Numerical quadrature and orthogonal rational functions

A. Bultheel*

Abstract

The purpose of these notes is to give an introduction to a generalization of Gauss and Szegő quadrature when these rules are not based on orthogonal polynomials, but on orthogonal rational functions with prefixed poles. This will of course require some knowledge of the theory of orthogonal rational functions (ORF) both on the unit circle and on (part of) the real line.

There are however many aspects that are intimately related to these ideas that we shall also sketch. This includes rational interpolation of Nevanlinna-Pick and Carathéodory type, moment problems, and many other aspects of what is generally known as Schur analysis. But also aspects of linear algebra like eigenvalue and inverse eigenvalue problems for structured matrices.

We shall start by briefly recalling the classical polynomial case before giving the rational generalization. In later sections we shall indicate some known results of the polynomial situation that deserve to be explored for a rational generalization. To illustrate the relevance of these rational extensions to the classical theory, and how they are used in current research, we shall conclude with some applications.

Contents

1 Introduction **5**
- 1.1 Polynomials and quadrature on $\mathbb{R}$ . . . 6
- 1.2 Laurent polynomials and quadrature on $\mathbb{T}$ . . . 9
- 1.3 Laurent polynomials and quadrature on $\mathbb{R}$ . . . 11
- 1.4 Objectives of this survey . . . 11

2 Orthogonal polynomials **12**
- 2.1 The Hankel case: orthogonality on $\mathbb{R}$ . . . 12
- 2.2 The Toeplitz case: orthogonality on $\mathbb{T}$ . . . 13

3 Moment problems and quadrature **16**
- 3.1 The real line and its subsets . . . 16
- 3.2 The unit circle . . . 17

*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, Marc Van Barel, Joris Van Deun, Patrick Van gucht. This is dedicated to the memory of Pablo González Vera.

4 Interpolation and quadrature **19**
4.1 Interpolation and quadrature on $\mathbb{R}$ . . . 19
4.2 Strong moment problems and quadrature . . . 19
4.3 Interpolation and quadrature on $\mathbb{T}$ . . . 20
4.4 Interpolation in the classes $\mathcal{B}$ and $\mathcal{C}$ . . . 21

5 Quadrature and linear algebra **23**
5.1 Quadrature on $\mathbb{R}$ and eigenvalue problems . . . 23
5.2 Quadrature on $\mathbb{T}$ and eigenvalue problems . . . 24

6 Towards the rational case **26**

7 Orthogonal rational functions and quadrature on $\mathbb{T}$ **28**
7.1 Euclidean spaces $\mathcal{L}_n$ for $\mathbb{T}$ . . . 28
7.2 Reproducing kernels . . . 29
7.3 Recurrence relations . . . 31
7.4 Quadrature . . . 35
7.5 Interpolation . . . 37
7.6 Interpolation and quadrature using the kernels . . . 39
7.7 Interpolatory product rules . . . 39

8 Orthogonal rational functions on the real line **41**
8.1 Euclidean spaces $\mathcal{L}_n$ for $\hat{\mathbb{R}}$ . . . 41
8.2 Recurrence relation . . . 43
8.3 Quadrature . . . 46
8.4 Quasi-orthogonal functions . . . 47
8.5 Interpolation . . . 48
8.6 Quadrature with kernels . . . 50
8.7 Pseudo-orthogonal functions . . . 50
8.8 The half-line and the interval . . . 50
8.9 More general situations . . . 52

9 Computational aspects **53**
9.1 The unit circle . . . 53
9.2 The real line . . . 55

10 Explicit examples of ORFs **57**
10.1 Rational Gauss-Chebyshev quadrature . . . 57
10.2 Rational Szegő quadrature with Chebyshev weights . . . 59
10.3 Rational modifications of a measure on $\mathbb{T}$ . . . 60
10.4 Near best rational approximation . . . 62
10.5 More connections between the circle and the real line . . . 63

11 Quadrature with preassigned nodes **64**
11.1 Rational Szegő QF with preassigned nodes . . . 64
11.1.1 Rational Szegő-Radau QF . . . 64
11.1.2 Rational Szegő-Lobatto QF . . . 65
11.1.3 Rational Szegő QF with more than two preassigned nodes . . . 67

11.2 Rational Gauss QF on $[-1,1]$ with preassigned nodes . . . 67
11.2.1 Rational Gauss-Radau QF . . . 67
11.2.2 Rational Gauss-Lobatto QF . . . 69
11.2.3 Relating QF on $\mathbb{T}$ and I with prescribed nodes . . . 70
11.3 Rational Szegő QF: maximal domain of exactness revisited . . . 71

12 Inverse eigenvalue problems **73**
12.1 The polynomial case . . . 73
12.2 The rational case . . . 74

13 Error bound **76**
13.1 An estimate on $\mathbb{R}$ . . . 76
13.2 An estimate on $\mathbb{T}$ . . . 77
13.3 A general estimate on $\mathbb{T}$ and $[-1,1]$ . . . 77
13.4 Anti-Gauss and anti-Szegő formulas . . . 78
13.5 Kronrod QF . . . 79

14 Denseness and convergence **81**
14.1 Results for $\mathbb{T}$ . . . 81
14.1.1 Different conditions . . . 81
14.1.2 Denseness for $\mathbb{T}$ . . . 81
14.1.3 Norm convergence . . . 82
14.1.4 Orthogonal polynomials w.r.t. varying measures . . . 83
14.1.5 Asymptotics under Szegő's condition . . . 84
14.1.6 Beyond Szegő's condition: ratio asymptotics . . . 85
14.1.7 The nth root asymptotics . . . 86
14.1.8 Convergence of interpolation and QF . . . 87
14.2 Results for $\mathbb{R}$ . . . 88
14.2.1 Denseness for $\mathbb{R}$ . . . 88
14.2.2 Asymptotics for the real line . . . 89
14.3 Convergence of modified approximants and moment problem . . . 91

15 Matrix valued versions **95**
15.1 Matrix valued ORFs . . . 95
15.2 Nevanlinna-Pick interpolation . . . 96
15.3 J-inner functions and Nehari interpolation . . . 97
15.4 Refined versions . . . 98

16 Some applications **99**
16.1 Rational Krylov methods . . . 99
16.2 Model reduction . . . 103
16.3 Linear prediction and system identification . . . 106

17 What we did not discuss... **110**

Appendix **112**

A Hardy, Lebesgue and other spaces **112**
A.1 The unit circle 112
A.2 The real line 114

1 Introduction

This survey will be about the role of rational functions in numerical quadrature. To embed this into common knowledge, we shall start by recalling some of the well known results about the role of polynomials in numerical quadrature. Polynomials are a special case of rational functions if all the poles are moved to infinity. On the other hand, to see in which direction some further development is possible, it is instructive to also highlight the area under study as being at the intersection of several diverse research areas. So it is also a breeding ground for cross fertilization.

In a first part (Sections 2-5) we shall give a brief account of well known results in numerical quadrature based on polynomials, in particular Gauss quadrature and Szegő quadrature. Thereafter we shall give the generalization and further details when the polynomials are generalized to rational functions. The polynomial introduction will also give the link with orthogonal polynomials, moment problems, Padé approximation, and linear algebra. In the rational case there will be finite poles, rational moment problems, and multipoint Padé or rational interpolation, and new eigenvalue problems for certain structured matrices.

We shall cover most of the theory in the case of the unit circle as well as for the real line. There are striking parallels yet also particular differences so that it is worthwhile to cover both of them.

Taking all the previous objectives into account, makes this survey rather ambitious and covering all aspects in detail is definitely impossible. The literature alone would need thousands of references. Therefore I will be obliged to be selective but we hope to give hints of what to read or where to start for many interesting trails that will not be explored in this text. For example we shall in a final section briefly mention some topics that we did not cover, but deserve further attention in the future to be explored and extended to the rational case. Also the list of references, although long already, will be far from exhaustive. If some topic is of interest to the reader, we invite him or her to consult the references of the references to explore that topic further. This author is aware of many omissions, but that is never intensional and reflects his personal interest and his limited knowledge.

An appendix A summarizes a very short introduction about what you should know about Hardy spaces in the disk and the upper half plane.

Some notation that will be used throughout this survey. By $\mathbb{R}$, $\mathbb{C}$ we denote the set of reals and complex numbers. The extended sets including ∞ are denoted with a hat: $\hat{\mathbb{R}}$ and $\hat{\mathbb{C}}$. The complex conjugate is denoted with a bar, the imaginary unit as i. The integers are denoted as $\mathbb{Z}$ and the set $\{0, 1, 2, \ldots\}$ is denoted as $\mathbb{N}$. The superscripts ^{T} and ^{H} denote transpose and complex conjugate transpose. The open unit disk, the unit circle and its exterior are denoted by $\mathbb{D}$, $\mathbb{T}$ and $\mathbb{E}$. The open upper half plane and the open lower half plane are denoted as $\mathbb{U}$ and $\mathbb{L}$. Hence the right half plane is denoted as $\mathrm{i}\mathbb{L}$. The inner product is denoted as $\langle f, g\rangle = \int \overline{f(z)} g(z) \mathrm{d}\mu(z)$ (note the complex conjugate is on the left factor). It should be clear from the context what μ is, otherwise, it will be indicated by a subscript. Furthermore the set of polynomials of degree at most n is denoted as $\mathcal{P}_n$ and $\mathcal{P}_{p,q} = \{\sum_{k=p}^{q} c_k z^k : c_k \in \mathbb{C}\}$ is a subset of the set of Laurent polynomials. The Kronecker delta $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. The $n \times n$ unit matrix with ones on the main diagonal is denoted as I_n and the $n \times n$ reversal matrix $\tilde{I}_n$ has ones along the main anti-diagonal, i.e., it has entries $\delta_{k,n+1-k}$, $k = 1, \ldots, n$. $\delta_a(z)$ is the Dirac delta function at a. The integral operator $\mathfrak{I}_\mu[f] = \int f \mathrm{d}\mu$ while an n-point quadrature formula is denoted as $\mathfrak{I}_n[f] = \int f \mathrm{d}\mu_n$ where $\mu_n = \sum_{k=1}^{n} w_k \delta_{x_k}$ where w_k are the weights and x_k are the nodes. The set $H(K)$

denotes the set of holomorphic functions and $C(K)$ is the set of continuous functions on $K \subset \mathbb{C}$. By x^{-1} we mean $1/x$ if x is a scalar, it is the inverse of x if x is a matrix.

Particular transforms are denoted by a fraktur font. The inverse of a transformation $\mathfrak{t}$ is denoted as $\mathfrak{t}^{inv}$. Some transforms that we shall use are the *Cayley transform* $w = \mathfrak{c}[z] = \mathrm{i}\frac{1-z}{1+z} \Leftrightarrow z = \mathfrak{c}^{inv}[w] = \frac{\mathrm{i}-w}{\mathrm{i}+w}$, a one-to-one map of $z \in \mathbb{U} \leftrightarrow w \in \mathbb{D}$ and $z \in \hat{\mathbb{R}} \leftrightarrow w \in \mathbb{T}$. The *Joukowski transform* extends the connection between $x = \cos\theta \in [-1,1]$ and $t = \mathrm{e}^{\mathrm{i}\theta} \in \mathbb{T}$ outside $[-1,1]$ as follows $w = \mathfrak{j}[z] = \frac{1}{2}(z+\frac{1}{z})$, $z \in \mathbb{D}$, $w \in \hat{\mathbb{C}}\backslash[-1,1]$. The inverse is not unique $z = \mathfrak{j}^{inv}[w] = w \pm \sqrt{w^2-1}$. We always assume that the sign is chosen such that $z \in \mathbb{D}$ if $w \in \hat{\mathbb{C}} \setminus [-1,1]$. The *Möbius transform* is defined as $w = \mathfrak{m}_c[z] = \frac{z-c}{1-\bar{c}z} \Leftrightarrow z = \mathfrak{m}^{inv}[w] = \frac{w+c}{1+\bar{c}w}$ which is a conformal map of $\mathbb{D}$ onto itself and $\mathbb{T}$ onto $\mathbb{T}$.

The acronym QF stand for quadrature formula, OP for orthogonal polynomial, POP and QOP are used for para-orthogonal and quasi-orthogonal polynomials. Similarly ORF, PORF, and QORF stand for corresponding notions for orthogonal rational functions. OPVM means orthogonal polynomials with respect to varying measures.

1.1 Polynomials and quadrature on $\mathbb{R}$

A one-dimensional integral

$$\mathfrak{I}_\mu[f] = \int f \mathrm{d}\mu = \int_a^b f(x)\mathrm{d}\mu(x), \quad -\infty \le a < b \le +\infty$$

can in most practical situations only be evaluated via numerical means. The idea is then that we can approximate the integrand f by some f_n that is 'easy to integrate' and then find an approximation for the integral as the integral of this approximation. Let us stick to the basics and consider the simplest case. Suppose that the support of μ is (part of) the real line and that f is a smooth function so that it can be easily approximated by a polynomial. The simplest possible choice is to take f_n as the interpolating polynomial of degree $n-1$ that interpolates f in n points, say $\mathbb{X}_n = \{x_k\}_{k=1}^n$. These interpolation points are called the *nodes* of the quadrature and they are collected in a (monic) *node polynomial* $\omega_n(x) = \prod_{k=1}^n (x - x_k)$. Then the interpolating polynomial can be expanded in the Lagrange basis $\{\ell_k\}$ as

$$f_n(x) = \sum_{k=1}^n \ell_k(x) f(x_k), \text{ where } \ell_k(x) = \frac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)}.$$

The $\ell_k(x)$ satisfy $\ell_k(x_j) = \delta_{jk}$ (Kronecker delta), $1 \le j,k \le n$. As a result the approximation of the integral can be written in the form of a classical (n-point) quadrature formula:

$$\mathfrak{I}_\mu[f] \approx \mathfrak{I}_\mu[f_n] \equiv \mathfrak{I}_n[f] = \sum_{k=1}^n w_k f(x_k) \text{ where } w_k = \int \ell_k(x)\mathrm{d}\mu(x).$$

These w_k are called the *weights* or *Christoffel numbers* corresponding to the nodes x_k for the measure μ. Nodes and weights are independent of the integrand f. The formula $\mathfrak{I}_n$ is an (n-point) *interpolatory quadrature formula*, and, up to rounding errors, it will integrate exactly any polynomial of degree $n-1$.

For simplicity we shall almost everywhere in this text assume that the measure μ is a finite positive Borel measure that lives either on a (possibly infinite interval on) the real line or on the unit circle. We shall assume it is normalized such that $\int \mathrm{d}\mu = 1$. Let's consider

the case where the support is a finite interval $[a, b]$, then a simple choice for the nodes, and hence also for the weights, is to take them equidistant in that interval. Then we get the well known *Newton-Cotes formulas* and the weights are called *Cotes numbers*. There are however serious drawbacks in using these formulas for large n. These reasons are analytical as well as numerical. The integration error, being the integral of the interpolation error, is given by

$$\mathcal{E}_n[f] = \int \frac{f^{(n)}(\xi(x))}{n!} \omega_n(x) \mathrm{d}\mu(x)$$

and depending on the higher derivatives of f, this can grow exponentially with n. From a numerical point of view, we may assume that the nodes and weights are exact. Indeed, they can be precomputed independently of f with all the necessary numerical care. However, the function values $f(x_k)$ are in most practical cases subject to rounding errors or measurement or approximation errors Δ_k. Suppose there is a common upper bound for these errors $|\Delta_k| \leq \varepsilon$, then the rounding error on the quadrature is bounded by

$$\left| \sum_{k=1}^{n} w_k \Delta_k \right| \leq \sum_{k=1}^{n} |w_k| \varepsilon.$$

The *Lebesgue constant* $\|\mathfrak{I}_n\| = \sum_{k=1}^{n} |w_k|$ will therefore be an important factor in the propagation of the error. Polya in 1933 [484] proved (see for example [509, Thm. 15.2.1])

Theorem 1.1 *For a finite interval $[a, b]$ and positive measure, assume that the quadrature formula $\mathfrak{I}_n[f] = \sum_{k=1}^{n} w_k f(x_k)$ with arbitrary nodes $\{x_k\}$ and weights $\{w_k\}$ converges for any polynomial f, then a necessary and sufficient condition for the convergence for any continuous function is that the sequence of Lebesgue constants $\sum_{k=1}^{n} |w_k|$ remains bounded.*

If we may assume furthermore that the quadrature will be exact for a constant function, then $\mathfrak{I}_n[1] = \sum_{k=1}^{n} w_k$. This means that if the Cotes numbers become large in absolute value, then they can only satisfy the latter constraint if they have different signs, and that of course will be the source of large rounding errors in the evaluation of the quadrature. Thus, even if theoretically the integration error goes to zero with n, it may not hold in practice because of numerical failure. It was originally proved by Steklov in 1916 [502] and later by Fejér in 1933 [249] and by several others that the following theorem holds (see [509, Thm. 15.2.2]).

Theorem 1.2 (Steklov-Fejér) *The Newton-Cotes quadrature $\mathfrak{I}_n$ for a finite interval and positive μ having nonnegative weights w_k will converge whenever the integral exists. I.e.,*

$$\lim_{n \to \infty} \mathfrak{I}_n[f] = \int_a^b f(x) \mathrm{d}\mu(x).$$

Newton-Cotes formulas will get larger weights with changing sign as n increases, making them not very useful. Indeed, the Lebesgue constants $\|\mathfrak{I}_n\|$ grow exponentially with n. If however we can produce quadrature formulas $\mathfrak{I}_n$ that are guaranteed to have positive weights and if they are exact for constants, then $\mathfrak{I}_n[1] = \sum_{k=1}^{n} w_k = \sum_{k=1}^{n} |w_k|$, and the sequence of Lebesgue constants will automatically be bounded.

These results have triggered many authors to go on a quest looking mainly for positive quadrature formulas, i.e., quadratures with only positive weights. Tchakaloff proved in 1934 [511] that such formulas always exist

Theorem 1.3 (Tchakaloff) *For any integral with positive measure on a finite interval, and for any positive integer d, there exists a positive quadrature formula that is exact for all polynomials of degree at most* $d-1$.

Of course instead of an interpolating polynomial to approximate f, any other choice can be made. Instead of interpolating polynomials, one might consider a least squares polynomial or even any other approximating function. But suppose we stick to polynomial interpolatory quadrature formulas, then fixing the nodes also fixes the weights, since they are integrals of the Lagrange basis. So, what would be an optimal choice for the nodes? One possible criterion is to maximize the order.

Definition 1.1 (order) *We say that a quadrature formula has (algebraic) order d if* $\mathfrak{I}_\mu[f] = \mathfrak{I}_n[f]$ *for all* $f \in \mathcal{P}_d$, *i.e., all polynomials* f *of degree at most* d *but there is at least one polynomial of degree* $d+1$ *for which equality does not hold. This is equivalent with saying that the first* $d+1$ *moments are matched:* $\mathfrak{I}_\mu[x^k] = m_k = \mathfrak{I}_n[x^k]$, $k = 0, \ldots, d$ *while* $m_{d+1} \neq \mathfrak{I}_n[x^{d+1}]$.

The order is also referred to as *degree of exactness*. The indication 'algebraic' refers to the fact that the definition is for algebraic polynomials. If there is no confusion possible, we shall drop the adjective.

Clearly an n-point interpolatory quadrature formula has order at least $n-1$. The following theorem explains how one may increase the order. See for example [186].

Theorem 1.4 *Let* ω_n *be the node polynomial of an interpolatory QF* $\mathfrak{I}_n$ *then its order will be* $n+k$ *with* $0 \leq k \leq n-1$ *if and only if*

$$\int x^i \omega_n(x) \mathrm{d}\mu(x) = 0, \quad i = 0, \ldots, k.$$

Assume that $\langle f(x), g(x) \rangle = \int \overline{f(\overline{x})} g(x) \mathrm{d}\mu(x)$ defines a genuine inner product. This will for example be the case when μ is positive on $\mathrm{supp}\,(\mu) = [a,b] \subseteq \mathbb{R}$. Then the previous relations express that ω_n should be orthogonal to $\mathcal{P}_k = \mathrm{span}\{1, x, \ldots, x^k\}$. Thus the maximal obtainable order is $n+(n-1) = 2n-1$ which is obtained if the nodes are the zeros of the nth orthogonal polynomial. In that case the quadrature is called a *Gauss quadrature formula*. It is known that the zeros of orthogonal polynomials are simple and are in (a,b). Moreover, Gauss quadrature formulas have an extra advantage because the Lagrange polynomials ℓ_j for an n-point formula are in $\mathcal{P}_{n-1}$, so that $\ell_j^2 \in \mathcal{P}_{2n-2} \subset \mathcal{P}_{2n-1}$ and thus ℓ_j^2 will be integrated exactly. So

$$\mathfrak{I}_n[\ell_j^2] = \sum_{k=1}^{n} w_k \ell_j^2(x_k) = w_j = \int \ell_j^2(x) \mathrm{d}\mu(x) > 0.$$

Thus the weights of Gauss quadrature formulas are all positive.

If we go for a suboptimal order, then the nodes will be zeros of a quasi-orthogonal polynomial, which is defined as follows.

Definition 1.2 (quasi-orthogonal) *A polynomial* ω_n *satisfying*

$$\left\langle x^i, \omega_n(x) \right\rangle \begin{cases} = 0, & i = 0, \ldots, n-l-1 \\ \neq 0, & i = n-l \end{cases}$$

is called quasi-orthogonal of order l.

If $l = 0$, quasi-orthogonal is just orthogonal and if the order is not mentioned, we assume that it is 1.

This links quadrature for a positive measure supported on $[a, b]$, $-\infty \leq a < b \leq \infty$ to the vast literature of the corresponding orthogonal polynomials. In the case where the functions involved are only real functions, then $\langle f(x), g(x)\rangle = \int f(x)g(x)\mathrm{d}\mu(x)$. For a complex function $f(z)$, the involution $f_*(z) = \overline{f(\overline{z})}$, $z \in \mathbb{C}$ defines a reflection in the real line such that $f_*(x) = \overline{f(x)}$ for all $x \in \mathbb{R}$ and it defines an analytic extension from $\mathbb{R}$ to the whole complex plane.

For an excellent survey on (Gauss) quadrature see [284].

1.2 Laurent polynomials and quadrature on $\mathbb{T}$

Quadrature formulas for a measure supported on (a subset of) the real line is certainly the most important type of quadratures investigated. Next to it we find the case where the measure is positive and supported on (a subset of) the complex unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. In that case, a reflection of $z = r\mathrm{e}^{\mathrm{i}\theta} \in \mathbb{C}$ in the unit circle is given by $1/\overline{z} = (1/r)\mathrm{e}^{-\mathrm{i}\theta}$, which we may assume to hold in the extended complex plane $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. So, for the unit circle, if we want $f_*(t) = \overline{f(t)}$ to hold for $t \in \mathbb{T}$ we should have $f_*(z) = \overline{f(1/\overline{z})}$ for $z \in \mathbb{C}$. Thus if μ is positive, then $\langle f, g\rangle = \int f_*(t)g(t)\mathrm{d}\mu(t)$ defines a genuine inner product. Note however that if f is a polynomial in z, then f_* is a polynomial in $1/z$. Hence if $f \in \mathcal{P}_p$ and $g \in \mathcal{P}_q$ are polynomials, then $\langle f, g\rangle$ is defined as an integral of a Laurent polynomial $f_*g \in \mathcal{P}_{-p,q} = \{p(z) = \sum_{j=-p}^{q} c_j z^j\}$.

This setting, like Fourier analysis, is closely related to integrals of periodic functions. Indeed, if we write $t = \mathrm{e}^{\mathrm{i}\theta}$ for $t \in \mathbb{T}$, then $f(t) = f(\mathrm{e}^{\mathrm{i}\theta}) = F(\theta)$ and $F(\theta)$ will be automatically 2π-periodic, while any function $g(x)$ periodic in $[-\pi, \pi]$ can be written as $g(x) = g(\cos\theta) = g\left(\frac{1}{2}(t + \frac{1}{t})\right) = G(t)$ where $x = \cos\theta$ and $t = \mathrm{e}^{\mathrm{i}\theta}$. Note that with this relation, Laurent polynomials in $t \in \mathbb{T}$ correspond to elements from $\mathrm{span}\{\cos k\,\theta, \sin k\,\theta : k \in \mathbb{N}\}$, i.e., to *trigonometric polynomials*.

Since in principle, we only need to know the integrand in the support of the measure, it is natural to require that the nodes of the quadrature are in that support. There are exceptions to this rule, for example if we use quadrature formulas to solve differential equations, in which case we use them as some kind of predictor. In general however, for integrals over $\mathbb{T}$ we expect the nodes to be on $\mathbb{T}$ as well. This has some implication. Indeed, if the node polynomial is $\omega_n(z) = \prod_{k=1}^{n}(z - t_k)$, then $\omega_n^*(z) = z^n\omega_{n*}(z) = k_n\omega_n(z)$ with $k_n = \prod_{k=1}^{n}(-t_k) \in \mathbb{T}$. A polynomial having this property is called k_n-*invariant* [355]. Since $\langle \omega_n, \omega_n\rangle \neq 0$, hence also $\langle z^n, \omega_n(z)\rangle \neq 0$ this implies that $\langle \omega_n^*(z), 1\rangle = \overline{k}_n \langle \omega_n(z), 1\rangle \neq 0$. Thus ω_n can not be orthogonal to 1.

The idea is the same as in the case of the real line. If the integrand f is smooth enough, we may approximate it with for example a Laurent polynomial in $\mathcal{P}_{p,q}$, $q-p = n-1$ interpolating in n nodes on $\mathbb{T}$. If the nodes are $\{t_k\}_{k=1}^n$, then this can be obtained by constructing the interpolating Laurent polynomial as

$$f_n(z) = \sum_{k=1}^{n} \ell_k(z)f(t_k), \quad \ell_k \in \mathcal{P}_{p,q} \text{ satisfying } \ell_k(t_j) = \delta_{kj}, \quad 1 \leq k, j \leq n$$

giving the weights $w_k = \int \ell_k(z)\mathrm{d}\mu(z)$. Alternatively we can find an algebraic polynomial $p_n \in \mathcal{P}_n$ interpolating $z^p f(z)$ in the selected nodes and then the Laurent polynomial $f_n(z) =$

$z^p p_n(z) \in \mathcal{P}_{p,q}$ will interpolate f in these nodes. From the latter construction, it is clear that we could restrict ourselves to the case $\mathcal{P}_{0,n-1}$ because shifting to $\mathcal{P}_{p,q}$, $q-p=n-1$ is simple.

For the same reasons as on the real line, we want the weights to be positive. Moreover, we want the nodes to be distributed in an optimal way, for example so as to maximize the dimension of the space $\mathcal{P}_{p,q}$ for which the quadrature formula is exact.

Because we are working with Laurent polynomials, we have to revise our notion of order of a quadrature formula. We shall immediately give some general form of the definition.

Definition 1.3 (order) *We say that a quadrature formula for an integral over $\mathbb{T}$ and $\mu > 0$ has (trigonometric) order d if $\mathfrak{I}_\mu[f] = \mathfrak{I}_n[f]$ for all Laurent polynomials $f \in \mathcal{P}_{p,q}$ with $\dim \mathcal{P}_{p,q} = q - p = d$ but not in $\mathcal{P}_{p-1,q}$ nor $\mathcal{P}_{p,q+1}$. This is equivalent with saying that a number of trigonometric moments are matched: $\mathfrak{I}_\mu[z^k] = m_k = \mathfrak{I}_n[z^k]$, $k = p, \ldots, q$ with $q - p = d$ but $m_{p-1} \neq \mathfrak{I}_n[x^{p-1}]$ and/or $m_{q+1} \neq \mathfrak{I}_n[x^{q+1}]$.*

Note that if $\mu > 0$, then $m_{-k} = \mathfrak{I}_\mu(z^{-k}) = \int \overline{t^k} \mathrm{d}\mu(t) = \overline{\int t^k \mathrm{d}\mu(t)} = \overline{m}_k$. However if the quadrature formula matches the moments m_k, $k = 0, \ldots, d$, this does not mean that then automatically also m_{-k}, $k = 0, \ldots, d$ will be matched. Anyway, an n-point interpolatory quadrature formula will have order $n-1$. If ω_n is the node polynomial, the $f_k(z) = z^k \omega_n(z) \in \mathcal{P}_{k,k+n}$ Because $\mathfrak{I}_n[f_k] = 0$, it will be integrated exactly if $\mathfrak{I}_\mu[f_k] = \int t^k \omega_n(t) \mathrm{d}\mu(t) = \langle z^{-k}, \omega_n(z) \rangle = 0$. Since ω_n is k_n-invariant, we know that $\langle 1, \omega_n \rangle \neq 0$. Suppose for simplicity that the interpolatory quadrature formula $\mathfrak{I}_n$ is based on a polynomial interpolation in $\mathcal{P}_{n-1}$. Then the following theorem holds.

Theorem 1.5 *The interpolatory quadrature formula $\mathfrak{I}_n$ with node polynomial ω_n for an integral over $\mathbb{T}$ with $\mu > 0$ will be exact in $\mathcal{P}_{-k,n-1}$ if and only if ω_n is orthogonal to $\mathcal{P}_{1,k} = \mathrm{span}\{z^1, \ldots, z^k\}$ and $1 \leq k \leq n-1$.*

Clearly the maximal order that can be obtained is $d = 2n-1$ when $k = n-1$. In that case the quadrature is called *Szegő quadrature formula*. This terminology was introduced in [355]. Just like the weights for Gaussian quadratures are guaranteed to be positive, also here we can show in the same way that the weights for Szegő quadratures are positive. Let us take for simplicity the symmetric case where the quadrature is exact in $\mathcal{P}_{n-1,n-1}$. If we choose $\ell_j \in \mathcal{P}_{n-1}$ to be Lagrange polynomials, then $|\ell_j(t)|^2 = \ell_{j*}(t)\ell_j(t) \in \mathcal{P}_{n-1,n-1}$ if $t \in \mathbb{T}$. Hence

$$\sum_{k=1}^{n} w_k |\ell_j(t_k)|^2 = w_j = \mathfrak{I}_\mu[|\ell_j|^2] > 0.$$

The notion of quasi-orthogonality has a unit circle analog called *para-orthogonality*.

