skip the proofs and formulate the result.

With the different weights we may associate the following parameters

| $i$ | $c$ | $d$ | $p$ | $q$ |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | $-1$ |
| 2 | 3/2 | 0 | 1 | 1 |
| 3 | 2 | 0 | 2 | 1 |

so that we have the following theorem [73, Thm 3.2]

**Theorem 7.1** *With a normalizing factor $N = \sqrt{2^i/\pi}\sqrt{1-|\beta_n|^2}$, the orthonormal basis function in $\mathcal{L}_n \setminus \mathcal{L}_{n-1}$ is given by*

$$\phi_n(x) = \frac{qN}{2z^{i-1}+q-3}\left(\frac{z^i B_{(n-1)^{\bar{*}}}(z)}{1-\beta_n z} - \frac{q}{(z-\beta_n)B_{n-1}(z)}\right), \quad n \geq 1,$$

*and $\phi_0 = \sqrt{p/\pi}$ where $x = \mathfrak{j}(z)$ and $\alpha_k = \mathfrak{j}(\beta_k)$.*

To construct the corresponding quadrature formula (which is called a *rational Gauss-Chebyshev qf*) we proceed in principle as before: i.e., find the zeros of the ORF $\phi_n$ or possibly of the PORF $Q_n(x,\tau) = \phi_n(x) + \tau\phi_n^*(x)$, $\tau \in \mathbb{T}$.

For the general setting of poles in $\hat{\mathbb{C}}^I$, this is further elaborated in [73]. It simplifies a bit when the poles are supposed to be real and hence in $\hat{\mathbb{R}}^I$. More details in this special case are found in [157]. The expression for the $\phi_n$ is as in the previous theorem, but with the sub-star in $B_{(n-1)^{\bar{*}}}$ can be removed. In the latter situation of real poles, letting the zeros $\xi_{nk}$ of $\phi_n$ be denoted as $\xi_{nk} = \cos\theta_{nk}$, then these can be obtained as follows [157, Thm. 4.1,Thm. 4.3].

**Theorem 7.2** *With the notation just introduced, then the $\theta_{nk}$ for the ith Chebyshev weight and for real poles in $\hat{\mathbb{R}}^I$ are solutions of*

$$f_n(\theta) - (n-c)\theta = \frac{\pi}{2}(2k-d), \quad f_n(\theta) = 2\sum_{j=1}^{n}{}' \arctan\frac{\sin\theta}{\cos\theta-\beta_j}\ .$$

*($\sum'$ means last term halved.) The weight $\lambda = \lambda_{nk}$ corresponding to a node $x = x_{nk}$ is given by*

$$\lambda = 2\pi(1-x^{i-1})(i+g_n(x))^{-1}, \quad g_n(x) = 2\sum_{k=1}^{n}{}' \frac{\sqrt{1-1/\alpha_k^2}}{1-x/\alpha_k}.$$

($\arctan(y/x)$ refers to the argument of the complex number $x + \mathbf{i}y$.) Note that the equation for the nodes looks complicated, but efficient algorithms can be designed because in between the poles, the functions $f_n$ are monotonically increasing. To compute the nodes of an $n$-point formula requires only $O(n^2)$ computations if all the poles are different. If there are only $m < n$ different nodes, this reduces to $O(mn)$ computations, and thus if there is only a finite, fixed number of poles that are repeatedly used, the complexity is only $O(n)$ as $n \to \infty$.

For the more general case of complex poles in $\hat{\mathbb{C}}^I$, it can not be guaranteed that all the nodes are in $I$ [71], and therefore one has to use the PORF, for which it can be proved that under some restrictions, its zeros are simple and in $I$. The computation of the nodes and weights is very similar to what has been discussed above. Of course one needs to introduce the parameter $\tau \in \mathbb{T}$ as well with some restrictions. We refer to [73] for more details and to [74] for more on the algorithmic aspects.

## 7.2 Rational Szegő quadrature with Chebyshev weights

We consider again the relation between a weight function $w$ on $I = [-1, 1]$ and a weight function $\mathring{w} = w(\cos\theta)|\sin\theta|$ which can also be considered as a weight function on $\mathbb{T}$ if we set $z = e^{\mathbf{i}\theta}$. For the *Jacobi weights* $w(x) = (1-x)^\alpha(1+x)^\beta$ corresponding to $\mathring{w}(\theta) = (1-\cos\theta)^a(1+\cos\theta)^b$, $a = \alpha + \frac{1}{2}$, $b = \beta + \frac{1}{2}$, and choosing $a, b \in \{0, 1\}$, we obtain the weight functions

$$\omega^{(1)}(\theta) = 1, \quad \omega^{(2)}(\theta) = \sin^2\theta, \quad \omega^{(3)}(\theta) = 1 + \cos\theta, \quad \omega^{(4)}(\theta) = 1 - \cos\theta.$$

As we have mentioned in the beginning of this section on rational Chebyshev quadrature, the ORF with respect to $\omega^{(1)}$ is given by the Takenaka-Malmquist basis

$$\phi_0 = 1, \quad \phi_n(z) = \sqrt{1-|\alpha_n|^2}\frac{zB_{n-1}(z)}{1-\overline{\alpha}_n z}, \quad n \geq 1.$$

**Exercise 7.3** Prove that the Takenaka-Malmquist functions are ORFs on $\mathbb{T}$ with respect to the Lebesgue measure. ◇

For the weight functions $\omega^{(i)}$, $i = 3, 4$, also explicit expressions for the ORF are known. We have [159, 72], [12, Thm. 3.1].

**Theorem 7.3** *Define $N_n(z) = 1 + zB'_{n-1}(z)/B_{n-1}(z)$ where $B_{n-1}$ is a Blaschke product. For $z = t \in \mathbb{T}$, this reduces to $N_n(t) = 1 + \sum_{k=1}^{n-1} \frac{1-|\alpha_k|}{|z-\alpha_k|^2}$. With weight function $\omega^{(i)}$, $i \in \{3, 4\}$, i.e., $\omega^{(i)}(\theta) = 1 + \nu_i \cos\theta$ with $\nu_i = (-1)^i$, then the ORF (not normalized) are given by*

$$\phi_0 = 1, \quad \phi_n(z) = \frac{1}{(z-\nu_i)^2}\left[a_n^{(i)} + z^2(z - b_n^{(i)})\frac{B_{n-1}(z)}{1-\overline{\alpha}_n z}\right], \quad n \geq 1$$

*where*

$$a_n^{(i)} = \frac{\nu_i B_{n-1}(\nu_i)}{(1-\nu_i\overline{\alpha}_n)N_n(\nu_i) + 1}, \quad b_n^{(i)} = \nu_i + \frac{1-\nu_i\overline{\alpha}_n}{B_{n-1}(\nu_i)}a_n^{(i)}.$$

Note that the factor $(z-\nu_i)^2$ in the denominator of $\phi_n$ cancels against the same factor in the numerator of the expression within the square brackets.

The case of weight function $\omega^{(2)}(\theta) = \sin^2\theta$ is much more complicated and is discussed in the paper [12].

**Exercise 7.4** Prove that $\omega^{(2)}(\theta) = \sin^2\theta = \frac{1}{2}|z+1|^2\omega^{(3)}(\theta), \quad z = e^{\mathrm{i}\theta}$. ◇

Thus $\omega^{(2)}$ can be considered as a rational modification of $\omega^{(3)}$. The construction of ORF for rational modifications will be considered below. For completeness, we give an indication of what the ORF for $\omega^{(2)}$ look like [12, Thm. 3.2].

**Theorem 7.4** *The ORF (not normalized) for $\omega^{(2)}$ have the following form*

$$\phi_0 = 1, \quad \phi_n(z) = \frac{1}{(z-1)^2}\left[d_n + e_n z + z^3(f_n + g_n z + z^2)\frac{B_{n-1}(z)}{1-\overline{\alpha}_n z}\right], \quad n \geq 1$$

*where the constants $d_n, e_n, f_n, g_n$ are the solution of a $4\times 4$ linear system of equations and $d_n \pm e_n \neq 0$.*

Again the factor $(z-\nu_i)^2$ in the denominator cancels against the same factor in the numerator of the expression within the square brackets.

To compute the nodes of the rational Szegő qf, we have to find the zeros of the PORF $Q_n(\cdot,\tau_n) = \phi_n + \tau_n\phi_n^*$. Using the explicit expressions for the $\phi_n$, it is possible to reformulate the problem such that the nodes are the zeros of a polynomial equation.

**Exercise 7.5** Use the (confluent) Christoffel-Darboux relations and the known formula $\lambda = 1/\sum_{k=0}^{n-1}|\phi_j(\xi)|^2$ for the weight corresponding to node $\xi$ in the $n$-point rational Szegő qf to come to the alternative expression

$$\lambda^{-1} = \frac{\xi|\xi-\alpha_n|}{1-|\alpha_n|^2}\begin{vmatrix} \overline{\phi_n(\xi)} & \overline{\phi_n^*(\xi)} \\ \phi_n^{*\prime}(\xi) & \phi_n'(\xi)\end{vmatrix}.$$

A prime denotes derivative. Note that the $\phi_j$ are here the *normalized* ORFs. ◇

Once the nodes are known, the previous formula allows (with a lot of tedious arithmetic) to find explicit expressions for the weights. For example, in the case of $\omega^{(1)}$, the weight corresponding to node $\xi$ is $\lambda = \frac{2\pi}{N_n(\xi)}$ where $N_n(z)$ is defined in Theorem 7.3. All the details for the different weights $w^{(i)}$, $i = 1,2,3,4$ are given in [12].

## 7.3 Rational modifications of a measure on $\mathbb{T}$

Some of the results above rely on the fact that once we know the ORF for a weight $w$ on $\mathbb{T}$, then one may derive the ORF corresponding to $\tilde{w}(z) = |r(z)|^2 w(z)$ where $r$ is a rational function (usually of low degree). In the polynomial case this was discussed in [99, 98]. Also [87, 90, 88, 89, 86, 53]. There are in fact two possibilities the so-called *Christoffel* ($r(z) = |z-\alpha|$) and *Geronimus* ($r(z) = 1/|z-\alpha|$) transformations. Often a third transformation is also considered namely $\mathrm{d}\tilde{w}(z) = \mathrm{d}w(z) + m\delta(z-\alpha)$, $\alpha \in \mathbb{T}$, which is known as *Uvarov transformation*. The reason why the latter fits in is the following. With the Christoffel transform one has $\langle f,g\rangle_{\tilde{w}} = \langle (z-\alpha)f(z), (z-\alpha)g(z)\rangle_w$. With the Geronimus transformation $\langle f,g\rangle_{\tilde{w}} = \left\langle \frac{f(z)}{z-\alpha}, \frac{g(z)}{z-\alpha}\right\rangle_w$, while for Uvarov $\langle f,g\rangle = \langle f,g\rangle + m\overline{f(\alpha)}g(\alpha)$.

Rather than formulating the general situation, we shall look at a particular case. We have [63, Thm. 5].

**Theorem 7.5** *Let $A = \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{D}$ and $C = \{\beta, \alpha_1, \ldots \alpha_n\} \subset \mathbb{D}$, and $\gamma \in \hat{\mathbb{D}} = \mathbb{D} \cup \mathbb{T}$. Set $\mathrm{d}\tilde{\mu}(z) = \left|\frac{z-\gamma}{1-\overline{\beta}z}\right|^2 \mathrm{d}\mu(z)$. Let $\phi_n(z;\tilde{\mu})$ be the ORF for the sequence $A$ and measure $\tilde{\mu}$ and let $\tilde{\phi}_{n+1}(z;\mu)$ be the ORF for the sequence $C$ and the measure $\mu$. Then for $n \geq 1$*

$$\left(\frac{z-\gamma}{1-\overline{\beta}z}\right)\phi_n(z;\tilde{\mu}) = \overline{\eta}_\beta \frac{1-\gamma\overline{\alpha}_n}{1-\beta\overline{\alpha}_n}\left[\tilde{\phi}_{n+1}(z;\mu) - \frac{\tilde{\phi}_{n+1}(\gamma;\mu)}{\tilde{k}_n(\gamma,\gamma;\mu)}\tilde{k}_n(z,\gamma;\mu)\right]$$

*where $\tilde{k}_n(z,y;\mu)$ is the reproducing kernel for $\tilde{\mathcal{L}}_n$ based on poles $C$ and measure $\mu$. $\eta_\beta \in \mathbb{T}$ is a normalizing factor: $\eta_\beta = \overline{\beta}/|\beta|$ if $\beta \neq 0$ and it is 1 otherwise.*

In [63, Thm. 6,7] this is specialized to the situations $\gamma \in \mathbb{T}$ and $\gamma \in \mathbb{D}$ respectively, where $\tilde{k}_n$ is then replaced by expressions derived from the Christoffel-Darboux relation.

Note that for $\gamma = 0$ and $\beta = re^{\mathbf{i}\theta}$, $r < 1$ and $\mathrm{d}\mu = \mathrm{d}\lambda$, the Lebesgue measure, then $\mathrm{d}\tilde{\mu}(z) = \frac{\mathrm{d}\lambda(z)}{1-2r\cos\theta+r^2}$ is proportional to the *Poisson weight* which was fist discussed e.g., by Waadeland [163] in connection with Szegő quadrature formulas.

It is often possible to directly connect a weight $\lambda$ corresponding to a node $\xi$ of the Rational Szegő qf with respect to measure $\mu$ and the corresponding weight for the measure $\tilde{\mu}$.

**Exercise 7.6** Prove that the ORF for the Poisson measure $\mathrm{d}\mathring{\mu}(z) = \frac{(1-|r|^2)\mathrm{d}z}{2\pi\mathbf{i}(z-r)(1-\overline{r}z)}$, $r \in \mathbb{T}$ are given by $\phi_n(z) = k_n\frac{(r-z)B_n(z)}{z-\alpha_n}$.
Check the special cases $r = 0$, $r = \alpha_n$, and the polynomial case: $r \neq 0$ but all $\alpha_k = 0$.
Find a polynomial whose zeros are the nodes of the (rational) Szegő qf. Check the special cases. See [19]. ◇

**Question 7.1** Rational modifications of the measure have been studied by F. Marcellán and his coworkers in a series of papers but only dealing with orthogonal polynomials. The computational procedure is a very elegant in terms of linear algebra. For example if $\tilde{\mu} = |z-\alpha|^2\mu$ is a measure on $\mathbb{T}$ (Christoffel transform) and if the corresponding Hessenberg matrices for the orthogonal polynomials are $\tilde{H}$ and $H$ respectively, then the transform from $H$ to $\tilde{H}$ is given by a QR step: $H - \alpha I = QR$ then $\tilde{H} - \alpha I = RQ$ as observed in [100, 94, 168]. How much of this type linear algebra techniques for these spectral transformations can be generalized to the rational case in a stable way? See also [105] for a stable implementation for the Hessenberg when it is stored in its factored form. ◯

## 7.4 Near best rational approximation

It is well known that for polynomial $L_\infty$ (minimax or best) approximation on $I = [-1, 1]$ one needs equi-oscillation of the error. Since the error for a polynomial $p_n$ interpolating in the points $x_0, \ldots, x_n$ is given by $|f(x) - p_n(x)| \leq \max_{t\in I} |f^{(n+1)}(t)|\frac{|x-x_0|\cdots|x-x_n|}{(n+1)!}$, a good approximate choice is to choose the interpolation points as the zeros of $T_{n+1}$, the Chebyshev polynomial of the first kind, since indeed, this polynomial is equi-oscillating. In some applications it is necessary to include the end points, in which case one uses the extrema of $T_n$, which are the zeros of $U_{n-1}$, the Chebyshev polynomial of the second kind and add the end points 1 and $-1$.

This can be generalized to the rational case as well. As before we define $\pi_m(x) = \prod_{k=1}^m (1 - x/\alpha_k)$ with $\alpha_k \in \hat{\mathbb{C}}^I$, then a rational interpolant in $\frac{p_n}{\pi_m}$ in the points $\{x_0, \ldots, x_n\} \subset I$

is obtained by making $p_n$ a polynomial interpolant of $\pi_m f$. Then the error satisfies

$$\left| f(x) - \frac{p_n(x)}{\pi_m(x)} \right| \leq \max_{t\in I} \left| [\pi_m(t) f(t)]^{(n+1)} \right| \frac{|x-x_0|\cdots|x-x_n|}{|\pi_m(x)|(n+1)!}.$$

Minimizing $\max_{x\in I} |(x-x_0)\cdots(x-x_n)/\pi_m(x)|$ means that the $x_k$ are the zeros of the rational Chebyshev ORF. The Chebyshev ORF for the weight $w^{(1)}(x) = (1-x^2)^{-1/2}$ and for $w^{(3)}(x) = (1-x^2)^{1/2}$ are given above. With $\alpha_k = \mathfrak{j}(\beta_k)$ and $x = \mathfrak{j}(z) = \frac{1}{2}(z+1/z)$, they are given by

$$\mathcal{T}_n(x) = \sqrt{\frac{1-|\beta_n|^2}{2\pi}} \left( \frac{z B_{(n-1)^\ddagger}(z)}{1-\beta_n z} + \frac{1}{(z-\beta_n)B_{n-1}(z)} \right) \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$$

and

$$\mathcal{U}_n(x) = \sqrt{\frac{1-|\beta_n|^2}{\pi}} \frac{z\sqrt{2}}{z^2-1} \left( \frac{z^2 B_{(n-1)^\ddagger}(z)}{1-\beta_n z} - \frac{1}{z(z-\beta_n)B_{n-1}(z)} \right) \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$$

for $n = 1, 2, \ldots$, $\mathcal{T}_0 = 1/\sqrt{\pi}$, $\mathcal{U}_0 = \sqrt{2/\pi}$.

Now suppose that we use $m$ finite poles $\alpha_j$, $j = 1, \ldots, m$ that are either real or appear in complex conjugate pairs and set $\alpha_k = \infty$ for $k > m$, then the zeros of $\mathcal{T}_n$ are real and simple and inside $I$ for $n > m$ and the extremes of $\mathcal{T}_n$ are the zeros of $\mathcal{U}_{n-1}$.

**Exercise 7.7** Prove that with the conditions just introduced,

$$\mathcal{T}_n'(x) = g_n(z)\mathcal{U}_{n-1}(x), \quad g_n(z) = n - m - \sum_{k=1}^{m} \frac{1-\beta_k^2}{2\beta_k} \frac{1}{x-\alpha_k}.$$

◇

Once this is known, we can apply either the algorithm based on Newton iteration described in [158] to find the zeros of $\mathcal{T}_n$ or $\mathcal{U}_{n-1}$. Alternatively one may use the eigenvalue routine described above. Note however that if non-real, but complex conjugate poles are used, the recurrence has to be slightly adapted, giving also rise to a slightly modified eigenvalue problem. It takes now the form $x\Phi_n(x)(I_n - S_n + D_n^{-1}J_n) = \Phi_n(x)J_n$, where and extra matrix $S_n$ appears, that has only the first superdiagonal nonzero, containing elements proportional to $1/Z_{(n-2)^\ddagger}(\alpha_{n-2})$. Also, to construct the rational interpolant, it is best to represent it in barycentric coordinates. This means that, if the interpolation points are $\{x_j : j = 0, n\}$ then we write it as $p_n/\pi_m$ with $p_n(x) = \sum_{j=0}^{n} w_j f(x_j) \prod_{k\neq j}(x - x_k)$ where $w_j = 1/t'(x_j)$ with $t$ the Chebyshev ORF delivering the interpolation points.