Definition 1.4 (para-orthogonality) *A polynomial ω_n satisfying*

$$\langle z^i, \omega_n \rangle \begin{cases} = 0, & i = 1, \ldots, n-l-1 \\ \neq 0, & i = 0, n-l \end{cases}$$

is called para-orthogonal of order l. If $l = 1$ we just call it para-orthogonal.

The maximal order of the quadrature formula is reached when the nodes are the zeros of a para-orthogonal polynomial. Note that it is not an orthogonal polynomial. It is known that the zeros of orthogonal polynomials are all inside the open unit disk [509], and not on the boundary as is required here. There is however a relation as the following theorem states [355, Thm. 6.1].

Theorem 1.6 *Let $\{\phi_n\}$ be a sequence of orthogonal polynomials on $\mathbb{T}$ for $\mu > 0$. Define $\phi_n^*(z) = z^n \phi_{n*}(z)$ then $\{q_n(z)\}$ will be a sequence of para-orthogonal polynomials if and only if there is a sequence of nonzero numbers c_n and a sequence of numbers $\kappa_n \in \mathbb{T}$ such that $q_n = c_n(\phi_n + \overline{\kappa}_n \phi_n^*)$.*

So again there is a link between quadrature formulas on the unit circle and orthogonal polynomials on the unit circle.

1.3 Laurent polynomials and quadrature on $\mathbb{R}$

In the case of the circle, we were 'forced' to use Laurent polynomials to be able to evaluate the inner product of algebraic polynomials as an integral of a Laurent polynomial. That was not needed for the case of the real line. However, it is perfectly possible, also in the case of the real line to approximate the integrand by an n-point interpolating Laurent polynomial from $\mathcal{P}_{p,q}$, $q-p = n-1$ and integrate it as an approximation (i.e., a QF) for the original integral. As we shall see later, this corresponds to solving a so called strong moment problem instead of a classical one. It still makes sense to maximize the order of such an n-point interpolatory QF by making it exact in a space of the form $\mathcal{P}_{p',q'} \supset \mathcal{P}_{p,q}$ with $q' - p' \geq n-1$ as large as possible.

1.4 Objectives of this survey

As a conclusion of the previous sections, we can say that the central theme in classical quadrature research is to find quadrature formulas that have nodes in the support of the measure, with positive weights and with the largest possible degree of exactness. As we shall see in subsequent sections this is related to the research on orthogonal polynomials, rational approximation and interpolation, moment problems, continued fractions, and fast algorithms for particular linear algebra problems. Our main objective is to illustrate how all of these areas can be extended to the case where (Laurent) polynomials are replaced by rational functions with prescribed poles. In this theory, the polynomials show up as a special case when all the poles of the rational functions are at infinity.

In the next four sections we briefly recall some of the links that have just been mentioned for the classical polynomial case. In the sections that then follow we show how this can be generalized to the rational case. It is impossible to cover all aspects extensively which is why we restrict ourselves to some essential exemplary results and refer to the literature for further details.

The step from polynomials to rational functions is just a small one in the amalgam of possible generalizations that were developed in parallel domains of operator theory, functional analysis and linear algebra. To illustrate this we add a short section on matrix forms of the scalar theory.

The basic scalar quadrature formulas are at the heart of the applications, but may be somewhat hidden below the vocabulary and peripheral complications that are specific for the domains in which they are applied. To show that the rational variant of classical quadrature formulas is still an active research field, we add a section on applications in systems theory especially for large scale problems (and this includes analysis of graphs, differential equations, signal processing, scattering theory, etc.).

2 Orthogonal polynomials

We have seen that orthogonal polynomials play an essential role in the construction of Gauss and Szegő quadratures. Let us recall some essentials to be able to place the sections that generalize to the rational case in proper perspective. All the material in this section is classical and can be found in many textbooks. Classics are for example [160, 262, 292, 297, 298, 429, 500, 509] and the more advanced [501]. More on orthogonal polynomials, but also on moment problems is found in [2, 316, 355, 497]. An excellent introduction for quadrature and orthogonal polynomials in the case of the real line is [291]. In fact Gautschi is one of the main contemporary contributors to the theory and computational aspects of orthogonal polynomials and numerical quadrature. A collection of his work with commentaries is about to appear [72].

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 nth orthogonal polynomial in an orthogonal polynomial sequence if $\langle \phi_n, \mathcal{P}_{n-1} \rangle = 0$. In principle we can start from any polynomial basis and apply the Gram-Schmidt algorithm to arrive at a sequence of OPs.

Assuming that our positive measure μ has an infinite support, this implies that all the Gram matrices

$$G_n = \left[\left\langle z^i, z^j \right\rangle\right]_{i,j=0}^{n}$$

are positive definite for all n, hence $\det(G_n) > 0$ for all n. In the case of the real line $\langle x^i, x^j \rangle = \int x^{i+j} \mathrm{d}\mu(x) = m_{i+j}$ which implies that the G_n have a Hankel structure. In the case of the unit circle, $\langle z^i, z^j \rangle = \int z^{j-i} \mathrm{d}\mu(z) = m_{j-i}$ and this shows that the Gram matrix has a Toeplitz structure.

Explicit expressions exist for the orthogonal polynomials:

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

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 OPs with positive leading coefficient equal to $\sqrt{\det G_{n-1} / \det G_n}$.

2.1 The Hankel case: orthogonality on $\mathbb{R}$

Because of the Hankel structure of the moment matrix, there exist short recurrences.

Indeed, $\langle xp(x), q(x) \rangle = \langle p(x), xq(x) \rangle$ implies that once $\phi_n \perp \mathcal{P}_{n-1}$, then $x\phi_n(x)$ will be automatically orthogonal to $\mathcal{P}_{n-2}$. This results in the well known three-term recurrence relation for $n = 0, 1, 2, \ldots$.

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

with

$$a_n = \langle x\phi_n(x), \phi_n(x) \rangle \quad \text{and} \quad b_n = \langle \phi_{n+1}(x), x\phi_n(x) \rangle = \sqrt{\det G_{n-1} \det G_{n+1}} / \det G_n > 0.$$

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

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

The recurrence relation is second order and hence it has two independent solutions. For example starting the three-term recurrence $b_{n-1}y_{n-1} + a_n y_n + b_n y_{n+1} = xy_n$ 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 ψ_k.[1] 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* (CD):

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

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

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

Brezinski [71] shows that the CD relation is equivalent with the three-term recurrence relation.

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 $\langle k_n(\cdot, z), p(\cdot)\rangle = p(z)$.

Also for these kernels we can find a determinant expression:

$$k_n(z, w) = \frac{-1}{\det G_n} \det \begin{bmatrix} G_n & \mathbf{w}^H \\ \mathbf{z} & 0 \end{bmatrix} = \mathbf{w}^H G_n^{-1}\mathbf{z}, \quad \begin{array}{rcl} \mathbf{z} & = & [1, z, \ldots, z^n], \\ \mathbf{w} & = & [1, w, \ldots, w^n]. \end{array}$$

2.2 The Toeplitz case: orthogonality on $\mathbb{T}$

It should be clear that in the Toeplitz case $\langle p, q\rangle = \langle q_*, p_*\rangle$ and also $\langle p(z), q(z)\rangle = \langle zp(z), zq(z)\rangle$, which can be combined as $\langle p, q\rangle = \langle q^*, p^*\rangle$ where $p^*(z) = z^n p_*(z)$ and $q^*(z) = z^n q_*(z)$ with $p, q \in \mathcal{P}_n$. Let us associate with any polynomial $p = \sum_{k=0}^{n} p_k z^k \in \mathcal{P}_n$ a stacking vector $\mathbf{p} = [p_0, p_1, \ldots, p_n]^T$, i.e., $p(z) = \mathbf{z}\mathbf{p}$ with $\mathbf{z}$ as above. Note that if $\mathbf{p}, \mathbf{p}_*, \mathbf{p}^*$ are the stacking vectors of the coefficients of p, p_*, p^*, then $\mathbf{p}_*$ is the complex conjugate of $\mathbf{p}$ and the entries

[1] Note that $\psi_k \in \mathcal{P}_{k-1}$ for $k > 0$.

of $\mathbf{p}^*$ are the entries of $\mathbf{p}_*$ in reverse order: $\mathbf{p}^* = \tilde{I}_n \mathbf{p}_*$, $\tilde{I}_n = [\delta_{k,n-k}]_{k=0}^n$. Because of the symmetry of the Toeplitz matrix with respect to its anti-diagonal $\tilde{I}_n G_n \tilde{I}_n = G_n^T$, and it follows also that

$$\langle p, q \rangle = \mathbf{p}^H G_n \mathbf{q} = \mathbf{q}_*^H G_n^T \mathbf{p}_* = \mathbf{q}^{*H} G_n \mathbf{p}^* = \langle q^*, p^* \rangle .$$

Therefore $\langle \phi_n, \mathcal{P}_{n-1} \rangle = 0$ implies $\langle \phi_n^*, z\mathcal{P}_{n-1} \rangle = 0$ so that one obtains the Szegő recurrence which is more easily written down for monic polynomials $\varphi_n(z) = \phi_n(z)/\kappa_n = \sqrt{\frac{\det G_n}{\det G_{n-1}}}\phi_n(z)$, $\varphi_0 = 1$, while for $n = 1, 2, \ldots$

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

with $\upsilon_n = -\frac{\det G_{n-2}}{\det G_{n-1}} \langle 1, z\varphi_{n-1}(z) \rangle = \varphi_n(0)$ which are called the *Szegő parameters* or *Schur parameters* (see section 4.4 below) or *reflection coefficients* [130, 213] or partial correlation (PARCOR) coefficients, but currently people start using the name of *Verblunsky coefficients* after [547] as promoted by B. Simon [498–500]. The norm of the monic polynomials is $\|\varphi_n\|^2 = \frac{|G|_n}{|G|_{n-1}} = m_0 \prod_{k=1}^{n-1}(1 - |a_k|^2)$. The positivity immediately implies that $|\upsilon_k| < 1$.

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.

Two independent solutions for the Szegő recurrence are obtained by starting from two independent initial vectors. So we may define

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

To get the recurrence for the orthonormal polynomials, one should replace the θ_k by $\tilde{\theta}_k = (1 - |\upsilon_k|^2)^{-1/2}\theta_k$.

Also here we call ψ_k *polynomials of the second kind* and the *determinant formula* is

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

The analog for the *Christoffel-Darboux relation* is in terms of monic OPs

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

or for the orthonormal polynomials $\phi_k = \varphi_k/\|\varphi_k\|$

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

The confluent version (i.e. $z = w = t \in \mathbb{T}$) becomes

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

As in the Hankel case, the kernels $k_n(z,w) = \sum_{k=0}^{n} \phi_k(z)\overline{\phi_k(w)}$ are *reproducing kernels* for the space $\mathcal{P}_n$. Also the orthogonal polynomials can be written as the determinant of the Gram matrix with the bottom row replaced by the powers of z as in the Hankel case.

Note that here we get some special results:

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

That such relations hold in the Toeplitz case is because the $n \times n$ leading and trailing submatrices of G_n are equal to G_{n-1}. In the Hankel case the analog is to consider the north-east and south-west $n \times n$ submatrices which correspond to a moment sequence shifted by one element.

3 Moment problems and quadrature

In a previous section we have shown a link between orthogonal polynomials and quadrature. The order of the quadrature corresponded to matching some of the moments of the measure. The converse is also of interest. Given an infinite sequence of moments, is it possible to recover the underlying measure? These are called moment problems. Now the quadrature formulas $\mathfrak{I}_n$, being approximations of the integral operator, can be constructed for an infinite sequence $n \in \mathbb{N}$ and it can be hoped that in some (weak-star or other) sense $\lim_{n\to\infty} \mathfrak{I}_n = \mathfrak{I}_\mu$. Thus quadrature formulas help to solve moment problems and analysis of the convergence of quadrature formulas is directly linked to the solutions of the corresponding moment problem.

There are several classical moment problems associated with polynomials. Given an infinite sequence of moments $\{m_k\}$, find conditions such that there exists a positive measure μ such that $\int z^k \mathrm{d}\mu(z) = m_k$ for all $k \in \mathbb{N}$. 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 several 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 text.

Another variation may be possible if one does not only consider polynomials but also *Laurent polynomials*. Then we need moments $m_k = \int z^k \mathrm{d}\mu(z)$ not only for $k \geq 0$ 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*.

Standard books on moment problems are for example [2, 383, 497]. The original papers on moment problems were closely related to continued fractions [516]. The recurrence relation for orthogonal polynomials is essentially the same as the recurrence relation for the convergents of the continued fraction.

We recall that we shall assume that $m_0 = \int \mathrm{d}\mu(z) = 1$.

3.1 The real line and its subsets

Hausdorff. [328]. If we are looking for a positive measure on [0,1], satisfying $m_k = \int x^k \mathrm{d}\mu(x)$ for given sequence $\{m_k : k \in \mathbb{N}\}$, then this is called the *Hausdorff moment problem*. Hausdorff showed that there is a σ-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 have positive integrals for all $n \in \mathbb{N}$ and $k = 0, \ldots, n$. This results in the condition that $(-1)^k \Delta^k[m_n] \geq 0$ for all $n, k = 0, 1, \ldots$ where Δ is the forward difference operator $\Delta[m_n] = m_{n+1} - m_n$. If you want a solution that not just a discrete measure with a finite number of mass points like our QFs, 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. [503]. If the positive measure should have 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 $\det(G_n)$ and $\det(G_n^{(1)})$ 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 μ 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. The *Carleman condition* $\sum_k m_k^{-1/2k} = +\infty$ is known to be sufficient for unique solvability (determinacy).

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 [363]. See also [63, 65–67, 217, 219, 443, 444, 446]. 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 $i, j = 0, \ldots, n$ and it is required that $(-1)^n \det G_n^{(-n)} > 0$ for $n = 0, 1, \ldots$ and $\det G_{n-1}^{(-n)} > 0$ for all $n = 1, 2, \ldots$.

Hamburger. [324–326]. 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 $\det G_n \geq 0$ for all n with strict inequality if one wants to have infinite support. The problem is *determinate* if the Carleman condition $\sum_k m_{2k}^{-1/2k} = +\infty$ is satisfied. Also here a *strong Hamburger moment problem* can be considered if moments m_k are prescribed for $k \in \mathbb{Z}$, i.e., not only considering polynomials but also Laurent polynomials. See [353, 361, 439, 445, 448, 452]. The strong problem has a solution if and only if $\det G_{n-1}^{(-n)} > 0$ and $\det G_n^{(-n)} > 0$ for $n = 0, 2, 4, \ldots$.

3.2 The unit circle

On the unit circle, the moments $m_k = \int_{\mathbb{T}} t^k \mathrm{d}\mu(t) = \int_{-\pi}^{\pi} \mathrm{e}^{ik\theta} \mathrm{d}\mu(t(\theta))$ are called *trigonometric moments* [2]. 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 trigonometric moment problem has a solution if and only if all the Gram matrices $G_n = [m_{i-j}]_{i,j=0}^n$ 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., [2, 497].

Note that unlike the real line case, there is not an extensive study of the moment problems as such for measures supported on a subarc or subarcs of the unit circle. There are however many papers related to approximation [494], orthogonal polynomials [295, 296], and

asymptotics [335] in such a situation. Since most of these papers are somewhat out of the scope of this survey, we do not give further references, but it might be worth looking into the quadrature aspects.

4 Interpolation and quadrature

Another vein in which quadrature can be situated is interpolation. We have seen already that interpolatory quadrature formulas are based on interpolating polynomials. But there are other links with interpolation.

4.1 Interpolation and quadrature on $\mathbb{R}$

Since the integral of polynomials in $\mathcal{P}_{2n-1}$ is completely defined by the moments $\{m_k\}_{k=0}^{2n-1}$, it follows that the quadrature formula will be exact in $\mathcal{P}_{2n-1}$ if the discrete measure of the quadrature formula i.e., $\mu_n(x) = \sum_{k=1}^n w_k \delta_{x_k}(x)$ has the same moments m_k as the integral for $0 \leq k \leq 2n-1$. Consider the Nevanlinna function defined by

$$S_\mu(z) = \int \frac{\mathrm{d}\mu(x)}{x-z}, \quad \operatorname{Im} z > 0.$$

This function is holomorphic in the upper half plane and has there a positive imaginary part with nontangential boundary limit on $\mathbb{R}$. This limiting value is the absolute continuous part of μ. With the involution $[S_\mu(z)]_* = \overline{S_\mu(\overline{z})}$ the function S_μ has an extension to the lower half plane and it holds that $\operatorname{Im} S_\mu(z)/\operatorname{Im} z \geq 0$, $z \in \mathbb{C} \setminus \mathbb{R}$. The Hamburger-Nevanlinna theorem [2, thm. 3.2.1] then says that the classical Hamburger moment problem is equivalent with finding a *Nevanlinna function* S_μ such that the following asymptotic expansion for $z \to \infty$ holds nontangentially (i.e. convergence inside the wedge $\epsilon \leq \arg z \leq \pi - \epsilon$, $\epsilon > 0$)

$$S_\mu(z) \sim -\left[\sum_{k=0}^{2n} \frac{m_k}{z^{k+1}}\right],$$

for all $n \in \mathbb{N}$. This is clear from $(x-z)^{-1} = -\sum_k x^k/z^{k+1}$. But since the Gauss QF gives

$$S_n(z) = \int \frac{\mathrm{d}\mu_n(x)}{x-z} = \sum_{k=1}^n \frac{w_k}{x_k - z} = -\frac{\psi_n(z)}{\phi_n(z)}, \quad \phi_n \in \mathcal{P}_n, \quad \psi_n \in \mathcal{P}_{n-1}$$

with $\phi_n(z) = \prod_{k=1}^n (z - x_k)$, the node polynomial, hence the monic orthogonal polynomial. The asymptotic expansion of S_n matches the moments m_k, $k = 0, \ldots, 2n-1$. Therefore the Gauss quadrature formula has a discrete measure that solves a truncated moment problem and delivers a (diagonal) Padé approximation at ∞ for the function S_μ.

When the support is $[0, \infty)$, this is often called the Stieltjes case and S_μ is called a Stieltjes transform. The above summary is essentially unchanged in the Stieltjes case.

4.2 Strong moment problems and quadrature

Instead of the classical Hamburger or Stieltjes problem, one may consider the *strong moment problem* in which case one has to match m_k for $k \in \mathbb{Z}$. See [9, 65–67, 216, 217, 219, 363, 441, 443, 445–447, 449, 512–514].

Consider (Markov-)Stieltjes functions for $\mu > 0$ on $[0, \infty)$ defined by $S_\mu(z) = \int_0^\infty \frac{\mathrm{d}\mu(x)}{z-x}$. They are analytic anywhere in $\mathbb{C} \setminus \operatorname{supp} \mu$. Moreover they map the upper half plane $\mathbb{U}$ to

$\mathbb{U} \cup \mathbb{R}$. There are two possible asymptotic expansions available for S_μ: one at 0 and one at ∞ [363]

$$S_{\mu,0}(z) \sim -\sum_{k=0}^{\infty} m_{-(k+1)} z^k, \ z \to 0 \ \text{ and } \ S_{\mu,\infty}(z) \sim \sum_{k=0}^{\infty} \frac{m_k}{z^{k+1}}, \ z \to \infty.$$

Now, orthogonal Laurent polynomials ϕ_n need be constructed that are in some sequence of nested spaces $\mathcal{P}_{p(n),q(n)}$ of dimension n, and their zeros are used to construct n-point QFs that are exact in spaces $\mathcal{P}_{p'(n),q'(n)} \supset \mathcal{P}_{p(n),q(n)}$ of dimension $2n-1$. The Stieltjes function for the discrete measure of the QF is

$$S_n(z) = \int \frac{\mathrm{d}\mu_n(x)}{x-z} = \sum_{k=1}^{n} \frac{w_k}{x_k - z} = -\frac{\psi_n(z)}{\phi_n(z)},$$

where ϕ_n is the nth orthogonal Laurent polynomial. This S_n will have expansions at 0 and ∞ that will match some of the initial terms of $S_{\mu,0}$ and of $S_{\mu,\infty}$, and the number of moments that are matched will be $2n$, which identifies S_n as a two-point Padé approximant for these expansions. See for example [89, 90, 126, 152, 172, 216, 218, 219].

Although the previous discussion of the Stieltjes case was for a measure with support on $\mathbb{R}^+$, there is no essential difference when treating the strong Hamburger moment problem, where $\operatorname{supp}\mu$ can be the whole real line. [352, 353, 356, 438, 440, 445, 448, 452].

4.3 Interpolation and quadrature on $\mathbb{T}$

Consider the Carathéodory function for the positive measure μ

$$\Omega_\mu(z) = \int D(t,z) \mathrm{d}\mu(t), \quad D(t,z) = \frac{t+z}{t-z}, \quad z \notin \mathbb{T}. \tag{4.1}$$

The real part is obtained using the Poisson kernel

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

and this is positive for all $z \in \mathbb{D}$ with boundary value on $\mathbb{T}$ the Radon-Nikodym derivative μ'. This Ω_μ belongs to the class $\mathcal{C}$ of Carathéodory functions

$$\mathcal{C} = \{f \in H(\mathbb{D}) : \operatorname{Re} f(\mathbb{D}) \geq 0\}$$

where $H(\mathbb{D})$ denotes the functions analytic in $\mathbb{D}$. If $\Omega_\mu \in H^1(\mathbb{D})$, then μ is absolutely continuous and if $\int \mathrm{d}\mu = 1$, then for $z \in \mathbb{D}$

$$D(t,z) = 1 + 2\sum_{k=1}^{\infty} z^k t^{-k} \text{ and hence } \Omega_\mu(z) = 1 + 2\sum_{k=1}^{\infty} m_{-k} z^k.$$

If $q_n = \phi_n + \tau\phi_n^*$ is a para-orthogonal polynomial, which is also the node polynomial of the Szegő quadrature and $p_n = \psi_n - \tau\psi_n^*$ the corresponding para-orthogonal polynomial of the second kind, and $\mu_n = \sum_{k=1}^n w_k \delta_{z_k}$ is the discrete measure of the Szegő quadrature, then

$$\Omega_n(z) = \int D(t,z) \mathrm{d}\mu_n(t) = \sum_{k=1}^{n} w_k D(z_k, z) = -\frac{p_n(z)}{q_n(z)}.$$

Similarly, for $z \in \mathbb{E}$, $\operatorname{Re}\Omega_\mu(z) = -[1 + \sum_{k=1}^{\infty} m_k z^{-k}] < 0$. The fact that $\mathfrak{I}_\mu[f] = \mathfrak{I}_n[f]$ for all $f \in \mathcal{P}_{-(n-1),n-1}$, is thus equivalent with the fact that in $\mathbb{D}$: $\Omega_\mu(z) - \Omega_n(z) = z^n h(z)$ with $h \in H(\mathbb{D})$ while in $\mathbb{E}$: $\Omega_\mu(z) - \Omega_n(z) = z^{-n} g(z)$ with $g \in H(\mathbb{E})$. This means that the rational function Ω_n is a two-point Padé approximant for the two series $1 + 2\sum_{k=1}^{\infty} m_{-k} z^k$, $z \in \mathbb{D}$ and $1 + 2\sum_{k=1}^{\infty} m_k z^{-k}$, $z \in \mathbb{E}$.

4.4 Interpolation in the classes $\mathcal{B}$ and $\mathcal{C}$

In the previous sections we considered functions S_μ analytic in $\mathbb{U}$ and having a positive imaginary part. Hence $\Omega_\mu(z) = -\mathrm{i}S_\mu(z)$ is still analytic with a positive real part. In the case of the circle, we considered functions $\Omega_\mu(z)$ analytic in $\mathbb{D}$ and having a positive real part. In both cases, we shall call the Ω_μ a Carathéodory function. By a Cayley transform (see Section A.2) we may map the upper half plane to the unit circle, thus, up to a factor i it maps the class of 'positive real' functions $\mathcal{C}$ to the class $\mathcal{B}$ of 'bounded analytic' functions and there is a one-to-one correspondence. Thus with every positive measure, we may not only associate a $\mathcal{C}$-function but also a $\mathcal{B}$-function and the interpolation properties for the $\mathcal{C}$-functions that we have illustrated in the previous sections, have corresponding interpolation properties for $\mathcal{B}$-functions. In the case of the circle, matching coefficients of the $\mathcal{C}$-function is the *Carathéodory coefficient problem* and corresponding interpolation of the $\mathcal{B}$-function is the *Schur coefficient problem*. The idea is that given a $\mathcal{B}$-function, one has to find a rational $\mathcal{B}$-function such that their initial Fourier coefficients match.

The Schur algorithm [2, 302, 491–493] was originally a method designed to check whether a function Γ belonged to the class $\mathcal{B}$ of the disk. In fact, if $\Omega \in \mathcal{C}$ and $\Omega(0) = 1$, then $\Gamma = [\Omega - 1]/[\Omega + 1] \in \mathcal{B}$ and $\Gamma(0) = 0$. It is a consequence of Schwartz's lemma [278, p. 1], [140, p. 191] that also $\Gamma_0(z) = \Gamma(z)/z \in \mathcal{B}$, and if we define the Möbius transform

$$\tilde{f} = \mathfrak{m}_c[f] = \frac{f - c}{1 - \overline{c}f} \quad \Leftrightarrow \quad f = \mathfrak{m}_c^{inv}[\tilde{f}] = \mathfrak{m}_{-c}[\tilde{f}] = \frac{\tilde{f} + c}{1 + \overline{c}\tilde{f}}$$

(which is a one-to-one conformal map of $\mathbb{D}$ onto itself) then all the functions

$$\Gamma_{k+1}(z) = \frac{1}{z}\mathfrak{m}_{\gamma_k}[\Gamma_k(z)], \quad \gamma_k = \Gamma_k(0), \quad k = 0, 1, \ldots$$

will be in $\mathcal{B}$. The algorithm will stop if Γ is rational because then at some point Γ_k will be constant. All the γ_k are in $\mathbb{D}$ if and only if $\Gamma \in \mathcal{B}$.

This algorithm is some kind of dual for the recurrence of the orthogonal polynomials since these γ_k are up to a sign change the same as the Verblunsky coefficients of the orthogonal polynomials for the μ defining Ω_μ. In terms of linear algebra the duality is obvious. Denote the Toeplitz moment matrix as T, the upper triangular matrix whose columns are the coefficients of the orthonormal polynomials as R, then orthogonality means that $TR = L$ with L lower triangular. The recurrence for the orthogonal polynomials computes the columns of R recursively, while the Schur algorithm does the same but for the columns of L. If the polynomials are orthonormal, then $R^H T R = I$ and thus R is an RR^H Cholesky type factorization of T^{-1} while $T = LL^H$ is a Cholesky factorization of T.

This is perhaps also the place to mention the link with J-unitary matrices. Since we started out with Γ as the ratio of two holomorphic functions, we may want to continue the recursion for numerator and denominator separately. Set

$$\Gamma_{k+1} = \frac{N_{k+1}(z)}{D_{k+1}(z)} \quad \text{and} \quad \Gamma_k(z) = \frac{N_k(z)}{D_k(z)},$$

then the previous Möbius transform can be written as

$$z^{k+1}\begin{bmatrix} D_{k+1} \\ N_{k+1}\end{bmatrix} = \tilde{t}_k(z) z^k \begin{bmatrix} D_k \\ N_k\end{bmatrix}, \quad \tilde{t}_k(z) = \frac{1}{\sqrt{1-|\gamma_k|^2}}\begin{bmatrix} 1 & -\overline{\gamma}_k \\ -\gamma_k & 1\end{bmatrix}\begin{bmatrix} z & 0 \\ 0 & 1\end{bmatrix}.$$

The factor $(1-|\gamma_k|^2)^{-1/2}$ is just to make the matrix $\tilde{t}_k$ J-unitary since $\tilde{t}_{k*}J\tilde{t}_k = J$ where $J = \operatorname{diag}(1,-1)$ and the substar for a matrix is taking the substar for every element of the transpose of the matrix. This implies that $\tilde{t}_k^{-1} = J\tilde{t}_{k*}J$. It is obvious that this has the same form as the matrix $\tilde{\theta}_k$ used in the recurrence for the polynomials if we set $\gamma_k = -\upsilon_k$. With this equality they are in fact equal since

$$\begin{bmatrix} \psi_n & \phi_n \\ \psi_n^* & \phi_n^*\end{bmatrix}\begin{bmatrix} 1 \\ \Omega_\mu\end{bmatrix} = \tilde{\Theta}_n \begin{bmatrix} 1 & 1 \\ -1 & 1\end{bmatrix}\begin{bmatrix} 1 \\ \Omega_\mu\end{bmatrix} = \tilde{\Theta}_n \begin{bmatrix} \Omega_\mu + 1 \\ \Omega_\mu - 1\end{bmatrix} = \begin{bmatrix} z^n D_n(z) \\ z^n N_n(z)\end{bmatrix}$$

does nothing but express the interpolating properties of ψ_n/ϕ_n. Thus up to a sign change, the Verblunsky coefficients and the Schur parameters are the same. In this form it also illustrates that interpolating a function $S = N/D \in \mathcal{B}$ by a rational function $S_n = N_n/D_n \in \mathcal{B}$ is equivalent with interpolating a function $\Omega \in \mathcal{C}$ by a rational function $-\psi_n/\phi_n \in \mathcal{C}$.