**Exercise 7.8** Suppose that the nodes $x_j$ are the zeros of $\mathcal{T}_{n+1}(x)$ with $m \leq n$ then prove that $\mathcal{T}_{n+1}'(x_j) = \frac{\pm g_{n+1}(z_j)}{\sin\theta_j}$ with $x_j = \cos\theta_j$, $z_j = e^{\mathbf{i}\theta_j}$ and $g_{n+1}(z_j) = n+1-m-\sum_{k=1}^{m} \frac{1-\beta_k^2}{2\beta_k(x_j-\alpha_k)}$ so that $w_j = (-1)^j \frac{\sin\theta_j}{g_{n+1}(z_j)}$ with $g_{n+1}(z_j) = \pi/\lambda_j$. Here $\lambda_j = 1/\sum_{k=0}^{n} |\mathcal{T}_k(x_j)|^2$ is the quadrature weight. ◇

For interpolation in the extremes, one takes $t(x) = (x^2-1)\mathcal{U}_{n-1}(x)$ with $m \leq n-2$. It can be verified that in that case $w_j = (-1)^j \delta_j / g_n(z_j)$ where $\delta_j = 1/2$ for $j = 0, n$ and 1 otherwise. Furthermore for $z_j \neq \pm 1$: $g_n(z_j) = \frac{\pi}{\lambda_j}(1-x_j^2) = \frac{\pi}{\lambda_j}\sin^2\theta_j$ where $\lambda_j = 1/\sum_{k=0}^{n-2} |\mathcal{U}_k(x_j)|^2$ is the quadrature weight.

For more details, consult [154, 158, 153, 152].

# 8 Quadrature with preassigned nodes

If we use as the nodes of the quadrature formula on $\mathbb{T}$ the zeros of a PORF, these depend on a parameter $\tau \in \mathbb{T}$ This means that we can vary this parameter to fix one of the nodes at a preassigned position. Similarly, for a quadrature on (part of) the real line, one may use the zeros of a QORF as nodes. These nodes also depend on a parameter $\tau$ which may be chosen again to place one of the nodes at a preassigned location. This, and some generalizations will be the subject of this section.

## 8.1 Rational Szegő qf with preassigned nodes

Consider the $n$-point rational quadrature formula approximating $\int_{\mathbb{T}} f(t)\mathrm{d}\mu(t)$ of the form $I_n(f) = \sum_{k=1}^n \lambda_{nk} f(\xi_{nk})$ whose nodes $\{\xi_{nk}\}_{k=1}^n$ are the zeros of the PORF $Q_n(z,\tau) = \phi_n(z) + \tau\phi_n^*(z)$. Whatever $\tau \in \mathbb{T}$ is, the formula will be exact in $\mathcal{L}_{n-1,n-1}$, the nodes $\xi_{nk}$ are on $\mathbb{T}$ and the weights $\lambda_{nk}$ are all positive.

### 8.1.1 Rational Szegő-Radau qf

Now we look for an $(n+1)$-point qf with one of its nodes fixed at $\xi \in \mathbb{T}$, the $n$ other nodes are free to maximize the domain of exactness. Clearly, if we take as nodes the zeros of the PORF $Q_{n+1}(z,\tau)$ then the nodes will depend on $\tau$. It is obvious that by choosing $\tau = \tau(\xi) = -\phi_{n+1}(\xi)/\phi_{n+1}^*(\xi)$ then $\xi$ will be a zero of $Q_{n+1}(z,\tau)$. Since $\tau(\xi)$ as a function of $\xi$ is a Blaschke product with winding number $n+1$, there are $n+1$ instances of $\xi \in \mathbb{T}$ where this value of $\tau$ is reached. In practice this means that we can rotate any of the $n+1$ (ordered) zeros of $Q_{n+1}(z,\tau)$ to the position $\xi$ by varying $\tau$ over $\mathbb{T}$.

This formula is thus a rational Szegő quadrature formula with a particular choice of the parameter $\tau = \tau(\xi)$. It has the form $\sum_{k=1}^{n+1} \lambda_{n+1,k}(\tau) f(\xi_{n+1,k}(\tau))$. One of the $\xi_{n+1,k} = \xi$ and the remaining nodes and all the weights are chosen to maximize the domain of exactness to $\mathcal{L}_{n,n}$. The formula has all positive weights and it is unique. This is called a rational *Szegő-Radau* formula.

**Exercise 8.1** With the particular choice of $\tau$ mentioned above, the PORF $Q_{n+1}(z,\tau)$ will be of the form $\frac{(z-\xi)q_n(z)}{\pi_{n+1}(z)}$ with $q_n \in \mathcal{P}_n$. Define $\tilde{Q}_n(z) = \frac{q_n(z)}{\pi_{n-1}(z)(1-\overline{\alpha}_{n+1}z)}$ and the measure $\mathrm{d}\tilde{\mu} = \frac{|z-\xi|^2}{|1-\overline{\alpha}_n z|^2}\mathrm{d}\mu(z)$ which is a rational modification of the measure $\mu$. Then prove that $\tilde{Q}_n$ is the $n$th PORF for $\tilde{\mu}$ for a sequence of poles $\tilde{A} = \{\alpha_1, \ldots, \alpha_{n-1}, \alpha_{n+1}\}$. This means that it is (1) $c$-invariant ($\tilde{Q}_n^*(z) = c\tilde{Q}_n(z)$ in $\tilde{\mathcal{L}}_n$), that (2) $\left\langle \tilde{Q}_n, f \right\rangle_{\tilde{\mu}} = 0$ for all $f \in \mathcal{L}_{n-1}(\alpha_{n+1})$, and (3) $\left\langle \tilde{Q}_n, 1 \right\rangle_{\tilde{\mu}} \left\langle \tilde{Q}_n, \tilde{B}_n \right\rangle_{\tilde{\mu}} \neq 0$.
Thus there must exist some $\tilde{\tau} \in \mathbb{T}$ such that $\tilde{Q}_n(z) = \tilde{\phi}_n(z) + \tilde{\tau}\tilde{\phi}_n^*(z)$ where $\tilde{\phi}_n$ is the $n$th ORF for measure $\tilde{\mu}$ and poles sequence $\tilde{A}$. $\diamond$

More details can be found in [41] where also the case of 2 prefixed nodes is considered, which we shall call rational a *Szegő-Lobatto* formula.

### 8.1.2 Rational Szegő-Lobatto qf

Since the Szegő-Radau formula is unique, with no free parameter left. All the freedom is used to maximize the domain of exactness. So to introduce some freedom again, we have to

reduce the dimension of this domain or equivalently increase the number of nodes. Recall that for the Radau case, we obtained the PORF $Q_n(z,\tau)$ by using the recurrence relation for the ORF (using the Schur parameters or Verblunsky coefficients $\{\varsigma_k\}_{k=1}^n \subset \mathbb{D}$), but we changed only the last Schur coefficient $\varsigma_{n+1} \in \mathbb{D}$ and replaced it by some $\tau \in \mathbb{T}$. So now we shall start from the $(n+2)$-point rational Szegő formula but modify the last two Schur parameters $\varsigma_{n+1}$ and $\varsigma_{n+2}$ in such a way that the two preassigned nodes are matched, that the nodes are all on $\mathbb{T}$, and at the same time use all the remaining freedom to maximize the domain of exactness.

In the $(n+2)$-point formula, two nodes are fixed, leaving us with $2n+2$ free parameters (nodes and weights). Hence it is reasonable to try and get exactness at least in the space $\mathcal{L}_{n,n}$ which has dimension $2n+1$. The problem was solved in the polynomial case by Jagels and Reichel [106]. The rational case is a relatively simple generalization. See [41]. Let $\xi_1$ and $\xi_2$ be the two prefixed nodes, and denote the remaining nodes as $z_j,\ j = 1,\ldots,n$. So the function in $\mathcal{L}_{n+2}$ whose zeros are these nodes has the form $R_{n+2}(z) = \frac{(z-\xi_1)(z-\xi_2)P_n(z)}{\pi_{n+2}(z)}$, $P_n(z) = \prod_{k=1}^n (z-z_j)$. The qf has the form $I_{n+2}\{f\} = A_1 f(\xi_1) + A_2 f(\xi_2) + \sum_{k=1}^n \lambda_j f(z_j)$. The following Theorem is proved in [41, Thm. 6.1].

**Theorem 8.1** *With the notation introduced above, the qf will be exact in $\mathcal{L}_{n,n}$ if and only if*

*(1) $I_{n+2}\{f\}$ is exact in $\mathcal{L}_{p,q}$ with $0 \le p,q \le n$ and $p+q = n+1$*
*(2) $\langle R_{n+2}, f\rangle = 0$ for all $f \in \mathcal{L}_n$ such that $f(\alpha_{n+1}) = f(\alpha_{n+2}) = 0$.*

**Proof.** The direction $\Rightarrow$ is straightforward. Just show that the integrand in $\langle R_{n+2}, f\rangle$ is in $\mathcal{L}_{n,n}$ and hence the qf formula is exact while its nodes are zeros of the integrand.
The direction $\Leftarrow$ is equally simple. Take an arbitrary $f \in \mathcal{L}_{n,n}$ and choose $L \in \mathcal{L}_{p,q}$ (note that its dimension is $n+2$) that interpolates $f$ in the $n+2$ nodes. The difference $g = f - L \in \mathcal{L}_{n,n}$ and vanishes in these nodes, so that $I_\mu\{f\} = I_\mu\{g+L\} = I_\mu\{g\} + I_\mu\{L\} = I_\mu\{g\} + I_{n+2}\{L\}$. It suffices to rewrite $I_\mu\{g\}$ as $\langle h, R_{n+2}\rangle$ and use the condition (2) to find that it is 0. □

**Remark 8.1** Compare the condition (2) for $R_{n+2}$ with the condition for a para-orthogonal function $Q_{n+1}$ that was characterized by the fact that it is orthogonal to all $f \in \mathcal{L}_n$ for which $f(\alpha_n) = 0$. In a sense we could call $Q_{n+2}$ para-orthogonal of second order.

Reorganizing the integral $\langle R_{n+2}, f\rangle = 0$, it means that $Q_n(z)$ is orthogonal to all polynomials in $\mathrm{span}\{z, z^2, \ldots, z^{n-1}\}$ with respect to the *complex* measure $\frac{(1-\xi_1\overline{z})(z-\xi_2)}{|\pi_n(z)|^2}\mathrm{d}\mu(z)$. However, since the measure is complex, nothing can be said about the location of its zeros.

The previous condition will thus lead to a system of linear equations (the orthogonality conditions) which will give the nodes and the weights we are looking for. An alternative approach is the following. See [39].

Now let us see how to construct the formula $I_{n+2}$ by manipulating the two Schur parameters $\varsigma_{n+1}$ and $\varsigma_{n+2}$. Whatever the $\{\varsigma_1,\ldots,\varsigma_n,\tilde{\varsigma}_{n+1}\} \subset \mathbb{D}$ are, we need the last one, $\tilde{\varsigma}_{n+2}$, to be on $\mathbb{T}$ to guarantee that its zeros are simple and on $\mathbb{T}$. So we generate the ORF $\phi_k$ for $k = 0,1,\ldots,n$ using the Schur parameters $\varsigma_k$, $k = 1,\ldots,n$. Next we replace $\varsigma_{n+1}$ by $\tilde{\varsigma}_{n+1} \in \mathbb{D}$ and define $\tilde{\phi}_{n+1}(z) = \frac{1-\overline{\alpha}_n z}{1-\overline{\alpha}_{n+1} z}[\zeta_{n+1}(z)\phi_n(z) + \tilde{\varsigma}_{n+1}\phi_n^*(z)] \in \mathcal{L}_{n+1}$ and finally compute the nodes as the zeros of $\tilde{\phi}_{n+2}(z) = \frac{1-\overline{\alpha}_{n+1} z}{1-\overline{\alpha}_{n+2} z}[\zeta_{n+2}(z)\phi_{n+1}(z) + \tilde{\varsigma}_{n+2}\phi_{n+1}^*(z)] \in \mathcal{L}_{n+2}$ Choosing $\tilde{\varsigma}_{n+1} \in \mathbb{D}$ guarantees that $\tilde{\phi}_{n+1}$ is an ORF with respect to some positive measure $\tilde{\mu}$ on $\mathbb{T}$. Now we have to choose $\tilde{\varsigma}_{n+1}$ and $\tilde{\varsigma}_{n+2}$ such that $\tilde{\phi}_{n+2}$ has zeros $\xi_1$ and $\xi_2$. Define $x_i = \zeta_{n+1}(\xi_i)$ and

$\tau_i = \zeta_n(\xi_i)\phi_n(\xi_i)/\phi_n^*(\xi_i)$, $i = 1, 2$. All these 4 numbers are in $\mathbb{T}$. The condition $\tilde{\phi}_{n+2}(\xi_i) = 0$ corresponds to $\overline{\tilde{\zeta}}_{n+1}\tilde{\zeta}_{n+2} - \tau_i\xi_i\tilde{\zeta}_{n+1} - \overline{\tau}_i\tilde{\zeta}_{n+2} + \xi_i = 0$, $i = 1, 2$. Set $i = 1$ and solve for $\tilde{\zeta}_{n+2}$ to get

$$\tilde{\zeta}_{n+2} = \xi_i \frac{\tau_1 - \tilde{\zeta}_{n+1}}{1 - \tau_1\overline{\tilde{\zeta}}_{n+1}}.$$

Plug this into the equation for $i = 2$ and we get

$$(\overline{\tau}_1\xi_1 - \overline{\tau}_2\xi_2)|\tilde{\zeta}_{n+1}|^2 + \overline{\tau}_1\overline{\tau}_2\tilde{\zeta}_{n+1}(\xi_2 - \xi_1) + (\xi_2 - \xi_1)\overline{\tilde{\zeta}}_{n+1} = \overline{\tau}_1\xi_2 - \overline{\tau}_2\xi_1.$$

The following cases can occur [39]:

1. $\overline{\tau}_1\xi_1 = \overline{\tau}_2\xi_2$. Then (because $\xi_1 \neq \xi_2$) the equation reduces to $\overline{\tau}_1\overline{\tau}_2\tilde{\zeta}_{n+1} + \overline{\tilde{\zeta}}_{n+1} = 0$. There are two (distinct) solutions: $\tilde{\zeta}_{n+1} = \tau_i$, $i = 1, 2$. Because of the linearity of the equation, all values on the line joining these two numbers will also satisfy the equation. However, since we need $\tilde{\zeta}_{n+1}$ to be in $\mathbb{D}$, we have to restrict ourselves to the segment from $\tau_1$ to $\tau_2$ inside $\mathbb{D}$.

2. $\overline{\tau}_1\xi_1 \neq \overline{\tau}_2\xi_2$. This is the more general case. The equation is now

$$|\tilde{\zeta}_{n+1}|^2 - \overline{\tau}_1\overline{\tau}_2\tilde{\zeta}_{n+1}c - \overline{\tilde{\zeta}}_{n+1}c = \frac{\overline{\tau}_1\xi_2 - \overline{\tau}_2\xi_1}{\overline{\tau}_1\xi_1 - \overline{\tau}_2\xi_2}$$

   with

$$c = \frac{\xi_1 - \xi_2}{\overline{\tau}_1\xi_1 - \overline{\tau}_2\xi_2} \quad \text{so that} \quad \frac{\overline{\tau}_1\xi_2 - \overline{\tau}_2\xi_1}{\overline{\tau}_1\xi_1 - \overline{\tau}_2\xi_2} = 1 - (\overline{\tau}_1 + \overline{\tau}_2)c$$

   Note $\overline{c} = \overline{\tau}_1\overline{\tau}_2 c$ and hence $|c|^2 + 1 - (\overline{\tau}_1 + \overline{\tau}_2)c = r^2$, $r = |1 - c\overline{\tau}_2|$. In other words

$$|\tilde{\zeta}_{n+1}|^2 - 2\mathrm{Re}\,(\tilde{\zeta}_{n+1}\overline{c}) + |c|^2 = |\tilde{\zeta}_{n+1} - c|^2 = r^2$$

   which is a circle $C$ with center $c$ and radius $r$. The values $\tau_1$ and $\tau_2$ are the intersection of the circle $C$ and $\mathbb{T}$. Again, since we need $\tilde{\zeta}_{n+1}$ to be in $\mathbb{D}$, we can only use the arc of $C$ from $\tau_1$ to $\tau_2$ that lies inside $\mathbb{D}$.
   In the exceptional case that $\tau_1 = \tau_2$, hence $c \in \mathbb{T}$ and $r = 0$. Then the solution is unique. If we introduce $\tilde{\zeta}_{n+1} = \tau_1$, then $\tilde{\phi}_{n+1}$ is a PORF for the measure $\mu$ and because of the definition of $\tau_1$ it means that $\xi_1$ and $\xi_2$ are both zeros of $\tilde{\phi}_{n+1}$. We have an $(n+1)$-point rational Szegő qf exact in $\mathcal{L}_{n,n}$ because it so happened that by choosing a $\tau$ that places one of its nodes at $\xi_1$ accidentally made $\xi_2$ coincide with another of its nodes.

### 8.1.3 Rational Szegő qf with more than two preassigned nodes

In the rational Szegő-Radau case, we fixed one node and the qf always existed and was unique.
In the rational Szegő-Lobatto case, we fixed two nodes, and we found generically infinitely many solutions depending on one parameter, restricted to some arc in $\mathbb{D}$.
The natural step is whether we can use that freedom in the latter to fix a third node. This will however not have a solution in general. One may see intuitively that if we want to fix the nodes $\xi_1$, $\xi_2$ and $\xi_3$, then the previous arguments will fix $\xi_1$ and $\xi_2$ with a parameter on an arc in $\mathbb{D}$. Using the same arguments, fixing $\xi_1$ and $\xi_2$ will require the free parameter to be on (another) arc in $\mathbb{D}$. Only if these arcs intersect one may hope for a solution to exist.

## 8.2 Rational Gauss qf on $[-1,1]$ with preassigned nodes

For the polynomial case, Gauss qf with preassigned nodes is well documented. An $n$-point Gauss formula has polynomial degree of exactness $2n-1$, nodes are inside the interval $I=[-1,1]$ and all the weights are positive. If you prescribe $m$ nodes and use the remaining ones to maximize the degree of exactness, then this degree will drop by 1 for every node that is fixed in advance. The nodes are found as the zeros of a quasi-orthogonal polynomial (QOP) of order $m$. That is a polynomial $P\in\mathcal{P}_n$ such that $\langle P,z^k\rangle=0$ for $k=0,\ldots,n-m-1$ but $\langle P,z^{n-m}\rangle\neq 0$. Clearly $m=0$ gives the usual Gauss qf, $m=1$ corresponds to the Gauss-Radau and $m=2$ to the Gauss-Lobatto qf[9]. Note however that the zeros of these POP may not be all in the interval and/or the corresponding weights may not be positive. An analysis of the 'good' and the 'bad' intervals as well as a modification of the computational algorithm via an eigenvalue problem can be found in [14]. This works for any finite, half infinite or infinite interval. The rational analog is however much more complicated and will be analyzed via a transformation of the previous rational Szegő qf.

### 8.2.1 Rational Gauss-Radau qf

Let us first suppose that the poles $\alpha_k$ are all in $\hat{\mathbb{R}}^I=\hat{\mathbb{R}}\setminus I$. If we fix just one node $\xi$ in an interpolatory $(n+1)$-point Gauss qf then it will not come as a surprise that we need to use as nodes the zeros of a QORF $Q_{n+1}(x)=\phi_{n+1}(x)+\tau_{n+1}\frac{Z_{n+1}(x)}{Z_n(x)}\phi_n(x)$, $\tau_{n+1}\in\mathbb{R}$. We just have to choose $\tau_{n+1}=-\frac{Z_n(\xi)\phi_{n+1}(\xi)}{Z_{n+1}(\xi)\phi_n(\xi)}$ to ensure that $\xi$ is one of its zeros. All its nodes will be simple and real. We only have to worry about them being in $I$. If we consider qf on $\mathbb{R}$, then we still need to ensure that none of the zeros coincides with one of the poles (degeneracy). If $\alpha_n\in\hat{\mathbb{R}}^I$, then all zeros are simple and in $(-1,1)$ but if $\alpha_n\notin\hat{\mathbb{R}}^I$, then it is only guaranteed that at least $n-1$ zeros will be in $(-1,1)$ [155, Thm. 4.1]. If the qf exists, it is exact in $\mathcal{L}_{n,n}$ and the solution is unique. This is very much like in the Szegő-Radau case. Details are found for example in [69], where it is also shown that these nodes are the zeros of an ORF of degree $n$ but for a rational modification of the measure. At the same time, it discusses the case where the poles are in $\hat{\mathbb{C}}^I$, i.e., they are not necessarily real.