For the real line, something similar holds. Again, because the Cayley transform is a one-to-one relation between the upper half plane and the unit disk, the following is obvious If $S = \mathrm{i}\Omega$ has a positive imaginary part in $\operatorname{Im} z > 0$, then

$$\Gamma = \mathfrak{c}^{inv}[S] = \frac{\mathrm{i} - S}{\mathrm{i} + S} = \frac{1-\Omega}{1+\Omega} \in \mathcal{B}, \quad \text{i.e. } |\Gamma(z)| \le 1, \quad \forall z \in \mathbb{U}.$$

However, in this case it is easier to work directly with the function S_μ or Ω_μ and do a continued fraction expansion at ∞, which goes as follows. If S_k is a Nevanlinna function (i.e., $\operatorname{Im} S_k(z)/\operatorname{Im} z \ge 0$), then

$$S_k(z) \sim -\frac{\omega_{k0}}{z} - \frac{\omega_{k1}}{z^2} - \cdots = -\frac{b_k^2}{z - a_k + S_{k+1}(z)}, \quad a_k = \frac{\omega_{k1}}{\omega_{k0}}, \quad b_k^2 = \omega_{k0}$$

where ω_{k1} is real and if $\omega_{k0} \neq 0$, it is positive, and S_{n+1} is again a Nevanlinna function [2, p. 111]. Setting $S_0 = S_\mu$ and repeating the previous rewriting step for $S_1, S_2, \ldots$ it will expanded S_0 into a continued fraction and all ω_{k0} will be positive for $k = 0, 1, \ldots$ if and only if S_0 is a Nevanlinna function, unless it is rational, and then the algorithm will stop with $\omega_{k0} = 0$.

$$S_\mu(z) = \frac{-b_0^2}{z - a_1} \underset{+}{-} \frac{-b_1^2}{z - a_2} \underset{+}{-} \frac{-b_2^2}{z - a_3} \underset{+}{-} \cdots.$$

The nth convergent of this continued fraction is a rational approximant of S_μ that matches the first $2n$ coefficients of the asymptotic expansion, and they are the S_n discussed above.

5 Quadrature and linear algebra

Much of the material concerning orthogonal polynomials has an equivalent in liner algebra, matrix factorization, matrix inversion, and conversely, subspace methods like Lanczos and Krylov type methods (see e.g., [307, 471]) rely on quadrature, Gram-Schmidt orthogonalization, and orthogonal polynomials. If the degree tends to infinity, then linear operators (e.g., of Hankel or Toeplitz type) will be involved. See e.g., [306]. We also refer to the applications section 16 of this survey.

5.1 Quadrature on $\mathbb{R}$ and eigenvalue problems

The three-term recurrence relation for the orthonormal polynomials can be written as

$$x[\phi_0, \phi_1(x), \ldots, \phi_{n-1}(x)] = [\phi_0, \phi_1(x), \ldots, \phi_{n-1}(x)]J_n + \mathbf{e}_n a_n \phi_n(x)$$

where $\mathbf{e}_n = [0, \ldots, 0, 1] \in \mathbb{R}^{1,n}$ and J_n is the truncated Jacobi matrix

$$J_n = \text{tridiag} \begin{pmatrix} a_1, a_2, \ldots, a_{n-1} \\ b_1, b_2, \ldots, b_{n-1}, b_n \\ a_1, a_2, \ldots, a_{n-1} \end{pmatrix}.$$

Setting $\boldsymbol{\phi}_n = [\phi_0, \phi_1, \ldots, \phi_{n-1}]$, then for x equal to a zero x_k of ϕ_n, it follows that $\boldsymbol{\phi}_n(x_k)$ is an eigenvector for J_n corresponding to the eigenvalue x_k.

Also the weights can be recovered from this eigenvalue decomposition. Therefore recall that Gauss quadrature formulas also relate to discrete orthogonality. Since for the product of orthogonal polynomials $\phi_i \phi_j \in \mathcal{P}_{2n-1}$ if $0 \leq i, j < n$, we get

$$\langle \phi_i, \phi_j \rangle = \sum_{k=1}^{n} w_k \overline{\phi_i(x_k)} \phi_j(x_k) = \|\phi_i\| \|\phi_j\| \delta_{i,j}.$$

This implies that the matrix Q_n with entries $[\sqrt{w_k}\phi_l(x_k)/\|\phi_j\|]_{k,l=0}^{n-1}$ is unitary an therefore

$$\sum_{k=0}^{n-1} \frac{\overline{\phi_k(x_i)}\phi_k(x_j)}{\|\phi_k\|^2} = \frac{\delta_{i,j}}{\sqrt{w_i w_j}}, \quad 1 \leq i, j \leq n$$

with in particular $\sum_{k=0}^{n-1} \frac{|\phi_k(x_j)|^2}{\|\phi_k\|^2} = \frac{1}{w_j}$, $j = 1, \ldots, n$. (See e.g., [292, p. 4].) By our normalization $\int \mathrm{d}\mu = 1$, $\phi_0 = 1$ and hence $1/w_j = \sum_{k=0}^{n-1} |\phi_k(x_j)|^2 = k_{n-1}(x_j, x_j)$ if the ϕ_k are orthonormal, otherwise w_j equals the modulus squared of the first component of the eigenvector for the eigenvalue x_j that has been normalized to length 1. This procedure is called the *Golub-Welsch algorithm* [308]. Many more details can be found in almost any handbook on numerical quadrature and for example in [186, 286, 504, 562]. See also the survey in [291, 293]. Summarizing the previous result, we have proved the following theorem [305, 306].

Theorem 5.1 *For the Gauss quadrature formula $\mathfrak{I}_n[f]$ and the truncated Jacobi matrix J_n, we have*

$$\mathfrak{I}_n[f] \equiv \sum_{j=1}^{n} w_j f(x_j) = \mathbf{e}_1^T f(J_n) \mathbf{e}_1 = [f(J_n)]_{11}.$$

Proof. Let J_n have a spectral decomposition $J_n = U_n X_n U_n^T$, with $X_n = \operatorname{diag}(x_1, \ldots, x_n)$, then we know that $[\sqrt{w_1}, \ldots, \sqrt{w_n}] = \mathbf{e}_1^T U_n$, and thus

$$\sum_{j=1}^{n} w_j f(x_j) = \mathbf{e}_1^T U_n f(X_n) U_n^T \mathbf{e}_1 = \mathbf{e}_1^T f(J_n) \mathbf{e}_1.$$

□

5.2 Quadrature on $\mathbb{T}$ and eigenvalue problems

A similar analysis can be made for the the case of the unit circle. Now however

$$z[\phi_0, \phi_1(z), \ldots, \phi_{n-1}(z)] = [\phi_0, \phi_1(z), \ldots, \phi_{n-1}(z)] H_n + \mathbf{e}_n c_n \phi_n(z)$$

where $\mathbf{e}_n = [0, \ldots, 0, 1] \in \mathbb{R}^{1,n}$ and H_n is a unitary upper Hessenberg matrix which is the $n \times n$ leading submatrix of

$$H_\infty = \begin{bmatrix} -\upsilon_1 & -c_1\upsilon_2 & -c_1c_2\upsilon_3 & -c_1c_2c_3\upsilon_4 & \cdots \\ c_1 & -\overline{\upsilon}_1\upsilon_2 & -\overline{\upsilon}_1c_2\upsilon_3 & -\overline{\upsilon}_1c_2c_3\upsilon_4 & \cdots \\ 0 & c_2 & -\overline{\upsilon}_2\upsilon_3 & -\overline{\upsilon}_2c_3\upsilon_4 & \cdots \\ 0 & 0 & c_3 & -\overline{\upsilon}_3\upsilon_4 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$

The numbers υ_n and c_n appear in the Szegő recurrence for the orthonormal polynomials ϕ_n

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

with $\upsilon_n = \phi_n(0)/\phi_n^*(0)$ and $c_n = \sqrt{1 - |\upsilon_n|^2}$.

Truncation of the infinite matrices leads to

$$z\boldsymbol{\phi}_n(z) = \boldsymbol{\phi}_n(z) H_n + [0, 0, \ldots, 0, c_n\phi_n(z)].$$

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 $c_n Q_n(z, \tau) = (1 + \overline{\upsilon}_n\tau)[z\phi_{n-1}(z) + \tilde{\tau}\phi_{n-1}^*(z)]$ with $\tilde{\tau} = \tilde{\zeta}_{\upsilon_n}(\tau) = \frac{\tau + \upsilon_n}{1 + \overline{\upsilon}_n\tau} \in \mathbb{T}$. Hence, we obtain the $Q_n(z, \tau)$, up to a constant multiple, by applying the recurrence with parameters $\upsilon_1, \upsilon_2, \ldots, \upsilon_{n-1}$ and then replace υ_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^τ then $z\boldsymbol{\phi}_n(z) = \boldsymbol{\phi}_n(z) H_n^\tau + [0, 0, \ldots, 0, cQ_n(z, \tau)]$ and if t_k is a zero of $Q_n(z, \tau)$, then

$$t_k \boldsymbol{\phi}_n(t_k) = \boldsymbol{\phi}_n(t_k) H_n^\tau$$

with $\boldsymbol{\phi}_n(z) = [\phi_0, \phi_1, \ldots, \phi_{n-1}]$ and H_n^τ is an $n \times n$ matrix, that is a unitary truncation because $\tau \in \mathbb{T}$. The t_k are eigenvalues of H_n^τ and the eigenvectors are $\boldsymbol{\phi}_n(t_k)$. As in the case of the real line, it is not difficult to prove that the weights are given by $1/w_j = \sum_{k=0}^{n-1} |\phi_k(t_j)|^2 = k_{n-1}(t_j, t_j) = \|\boldsymbol{\phi}_n(t_j)\|^2$. Thus w_j is the modulus squared of the first component of the normalized eigenvector for the node t_j, which gives the same result as

in the previous section: $\mathfrak{I}_n[f] = \mathbf{e}_1^T H_n^\tau \mathbf{e}_1$. Practical algorithms can be found in modern textbooks on linear algebra. But see e.g., also [19, 313, 314, 317].

The Hessenberg matrix is rather dense, but it has a lot of structure. 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 orthogonalizes that basis instead. It is also the natural habitat where the polynomials ϕ_k and the reciprocal polynomials ϕ_{k*} coexist. Moreover, the algebraic polynomials do *not* form a basis for the Hilbert space that is involved but the Laurent polynomials do as already remarked in [558]. 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 [153–156], [500, Part 1], [19, 139, 558]).

$$C_\infty = \begin{pmatrix} -\upsilon_1 & -c_1\upsilon_2 & c_1c_2 & 0 & 0 & 0 & 0 & \cdots \\ c_1 & -\overline{\upsilon}_1\upsilon_2 & \overline{\upsilon}_1c_2 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -c_2\upsilon_3 & -\overline{\upsilon}_2\upsilon_3 & -c_3\upsilon_4 & c_3c_4 & 0 & 0 & \cdots \\ 0 & c_2c_3 & \overline{\upsilon}_2c_3 & -\overline{\upsilon}_3\upsilon_4 & \overline{\upsilon}_3c_4 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -c_4\upsilon_5 & -\overline{\upsilon}_4\upsilon_5 & -c_5\upsilon_6 & c_5c_6 & \cdots \\ 0 & 0 & 0 & c_4c_5 & \overline{\upsilon}_4c_5 & -\overline{\upsilon}_5\upsilon_6 & \overline{\upsilon}_5c_6 & \cdots \\ 0 & 0 & 0 & 0 & 0 & -c_6\upsilon_7 & -\overline{\upsilon}_6\upsilon_7 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{5.1}$$

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

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

By truncating the infinite matrix C_∞, we can again solve an eigenvalue problem to get the nodes and the weights of the Szegő quadrature formula. The para-orthogonal polynomial $Q_n(z, \tau)$ is obtained as before by replacing the nth Verblunsky coefficient υ_n by $\tilde{\tau}$ giving the matrix C_n^τ.

Theorem 5.2 *The zeros of the para-orthogonal polynomials $Q_n(z, \tau)$ are the eigenvalues of C_n^τ whose (right) eigenvectors are $\boldsymbol{\chi}_n(z_j)$, $j = 1, \ldots, n$ with $\boldsymbol{\chi}_n = [\chi_0, \ldots, \chi_{n-1}]^H$ and the corresponding weights are given by the squared modulus of the first component of the normalized eigenvector.*

Instead of the basis $\{1, z, z^2, \ldots\}$ or $\{1, z^{-1}, z, z^{-2}, z^2, \ldots\}$ one my choose an arbitrary nested ordering of the basis $\{z^k : k \in \mathbb{Z}\}$ [172]. This gives a matrix representation of the shift (multiplication with z) operator that has a so-called snake-shape form [171].

For the practical implementation, there exist many efficient algorithms based on the fact that the information is stored in a sequence of factors Θ_k. To name just a few [22, 23, 171, 315, 545, 556, 557].

6 Towards the rational case

The eventual purpose of this survey is to generalize the polynomials that appeared in the previous sections to rational functions with some prescribed sequence of (finite or infinite) poles. 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 $n+1$ ($\mathcal{L}_{-1} = \varnothing$). In the previous sections we had $\mathcal{L}_n = \mathcal{P}_n$. From now on we fix a sequence of denominators $\{\pi_n\}$ with $\pi_0 = 1$ and $\pi_n = \pi_{n-1}\varpi_n$ with $\varpi_n \in \mathcal{P}_1$ and hence $\pi_n \in \mathcal{P}_n$. The subspaces are rational functions of the form $\mathcal{L}_n = \{\frac{p_n}{\pi_n} : p_n \in \mathcal{P}_n\}$. Clearly when $\varpi_n = 1$ for all n, then $\pi_n = 1$ for all n, and hence $\mathcal{L}_n = \mathcal{P}_n$, which illustrates that the polynomials appear as a special case.

In the polynomial case, it is known that $\mathcal{P}_\infty$ is dense in the space of continuous functions. So what will happen for $n \to \infty$ is clearly related to the denseness problem. This is also connected to the solution of the moment problem and the proof of Favard's theorem, and to basically all the problems where one needs the dimension of the spaces $\mathcal{L}_n$ to go to infinity. The type of convergence one wants to prove will also fix the type of denseness one need to consider. One might consider completeness in the set of continuous functions or in the L^p or H^p spaces with respect to the Lebesgue measure or with respect to the measure μ involved etc. But for the rational case it will of course also depend on the asymptotic behavior of the denominators π_n, i.e., the distribution of the poles. Before we can give some illustrations of these conditions, we shall first introduce the notation and the finite dimensional Euclidean spaces, i.e., when n is finite.

Early traces of these rational generalization come from different sides. In connection with Blaschke products and rational approximation, there is the Takenaka-Malmquist [555] basis forming an orthogonal system for the Lebesgue measure on the unit circle. In systems theory the so called Kautz and Laguerre basis were introducing one pole or a pair of complex conjugate poles in basis functions orthogonal with respect to the Lebesgue measure. The support of the measure was on the real line (or rather on the imaginary axis in the systems theory literature) for continuous time systems and on the unit circle in the case of discrete time systems. Only at the end of the 20th century was this generalized to an arbitrary sequence of poles. See more references in Section 10.

In the theory of moment problems, Njåstad started studying so called *extended moment problems*. The idea was that in the strong moment problems there were two series that were approximated simultaneously by two-point Padé approximants. When this is extended to more than two series that are expansions in a finite number, but more than two points, then this corresponds to have only a finite number of poles each of which is repeated an infinite number of times (e.g. in a cyclic manner), then approximating these series simultaneously is called *multipoint Padé approximation* [311, 332, 438, 439, 442, 448, 449].

For the trigonometric moment problem, seen as equivalent to interpolating Schur functions, a generalization of the Schur algorithm was given by Nevanlinna and Pick [430–433, 477–479]. This is a multipoint generalization of the Schur test to check whether a given function was a Schur function by computing its Schur coefficients and checking whether they were all bounded by 1 (See Section 4.4). While Schur's algorithm evaluates the iterates always at the origin, the Nevanlinna-Pick version evaluates at a sequence of points that are selected in $\mathbb{D}$.

In system, circuit, scattering theory and signal processing, the same algorithm did show up [204, 209]. It was used for modeling the system, circuit, medium or signal, and it is of great importance that at the same time it did guarantee stability of the model. Since the approximating model in linear systems is a rational function, checking stability reduces to checking the location of the zeros of a polynomial. Hence it can be seen as a generalization of the Schur-Cohn test to check whether poles of a polynomial are inside or outside the disk $\mathbb{D}$ which is the discrete time analog of the Routh-Hurwitz test for continuous systems to check whether poles are in some half plane (usually the right half plane) [334]. See [131] for an old survey of rational approximation problems in systems theory. We shall come back to these algorithms and the applications in later Sections 15 and 16.

Further generalizations were considered recently by Fialkow and coworkers. The moment problems they consider are associated with planar measures. So they consider multivariate moment and interpolation problems and related cubature formulas. This includes the extension problem (i.e. how to extend a finite set of moments such that the extended sequence of moments still allows for a positive measure as solution). See their paper [178] and its references for an early paper, but the multivariate problem is still an active domain of research [250–253] to name just a few of their recent papers. We shall however not consider this any further in this survey.

7 Orthogonal rational functions and quadrature on $\mathbb{T}$

In our previous polynomial approach we have given the case of the real line before the circle case because Gauss quadrature formulas are probably best known to a broader public. However, as e.g. already suggested by the many different moment problems, the case of the real line has many more subcases. For this and for other reasons, the real line case is not always the simplest for the rational generalization. Therefore, in the next few sections about the rational generalization we start with the case of the unit circle. In later sections we shall start with either the circle or the real line, depending on what we think to be the simplest. Because the next sections are at the core of the upgrading step from the polynomial to the rational case, we shall include some very brief and sketchy ideas of the proofs whenever this is not too involved. They can be skipped without any harm if so desired.

7.1 Euclidean spaces $\mathcal{L}_n$ for $\mathbb{T}$

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 ∞, 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. So let's stick to that notation. Thus the rational functions have denominators that are of the form $\prod_k(1-\overline{\alpha}_k z)$. However, one might also allow such a sequence of poles anywhere 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 $\alpha_k \in \mathbb{T}$). The latter situation is only a slight complication over our assumption that the α_k are all in $\mathbb{D}$. It is for example a standard assumption in the work by A. Lasarow who has generalized our scalar development to the matrix case. This matrix generalization will be sketched briefly in Section 15. See for example [265–268, 270, 271, 377, 384, 386].

Thus for simplicity reasons, we choose a sequence of points $\mathbb{A} = \{\alpha_k : k = 0, 1, \ldots\} \subset \mathbb{D}$ with $\alpha_0 = 0$ and define the denominators as $\pi_n(z) = \prod_{k=0}^{n}(1-\overline{\alpha}_k z)$. This means that we will consider rational functions in $\mathcal{L}_n$ whose poles will be in the set $\mathbb{A}_{n*} = \{1/\overline{\alpha}_k : k = 1, \ldots, n\}$.

$$\mathcal{L}_n = \left\{ \frac{p_n}{\pi_n} : p_n \in \mathcal{P}_n, \pi_n(z) = \prod_{k=0}^{n}(1-\overline{\alpha}_k z) \right\}, \quad n = 0, 1, \ldots. \tag{7.1}$$

As before we set $\mathcal{L}_{n*} = \{f_* : f \in \mathcal{L}_n\}$ and we define $\mathcal{L}_{p,q} = \mathcal{L}_{p*} \cdot \mathcal{L}_q = \mathcal{L}_{p*} + \mathcal{L}_q$. The inner product will be defined as before by $\langle f, g\rangle = \int f_*(t)g(t)\mathrm{d}\mu(t) = \int \overline{f(1/\bar{t})}g(t)\mathrm{d}\mu(t)$ and we assume $m_0 = \int \mathrm{d}\mu(t) = 1$. One can choose different sets of basis functions for $\mathcal{L}_n$. If all the α_k are distinct, then the basis $\{(1-\overline{\alpha}_k z)^{-1} : k = 0, 1, \ldots\}$ has some advantages and may give simple formulas. For example, the Gram matrix $G_n = [\langle z^i, z^j\rangle]$ for $\mathcal{L}_n$ which was Toeplitz in the polynomial case will now be a *Pick matrix* with entries

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

where $\Omega_\mu(z)$ is the Carathéodory function associated with μ and $D(t, z)$ the Riesz-Herglotz kernel as these were defined in (4.1). However, if the point α_k appears more than once, we also need basis functions of the form $z^j/(1-\overline{\alpha}_k z)^n$ which does complicate matters. This

implies that in the Gram matrix also elements of the form $\Omega_\mu^{(j-1)}(\alpha_k)$ (i.e., derivatives of Ω_μ of order $j-1$ evaluated in α_k) will show up. It should be clear though that the inner product in $\mathcal{L}_n$ will be completely defined by such values:

Theorem 7.1 *Suppose $\mathbb{A}_n$ consists of $m \le n$ mutually different points α_i, $i = 0, \dots, m$ with α_i repeated n_i times, then the inner product in $\mathcal{L}_n$ will be completely defined by the n values $\{\Omega_\mu^{(j)}(\alpha_i) : j = 0, \dots, n_i - 1; i = 0, \dots, m\}$ and their complex conjugates.*

This immediately implies that if we find a rational function $\Omega_n \in \mathcal{L}_n$ for which $\Omega_\mu^{(j)}(\alpha_i) = \Omega_n^{(j)}(\alpha_i)$ for $j = 0, \dots, n_i - 1; i = 0, \dots, m$, then it will be the Carathéodory function of a discrete measure defining the quadrature formula that is exact in $\mathcal{L}_{(n-1)*} \cdot \mathcal{L}_{n-1} = \mathcal{L}_{n-1,n-1}$.

A Pick matrix is not so much different from a Toeplitz matrix because

Theorem 7.2 *If $\boldsymbol{w}_n = [w_0, \dots, w_n]^T$ is a basis for $\mathcal{L}_n$, with Gram matrix $G_n(\boldsymbol{w}_n) = \langle \boldsymbol{w}_n, \boldsymbol{w}_n \rangle = [\langle w_k, w_l \rangle]_{k,l=0}^n$, then there is a Toeplitz matrix T_n and an invertible matrix V such that $VG_n(\boldsymbol{w}_n)V^H = T_n$.*

Proof. Note that we may always choose the basis $\{\ell_k(z) = z^k/\pi_n(z) : k = 0, \dots, n\}$ with $\pi_n(z)$ as in (7.1). Hence the Gram matrix for this basis is Toeplitz, with entries that 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. □

However, we prefer to work with the basis of Blaschke products[2]

$$B_0 = 1, \quad B_n(z) = \prod_{k=0}^{n} \zeta_k(z), \quad \zeta_k(z) = s_k \frac{z - \alpha_k}{1 - \overline{\alpha}_k z}, \quad s_k \in \mathbb{T}, \qquad n = 0, 1, \dots. \tag{7.2}$$

Note that $\zeta_k(z) = s_k \mathfrak{m}_{\alpha_k}(z)$ where $\mathfrak{m}_\alpha[z] = (z - \alpha)(1 - \overline{\alpha}z)^{-1}$ with $\alpha \in \mathbb{D}$, is a *Möbius transformation* transforming the unit disk into itself and the circle into itself. It will show up several times more in this context.

We now are set to introduce the general ORF for a measure supported on $\mathbb{T}$. However, it will be simpler to start with the reproducing kernels.

All the results of this part (and more) can be found in the monograph [105]. Some shorter, partial summaries are given in [100, 116, 118].

7.2 Reproducing kernels

Suppose $\{e_0, \dots, e_n\}$ is a basis for $\mathcal{L}_n$, orthonormal with respect to μ. 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., $\langle k_n(\cdot, w), f(\cdot) \rangle = f(w)$ for all $f \in \mathcal{L}_n$ and $w \in \mathbb{D}$. If $f \in \mathcal{L}_m$ with $m > n$, then $\langle k_n(\cdot, w), f(\cdot) \rangle$ 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 $k = 0, \dots, n$ and (note $|B_n(t)| = 1$ for $t \in \mathbb{T}$)

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

[2] To make convergence possible for $n \to \infty$, the s_k in the definition of the Blaschke products are usually defined as $-\overline{\alpha}_k/|\alpha_k|$ if $\alpha_k \neq 0$ and defined as 1 otherwise. For computations with finite n, we may remove them, i.e., just set $s_k = 1$ for all k.

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 7.3 *If $\boldsymbol{e}_n(z) = [e_0(z),\ldots,e_n(z)]$ represents any basis for $\mathcal{L}_n$, then the reproducing kernel is given by*

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

where $G_n = \langle \boldsymbol{e}_n, \boldsymbol{e}_n\rangle$ is the Gram matrix for $\boldsymbol{e}_n = [e_0,\ldots,e_n]$.

Proof. Let V be a matrix transforming the basis $\boldsymbol{e}_n$ into the orthonormal basis $\boldsymbol{f}_n = \boldsymbol{e}_n V$. Then $I = \langle \boldsymbol{f}_n, \boldsymbol{f}_n\rangle = V^H \langle \boldsymbol{e}_n, \boldsymbol{e}_n\rangle_\mu V = V^H G_n V$, so that $G_n^{-1} = VV^H$ and thus $k_n(z,w) = \boldsymbol{f}_n(z)\boldsymbol{f}_n(w)^H = \boldsymbol{e}_n(z)VV^H\boldsymbol{e}_n(w)^H = \boldsymbol{e}_n(z)G_n^{-1}\boldsymbol{e}_n(w)^H$, which proves the theorem. □

From now on, we shall use the notation ϕ_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 κ_n. Since ϕ_n is uniquely defined up to a constant of modulus 1, we can fix ϕ_n uniquely, for example by assuming that $\kappa_n > 0$, which we shall do throughout this text, 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 space $\mathcal{L}_n$ to which f belongs, i.e., on n and on the $\{\alpha_k : k = 0,\ldots,n\}$. So a notation like $f^{[n]}$ or even $f^{[\mathbb{A};n]}$ 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: [105, Thm. 2.2.3], [84, Lemma 1].

Corollary 7.4 *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 $k_n(\alpha_n,\alpha_n) = \kappa_n^2$.*

The following Christoffel-Darboux (CD) relations hold [105, Thm. 3.1.3], [84, Thm. 3]:

Theorem 7.5 *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:

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

with $P(t,z)$ the Poisson kernel given in (A.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 7.3) 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 7.6 *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 $z \in \mathbb{D}$, $(\mathbb{T}, \mathbb{E})$.*

7.3 Recurrence relations

The kernels satisfy a fundamental recurrence relation [105, Thm. 3.2.1], [84, Thm. 4] (but see also [212, 213]).

Theorem 7.7 *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

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

where

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

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 $\rho_n(w)$ 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

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

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 ϕ_n [105, Thm. 4.4.1], [84, Thm. 7].

Theorem 7.8 *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:

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

with

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

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

$$\upsilon_n = s_{n-1}\frac{\delta_n}{\varepsilon_n} = \overline{\eta_n^1}\frac{\phi_n(\alpha_{n-1})}{\phi_n^*(\alpha_{n-1})} \in \mathbb{D},$$

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

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 $\upsilon_n \in \mathbb{D}$. 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

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

or, disregarding the normalizations, also

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

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

Note that if all the α_k are zero, then we recover the polynomial case. The recurrence relation is just the Szegő recurrence and the υ_n being then equal to $\phi_n(0)/\phi_n^*(0)$, are the Szegő or Schur or *Verblunsky coefficients*. Therefore we shall keep the name and call also in the rational case the υ_k the (generalized or rational) Verblunsky coefficients.

Since this is derived from the CD relation, the first recurrence does not depend on the fact that the α_k are in $\mathbb{D}$. It is not so difficult to generalize the whole theory to the case where α_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 υ_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 α_k are in $\mathbb{D}$ or all in $\mathbb{E}$, then the system is non-degenerate and $\upsilon_n \in \mathbb{D}$.

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.

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

Then it can be proved that $(\psi_n, -\psi_n^*)$ is another solution of the recurrence relation for (ϕ_n, ϕ_n^*), which corresponds to the initial condition $(1, -1)$. In other words (ψ_n, ψ_n^*) satisfies the same recurrence relation as (ϕ_n, ϕ_n^*), but with υ_n replaced by $-\upsilon_n$ [105, Sec. 4.2].

Theorem 7.9 *With the notation of theorem 7.8 we have*

$$\begin{bmatrix} \phi_n(z) & \psi_n(z) \\ \phi_n^*(z) & -\psi_n^*(z) \end{bmatrix} = \tilde{t}_n(z) \begin{bmatrix} \phi_{n-1}(z) & \psi_{n-1}(z) \\ \phi_{n-1}^*(z) & -\psi_{n-1}^*(z) \end{bmatrix}. \tag{7.3}$$

Proof. We only have to prove the relation for the ψ_n. From the definition and the previous observation we have for $n > 1$

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

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

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

It can be shown that the integral equals

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

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 [105, Thm. 4.3.1], [110, Sec. 4].

Theorem 7.10 *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 (ϕ_n, ϕ_n^*) or (ψ_n, ψ_n^*) gives several identities which we leave for the reader to discover. We just give one example:

$$\psi_n(z)\phi_n^*(z) + \psi_n^*(z)\phi_n(z) = 2B_n(z)P(z,\alpha_n) \tag{7.4}$$

with $P(z,w)$ the Poisson kernel (A.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 (7.3). 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 ϕ_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 [105, Thm. 8.1.4], [91]. (See also [92] for the case when the poles are on $\mathbb{T}$ rather than in $\mathbb{D}$. But this is based on a three-term recurrence relation and is therefore essentially a translation of the case of the real line.)

Theorem 7.11 (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 ϕ_n are generated by a recurrence relation of the 3rd form given in Theorem 7.8 with all $\upsilon_n \in \mathbb{D}$, and the unimodular constants such that $\phi_n^*(\alpha_n) > 0$, then these ϕ_n will form a sequence of ORFs with respect to a positive measure on $\mathbb{T}$. If the Blaschke product of the α_k diverges, the measure is unique.*

Proof. We give a sketch of the proof. It is easily proved by induction that the ϕ_n^* generated by the recursion are indeed superstar conjugates of the ϕ_n and the parameters λ_n are given by their expressions of Theorem 7.8. 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 $0 \leq m \leq n$

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

Thus $\phi_n \perp_{w_n} \mathcal{L}_{n-1}$, and because an inverse recurrence (obtained by multiplying (7.3) 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 ($\int w_n \mathrm{d}\lambda = 1$), so that there is a subsequence w_{n_k} which converges weakly. Thus there is some μ such that $\lim_{n\to\infty} \int f(t) w_{n_k}(t) \mathrm{d}\lambda(t) = \int f(t) \mathrm{d}\mu(t)$ for all functions f continuous on $\mathbb{T}$. If $\sum(1-|\alpha_k|) = \infty$, then a denseness result (see Section 14) shows that the measure μ 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 [101], but to the best of my knowledge, a satisfactory answer is still missing.

Because the functions of the second kind satisfy a recurrence relation of this form (the υ_n have to be replaced by $-\upsilon_n$), 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 ψ_n the same properties as we can attribute to the ϕ_n (location of the zeros, CD relations, etc.). If $\Omega_\mu \in \mathcal{C}$ is the Riesz-Herglotz transform of μ, then clearly $1/\Omega_\mu$ is also in $\mathcal{C}$. So it can be written as the Riesz-Herglotz transform of some measure ν. It can be shown that if the ϕ_n are orthogonal w.r.t. μ, then the associated functions of the second kind ψ_n, will be orthogonal w.r.t. the associated ν.

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.

7.4 Quadrature

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

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 $\mathcal{L}_{p*} = \{f : f_* \in \mathcal{L}_p\}$. It is not difficult to show that this f_n shall interpolate f in the points

$\{t_{ni}\}_{i=1}^n \subset \mathbb{T}$ if

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

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

Thus the quadrature formula is $\mathfrak{I}_n[f] = \sum_{k=1}^n w_{n,k} f(t_{nk})$, $w_{n,k} = \mathfrak{I}_\mu[L_{n,k}]$. 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 t_{nk} in a particular way. We cannot choose them as the zeros of the ORFs ϕ_n since these, just like orthogonal polynomials, have all their zeros inside $\mathbb{D}$, and not on $\mathbb{T}$. However *para-orthogonal rational functions* (PORFs) will have zeros on $\mathbb{T}$. First we define $\mathcal{L}_{n-1}(\alpha_n) = \{f \in \mathcal{L}_{n-1} : f(\alpha_n) = 0\}$.

Definition 7.1 *We say that $Q_n \in \mathcal{L}_n$ is a PORF if it is orthogonal to $\mathcal{L}_{n-1}(\alpha_n)$ and $\langle 1, Q_n\rangle \neq 0 \neq \langle B_n, Q_n\rangle$.*

The following result holds [112, Thm. 4].

Theorem 7.12 *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 ξ 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 $\langle f, \phi_n\rangle = 0$ 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 $\langle f, \phi_n^*\rangle = \langle B_n g_*, B_n\phi_{n*}\rangle = \langle \phi_n, g\rangle = 0$. On the other hand $\langle 1, Q_n\rangle = \tau\,\langle 1, \phi_n^*\rangle = \tau\,\langle \phi_n, B_n\rangle \neq 0$, 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 $\{t_{n1}, \ldots t_{nn}\}$ of Q_n. Then simply writing it out shows that the interpolation error can be written as

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

Thus, because of the para-orthogonality of Q_n, we get an integration error $\mathfrak{I}_\mu[E_n] = \langle S, Q_n\rangle = 0$, 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 $\mathcal{L}_{n-1,n}$ [112]. In this sense the quadrature formula is optimal and we shall refer to it as the *rational Szegő quadrature formula*. We denote the (discrete) measure having masses w_{nk} in the points t_{nk} by $\mu_n = \sum_{k=1}^n w_{nk}\delta_{t_{nk}}$ so that by definition $\mathfrak{I}_n[f] = \mathfrak{I}_{\mu_n}[f]$.

To note the relation between reproducing kernels and PORFs, one may prove that if $t \in \mathbb{T}$, then $c(1-\zeta_n(z)\overline{\zeta_n(t)})k_{n-1}(z,t)$ is a PORF where $c = -1/\overline{\phi_n(t)}$ and $\overline{\tau} = -\phi_n^*(t)/\phi_n(t) \in \mathbb{T}$.

Clearly $z = t$ is a zero of this PORF. Thus we can fix one of the zeros at $t \in \mathbb{T}$ in this representation. See also [128].

It can also be proved that like for successive orthogonal polynomials, and the para-orthogonal polynomials [303], an interlacing property holds for the zeros of PORFs [124].

Theorem 7.13 *The zeros of two consecutive PORFs Q_n and Q_{n-1} are all simple and on $\mathbb{T}$ and they interlace.*

By using the recurrence, it follows that a PORF $Q_n(z, \tau)$ can also be written in terms of ϕ_{n-1}:

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

with $\tilde{\tau} = \mathfrak{m}^{inv}_{\upsilon_n}[\tau] = \frac{\tau + \upsilon_n}{1 + \tau\overline{\upsilon}_n} \in \mathbb{T}$ (disregarding normalization).

7.5 Interpolation

We have used interpolation to construct a quadrature formula and found that integration with respect to μ and μ_n, the discrete measure of the QF, is the same in $\mathcal{L}_{n-1,n-1}$. This happens if and only if the initial terms in the expansions of their Riesz-Herglotz transforms in the points $\{\alpha_0, \ldots, \alpha_n\}$ (i.e. function values and possibly values of the derivatives, taking multiplicity of the points into account). This is because the respective Gram matrices depend only on these values. So let us first find out what the Riesz-Herglotz transform of μ_n is. We have the following theorem [98, Thm. 2.3], [109].

Theorem 7.14 *Set $\Omega_n(z) = \int D(t,z)\mathrm{d}\mu_n(t) = -\frac{P_n(z)}{Q_n(z)}$ where $Q_n \in \mathcal{L}_n$ is a PORF. Furthermore let ϕ_n be the ORF and ψ_n ORF of the second kind. Then for $n \geq 1$, there is some $P_n \in \mathcal{L}_n$ such that*

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

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

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 ψ_n are orthogonal.

The previous theorem has an interesting corollary [109].

Corollary 7.15 *The weights of the rational Szegő formula corresponding to the nodes t_{nk} are given by*

$$w_{nk} = \frac{1}{2t_{nk}}\frac{P_n(t_{nk})}{Q'_n(t_{nk})} = \frac{1}{\sum_{i=0}^{n-1}|\phi_i(t_{nk})|^2} > 0.$$

Proof. Recall $\Omega_n(z) = -P_n(z)/Q_n(z) = \mathfrak{I}_n[D(\cdot, z)] = \sum_{k=1}^n w_{nk}D(t_{nk}, z)$ and $Q_n(t_{nk}) = 0$. Thus after multiplying by $(z - t_{nk})$ we have

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

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

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

with $P(t, z)$ the Poisson kernel (A.1), while the determinant formula leads to $P_n(t_{nk}) = 2P(t_{nk}, \alpha_n)/\overline{\phi_n(t_{nk})}$, so the corollary is proved. □

We can now derive interpolation properties for $\Omega_n = -P_n/Q_n$ (see [109]):

Theorem 7.16 *Let Ω_n and Ω_μ be the Riesz-Herglotz transforms of μ_n and μ respectively, then for $z \in \mathbb{D}$: $\Omega_\mu(z) - \Omega_n(z) = zB_{n-1}(z)h(z)$, for $n \geq 1$ with h analytic in $\mathbb{D}$.*

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 α_n [125, Thm. 5.4].

Theorem 7.17 *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 $\langle \phi_n, B_{n-1}\rangle = 0$. The second relation is similarly proved since now for $z = 0$ we find $\langle \phi_n^*, B_n\rangle = 0$. □

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 α_n, one might expect that the absolute continuous measure $w_n(t)\mathrm{d}\lambda(t)$ gives the same integrals in $\mathcal{L}_{n,n}$ as $\mathrm{d}\mu(t)$, and it does indeed [125, Thm. 5.6].

Theorem 7.18 *The inner product in $\mathcal{L}_n$ is the same for the measure μ 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 α_n. Thus ϕ_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 ϕ_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.

7.6 Interpolation and quadrature using the kernels

After we formulated the recurrence relation for the kernels in Theorem 7.7, we gave the recurrence for the ORF in Theorem 7.8, 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, or one may want to check the literature [105].

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 7.7. By our remark after Theorem 7.11, 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 ψ_k are the ORF. The interpolation properties are a bit harder to obtain because they require a parameterized (in w) Riesz-Herglotz transform:

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

with $D(t,z)$ 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. $\mathrm{d}\mu(t)$.

7.7 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 ϕ_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 ϕ_n and $\tilde{\phi}_n$, 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 $2n+1$. However, it is also possible to use this approach to generate ORF and rational Szegő QF for sequences of poles that are in $\mathbb{C} \setminus \mathbb{T}$, where the order of the poles can be arbitrary. By this I mean that there need not be an alternation where the poles with even index are in $\mathbb{D}$ and the poles with odd index in $\mathbb{E}$ or vice versa.

For the details we refer to [106–108, 123].

8 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 what has been said before for the unit circle case. We do want to avoid repetition and skip this case. 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, again a situation that we do not discuss in detail here. Note that this relation between the circle and the line is somewhat more subtle than just applying the Cayley transform to the variables. If one does that, then polynomials on $\mathbb{T}$ are mapped onto rational functions that are essentially polynomials in $\frac{1-z}{1+z}$. We shall have a look at measures supported on the whole real line, on a half-line or on an interval, mostly choosing the poles outside that support, except when the support is the whole real line. 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 on $\mathbb{R}$ will not be discussed here.

8.1 Euclidean spaces $\mathcal{L}_n$ for $\hat{\mathbb{R}}$

For polynomials, 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 ∞, the measure may well be positive at infinity. If the integral is on a half line, then ∞ is a boundary point, and if the integral is over a finite interval such as e.g. $[0, 1]$, then of course ∞ is again far away from the support as in the case of the circle. So there are many possible situations to consider, which also correspond to many generalizations of the different classical moment problems. One may wonder what it means to take poles "in the neighborhood of ∞" in this case. Should they be on $\mathbb{R}$ or can they be complex? If the support of μ is all of $\hat{\mathbb{R}}$, then one can still 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 or complex) 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". They are real, yet outside the support of the measure. This is related to 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 in a neighborhood of ∞, 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 like Carathéodory functions, Nevanlinna, Riesz, Markov or Cauchy transforms or bounded analytic functions for a half plane. They are all in some sense equivalent instances of the same problem.

It should also be stressed that when poles are located in the support of μ, e.g., when that is $\hat{\mathbb{R}}$ and the poles are real, the integral of $f \in \mathcal{L}$ should exist and, if we consider an inner product, then we need to assume that also integrands of the form $f_{\bar{*}}g \in \mathcal{L}_{\bar{*}} \cdot \mathcal{L}$, if $f, g \in \mathcal{L}$, should be integrable. If each one of the poles is real and repeated an infinite number of times, then $\mathcal{L}_{n\bar{*}} \cdot \mathcal{L}_n \subset \mathcal{L}_{\bar{*}} \cdot \mathcal{L} = \mathcal{L}$ but this is not true in general. This makes rational moment problems more complicated than the polynomial case. Let $\{b_k : k = 0, 1, \ldots\}$ be a basis for $\mathcal{L}_\infty$, then one may consider moment problems where rational moments are given of the form $m_k = \int b_k(x) \mathrm{d}\mu(x)$, $k = 0, 1, 2, \ldots$, or one may consider a strong version where we know m_k, $k \in \mathbb{Z}$ where $b_{-k}(z)$ for $k > 0$ is defined as $b_{k\bar{*}}(z) = \overline{b_k(\overline{z})}$. Thus we have a moment problem in $\mathcal{L}$ and a strong version requiring moments in $\mathcal{L}_{\bar{*}} \cdot \mathcal{L}$.

Let us be more specific about one of all these possible settings.

We select a sequence $\mathbb{A} = \{\alpha_k \in \mathbb{R} \setminus \{0\} : k = 1, 2, \ldots\}$, set $\alpha_0 = \infty$ and define a basis[3]

$$b_0 = 1, \quad b_n(z) = b_{n-1}(z) Z_n(z) = \prod_{k=0}^{n} Z_k(z), \quad Z_k(z) = \frac{z}{1 - z/\alpha_k}, \quad n = 1, 2, \ldots$$

Recall the notation $f_{\bar{*}}(z) = \overline{f(\overline{z})}$. Note that by choosing all α_k on $\mathbb{R}$ we get $Z_{k\bar{*}}(z) = Z_k(z)$, hence also $b_{n\bar{*}}(z) = b_n(z)$ for any $z \in \mathbb{C}$. This will simplify things, but note that it will not be true anymore if the α_k can be complex.

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

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

If we allow that μ can have mass at infinity, then to define the analog of Ω_μ, we have to replace the Riesz-Herglotz kernel by the *Nevanlinna kernel*. See Appendix A.2 on the definitions introduced here. It is defined as

$$D(x, z) = \frac{1}{\mathrm{i}} \frac{1 + xz}{x - z}.$$

and the *Poisson kernel* is $P(x, z) = \frac{\operatorname{Im} z}{|x - z|^2}$. We define $\mathrm{d}\mathring{\mu}(x) = \frac{\mathrm{d}\mu(x)}{1+x^2}$.

The class of Carathéodory functions for the upper half plane $\mathbb{U}$ is defined as $\mathcal{C}(\mathbb{U}) = \{f \in H(\mathbb{U}) : \operatorname{Re} f(\mathbb{U}) \geq 0\}$. All $\mathcal{C}(\mathbb{U})$ functions have the following Nevanlinna representation

$$\Omega_{\mathring{\mu}}(z) = \mathrm{i}c + \int D(x, z) \mathrm{d}\mathring{\mu}(x), \quad c \in \mathbb{R}, \ x \in \mathbb{R}, \quad z \in \mathbb{U}.$$

The real part is given by

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

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

By Fatou's theorem,

$$\lim_{y \downarrow 0} \operatorname{Re} \ \Omega_{\mathring{\mu}}(x + \mathrm{i}y) = \mu'(x), \quad \forall x \in \operatorname{supp} \mu.$$

[3]We could have chosen as in the case of the circle to denote the poles as $1/\alpha_k$ instead of α_k as we do here. However, that would involve quite often evaluations of functions at $1/\alpha_k$. So for purely notational convenience, we switch the notation. A polynomial $p \in \mathcal{P}_n$ having a zero of order k at infinity means that its degree is $n - k$.

8.2 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 nth 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 (since we chose real poles) and hence that $\phi_{n^*} = \phi_n$. If complex poles are used, then this complicates the formulas considerably.

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

We can now give the theorem generalizing the three-term recurrence relation for polynomials [105, Thm. 11.1.2].

Theorem 8.1 *The following recurrence relation for the ORF holds if and only if ϕ_n and ϕ_{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]

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

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 ϕ_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 ϕ_n is singular and that

$$E_n = \frac{1}{\kappa_{n-1}}\left[\kappa_n + \frac{\kappa'_n}{Z_n(\alpha_{n-1}))}\right] = \frac{1}{\kappa_{n-1}}\left[\kappa_n + \kappa'_n D_{n+1}\right]. \tag{8.1}$$

It is somewhat harder to prove that $E_n = -C_n E_{n-1}$, $n \geq 2$ [105, Lemma 11.1.6].
If the whole sequence ϕ_n is nonsingular, then one may start the recurrence with $\phi_0 = 1$ and $\phi_{-1} = 0$ if it is assumed that $\int \mathrm{d}\mathring{\mu}(t) = 1$ and $\alpha_0 = \infty$ so that $Z_0(z) = z$.

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

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

Remark 8.1 If $\phi_{n-1} = p_{n-1}/\pi_{n-1}$ is regular, then A_n exists and if ϕ_n regular, then $E_n \neq 0$. Note that A_n can still exist if ϕ_{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. 8.10 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.

Remark 8.2 As a special case we get 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 $A_nC_n \neq 0$.

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., [89, 90, 126, 167, 172, 216, 218, 219, 333, 356, 359, 451]. This has been extended to the rational case in [114]. It involves technicalities that we will not include in this survey.

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]

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

Note that these ψ_k belong to $\mathcal{L}_k$ only for $k \geq 1$. One could e.g., generate σ_k from the recurrence relation for $k \geq 1$ starting with initial conditions $\sigma_{-1} = 1$, $\sigma_0 = 0$. In that case $\sigma_k \in \mathcal{L}_k(\infty)$. This means that they are in $\mathcal{L}_k$ but the numerator is of degree $k-1$ at most.

The recursion for ϕ_n and ψ_n can be written simultaneously as

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

As expressions of the form $\frac{\phi_n}{Z_n}$ will occur often, it is advantageous to introduce the notation $\tilde{\phi}_n(z) = \frac{z\phi_n(z)}{Z_n(z)} = (1 - \frac{z}{\alpha_n})\phi_n(z)$. If a similar notation is introduced for the ψ_n, then the previous relation can be rewritten as

$$\begin{bmatrix} \tilde{\phi}_n(z) & \tilde{\psi}_n(z) \\ \tilde{\phi}_{n-1}(z) & \tilde{\psi}_{n-1}(z) \end{bmatrix} = \begin{bmatrix} E_nZ_{n-1}(z) + B_n & C_n \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\phi}_{n-1}(z) & \tilde{\psi}_{n-1}(z) \\ \tilde{\phi}_{n-2}(z) & \tilde{\psi}_{n-2}(z) \end{bmatrix}.$$

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

[6] If the coefficients of ϕ_n are real, then the coefficients of ψ_n will be purely imaginary because of the factor i in $D(x,z)$.

Theorem 8.2 *For a regular sequence of ORFs it holds that*

$$\tilde{\phi}_n(z)\tilde{\psi}_{n-1}(z) - \tilde{\psi}_n(z)\tilde{\phi}_{n-1}(z) = \mathrm{i}(1+z^2)E_n$$

This is a special case of the *Liouville-Ostrogradskii* formula which can be proved independently of the regularity condition [105, Thm. 11.2.3].

Theorem 8.3 *For the ORFs ϕ_k and the second kind functions ψ_k it holds that for $n \geq 1$*

$$\frac{\tilde{\phi}_n(z)\tilde{\psi}_{n-1}(w) - \tilde{\psi}_n(w)\tilde{\phi}_{n-1}(z)}{w-z} = E_n \left[\sum_{k=1}^{n-1} \phi_k(z)\psi_k(w) - D(z,w) \right].$$

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 [105, Thm. 11.3.1].

Theorem 8.4 *The Christoffel-Darboux relation is*

$$\frac{\tilde{\phi}_n(x)\tilde{\phi}_{n-1}(y) - \tilde{\phi}_n(y)\tilde{\phi}_{n-1}(x)}{x-y} = E_n \sum_{k=0}^{n-1} \phi_k(x)\phi_k(y) = E_n k_{n-1}(x,y)$$

with $k_{n-1}(x,y)$ the reproducing kernel for $\mathcal{L}_{n-1}$.
The confluent form ($y \to x$) is

$$\tilde{\phi}_n'(x)\tilde{\phi}_{n-1}(x) - \tilde{\phi}_n(x)\tilde{\phi}_{n-1}'(x) = E_n k_{n-1}(x,x).$$

A simple proof mimics the proof in the polynomial case. 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 the support 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* [105, Thm. 11.4.1].

Theorem 8.5 (Green's formula) *If $\chi_n(z;s)$ and $\chi_n(w;t)$ are complex solutions of the recurrence relation, then, using again the tilde-notation*

$$\frac{\tilde{\chi}_n(z;s)\overline{\tilde{\chi}_{n-1}(w;t)} - \overline{\tilde{\chi}_n(w;t)}\tilde{\chi}_{n-1}(z;s)}{z-\overline{w}} = E_n \left[\sum_{k=1}^{n-1} \chi_k(z;s)\overline{\chi_k(w;t)} + X \right]$$

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. For example [105, Thm. 11.4.2]:

Corollary 8.6

$$\frac{\tilde{\psi}_n(z)\overline{\tilde{\psi}_{n-1}(w)} - \overline{\tilde{\psi}_n(w)}\tilde{\psi}_{n-1}(z)}{z-\overline{w}} = E_n\left[\sum_{k=0}^{n-1}\psi_k(z)\overline{\psi_k(w)} + 1\right]$$

and

$$\frac{\tilde{\phi}_n(z)\overline{\tilde{\psi}_{n-1}(w)} - \overline{\tilde{\psi}_n(w)}\tilde{\phi}_{n-1}(z)}{z-\overline{w}} = E_n\left[\sum_{k=0}^{n-1}\phi_k(z)\overline{\psi_k(w)} + D(z,\overline{w})\right]$$

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 8.7 (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 ORFs with respect to some bounded positive measure μ with real support.*

In the case of the real line with a cyclic repetition of a finite number of different poles see [331]. For a general sequence of real poles, a proof can be found in [105, Sec. 11.9]. A general proof when the poles can be anywhere in the complex plane is given in [191]. If the measure is supported on $\mathbb{T}$ and the poles of the ORF are all on $\mathbb{T}$, then there is also a three-term recurrence relation and also a Favard theorem can be proved [92].

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

8.3 Quadrature

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

The fundamental Lagrange interpolating basis functions are

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

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

$$\mathfrak{I}_{\mathring{\mu}}[f] = \int f(x)\mathrm{d}\mathring{\mu}(x) \approx \mathfrak{I}_n[f] = \mathfrak{I}_{\mathring{\mu}}[f_n] = \sum_{k=1}^n w_{nk}f(x_{nk}), \quad w_{nk} = \mathfrak{I}_{\mathring{\mu}}[L_{nk}]$$

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

The challenge is to choose the quadrature nodes $\{x_{nk}\}$ 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

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

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 n zeros of ϕ_n are all in $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ and are simple. This can be guaranteed for the numerator of ϕ_n but there could be coincidence with a zero of the denominator, i.e. with one of the α_k, $k = 0, \ldots, n$. If such a cancellation occurs, we can use quasi-orthogonal polynomials to be considered in the next section.

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 [105, Sec. 11.6]:

Theorem 8.8 *If ϕ_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 $f(z) = L_{ni}(z)[L_{ni}(z) - 1] \in \mathcal{L}_{n-1,n-1} \subset \mathcal{L}_{n-1,n}$. Denote the nodes as x_{ni}, $i = 1, \ldots, n$. Since $L_{ni}(x_{nk}) = \delta_{ik}$, it is clear that $f(z)$ vanishes at all points $\{x_{nk}\}_{k=1}^n$. Because the quadrature is exact in $\mathcal{L}_{n-1,n}$, $\mathfrak{I}_{\mathring{\mu}}[f] = \mathfrak{I}_n[f] = 0$. Thus the weights of the QF are given by $w_{ni} = \mathfrak{I}_{\mathring{\mu}}[L_{ni}] = \mathfrak{I}_{\mathring{\mu}}[L_{ni}^2] > 0$. □

We shall give expressions for the weights in a later section.

8.4 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 $\mathbb{R}$ (still assuming the poles are in $\hat{\mathbb{R}}$). 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 ϕ_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 [105, Thm. 11.6.1]:

Theorem 8.9 *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. With some appropriate interpretation, the results will also apply to the case $\tau = 0$ discussed in the previous section.

The main problem is whether the QORF Q_n has n simple zeros on $\hat{\mathbb{R}}$. These are the zeros of the numerator polynomials that do not coincide with a pole. The answer is that we can always use the parameter τ of the QORF to achieve this.

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 can then prove that $Q_n(z)$ has n simple zeros on $\hat{\mathbb{R}}$ [105, Sec. 11.5].

Theorem 8.10 *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 $\hat{\mathbb{R}}$. If ϕ_n is regular, then for all except $n+1$ values of τ all the zeros of p_n will differ from the $\{\alpha_k\}_{k=0}^n$.*

Proof. It can be shown that if ϕ_n is regular (i.e., its numerator does not vanish at α_{n-1}), then the numerators of ϕ_n and ϕ_{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 $\tau \in \hat{\mathbb{R}}$. □

If τ is not one of these forbidden points it will be called *regular* and also the $Q_n(z, \tau)$ will be called *regular*. Consequently, there are at most $n + 1$ values of τ for which a zero of the numerator of $Q_n(z)$ will coincide with one of the poles $\{\alpha_i\}_{i=0}^n$, which means that if ϕ_n is regular, then, except for at most $n + 1$ values of τ, the QORF $Q_n(z, \tau)$ has n simple zeros on $\hat{\mathbb{R}}$. As a particular case, we can say that if ϕ_n is regular and if $\tau = 0$ is a regular value, then ϕ_n will have n simple zeros in $\hat{\mathbb{R}}$.

8.5 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 $\mathbb{A}_n = \{\alpha_k\}_{k=1}^n$ or at ∞ and $\Omega_\mu^{(k)}$ will be continuous at $\alpha \in \mathbb{A}_n$ for $k = 0, 1, \ldots \alpha^{\#} - 1$ where $\alpha^{\#}$ is the multiplicity of α in $\mathbb{A}_n$.

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 $\alpha \in \mathbb{A}_n \cup \{\mathrm{i}\}$ and $k = 0, \ldots, 2\alpha^{\#} - 1$.

Recall that the Ω-functions were defined as a Riesz-Herglotz-Nevanlinna transform of the measure and they are only defined outside the support of the measure. Whenever we write the value of this function in a point on $\mathbb{R}$ we shall always mean the nontangential limit in that point.

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 [105, Thm. 11.10.3]. Thus that Ω_n is an Hermite interpolant for $\Omega_{\mathring{\mu}}$ in the points $\{\mathrm{i}, -\mathrm{i}, \alpha_1, \alpha_1, \alpha_2, \alpha_2, \ldots, \alpha_{n-1}, \alpha_{n-1}\}$ if

the nodes are the zeros of Q_n. The interpolation in the point i is a consequence of $D(x,\mathrm{i})=1$ and thus $\int \mathrm{d}\mathring{\mu}=1=\Omega_{\mathring{\mu}}(\mathrm{i})=\overline{\Omega_{\mathring{\mu}}(-\mathrm{i})}$. If $\tau=0$ and thus $Q_n=\phi_n$ and, as will be shown below, $\Omega_n=-\psi_n/\phi_n$, then α_n is an extra interpolation point.

An explicit expression for Ω_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 x_{nk} are zeros of Q_n will integrate exactly, so that this is equal to $-[\sum_k D(x_{nk},z)w_{nk}]\frac{Z_{n-1}(z)}{Z_n(z)}Q_n(z)$ 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 $\tau=0$ and $=-\frac{\psi_{n-1}(z)}{\phi_{n-1}(z)}$ if $\tau=\infty$).

We are now ready to give alternative expressions for the weights of the QF [105, Thm. 11.6.3 and 4].

Theorem 8.11 *The weights of the interpolating QF whose nodes are the zeros $\{x_k\}$ of $Q_n(z)$ are given by*

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

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 $\mathrm{i}/(1+x_k^2)$.
For the second formula, set as before $\tilde{\phi}_n(x)=(1-\frac{x}{\alpha_n})\phi_n(x)$ and define also $\tilde{Q}_n(x)=(1-\frac{x}{\alpha_n})Q_n(x)=\tilde{\phi}_n(x)+\tau\tilde{\phi}_{n-1}(x)$, then $Q_n(x_k)=0$ implies $\tau=-\tilde{\phi}_n(x_k)/\tilde{\phi}_{n-1}(x_k)$. Then the CD relation gives

$$\tilde{\phi}_{n-1}(x_k)[-\tilde{Q}_n(x)]=\tilde{\phi}_{n-1}(x_k)[\tilde{Q}_n(x_k)-\tilde{Q}_n(x)]=E_nk_{n-1}(x,x_k)(x_k-x).$$

Divide by (x_k-x) and let $x\to x_k$ to get

$$-\tilde{\phi}_{n-1}(x_k)\tilde{Q}_n'(x_k)=E_nk_{n-1}(x_k,x_k).$$

Similarly, using the (complex conjugate of) second formula of Corollary 8.6, we can get

$$\tilde{\phi}_{n-1}(x_k)\tilde{P}_n(x)=E_n\left[\sum_{j=0}^{n-1}\psi_j(x)\phi_j(x_k)-D(x_k,x)\right](x_k-x).$$

Divide by (x_k-x) and let $x\to x_k$ to get

$$\tilde{\phi}_{n-1}(x_k)\tilde{P}_n(x_k)=-\mathrm{i}E_n(1+x_k^2).$$

A division then gives the result because $\frac{\tilde{P}_n(x_k)}{\tilde{Q}_n'(x_k)}=\frac{P_n(x_k)}{Q_n'(x_k)}$. □

[7] Again, if the coefficients of ϕ_k's are real, the coefficients of ψ_k's and hence of P_n are purely imaginary.

8.6 Quadrature with kernels

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

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

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 [105, Thm. 11.10.5]. 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.

Remark 8.3 It is not clear whether a Carathéodory function can be found in this case of the kernels. Neither is it clear how much can be generalized when the second variable in the kernel is not i but a more general $w \in \mathbb{C}$. (see also [105].)

8.7 Pseudo-orthogonal functions

In the case of the circle, the PORF were defined as $\phi_n + \tau\phi_n^*$, with ϕ_n an ORF and $\tau \in \mathbb{T}$. For the case of the real line, the QORF were defined as $\phi_n + \tau_n \frac{Z_n}{Z_{n-1}}\phi_{n-1}$, $\tau \in \mathbb{R}$ and ϕ_n and ϕ_{n-1} successive ORFs. The factor Z_n/Z_{n-1} arranged for the fact that the second term and hence the QORF to be orthogonal to $\mathcal{L}_{n-1}(\alpha_n)$, a subspace if dimension $n-1$.

In the polynomial case, $Z_n = Z_{n-1}$ and the extra factor will disappear by cancellation. One may wonder if it would not be a simpler choice to define pseudo-ORF by $\phi_n + \tau\phi_{n-1}$, $\tau \in \hat{\mathbb{R}}$. These will still be in $\mathcal{L}_n$ and they will be orthogonal to an $n-1$ dimensional subspace $\mathcal{L}_{n-2}$. It turns out that these pseudo-ORF have quite similar properties as the QORF. For appropriate choice of τ they have simple zeros that are in the appropriate places and one has separation of zeros, etc. In short, they can be used to construct QF with positive weights that are exact in a space of maximal dimension which is $\mathcal{L}_n \cdot \mathcal{L}_{n-2}$ instead of $\mathcal{L}_{n-1} \cdot \mathcal{L}_{n-1}$ as for the QORF. There are some practical computational issues though that make them somewhat less attractive from that point of view. For the Stieltjes case ($\operatorname{supp}\mu = [0,\infty)$) the properties of pseudo-ORF and their application in the construction have been investigate in [120].