A study of ORF on the positive real line and poles on $[-\infty,0)]$ was undertaken in [36]. The corresponding quadrature formulas based on QORFs, hence including the rational Gauss-Radau case, was given in [43]. There it is shown that in general it can only be guaranteed that at least $n-1$ simple zeros are in $(0,\infty)$ and the weights corresponding to positive nodes are positive. The values of $\tau$ for which all the zeros are in $(0,\infty)$ are characterized.

**Exercise 8.2** Suppose the poles $\alpha_k$ are all real and outside the support of the measure $\mu$. The QORF $Q_n(z,\tau)=\phi_n(x)+\tau\frac{Z_n(z)}{Z_{n-1}(x)}\phi_{n-1}(x)$ will have $n$ zeros that depend on $\tau$. This means that if $\xi$ is a zero, then the value $\tau=\tau(\xi)=-\frac{\phi_n(x)Z_{n-1}(x)}{\phi_{n-1}(x)Z_n(x)}$ will be obtained for $n$ different values of $\xi$. Set $\phi_k=p_k/\pi_k$, then the poles of $\tau(\xi)$ will be the zeros of $p_{n-1}(x)(1-x/\alpha_{n-1})$ and the interlacing zeros are the zeros of $p_n(x)(1-x/\alpha_n)$. The function is monotone in between the poles and goes from $-\infty$ to $+\infty$ (or vice versa, depending on normalization). Make a sketch of the graph and derive for what values of $\tau$, all the zeros will be in the support of the measure. $\diamond$

[9] These names are classically used for preassigned nodes in one or both end points, but we use this also if the prefixed nodes are not end points of the interval.

### 8.2.2 Qf on $\mathbb{T}$ and $I$ with prescribed nodes

To the best of my knowledge, not a general study has been made of the rational Gauss-Lobatto qf, i.e., rational Gauss qf with respect to a measure supported on (a subset of) the real line and with two prescribed nodes, and a maximal domain of validity.

**Question 8.1** What can be said about rational Gauss qf with two prescribed nodes, and support of the measure on $\hat{\mathbb{R}}$, $\mathbb{R}_+$ of $I = [-1, 1]$? ◯

What has been investigated, for the polynomial case [15, 16, 17] as for the rational case [67, 68]. The idea is rather simple. Consider the case of a symmetric measure $\mathring{\mu}$ on $\mathbb{T}$ and suppose also the poles $\alpha_k \in \hat{\mathbb{C}}^I = \hat{\mathbb{C}} \setminus [-1, 1]$ are symmetric with respect to $\mathbb{R}$, i.e., they appear either in complex conjugate pairs or are real on $\mathbb{T}$. Then also the zeros $z_k \in \mathbb{T}$ of a QORF $\mathring{Q}_n(z, \mathring{\tau})$ will be symmetric with respect to $\mathbb{R}$ if $\mathring{\tau} = \pm 1$ and the weights for symmetric nodes are equal. In particular $z = 1$ will be a zero of $\mathring{Q}_n(z, \mathring{\tau})$ if $\mathring{\tau} = -1$ and $z = -1$ is a zero if $\mathring{\tau} = (-1)^{n+1}$.

On $I = [-1, 1]$ we associate the measure $\mu$ as follows: $\mathring{\mu}'(e^{\mathbf{i}\theta}) = |\sin(\theta)|\mu'(\cos\theta)$. The poles are given as $\beta_k = \mathfrak{j}(\alpha_k)$ We can project the nodes of a $2n$-point rational Szegő formula from $\mathbb{T}$ on the interval $I = [-1, 1]$, i.e., if $z_k = e^{\mathbf{i}\theta_k} \in \mathbb{T}$ is a node for the Szegő qf, then $x_k = \cos\theta_k$ will be a node for the corresponding rational $n$-point Gauss qf and the weights also match. This result holds in both directions [67, Thm. 3.2,3.3].

In the case of a $(2n + 1)$-point rational Szegő qf, then one of the nodes should be $+1$ or $-1$. The 'projected' formula will then be a rational Gauss-Radau qf for $I$.

If a $(2n + 2)$-point Szegő-Lobatto qf is used with prescribed nodes $+1$ and $-1$, then the 'projected' formula will be a rational Gauss-Lobatto formula for $I$ and vice versa [67, Thm. 3.4].

See also [68] for a slightly more general discussion.

# 9 Error bound

There are many different ways to estimate the quadrature error $E_n\{f\} = I_\mu\{f\} - I_n\{f\}$. A survey can be found in [31]. We give some examples.

## 9.1 An estimate on $\mathbb{R}$

One possible expression can be found on the basis of an alternative derivation of the qf. Indeed, it is possible to construct a rational interpolant $R_{2n-1} \in \mathcal{L}_{n,n-1} = \mathcal{L}_n \cdot \mathcal{L}_{n-1}$ that interpolates $f$ twice in each point of the set $\{t_1, \ldots, t_n\}$, i.e., it interpolates the function value as well as the derivative. Thus $R_{2n-1}(t_k) = f(t_k)$ and $R'_{2n-1}(t_k) = f'(t_k)$, $k = 1, \ldots, n$. It is clear that it is of the form $R_{2n-1}(x) = P_{2n-1}/D_{2n-1}$, $P_{2n-1} \in \mathcal{P}_{2n-1}$ and $D_{2n-1} = \pi_n\pi_{n-1}$, $\pi_k(x) = \prod_{j=1}^k (1 - x/\alpha_j)$. The numerator satisfies $P_{2n-1}(x) = \sum_{j=1}^n [H_j(x)f(t_j) + \tilde{H}_j f'(t_j)]$ with $H_j(x) = L_j(x)D_{2n-1}(t_j) + \tilde{L}_j(x)D'_{2n-1}(t_j)$ and $\tilde{H}_j(x) = \tilde{L}_j(x)D_{2n-1}(t_j)$ where the Lagrange polynomials are given by $L_j(t_k) = \delta_{jk}$, $L'_j(t_k) = 0$ while $\tilde{L}_j(t_k) = 0$, $\tilde{L}'_j(t_k) = \delta_{jk}$. The corresponding interpolatory qf is therefore $I_n\{f\} = \sum_{j=1}^n [A_j f(t_j) + \tilde{A}_j f'(t_j)]$ with $A_j = I_\mu\{H_j/D_{2n-1}\}$ and $\tilde{A}_j = I_\mu\{\tilde{H}_j/D_{2n-1}\}$. If we take as interpolation nodes $t_j$ the zeros of the $n$th ORF $\phi_n$, then if follows from orthogonality that $\tilde{A}_j = 0$ and the qf becomes a rational Gauss qf. The advantage is that we can now integrate the interpolation error to find the integration error, namely

$$E_n\{f\} = I_\mu\{f - R_{2n-1}\} = \frac{(fD_{2n-1})^{(2n)}(\theta(x))}{(2n)!} \int \phi_n^2(x)(1 - x/\alpha_n)\mathrm{d}\mu(x).$$

For a measure supported on $[0, +\infty)$ and poles satisfying the density result in the appendix In combination with the density result given in the appendix C.2, it follows that the qf will converge for any $f$ continuous on $[0, \infty)$ and for which the limit as $x \to +\infty$ is finite.

As a consequence one may prove [33, Thm. 9.1]

**Theorem 9.1** *Under the conditions listed above and with $f$ analytic in a connected domain, containing $[0, +\infty)$ whose boundary is a rectifiable curve $\Gamma$, then*

$$|E_n\{f\}| \leq \frac{C(\Gamma, f)}{\inf_{z\in\Gamma} |\lambda_n(z)\phi_n^2(z)|} \int_0^\infty \lambda_n(x)\phi_n^2(x)\mathrm{d}\mu(x)$$

*$\lambda_n(x) = 1 - x/\alpha_n$, $C(\Gamma, f)$ is a constant depending on $\Gamma$ and $f$.*

To obtain this result, an expression for the corresponding interpolation error of the Carathéodory function $\Omega_\mu(z) = \Omega_n(z)$ is used.

## 9.2 An estimate on $\mathbb{T}$

Arguments along the same lines can be derived for estimates of the integration error for rational Szegő qf on $\mathbb{T}$.

Another approach gives rates of convergence via root asymptotics (i.e. rate of convergence). With the sequence $A_n = (\alpha_k)_{k=1}^n$ associate a measure $\nu_n = \sum_{k=1}^n \delta_{\alpha_j}$ Suppose that this converges to $\nu^A$ in a weak star sense. Set $\hat{z} = 1/\overline{z}$ for any complex number $z$,

$A = \{\alpha_k\}_{k=1}^{\infty}$ and $\hat{A} = \{\hat{\alpha}_k\}_{k=1}^{\infty}$. Then it can be shown that the Blaschke products whose zeros are these $\alpha_k$ satisfy

$$\lim_{n\to\infty} |B_n(z)|^{1/n} = \exp\{\lambda(z)\}, \quad \lim_{n\to\infty} |B_n(z)|^{-1/n} = \exp\{\lambda(\hat{z})\}, \quad \hat{z} = 1/\overline{z}$$

for $z \in \mathbb{C} \setminus \left(\{0\} \cup \operatorname{supp}(\nu^A) \cup \operatorname{supp}(\nu^{\hat{A}})\right)$ where

$$\lambda(z) = \int \log \left| \frac{x - z}{1 - \overline{z}x} \right| \mathrm{d}\nu^A(x).$$

This can be used in combination with the asymptotics for the PORF and the approximation of the Carathéodory functions to arrive at the following theorem [23, Thm. 5.4].

**Theorem 9.2** *Let $f$ be a function such that it is analytic in a closed region $G$ with $\mathbb{T} \subset G$ and $G \cap (A \cup \hat{A}) = \varnothing$ and such that its boundary $\Gamma$ is a finite set of rectifiable curves. Suppose further $\sum(1 - |\alpha_k|) = \infty$. Then the interpolation error $E_n$ for the rational Szegő qf $I_n$ satisfies*

$$|E_n\{f\}|^{1/n} < \gamma < 1$$

*where $\gamma = \max\{\gamma_1, \gamma_2\}$ with $\gamma_1 = \max_{z\in\Gamma\cap\mathbb{D}} \exp\{\lambda(z)\}$ and $\gamma_2 = \max_{z\in\Gamma\cap\mathbb{E}} \exp\{\lambda(\hat{z})\}$.*

## 9.3 A general estimate

A general result can be found in the work of de la Calle Ysern [59, 58]. The main result for $\mathbb{T}$ is as follows [59, Thm. 1]. Let $I_\mu = \int_{\mathbb{T}} f(e^{\mathbf{i}\theta})\rho(\theta)\mathrm{d}\theta$ with $\rho$ possibly a *complex* measure, $\pi_n(z) = \prod_{k=1}^{n}(1 - \overline{\alpha}_k z)$, $I_n$ is a rational qf exact in $\mathcal{L}_{p,q}$, $p, q < n$, $p + q \geq n - 1$, and $E_n = I_\mu - I_n$. Given a function $f$ analytic in the neighborhood of $\mathbb{T}$ let $T_r$ be a circle with radius $r < 1$ and $T_R$ a circle with radius $R > 1$ and such that $T_r \cup T_R \subset U$ where $U$ is the domain of analyticity of $\pi_{p*}\pi_q f$.

**Theorem 9.3** *With the notation just introduced, the quadrature error can be bounded by*

$$|E_n\{f\}| \leq \|\pi_{p*}\pi_q f\|_{T_r\cup T_R}(\|\rho\| + \|I_n\|) \times \left( \frac{r^{p+1}}{1 - r^2}\sqrt{g_p(r)g_q(r)} + \frac{R^{1-q}}{R^2 - 1}\sqrt{g_p(1/R)g_q(1/R)} \right)$$

*where*

$$g_m(s) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\mathrm{d}t}{|\pi_m(se^{\mathbf{i}t})|^2}, \quad |s| < 1.$$

**Corollary 9.4** *Suppose $\int_0^{2\pi} \frac{|\rho(\theta)|^2}{w(\theta)}\mathrm{d}\theta < \infty$ for some weight function $w$ and suppose that $\sum_{j=1}^{\infty}(1 - |\alpha_j|) = \infty$. Then for any function $f$ continuous on $\mathbb{T}$, and $I_n$ an interpolatory qf with weights $A_j$ whose nodes $z_j$ are the zeros of the QORF w.r.t. $w$*

$$\lim_{n\to\infty} I_n\{f\} = \int_0^{2\pi} f(e^{\mathbf{i}\theta})\rho(\theta)\mathrm{d}\theta$$

*and*

$$\lim_{n\to\infty} \sum_{j=1}^{n} |A_j| f(z_j) = \int_0^{2\pi} f(e^{\mathbf{i}\theta})|\rho(\theta)|\mathrm{d}\theta.$$

Similar results can be obtained for integrals over $I = [-1, 1]$. The circle $T_r$ is replaced by its inverse Joukowski map [58, Thm. 1]. Now $\pi_n(x) = \prod_{k=1}^{n}(1 - x/\beta_k)$ (the poles are $\beta_k$), set $\alpha_k = \mathrm{j}^{inv}(\beta_k)$ and $\tilde{\pi}_n(z) = \prod_{i-1}^{n}(1 - \alpha_i z)$, and let $U$ be the domain of analyticity of $\pi_m f$ and $T_r = \{z \in \mathbb{C} : |\mathrm{j}^{inv}(z)| = r\}$ with $r < 1$ such that $T_r \subset U$.

**Theorem 9.5** *Let $f$ be analytic in a neighborhood of $I$ and $I_n$ is an $n$-point qf exact in $\mathcal{L}_{m+1}$ with $n \leq m + 1 \leq 2n$. Then the quadrature error is bounded by*

$$|E_n\{f\}| \leq \|\pi_m f\|_{T_r} \frac{2r^{m+1}}{1 - r^2}(\|\mu\| + \|I_n\|) \prod_{i=1}^{m} |1 + \alpha_i^2| \times g(r)$$

*with*

$$g(r) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}t}{|\tilde{\pi}_m(re^{\mathbf{i}t})|^2}.$$

**Corollary 9.6** *If $\sum_{k-1}^{n}(1 - |\alpha_k|) = \infty$ then the above interpolatory qf will converge to the integral of any function $f$ continuous on $I$ if and only if $\|I_n\| = \sum_k |A_k| \leq M$.*

# 10 Some applications

To illustrate the use of ORF, we pick two applications.

## 10.1 Rational Krylov methods

Given a linear operator $A \in \mathbb{C}^{N\times N}$ (usually large scale and Hermitian), one may project it onto a lower dimensional subspace to find the solution of a system $Ax = b$ or to find its eigenvalues. Most often this subspace is a Krylov subspace which is of the form $K_n(A, v) = \text{span}\{v_0, v_1, \ldots, v_n\}$ with $v_k = A^k v$ for some starting vector $v$. To find the projection, the basis $\{v_k\}$ is orthogonalized to give orthogonal vectors $\{q_k = \phi_k(A)v\}$ instead where the $\phi_k$ is an orthogonal polynomial of degree $k$, orthogonal to $\mathcal{P}_{k-1}$ with respect to a linear functional $M$ which is defined by the moments $M(z^k) = m_k = v^H A^k v$, $k = 0, 1, 2, \ldots$. This corresponds to orthogonal polynomials on $\mathbb{R}$. The computation is efficient (but not always very stable) because there is the well known three-term recurrence relation, characterized by a tridiagonal Jacobi matrix.

The projection (restriction) of $A$ on $K_n$ is obtained as $Q_n^H A Q_n = J_n$ with $Q_n = [q_0, \ldots, q_n]$ and $J_n$ the truncated Jacobi matrix. The eigenvalues of $J_n$ are called Ritz values and approximate the eigenvalues of $A$. This corresponds to the power method for eigenvalue computation and we will therefore have converge for the dominant eigenvalue of $A$.

To approximate a different eigenvalue (e.g., near $\alpha$), one may use the idea of the inverse power method with shift. Hence it is better to use a rational Krylov space to project on, spanned by $v_k = B^k v$ with $B = (I - \alpha^{-1}A)^{-1}$ or more generally $v_k = (I - \beta_k A)(I - \alpha_k^{-1}A)^{-1}v_{k-1}$, $k = 0, 1, \ldots$. Intuitively, it will enforce convergence to eigenvalues in the neighborhood of $\alpha_k$ and diminish the influence of eigenvalues in the neighborhood of $\beta_k$.

Obviously, an orthogonal basis can now be obtained by a set of ORFs $\phi_k$ instead of orthogonal polynomials. So, we define as before

$$Z_k(z) = \tfrac{z}{1-z/\alpha_k} \text{ and } b_k(z) = \textstyle\prod_{j=1}^k Z_j(z),\ k \in \mathbb{N}, \text{ and } \mathcal{L}_n = \text{span}\{b_0, \ldots, b_n\},\ n \in \mathbb{N}.$$

Note that if $A$ is a matrix, then $A(I - \alpha^{-1}A)^{-1} = (I - \alpha^{-1}A)^{-1}A$ so that we may write $Z_k(A) = \frac{A}{I-\alpha_k^{-1}A}$ without ambiguity.

Now suppose that $A \in \mathbb{C}^{N\times N}$ is Hermitian positive definite. Hence it has an eigenvalue decomposition $A = U^H \Lambda U$ with $U$ unitary and $\Lambda$ diagonal containing the eigenvalues of $A$, namely $\{\lambda_j\}_{j=0}^{N-1} \subset (0, \infty)$ and $U = [u_0, \ldots, u_{N-1}]$ contains the eigenvectors. Define a discrete measure $\mu$ having weights $w_j = |v^H u_j|^2$ at the nodes $\lambda_j$. Thus

$$\mu = \sum_{j=0}^{N-1} w_j \delta_{\lambda_j}, \quad w_j = |v^H u_j|^2.$$

We set $\mathbf{b}(z) = [b_0, b_1(z), \ldots, b_{N-1}(z)]$ so that a function $r(z) = \sum_{k=0}^{N-1} r_k b_k(z) \in \mathcal{L}_{N-1}$ can be written as $r(z) = \mathbf{b}(z)\mathbf{r}$ with $\mathbf{r} = [r_0, \ldots, r_N]^T \in \mathbb{C}^N$ the coordinate vector of $r$ w.r.t. the basis $\mathbf{b}$. Note that the coordinate vector of $b_k$ w.r.t. $\mathbf{b}$ is $\mathbf{e}_k$, the $k$th unit vector. Define $V = [b_j(\lambda_i)]_{i,j=0}^{N-1}$ then the Gram matrix for $\mu$ w.r.t. the basis $\mathbf{b}$ is given by $G = V^H W V$ with $W = \text{diag}(w_0, \ldots, w_{N-1})$. Indeed,

$$G_{i,j} = \mathbf{e}_i^T G \mathbf{e}_j = \sum_{k=0}^{N-1} w_k \overline{b_i(\lambda_k)} b_j(\lambda_k) = \langle b_i, b_j \rangle_\mu .$$

Hence the inner product w.r.t. the discrete measure $\mu$ for rational functions in $\mathcal{L}_{N-1}$ can be written as

$$f(z) = \mathbf{b}(z)\mathbf{f}, \quad g(z) = \mathbf{b}(z)\mathbf{g} \in \mathcal{L}_{N-1} \quad \Rightarrow \quad \langle f, g \rangle_\mu = \mathbf{f}^H G \mathbf{g} = (V\mathbf{f})^H W (V\mathbf{g}).$$

Let the corresponding ORFs with respect to the discrete inner product $\langle \cdot, \cdot \rangle_\mu$ be denoted as $\phi_k$, then an orthonormal basis for the space $\text{span}\{v, v_1, \ldots, v_{N-1}\}$ with $v_k = b_k(A)v$ is now given as $q_k = \phi_k(A)v$, with $k = 0, \ldots, n-1$. From our general theory, we know that $w_j = 1/\sum_{k=0}^{N-1} |\phi_k(\lambda_j)|^2$.