8.8 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 α_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 $+\infty$, then 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 ϕ_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 ϕ_n (set $\sigma_{-1} = -1$ and $\sigma_0 = 0$). Note that these are simplified versions of the functions of the second kind ψ_k that we introduced before. This is made possible because there is no mass point at $+\infty$. Note also that the factor i 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 (we use again the tilde-notation, recall also $\alpha_0 = \alpha_{-1} = \infty$)

$$\begin{bmatrix} \tilde{\phi}_n & \tilde{\sigma}_n \\ \tilde{\phi}_{n-1} & \tilde{\sigma}_{n-1} \end{bmatrix} = \begin{bmatrix} b_n & a_n \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\phi}_{n-1} & \tilde{\sigma}_{n-1} \\ \tilde{\phi}_{n-2} & \tilde{\sigma}_{n-2} \end{bmatrix}, \quad \begin{bmatrix} \tilde{\phi}_0 & \tilde{\sigma}_0 \\ \tilde{\phi}_{-1} & \tilde{\sigma}_{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We have set $a_n = C_n$ and $b_n = E_n Z_{n-1} + B_n$.

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

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

Assuming that the poles are alternating between the points near $-\infty$ and 0 like $-\infty < \alpha_{2m} \le \alpha^- < \alpha^+ \le \alpha_{2m+1} < 0$, then it can be shown along these lines that on (α^-, α^+), 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 Ω_n interpolates Ω_μ 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 ϕ_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 [114, 120].

For the interval, say $I = [-1, 1]$, the poles can be chosen anywhere in $\hat{\mathbb{C}} \setminus I$. In principle, this is not essentially different from the general case where the measure is supported on the whole real line. Much can be inherited though from the results on the unit circle. Indeed, following Szegő [509] (see also [81]) for the polynomial case 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 μ 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. [200, 539]. 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 τ is real (i.e., ± 1) and hence the transformed nodes will be real (and inside 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 of I (it is a *Gauss-Radau* formula) and if $\tau = -1$ and n is even, both end points ± 1 will be nodes and we get a *Gauss-Lobatto* formula [197]. More about these QFs with boundary node(s) 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 [194].

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 [534].

For simple interpolatory rational QF on an interval, possibly with a complex weight function see [141, 518].

8.9 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 we do need the complex conjugation in the inner product $\langle f, g\rangle = \int_{-1}^{1} \overline{f(\overline{x})} g(x) \mathrm{d}\mu(x)$. 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 ϕ_n vanishes at α_{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 three-term recurrence relation has been proved in [194]. See also [191, 196] for related results.

Also rational Gauss QF can be obtained using zeros of either QORFs (which are defined to be functions in $\mathcal{L}_n \backslash \mathcal{L}_{n-1}$ and orthogonal to $\mathcal{L}_{n-1}(\overline{\alpha}_{n-1})$), or of 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 $\phi_n(x) = \frac{b_n(x)}{\overline{b_n(\overline{x})}} \overline{\phi_n(\overline{x})} = \frac{b_n(x)}{b_{n^{\bar{*}}}(x)} \phi_{n^{\bar{*}}}(x)$. In both cases, their zeros depend on a complex parameter τ 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 [198].

9 Computational aspects

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

However, in practice, this technique is almost never used. Since the *Golub-Welsch algorithm* [308] it is common practice to compute weights and nodes via an eigenvalue problem. We shall now adapt this to the rational case.

9.1 The unit circle

Let us find out how the polynomial eigenvalue problems for unitary Hessenberg or CMV matrices can be adapted to the rational case. For more details, please consult [75, 546]. If we relax the normalization of the ORF a bit, we may rewrite the recurrence of Theorem 7.8 as follows:

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

with

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

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 $H = H_\infty$, namely $\tilde{\zeta}_D(H_\infty)$ where $D = D_\infty$ is the diagonal matrix whose entries are the α_k and $\tilde{\zeta}_D(H) = \eta_D^{-1}(H+D)(I+D^H H)^{-1}\eta_{D^H}$ where the superscript ^{H} refers to the adjoint and $\eta_D = \sqrt{I - DD^H}$. Note that $\tilde{\zeta}_D(H)$ is a matrix version of a Möbius transform: $\tilde{\zeta}_D(H) = \mathfrak{m}_D^{-1}[H] = \mathfrak{m}_{-D}[H]$. Again the transformation of the original basis is clearly reflected here. The polynomial case uses powers, i.e., products of the variable z, but in the rational case, these are replaced by a product of different factors of the form $\zeta = \mathfrak{m}_\alpha[z] = \frac{z-\alpha}{1-\overline{\alpha}z}$ whose inverse is $z = \mathfrak{m}_{-\alpha}[\zeta] = \tilde{\zeta}_\alpha(\zeta) = \frac{\zeta+\alpha}{1+\overline{\alpha}\zeta}$, $\alpha \in \{\alpha_1, \alpha_2, \ldots\}$.

For the nodes of the quadrature, we need the zeros of PORFs $Q_n(z,\tau) = \phi_n(z) + \tau\phi_n^*(z)$, $\tau \in \mathbb{T}$. By using the recurrence relation we can rewrite this in terms of ϕ_{n-1} as

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

In concrete terms, this means that we can compute $Q_n(z,\tau)$ by using the recurrence for the ORF with parameters $\upsilon_1, \upsilon_2, \ldots, \upsilon_{n-1}$, but use $\tilde{\tau} = \tilde{\zeta}_{\upsilon_n}(\tau) \in \mathbb{T}$ instead of υ_n. As in the polynomial case we denote the corresponding Hessenberg matrix by H_n^τ. Then the unitary truncation of $\tilde{\zeta}_D(H)$ is given by $V_n^\tau = \tilde{\zeta}_D(H_n^\tau)$. So we have the following theorem [75].

Theorem 9.1 *The zeros t_{nk} of the PORF $Q_n(z,\tau)$ are the eigenvalues of $V_n^\tau = \tilde{\zeta}_D(H_n^\tau)$. Equivalently, the t_{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 $\boldsymbol{\phi}_n(t_{nk})\eta_{D_n}^{-1/2}$ with $\boldsymbol{\phi}_n = [\phi_0, \ldots, \phi_{n-1}]$.*

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

$$\chi_{2n} = B^e_{n*}\phi^*_{2n}, \quad \chi_{2n+1} = B^o_{n*}\phi^*_{2n+1}, \quad n \geq 0,$$

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

$$\mathcal{X}_{2n} = B^e_{n*}\mathcal{L}_{2n} = \text{span}\{B^e_{0*}, B^o_1, B^e_{1*}, \ldots, B^e_{(n-1)*}, B^o_n\}$$

and

$$\mathcal{X}_{2n+1} = B^e_{(n+1)*}\mathcal{L}_{2n+1} = \text{span}\{B^e_{0*}, B^o_1, B^e_{1*}, \ldots, B^o_n, B^e_{n*}\},$$

i.e., $\mathcal{X}_{2n}, \mathcal{X}_{2n+1}$ are the sets of ORF whose poles are in $\{\hat{\alpha}_1, \alpha_2, \hat{\alpha}_3, \ldots, \alpha_{2n-2}, \hat{\alpha}_{2n-1}\}$ and in $\{\hat{\alpha}_1, \alpha_2, \hat{\alpha}_3, \ldots, \alpha_{2n-1}, \hat{\alpha}_{2n-1}, \alpha_{2n}\}$ respectively (repeated according to multiplicity and $\hat{\alpha}_k = 1/\overline{\alpha}_k$). Moreover these are orthogonal, because the orthogonality of the χ_k is equivalent to the orthogonality of the ϕ_k.

Note that $\mathcal{X}_{2n} = \mathcal{L}_{n,n} = \mathcal{L}_{n*} \cdot \mathcal{L}_n = \mathcal{L}_{n*} + \mathcal{L}_n$ and $\mathcal{X}_{2n+1} = \mathcal{L}_{n,n+1} = \mathcal{L}_{n*} \cdot \mathcal{L}_{n+1} = \mathcal{L}_{n*} + \mathcal{L}_{n+1}$ where we have used $\mathcal{L}_{k*} = \{f : f_* \in \mathcal{L}_k\}$. These spaces $\mathcal{L}_{p,q}$ 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^2_μ, we get again a simple matrix representation of the multiplication operator, given as an operator Möbius transform of the C_∞ matrix, namely $\tilde{\zeta}_{D_\infty}(C_\infty)$, the same transformation used before.

With the same arguments as used for the Hessenberg representation we can truncate and modify the last Schur parameter υ_n into $\tilde{\zeta}_{\upsilon_n}(\tau)$, $\tau \in \mathbb{T}$, so that the zeros of the quadrature formula are given as follows [75].

Theorem 9.2 *The zeros t_{nk} of the PORF $Q_n(z, \tau)$ are the eigenvalues of $U^\tau_n = \tilde{\zeta}_{D_n}(C^\tau_n)$ with (right) eigenvectors $\boldsymbol{\chi}_n(t_{nk})^H = [\chi_0(t_{nk}), \chi_1(t_{nk}), \ldots, \chi_{n-1}(t_{nk})]^H$.*
Equivalently, these zeros are the generalized eigenvalues of the pencil $(D_n + C^\tau_n, I_n + D^H_n C^\tau_n)$ and the corresponding (left) eigenvectors are given by $\boldsymbol{\chi}_n(t_{nk})\eta^{-1}_{D_n}$.

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

$$U^\tau_n = \tilde{\zeta}_{D_n}(C^\tau_n) = \eta_{D_n}([C^\tau_n]^H + D^H_n)^{-1}(I_n + [C^\tau_n]^H D_n)\eta^{-1}_{D_n}.$$

This implies that the eigenvalue problem $U^\tau_n E_n = E_n \Lambda_n$ is also

$$\begin{aligned}
\eta_{D_n}((C^\tau_n)^H + D^H_n)^{-1}(I_n + (C^\tau_n)^H D_n)\eta^{-1}_{D_n} E_n &= E_n \Lambda_n \\
((C^\tau_n)^H + D^H_n)^{-1}(I_n + (C^\tau_n)^H D_n) F_n &= F_n \Lambda_n, \quad F_n = \eta^{-1}_{D_n} E_n \\
((C^\tau_{on})^H + C^\tau_{en} D^H_n)^{-1}(C^\tau_{en} + (C^\tau_{on})^H D_n) F_n &= F_n \Lambda_n,
\end{aligned}$$

since also C^τ_{on} and C^τ_{en} are unitary truncations of the matrices $C_{o\infty}$ and $C_{e\infty}$. This means that we also have the following [75].

Theorem 9.3 *The zeros t_{nk} of the PORF are given as generalized eigenvalues of the pencil $(C^\tau_{en} + (C^\tau_{on})^H D_n, (C^\tau_{on})^H + C^\tau_{en} D^H_n)$. The corresponding (right) eigenvectors are $\boldsymbol{\chi}_n(t_{nk})^H = \{[\chi_0(t_{nk}), \chi_1(t_{nk}), \ldots, \chi_{n-1}(t_{nk})]\eta^{-1}_{D_n}\}^H$.*

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

The previous theorems tell us how to compute the nodes of the rational Szegő QF. The corresponding weights w_{nk} are given (as in the polynomial case) by the values that we had before $w_{nk} = 1/\sum_{j=0}^{n-1} |\phi_j(t_{nk})|^2$ where ϕ_k are the orthonormal ORF. This is the inverse of the squared norm of the eigenvectors in the previous theorems. Since these eigenvectors $\boldsymbol{e}_{nk}$ are only defined up to a multiplicative constant, we can choose a normalized eigenvector $\boldsymbol{f}_{nk} = \boldsymbol{e}_{nk}/\|\boldsymbol{e}_{nk}\|$ and for example with the eigenvector $\boldsymbol{\phi}_n(t_{nk})$ we should have $\boldsymbol{f}_{nk} = c\boldsymbol{\phi}_n(t_{nk})$ with $c = \|\boldsymbol{\phi}_n(t_{nk})\|^{-1} = w_{nk}^{1/2}$. Because ϕ_0 was supposed to be 1 by our normalization of the measure, it follows that w_{nk} is the square of the modulus of the first element in the normalized eigenvector. For the cases with the ORF χ_k it suffices to note that $\sum_{j=0}^{n-1} |\chi_j(t_{nk})|^2 = \sum_{j=0}^{n-1} |\phi_j(t_{nk})|^2$ to come to the same conclusion.

Remark 9.1 In a situation where the measure on $\mathbb{T}$ is symmetric (meaning that $\mathrm{d}\mu(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{d}\mu(\mathrm{e}^{-\mathrm{i}\theta})$) and if also the poles are chosen in a symmetric way, (meaning that they are either real or complex conjugate), then it might be possible to divide the complexity of the computations by a factor 2. See for example [78] for the polynomial case. About the same effect can result if complex arithmetic can be replaced by real arithmetic by introducing the complex conjugate pairs in 2×2 blocks, i.e., working with a Hessenberg or a CMV matrix matrix having 2×2 blocks. It has not been cleared out yet how to achieve the latter. An attempt is made in [137].

Remark 9.2 The functions χ_k are also a generalization of a biorthogonal system of (trigonometric) polynomials as discussed originally by Szegő in [508] but more extensively in [151, 169, 170, 173]. Indeed, $\rho_n = \chi_{2n} \in \mathcal{X}_{2n} = \mathcal{L}_{n-1,n}$ and it is orthogonal to $\mathcal{X}_{2n-1}$ and hence also to $\mathcal{X}_{2n-2}$. Therefore $\langle \rho_n, \rho_m \rangle = \delta_{n,m}\|\rho_n\|^2$. Similarly $\sigma_n = \chi_{2n+1} \in \mathcal{M}_{2n+1} = \mathcal{L}_{n,n}$ and it is orthogonal to $\mathcal{X}_{2n}$ and hence also to $\mathcal{X}_{2n-1}$. Therefore $\langle \sigma_n, \sigma_m \rangle = \delta_{n,m}\|\sigma_n\|^2$. Furthermore $\langle \rho_n, \sigma_m \rangle = 0$ for all $n, m = 0, 1, \ldots$.

9.2 The real line

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 $\alpha_0 = \infty$)

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

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

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 ϕ_n, and we suppose that it is nonzero and not in $\{\alpha_0, \ldots, \alpha_{n-1}\}$, then it can be rewritten as

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

which says that ξ is a generalized eigenvalue of the pencil $(J_n, I_n + D_n^{-1}J_n)$ and $\boldsymbol{\phi}_n(\xi)$ is the corresponding eigenvector. The quadrature weight at ξ can be written once more as $\|\boldsymbol{\phi}_n(\xi)\|^{-2}$ where it is supposed that ϕ_0, the first component of $\boldsymbol{\phi}_n(\xi)$, 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 $\boldsymbol{\phi}_\infty$.

An ad hoc interpretation, (which could be turned into a hard argument) is as follows. If J_∞ 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. This has been worked out in more detail in the previous subsection for the case of the circle. Similar results can be obtained for the case of the real line.

The original approach can be found in [113]

Remark 9.3 This approach is based on the existence of the three-term recurrence relation. This only holds when the system is regular. If there are singular ϕ_k in the sequence, then some efforts have very recently been undertaken to cure the problem. See [486, 488].

Remark 9.4 In the case of poles anywhere in the complex plane, a similar generalized eigenvalue problem can be considered. There are however even more extra conditions needed to guarantee the existence of the recurrence relation. Although more complicated, most of the results can be generalized. See for example [198, 530].

If one wants to use the zeros of a QORF as the nodes, it can be shown that assuming regularity, there exist constants $\check{E}_n, \check{B}_n, \check{C}_n$, depending on τ such that

$$\tilde{Q}_n(x, \tau) = (\check{E}_n Z_{n-1}(x) + \check{B}_n)\tilde{\phi}_n(x) + \check{C}_n\tilde{\phi}_{n-1}(x)$$

where we used again the tilde-notation meaning that $\tilde{X}_k(x) = (1 - \frac{x}{\alpha_k})X_k(x)$ for $X \in \{Q, \phi\}$. See [195]. This means that we can compute the zeros of Q_n and the corresponding weights of the QF again by solving a generalized eigenvalue problem of the same form as above, but where the bottom entries of J_n should be modified.

10 Explicit examples of ORFs

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 [554] Takenaka-Malmquist functions (after [510] and [404]). They are given by

$$B_0 = 1, \quad \frac{\sqrt{1-|\alpha_k|^2}}{1-\overline{\alpha}_k z}\tilde{B}_{k-1}(z) \quad \text{with } \tilde{B}_k = \zeta_0 B_k = \prod_{j=0}^{k} \zeta_k, \quad k \geq 1, \quad \zeta_0(z) = z. \tag{10.1}$$

Some early investigation of these functions is found in the Russian literature in papers by M.M. Djrbashian [220–225]. (An English translation of the fundamental paper [223] is available as [426].) Some spectral properties of the corresponding Gram matrix can be found in [76]. Later this was studied by V. Pan [392, 461–470]. Traces of ORF for the Lebesgue measure are also found in system theory literature under different names: Kautz systems [371] (one finite pole repeated), or Laguerre systems [391] (a pair of complex conjugate poles repeated). More general rational functions in identification were only introduced later in the last decade of the previous century [434–437]. In circuit theory these are known in the context of Darlington synthesis [28, 181, 213]. See also the application in Section 16.3.

Another easy exception is given by the Poisson weight: $\mathrm{d}\mu(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\frac{1-|r|^2}{|\mathrm{e}^{\mathrm{i}\theta}-r|^2}\mathrm{d}\theta$ where $0 \leq r < 1$. The recurrence parameters are independent of the chosen α_k and are all zero except the first which is $-r$. This can be seen as a rational modification of the Lebesgue measure $\mathrm{d}\lambda(\mathrm{e}^{\mathrm{i}\theta}) = \frac{\mathrm{d}\theta}{2\pi}$. It was first considered in [550]. More on rational modifications follows below in Section 10.3.

Somewhat less trivial, yet important examples are given by rational versions of the Chebyshev polynomials that are well known orthogonal polynomials on $[-1, 1]$ with respect to weights $(1-x)^\alpha(1+x)^\beta$, $\alpha, \beta \in \{\pm 1/2\}$. See also [70, 419, 420, 422] for the weight $(1-x^2)^{-1/2}$. We will consider the problem on the real line first and then do the analog for the unit circle.

10.1 Rational Gauss-Chebyshev quadrature

The idea is to consider ORF on $I = [-1, 1]$ with poles $\mathbb{A} = \{\alpha_1, \alpha_2, \ldots\} \subset \hat{\mathbb{C}}^I$ where $\hat{\mathbb{C}}^I = \hat{\mathbb{C}} \setminus I = (\mathbb{C} \cup \{\infty\}) \setminus [-1, 1]$ denotes the cut Riemann sphere. These ORFs 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 $z \in \mathbb{D}$.

If $\mathbb{A} = \{\alpha_1, \alpha_2, \ldots\}$ is a set of poles in $\hat{\mathbb{C}}^I$, then these correspond to a set $\mathbb{B} = \{\beta_1, \beta_2, \ldots\} \subset \mathbb{D}$ where $\beta_k = \mathfrak{j}^{inv}[\alpha_k]$. With these poles in $\mathbb{B}$, 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 as before $f_{\bar{*}}(z) = \overline{f(\overline{z})}$ 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 $\mathcal{L}_{n^{\bar{*}}} = \mathcal{L}_n^{\bar{*}} = \mathcal{L}_n$. It is then not difficult to check that $Z_k(\mathfrak{j}[z]) = E\zeta_{k^{\bar{*}}}(z) + F + E/\zeta_k(z)$ where E and F are constants depending on β_k.

Let us now consider the following weights on I.

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

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 [509, p. 31] in the case that all $\alpha_k = \infty$ and $w(x) = w(\cos\theta) = |g(\mathrm{e}^{\mathrm{i}\theta})|^2$ 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*.

One may check that with $\alpha_k \in \hat{\mathbb{R}}^I$, $\zeta_k(z) = \frac{z-\beta_k}{1-\bar{\beta}_k z}$, $x = \mathfrak{j}[z]$, i.e., $x = \cos\theta = \mathfrak{j}[\mathrm{e}^{\mathrm{i}\theta}]$, $z = \mathrm{e}^{\mathrm{i}\theta}$ and $\alpha_k = \mathfrak{j}[\beta_k]$, it follows that $\zeta_k(z) = \exp\left\{2\mathrm{i}\arctan\frac{\sin\theta}{\cos\theta - \beta_k} - \mathrm{i}\theta\right\}$.

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 $w^{(i)}$, $i = 1, 2, 3$ we may associate parameters as in the table

i	c_i	d_i	p_i	q_i
1	1	1	1	-1
2	3/2	0	1	1
3	2	0	2	1

so that we have the following theorem [203, Thm. 3.2]

Theorem 10.1 *With a normalizing factor $N_i = \sqrt{2^i/\pi}\sqrt{1-|\beta_n|^2}$, the orthonormal basis function in $\mathcal{L}_n \setminus \mathcal{L}_{n-1}$ for the weight $w^{(i)}$ ($i \in \{1,2,3\}$) is given by*

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

and $\phi_0^{(i)} = \sqrt{p_i/\pi}$ 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 $\phi_n^{(i)}$ or possibly of the PORF $Q_n^{(i)}(x,\tau) = \phi_n^{(i)}(x) + \tau(\phi_n^{(i)})^*(x)$, $\tau \in \mathbb{T}$.

For the general setting of poles in $\hat{\mathbb{C}}^I$, this is further elaborated in [203]. 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 [535]. The expression for the ϕ_n is as in the previous theorem, but the sub-star in $B_{(n-1)^*}$ can be removed. In the latter situation of real poles, letting the zeros x_{nk} of ϕ_n be denoted as $x_{nk} = \cos\theta_{nk}$, then these can be obtained as follows [535, Thm. 4.1,Thm. 4.3].

Theorem 10.2 *Consider Gauss-Chebyshev QFs for any of the Chebyshev weights $w^{(i)}$, $i = 1, 2, 3$ with their associated parameters of the previous table and assume rational functions with real poles in $\hat{\mathbb{R}}^I$. Denote the nodes as $x_{nk} = \cos(\theta_{nk})$, then the θ_{nk}, $k = 1, \ldots, n$ for the ith Chebyshev weight are solutions of the equations (depending on n, k and i)*

$$f_n(\theta) - (n - c_i)\theta = \frac{\pi}{2}(2k - d_i) \quad \textit{with} \quad f_n(\theta) = 2\sum_{j=1}^{n}{}' \arctan\frac{\sin\theta}{\cos\theta - \beta_j} \ .$$

(The prime in $\sum'$ means that the last term is halved.) The quadrature weight $w = w_{nk}$ corresponding to a node $x = x_{nk}$ is given by

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

($\arctan(y/x)$ refers to the argument of the complex number $x + \mathrm{i}y$.) The equation for the nodes looks complicated, but efficient algorithms can be designed because in between the poles, the function f_n is monotonically increasing. Computing 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 [201], 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 [203] for more details and to [202] for more on the algorithmic aspects.

10.2 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 $z = \mathrm{e}^{\mathrm{i}\theta}$. For the *Chebyshev weights* $w(x) = (1 - x)^\alpha(1 + x)^\beta$, $\alpha, \beta \in \{\pm 1/2\}$ 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, the ORF with respect to $\omega^{(1)}$ is given by the Takenaka-Malmquist basis (10.1). For the weight functions $\omega^{(i)}$, $i = 3, 4$, also explicit expressions for the ORF are known. We have [200, 539], [77, Thm. 3.1].

Theorem 10.3 *Define $N_n(z) = 1 + zB'_{n-1}(z)/B_{n-1}(z)$ where B_{n-1} is a Blaschke product (with zeros α_k defining the rational functions) and the prime means derivative. 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

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

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 ϕ_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 [77].

Note that $\omega^{(2)}(\theta) = \sin^2\theta = \frac{1}{2}|z+1|^2\omega^{(3)}(\theta)$, $z = \mathrm{e}^{\mathrm{i}\theta}$ and hence that $\omega^{(2)}$ can be considered as a rational modification of $\omega^{(3)}$. The construction of ORF for rational modifications will be considered below in Section 10.3. For completeness, we give an indication of what the ORF for $\omega^{(2)}$ look like [77, Thm. 3.2].

Theorem 10.4 *The ORF (not normalized) for $\omega^{(2)}$ have the following form*

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

where the constants d_n, e_n, f_n, g_n are the solution of a 4×4 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 ϕ_n, it is possible to reformulate the problem such that the nodes are the zeros of a polynomial equation.

Using the (confluent) Christoffel-Darboux relations and assuming the ϕ_k are orthonormal then the known formula $w_\xi = 1/\sum_{k=0}^{n-1}|\phi_j(\xi)|^2$ for the weight corresponding to node ξ in the n-point rational Szegő QF can be rewritten as (a prime denotes derivative)

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

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 ξ is $w_\xi = \frac{2\pi}{N_n(\xi)}$ where $N_n(z)$ is defined in Theorem 10.3. All the details for the different weights $w^{(i)}$, $i = 1,2,3,4$ are given in [77]. For the polynomial case see [182, 184].

10.3 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 [300, 301]. See also [157, 183, 185, 279–283, 405]. 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 $\mathrm{d}\tilde{w}(z) = \mathrm{d}w(z) + m\delta_\alpha(z) + \overline{m}\delta_{\hat{\alpha}}(z)$, $\hat{\alpha} = 1/\overline{\alpha}$, $\alpha \in \mathbb{C} \setminus \{0\}$, $m \in \mathbb{C}$ which is known as *Uvarov transformation*. With the Christoffel transform one has $\langle f, g\rangle_C = \langle (z-\alpha)f(z), (z-\alpha)g(z)\rangle_w$ and with the Geronimus transformation $\langle f, g\rangle_G = \left\langle \frac{f(z)}{z-\alpha}, \frac{g(z)}{z-\alpha} \right\rangle_w$.

Rather than formulating the general situation, we shall look at a particular case. We have [189, Thm. 5].

Theorem 10.5 *Let* $\mathbb{A} = \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{D}$ *and* $\mathbb{A}_\beta = \{\beta, \alpha_1, \ldots \alpha_n\} \subset \mathbb{D}$, *and* $\gamma \in \hat{\mathbb{D}} = \mathbb{D} \cup \mathbb{T}$. *Set* $\mathrm{d}\dot{\mu}(z) = \left|\frac{z-\gamma}{1-\overline{\beta}z}\right|^2 \mathrm{d}\mu(z)$. *Let* $\phi_n(z;\dot{\mu})$ *be the ORF for the sequence* $\mathbb{A}$ *and measure* $\dot{\mu}$ *and let* $\dot{\phi}_{n+1}(z;\mu)$ *be the ORF for the sequence* $\mathbb{A}_\beta$ *and the measure* μ. *Then for* $n \geq 1$

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

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

In [189, 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 $\beta = r\mathrm{e}^{\mathrm{i}\theta}$, $r < 1$ and $\mathrm{d}\mu = \mathrm{d}\lambda$, the Lebesgue measure, then $\mathrm{d}\dot{\mu}(z) = \frac{\mathrm{d}\lambda(z)}{1-2r\cos\theta+r^2}$ is proportional to the *Poisson weight* that we mentioned before and which was fist discussed by Waadeland [550] in connection with Szegő quadrature formulas.

It is often possible to directly connect a weight w_ξ corresponding to a node ξ of the rational Szegő QF with respect to measure μ and the corresponding weight for the measure $\dot{\mu}$. Using this idea, it is not difficult to prove that the ORF for the Poisson measure $\mathrm{d}\mathring{\mu}(t) = \frac{(1-|r|^2)\mathrm{d}t}{2\pi\mathrm{i}(t-r)(1-\overline{r}t)}$, $t \in \mathbb{T}$ are given by $\phi_n(z) = k_n\frac{(r-z)B_n(z)}{z-\alpha_n}$. One may want to check the special cases $r = 0$, $r = \alpha_n$, and the polynomial case: $r \neq 0$ but all $\alpha_k = 0$. See [119].

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 very elegant in terms of linear algebra. For example if $\dot{\mu} = |z-\alpha|^2\mu$ is a measure on $\mathbb{T}$ (Christoffel transform) and if the corresponding Hessenberg matrices for the orthogonal polynomials for $\dot{\mu}$ and μ are $\dot{H}$ and H respectively, then the transform from H to $\tilde{H}$ is given by a QR step: $H - \alpha I = QR$ then $\dot{H} - \alpha I = RQ$ as observed in [558]. For the real line, the QR steps can be replaced by LR or RL steps for a linear factor, [256, 275, 304, 370, 558] which is again a QR step for a quadratic factor [293].

It is currently still unclear how much of this type of linear algebra techniques for these spectral transformations can be generalized to the rational case in a stable way. The adaptation needed is that the Jacobi or unitary Hessenberg matrix from the polynomial case become more complicate to take the set of poles α_k into account. See also [345] for a stable implementation for Geronimus transform of a Hessenberg matrix (polynomial case) when the latter is stored in its factored form.

10.4 Near best rational approximation

It is well known that for polynomial L_∞ (i.e. 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 approximating choice is to take as interpolation points 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

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

and

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

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 α_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 $\mathcal{U}_{n-1}$.

With the conditions just introduced, it can be shown that

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

Once this is known, we can apply either the algorithm based on Newton iteration described in [528, 536] 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, 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 $x\boldsymbol{\phi}_n(x)(I_n - S_n + D_n^{-1} J_n) = \boldsymbol{\phi}_n(x) J_n$, where an extra matrix S_n appears, that is a zero matrix, except for its first superdiagonal that contains elements proportional to $1/Z_{(n-2)^\bar{*}}(\alpha_{n-2})$. Note that for real poles $S_n = 0$. Also, to construct the rational interpolant, it is best to represent it in barycentric coordinates. This means that, if the interpolation points are $\{x_j : j = 0, \ldots, n\}$ then we write the rational interpolant as p_n/π_m with $p_n(x) = \sum_{j=0}^n \lambda_j f(x_j) \prod_{k\neq j} (x - x_k)$ where $\lambda_j = 1/t'(x_j)$ with t the Chebyshev ORF delivering the interpolation points.

Suppose that the nodes x_j are the zeros of $\mathcal{T}_{n+1}(x)$ with $m \leq n$ then it can be proved that $\mathcal{T}'_{n+1}(x_j) = \frac{\pm g_{n+1}(z_j)}{\sin\theta_j}$ with $x_j = \cos\theta_j$, $z_j = \mathrm{e}^{\mathrm{i}\theta_j}$ 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 $\lambda_j = (-1)^j \frac{\sin\theta_j}{g_{n+1}(z_j)}$ with $g_{n+1}(z_j) = \pi/w_j$. Here $w_j = 1/\sum_{k=0}^{n} |\mathcal{T}_k(x_j)|^2$ is the quadrature weight. See also [56, 375] for unweighted quadrature based on barycentric interpolation.

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 $\lambda_j = (-1)^j\delta_j/g_n(z_j)$ where $\delta_j = 1/2$ for $j = 0$ and $j = n$ and 1 otherwise. Furthermore for $z_j \neq \pm 1$: $g_n(z_j) = \frac{\pi}{w_j}(1-x_j^2) = \frac{\pi}{w_j}\sin^2\theta_j$ where $w_j = 1/\sum_{k=0}^{n-2} |\mathcal{U}_k(x_j)|^2$ is the quadrature weight.

From this discussion it is clear that the same algorithms used for the computation of the rational Chebyshev quadratures can also be applied to solve this near best approximation problem.

For more details, consult [528, 529, 536].

10.5 More connections between the circle and the real line

Since several results of this section rely on mappings between the circle and an interval, we also mention another connection that can be made between an integral over $\mathbb{R}$ and over an interval, and hence also over the unit circle. Consider an integral of a 2π-periodic function $f(e^{\mathrm{i}\theta})$ of the form $\int_{\mathbb{R}} f(e^{\mathrm{i}\theta})w(\theta)\mathrm{d}\theta$ where $w > 0$. Under some mild conditions we can then write

$$\int_{\mathbb{R}} f(e^{\mathrm{i}\theta})w(\theta)\mathrm{d}\theta = \sum_{j\in\mathbb{Z}} \int_{-\pi}^{\pi} f(e^{\mathrm{i}\theta})w(\theta+2\pi j)\mathrm{d}\theta = \int_{-\pi}^{\pi} f(e^{\mathrm{i}\theta})\omega(\theta)\mathrm{d}\theta,$$

where

$$\omega(\theta) = \sum_{j\in\mathbb{Z}} w(\theta+2\pi j)$$

is a 2π-periodic weight.

In some cases, this can result in an $\omega(\theta)$ that is known. An example: when $w(\theta) = \sqrt{2/\pi}\exp\{-tx^2\}$. If we set $q = \sqrt{\exp\{-1/t\}}$, then it is known that

$$\omega(\theta) = \frac{1}{\sqrt{2\pi\log(1/q)}} \sum_{j\in\mathbb{Z}} \exp\left(-\frac{(\theta-2\pi j)^2}{2\log(1/q)}\right).$$

This is not an explicit formula, but it is recognized as the weight related to the so-called *Rogers-Szegő polynomials* [36, 174, 177] which are defined by

$$H_n(z) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k, \quad \begin{bmatrix} n \\ k \end{bmatrix} = \frac{(n)_q}{(k)_q(n-k)_q},$$

with shifted factorials defined as $(n)_q = \prod_{j=1}^{n}(1-q^j)$, and the moments are known:

$$m_k = \int_{-\pi}^{\pi} e^{-\mathrm{i}k\theta}\omega(\theta)\mathrm{d}\theta = q^{\frac{k^2}{2}}, \quad k \in \mathbb{Z}.$$

See also [175, 310] for another example illustrating this technique.

So far, there is no rational analog available but it is clear that in principle the same techniques can be applied in the rational case as well.

11 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 τ 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. The classical polynomial situations of Gauss-Radau [487] and Gauss-Lobatto [395, les 24, §207] will drop out as special cases.

11.1 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 $\mathfrak{I}_n[f] = \sum_{k=1}^n w_{nk} f(t_{nk})$ whose nodes $\{t_{nk}\}_{k=1}^n$ 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 t_{nk} are on $\mathbb{T}$ and the weights w_{nk} are all positive.

11.1.1 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 τ. It is obvious that by choosing $\tau = \tau(\xi) = -\phi_{n+1}(\xi)/\phi_{n+1}^*(\xi)$ then ξ will be a zero of $Q_{n+1}(z,\tau)$. Since $\tau(\xi)$ as a function of ξ is a Blaschke product with winding number $n+1$, there are $n+1$ instances of $\xi \in \mathbb{T}$ where this value of τ 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 ξ by varying τ over $\mathbb{T}$.

This formula is thus a rational Szegő QF with a particular choice of the parameter $\tau = \tau(\xi)$. It has the form $\sum_{k=1}^{n+1} w_{n+1,k}(\tau) f(t_{n+1,k}(\tau))$. One of the $t_{n+1,k} = \xi$ 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.

There is another way to look at this formula. With the particular choice of τ mentioned above, the PORF $Q_{n+1}(z,\tau)$ will be of the form

$$Q_{n+1}(z,\tau) = \frac{(z-\xi)q_n(z)}{\pi_{n+1}(z)} \quad \text{with} \quad q_n \in \mathcal{P}_n.$$

Define the space of rational functions $\check{\mathcal{L}}_n$ based on the set of poles $\check{\mathbb{A}} = \{\alpha_1, \ldots, \alpha_{n-1}, \alpha_{n+1}\}$ and let $\check{\phi}_n \in \check{\mathcal{L}}_n \setminus \check{\mathcal{L}}_{n-1}$ be the ORF for the measure $\mathrm{d}\check{\mu}(t) = \frac{|t-\xi|^2}{|1-\overline{\alpha}_n t|^2}\mathrm{d}\mu(t)$ which is a rational modification of the measure μ. Then it can be proved that $\check{Q}_n$ defined by

$$\check{Q}_n(z) = \frac{q_n(z)}{\pi_{n-1}(z)(1-\overline{\alpha}_{n+1}z)}$$

is the nth PORF for $\check{\mu}$. Thus there must exist some $\check{\tau} \in \mathbb{T}$ such that $\check{Q}_n(z) = \check{\phi}_n(z) + \check{\tau}\check{\phi}_n^*(z)$. This means that the n free nodes and weights can be computed as for an n-point Szegő QF constructed (with a particular choice of $\check{\tau} \in \mathbb{T}$) exact in $\check{\mathcal{L}}_{n-1,n-1} = \mathcal{L}_{n-1,n-1}$ for a rational modification $\check{\mu}$ of the measure μ.

More details can be found in [122] where also the case of 2 prefixed nodes is considered, which we shall call a rational *Szegő-Lobatto* formula.

11.1.2 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 Verblunsky coefficients $\{\upsilon_k\}_{k=1}^n \subset \mathbb{D}$), but we changed only the last coefficient $\upsilon_{n+1} \in \mathbb{D}$ 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 Verblunsky parameters υ_{n+1} and υ_{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 [349]. The rational case is a relatively simple generalization. See [122]. Let ξ_1 and ξ_2 be the two prefixed nodes, and denote the remaining nodes as $t_j,\ j = 1,\ldots,n$. So the function in $\mathcal{L}_{n+2}$ whose zeros are these nodes has the form $R_{n+2}(z) = \frac{(z-\xi_1)(z-\xi_2)P_n(z)}{\pi_{n+2}(z)}$, $P_n(z) = \prod_{k=1}^n (z-t_j)$. The QF has the form $\mathfrak{I}_{n+2}[f] = A_1 f(\xi_1) + A_2 f(\xi_2) + \sum_{k=1}^n w_j f(t_j)$. The following theorem is proved in [122, Thm. 6.1].

Theorem 11.1 *The QF $\mathfrak{I}_{n+2}$ just introduced will be exact in $\mathcal{L}_{n,n}$ if and only if*

(1) $\mathfrak{I}_{n+2}[f]$ 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 $f \in \mathcal{L}_n$ that satisfy $f(\alpha_{n+1}) = f(\alpha_{n+2}) = 0$.

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 $g = f - L \in \mathcal{L}_{n,n}$ vanishes in these nodes, so that $\mathfrak{I}_\mu[f] = \mathfrak{I}_\mu[g+L] = \mathfrak{I}_\mu[g] + \mathfrak{I}_\mu[L] = I_\mu[g] + \mathfrak{I}_{n+2}[L]$. It suffices to rewrite $\mathfrak{I}_\mu[g]$ as $\langle h, R_{n+2}\rangle$ and use the condition (2) to find that it is 0. □

Remark 11.1 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 $\operatorname{span}\{z, z^2, \ldots, z^{n-1}\}$ 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 given in [97] which we shall currently explain.

Let us see how to construct the formula $\mathfrak{I}_{n+2}$ by manipulating the two Verblunsky parameters υ_{n+1} and υ_{n+2} and change them into $\check{\upsilon}_{n+1}$ and $\check{\upsilon}_{n+2}$. Whatever the $\{\upsilon_1,\ldots,\upsilon_n,\check{\upsilon}_{n+1}\} \subset \mathbb{D}$ are, we need the last one, $\check{\upsilon}_{n+2}$, to be on $\mathbb{T}$ to guarantee that we obtain a rational function whose zeros are simple and on $\mathbb{T}$. So we generate the ORF ϕ_k for $k = 0,1,\ldots,n$ using the Schur parameters υ_k, $k = 1,\ldots,n$. Next we replace υ_{n+1} by $\check{\upsilon}_{n+1} \in \mathbb{D}$ and define

$$\check{\phi}_{n+1}(z) = \frac{1-\overline{\alpha}_n z}{1-\overline{\alpha}_{n+1} z}[\zeta_{n+1}(z)\phi_n(z) + \check{\upsilon}_{n+1}\phi_n^*(z)] \in \mathcal{L}_{n+1}$$

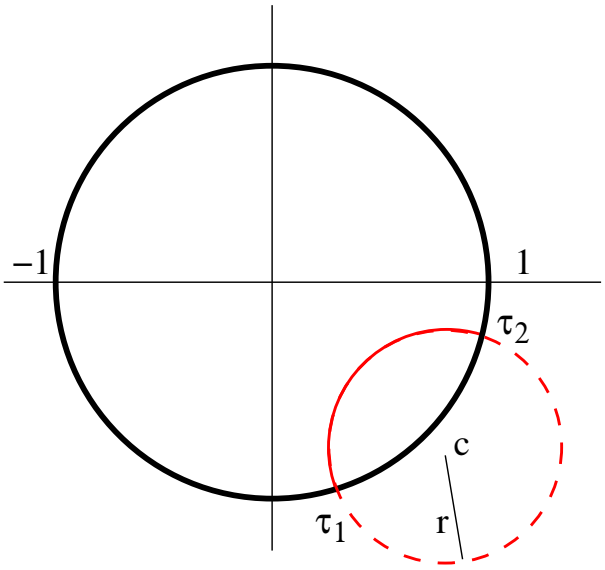


Figure 1: Rational Szegő-Lobatto quadrature. The coefficient $\check{\upsilon}_{n+1}$ lies on the circle C with center c and radius r which is the smaller circle. The large circle is the unit circle $\mathbb{T}$. Only the part of C that is in $\mathbb{D}$ gives a 'good' selection of $\check{\upsilon}_{n+1}$. Note that $C \cap \mathbb{T} = \{\tau_1, \tau_2\}$.

and finally compute the nodes as the zeros of

$$\check{\phi}_{n+2}(z) = \frac{1-\overline{\alpha}_{n+1}z}{1-\overline{\alpha}_{n+2}z}[\zeta_{n+2}(z)\check{\phi}_{n+1}(z) + \check{\upsilon}_{n+2}\check{\phi}^*_{n+1}(z)] \in \mathcal{L}_{n+2}.$$

Choosing $\check{\upsilon}_{n+1} \in \mathbb{D}$ guarantees that $\check{\phi}_{n+1}$ is an ORF with respect to some positive measure $\check{\mu}$ on $\mathbb{T}$. Now we have to choose $\check{\upsilon}_{n+1}$ and $\check{\upsilon}_{n+2}$ such that $\check{\phi}_{n+2}$ has zeros ξ_1 and ξ_2. Define $x_i = \zeta_{n+1}(\xi_i)$ and $\tau_i = \zeta_n(\xi_i)\phi_n(\xi_i)/\phi^*_n(\xi_i)$, $i = 1, 2$ (recall the definition of ζ_n from (7.2)). All these 4 numbers x_i, τ_i, $i = 1, 2$ are on $\mathbb{T}$. The condition $\check{\phi}_{n+2}(\xi_i) = 0$ corresponds to $\overline{\check{\upsilon}}_{n+1}\check{\upsilon}_{n+2} - \tau_i\xi_i\check{\upsilon}_{n+1} - \overline{\tau}_i\check{\upsilon}_{n+2} + \xi_i = 0$, $i = 1, 2$. Set $i = 1$ and solve for $\check{\upsilon}_{n+2}$ to get

$$\check{\upsilon}_{n+2} = \xi_i\frac{\tau_1 - \check{\upsilon}_{n+1}}{1 - \tau_1\overline{\check{\upsilon}}_{n+1}} = \xi_i\mathfrak{m}^{-1}_{\check{\upsilon}_{n+1}}[\tau_1].$$

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

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

or, if $\overline{\tau}_1\xi_1 \neq \overline{\tau}_2\xi_2$,

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

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

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

which is a circle C with center c and radius r. The values τ_1 and τ_2 are the intersections of the circle C and $\mathbb{T}$. Again, since we need $\check{\upsilon}_{n+1}$ to be in $\mathbb{D}$, we can only use the arc of C

between τ_1 and τ_2 that lies inside $\mathbb{D}$, i.e., $C \cap \mathbb{D}$.
If $\overline{\tau}_1\xi_1 = \overline{\tau}_2\xi_2$ the circle becomes a straight line (circle with infinite radius) and the conclusion is the same.
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 $\check{\upsilon}_{n+1} = \tau_1$, then $\check{\phi}_{n+1}$ is a PORF for the measure μ and because of the definition of τ_1 it means that ξ_1 and ξ_2 are both zeros of $\check{\phi}_{n+1}$. We have an $(n+1)$-point rational Szegő QF exact in $\mathcal{L}_{n,n}$ because it so happened that by choosing a τ that places one of its nodes at ξ_1 accidentally made ξ_2 coincide with another of its nodes.

11.1.3 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 circular arc in $\mathbb{D}$.

It is a natural step to ask 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 ξ_1, ξ_2 and ξ_3, then the previous arguments will fix ξ_1 and ξ_2 with a parameter on an arc in $\mathbb{D}$. Using the same arguments, fixing ξ_1 and ξ_3 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.

11.2 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 one prescribes m nodes and uses 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 other 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 QOP 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 can be found in [50] and in [79, 80]. In [79] a modification of the computational algorithm via an eigenvalue problem can be found. This works for any finite, half infinite or infinite interval. But that is for the polynomial case. The rational analog is however more complicated.

11.2.1 Rational Gauss-Radau QF

Let us first suppose that the poles α_k are all in $\hat{\mathbb{R}}^I = \hat{\mathbb{R}} \setminus I$ with $I = [-1, 1]$. If we fix just one node ξ 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 ξ 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

[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.

consider a QF on $\mathbb{R}$, then we still need to ensure that none of the zeros coincides with one of the poles (degeneracy). If the real poles are outside I, i.e., $\{\alpha_k\} \subset \hat{\mathbb{R}}^I$, then all zeros are simple and in $(-1,1)$ but otherwise, it is only guaranteed that at least $n-1$ zeros will be in $(-1,1)$ [532, 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 [198], 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 generalized eigenvalue problem is proposed as a method to compute the nodes.

A study of ORFs on the positive real line and poles on $[-\infty, 0]$ was undertaken in [114]. The corresponding quadrature formulas based on QORFs, hence including the rational Gauss-Radau case, was given in [120]. 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 τ for which all the zeros are in $(0,\infty)$ are characterized.

We can sketch the idea for the interval $I = [-1,1]$ as follows (see Figure 2). Suppose the poles α_k are all real and outside the support of the measure μ but similar observations hold for general complex poles outside I. The QORF $Q_n(z,\tau) = \phi_n(x) + \tau\frac{Z_n(x)}{Z_{n-1}(x)}\phi_{n-1}(x)$ will have n zeros that depend on τ. This means that if $x = \xi$ has to be a zero, then the value $\tau = \tau(\xi)$ that we have to choose will be obtained by the function

$$\tau(x) = -\frac{\phi_n(x)Z_{n-1}(x)}{\phi_{n-1}(x)Z_n(x)} = -\frac{\tilde{\phi}_n(x)}{\tilde{\phi}_{n-1}(x)} = -\frac{p_n(x)}{p_{n-1}(x)(1-\frac{x}{\alpha_{n-1}})},$$

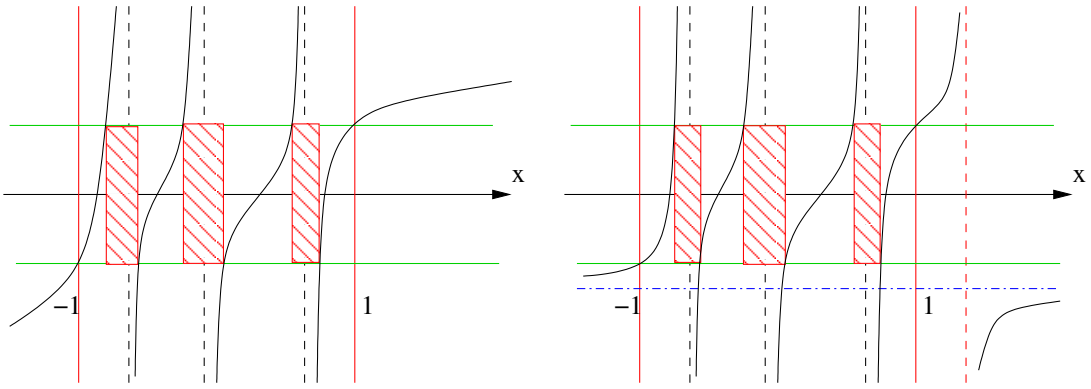
where as before

$$\tilde{\phi}_k(x) = (1-\frac{x}{\alpha_k})\phi_k(x) = \frac{(1-\frac{x}{\alpha_k})p_k(x)}{\pi_k(x)},$$

for n different values of x. The poles of $\tau(x)$ are the zeros of $p_{n-1}(x)(1-x/\alpha_{n-1})$ and the zeros of the numerator $p_n(x)$ interlace. The function $\tau(x)$ is monotone in between the poles and goes from $-\infty$ to $+\infty$ (or vice versa, depending on normalization). A sketch of this graph then easily shows for what values of τ, all the zeros will be in the support of the measure. One should have $\tau(-1) \le \tau(x) \le \tau(1)$. For $\tau = \tau(-1)$, the leftmost of the zeros of $Q_n(x,\tau)$ will be in -1, the others are in $(-1,1)$. This corresponds a left Radau QF where the left endpoint -1 of the interval is prefixed. For $\tau = \tau(1)$, the rightmost of the zeros of $Q_n(x,\tau)$ will be in 1, the others are in $(-1,1)$. This corresponds a right Radau QF where the right endpoint 1 of the interval is prefixed. The interval $[-1,1]$ can be partitioned in a succession of 'good' and 'bad' intervals. If we select our prefixed node in a 'good' interval, all the other nodes will be inside $[-1,1]$ and the weights of the QF will be positive. If the prefixed node is in a 'bad' interval, one of the other nodes will be forced to be outside $[-1,1]$. The 'good' intervals lie around the zeros of $\tau(x)$, i.e., around the zeros of $p_n(x)$, which are the Gauss nodes. If x_k^G are the Gauss nodes, x_k^{LR} the left Radau nodes and x_k^{RR} the right Radau nodes, then

$$-1 = x_1^{LR} < x_1^G < x_1^{RR} < x_2^{LR} < x_2^G < x_2^{RR} < \cdots < x_n^{LR} < x_n^G < x_n^{RR} = 1$$

and the 'good' intervals are $[x_k^{LR}, x_k^{RR}]$ with the 'bad' ones in between.



$n = 3,\ \alpha_n = \infty$ $\qquad$ $n = 3,\ \alpha_n \neq \infty$

Figure 2: The curve is $\tau(x) = -\tilde{\phi}_n(x)/\tilde{\phi}_{n-1}(x)$. The vertical asymptotes correspond to the zeros of p_{n-1}. Choose the prefixed node $x = \xi$ on the x-axis and find the point $P = (\xi, \tau(\xi))$. Though P draw a horizontal line, which intersects the curve in n points which are the zeros of $Q_n(x, \tau(\xi))$. If ξ is too close to a vertical asymptote (shaded region), one node will be outside the interval I. The 'good' regions are of the form $[x_k^{LR}, x_k^{RR}]$. They contain the Gauss nodes x_k^G as the intersections of the curve and the x-axis.

11.2.2 Rational Gauss-Lobatto QF

For the special case of the Chebyshev measure $\mathrm{d}\mu(x) = (1 - x^2)^{-1/2}\mathrm{d}\lambda(x)$ on $[-1, 1]$, a rational Gauss-Lobatto QF was considered in [421] when the α_k are complex conjugate pairs in $\mathbb{C} \setminus [-1, 1]$.

A general study of the rational Gauss-Lobatto QFs involves the existence and computation of rational Gauss QFs with respect to a measure supported on a real interval and with two prescribed nodes (not necessarily on the boundary of the support) and a maximal domain of validity was undertaken in [195] where the poles could be anywhere in $\hat{\mathbb{C}}^I = \hat{\mathbb{C}} \setminus [-1, 1]$.

To describe the 'good' and the 'bad' intervals in this case we need some extra notation denoting nodes of classical rational Gauss, (left and right) Gauss-Lobatto, and classical Gauss-Lobatto formulas both for n and $n - 1$ point formulas.

By $\{x_{nk}^G\}_{k=1}^n$ we denote the n-point rational Gauss (G) nodes, $\{x_{nk}^{RL}\}_{k=1}^n$ are the n-point right rational Gauss-Lobatto (RL) nodes (i.e., $x_{nn}^{RL} = 1$ and another node fixed in $(-1, 1)$), $\{x_{nk}^{LL}\}_{k=1}^n$ are the n-point left rational Gauss-Lobatto (LL) nodes (i.e., $x_{n1}^{LL} = -1$ and another node fixed in $(-1, 1)$), and $\{x_{nk}^{CL}\}_{k=1}^n$ are the n-point classical rational Gauss-Lobatto (CL) nodes (i.e., $x_{n1}^{CL} = -1$ and $x_{nn}^{CL} = 1$ are fixed). The more general rational Gauss-Lobatto nodes $\{x_{nk}^L\}_{k=1}^n$ allow two fixed nodes in $(-1, 1)$. It holds that

$$-1 = x_{n1}^{CL} < x_{n1}^G < \cdots < x_{nk}^{CL} < x_{n-1,k}^G < \cdots < x_{n,n-1}^{CL} < x_{n-1,n-1}^G < x_{nn}^{CL} = 1.$$

Left and right rational Gauss-Lobatto QF are complementary in the following sense

Theorem 11.2 *Suppose we fix $\xi \in (-1, 1)$ then we have the following possibilities:*

- $x_{nk}^{CL} < \xi < x_{n-1,k}^G$ *then there exists an n-point RL QF.*
- $x_{n-1,k-1}^G < \xi < x_{n,k}^{CL}$ *then there exists an n-point LL QF.*
- $\xi = x_{n,k}^{CL}$ *then we have the n-point CL QF.*

- $\xi = x^G_{n-1,k}$ *then we have the* $(n-1)$*-point G QF.*

A general rational Gauss-Lobatto formula with two internal nodes ξ and $\check{\xi}$ fixed exists when either

$$x^{CL}_{nk} < \xi < x^G_{n-1,k} \quad \text{and} \quad \check{x}_{nl}(\xi)^G < \check{\xi} \leq \hat{x}^{RL}_{nl}(\xi)$$

or

$$x^G_{n-1,k-1} < \xi < x^{CL}_{n,k} \quad \text{and} \quad \hat{x}^{LL}_{nl}(\xi) \leq \check{\xi} < \check{x}^G_{nl}(\xi),$$

where $\{\check{x}^G_{nk}(\xi)\} = \{x^G_{nk}\} \cup \{\xi\}$ and $\{\hat{x}^{LL}_{nk}(\xi)\}$ are the nodes of an LL QF with ξ fixed and $\{\hat{x}^{RL}_{nk}(\xi)\}$ are the nodes of an RL QF with ξ fixed.

In all these cases, the computations can be done by solving a generalized eigenvalue problem with tridiagonal matrices.

The nodes are the zeros of a QORF $Q_n(x)$ of second order, i.e., it should be orthogonal to $\mathcal{L}_{n-2}(\alpha_n)$ (recall that a (first order) QORF was orthogonal to $\mathcal{L}_{n-1}(\alpha_n)$). To guarantee existence, one possibly needs to replace α_n by a different value. The QFs fixing two nodes are exact in $\mathcal{L}_{n-1} \cdot \mathcal{L}_{n-2}$. One might think of deriving similar formulas exact in $\mathcal{L}_n \cdot \mathcal{L}_{n-3}$.

11.2.3 Relating QF on $\mathbb{T}$ and I with prescribed nodes

By the Joukowski transform mapping the circle $\mathbb{T}$ to the interval $I = [-1, 1]$, it is clear that one may expect a relation between the Szegő QF with prescribed nodes on the circle and the Gauss QF with prescribed nodes on the interval.

Some of these relations exist in the literature. They result from the relation between a QF on the circle and its 'projection' on the interval $I = [-1, 1]$. This has been investigated, for the polynomial case [81–83, 176] as well as for the rational case [196, 197].

The idea is rather simple. Consider 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 they are real. Then also the zeros $t_k \in \mathbb{T}$ of a QORF $\mathring{Q}_n(z, \mathring{\tau})$ will be symmetric with respect to $\mathbb{R}$ (and on $\mathbb{T}$) if $\mathring{\tau} = \pm 1$ and the weights for symmetric nodes are equal. In particular $t = 1$ will be a zero of $\mathring{Q}_n(z, \mathring{\tau})$ if $\mathring{\tau} = -1$ and $t = -1$ is a zero if $\mathring{\tau} = (-1)^{n+1}$.

For simplicity assume that $\mathring{\mu}$ is absolutely continuous, then on $I = [-1, 1]$ we associate the 'projected' measure μ as follows: $\mathring{\mu}'(\mathrm{e}^{\mathrm{i}\theta}) = |\sin(\theta)|\mu'(\cos\theta)$. The poles are given as $\beta_k = \mathsf{j}[\alpha_k]$ (as before j denotes the Joukowski transform). We can project the nodes of a $2n$-point rational Szegő formula from $\mathbb{T}$ on the interval $I = [-1, 1]$, i.e., if $t_k = \mathrm{e}^{\mathrm{i}\theta_k} \in \mathbb{T}$ is a node for the Szegő QF for $\mathring{\mu}$, then $x_k = \cos\theta_k$ will be a node for the corresponding rational n-point Gauss QF for μ and the weights also match. This result holds in both directions [196, Thm. 3.2,3.3].

In the case of a $(2n + 1)$-point rational Szegő QF for $\mathring{\mu}$, then one of the nodes should be $+1$ or -1. The 'projected' formula will then be a rational Gauss-Radau QF for I with measure μ.

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 [196, Thm. 3.4].

See also [197] for a slightly more general discussion.

Another idea from Gauss quadrature that has not been generalized to the rational case is what is proposed by Gautschi in [289] as being generalized Gauss-Radau and Gauss-Lobatto formulas where double nodes are imposed at one or both end points of the interval. Later it

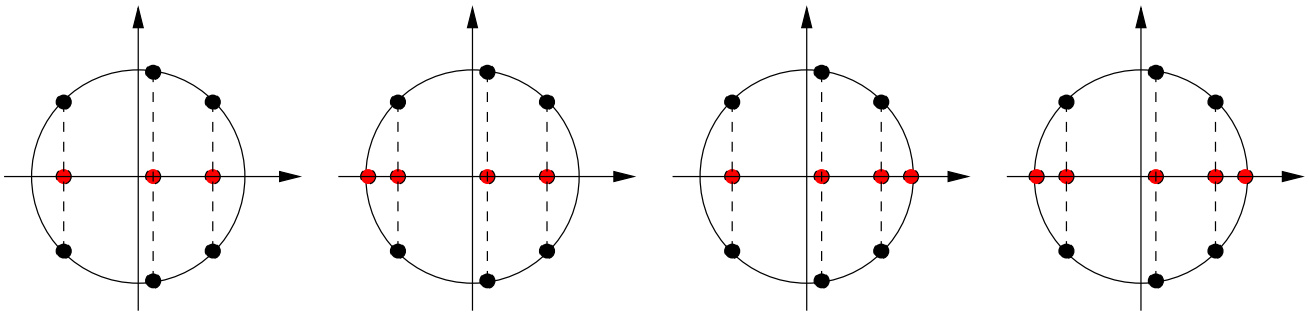


Figure 3: The 'projection' of a symmetric (rational) Szegő quadrature from $\mathbb{T}$ onto the interval $I = [-1, 1]$

was proved in [364] that the weights of such a QF are positive. More generally (Gaussian) QFs with multiple nodes, the so-called *Gauss-Turán* QFs [515], have received considerable attention in the literature. See for an example [304, 379, 416, 417, 473, 495, 505, 549]. The node polynomial $N_n(x)$ is s-orthogonal, meaning that $[N_n(x)]^{2s+1}$ is orthogonal to $\mathcal{P}_{n-1}$ where s is the multiplicity of each of the nodes. This n-point QF is exact in $\mathcal{P}_{2(s+1)n-1}$ but positive weights are not guaranteed. The group in Niš with Milovanović, Spalević and others is continuing this track of research and expanding the theory and constructing error bounds in the polynomial case. Rational analogs in the sense of this paper have not been considered so far.

11.3 Rational Szegő QF: maximal domain of exactness revisited

The n-point rational Gauss QF was exact in $\mathcal{L}_n \cdot \mathcal{L}_{n-1}$, a space of dimension $2n$, but there was no freedom left. It was unique. If we wanted a free parameter, we had to use the zeros of QORFs introducing a parameter $\tau \in \mathbb{R}$, but reducing the domain of exactness to $\mathcal{L}_{n-1} \cdot \mathcal{L}_{n-1}$, which has dimension $2n - 1$.