The rational Krylov matrix $K = [v_0, v_1, \ldots, v_{N-1}]$ with $v_k = b_k(A)v$ can be written as

$$K = \mathbf{b}(A)v = U \text{diag}(U^H v) V.$$

$U^H v$ is the vector of coordinates of $v$ in the basis $\{u_0, \ldots, u_{N-1}\}$. Note that $K^H K = V^H W V = G$. Now set $\phi_k(z) = \mathbf{b}(z)\mathbf{f}_k$, then $q_k = \phi_k(A)v = \mathbf{b}(A)v\mathbf{f}_k = K\mathbf{f}_k$.

$$\langle \phi_k, \phi_l \rangle_\mu = \mathbf{f}_k^H \left\langle \mathbf{b}^H, \mathbf{b} \right\rangle_\mu \mathbf{f}_l = \mathbf{f}_k^H G \mathbf{f}_l = (K\mathbf{f}_k)^H (K\mathbf{f}_l) = q_k^H q_l = (V\mathbf{f}_k)^H W (V\mathbf{f}_l).$$

Thus defining $\Phi = ]\mathbf{f}_0, \ldots, \mathbf{f}_{N-1}]$ (which is an upper triangular matrix), then $Q = [q_0, \ldots, q_{N-1}] = K\Phi$ with $Q^H Q = \Phi^H K^H K \Phi = I$.

*Assuming regularity* of the ORFs, there is the three term recurrence relation and hence the orthogonal basis can be computed recursively using a relation of the form

$$Aq_k = c_{k-1}(I - \alpha_{k-1}^{-1} A)q_{k-1} + d_k(I - \alpha_k^{-1} A)q_k + c_k(I - \alpha_{k+1}^{-1} A)q_{k+1}.$$

Taking the first $n$ columns i.e., $Q_n = [q_0, \ldots, q_{n-1}]$,

$$AQ_n = Q_n J_n - AQ_n D_n^{-1} J_n \Rightarrow AQ_n(I_n + D_n^{-1} J_n) = Q_n J_n + \text{rank } 1$$

Then with $A_n = Q_n^H A Q_n$ it follows that

$$A_n(I_n + D_n^{-1} J_n) = J_n + \text{rank } 1.$$

Neglecting the last term of rank 1, it follows that the approximate eigenvalues of $A$, given as eigenvalues of $A_n$ are obtained by solving the generalized eigenvalue problem for the pencil $(I_n + D_n^{-1} J_n, J_n)$.

For further details about how to compute the coefficients $c_k$ and $d_k$ of the recurrence and about the choice of the $\alpha_k$ (one has to avoid 'forbidden poles') consult [60]. About the speed of convergence for the different eigenvalues, an analysis based on potential theory see [54].

## 10.2 Identification and model reduction

Another area where the ORFs are often used is in systems theory. See for example the book [102] and its references as well as many papers by Ninness and coworkers [123, 124, 2, 3, 125, 122, 164, 144, 11].

The problem is as follows. Suppose we have a discrete linear system with input $u = \{u_i\}$, output $y = \{y_i\}$ and with impulse response $g = \{g_i\}$. The output $y$ is a convolution of the input $u$ and the impulse response $h$. Assuming certain conditions, this corresponds in the frequency domain ($z$-transform, e.g. $U(z) = \sum_k u_k z^k$, $z = e^{\mathbf{i}\theta}$ etc.) to a relation $Y(z) = G(z)U(z)$, where $G$ is the transfer function.

Suppose we know the input $u$ and observe the output $y$ (possibly corrupted by noise), but we do our observations in the frequency domain. This means that we know $(U(z_k), Y(z_k))$ thus $G(z_k)$ for a number of points $\{z_1, \ldots, z_N\} \subset \mathbb{T}$. We want to approximate $G$ by a rational function $G_n \in \mathcal{L}_n$. Because the transfer function should be stable it has to be analytic *outside* $\mathbb{D}$.

Suppose we have some weights $w_k^2$ associated with the measurements. Then we want to minimize the discrete least squares error

$$\|G - G_n\|_w^2 = \sum_{k=1}^{N} w_k^2 |G(z_k) - G_n(z_k)|^2.$$

if the observed $Y$ is corrupted and we actually measure $Y = GU + V$, then we just want to find $G_n$ such that $Y_n = G_n U + V$ will minimize $\|Y - Y_n\|_w = \|(G - G_n)U\|_w = \|G - G_n\|_{\tilde{w}}$ where $\tilde{w}_k = w_k U(z_k)$.

In classical papers one used the so-called *Laguerre basis*, which was actually an orthogonalization of the basis $(\frac{1-\overline{\alpha}z}{z-\alpha})^k$ w.r.t. the Lebesgue measure on $\mathbb{T}$ [115]. Later one used the *Kautz basis* which involved not just one $\alpha$, but a pair of complex conjugate poles $(\alpha, \overline{\alpha})$ [165]. Then in [104] a cyclic repetition of a finite number of poles was used. The more general Takenaka-Malmquist basis with a general sequence of poles (but still orthogonal with respect to a continuous Lebesgue measure) was promoted by Ninness et al.

However, although the use of these bases meant an improvement, it is still is not the proper choice to solve the *weighted* and *discrete* least squares problem. Therefore orthogonality should be w.r.t. the discrete and weighted measure. Because the measure is discrete, we can actually evaluate all the integrals and hence all the inner products needed to evaluate the recursion for the ORF. This results in a very efficient algorithm as long as the degree is not too high (see [161]).

Heuberger at al. also used state space realizations for the rational approximants in the case of cyclically repetition of a finite set of poles (they all it a *Hambo basis*, see also [164]). This kind of realization is particularly interesting when the input and output are vectors (MIMO systems). It is then describes in discrete time as

$$\begin{array}{ll} x_{k+1} = Ax_k + Bu_k & \\ y_k = Cx_k + Du_k & \text{or} \end{array} \quad \begin{array}{l} zX(z) = AX(z) + BU(z) \\ Y(z) = CX(z) + DU(z) \end{array}$$

or $Y(z) = [C(zI - A)^{-1} + D]U(z) \equiv G(z)U(z)$. Generalizing this idea to the general rational basis is not very practical because the size of the matrices $(A, B, C, D)$ grows as the degree of the rational approximant increases [160].

There is however an alternative (linear algebra) approach to the construction of ORFs for a discrete measure on $\mathbb{T}$. This is e.g. explained in [47].

Denote $\Phi(z) = [\phi_0, \phi_1, \ldots, \phi_n]$ and $A = \operatorname{diag}(\alpha_0, \ldots, \alpha_n)$. This $A$ contains the poles $\{\alpha_k\}$ of our ORF system since these have to be inside $\mathbb{D}$. Then there is an Hessenberg matrix $H$ such that [47, Lemma 3.3]

$$\Phi(z) = \Phi(z)(zI - A)H + [0, \ldots, 0, \eta_{n+1,n}(z - \alpha_{n+1})\phi_{n+1}(z)]$$

Let $\mathbf{w}$ be the vector with weights $w_1, \ldots, w_n$ and set $Z = \operatorname{diag}(z_1, \ldots, z_n)$, the points on $\mathbb{T}$ supporting the discrete measure. Then [47, Thm. 3.6] there exists a unitary matrix $Q$ such that

$$Q^H [\mathbf{w}|Z] \left[\begin{array}{c|c} 1 & \\ \hline & Q \end{array}\right] = [\mathbf{w}_1|S + A]$$

where $S = H^{-1}$ is a semi-separable matrix and $\mathbf{w}_1$ is the vector $\|\mathbf{w}\|\mathbf{e}_1$. If the $z_k$ are on $\mathbb{R}$ then $S+A$ is Hermitian and if $z_k \in \mathbb{T}$ then $S+A$ is unitary. And this is an inverse eigenvalue problem: we know $\mathbf{w}$ and $Z$ and we have to find a matrix of the form $S + A$ using unitary transformations since $S + A = Q^H Z Q$ and the eigenvalues of this $S + A$ have to be the $z_k$ (the zeros of the ORF). There exists a fast algorithm to do this as explained in the cited paper.

**Question 10.1** The previous algorithm involves complex poles. In most practical situations the data are real and the poles are complex conjugate. So it is expected that combining such a pair of complex conjugate poles, an algorithm can be designed that works in real arithmetic. For the moment, there is not a nice clean way to achieve this [48]. ◯

In a somewhat different vain of linear prediction theory, of which the ideas go back to Wiener and Toeplitz, one does not know the input, or it is assumed to be white noise or a pulse train (like e.g., in speech processing). Then one wants to predict the next simple from its past:

$$y_k = -\sum_{j=1}^{N} a_j y_{k-j} \quad \forall k \in \mathbb{Z} \quad \text{or} \quad A(z)Y(z) = 0.$$

Where $A(z) = \sum_{j=0}^{N} a_j z^j$ with $a_0 = 1$, $Y(z) = \sum_{k\in\mathbb{Z}} y_k z^k$ so that $z$ acts a backward shift operator: $zy_k = y_{k-1}$ because $zY(z) = \sum_{k\in\mathbb{Z}} y_{k-1} z^k$.

Thus the problem is to minimize $\int_{\mathbb{T}} |A(z)Y(z)|^2 \mathrm{d}\lambda(z)$, or with $\mathrm{d}\mu(z) = |Y(z)|^2 \mathrm{d}\lambda(z)$, we have to minimize the norm of the polynomial $A(z)$ in w.r.t. $\mu$, given the fact that its constant term is 1.

The solution to this problem is known, and the answer is that $A$ is the reciprocal $\phi_N^*$ of the $N$th monic orthogonal polynomial for $\mu$. In practice this is done in a recursive way computing the successive polynomials $\phi_n$, $n = 1, 2, \ldots$. The polynomials $\mathcal{P}_n$ reach an $n + 1$ dimensional subspace of the past. However, if the true model has some singularities close to $\mathbb{T}$, then the convergence may be very slow. In such cases, it will be an advantage, even if we have only a vague idea of where these poles are (usually this is quite obvious from the shape of the spectrum $|Y(z)|^2$) to use a space of rational functions (with prescribed poles that we should choose in the neighborhood of the probable singularities).

Also for rational functions $A$ of degree $n$, the one with minimal norm will be the $n$th orthogonal one with monic normalization.

Where the polynomials modelled the signal by a filter $G(z) = 1/\phi_n^*(z)$, giving an autoregressive (AR) model, the rational case gives a filter that has some zeros (transmission zeros) which are the poles that we have chosen for the ORFs. This gives an autoregressive moving average (ARMA) model.

A comparison with the linearized rational approach mentioned in item 2 of Section 11 can be found in [47].

# 11 What we did not discuss

1. Gauss-Kronrod formulas: These are qf that start from an $n$-point Gauss qf with DoE $2n-1$ and add $n+1$ new points such that the DoE of the new formula is $3n+1$. These $n+1$ new points are the zeros of Stieltjes polynomials. If $\phi_n$ is the $n$th orthogonal polynomial on $I$ with respect to $\mu$, then the Stieltjes polynomial $\sigma_{n+1}$ whose zeros are the $n+1$ new nodes in the Kronrod extension is defined by the orthogonality conditions that $\phi_n\sigma_{n+1}$ is orthogonal to $\mathcal{P}_n$ on $I$ w.r.t. $\mu$. It would be interesting to extend this to the rational case. These formulas are mostly used to get error estimates.

2. Discrete least squares problems as inverse eigenvalue problems. There is a vast literature for the construction of recurrence relations for polynomials orthogonal with respect to a discrete measure $\mu(x) = \sum_k m_k \delta_{x_k}$. One starts from a set of interpolation data; $(x_k, f_k, w_k)_{k=1}^N$. The idea is to find the function $f$ such that $f(x_k) = f_k$ and this is done in a discrete least squares sense: $\min_{a_k} \|g_n - f\|_\mu$ where $g_n$ is an approximant of degree $n < N$ and $\|h\|_\mu^2 = \sum_k w_k^2 |h(x_k)|^2$. The approximant $g_n(x)$ can be polynomial or even rational with prescribed poles. The points $x_k$ can be on $\mathbb{R}$ or on $\mathbb{T}$. It is best to express $g_n$ as a linear combination of polynomials orthogonal with respect to the discrete measure $\mu$. The problem is efficiently solved as an *inverse eigenvalue problem*. Indeed, we know the discrete measure (i.e., weights $w_k$ and nodes $x_k$) and the problem is to find the recurrence matrix $R$ (i.e., the Jacobi or Hessenberg matrix) that will give the recurrence relation for the orthogonal polynomials $\phi_k$, $k = 0, \ldots, n$ such that $\phi_n$ has the nodes as zeros $x_j$ (i.e., these are the eigenvalues or the $R$) and the weights are given by the first component of the corresponding normalized eigenvector. For the rational case, the $R$ matrix can be formulated in terms of a semi-separable matrix plus a diagonal [150, 48, 151]. The interpolation data can be introduced one triple at a time, each time requiring a very efficient update of the $R$ matrix. Mostly on the linearized rational case see [149, 44, 47]. There are some applications in system identification in the frequency domain [45, 46, 137, 138].

3. Rational moment problems, and limiting disks. There is a difference between $\mathbb{T}$ and $\mathbb{R}$ in the following sense. We can consider on $\mathbb{T}$ with poles in $\alpha_k \in \mathbb{D}$ the moments $m_k = \int \frac{\mathrm{d}\hat{\mu}(z)}{\pi_k(z)}$, $\pi_k(z) = \prod_{j=1}^k (1 - \overline{\alpha}_k z)$, then this defines integrals of functions in $\mathcal{L}_\infty$. For an inner product, we need to define integrals of functions in $\mathcal{L}_{\infty,\infty}$ which involves all moments of the form $m_{k,l} = \int \frac{\mathrm{d}\mu(z)}{\pi_{k*}(z)\pi_l(z)}$. However, assuming an Hermitian inner product i.e., $m_k = m_{0,k} = \langle 1, \pi_k^{-1} \rangle = \overline{\langle \pi_{k*}^{-1}, 1 \rangle} = \overline{m}_{k,0}$, then using e.g. partial fraction decomposition, all $m_{k,l}$ can be computed from $m_k$.

   On $\mathbb{R}$ however, we need all moments $m_{k,l} = \int b_n(x) b_m(x) \mathrm{d}\mu(x)$ for an inner product on $\mathcal{L}_\infty$ if the poles are real or when they are arbitrarily complex. Indeed for real poles then $b_n(x)b_m(x)$ with $n \leq m$, this function duplicates the multiplicity of the poles for $b_n$. Only if there is a finite number of different poles which are each repeated an infinite number of times, only the knowledge of the $m_{0,k}$ will suffice to define an inner product in $\mathcal{L}_\infty$. So here one may consider a moment problem for $\mathcal{L}_\infty$ and a (different) moment problem for $\mathcal{L}_{\infty,\infty}$.

   The study of the moment problem being determinate or indeterminate is by considering truncated moment problems (i.e. quadratures for nested subspaces) and characterizing the measures via the corresponding Carathéodory function $\Omega(z)$. Since the solutions

for $\mathbb{T}$ depend on the parameter $\tau \in \mathbb{T}$ of the PORFs, by varying this $\tau$ one obtains for fixed $z$ a circle, which is the boundary of a disk. If the limiting radius of these (nested) disks, goes to 0, there is a unique solution, otherwise there are infinitely many, corresponding to the limiting disk.

For $\mathbb{R}$ something similar happens. The QORF depend on a parameter $\tau \in \mathbb{R}$, hence also the approximations to $\Omega$. Varying $\tau$ describes a circle for fixed $z$, leading to nested disks as $n \to \infty$. A limiting point gives a unique solution, otherwise the moment problem is indeterminate.

See [27, 34, 38, 40, 42, 18, 24, 32, 30].

4. Associated rational functions: these are rational functions obtained by shifting the recurrence coefficients as well as the poles. They are again ORFS but for a different measure. See [61, 62, 70]. They were investigated originally in the context of positive interpolatory quadrature formulas by Peherstorfer [136, 134, 133, 135].

5. Other rational quadrature formulas. Gautschi also has quadrature formulas related to the integration of rational functions. His work is however different from our approach, although the results could well be compared. See [91, 92, 93, 95].

6. Integration of periodic functions over $\mathbb{R}$. If $f$ is $2\pi$ periodic and one wants to compute the integral $\int_{\mathbb{R}} f(x)\sigma(x)\mathrm{d}x$, then it can be rewritten as $\int_{-\pi}^{\pi} f(x)\omega(x)\mathrm{d}x$ with $\omega(x) = \sum_{j\in\mathbb{Z}} \sigma(x+2\pi j)$. If we suppose that the latter sum converges. This has been considered for polynomial qf in the paper [57]. It might be interesting to explore this in the rational case.

7. Biorthogonal systems. Let $\{\phi_k\}$ be orthogonal polynomials on $\mathbb{T}$ w.r.t. weight function $\omega$ and $\gamma \in \{0, 1/2\}$. Define $f^{(\gamma)}_{n+1-2\gamma}$ and $g^{(\gamma)}_{n+1-2\gamma}$ by

$$c_n z^{-(n+\gamma)}\phi_{2n+1}(z) = f^{(\gamma)}_{n+1-2\gamma}(z) + \mathbf{i}g^{(\gamma)}_{n+1-2\gamma}(z)$$

for a particular constant $c_n$ and such that both $f^{(\gamma)}_{n+1-2\gamma}$ and $g^{(\gamma)}_{n+1-2\gamma}$ are real, then these form a biorthogonal system in the sense that $\left\langle f^{(\gamma)}_m, f^{(\gamma)}_n \right\rangle_\omega = p_n\delta_{m,n}$, $\left\langle g^{(\gamma)}_m, g^{(\gamma)}_n \right\rangle_\omega = q_n\delta_{m,n}$, $\left\langle f^{(\gamma)}_m, g^{(\gamma)}_n \right\rangle_\omega = 0$. For the case $\gamma = 0$ see [145] and [55] and for the more general treatment [56]. It might be interesting to see how much could be generalized to the rational case for particular situations of the poles.

8. ORFs for a measure supported on one or more intervals on $\mathbb{T}$ or $\mathbb{R}$

9. The relation to continued fractions

10. Extension to the matrix valued case. Many of the results that were given here for the scalar case have been extended to the matrix case. These are often studied in connection with matrix moment problems and matrix valued interpolation. [81, 80, 83, 112, 82, 113, 114, 84, 37] and many other authors: H. Dym, P. Dewilde, J. Ball, J.W. Helton, E. Frazho, D. Alpay, B. Fritzsche, B. Kirstein, etc. who developed a field of mathematics they call 'Schur analysis'. The applications are usually in systems theory, electrical engineering or stochastic processes. The mathematical background is often operator theory and/or linear algebra.

11. ...

# A Moment problems (polynomial case)

There are many classical moment problems associated with polynomials. Given an infinite sequence of moments $\{m_k\}$, find conditions such that there exists a positive measure $\mu$ such that $\int z^k \mathrm{d}\mu(z) = m_k$ for all $k$. In particular, if a solution exists characterize all possible measures that solve the moment problem, and find conditions such that the solution is unique.

If there is more than 1 solution to the problem, the moment problem is said to be *indeterminate*, and *determinate* if there is a unique solution.

There are also possible conditions on the support of the measure. The situations that have been studied most are the cases where the support is a compact interval of the real line (without loss of generality the interval $[0, 1]$), or a half line (without loss of generality $\mathbb{R}^+ = [0, \infty)$), or the whole real line $\mathbb{R}$. Some more recent results deal with the case where the support consists of several disjoint intervals. We shall not go into this for this introductory course.

Another variation may be possible if one does not only consider polynomials but also *Laurent polynomials*. Then we need moments $m_k = \int z^k \mathrm{d}\mu(z)$ for nonnegative $k$ but also for negative integers $k$. This case where $k \in \mathbb{Z}$ instead of $k \in \mathbb{N}$ is sometimes called a *strong* moment problem.

If only a finite number of moments is given, the problem is called a *truncated moment problem*.

In the sequel, we shall always assume that $m_0 = \int \mathrm{d}\mu(z) = 1$.

## A.1 The real line and its subsets

Hausdorff. If we are looking for a positive measure on [0,1], satisfying $m_k = \int x^k \mathrm{d}\mu(x)$ for $k \in \mathbb{N}$, then this is called the *Hausdorff moment problem*. Hausdorff showed that there is a $\sigma$-additive probability measure as a solution to the problem if the sequence $\{m_k\}$ is completely positive. Since we are looking for a positive measure, we should require that the integrals of polynomials positive on $[0, 1]$ should be positive. For this it is sufficient that the positive polynomials $x^k(1-x)^{n-k}$ should be positive for all $n \in \mathbb{N}$ and $k = 0, \ldots, n$. This results in that condition that $(-1)^k \Delta^k m_n \geq 0$ for all $k = 0, 1, \ldots$ where $\Delta$ is the forward difference operator $\Delta m_n = m_{n+1} - m_n$. If you want a solution with an infinite support, one needs strict inequalities. If there is such a solution, it will automatically be unique by the F. Riesz representation theorem because the polynomials are uniformly dense in the space $C([0, 1])$ so that the linear functional on bounded continuous functions can be extended uniquely to a positive linear functional on all Borel measurable functions. We shall not pursue this further.

Stieltjes. If the positive measure should have an infinite support on $\mathbb{R}^+$, one obtains the most famous *Stieltjes moment problem*. Since we want the linear functional (i.e., the measure) to be positive, All polynomials that are nonnegative on $\mathbb{R}^+$ should have a positive integral. Any such polynomial can be written as $(p_n(x))^2$ or $x(p_n(x))^2$. This means that the Hankel matrices $G_n = [m_{ij}]_{i,j=0}^n$ with $m_{ij} = m_{i+j}$ as well as the shifted Hankel matrices $G_n^{(1)} = [m_{ij}^{(1)}]_{i,j=0}^n$ with $m_{ij}^{(1)} = m_{i+j+1}$ have to be positive definite, or equivalently all the determinants $\Delta_n = \det(G_n)$ and $\Delta_n^{(1)} = \det(G_n^{(1)})$ of these matrices have to be positive for $n = 0, 1, \ldots$. This is the well known *Stieltjes condition* which is necessary and sufficient for the solvability of the moment problem with a $\mu$ having infinite support. If from some point on all the determinants are zero, there is a solution with only a finite number of mass points.

A sufficient condition for unique solvability (determinacy) is known to be the *Carleman condition*, which is $\sum_n m_n^{-1/2n} = +\infty$.

If we do not restrict ourselves to polynomials but also include negative powers so that the $m_k$ are prescribed for all integer $k \in \mathbb{Z}$, we speak of a *strong Stieltjes moment problem* (or a two-point Stieltjes moment problem). This notion was introduced by Jones, Thron, Waadeland in [110]. See also [129], [130], [128], [6], [8], [7], [75], [76], [5]. Based on a similar argument, it can be proved that there is a solution with infinite support if and only if some Hankel determinants have prescribed signs. We now need Hankel matrices of the form $G_n^{(k)}$ whose entries are $m_{k+i+j}$ for $i, j = 0, \ldots, n$. Denoting their determinants as $\Delta_n^{(k)}$, one needs $(-1)^n \Delta_n^{(-n)} > 0$ for $n = 0, 1, \ldots$ and $\Delta_{n-1}^{(-n)} > 0$ for all $n = 1, 2, \ldots$.

Hamburger. In the *Hamburger moment problem*, the measure is allowed to have support on the whole real line. Again we can apply the positivity condition, as above, but it will be simpler because a polynomial will now only be guaranteed to be nonnegative if it is the square of a polynomial and hence the solvability condition reads $\Delta_n \geq 0$ for all $n$ with strict inequality if one wants to have infinite support. The problem is *determinate* if the Carleman condition $\sum_n m_{2n}^{-1/2n} = +\infty$ is satisfied. Also here a *strong Hamburger moment problem* can be considered if moments $m_n$ are prescribed for $n \in \mathbb{Z}$, i.e., not only considering polynomials but also Laurent polynomials. See [109], [107], [126], [132], [127], [131]. The strong problem has a solution iff $\Delta_{n-1}^{(-n)} > 0$ and $\Delta_n^{(-n)} > 0$ for $n = 0, 2, 4, \ldots$.

## A.2 The unit circle

On the unit circle, the moments $m_k = \int_{\mathbb{T}} z^{-k} \mathrm{d}\mu(z) = \int_{-\pi}^{\pi} e^{-ikt} \mathrm{d}\mu(t)$ are called *trigonometric moments*. Note that for a real measure $m_{-k} = \overline{m}_k$, $k \in \mathbb{Z}$. Hence, if the moments are given for $n = 0, 1, 2, \ldots$, then they are automatically known for $n = -1, -2, -3, \ldots$ as well. Thus the strong moment problem (related to Laurent polynomials) is identical to the usual problem (related to algebraic polynomials).

The partial Gram matrices $G_n = [m_{ij}]_{i,j=0}^n$ now satisfy $m_{ij} = m_{i-j}$ and hence they are Hermitian Toeplitz matrices. The trigonometric moment problem has a solution if and only if all $G_n$ are positive semidefinite, $n = 0, 1, 2, \ldots$. If a solution exists, it is unique, i.e., the trigonometric moment problem is *determinate*. See e.g., [140, 4].

# B Orthogonal polynomials (a crash course)

## B.1 The orthogonal polynomials

With the inner products defined above we can immediately associate orthogonal polynomials. Suppose $\mathcal{P}_n$ denotes the set of polynomials of degree $n$ or lower and $\mathcal{P}$ is the set of all polynomials. A polynomial $\phi_n \in \mathcal{P}_n \setminus \mathcal{P}_{n-1}$ is the $n$th orthogonal polynomial in an orthogonal polynomial sequence (OPS) if $\langle \phi_n, \mathcal{P}_{n-1} \rangle_\mu = 0$. In principle one can start from any polynomial basis and apply the Gram-Schmidt algorithm to arrive at an OPS.

Explicit expressions exist:

$$\phi_n(z) = \frac{1}{\sqrt{\Delta_{n-1}\Delta_n}} G_n(z), \quad n = 0, 1, 2, \ldots \quad (\Delta_{-1} = 1)$$

where $G_n(z)$ is the same as $G_n$ but with the last row replaced by $\mathbf{z} = [1, z, \ldots, z^n]$. These are the normalized OP's with positive leading coefficient equal to $\sqrt{\Delta_{n-1}/\Delta_n}$.

### B.1.1 Recurrence

Because of the special structure of the moment matrix, there exist short recurrences. For example, in the *Hankel case* $\langle xp(x), q(x) \rangle_\mu = \langle p(x), xq(x) \rangle_\mu$ implies that $x\phi_n(x)$ will be automatically orthogonal to $\mathcal{P}_{n-2}$. This results in the well known three-term recurrence relation.

$$x\phi_n(x) = b_{n-1}\phi_{n-1}(x) + a_n\phi_n(x) + b_n\phi_{n+1}(x), \quad n = 0, 1, \ldots, \quad \phi_0 = 1/\sqrt{m_0}, \phi_{-1} = 0$$

with $a_n = \langle x\phi_n(x), \phi_n(x) \rangle_\mu \in \mathbb{R}$ and $b_n = \sqrt{\Delta_{n-1}/\Delta_n} > 0$. These coefficients are the elements in a three-diagonal *Jacobi matrix*. There is also a converse to this. If we have a three-term recurrence relation of this form and $a_k > 0$ and $b_k \in \mathbb{R}$, then the sequence of polynomials generated will be orthogonal with respect to some positive measure. This is called a *Favard theorem*.

Defining the monic polynomials $\varphi_k(x) = \phi_k(x)/\kappa_k$ if $\phi_k(x) = \kappa_k x^k + \kappa'_k x^{k-1} + \cdots$, then these satisfy

$$x\varphi_n(x) = b_{n-1}^2\varphi_{n-1}(x) + a_n\varphi_n(x) + \varphi_{n+1}(x), \quad n = 0, 1, \ldots, \quad \varphi_0 = 1, \varphi_{-1} = 0, b_n = \frac{\kappa_n}{\kappa_{n+1}}.$$

In the *Toeplitz case* one usually generalizes the operation of taking the complex conjugate on the unit circle to a more general involution on the polynomials $\sum_k p_k z^k \mapsto \sum \overline{p}_k 1/z^k$, which we shall also extend to all functions as $f(z) \mapsto f_*(z) = \overline{f(1/\overline{z})}$. It should be clear that then $\langle p, q \rangle = \langle q_*, p_* \rangle_\mu$ and also $\langle p(t), q(t) \rangle_\mu = \langle tp(t), tq(t) \rangle_\mu$, which can be combined as $\langle p, q \rangle_\mu = \langle q^*, p^* \rangle_\mu$ where $p^*(z) = z^n p_*(z)$ and $q^*(z) = z^n q_*(z)$. Because of the symmetry of the Toeplitz matrix with respect to its anti-diagonal it follows also that

$$\langle p, q \rangle_\mu = \mathbf{p}^H G_n \mathbf{q} = \mathbf{q}_*^H G_n \mathbf{p}_*$$

Therefore $\langle \phi_n, \mathcal{P}_{n-1} \rangle_\mu = 0$ implies $\langle \phi_n^*, z\mathcal{P}_{n-1} \rangle_\mu = 0$ so that one obtains the Szegő recurrence which is more easily written down for monic polynomials $\varphi_n(z) = \phi_n(z)/\kappa_n = \sqrt{\frac{\Delta_n}{\Delta_{n-1}}}\phi_n(z)$, $\varphi_0 = 1$

$$\begin{bmatrix} \varphi_{n+1}(z) \\ \varphi_{n+1}^*(z) \end{bmatrix} = \begin{bmatrix} z\varphi_n(z) - a_k\varphi_n^*(z) \\ \varphi_n^*(z) - \overline{a}_n z\varphi_n(z) \end{bmatrix} = \begin{bmatrix} 1 & -a_n \\ -\overline{a}_n & 1 \end{bmatrix} \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi_n(z) \\ \varphi_n^*(z) \end{bmatrix},$$

with $a_n = \frac{\Delta_{n-1}}{\Delta_n} \langle 1, z\varphi_n(z)\rangle_\mu = -\varphi_{n+1}(0)$ which are called the Szegő parameters or Verblunsky coefficients. The norm of the monic polynomials is $\|\varphi_n\|^2 = \frac{\Delta_n}{\Delta_{n-1}} = m_0 \prod_{k=1}^{n-1}(1-|a_k|^2)$. The positivity immediately implies that $|a_k| < 1$. Also here there is a converse: if we have a recurrence of this form with all $|a_k| < 1$ then the polynomial sequence generated is orthogonal with respect to some positive measure on $\mathbb{T}$.

Also in this case, there is a Favard-type theorem, i.e., if the polynomials satisfy a recursion of the form given above, then they are orthogonal with respect to some positive measure on the circle.

### B.1.2 Related polynomials

The recurrence relations in both cases allow for two independent solutions.

case $\mathbb{R}$ For example starting the three-term recurrence $b_{n-1}y_n + a_n y_n + b_n y_{n+1} = x y_n$ with $(y_{-1}, y_0) = (0, 1)$ gives the orthogonal polynomials or polynomials of the first kind, but with $(y_{-1}, y_0) = (1, 0)$ (and $b_{-1} = -1$) we get another independent sequence of polynomials. These are called *polynomials of the second kind* which we shall denote as $\psi_k$. This writes as

$$\begin{bmatrix} \psi_n & \phi_n \\ \psi_{n+1} & \phi_{n+1} \end{bmatrix} = \theta_n\theta_{n-1}\cdots\theta_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \theta_k(x) = \begin{bmatrix} 1 & 0 \\ 0 & b_k^{-1} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -b_{k-1} & x - a_k \end{bmatrix}.$$

Taking determinants gives the *Liouville-Ostrogradskii formula* or *determinant formula*

$$\phi_n\psi_{n+1} - \phi_{n+1}\psi_n = \frac{1}{b_n}.$$

As a special case of the *Green's formula* one obtains the *Christoffel-Darboux relation* (C-D):

$$\sum_{k=0}^{n} \phi_k(x)\phi_k(y) = \frac{b_n[\phi_n(x)\phi_{n+1}(y) - \phi_{n+1}(x)\phi_n(y)]}{y - x}.$$

If $y \to x \in \mathbb{R}$, we get the confluent formula

$$\sum_{k=0}^{n} \phi_k(x)\phi_k(y) = b_n[\phi_n(x)\phi'_{n+1}(x) - \phi_{n+1}(x)\phi'_n(x)].$$

Brezinski [9] shows that the C-D relation is equivalent with the three-term recurrence relation.

case $\mathbb{T}$ For the case of the unit circle, the previous Szegő recurrence is applied to a starting matrix whose columns give two independent solutions:

$$\begin{bmatrix} \psi_n & \varphi_n \\ \psi_n^* & \varphi_n^* \end{bmatrix} = \theta_n\theta_{n-1}\cdots\theta_0 \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \theta_k(z) = \begin{bmatrix} 1 & -a_k \\ -\overline{a}_k & 1 \end{bmatrix} \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}.$$

Also here we call $\psi_k$ *polynomials of the second kind* and the *determinant formula* is

$$\psi_n(z)\varphi_n^*(z) - \varphi_n(z)\psi_n^*(z) = 2z^{n+1}\prod_{k=0}^{n-1}(1 + |a_k|^2).$$

The analog for the *Christoffel-Darboux relation* is

$$\sum_{k=0}^{n} \frac{\varphi_k(z)\overline{\varphi_k(w)}}{\|\varphi_k\|^2} = \frac{\varphi_n^*(z)\overline{\varphi_n^*(w)} - z\overline{w}\varphi_n(z)\overline{\varphi_n(w)}}{\|\varphi_n\|^2(1 - z\overline{w})}.$$

or

$$\sum_{k=0}^{n} \phi_k(z)\overline{\phi_k(w)} = \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - z\overline{w}\phi_n(z)\overline{\phi_n(w)}}{1 - z\overline{w}}$$
$$= \frac{\phi_{n+1}^*(z)\overline{\phi_{n+1}^*(w)} - \phi_{n+1}(z)\overline{\phi_{n+1}(w)}}{1 - z\overline{w}}.$$

The confluent version (if $z = w = t \in \mathbb{T}$) results in

$$\sum_{k=0}^{n} |\phi_k(t)|^2 = t[\phi'_{n+1}(t)\overline{\phi_{n+1}(t)} - \phi^{*\prime}_{n+1}(t)\overline{\phi^*_{n+1}(t)}].$$

## B.2 Reproducing kernels

Kernels of the form $k_n(z, w) = \sum_{k=0}^{n} \phi_k(z)\overline{\phi_k(w)}$ are called *reproducing kernels* for the space $\mathcal{P}_n$ because any polynomial $p \in \mathcal{P}_n$ will be "reproduced" in the sense that $\langle k_n(z, \cdot), p\rangle_\mu = p(z)$.

Also for these kernels we can find a determinant expression:

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

This holds in both cases, but only for the *case of the circle*, $G_n$ is Toeplitz and we get some special results:

$$k_n(z, 0) = \kappa_n \phi_n^*(z), \quad k_n(0, 0) = \kappa_n^2.$$

For the *case of the real line*, $G_n$ is Hankel, and formulas like the previous ones are not possible.

## B.3 Hardy, Lebesgue and other spaces

Let $\lambda$ denote the normalized Lebesgue measure on $\mathbb{T}$, i.e., $\mathrm{d}\lambda(t) = \frac{\mathrm{d}t}{t} = \frac{\mathrm{d}\theta}{2\pi}$ if $t = e^{i\theta}$, and let $H(\mathbb{D})$ denote the space of functions analytic in the open unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. Then $H^p = H^p(\mathbb{D}, \mathrm{d}\lambda)$ for $1, p < \infty$ is the *Hardy space*

$$H^p(\mathbb{D}, \mathrm{d}\lambda) = \{f \in H(\mathbb{D}) : \|f\|_p = \sup_{0<r<1} \left[\int |f(rt)|^p \mathrm{d}\lambda(t)\right]^{1/p} < \infty\}.$$

If $p = \infty$ then $H^\infty = H^\infty(\mathbb{D}, \mathrm{d}\lambda)$ is the set of *bounded analytic functions*

$$H^\infty(\mathbb{D}, \mathrm{d}\lambda) = \{f \in H(\mathbb{D}) : \|f\|_\infty = \sup_{0<r<1} |f(rt)| < \infty\}.$$

With a function $f \in H^p$ one may associate a *boundary function*, which is the radial limit of $f(rt)$ as $r \to 1$. By *Fatou's theorem*, the radial limit can be replaced by any nontangential limit. The boundary function is in the *Lebesgue space* $L^2(\mathbb{T}, \mathrm{d}\lambda)$ (same norm as $H^2$). Conversely, if one knows the boundary function in $L^p$ to be continuous, one may recover the Hardy function in the disk using the *Poisson kernel*[10]

$$P(t, z) = \frac{t(1 - |z|^2)}{(t - z)(1 - \overline{z}t)}.$$

[10]Note that for $t = e^{i\theta} \in \mathbb{T}$, then $P(t, z) = \frac{1-|z|^2}{|t-z|^2} = \mathrm{Re}\left(\frac{t+z}{t-z}\right)$ and if we also take $z = re^{i\sigma}$, then $P(t, z) = \frac{1-r^2}{1-2r\cos(\theta-\sigma)+r^2} = P_r(\theta - \sigma)$ with $P_r(\eta) = \frac{1-r^2}{1-2r\cos\eta+r^2}$.

The construction is as follows:

$$u(z) = \int P(t,z) f(t) \mathrm{d}\lambda(t) = \mathrm{Re}\left( \int D(t,z) f(t) \mathrm{d}\lambda(t) \right), \quad D(t,z) = \frac{t+z}{t-z},$$

is an harmonic function in $\mathbb{D}$ and it can be completed (with an imaginary part) to an analytic function that is continuous in the closed disk $\overline{\mathbb{D}} = \mathbb{D} \cup \mathbb{T}$ which coincides on $\mathbb{T}$ with the boundary function $f$.

The space $H^2$ is a Hilbert space with inner product $\langle f, g \rangle_\lambda = \int \overline{f(t)} g(t) \mathrm{d}\lambda(t)$. The orthogonal polynomials are trivial and are simply given by $\{1, z, z^2, \ldots\}$. The Verblunsky coefficients are all zero and the reproducing kernel for $\mathcal{P}_n$ is $k_n(z,w) = \sum_{k=0}^n (\overline{w}z)^k$. As $n \to \infty$ we obtain the *Szegő kernel* $s(z,w) = s_w(z) = \lim_{n\to\infty} \sum_{k=0}^n (\overline{w}z)^k = \frac{1}{1-z\overline{w}}$. If is a kind of reproducing kernel for $H^2(\mathbb{D}, \mathrm{d}\lambda)$ since $\langle f, s_w \rangle_\lambda = f(w)$ for all $f \in H^2(\mathbb{D}, \mathrm{d}\lambda)$.

For general treatises on orthogonal polynomials we refer to [4], [31], [79], [96, 97], [101], [108], [121], [140], [143], [147], [141].

Polynomials may be viewed as rational functions whose poles are all fixed at infinity. By fixing a sequence of poles $\{\gamma_n\}$ in the extended complex plane (i.e., on the Riemann sphere), we obtain a theory of orthogonal rational functions. The poles can in principle be taken anywhere in the extended plane. Some of the $\gamma_k$ can be repeated, possibly an infinite number of times. The sequence is fixed once and for all, and the order in which the $\gamma_k$ occur (possible repetitions included) is also given.

# C Density and the proof of Favard's theorem

## C.1 Density for $\mathbb{T}$

For the density of the Blaschke products in $L_p(\mathbb{T})$, one can easily adapt a result of [1, p.244] to find the following.