The n-point rational Szegő QF was exact in $\mathcal{L}_{(n-1)*} \cdot \mathcal{L}_{n-1}$, a space of dimension $2n - 1$, but there were infinitely many solutions because of the free parameter $\tau \in \mathbb{T}$ of the PORF. Hence one may wonder whether also the rational Szegő QF cannot be exact in a space of dimension $2n$ by fixing the parameter $\tau \in \mathbb{T}$. That space cannot be of the form $\mathcal{L}_{n*} \cdot \mathcal{L}_{n-1}$ or $\mathcal{L}_{(n-1)*} \cdot \mathcal{L}_n$, since it was proved that $\mathcal{L}_{(n-1)*} \cdot \mathcal{L}_{n-1}$ was the maximal one of this form. And indeed, there is a space of dimension $2n$ in which it is exact. This was first stated in the polynomial case [450] and for the rational case in [122].

We know that the partial Blaschke products $\{B_k\}_{k=0}^n$ form a basis for $\mathcal{L}_n$ while $\{B_{-k}\}_{k=1}^n$ where $B_{-k} = B_{k*} = 1/B_k$ is a basis for $\mathcal{L}_{n*}$. So for $\mathcal{L}_{n,n} = \mathcal{L}_{n*} \cdot \mathcal{L}_n$ we may introduce a change of basis: set $\rho_0 = B_0 = 1$ and

$$\begin{bmatrix} \rho_k \\ \rho_{-k} \end{bmatrix} = \begin{bmatrix} \overline{s}_k & s_k \\ \overline{s}_k & -s_k \end{bmatrix} \begin{bmatrix} B_k \\ B_{-k} \end{bmatrix}, \quad s_k \in \mathbb{T}, \quad k = 1, 2, \ldots.$$

Then consider two different PORFs $Q_n^{\pm}(z) = \phi_n(z) \pm \tau_n \phi_n^*(z)$ and use their zeros as nodes

to construct two rational Szegő QFs $\mathfrak{I}_n^{\pm}$. Both of these will be exact in $\mathcal{L}_{n-1,n-1} = \mathcal{L}_{(n-1)*} + \mathcal{L}_{n-1} = \text{span}\{\rho_{-(n-1)}, \ldots, \rho_{n-1}\}$. However if we define the spaces $\mathcal{M}_n^{+} = \text{span}\{\rho_{-(n-1)}, \ldots, \rho_n\}$ resp. $\mathcal{M}_n^{-} = \text{span}\{\rho_{-n}, \ldots, \rho_{n-1}\}$, each of these with dimension $2n$ and $\mathcal{L}_{n-1,n-1} \subset \mathcal{M}_n^{\pm} \subset \mathcal{L}_{n,n}$. Then the following can be shown [122].

Theorem 11.3 *Let $\mathfrak{I}_n^{\pm}$ be the rational Szegő QF whose nodes are the zeros of the PORF $Q_n^{\pm}(z) = \phi_n(z) \pm \tau_n \phi_n^*(z)$, $\tau_n \in \mathbb{T}$. Then $\mathfrak{I}_n^{\pm}$ will be exact in the spaces $\mathcal{M}_n^{\pm}$ of dimension $2n$ defined above on condition that $s_n = Q_n^{\pm}(\alpha_n)/|Q_n^{\pm}(\alpha_n)|$ is chosen in the definition of the ρ_k.*

12 Inverse eigenvalue problems

As we have seen, the computation of an optimal QF can be reduced to the solution of an eigenvalue problem. To this eigenvalue problem, there is also an inverse that relates to discrete least squares approximation. These problems have a long tradition. See for example [164]. The idea is to recover properties of a dynamical system, given its spectral data. In the case of an interval on $\mathbb{R}$, the problem is whether it is possible to recover the Jacobi matrix, given the nodes and weights of the quadrature formula. In the circle case, one has to recover a Hessenberg matrix.

12.1 The polynomial case

Suppose there is a function f that we can observe as $f(z_k)$ in a set of points $\{z_k\}$ with in each point a weight $\omega_k > 0$ reflecting the confidence that we have in the observation. The set of interpolation data; (z_k, f_k, ω_k), $k = 0, 1, \ldots$ comes available while we are computing. The idea is that for an (extended) Chebyshev system $\{u_0, u_1, \ldots\}$, we want to find the best approximant $g_n \in \mathcal{L}_n = \text{span}\{u_0, \ldots, u_n\}$ that minimizes

$$\|f - g_n\|^2 = \sum_{k=0}^{n} w_k |f(z_k) - g_n(z_k)|^2, \quad w_k = \omega_k^2.$$

That is a (discrete) least squares approximation in the Euclidean space $\mathcal{L}_n$ with inner product

$$\langle f, g \rangle = \sum_{k=0}^{N} w_k \overline{f(z_k)} g(z_k).$$

The approximation should be recursive, i.e., if g_n is not good enough, we add an extra triple $(z_{n+1}, f_{n+1}, \omega_{n+1})$ and compute the next approximant as $g_{n+1} = g_n + h_{n+1}$ with $h_{n+1} \in \mathcal{L}_{n+1} \setminus \mathcal{L}_n$. This means that at the moment that we compute g_n, only the first $n + 1$ data triples $(x_k, f_k, \omega_k)_{k=0}^n$ are available, so that we approximate the inner product with

$$\langle f, g \rangle = \sum_{k=0}^{n} w_k \overline{f(z_k)} g(z_k).$$

It is best to express g_n as a linear combination of the orthogonal basis $\{\phi_k\}_{k=0}^n$ for $\mathcal{L}_n$, thus $g_n(z) = \sum_{j=0}^n a_j \phi_j(z)$. Obviously, one can then find the coefficients by evaluating a QF $a_j = \langle \phi_j, f \rangle = \sum_{k=0}^n w_k f(z_k) \overline{\phi_j(z_k)}$, $j = 0, \ldots, n$, and $g_{n+1}(z) = g_n(z) + a_{n+1}\phi_{n+1}(z)$. Thus here we already know the nodes and weights of the quadrature, but we do not know the orthogonal basis functions $\{\phi_j\}$. Note that the inner product, and thus also the orthogonal basis changes with n. The problem is reduced to finding and updating the orthogonal basis, or even better the values of $\phi_k(z_j)$, $j, k = 0, \ldots, n$ and to update them as n increases by 1.

There is a vast literature for the construction of recurrence relations for such OPs. In general this procedure requires $O(n^3)$ computations, but in the special case where the points are on $\mathbb{R}$ or on $\mathbb{T}$, this reduces to $O(n^2)$.

This is an *inverse eigenvalue problem*. Indeed, we know the discrete measure (i.e., weights w_k and nodes z_k of a QF) and the problem is to find the matrix S_n characterizing the recurrence (i.e., in the polynomial case the Jacobi or Hessenberg matrix representing the

shift operator in $\mathcal{L}_n = \mathcal{P}_n$). So we know the eigenvalues and the first components of the normalized eigenvectors and we have to find the matrix.

Define $\mathbf{q}_{n,j} = D_n(\mathbf{w})[\phi_j(z_0), \ldots, \phi_j(z_n)]^T$ with $D_n(\mathbf{w}) = \operatorname{diag}(\omega_0, \ldots, \omega_n)$, then $\mathbf{q}_{n,i}^H \mathbf{q}_{n,j} = \delta_{ij}$. Thus, the matrix $Q_n = [\mathbf{q}_{n,0} \cdots \mathbf{q}_{n,n}]$ is unitary[10].

For example in the case of $\mathbb{R}$, the data points z_k are all on $\mathbb{R}$ and the inverse eigenvalue problem is to find a tridiagonal Jacobi matrix S_n and a unitary matrix Q_n such that $D_n(\mathbf{z})Q_n = Q_n S_n$ where $D_n(\mathbf{z}) = \operatorname{diag}(z_0, \ldots, z_n)$ with $Q_n^H \mathbf{w}_n = \|\mathbf{w}_n\|\mathbf{e}_0$ where $\mathbf{w}_n = [\omega_0, \ldots, \omega_n]^T$ and $\mathbf{e}_0 = [1, 0, \ldots, 0]^T$. In the case of $\mathbb{T}$, the problem is similar, except that the z_k are all on $\mathbb{T}$ and S_n is a unitary Hessenberg matrix.

Suppose we kick start with the data $\mathbf{w}_n$ and $D_n(\mathbf{z})$ and arrange this as

$$M_n = [\mathbf{w}_n | D_n(\mathbf{z})] = \left[\begin{array}{c|ccc} \omega_0 & z_0 & & \\ \vdots & & \ddots & \\ \omega_n & & & z_n \end{array}\right] \in \mathbb{C}^{(n+1)\times(n+2)}.$$

By orthogonal similarity transformations Q_n, e.g. by a sequence of Givens transformations, this is brought into the form

$$Q_n^H M_n \left[\begin{array}{c|c} 1 & \\ \hline & Q_n \end{array}\right] = \left[\begin{array}{c|cccc} \|\mathbf{w}_n\| & * & \cdots & \cdots & * & * \\ 0 & * & \cdots & \cdots & * & * \\ \vdots & & * & \cdots & * & * \\ \vdots & & & \ddots & \vdots & \vdots \\ 0 & & & & * & * \end{array}\right] = \left[\begin{array}{c|c} \|\mathbf{w}_n\|\mathbf{e}_0 & H_n \end{array}\right]$$

The matrix H_n is the unitary Hessenberg matrix representing the shift in the case of $\mathbb{T}$ (since the z_k are of modulus 1, and $H_n = Q_n^H D_n(\mathbf{z}) Q_n$), and it reduces for symmetry reasons to a tridiagonal Jacobi matrix in the case of $\mathbb{R}$ (the z_k are real and hence H_n is Hermitian, thus tridiagonal). The matrix Q_n is the one defined above. The coefficient vector $\mathbf{a}_n$ of the expansion $g_n = \sum a_n \phi_k = [\phi_0, \ldots, \phi_n]\mathbf{a}_n$ is given by the product $\mathbf{a}_n = Q_n \mathbf{f}_n$ with $\mathbf{f}_n = D_n(\mathbf{w})[f(z_0), \ldots, f(z_n)]^H$. Adding a new measurement extends the matrix with one row and one column, placing a nonzero element ω_{n+1} in the first column and an entry z_{n+1} on the diagonal. Again a new sequence of Givens transforms can bring this in the same upper trapezoidal form. So we can start with $n = 1$ and extend the matrices to $n = 2, 3, \ldots$.

Efficient algorithms exist to solve the problem recursively for successive n. These are the algorithms we have mentioned earlier in the polynomial case [18, 20, 21, 62, 68, 164, 165, 236, 237, 244, 305, 306, 312, 321, 457, 519–522, 524]. More general problems and algorithms for the inverse problem when the eigenvalues of two submatrices are known and the whole Jacobi matrix has to be reconstructed etc. see for example [45–49, 166, 351, 393, 394, 496, 564].

12.2 The rational case

For the rational discrete least squares problem, the problem can also be linearized so that the numerator and denominator polynomials are stored in a polynomial vector and similar,

[10]Note that the rows of Q_n are the normalized eigenvectors $\boldsymbol{\phi}_{n+1}(z_k) = [\phi_0(z_k), \ldots, \phi_n(z_k)]$ of the Golub-Welsch algorithm or its circle analog if the z_k are the eigenvalues.

more generalized algorithms apply. See [132, 135, 208, 523]. There are some applications in system identification in the frequency domain [133, 134, 482, 483].

These methods can be easily extended to Laurent polynomials, multivariate polynomials, or polyanalytic polynomials, i.e., complex polynomials in z and $\overline{z}$. However, all these approaches are essentially working in a (vector) polynomial context and are not really rational.

For the rational case, when the orthogonal polynomials are replaced by orthogonal rational functions, the representation of the shift operator as an S matrix is more complicated and the algorithm can be formulated in terms of a semi-separable matrix plus a diagonal [137, 525, 526].

Suppose for simplicity that all the poles are simple and we orthogonalize the basis $\{1, (z-\alpha_1)^{-1}, \ldots, (z-\alpha_n)^{-1}\}$. The ORF are of the form $\phi_k = p_k/\pi_k$ with $p_k \in \mathcal{P}_k$ and $\pi_k(x) = \prod_{j=1}^{k}(z-\alpha_j)$, $k \geq 1$. With as before $\boldsymbol{\phi}_{n+1} = [\phi_0, \ldots, \phi_n]$, the recursion reads in $\mathcal{L}_n$

$$\boldsymbol{\phi}_{n+1}(z)(zI_n - D_n(\mathbb{A}))H_n = \boldsymbol{\phi}_{n+1}(z), \quad D_n(\mathbb{A}) = \operatorname{diag}(\alpha_0, \ldots, \alpha_n)$$

with H_n the upper Hessenberg matrix. Set $S_n = H_n^{-1}$, then

$$z\boldsymbol{\phi}_{n+1}(z) = \boldsymbol{\phi}_{n+1}(z)(S_n + D_n(\mathbb{A})).$$

Hence when as before $Q_n = [\omega_i \phi_j(z_i)]_{i,j=0}^{n}$, this is $D_n(\mathbf{z})Q_n^H = Q_n(S_n + D_n(\mathbb{A}))$. The unitary upper Hessenberg matrix had a lot of structure and as we have seen, it can be stored as a sequence of 2×2 elementary Θ_k matrices. Hence its inverse also has a lot of structure and can also be applied and stored as a sequence of elementary matrices. It is a *semiseparable matrix*, i.e., a structured matrix with a special rank structure. Thus the inverse eigenvalue problem is here to find the unitary matrix Q_n and the semiseparable matrix S_n, given $\mathbf{w}_n$, $D_n(\mathbb{A})$, and $D_n(\mathbf{z})$. For the practical algorithm we refer to the literature [243]. But to treat the general case, and/or the case where the z_k consist of reals and complex conjugate pairs (see [137]) are still not completely solved in a satisfactory way and need further attention. For a generalized Hessenberg structure in a general situation see [403].

These spectral problems and more generally problems related to dense matrices but with a special structure so that all its information can be stored in a sparse way is the subject of a flourishing domain in (numerical) linear algebra. Semiseparable matrices as the inverse of an Hessenberg matrix are a special case. The notion of what semiseparable exactly means may differ from publication to publication or from author to author [543]. One may want to consult the recent books [207, 458–460, 527, 542, 544, 545] and the many references they contain.

13 Error bound

There are many ways to find theoretical error bounds for usual Gauss or Szegő QFs and all their variants which hold in general or for particular classes of integrands. Likewise there are many heuristic or practical ways to find the accuracy of the computed quadrature result. To estimate the rational quadrature error $\mathcal{E}_n[f] = \mathfrak{I}_\mu[f] - \mathfrak{I}_n[f]$ and its asymptotic behavior many approaches could be taken as well. Several can be found in [118]. We give some examples. These are only rough expressions and much finer analysis can give better estimates.

13.1 An estimate on $\mathbb{R}$

One possible expression can be found on the basis of an alternative derivation of the Gauss QF with nodes $\{x_1, \ldots, x_n\}$. 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 node point $\{x_1, \ldots, x_n\}$, i.e., it interpolates the function value as well as the derivative. Thus $R_{2n-1}(x_k) = f(x_k)$ and $R'_{2n-1}(x_k) = f'(x_k)$, $k = 1, \ldots, n$. It is clear that it is of the form $R_{2n-1}(x) = P_{2n-1}/\pi_{n,n-1}$, $P_{2n-1} \in \mathcal{P}_{2n-1}$ and $\pi_{n-1,n} = \pi_n \pi_{n-1}$, $\pi_k(x) = \prod_{j=1}^k (1 - x/\alpha_j)$. The numerator satisfies $P_{2n-1}(x) = \sum_{j=1}^n [H_j(x) f(t_j) + \tilde{H}_j(x) f'(x_j)]$ with $H_j(x) = L_j(x)\pi_{n-1,n}(x_j) + \tilde{L}_j(x)\pi'_{n-1,n}(x_j)$ and $\tilde{H}_j(x) = \tilde{L}_j(x)\pi_{n-1,n}(x_j)$ where the Lagrange polynomials are given by $L_j(x_k) = \delta_{jk}$ and $L'_j(x_k) = 0$, while $\tilde{L}_j(x_k) = 0$ and $\tilde{L}'_j(x_k) = \delta_{jk}$. The corresponding interpolatory QF is therefore $\mathfrak{I}_n[f] = \sum_{j=1}^n [w_j f(x_j) + \tilde{w}_j f'(x_j)]$ with $w_j = \mathfrak{I}_\mu[H_j/\pi_{n-1,n}]$ and $\tilde{w}_j = \mathfrak{I}_\mu[\tilde{H}_j/\pi_{n-1,n}]$. If we take as interpolation nodes x_j the zeros of the nth ORF ϕ_n, then if follows from orthogonality that $\tilde{w}_j = 0$ 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

$$\mathcal{E}_n[f] = \mathfrak{I}_\mu[f - R_{2n-1}] = \frac{(f\pi_{n-1,n})^{(2n)}(\theta(x))}{(2n)!} \int \phi_n^2(x)(1 - x/\alpha_n)\mathrm{d}\mu(x).$$

For a measure supported on $[0, +\infty)$ and poles satisfying the condition that makes the Blaschke product diverge, and hence that guarantees denseness of the rationals given in the Appendix 14.2.1, it follows that the QF will converge for any integrand f continuous on $[0, \infty)$ and for which the limit as $x \to +\infty$ is finite.

As a consequence one may prove [111, Thm. 5.3]

Theorem 13.1 *Under the conditions listed above and with f analytic in a connected domain, containing $[0, +\infty)$ whose boundary is a rectifiable curve Γ, then*

$$|\mathcal{E}_n[f]| \leq \frac{c(\Gamma, f)}{\inf_{z\in\Gamma} |(1 - z/\alpha_n)\phi_n^2(z)|} \int_0^\infty (1 - x/\alpha_n)\, \phi_n^2(x)\mathrm{d}\mu(x)$$

$c(\Gamma, f)$ is a constant depending on Γ 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.

13.2 An estimate on $\mathbb{T}$

Arguments along the same lines as in the previous section can be derived for estimates of the integration error for rational Szegő QF on $\mathbb{T}$. There will be an extra term showing up. We have from [104, 118] the following. If f is analytic in a region G containing $\mathbb{T}$ but excluding the points α_k and their reflections $1/\overline{\alpha}_k$, then the Szegő QF $\mathfrak{I}_n$ based on the zeros of the PORF $Q_n(z,\tau)$ satisfies

$$\mathcal{E}_n[f] = [\mathfrak{I}_\mu - \mathfrak{I}_n][f] = \frac{1}{2\pi \mathrm{i}} \int_\Gamma [\Omega_\mu(z) - \Omega_n(z)] \frac{-f(z)}{2z} \mathrm{d}z,$$

with $\Gamma = \partial G$ and

$$\Omega_\mu(z) - \Omega_n(z) = \frac{2zB_{n-1}(z)}{X_n^2(z)} \left(\mathfrak{I}_\mu \left[\frac{X_n^2(\cdot)}{(\cdot - z)B_{n-1}(\cdot)} \right] + D_n \right),$$

$X_n = Q_n(z,\tau)(1 - \overline{\alpha}_n z)$ and $D_n = \mathfrak{I}_\mu[X_n]$. The last term D_n is caused by the para-orthogonality of Q_n, lacking one orthogonality condition. Hence we may write

Theorem 13.2 *Let f be analytic in an region G containing $\mathbb{T}$ but none of the points in $\mathbb{A}_n \cup \hat{\mathbb{A}}_n$ where $\mathbb{A}_n = \{\alpha_k : k = 1, \ldots, n\}$ and $\hat{\mathbb{A}}_n = \{\hat{\alpha}_k = 1/\overline{\alpha}_k : k = 1, \ldots, n\}$. Let $\mathfrak{I}_n[f] = \sum_{k=1}^n w_k f(t_k) = \int f(t) \mathrm{d}\mu_n(t)$ be an n-point rational Szegő QF exact in $\mathcal{L}_{n-1,n-1}$, and define with $D(t,z) = \frac{t+z}{t-z}$,*

$$\Omega_n(z) = \sum_{k=1}^n w_k D(t_k, z) = \int D(t,z) \mathrm{d}\mu_n(t) \quad \textit{and} \quad \Omega_\mu(z) = \int D(t,z) \mathrm{d}\mu(t),$$

then

$$|\mathcal{E}_n[f]| = |(\mathfrak{I}_\mu - \mathfrak{I}_n)[f]| \leq \frac{1}{4\pi} \sup_{z \in G} \frac{|f(z)|}{|z|} \int_\Gamma |\Omega_\mu(z) - \Omega_n(z)| |\mathrm{d}z|$$

and

$$|\Omega_\mu(z) - \Omega_n(z)| \leq \frac{2|zB_{n-1}(z)|}{|X_n(z)|^2} M_n \left(M_n \int_{\mathbb{T}} \frac{\mathrm{d}\mu(t)}{|t - z|} + 1 \right)$$

where $\Gamma = \partial G$, $X_n(z) = Q_n(z,\tau)(1 - \overline{\alpha}_n z)$, $M_n = \sup_{t \in \mathbb{T}} |X_n(t)|$, and B_{n-1} is the familiar Blaschke product.

It clearly explains how the interpolation error $\Omega_\mu - \Omega_n$ can be related to the error of the quadrature formula.

13.3 A general estimate on $\mathbb{T}$ and $[-1, 1]$

A general result can be found in the work of de la Calle Ysern [145, 147]. The main result for $\mathbb{T}$ is as follows [145, Thm. 1]. Let $\mathfrak{I}_\rho[f] = \int_{\mathbb{T}} f(\mathrm{e}^{\mathrm{i}\theta}) \rho(\theta) \mathrm{d}\theta$ with ρ possibly a *complex* measure, $\pi_n(z) = \prod_{k=1}^n (1 - \overline{\alpha}_k z)$, $\mathfrak{I}_n[f] = \sum_{k=1}^n A_k f(t_k)$ is a rational QF for $\mathfrak{I}_\rho$ exact in $\mathcal{L}_{p,q}$, $p, q < n$, $p + q = m \geq n - 1$, and $\mathcal{E}_n = \mathfrak{I}_\rho - \mathfrak{I}_n$. 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 13.3 *With the notation just introduced, the quadrature error can be bounded by*

$$|\mathcal{E}_n[f]| \le \|\pi_{p*}\pi_q f\|_{T_r \cup T_R}(\|\rho\| + \|\mathfrak{I}_n\|) \times \left(\frac{r^{p+1}}{1-r^2}\sqrt{g_p(r)g_q(r)} + \frac{R^{1-q}}{R^2-1}\sqrt{g_p(1/R)g_q(1/R)} \right)$$

where

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

$\|\rho\| = \int |\rho(\theta)|\mathrm{d}\theta$ *and* $\|\mathfrak{I}_n\| = \sum_{k=1}^n |A_k|$ *is the Lebesgue constant.*

Although we shall discuss some ideas about convergence later in Section 14, the following is a direct consequence of the previous theorem [147].

Corollary 13.4 *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* $\mathfrak{I}_n$ *an interpolatory QF for* $\mathfrak{I}_\rho$ *with weights* A_j *whose nodes* z_j *are the zeros of the PORF w.r.t. the weight* w

$$\lim_{n\to\infty} \mathfrak{I}_n[f] = \int_0^{2\pi} f(\mathrm{e}^{\mathrm{i}\theta})\rho(\theta)\mathrm{d}\theta$$

and

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

Similar results can be obtained for integrals over $I = [-1, 1]$. The circle T_r is replaced by its inverse Joukowski map [145, Thm. 1]. Now $\pi_n(x) = \prod_{k=1}^n (1 - x/\beta_k)$ (the poles are β_k), set $\alpha_k = \mathfrak{j}^{inv}(\beta_k)$ 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 $T_r = \{z \in \mathbb{C} : |\mathfrak{j}^{inv}(z)| = r\}$ with $r < 1$ such that $T_r \subset U$.

Theorem 13.5 *Let* f *be analytic in a neighborhood of* I *and* $\mathfrak{I}_n$ *is an* n*-point QF exact in* $\mathcal{L}_{m+1}$ *with* $n \le m+1 \le 2n$. *Then the quadrature error is bounded by*

$$|\mathcal{E}_n[f]| \le \|\pi_m f\|_{T_r} \frac{2r^{m+1}}{1-r^2}(\|\mu\| + \|\mathfrak{I}_n\|) \prod_{i=1}^m |1 + \alpha_i^2| \times g(r)$$

with

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

Corollary 13.6 *If* $\sum_k (1 - |\alpha_k|) = \infty$ *then the above interpolatory QF will converge to the integral of any function* f *continuous on* I *if and only if* $\|\mathfrak{I}_n\| = \sum_k |A_k| \le M$.

13.4 Anti-Gauss and anti-Szegő formulas

Anti-Gauss QF were proposed by Laurie [389]. To estimate the error $\mathcal{E}_n[f] = (\mathfrak{I}_\mu - \mathfrak{I}_n)[f]$, another $(n+1)$-point QF $\hat{\mathfrak{I}}_{n+1}$, the so called anti-Gauss QF, is constructed whose error has the same magnitude but the opposite sign. Thus

$$(\mathfrak{I}_\mu - \hat{\mathfrak{I}}_{n+1})[f] = -(\mathfrak{I}_\mu - \mathfrak{I}_n)[f], \quad \forall \quad f \in \mathcal{P}_{2n+1}$$

from which we get $\mathfrak{I}_\mu = \frac{1}{2}(\mathfrak{I}_n + \hat{\mathfrak{I}}_{n+1})$, a $(2n+1)$-point averaged Gauss QF, exact in $\mathcal{P}_{2n+1}$. Because $\hat{\mathfrak{I}}_{n+1} = 2\mathfrak{I}_\mu - \mathfrak{I}_n$, the QF $\hat{\mathfrak{I}}_{n+1}$ turns out to be a Gauss QF for the measure $\hat{\mu} = 2\mu - \mu_n$. Its nodes are the zeros of the $(n+1)$th OP $\hat{\phi}_{n+1}$ for $\hat{\mu}$. The analysis of the problem shows that the sequence $\{\hat{\phi}_k\}$ is the same as the original set of OP $\{\phi_k\}$ for $k = 0, 1, \ldots, n$ but that $\hat{\phi}_{n+1}$ should be computed with a particular choice of the recursion coefficients. Hence this results in a very efficient computational procedure. If $\{x_k\}_{k=1}^n$ are the nodes of $\mathfrak{I}_n$ and $\{\hat{x}_k\}_{k=1}^{n+1}$ are the nodes of $\hat{\mathfrak{I}}_{n+1}$ then they interlace

$$\hat{x}_1 < x_1 < \hat{x}_2 < x_2 < \cdots < x_n < \hat{x}_{n+1}.$$

and all the weights $\{\hat{w}_k\}_{k=1}^{n+1}$ of the anti-Gauss QF $\hat{\mathfrak{I}}_{n+1}$ are positive. Not all the nodes need be inside the support of the measure though. Suppose $\operatorname{supp}\mu = [a, b]$. Clearly $\{\hat{x}_k\}_{k=2}^n \subset (a, b)$ but for $\hat{x}_1$ and $\hat{x}_{n+1}$ this is not guaranteed in general. Suppose we use the monic OP $\varphi_{n+1}(x) = (x - a_n)\varphi_n(x) - b_{n-1}^2\varphi_{n-1}(x)$ then $\hat{x}_1 \in [a, b]$ if $\frac{\varphi_{n+1}(a)}{\varphi_{n-1}(a)} \geq b_{n-1}^2$ and $\hat{x}_{n+1} \in [a, b]$ if $\frac{\varphi_{n+1}(b)}{\varphi_{n-1}(b)} \geq b_{n-1}^2$.

Inspired by the anti-Gauss formulas, Kim and Reichel [374] proposed an anti-Szegő quadrature to estimate the error of a Szegő QF. The idea is that with a Szegő QF $\mathfrak{I}_n$ one can associate a second (anti-)Szegő QF $\hat{\mathfrak{I}}_n$, each exact in $\mathcal{P}_{-(n-1),n-1}$ such that a particular linear combination will be exact in $\mathcal{P}_{-n,n}$. They consider a slightly more general situation using an extra parameter c and they start out with the condition $\mathfrak{I}_\mu - \hat{\mathfrak{I}}_n = -c(\mathfrak{I}_\mu - \mathfrak{I}_n)$ which should hold in $\mathcal{P}_{-n,n}$. They proved the following.

Theorem 13.7 *Let $\{\upsilon_k\}$ be the Verblunsky coefficients of the measure μ and $\{\phi_k\}$ the OPs. The Szegő QF $\mathfrak{I}_n$ uses the zeros of $z\phi_{n-1} + \tau\phi_{n-1}^*$. Choose*

$$c = \frac{1 - |\upsilon_n|^2}{1 + |\upsilon_n|^2 - 2\mathrm{Re}\,(\upsilon_n\overline{\tau})} > 0$$

and define $\hat{\tau} = \upsilon_n + c(\upsilon_n + \tau) \in \mathbb{T}$. Construct the Szegő QF $\hat{\mathfrak{I}}_n$ whose nodes are the zeros of $z\phi_{n-1} + \hat{\tau}\phi_{n-1}^*$, then $\frac{c}{1+c}\mathfrak{I}_n + \frac{1}{1+c}\hat{\mathfrak{I}}_n$ is exact in $\mathcal{P}_{-n,n}$.

The parameter c can be avoided and chosen equal to 1 if we exploit the freedom of choosing τ in the construction of $\mathfrak{I}_n$. Indeed, if τ is such that $|2\upsilon_n - \tau| = 1$ then choosing $\hat{\tau} = 2\upsilon_n - \tau$ will result in $\frac{1}{2}(\mathfrak{I}_n + \hat{\mathfrak{I}}_n)$ being exact in $\mathcal{P}_{-n,n}$. If τ is fixed and cannot be chosen freely, and if we insist on taking $c = 1$, then $\hat{\mathfrak{I}}_n$ has to be replaced by an $(n+1)$-point Szegő QF using as nodes the zeros of $z\phi_n + \tilde{\tau}\phi_n^*$.

At the moment, there is no rational analog yet.

13.5 Kronrod QF

The idea of Kronrod formulas follows approximately the same idea as in the previous section. How to find an error estimate by computing the same integral with another QF that has an higher degree of exactness than the Gauss QF, preferably keeping the extra cost minimal. The latter usually means that one should re-use (part of) the work done for the original QF. If $\mathfrak{I}_n$ is an n-point Gauss QF, then it has order $2n-1$, so that one should add to the n Gauss nodes at least $n+1$ extra points to get an interpolatory QF of order $2n$. The original paper by Kronrod is [381]. Early surveys are given by Gautschi [287] and Notaris [454].