**Theorem C.1** *Define the Blaschke products as before for $n \geq 0$ and $B_{-n} = 1/B_n$ for $n \geq 1$. Then the system $\{B_n\}_{n\in\mathbb{Z}}$ is complete in $L_p(\mathbb{T})$, for any $p \geq 1$ as well as in the space $C(\mathbb{T})$ of continuous functions on $\mathbb{T}$ if and only if $\sum(1-|\alpha_k|) = \infty$ (the Blaschke product diverges). Also the system $\{B_n\}_{n=0}^{\infty}$ is complete in $H_p$ for any $p \geq 1$ if and only if $\sum(1 - |\alpha_k|) = \infty$ (the Blaschke product diverges).*

Walsh [167, p.305-306] states

**Theorem C.2** *If the Blaschke product diverges and $f \in H_2$, then $f_n \in \mathcal{L}_n$ which interpolates $f$ in $\{w, \alpha_1, \ldots, \alpha_n\} \subset \mathbb{D}$ ($w$ is an arbitrary but fixed number in $\mathbb{D}$) will converge to $f$ uniformly on compact subsets of $\mathbb{D}$, and if $f$ is continuous up to $\mathbb{T}$, we also have uniform convergence on $\mathbb{T}$.*

For a general probability measure, Szegő's condition $\log \mu' \in L_1$ plays a role. For example, from Walsh [167, p.116,186,50,92,144] we may conclude that the polynomials are dense in $L_p(\mu)$ if and only if $\log \mu' \notin L_1$. For the rational case, the situation seems to be less simple.

**Theorem C.3** *(1) If $\sum_k(1 - |\alpha_k|) = \infty$ then $\log \mu' \notin L_1 \Leftrightarrow \{B_n\}_{n\geq 0}$ is complete in $L_2(\mu)$*
*(2) If $\log \mu' \in L_1$ then $\sum_k(1 - |\alpha_k|) = \infty \Leftrightarrow \{B_n\}_{n\geq 0}$ is complete in $H_2(\mu)$*
*(3) If $\sum_k(1 - |\alpha_k|) = \infty$ then $\{B_n\}_{n\in\mathbb{Z}}$ is complete in $L_2(\mu)$.*

Note that we have no equivalence anymore. The divergence of the Blaschke product implies completeness of $\{B_n\}_{n=0}^{\infty}$ in $H_2(\mu)$ and of $\{B_n\}_{n\in\mathbb{Z}}$ in $L_2(\mu)$, but the converse need not be true. We leave it to the reader to look up the proofs of the above theorems. The precise characterization of completeness of the Blaschke products is not totally cleared out.

## C.2 Density for $\mathbb{R}$

For the case of a measure with support on a part of $\mathbb{R}$, these are several possibilities. Among the standard one we find all of $\hat{\mathbb{R}}$, a half line $\mathbb{R}_+$, and an interval $I = [-1, 1]$. Besides this, there are many possibilities on where to place the poles of the rational functions. They can be real or complex, and inside or outside the support of the measure. For the polynomial case, this is clear and density is closely connected with the solvability of the corresponding moment problem. You may want to measure density in the space of continuous functions, in $L_p$ of $H_p$ with respect to the Lebesgue measure of with respect to the measure $\mu$, etc. Since we have not discussed in detail all the rational moment problems here, we cannot go into the details of this problem. It depends not only on the measure but also on the location and asymptotics of the poles.

Just an example from [116].

**Theorem C.4** *Let $\mu$ be supported on $[0, \infty)$ and all the poles $\alpha_k$ are in $\mathbb{C} \setminus [0, \infty)$. If*

$$\sum_{\mathrm{Im}\,(\sqrt{\alpha_k})>0} \frac{\mathrm{Im}\,(\sqrt{\alpha_k})}{1 + |\alpha_k|} = +\infty$$

*then if all the $\alpha_k$ are different (simple poles) the system*

$$\{1, (1 - x/\alpha_1)^{-1}, (1 - x/\alpha_2)^{-1}, \ldots\}$$

*is dense in uniform norm in the space of continuous functions on $[0, \infty)$ satisfying $\lim_{x\to+\infty} f(x) < \infty$.*

This and other results are direct consequences of what can be found in the book by Achieser [1, p. 254].

**Question C.1** Is there a global and systematic approach or result that catches all the results for the density of rational functions and measure on (parts of) $\hat{\mathbb{R}}$? ◯

## C.3 Proof of Favard theorem on $\mathbb{T}$

It is easily proved by induction that the $\phi_n^*$ generated by the recursion are indeed superstar conjugates of the $\phi_n$ that it generates and the parameters $\lambda_n$ are given by their expressions of Theorem 4.7. Moreover, the $\{\phi_k\}_{k=0}^n$ are a set of ORF with respect to the weight $w_n(t) = P(t, \alpha_n)/|\phi_n|^2$. This is shown as follows. Note that for $0 \le m \le n$

$$\langle \phi_n, \phi_m \rangle_{w_n} = \int \frac{\phi_{m*}(t)}{\phi_{n*}(t)} P(t, \alpha_n) d\lambda(t) = \left. \frac{B_n(z)}{B_m(z)} \frac{\phi_m^*(z)}{\phi_n^*(z)} \right|_{z=\alpha_n} = \delta_{nm}.$$

Thus $\phi_n \perp_{w_n} \mathcal{L}_{n-1}$, and because an inverse recurrence (obtained by multiplying (4.4) from the left with the inverse of the matrix $\tilde{t}_n(z)$), defines all the previous ORF uniquely, the orthonormality with respect to $w_n$ is proved.

Now for $n \to \infty$, we have a sequence of weights $w_n$, uniformly bounded ($\int w_n d\lambda = 1$), so that there is a subsequence $w_{n_k}$ which converges weakly. Thus there is some $\mu$ such that $\lim_{n\to\infty} \int f(t) w_{n_k}(t) d\lambda(t) = \int f(t) d\mu(t)$ for all functions $f$ continuous on $\mathbb{T}$. If $\sum(1-|\alpha_k|) = \infty$, then a previous completeness result shows that the measure $\mu$ is unique because the Riesz representation of a linear functional on $C(\mathbb{T})$ is unique.

Since the reproducing kernels also satisfy a recurrence relation, one may wonder whether kernels satisfying such a recursion will be reproducing with respect to some measure. A partial answer can be found in [21].

## C.4 Proof of Favard theorem on $\mathbb{R}$

There are several proofs available for the Favard theorem in this case. For the case of a measure on $\hat{\mathbb{R}}$ and poles in $\hat{\mathbb{R}}$, a proof can be found in [28, Sec. 11.9] and [22]. For arbitrary poles in the complex plane, a proof was given in [64]. The proofs as well as the conditions imposed are tedious, long and very technical and will not be repeated here.

# D Asymptotics for $\mathbb{T}$

Density problems and solution of rational moment problems are clearly related to the asymptotic behavior of the ORF, the recurrence coefficients, etc. We shall give some ideas of the problems and complications that one may expect. We shall do it only for the case of the unit circle.

## D.1 Orthogonal polynomials w.r.t. varying measures

For the convergence of the ORF we shall rely on convergence results of orthogonal polynomials with respect to varying measures OPVM. Let $\mathcal{P}_m$ we the space of polynomials of at most degree $m$ and construct a sequence of OPVM for measure $d\mu(t)/|\pi_n(t)|^2$ with $\pi_n(t) = \prod_{k=1}^n (1-\overline{\alpha}_k z)$. We denote the orthonormal ones by $\phi_{n,k}(z) = v_{n,k} z^k + \cdots$, $v_{n,k} > 0$. By our general theory they satisfy the recurrence $\phi_{n,m}(z) = e_{n,m}[z\phi_{n,m-1}(z) + \overline{\lambda}_{n,m}\phi^*_{n,m-1}(z)]$ with $\lambda_{n,m} = \overline{\phi_{n,m}(0)}/v_{n,m}$ and $e_{n,m} = (1-|\lambda_{n,m}|^2)^{-1/2} = v_{n,m}/v_{n,m-1}$. Here $\phi^*_{n,m}(z) = z^n\overline{\phi_{n,m}(1/\overline{z})}$. From these OP we can construct a set of ORF for $\mathcal{L}_n$, but as $n$ increases, the whole set of ORF will change.

**Exercise D.1** Denote $f_{n,m}(z) = t_{n,m}\phi_{n,m}(z)/\pi_n(z)$ with $t_{n,m} \in \mathbb{T}$ for an appropriate normalization. Prove that the $\{f_{n,m} : m = 0, \ldots, n\}$ is a set of orthonormal rational functions for $\mathcal{L}_n$ with respect to $\mu$. They are obtained by orthonormalizing the basis $\{z^k/\pi_n(z) : k = 0, \ldots, n\}$. Show also that

$$h_n(z) = \frac{z\phi_{n,n}(z) - \frac{\phi_{n,n}(\alpha_n)}{\phi^*_{n,n}(\alpha_n)}\alpha_n\phi^{\bar{*}}_{n,n}(z)}{(z-\alpha_n)\pi_n(z)} \in \mathcal{L}_n \setminus \mathcal{L}_{n-1} \tag{4.1}$$

and that $\langle h_n, g\rangle_\mu = 0$ for all $g \in \mathcal{L}_{n-1}$, so that there is some constant $c_n$ with $|c_n| = |\phi^*_n(0)|$ such that $h_n = c_n\phi_n$. $\diamond$

We need the following conditions for the probability measure $\mu$ and the point set $A = \{\alpha_1, \alpha_2, \ldots\}$:

1. $\mu' > 0$ a.e. $(\lambda)$ (Erdős-Turán condition)
2. $\sum_{n=1}^\infty (1-|\alpha_n|) = \infty$ (BD = Blaschke divergence condition)

The first condition is denoted as $\mu \in ET$. If it satisfies the stronger Szegő condition $\log \mu' \in L_1$, we denote it by $\mu \in SZ$. The second condition is denoted as $A \in BD$, while if it is replaced by the stronger condition that $A$ is compactly contained in $\mathbb{D}$ (i.e., the $\alpha_k$ stay away from the boundary), we denote it as $A \in CC$. The following results are borrowed from the work of Guillermo Lopez [117, 118, 119].

**Theorem D.1** *If $\mu \in ET$ and $A \in BD$ then with our previous notation*

1. $\lim_{n\to\infty} \lambda_{n,n+k+1} = 0$.
2. $\lim_{n\to\infty} v_{n,n+k+1}/v_{n,n+k} = 1$.
3. $\lim_{n\to\infty} \phi_{n,n+k+1}(z)/\phi_{n,n+k}(z) = z$ *locally uniformly in* $\mathbb{E}$.
4. $\lim_{n\to\infty} \phi^*_{n,n+k+1}(z)/\phi^{\bar{*}}_{n,n+k}(z) = 1$ *locally uniformly in* $\mathbb{D}$.

5. $\lim_{n\to\infty} \phi^*_{n,n+k}(z)/\phi_{n,n+k}(z) = 0$ *locally uniformly in* $\mathbb{E}$.

*If* $\mu \in SZ$, $A \in CC$ *and* $\sigma$ *is the outer spectral factor of* $\mu$, *then*

6. $\lim_{n\to\infty} \phi^*_{n,n+k}(z)/\pi_n(z) = 1/\sigma(z)$ *locally uniformly in* $\mathbb{D}$.

## D.2 Szegő's condition and asymptotics

We can now show that the following holds.

**Theorem D.2** *If* $\mu \in SZ$ *and* $A \in BD$, *then we have the following (l.u. means locally uniformly)*

$$\lim_{n\to\infty} k_n(z,w) = \frac{1}{\overline{\sigma(w)}(1-\overline{w}z)\sigma(z)}, \quad l.u.\ (z,w) \in \mathbb{D}\times\mathbb{D},$$

$$\lim_{n\to\infty} (1-\overline{\alpha}_n z)\frac{\phi^*_n(z)}{\phi^*_n(0)} = \frac{\sigma(0)}{\sigma(z)}, \quad l.u.\ z \in \mathbb{D},$$

$$\lim_{n\to\infty} \phi_n(z) = \lim_{n\to\infty} \frac{\phi_n(z)\phi^*_n(0)}{\phi^*_n(z)} = \lim_{n\to\infty} \frac{\phi_n(z)}{\phi^*_n(z)} = 0, \quad l.u.\ z \in \mathbb{D}.$$

*If* $\mu \in SZ$, $\sigma(0) > 0$, *and* $A \in CC$ *then*

$$\lim_{n\to\infty} \eta_n \phi^*_n(z)\frac{1-\overline{\alpha}_n z}{\sqrt{1-|\alpha_n|^2}} = \frac{1}{\sigma(z)}, \quad \eta_n = \frac{|\phi^*_n(0)|}{\phi^*_n(0)} \in \mathbb{T}, \quad l.u.\ z \in \mathbb{D}.$$

**Proof.** For the first relation, note that since the $f_{n,k}$ form an orthonormal basis for $\mathcal{L}_n$, we have

$$k_n(z,w) = \sum_{k=0}^{n} f_{n,k}(z)\overline{f_{n,k}(w)} = \frac{\sum_{k=0}^{n} \phi_{n,k}(z)\overline{\phi_{n,k}(w)}}{\pi_n(z)\overline{\pi_n(w)}}.$$

Then use the CD relation for the $\phi_{nk}$ and divide by $f^*_{n,n}(z)\overline{f^*_{n,n}(w)}$:

$$\frac{k_n(z,w)}{f^*_{n,n}(z)\overline{f^*_{n,n}(w)}} = \frac{1}{1-\overline{w}z} - \frac{z\overline{w}}{1-\overline{w}z}\left(\frac{\phi_{n,n}(z)}{\phi^{\bar{*}}_{n,n}(z)}\right)\overline{\left(\frac{\phi_{n,n}(w)}{\phi^{\bar{*}}_{n,n}(w)}\right)}. \tag{4.2}$$

The last two factors go to 0, while for some $\eta_n \in \mathbb{T}$, $\lim_{n\to\infty} \eta_n f^*_{n,n}(z) = 1/\sigma(z)$, which proves the first result.
Since $h_n = c_n\phi_n$, we have by our expression for $h_n$

$$(1-\overline{\alpha}_n z)\phi^*_n(z) = \frac{1}{\overline{c}_n}\frac{\phi^{\bar{*}}_{n,n}(z)}{\pi_n(z)}\left[1-\overline{\eta}_n\overline{\alpha}_n z\frac{\phi_{n,n}(z)}{\phi^{\bar{*}}_{n,n}(z)}\right], \quad \eta_n = \frac{\phi_{n,n}(\alpha_n)}{\phi^{\bar{*}}_{n,n}(\alpha_n)}. \tag{4.3}$$

Recall that $\phi_{n,n}(z)/\phi^{\bar{*}}_{n,n}(z) \to 0$ l.u. in $\mathbb{D}$, and being a Blaschke product it is bounded by 1 in $\mathbb{D}\cup\mathbb{T}$ so that $|\eta_n| \le 1$. Hence the term between square brackets goes to 1 as $n \to \infty$. On the other hand, $\phi^{\bar{*}}_{n,n}(z)/\pi_n(z)$ converges to $1/\sigma(z)$ l.u. in $\mathbb{D}$, up to some normalizing constant. But this constant will cancel against the same constant in the denominator which is obtained by setting $z = 0$. This proves the second result.
Taking $z = w$ in the first relation implies $\sum_{k=1}^{n} |\phi_k(z)|^2$ converges for $n \to \infty$. Thus $\lim_n \phi_n(z) = 0$. When multiplying the inverse of the second relation of this theorem with $\phi_n(z)$, and noting that $(1-\overline{\alpha}_n z)^{-1}$ is uniformly bounded if $z$ is in a compact subset of $\mathbb{D}$,

we also get the second limit.
For the third one, we note that $|\phi_n^*(0)|$ is uniformly bounded away from 0, because the CD relation implies $|\phi_n^*(0)|^2 \geq 1$.
For the last limit, note that $\eta_n$ is used to normalize $\phi_n^*$ to make it positive in $z = 0$. By the CD relation for $z = w = 0$

$$\frac{|\phi_n^*(0)|^2}{1-|\alpha_n|^2} = k_{n-1}(0,0) + \frac{|\phi_n(0)|^2}{1-|\alpha_n|^2}.$$

Because $\phi_n(0) \to 0$, and the $\alpha_k$ stay away from the circle, the last term goes to 0, and we know that the second one goes to $|\sigma(0)|^{-2}$. Thus $|\phi_n^*(0)|/\sqrt{1-|\alpha_n|^2} \to 1/\sigma(0)$. In combination with the second limit of this theorem, the result follows. □

**Remark D.1** For the case of a measure $\tilde{\mu}$ on $I = [-1,1]$ and poles $\beta_k$ in $\hat{\mathbb{C}}^I$, but staying away form $I$, denote $\cos\theta = x = \mathfrak{j}(z) = \mathfrak{j}(e^{\mathbf{i}\theta})$, $\mu'(z) = \tilde{\mu}'(\cos\theta)|\sin\theta|$ and $\alpha_k = \mathfrak{j}^{inv}(\beta_k)$, and $\tilde{\phi}_k$ the ORF for $\tilde{\mu}$ on $I$ and poles $\{\beta_k\}_k$, then if $\int_I \frac{\tilde{\mu}'(x)}{\sqrt{1-x^2}}\mathrm{d}x > -\infty$ (equivalent with $\mu \in SZ$) then

$$\lim_{n\to\infty} \lambda_n B_n(z) \frac{1-\overline{\beta}_n x}{\sqrt{1-|\beta_n|^2}} \tilde{\phi}_n(x) = \frac{1}{\sqrt{2\pi}\sigma(z)}$$

for an appropriate unimodular constant $\lambda_n$ and $\sigma$ the spectral factor of $\mu$. For this and other related results see [66, Thm. 4.3].

When $\mu \notin SZ$, then $\sigma$ need not be defined, and the previous relations can not be obtained. A way to avoid the $\sigma$ is to consider ratio asymptotics, so that the $\sigma$ cancel out. These are weaker results in the sense that if $\mu \in SZ$, then the ratio asymptotics are almost immediately obtained from the stronger asymptotics that were previously obtained. We use again the OPVM with measure $\mu/|\pi_n|^2$. Denoting the reproducing kernel for $\mathcal{P}_n$ w.r.t. this measure as $k_{n,n}(z,w)$, it is easily seen that $k_{n,n}(z,0) = v_n \phi_{n,n}^{\bar{*}}(z)$. Also the usual rational kernel for $\mathcal{L}_n$ satisfies $k_n(z,0) = k_{n,n}(z,0)/\pi_n(z)$.

**Exercise D.2** Applying Theorem 4.16 to the OPVM, prove that the inner product in $\mathcal{L}_n$ w.r.t. $d\mu(t)$ and w.r.t. $v_n^2/|k_n(t,0)|^2 d\lambda(t)$ is the same. Of course, this is also a direct consequence of the results in section 4.7. ◇

**Theorem D.3** *If $\mu \in ET$ and $A \in BD$, then*

$$\lim_{n\to\infty} \frac{k_n(z,0)}{k_{n+1}(z,0)} = 1, \quad l.u.\ z \in \mathbb{D}.$$

*and with the superstar referring to $z$,*

$$\lim_{n\to\infty} k_n^*(z,0)/k_n(z,0) = 0, \quad l.u.\ z \in \mathbb{D}.$$

**Proof.** Set $g_n(z) = k_n(z,0)/k_{n+1}(z,0) - v_n/v_{n+1}$, then by the previous exercise we see that

$$\int |g_n(t)|^2 d\lambda(t) = 2\frac{v_n^2}{v_{n+1}^2}\left(1 - \frac{v_n}{v_{n+1}}\right). \tag{4.4}$$

On the other hand, because $g_n$ is analytic in $\mathbb{D}\cup\mathbb{T}$, we have by the Poisson formula $g_n(z) = \int P(t,z)g_n(t)d\lambda(t)$. For $t\in\mathbb{T}$, $|P(t,z)|\le M$ holds uniformly for $z$ in a compact subset of $\mathbb{D}$. Thus we also have $|g_n(z)|\le M\int|g_n(t)|d\lambda(t)$ and also $|g_n(z)|^2\le M^2\int|g_n(t)|^2d\lambda(t)$ so that

$$|g_n(z)|^2 \le 2M^2\frac{v_n^2}{v_{n+1}^2}\left(1-\frac{v_n}{v_{n+1}}\right). \tag{4.5}$$

Thus $\frac{k_n(z,0)}{k_{n+1}(z,0)}\to 1$ iff $\frac{v_n}{v_{n+1}}\to 1$. Obviously $k_{n-1}(0,0)/k_n(0,0)\le 1$, but using the CD relation we also have

$$\frac{k_{n-1}(0,0)}{k_n(0,0)} = \frac{|\phi_n^*(0)|^2-|\phi_n(0)|^2}{|\phi_n^*(0)|^2-|\alpha_n|^2|\phi_n(0)|^2} \ge \frac{|\phi_n^*(0)|^2-|\phi_n(0)|^2}{|\phi_n^*(0)|^2-|\phi_n(0)|^2} = 1.$$

So that $\lim_{n\to\infty} v_n^2/v_{n+1}^2 = 1$, which proves the first part.

For the second relation, note that $k_n(z,0) = \phi_{n,n}^{\bar *}(z)v_n\eta_n/\pi_n(z)$ and hence $k_n^*(z,0) = \phi_{n,n}(z)v_n\overline{\eta}_n/\pi_n(z)$ with $\eta_n\in\mathbb{T}$. Take the ratio and use $\phi_{n,n}(z)/\phi_{n,n}^{\bar *}(z)\to 0$ l.u. in $\mathbb{D}$, to conclude the proof. □

**Question D.1** Can we replace the value 0 by some arbitrary complex number $w$ in the previous theorem? What would the be the limit? ◯

**Theorem D.4** *If $\mu\in ET$ and $A\in CC$, then*

$$\lim_{n\to\infty}\frac{\zeta_n(z)\phi_n(z)}{\phi_n^*(z)} = \lim_{n\to\infty}\frac{\phi_n(z)}{\phi_n^*(z)} = \lim_{n\to\infty}\lambda_n = 0, \quad \textit{l.u. in } \mathbb{D}.$$

*Denote $\Phi_n^*(z) = \phi_n^*(z)/\phi_n^*(0)$, then under the previous conditions*

$$\lim_{n\to\infty}\frac{\Phi_{n+1}^*(z)(1-\overline{\alpha}_{n+1}z)}{\Phi_n^*(z)(1-\overline{\alpha}_n z)} = 1, \quad \textit{l.u. in } \mathbb{D}.$$

**Proof.** Repeat the proof of the second part in Theorem D.2. I.e., use exercise 4.1 giving $\phi_n = c_n h_n$ and its superstar conjugate, to find expressions for $(z-\alpha_n)\phi_n(z)$ and $(1-\overline{\alpha}_n z)\phi_n^*(z)$ and take their ratio. Then using $\phi_{n,n}(z)/\phi_{n,n}^{\bar *}(z)\to 0$ from theorem D.1 leads to the first conclusion. Note that we need $\phi_{n,n}(\alpha_n)/\phi_{n,n}^{\bar *}(\alpha_n)$ to go to zero, which can only be guaranteed when $A\in CC$. It is also needed for the convergence of $\phi_n(0)/\phi_n^*(0)$ because this will follow form Theorem D.3 or the previous exercise if $A$ is in a compact subset of $\mathbb{D}$. For the last equality note that $|\lambda_n| = |\phi_n(\alpha_{n-1})/\phi_n^*(\alpha_{n-1})|$.

For the second formula, write the second CD formula for $k_n(z,0)$ and its superstar conjugate and eliminate $\phi_n(z)$ to obtain

$$\Phi_n^*(z)(1-\overline{\alpha}_n z) = \frac{\phi_n^*(z)(1-\overline{\alpha}_n z)}{\phi_n^*(0)} = \frac{k_n(z,0)}{v_n^2}\left[1-\overline{\alpha}_n z\frac{\overline{\phi_n(0)}}{\phi_n^*(0)}\frac{k_n^*(z,0)}{k_n(z,0)}\right].$$

Note that the term in square brackets goes to 1 l.u. in $\mathbb{D}$ as $n\to\infty$. Rewrite this for $n$ replaced by $n+1$ and take their ratio. Because $k_{n+1}(z,0)/k_n(z,0)$ and $v_{n+1}/v_n$ go to 1 as $n\to\infty$, the second result follows. □

## D.3 Szegő's problem

This problem is essentially the problem of linear prediction. Recall that the linear prediction problem can be formulated as finding $f \in H_2(\mu)$ such that $\|f\|_\mu^2$ is minimal with the side condition that $f(0) = 1$. This $f$ represented the prediction error. (Note that we have replaced the $z$ of the first section by $1/z$ to be in line with our discussion of ORF, which is obviously a matter of convention.) We consider a slightly more general situation where we replace the constraint by $f(w) = 1$ for some $w \in \mathbb{D}$. This is a problem that can be solved in any reproducing kernel Hilbert space.

**Theorem D.5** *Let $H$ be a reproducing kernel Hilbert space with kernel $k(z, w)$, then the minimum of $\|f\|_\mu^2$ with $f(w) = 1$ is obtained for $f(z) = K(z) \equiv k(z, w)/k(w, w)$ and the minimum is $1/k(w, w)$.*

**Proof.** Obviously $\|K(z)\|_\mu^2 = 1/k(w, w)$. For any other $f$ satisfying $f(w) = 1$ for which $\|f\|_\mu^2 = m$ we have $0 \leq \|f - K\|_\mu^2 = \|f\|_\mu^2 + 1/k(w, w) - 2\mathrm{Re}\,\langle f, K\rangle_\mu = m - 1/k(w, w)$. Thus $m \geq 1/k(w, w)$. □

From the polynomial case, it is known that the reproducing kernel for $H_2(\mu)$ is given by the Szegő kernel $s(z, w) = [(1 - \overline{w}z)\sigma(z)\overline{\sigma(w)}]^{-1}$ where $\sigma$ is the outer spectral factor of $\mu$, which we suppose to be normalized by $\sigma(0) > 0$. Thus we have a completely predictable sequence if the minimum $|\sigma(0)|^2$ of $\|f\|_\mu^2$, $f(0) = 1$ is zero, i.e. since $|\sigma(0)|^2 = \exp\{\int \log \mu'(t) d\lambda(t)\}$, if $\int \log \mu' d\lambda = -\infty$, and thus $\log \mu' \notin L_1$. If we do not have complete predictability, it is still a valuable objective to find the best possible predictor by minimizing the prediction error.

Instead of minimizing over the complete past (i.e. all of $H_2(\mu)$), we may be less ambitious and start by minimizing over a finite dimensional subspace, e.g. $\mathcal{L}_n$. There the minimum of the general problem is $k_n(z, w)/k_n(w, w)$.

The advantage of computing these approximants using ORF over the computation with OPVM is that they are easily computed recursively. If the approximation is not good enough for a certain $n$, then increasing $n$ by 1 requires just one more step of the recurrence relation. For the OPVM, increasing $n$ by 1, would mean that we have to start the computations all over again. Note also that we have at every step an estimate of the prediction error which is $1/k_n(0, 0)$ or $1/k_n(w, w)$ for the general problem. Thus if the $|\phi_n(w)|$ do not go to zero fast enough, then $k_n(w, w)$ will go to $\infty$ and the error will go to zero, but under the conditions of our theorems, the $k_n(w, w)$ will converge, so that the error will not go to zero. It will be bounded from below by $(1 - |w|^2)|\sigma(w)|^2$.

We also note the following result

**Theorem D.6** *Let $\mu \in SZ$, $w \in \mathbb{D}$ a fixed number, and let $s_w(z) = s(z, w)$ be the Szegő kernel. Consider the problem $\min \|f - s_w\|_\mu^2$ over all $f \in \mathcal{L}_n$, then the solution is $k_n(z, w)$ and the minimum is $s_w(w) - k_n(w, w)$. If $A \in BD$ then $\lim_{n\to\infty} \|k_n(t, w) - s_w(t)\|_\mu = 0$.*

**Proof.** Suppose $f(w) = a$, then $\|s_w - f\|_\mu^2 = \|s_w\|_\mu^2 + \|f\|_\mu^2 - 2\mathrm{Re} f(w)$. Thus we have to minimize $\|f\|_\mu^2 - 2\mathrm{Re}\, a$ over $f \in \mathcal{L}_n$ with $f(w) = a$ and minimize the result over all possible $a$. In other words, we have to find the infimum over $a$ of $|a|^2/k_n(w, w) - 2\mathrm{Re}\, a$ and this is given by $a = k_n(w, w)$. This proves the first part. If $A \in BD$ then as $n \to \infty$, $\mathcal{L}_n$ becomes dense in $H_2(\mu)$, which means that the error in $H_2(\mu)$ goes to 0. □

# References


[1] N.I. Achieser. *Theory of approximation*. Frederick Ungar Publ. Co., New York, 1956.

[2] H. Akçay and B. Ninness. Rational basis functions for robust identification from frequency and time domain measurements. *Automatica*, 34(9):1101–1117, 1998.

[3] H. Akçay and B. Ninness. Orthonormal basis functions for continuous-time systems and $L_p$ convergence. *Math. Control Signals Systems*, 12:295–305, 1999.

[4] N.I. Akhiezer. *The classical moment problem and some related questions in analysis*. Hafner, New York, 1965.

[5] C. Bonan-Hamada. *Orthogonal Laurent polynomials and indeterminate strong Stieltjes moment problems*. PhD thesis, University of Colorado, Boulder, 1994.

[6] C. Bonan-Hamada, W.B. Jones, and O. Njåstad. Natural solutions of strong Stieltjes moment problems derived from PC-fractions. In W.B. Jones and A.S. Ranga, editors, *Orthogonal functions, moment theory and continued fractions*, volume 199 of *Lecture Notes in Pure and Appl. Math.*, pages 15–30. Marcel Dekker, 1998.

[7] C. Bonan-Hamada, W.B. Jones, and O. Njåstad. Zeros of orthogonal Laurent polynomials and solutions of the strong Stieltjes moment problem. *J. Comput. Appl. Math.*, 235(4):895–903, 2010.

[8] C. Bonan-Hamada, W.B. Jones, O. Njåstad, and W.J. Thron. A class of indeterminate strong Stieltjes moment problems with discrete distributions. In W.B. Jones and A.S. Ranga, editors, *Orthogonal functions, moment theory and continued fractions*, volume 199 of *Lecture Notes in Pure and Appl. Math.*, pages 31–55. Marcel Dekker, 1998.

[9] C. Brezinski. A direct proof of the Christoffel-Darboux identity and its equivalence to the recurrence relationship. *J. Comput. Appl. Math.*, 32:17–25, 1990.

[10] A. Bultheel and M.J. Cantero. A matricial computation of rational quadrature formulas on the unit circle. *Numer. Algorithms*, 52(1):47–68, 2009. Online since 13 December 2008.

[11] A. Bultheel and P. Carrette. Algebraic and spectral properties of general Toeplitz matrices. *SIAM J. Control Optim.*, 41(5):1413–1439, 2003.

[12] A. Bultheel, R. Cruz-Barroso, K. Deckers, and P. González-Vera. Rational Szegő quadratures associated with Chebyshev weight functions. *Math. Comp.*, 78:1031–1059, 2009. Online since December 9, 2008.

[13] A. Bultheel, R. Cruz-Barroso, P. González-Vera, and F. Perdomo-Pío. On the computation of symmetric Szegő-type quadrature formulas. In A. Elipe, J.I. Montijano, and L. Rández, editors, *Proceedings of the MCalv6(5) Conference, Zaragoza, 7-9 September 2009*, volume 33 of *Monografías de la Real Academia de Ciencias Exactas Físicas Químicas y Naturales de Zaragoza*, pages 177–196, 2010.

[14] A. Bultheel, R. Cruz-Barroso, and Marc Van Barel. On Gauss-type quadrature formulas with prescribed nodes anywhere on the real line. *Calcolo*, 47(1):21–48, 2010. (Published online since 29 Sept. 2009).

[15] A. Bultheel, L. Daruis, and P. González-Vera. A connection between quadrature formulas on the unit circle and the interval $[-1, 1]$. *J. Comput. Appl. Math.*, 132(1):1–14, 2001.

[16] A. Bultheel, L. Daruis, and P. González-Vera. Positive interpolatory quadrature formulas and para-orthogonal polynomials. *J. Comput. Appl. Math.*, 179(1-2):97–119, 2005.

[17] A. Bultheel, L. Daruis, and P. González-Vera. Quadrature formulas on the unit circle with prescribed nodes and maximal domain of validity. *J. Comput. Appl. Math.*, 231(2):948–963, 2009. Online since 20 May 2009.

[18] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. A moment problem associated to rational Szegő functions. *Numer. Algorithms*, 3:91–104, 1992.

[19] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Quadrature formulas on the unit circle based on rational functions. *J. Comput. Appl. Math.*, 50:159–170, 1994.

[20] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Recurrence relations for orthogonal functions. In S.C. Cooper and W.J. Thron, editors, *Continued Fractions and Orthogonal Functions*, volume 154 of *Lecture Notes in Pure and Appl. Math.*, pages 24–46. Marcel Dekker, 1994.

[21] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Favard theorem for reproducing kernels. *J. Comput. Appl. Math.*, 57(1/2):57–76, 1995.

[22] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. A Favard theorem for rational functions with poles on the unit circle. *East J. Approx.*, 3:21–37, 1997.

[23] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Rates of convergence of multipoint rational approximants and quadrature formulas on the unit circle. *J. Comput. Appl. Math.*, 77:77–102, 1997.

[24] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. A rational moment problem on the unit circle. *Methods Appl. Anal.*, 4(3):283–310, 1997.

[25] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions and interpolatory product rules on the unit circle. I. Recurrence and interpolation. *Analysis*, 18:167–183, 1998.

[26] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions and interpolatory product rules on the unit circle. II. Quadrature and convergence. *Analysis*, 18:185–200, 1998.

[27] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Elements of a theory of orthogonal rational functions. *Rev. Acad. Canaria Cienc.*, 9(1-2):127–152, 1999.

[28] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. *Orthogonal rational functions*, volume 5 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 1999.

[29] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions and interpolatory product rules on the unit circle. III. Convergence of general sequences. *Analysis*, 20:99–120, 2000.

[30] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Determinacy of a rational moment problem. *J. Comput. Appl. Math.*, 133(1-2):241–252, 2001.

[31] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Quadrature and orthogonal rational functions. *J. Comput. Appl. Math.*, 127(1-2):67–91, 2001. Invited paper.

[32] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. A rational Stieltjes moment problem. *Appl. Math. Comput.*, 128(2-3):217–235, 2002.

[33] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions and quadrature on the real half line. *J. Complexity*, 19(3):212–230, 2003.

[34] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Positivity of continued fractions associated with rational Stieltjes moment problems. *Rocky Mountain J. Math.*, 33(2):609–627, 2003.

[35] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Separation of zeros of para-orthogonal rational functions. *Rev. Acad. Canaria Cienc.*, 16(1-2):9–14, 2004.

[36] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions on the real half line with poles in $[-\infty, 0]$. *J. Comput. Appl. Math.*, 179(1-2):121–155, 2005.

[37] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions on the unit circle: from the scalar to the matrix case. In W. Van Assche and F. Marcellán, editors, *Orthogonal Polynomials and Special Functions*, volume 1883 of *Lecture Notes in Math.*, pages 187–228. Springer Verlag, 2006.

[38] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. An indeterminate rational moment problem and Carathéodory functions. *J. Comput. Appl. Math.*, 219(2):359–369, 2008. Online since 11-July-2008.

[39] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Computation of rational Szegő-Lobatto quadrature formulas. *Appl. Numer. Math.*, 60(12):1251–1263, 2010. Online since 25 May 2010.

[40] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Growth properties of Nevanlinna matrices for rational moment problems. *J. Approx. Theory*, 162(12):2184–2201, 2010. Online since 14 July 2010.

[41] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Rational quadrature formulas on the unit circle with prescribed nodes and maximal domain of validity. *IMA J. Numer. Anal.*, 30(4):940–963, 2010. Online since August 21, 2009.

[42] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Natural solutions of rational Stieltjes moment problems. *J. Math. Anal. Appl.*, 377(2):571–583, 2011. Online since November 26, 2010.

[43] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Quadratures associated with pseudo-orthogonal rational functions on the real half line with poles in $[-\infty, 0]$. *J. Comput. Appl. Math.*, 2012. Online since 12 July 2012.

[44] A. Bultheel and M. Van Barel. Vector orthogonal polynomials and least squares approximation. *SIAM J. Matrix Anal. Appl.*, 16(3):863–885, 1995.

[45] A. Bultheel, M. Van Barel, Y. Rolain, and R. Pintelon. Numerically robust transfer function modelling from noisy frequency domain data. *IEEE Trans. Automat. Control*, 50(11):1835–1839, 2005.

[46] A. Bultheel, M. Van Barel, Y. Rolain, and J. Schoukens. Numerically robust rational approximation for identification. In IEEE, editor, *40th IEEE Conference on Decision and Control, Orlando FL, December 4-7, 2001*, volume 5, pages 4770–4775. IEEE, 2001. Paper PrP06-4.

[47] A. Bultheel, M. Van Barel, and P. Van gucht. Orthogonal bases in discrete least squares rational approximation. *J. Comput. Appl. Math.*, 164-165:175–194, 2004.

[48] A. Bultheel, P. Van gucht, and M. Van Barel. Orthogonal rational functions with real coefficients and semiseparable matrices. *J. Comput. Appl. Math.*, 233(5):1192–1201, 2010. Online since 17 March 2009.

[49] M.J. Cantero, L. Moral, and L. Velázquez. Measures and para-orthogonal polynomials on the unit circle. *East J. Approx.*, 8:447–464, 2002.

[50] M.J. Cantero, L. Moral, and L. Velázquez. Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.*, 362:29–56, 2003.

[51] M.J. Cantero, L. Moral, and L. Velázquez. Minimal representations of unitary operators and orthogonal polynomials on the unit circle. *Linear Algebra Appl.*, 408:40–65, 2005.

[52] M.J. Cantero, L. Moral, and L. Velázquez. Measures on the unit circle and unitary truncations of unitary operators. *J. Approx. Theory*, 139(1-2):430–468, 2006.

[53] K. Castillo, L. Garza, and F. Marcellán. Linear spectral transformations, Hessenberg matrices, and orthogonal polynomials. *Rend. Circ. Mat. Palermo Serie II Suppl. 82*, pages 1–23, 2010.

[54] A. Chesnokov, K. Deckers, and M. Van Barel. A numerical solution of the constrained weighted energy problem. *J. Comput. Appl. Math.*, 235(4):950–965, 2010.

[55] R. Cruz-Barroso, L. Daruis, P. González-vera, and O. Njåstad. Quadrature rules for periodic integrands. Bi-orthogonality and para-orthogonality. *Annales Mathematicae et Informaticae*, 32:5–44, 2005.

[56] R. Cruz-Barroso, P. González-Vera, and O. Njåstad. On bi-orthogonal systems of trigonometric functions and quadrature formulas for periodic integrands. *Numer. Algorithms*, 44:309–333, 2007.

[57] R. Cruz-Barroso, P. González-Vera, and F. Perdomo-Pío. An application of Szegő polynomials to the computation of certain weighted integrals on the real line. *Numer. Algorithms*, 52(3):273–293, 2009.

[58] B. de la Calle Ysern. Error bounds for rational quadrature formulae of analytic functions. *Numer. Math.*, 101:251–271, 2005.

[59] B. de la Calle Ysern and P. González-Vera. Rational quadrature formulas on the unit circle with arbitrary poles. *Numer. Math.*, 107(4):559–587, 2007.

[60] K. Deckers and A. Bultheel. Rational Krylov sequences and orthogonal rational functions. Technical Report TW499, Department of Computer Science, K.U.Leuven, August 2007.

[61] K. Deckers and A. Bultheel. Associated rational functions based on a three-term recurrence relation for orthogonal rational functions. *IAENG Int. J. Appl. Math.*, 38:214–222, 2008.

[62] K. Deckers and A. Bultheel. Orthogonal rational functions, associated rational functions and functions of the second kind. In S.I. Ao, L. Gelman, D.W.L. Hukins, A. Hunter, and A.M. Korsunsky, editors, *Proceedings of the World Congress on Engineering 2008, vol. 2*, Lecture Notes in Engineering and Computer Science, pages 838–843, London, United Kingdom, 2008. Newswood Limited, International Association of Engineers. CDROM, isbn 978-988-17012-3-7; Print: isbn: 978-988-98671-9-5.

[63] K. Deckers and A. Bultheel. Orthogonal rational functions and rational modifications of a measure on the unit circle. *J. Approx. Theory*, 157:1–18, 2009. Online since 13 October 2008.

[64] K. Deckers and A. Bultheel. Orthogonal rational functions with complex poles: The Favard theorem. *J. Math. Anal. Appl.*, 356(2):764–768, August 2009.

[65] K. Deckers and A. Bultheel. Recurrence and asymptotics for orthogonal rational functions on an interval. *IMA J. Numer. Anal.*, 29:1–23, 2009. Online since January 29, 2008.

[66] K. Deckers and A. Bultheel. Rational interpolation I: Least squares convergence. *J. Math. Anal. Appl.*, 395(2):455–464, 2012. published online 26 May 2012.

[67] K. Deckers, A. Bultheel, R. Cruz-Barroso, and F. Perdomo-Pío. Positive rational interpolatory quadrature formulas on the unit circle and the interval. *Appl. Numer. Math.*, 60(12):1286–1299, 2010. Online since 30 March 2010.

[68] K. Deckers, A. Bultheel, and F. Perdomo-Pío. Rational Gauss-Radau and rational Szegő-Lobatto quadrature on the interval and the unit circle respectively. *Jaen J. Approx.*, 3(1):15–66, 2011.

[69] K. Deckers, A. Bultheel, and J. Van Deun. A generalized eigenvalue problem for quasi-orthogonal rational functions. *Numer. Math.*, 117(3):463–506, 2010. Online since 18 Jan. 2011.

[70] K. Deckers, M.J. Cantero, L. Moral, and L Velázquez. An extension of the associated rational functions on the unit circle. *J. Approx. Theory*, 163(4):524–546, 2011.

[71] K. Deckers, J. Van Deun, and A. Bultheel. Computing rational Gauss-Chebyshev quadrature formulas with complex poles. In B. H. V. Topping, G. Montero, and R. Montenegro, editors, *Proceedings of the Fifth International Conference on Engineering Computational Technology*, Kippen, Stirlingshire, United Kingdom, 2006. Civil-Comp Press. Paper 30.

[72] K. Deckers, J. Van Deun, and A. Bultheel. An extended relation between orthogonal rational functions on the unit circle and the interval $[-1, 1]$. *J. Math. Anal. Appl.*, 334(2):1260–1275, 2007.

[73] K. Deckers, J. Van Deun, and A. Bultheel. Rational Gauss-Chebyshev quadrature formulas for complex poles outside $[-1, 1]$. *Math. Comp.*, 77(262):967–983, 2008. Online since 28 September 2007.

[74] K. Deckers, J. Van Deun, and A. Bultheel. Computing rational Gauss-Chebyshev quadrature formulas with complex poles: the algorithm. *Adv. Eng. Softw.*, 40(8):707–717, 2009. Online 7 January 2009.

[75] C. Díaz-Mendoza, P. González-Vera, and M. Jiménez-Paiz. Orthogonal Laurent polynomials corresponding to certain strong Stieltjes distributions with applications to numerical quadratures. *Math. Comp.*, 75:281–305, 2005.

[76] C. Díaz-Mendoza, P. Gonz'alez-Vera, M. Jiménez-Paiz, and F. Cala-Rodriguez. Orthogonal Laurent polynomials corresponding to certain strong Stieltjes distributions with applications to numerical quadratures. *Math. Comp.*, 75(253):281–305, 2006.

[77] C. Díaz-Mendoza, P. González-Vera, M. Jiménez Paiz, and O. Njåstad. Orthogonality and recurrence for ordered laurent polynomial sequences. *J. Comput. Appl. Math.*, 235(4):982–997, 2010.

[78] P.L. Duren. *The theory of $H^p$ spaces*, volume 38 of *Pure and Applied Mathematics*. Academic Press, New York, 1970.

[79] G. Freud. *Orthogonal polynomials*. Pergamon Press, Oxford, 1971.

[80] B. Fritzsche, B. Kirstein, and A. Lasarow. On a moment problem for rational matrix-valued functions. *Linear Algebra Appl.*, 372:1–31, 2003.

[81] B. Fritzsche, B. Kirstein, and A. Lasarow. Orthogonal rational matrix-valued functions on the unit circle. *Math. Nachr.*, 278(5):525–553, 2005.

[82] B. Fritzsche, B. Kirstein, and A. Lasarow. Szegő pairs of orthogonal rational matrix-valued functions on the unit circle. In M. Langer, A. Luger, and H. Woracek, editors, *Operator theory and indefinite inner product spaces*, volume 163 of *Oper. Theory: Adv. Appl.*, pages 163–189. Birkhäuser Verlag, 2005.

[83] B. Fritzsche, B. Kirstein, and A. Lasarow. Orthogonal rational matrix-valued functions on the unit circle: Recurrence relations and a Favard-type theorem. *Math. Nachr.*, 279:513–542, 2006.

[84] B Fritzsche, B Kirstein, and Andreas Lasarow. Orthogonal rational matrix-valued functions on the unit circle. *Math. Nachr.*, 278(5):525–553, 2005.

[85] J. B. Garnett. *Bounded analytic functions*. Academic Press, New York, 1981.

[86] L. Garza, , F. Hernández, and F. Marcellán. Orthogonal polynomials and measures on the unit circle. The Geronimus transformations. *J. Comput. Appl. Math.*, 233(5):1220–1231, 2010.

[87] L. Garza and F. Marcellán. Szegő transformations and $N$th order associated polynomials on the unit circle. *Comput. Math. Appl.*, 57(10):1659–1671, 2009.

[88] L. Garza and F. Marcellán. Verblunski parameters and linear spectral transformations. *J. Math. Anal. Appl.*, 16:69–89, 2009.

[89] L. Garza and F. Marcellán. Quadrature rules on the unit circle. a survey. In J. Arvesú, F. Marcellán, and A. Martínez-Finkelstein, editors, *Recent trend is orthogonal polynomials and approximation theory*, volume 507 of *Contemporary Mathematics*, pages 113–139. AMS, 2010.

[90] L.E. Garza Gaone. *Transformatciones espectrales, funciones de Carathéodory y polynomios ortogonales en la circunferencia unidad*. PhD thesis, Univ. Carlos III de Madrid, February 2009.

[91] W. Gautschi. Gauss-type quadrature rules for rational functions. In H. Brass and G. Hämmerlin, editors, *Numerical integration IV*, volume 112 of *Internat. Ser. of Numer. Math.*, pages 111–130, Basel, 1993. Birkhäuser Verlag.

[92] W. Gautschi. On the computation of generalized Fermi-Dirac and Bose-Einstein integrals. *Computer Physics Comm.*, 74:233–238, 1993.

[93] W. Gautschi. Algorithm 793: GQRAT - Gauss quadrature for rational functions. *ACM Trans. Math. Software*, 25:213–139, 1999.

[94] W. Gautschi. The interplay between classical analysis and (numerical) linear algebra: a tribute to Gene H. Golub. *Electron. Trans. Numer. Anal.*, i13:119–147, 2002.

[95] W. Gautschi, L. Gori, and M.L. Lo Cascio. Quadrature rules for rational functions. *Numer. Math.*, 86(4):617–633, 2000.

[96] Ya. Geronimus. *Polynomials orthogonal on a circle and their applications*, volume 3 of *Transl. Math. Monographs*, pages 1–78. Amer. Math. Soc., 1954.

[97] Ya. Geronimus. *Orthogonal polynomials*. Consultants Bureau, New York, 1961.

[98] E. Godoy and F. Marcellan. Orthogonal polynomials on the unit circle: Distribution of zeros. *J. Comput. Appl. Math.*, 37(1-3):195–208, 1991.

[99] E. Godoy and F. Marcellan. Orthogonal polynomials and rational modification of measures. *Canad. J. Math.*, 45:930–943, 1993.

[100] G.H. Golub and J. Kautsky. Calculation of Gauss quadratures with multiple free and fixed knots. *Numer. Math.*, 41:147–163, 1983.

[101] U. Grenander and G. Szegő. *Toeplitz forms and their applications*. University of California Press, Berkley, 1958.

[102] P.S.C. Heuberger, P.M.J. Van den Hof, and B. Wahlberg, editors. *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer Verlag, 2005.

[103] K. Hoffman. *Banach spaces of analytic functions*. Prentice-Hall, Englewood Cliffs, 1962.

[104] P. Hueberger, P. Van den Hof, and O.H. Bosgra. A generalized orthonormal basis for linear dynamical systems. *IEEE Trans. Automat. Control*, 40(3):451–465, 1995.

[105] M. Humet and M. Van Barel. Algorithms for the Geronimus transformation for orthogonal polynomials on the unit circle. TW Reports TW608, Department of Computer Science, KU Leuven, January 2012.

[106] C. Jagels and L. Reichel. Szegő-Lobatto quadrature rules. *J. Comput. Appl. Math.*, 200(1):116–126, 2007.

[107] W.B. Jones, O. Njåstad, and W.J. Thron. Continued fractions and strong Hamburger moment problems. *Proc. London Math. Soc., Ser III*, 47(3):363–384, 1983.

[108] W.B. Jones, O. Njåstad, and W.J. Thron. Moment theory, orthogonal polynomials, quadrature and continued fractions associated with the unit circle. *Bull. London Math. Soc.*, 21:113–152, 1989.

[109] W.B. Jones and W.J. Thron. Orthogonal Laurent polynomials and Gaussian quadrature. In K. Gustafson and W.P. Reinhardt, editors, *Quantum mechanics in mathematics, chemistry and physics*, pages 449–455, New York, 1984. Plenum.

[110] W.B. Jones, W.J. Thron, and H. Waadeland. A strong Stieltjes moment problem. *Trans. Amer. Math. Soc.*, 206:503–528, 1980.

[111] P. Koosis. *Introduction to $H_p$ spaces*, volume 115 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2nd edition, 1998.

[112] H. Langer and A. Lasarow. Solution of a multiple Nevanlinna-Pick problem via orthogonal rational functions. *J. Math. Anal. Appl.*, 293(2):605–632, 2004.

[113] A. Lasarow. Dual pairs of orthogonal systems of rational matrix-valued functions on the unit circle. *Analysis*, 26:209–244, 2006.

[114] A. Lasarow. Solution of a multiple Nevanlinna-Pick problem for Carathéodory functions via orthogonal rational functions in the matrix case. *Math. Nachr.*, 281(1):75–108, 2008.

[115] Y.W. Lee. Synthesis of electrical networks by means of the Fourier transforms of Laguerre functions. *J. Math. Phys.*, 11:83–113, 1933.

[116] G. López-Lagomasino and J. Illán-González. Quadrature formulas for unbounded intervals. *Rev. Cienc. Mat.*, 3:29–47, 1982. (In Spanish).

[117] G.L. Lopes [López-Lagomasino]. On the asymptotics of the ratio of orthogonal polynomials and convergence of multipoint Padé approximants. *Math. USSR-Sb.*, 56:207–219, 1987.

[118] G.L. López [López-Lagomasino]. Szegő's theorem for orthogonal polynomials with respect to varying measures. In M. Alfaro et al., editors, *Orthogonal polynomials and their applications*, volume 1329 of *Lecture Notes in Math.*, pages 255–260. Springer, 1988.

[119] G.L. López [López-Lagomasino]. Asymptotics of polynomials orthogonal with respect to varying measures. *Constr. Approx.*, 5:199–219, 1989.

[120] F. Malmquist. Sur la détermination d'une classe de fonctions analytiques par leurs valeurs dans un ensemble donné de points. In *C.R. 6ième Cong. Math. Scand. (Kopenhagen, 1925)*, pages 253–259, Copenhagen, 1926. Gjellerups.

[121] P. Nevai. Géza Freud, orthogonal polynomials and Christoffel functions. A case study. *J. Approx. Theory*, 48:3–167, 1986.

[122] B. Ninness and F. Gustafsson. A unifying construction of orthonormal bases for system identification. In *Proceedings of the IEEE Conference on Decision and Control Orlando Florida December 1994*, pages 3388–3393, 1994.

[123] B. Ninness and F. Gustafsson. A unified construction of orthogonal bases for system identification. *IEEE Trans. Automat. Control*, 42(2):515–522, 1997.

[124] B. Ninness, H. Hjalmarsson, and F. Gustafsson. The fundamental role of general orthogonal bases in system identification. *IEEE Trans. Automat. Control*, 44(7):1384–1407, 1999.

[125] B. Ninness, H. Hjalmarsson, and F. Gustafsson. Generalised Fourier and Toeplitz results for rational orthonormal bases. *SIAM J. Control Optim.*, 37:429–460, 1999.

[126] O. Njåstad. An extended Hamburger moment problem. *Proc. Edinburgh Math. Soc.*, 28:167–183, 1985.

[127] O. Njåstad. Unique solvability of an extended Hamburger moment problem. *J. Math. Anal. Appl.*, 124:502–519, 1987.

[128] O. Njåstad. Solutions of the Stieltjes moment problem. *Comm. Anal. Th. Continued Fractions*, 3:32–51, 1994.

[129] O. Njåstad. Extremal solutions of the strong Stieltjes moment problem. *J. Comput. Appl. Math.*, 65:309–318, 1995.

[130] O. Njåstad. Solutions of the strong Stieltjes moment problem. *Methods Appl. Anal.*, 2:320–347, 1995.

[131] O. Njåstad. Solutions of the strong Hamburger moment problem. *J. Math. Anal. Appl.*, 197:227–248, 1996.

[132] O. Njåstad and W.J. Thron. Unique solvability of the strong Hamburger moment problem. *J. Austral. Math. Soc. (Series A)*, 40:5–19, 1986.

[133] F. Peherstorfer. On positive quadrature formulas. In *International Series on Numerical Mathematics*, volume 112, pages 297–313. Birkhäuser Verlag, 1993.

[134] F. Peherstorfer. A special class of polynomials orthogonal on the unit circle including the associated polynomials. *Constr. Approx.*, 12:161–185, 1996.

[135] F. Peherstorfer. Positive quadrature formulas III: asympotics of weights. *Math. Comp.*, 77:2241–2259, 2008.

[136] F. Peherstorfer and R. Steinbauer. Perturbation of orthogonal polynomials on the unit circle – a survey. In M. Alfaro, A. García, C. Jagels, and F. Marcellan, editors, *Orthogonal polynomials on the unit circle: theory and applications*, pages 97–119. Universidad Carlos III, Madrid, 1994.

[137] R. Pintelon, Y. Rolain, A. Bultheel, and M. Van Barel. Numerically robust frequency domain identification of multivariable systems. In P. Sas and B. Van Hal, editors, *Proceedings of the 2002 International Conference on Modal Analysis, Noise and Vibration Engineering, Leuven 2002, September 16-18*, pages 1316–1321, 2002.

[138] R. Pintelon, Y. Rolain, A. Bultheel, and M. Van Barel. Frequency domain identification of multivariabe systems using vector orthogonal polynomials. In B. De Moor, B. Motmans, J. Willems, P. Van Dooren, and V. Blondel, editors, *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems*. K.U.Leuven, Dept. ESAT, 2004. Cdrom.

[139] W. Rudin. *Real and complex analysis*. McGraw-Hill, New York, 3rd edition, 1987.

[140] J.A. Shohat and J.D. Tamarkin. *The problem of moments*, volume 1 of *Math. Surveys*. Amer. Math. Soc., Providence, R.I., 1943.

[141] B. Simon. *Orthogonal polynomials on the unit circle*, volume 54 of *Colloquium Publications*. AMS, 2005.

[142] B. Simon. *Orthogonal polynomials on the unit circle. Part 1: Classical theory*, volume 54 of *Colloquium Publications*. AMS, 2005.

[143] H. Stahl and V. Totik. *General orthogonal polynomials*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992.

[144] Z. Szabó, J. Bokor, and F. Schipp. Identification of rational approximate models in $H_\infty$ using generalized orthogonal basis. *IEEE Trans. Automat. Control*, 44(1):153–158, 1999.

[145] G. Szegő. On bi-orthogonal systems of trigonometric polynomials. *Magyar Tud. Alcad. Kutato Int. Közl.*, 8:255–273, 1963.

[146] G. Szegő. *Orthogonal polynomials*, volume 33 of *Amer. Math. Soc. Colloq. Publ.* Amer. Math. Soc., Providence, Rhode Island, 3rd edition, 1967. First edition 1939.

[147] G. Szegő. *Orthogonal polynomials*, volume 33 of *Amer. Math. Soc. Colloq. Publ.* Amer. Math. Soc., Providence, Rhode Island, 4th edition, 1975. First edition 1939.

[148] S. Takenaka. On the orthogonal functions and a new formula of interpolation. *Japanese J. Math.*, 2:129–145, 1925.

[149] M. Van Barel and A. Bultheel. Discrete linearized least squares approximation on the unit circle. *J. Comput. Appl. Math.*, 50:545–563, 1994.

[150] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. Orthogonal rational functions and structured matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):810–829, 2005.

[151] Marc Van Barel and Matthias Humet. Orthogonal functions and inverse eigenvalue problems. 4èmes Journées Approximation, International conference on constructive complex approximation, Lille, France, 28-29 June 2012, June 2012.

[152] J. Van Deun. Eigenvalue problems to compute almost optimal points for rational interpolation with prescribed poles. *Numer. Algorithms*, 45(1):89–99, 2007.

[153] J. Van Deun. Electrostatics and ghost poles in near best fixed poles rational interpolation. *Electron. Trans. Numer. Anal.*, 26:439–452, 2007.

[154] J. Van Deun. Computing near-best fixed pole rational interpolants. *J. Comput. Appl. Math.*, 235:1077–1084, 2010.

[155] J. Van Deun and A. Bultheel. Orthogonal rational functions and quadrature on an interval. *J. Comput. Appl. Math.*, 153(1-2):487–495, 2003.

[156] J. Van Deun and A. Bultheel. The computation of orthogonal rational functions on an interval. *J. Comput. Appl. Math.*, 179(1-2):355–373, 2005.

[157] J. Van Deun, A. Bultheel, and P. González-Vera. On computing rational Gauss-Chebyshev quadrature formulas. *Math. Comp.*, 75(253):307–326, 2006.

[158] J. Van Deun, K. Deckers, A. Bultheel, and J.A.C. Weideman. Algorithm 882: Near best fixed pole rational interpolation with applications in spectral methods. *ACM Trans. Math. Software*, 35(2):article no. 14, pp. 1–21, 2008.

[159] P. Van gucht and A. Bultheel. A relation between orthogonal rational functions on the unit circle and the interval $[-1, 1]$. *Comm. Anal. Th. Continued Fractions*, 8:170–182, 2000.

[160] P. Van gucht and A. Bultheel. State space representation for arbitrary orthogonal rational functions. *Systems Control Lett.*, 49(2):91–98, 2003.

[161] P. Van gucht and A. Bultheel. Using orthogonal rational functions for system identification: numerical aspects. *IEEE Trans. Automat. Control*, 48(4):705–709, 2003.

[162] L. Velázquez. Spectral methods for orthogonal rational functions. *J. Funct. Anal.*, 254(4):954–986, 2008.

[163] H. Waadeland. A Szegő quadrature formula for the Poisson integral. In C. Brezinski and U. Kulish, editors, *Computational and Applied Mathematics I*, pages 479–486. Elsevier, 1992.

[164] B. Wahlberg. Orthogonal rational functions: a transformation analysis. *SIAM Rev.*, 45(4):689–705, 2003.

[165] B. Wahlberg and P.M. Mäkilä. On approximation of stable linear dynamical systems using Laguerre and Kautz functions. *Automatica*, 32(5):693–708, 1996.

[166] J.L. Walsh. Interpolation and functions analytic interior to the unit circle. *Trans. Amer. Math. Soc.*, 34:523–556, 1932.

[167] J.L. Walsh. *Interpolation and approximation*, volume 20 of *Amer. Math. Soc. Colloq. Publ.* Amer. Math. Soc., Providence, Rhode Island, 3rd edition, 1960. First edition 1935.

[168] D.S. Watkins. Some perspectives on the eigenvalues problem. *SIAM Rev.*, 35:430–471, 1993.