Let $\mathfrak{I}_n[f] = \sum_{k=1}^n w_k f(x_k)$ be a Gauss QF. An optimal extension with r points is a QF of the form

$$\hat{\mathfrak{I}}_{n+r}[f] = \sum_{k=1}^{n} \hat{w}_k f(x_k) + \sum_{k=n+1}^{n+r} \hat{w}_k f(\hat{x}_k)$$

such that this is exact for all $f \in \mathcal{P}_{2n-1+s(r)}$ with $s(r)$ as large as possible. If $\omega_n(x) = \prod_{k=1}^n (x-x_k)$ is the node polynomial for $\mathfrak{I}_n$, then $\hat{\omega}_r(x) = \prod_{k=n+1}^{n+r}(x-\hat{x}_k)$ should be a Stieltjes polynomial, i.e., it should satisfy the orthogonality condition $\int \omega_n(x)\hat{\omega}_r(x)x^s \mathrm{d}\mu(x) = 0$, $s = 0, \ldots, r-1$. In other words, $\hat{\omega}_r$ is orthogonal with respect to the *signed* measure $\mathrm{d}\hat{\mu}(x) = \hat{\omega}_r(x)\mathrm{d}\mu(x)$. The measure being signed, there is little to say about positivity of the weights and the location of the nodes in general. The case that is mostly considered in the literature is $r = n+1$, aiming at an interpolatory QF with $2n+1$ points of order $3n+1$. If the zeros of ω_n and $\hat{\omega}_{n+1}$ interlace and are all in the interval of integration, then the weights will all be positive, and the order will be at least $3n+1$, but such cases can only be proven for particular measures. For some computational aspects see e.g. [17, 149, 423]. We also note that in the polynomial case, there also exist Radau-Kronrod and Lobatto-Kronrod QFs [390].

A similar approach exists for the unit circle. In that case, the extension consists of nodes that are zeros of para-orthogonal Stieltjes polynomials [146, 148]. One may force the additional nodes to be on $\mathbb{T}$, but neither the interlacing property nor the positivity of the weights can be guaranteed.

Generalization of the Kronrod ideas to the setting of rational functions is for the moment a largely unexplored domain.

14 Denseness and convergence

The quadrature formulas and the rational interpolation problems that were solved give approximations of the integral, and of the function interpolated. The discrete measure of the quadrature is an approximation of the measure used in the integral, which corresponds to an interpolation of the corresponding Carathéodory functions or equivalently of the associated Schur functions, So one might wonder whether we will recover the exact integral, the original Carathéodory or Schur function, or the original measure if we let n tend to infinity. In view of the convergence of the integral, a first problem is to find out which functions will be in the closure of the space of rational functions that were considered. It will clearly depend (1) on the asymptotic distribution of the points α_k that we have used to define the poles and (2) on the topology of the space.

Denseness problems and solution of rational moment problems are thus 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.

It might be useful to read Appendix A first. It gives some introductory notes on Hardy and Hilbert spaces.

14.1 Results for $\mathbb{T}$

As before, we will use $\pi_n(z) = \prod_{k=1}^{n}(1-\overline{\alpha}_k z)$, and by λ we refer to the normalized Lebesgue measure on $\mathbb{T}$. The Radon-Nikodym derivative of μ with respect to λ is denoted as μ'.

14.1.1 Different conditions

We need the following conditions for the measure μ (still assuming $\int \mathrm{d}\mu = 1$),

1. $\mu' > 0$ a.e. (λ) (Erdős-Turán condition), denoted as $\mu \in ET$,
2. $\log \mu' \in L^1(\lambda)$ (Szegő's condition), denoted as $\mu \in SZ$,

and the point set $\mathbb{A} = \{\alpha_1, \alpha_2, \ldots\}$:

1. $\sum_{n=1}^{\infty}(1 - |\alpha_n|) = \infty$ (Blaschke divergence condition), denoted as $\mathbb{A} \in BD$
2. $|\alpha_k| \leq r < 1$, $\forall k \in \mathbb{N}$ ($\mathbb{A}$ compactly included in $\mathbb{D}$), denoted as $\mathbb{A} \in CI$.

Note $SZ \subset ET$ and $CI \subset BD$. It is clear that these and other conditions can be imposed leading to different convergence results with subtle differences. We give below some illustrations of what is possible.

14.1.2 Denseness for $\mathbb{T}$

For the denseness of the sequence of partial Blaschke products in $L_p(\lambda)$, one can easily adapt a result of [1, p. 244] to find the following.

Theorem 14.1 *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 $L^p(\lambda)$, for any $p \geq 1$ as well as in the space $C(\mathbb{T})$ of continuous functions on $\mathbb{T}$ if and only if $\mathbb{A} \in BD$.*
Also the system $\{B_n\}_{n=0}^{\infty}$ is complete in H_p for any $p \geq 1$ if and only if $\mathbb{A} \in BD$.

This immediately implies (see Walsh [554, p. 305-306])

Theorem 14.2 *If $\mathbb{A} \in BD$ and $f \in H^2(\lambda)$, then $f_n \in \mathcal{L}_n$ which interpolates f in the points $\{w, \alpha_1, \ldots, \alpha_n\} \subset \mathbb{D}$ (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 μ, Szegő's condition plays a crucial role in this context. For example, from Walsh [554, p. 116, 186, 50, 92, 144] we may conclude that the polynomials are dense in $L^p(\mu)$ if and only if $\mu \notin SZ$. For the rational case, the situation seems to be less simple. See [105, Sec. 7.2].

Theorem 14.3 *With the above notation:*
(1) If $\mathbb{A} \in CI$ then $\mu \notin SZ \Leftrightarrow \{B_n\}_{n\geq 0}$ is complete in $L^2(\mu)$
(2) If $\mu \in SZ$ then $\mathbb{A} \in BD \Leftrightarrow \{B_n\}_{n\geq 0}$ is complete in $H^2(\mu)$
(3) If $\mathbb{A} \in BD$ then $\{B_n\}_{n\in\mathbb{Z}}$ is complete in $L^2(\mu)$.

14.1.3 Norm convergence

Convergence on $\mathbb{T}$ is most often treated in some norm. We start with a generalized form of *Szegő's problem*. The original problem was to find the function in $H^2(\mu)$ taking the value 1 in the origin and with the least $H^2(\mu)$ norm. This is equivalent with the orthogonal projection of 1 onto $zH^2(\mu)$, or with the projection of $1/z$ onto $H^2(\mu)$. When restricting to the finite dimensional subspaces $\mathcal{P}_n \subset H^2(\mu)$ then the answer is given by the orthogonal polynomial normalized with constant term 1 or its complex conjugate ϕ_n^*/κ_n if $\phi(z) = \kappa_n z^n + \cdots$, $\kappa_n > 0$ and $\phi_n^*(z) = z^n \phi_{n*}(z)$. The minimal norm squared is κ_n^{-2}. Recall from Section A.1 that $\sigma(z) = \exp\{\frac{1}{2}\int D(t,z) \log \mu'(t) \mathrm{d}\lambda(t)\}$ is the outer spectral factor of the measure. Then, taking $n \to \infty$ the previous approximant converges to $\overline{\sigma(0)}/\sigma(z)$ with minimal norm $|\sigma(0)|^2 = \exp\{\int \log \mu' \mathrm{d}\lambda\}$. The latter is zero when $\int \log \mu' \mathrm{d}\lambda = -\infty$, i.e., when the polynomials are dense in $H^2(\mu)$. See [316, p. 44], [262, p. 200]. But Grenander and Szegő immediately generalized this to the case where the constraint to take the value 1 is moved from the origin to some arbitrary point $w \in \mathbb{D}$. Again restricting to the subspace $\mathcal{P}_n$, the solution is given by a normalized reproducing kernel: $k_n(z,w)/k_n(w,w)$ with minimal norm squared $1/k_n(w,w)$. The latter, also known as the *Christoffel function*, converges to $(1-|w|^2)|\sigma(w)|^2$ and the kernel $k_n(z,w)$ to the *Szegő kernel* $s_w(z)$ defined by

$$\frac{1}{s_w(z)} = (1 - \overline{w}z)\sigma(z)\overline{\sigma(w)} \tag{14.1}$$

where it is assumed that $\log \mu' \in L^1$ so that $\sigma(z)$ is well defined. See [316, Section 3.2].

These results are readily generalized to the rational case where the subspaces $\mathcal{P}_n$ are replaced by rational subspaces $\mathcal{L}_n$. We have for example [105, Thm. 2.3.8 and 9.1.1]:

Theorem 14.4 *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 $\mathbb{A} \in BD$ then $\lim_{n\to\infty} \|k_n(t,w) - s_w(t)\|_\mu = 0$ and*

$$\lim_{n\to\infty} k_n(w,w) = s_w(w) \quad \textit{and} \quad \lim_{n\to\infty} \phi_n(w) = 0,$$

pointwise in $\mathbb{D}$.

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 $\mathbb{A} \in BD$ 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. Because of the denseness, $k_n(w,w)$ converges in $\mathbb{D}$ and the last line follows. □

For more related results about L^2 convergence, see [105, Sec. 9.1].

For stronger convergence results we shall have to rely on properties that were proved for polynomials with respect to a varying measure.

14.1.4 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 degree m at most and construct a sequence of OPVM for measure $\mathrm{d}\mu(t)/|\pi_n(t)|^2$. We denote the orthonormal ones by $\phi_{n,k}(z) = \kappa_{n,k} z^k + \cdots$, $\kappa_{n,k} > 0$. By our general theory they satisfy the recurrence $\phi_{n,m}(z) = e_{n,m}[z\phi_{n,m-1}(z) + \upsilon_{n,m}\phi^*_{n,m-1}(z)]$ with $\upsilon_{n,m} = \phi_{n,m}(0)/\kappa_{n,m}$ and $e_{n,m} = (1 - |\upsilon_{n,m}|^2)^{-1/2} = \kappa_{n,m}/\kappa_{n,m-1}$. The $\phi_{n,m}$ are polynomials orthogonal on $\mathbb{T}$, and thus they are a special case of the ORFs with all $\alpha_k = 0$. Hence the substar and superstar are well defined. However to make the distinction between $\phi^*_m(z) = B_m(z)\overline{\phi_m(1/\overline{z})} \in \mathcal{L}_m$ and $\phi^{\bar{*}}_{n,m}(z) = z^m\overline{\phi_{n,m}(1/\overline{z})} \in \mathcal{P}_n$ we introduce a different notation for the superstar. From these OPs we can construct a set of ORFs for $\mathcal{L}_n$, but as n increases, the whole set of ORFs will change.

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. It should be clear that $\{f_{n,m} : m = 0, \ldots, n\}$ is a set of orthonormal rational functions for $\mathcal{L}_n$ with respect to μ. They are obtained by orthonormalizing the basis $\{z^k/\pi_n(z) : k = 0, \ldots, n\}$. Then one may show that there is some normalizing constant c_n such that

$$c_n\phi_n(z) = \frac{z\phi_{n,n}(z) - \frac{\phi_{n,n}(\alpha_n)}{\phi^{\bar{*}}_{n,n}(\alpha_n)}\alpha_n\phi^{\bar{*}}_{n,n}(z)}{(z-\alpha_n)\pi_n(z)} \in \mathcal{L}_n \setminus \mathcal{L}_{n-1} \tag{14.2}$$

since it is not difficult to check that the middle expression is orthogonal $\mathcal{L}_{n-1}$.

The following results are borrowed from the work of Guillermo Lopez [397, 401, 402].

Theorem 14.5 *If $\mu \in ET$ and $\mathbb{A} \in BD$ then with our previous notation*

1. $\lim_{n\to\infty} \upsilon_{n,n+k+1} = 0$.
2. $\lim_{n\to\infty} \kappa_{n,n+k+1}/\kappa_{n,n+k} = 1$.
3. $\lim_{n\to\infty} \phi_{n,n+k+1}(z)/\phi_{n,n+k}(z) = z$ *locally uniformly in* $\mathbb{E}$.
4. $\lim_{n\to\infty} \phi^{\bar{*}}_{n,n+k+1}(z)/\phi^{\bar{*}}_{n,n+k}(z) = 1$ *locally uniformly in* $\mathbb{D}$.
5. $\lim_{n\to\infty} \phi^{\bar{*}}_{n,n+k}(z)/\phi_{n,n+k}(z) = 0$ *locally uniformly in* $\mathbb{E}$.

If $\mu \in SZ$, $\mathbb{A} \in CI$ and σ is the outer spectral factor of μ, then

6. $\lim_{n\to\infty} \phi^{\bar{*}}_{n,n+k}(z)/\pi_n(z) = 1/\sigma(z)$ *locally uniformly in* $\mathbb{D}$.

Much of the work of Guillermo Lopez is closely connected to what is given in this survey, but it is mostly placed in this context of varying measures. An excellent survey of his work was given by de la Calle Ysern [144] with a long list of references.

14.1.5 Asymptotics under Szegő's condition

To illustrate how the OPVM can be used and where the conditions are needed, we include some of the proofs. The following results hold [105, Thm. 9.6.{4-6,9}] (although there, slightly better results are formulated using Carleman type conditions for the rational moments).

Theorem 14.6 *If $\mu \in SZ$ and $\mathbb{A} \in BD$, then we have the following (l.u. means locally uniformly, i.e., uniformly in compact subsets)*

$$\lim_{n\to\infty} k_n(z,w) = s_w(z) = \frac{1}{\overline{\sigma(w)}(1-\overline{w}z)\sigma(z)}, \quad l.u.\ (z,w) \in \mathbb{D}\times\mathbb{D},$$

$$\lim_{n\to\infty} (1-\overline{\alpha}_n z)\frac{\phi_n^*(z)}{\phi_n^*(0)} = \frac{\sigma(0)}{\sigma(z)}, \quad l.u.\ z \in \mathbb{D},$$

$$\lim_{n\to\infty} \phi_n(z) = \lim_{n\to\infty} \frac{\phi_n(z)\phi_n^*(0)}{\phi_n^*(z)} = \lim_{n\to\infty} \frac{\phi_n(z)}{\phi_n^*(z)} = 0, \quad l.u.\ z \in \mathbb{D}.$$

If $\mu \in SZ$, $\sigma(0) > 0$, and $\mathbb{A} \in CI$ 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 ϕ_{nk} and divide by $f_{n,n}^*(z)\overline{f_{n,n}^*(w)}$:

$$\frac{k_n(z,w)}{f_{n,n}^*(z)\overline{f_{n,n}^*(w)}} = \frac{1}{1-\overline{w}z} - \frac{z\overline{w}}{1-\overline{w}z}\left(\frac{\phi_{n,n}(z)}{\phi_{n,n}^{\bar{*}}(z)}\right)\overline{\left(\frac{\phi_{n,n}(w)}{\phi_{n,n}^{\bar{*}}(w)}\right)}. \tag{14.3}$$

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.
By taking the superstar of (14.2)

$$(1-\overline{\alpha}_n z)\phi_n^*(z) = \frac{1}{\overline{c}_n}\frac{\phi_{n,n}^{\bar{*}}(z)}{\pi_n(z)}\left[1-\overline{\eta}_n\overline{\alpha}_n z\frac{\phi_{n,n}(z)}{\phi_{n,n}^{\bar{*}}(z)}\right], \quad \eta_n = \frac{\phi_{n,n}(\alpha_n)}{\phi_{n,n}^{\bar{*}}(\alpha_n)}. \tag{14.4}$$

Recall that $\phi_{n,n}(z)/\phi_{n,n}^{\bar{*}}(z) \to 0$ 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, $\phi_{n,n}^{\bar{*}}(z)/\pi_n(z)$ 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 that $\sum_{k=1}^{n} |\phi_k(z)|^2$ 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 η_n is used to normalize ϕ_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 α_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. □

With stronger conditions we can have convergence in the closed disk $\mathbb{D} \cup \mathbb{T}$ [136].

14.1.6 Beyond Szegő's condition: ratio asymptotics

When $\mu \notin SZ$, then we should avoid σ. This is realized by considering ratio asymptotics, so that the σ cancel out. These are weaker results in the sense that if $\mu \notin SZ$, then the ratio asymptotics are almost immediately obtained from the stronger asymptotics that were previously obtained. We use again the OPVM with measure $\mathrm{d}\mu = \mathrm{d}\mu/|\pi_n|^2$. Denoting the reproducing kernel for $\mathcal{P}_n$ w.r.t. this measure as $k_{n,n}(z,w)$, it is easily seen that $k_{n,n}(z,0) = \kappa_{n,n}\phi_{n,n}^{\bar{*}}(z)$. Also the usual rational kernel for $\mathcal{L}_n$ satisfies $k_n(z,0) = k_{n,n}(z,0)/\pi_n(z)$.

As a starting exercise, one may want to prove a property about the equivalence of the inner product in $\mathcal{L}_n$:

$$\int \overline{f(t)} g(t) \mathrm{d}\mu(t) = \int \overline{f(t)} g(t) \frac{\kappa_{n,n}^2}{|k_n(t,0)|^2} \mathrm{d}\lambda(t) \tag{14.5}$$

That follows by applying Theorem 7.18 to the OPVM but it is also a direct consequence of the results in Section 7.6.

We are then ready for the following theorem [105, Thm. 9.8.8-9].

Theorem 14.7 *If $\mu \in ET$ and $\mathbb{A} \in BD$, then*

$$\lim_{n\to\infty} \frac{k_n(z,0)}{k_{n+1}(z,0)} = 1, \quad \textit{l.u. } z \in \mathbb{D}.$$

and with the superstar referring to z,

$$\lim_{n\to\infty} k_n^*(z,0)/k_n(z,0) = 0, \quad \textit{l.u. } z \in \mathbb{D}.$$

Proof. Set $g_n(z) = k_n(z,0)/k_{n+1}(z,0) - v_n/v_{n+1}$, then by (14.5) we see that

$$\int |g_n(t)|^2 \mathrm{d}\lambda(t) = 2\frac{v_n^2}{v_{n+1}^2}\left(1 - \frac{v_n}{v_{n+1}}\right). \tag{14.6}$$

On the other hand, because g_n is analytic in $\mathbb{D} \cup \mathbb{T}$, we have by the Poisson integral $g_n(z) = \int P(t,z) g_n(t) \mathrm{d}\lambda(t)$. For $t \in \mathbb{T}$, $|P(t,z)| \leq M$ holds uniformly for z in a compact subset of $\mathbb{D}$. Thus we also have $|g_n(z)| \leq M \int |g_n(t)| \mathrm{d}\lambda(t)$ and also $|g_n(z)|^2 \leq M^2 \int |g_n(t)|^2 \mathrm{d}\lambda(t)$ so that

$$|g_n(z)|^2 \leq 2M^2 \frac{v_n^2}{v_{n+1}^2}\left(1 - \frac{v_n}{v_{n+1}}\right). \tag{14.7}$$

Thus $\frac{k_n(z,0)}{k_{n+1}(z,0)} \to 1$ iff $\frac{\upsilon_n}{\upsilon_{n+1}} \to 1$. Obviously $k_{n-1}(0,0)/k_n(0,0) \leq 1$, but using the CD relation we also have

$$\frac{k_{n-1}(0,0)}{k_n(0,0)} = \frac{|\phi_n^*(0)|^2 - |\phi_n(0)|^2}{|\phi_n^*(0)|^2 - |\alpha_n|^2|\phi_n(0)|^2} \geq \frac{|\phi_n^*(0)|^2 - |\phi_n(0)|^2}{|\phi_n^*(0)|^2 - |\phi_n(0)|^2} = 1.$$

So that $\lim_{n\to\infty} \upsilon_n^2/\upsilon_{n+1}^2 = 1$, which proves the first part.

For the second relation, note that $k_n(z,0) = \phi_{n,n}^{\bar{*}}(z)\upsilon_n\eta_n/\pi_n(z)$ and hence $k_n^*(z,0) = \phi_{n,n}(z)\upsilon_n\overline{\eta}_n/\pi_n(z)$ 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. □

The ratio asymptotics for the ORFs can now also be given [105, Thm. 9.8.14-15].

Theorem 14.8 *If $\mu \in ET$ and $\mathbb{A} \in CI$, then (υ_n is a Verblunsky coefficient)*

$$\lim_{n\to\infty} \frac{\zeta_n(z)\phi_n(z)}{\phi_n^*(z)} = \lim_{n\to\infty} \frac{\phi_n(z)}{\phi_n^*(z)} = \lim_{n\to\infty} \upsilon_n = 0, \quad \textit{l.u. in } \mathbb{D}.$$

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 14.6. I.e., use equation (14.2) and its superstar conjugate (14.4), 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 14.5 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 $\mathbb{A} \in CI$. It is also needed for the convergence of $\phi_n(0)/\phi_n^*(0)$ because this will follow form Theorem 14.7 or (14.5) of the previous exercise if $\mathbb{A}$ is in a compact subset of $\mathbb{D}$. For the last equality note that $|\upsilon_n| = |\phi_n(\alpha_{n-1})/\overline{\phi_n^*(\alpha_{n-1})}|$.

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

$$\Phi_n^*(z)(1-\overline{\alpha}_n z) = \frac{\phi_n^*(z)(1-\overline{\alpha}_n z)}{\phi_n^*(0)} = \frac{k_n(z,0)}{\upsilon_n^2}\left[1 - \overline{\alpha}_n z \frac{\overline{\phi_n(0)}}{\phi_n^*(0)}\frac{k_n^*(z,0)}{k_n(z,0)}\right].$$

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 $\upsilon_{n+1}/\upsilon_n$ go to 1 as $n \to \infty$, the second result follows. □

14.1.7 The nth root asymptotics

The weakest form of asymptotics are expressions is the nth root asymptotics in the sense that strong asymptotics imply ratio asymptotics and these imply the weakest form. The latter are usually derived in the context of potential theory and convergence in capacity [501]. There is an extensive literature where these techniques are applied in rational approximation and orthogonal polynomials. We give just a couple of examples for the rational case.

To define the distribution of the poles, i.e., of the sequence $\mathbb{A}$ we introduce a counting measure $\nu_n^{\mathbb{A}} = \frac{1}{n}\sum_{k=1}^n \delta_{\alpha_k}$. Its *logarithmic potential* is

$$V_{\nu_n^{\mathbb{A}}}(z) = -\int \log|z-x|\mathrm{d}\nu_n^{\mathbb{A}}(x) = -\frac{1}{n}\sum_{k=1}^n \log|z-\alpha_k|.$$

Note that $|\pi_n^*(z)|^{1/n} = \exp\{-V_{\nu_n^{\mathbb{A}}}(z)\}$. Set $\hat{z} = 1/\overline{z}$ for any complex z. Then if we have weak star convergence $\nu_n^{\mathbb{A}} \xrightarrow[n]{*} \nu^{\mathbb{A}}$ with compact support, then [105, Lemma 9.9.1]

$$\limsup_{n\to\infty} |B_n(z)|^{1/n} \le \exp\{\omega(z)\} \quad \text{and} \quad \limsup_{n\to\infty} |B_n(z)|^{-1/n} \le \exp\{\omega(\hat{z})\}$$

where $z \in \mathbb{C}$ and

$$\omega(z) = \int \log|\zeta_z(t)|\mathrm{d}\nu^{\mathbb{A}}(t), \quad \zeta_z(t) = \mathfrak{m}_z[t] = \frac{t-z}{1-\overline{z}t}.$$

The lim sup can be replaced by lim and the inequalities become equalities if $z \notin \operatorname{supp}\nu^{\mathbb{A}} \cup \operatorname{supp}\nu^{\hat{\mathbb{A}}}$. This implies then relatively easily [105, Thm. 9.9.2-5] and [121].

Theorem 14.9 *Suppose $\mu \in ET$ and $\mathbb{A} \in CI$ then with the notation just introduced the root asymptotics for the ORF are*

$$\lim_{n\to\infty} |\phi_n^*(z)|^{1/n} = 1, \quad l.u.\ z \in \mathbb{D}$$

and

$$\lim_{n\to\infty} |\phi_n(z)|^{1/n} = \exp\{\omega(z)\} \quad l.u.\ z \in \mathbb{E} \setminus \operatorname{supp}\nu^{\hat{\mathbb{A}}}.$$

If $\mu \in SZ$ and $\mathbb{A} \in CI$ then the PORF $Q_n(z,\tau)$ satisfy

$$\lim_{n\to\infty} |Q_n(z,\tau)|^{1/n} = 1, \quad l.u.\ z \in \mathbb{D}$$

and $\lim_{n\to\infty} \|Q_n\|_\infty^{1/n} = 1$ and

$$\lim_{n\to\infty} |Q_n(z,\tau)|^{1/n} = \exp\{\omega(z)\} \quad l.u.\ z \in \mathbb{E} \setminus \operatorname{supp}\nu^{\hat{\mathbb{A}}}.$$

14.1.8 Convergence of interpolation and QF

Convergence results do follow immediately from the asymptotics that were just obtained. Because of the close connection between quadrature and the connected interpolation problems, we can treat them simultaneously.

As a first example we give a convergence result from [98, Thm. 4.3] and [119, Thm. 6].

Theorem 14.10 *Suppose $\mathbb{A} \in CI$ then with $\Omega_\mu(z) = \int D(t,z)\mathrm{d}\mu(t)$ and ϕ_n the ORF and ψ_n the functions of the second kind, then*

$$\begin{aligned}
\lim_{n\to\infty} \frac{\psi_n^*(z)}{\phi_n^*(z)} &= -\Omega_\mu(z) \quad l.u.\ z \in \mathbb{D} \\
\lim_{n\to\infty} \frac{\psi_n(z)}{\phi_n(z)} &= \Omega_\mu(z) \quad l.u.\ z \in \mathbb{E} \\
\lim_{n\to\infty} \frac{P_n(z,\tau)}{Q_n(z,\tau)} &= \Omega_\mu(z) \quad l.u.\ z \in \mathbb{D} \cup \mathbb{E}
\end{aligned}$$

and if $\mathfrak{I}_n$ is the n-point rational Szegő QF then $\lim_{n\to\infty} \mathfrak{I}_n[f] = \mathfrak{I}_\mu[f]$ for any μ-integrable function $f \in L^1(\mu)$.

We get rates of convergence via nth root asymptotics. The distribution of the $\mathbb{A} = \{\alpha_k : k \geq 1\}$ is characterized by $\nu^{\mathbb{A}}$ and the logarithmic potential was defined for $z \in \mathbb{C} \setminus \left(\{0\} \cup \operatorname{supp}(\nu^{\mathbb{A}}) \cup \operatorname{supp}(\nu^{\hat{\mathbb{A}}})\right)$ as

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

Also recall $\hat{\mathbb{A}} = \{1/\overline{\alpha} : \alpha \in \mathbb{A}\}$. The following theorem holds concerning the convergence of the rational Szegő QF [121, Thm. 5.4] and the Carathéodory interpolation problem [121, Thm. 4.3].

Theorem 14.11 *Let f be a function such that it is analytic in a closed region G with $\mathbb{T} \subset G$ and $G \cap (\mathbb{A} \cup \hat{\mathbb{A}}) = \varnothing$ and such that its boundary Γ is a finite set of rectifiable curves. Suppose further $\sum(1 - |\alpha_k|) = \infty$. Then with $\Omega_\mu(z) = \int D(t, z)\mathrm{d}\mu(t)$ we have*

$$\begin{aligned}
\lim_{n\to\infty} \left| \frac{\psi_n(z)}{\phi_n(z)} + \Omega_\mu(z) \right|^{1/n} &\leq \exp\{\omega(z)\} < 1, \quad z \in \mathbb{D} \\
\lim_{n\to\infty} \left| \frac{\psi_n^*(z)}{\phi_n^*(z)} - \Omega_\mu(z) \right|^{1/n} &\leq \exp\{\omega(\hat{z})\} < 1, \quad z \in \mathbb{E} \\
\lim_{n\to\infty} \frac{P_n(z, \tau)}{Q_n(z, \tau)} &= \Omega_\mu(z) \quad \textit{l.u. } z \in \mathbb{D} \cup \mathbb{E}
\end{aligned}$$

and the integration error $\mathcal{E}_n$ of the rational Szegő QF $\mathfrak{I}_n$ satisfies for $f \in L^1(\mu)$

$$|\mathcal{E}_n[f]|^{1/n} < \gamma < 1$$

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

For similar but more advanced convergence results see [104].

14.2 Results for $\mathbb{R}$

For the case of a measure with support on a part of $\mathbb{R}$, there are several possibilities. As before, we can consider the standard ones: 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. On top of that, there are also the many ways in which one may want to prove convergence. We shall restrict this section to some examples of what could be expected.

14.2.1 Denseness for $\mathbb{R}$

Just an example from [399].

Theorem 14.12 *Let μ be supported on $[0, \infty)$ and all the poles α_k are simple and in $\mathbb{C} \setminus [0, \infty)$. If*

$$\sum_{\operatorname{Im}(\sqrt{\alpha_k}) > 0} \frac{\operatorname{Im}(\sqrt{\alpha_k})}{1 + |\alpha_k|} = +\infty \tag{14.8}$$

then if all the α_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 $[0,\infty)$ that satisfy the condition $\lim_{x\to+\infty} f(x) < \infty$.

In the case of real poles and $\operatorname{supp}\mu \subset [-1,1]$ we have

Theorem 14.13 *Let μ be supported on $[-1,1]$ and all the poles $\alpha_k \in \mathbb{R}\setminus[-1,1]$. If*

$$\sum_{k=1}^{\infty} \sqrt{\alpha_k^{-2}-1} = +\infty$$

then if all the α_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 $[-1,1]$.

These and other results are direct consequences of what can be found in the book by Achieser [1, p. 254] and also [69, Chap. 4].

14.2.2 Asymptotics for the real line

Depending on what the support of the measure is and where the poles are chosen, the different conditions that we formulated in the case of the circle may translate in different ways to the situation at hand for the real line. We shall not cover all possibilities but just mention a few characteristic examples.

The interval $I = [-1,1]$

Many results for asymptotics on an interval $I = [-1,1]$ can be derived form the results known on $\mathbb{T}$ using the Joukowski transform.

For the case of a measure μ on I and poles α_k in $\hat{\mathbb{C}}^I = \hat{\mathbb{C}}\setminus I$, denote $z = \mathrm{e}^{\mathrm{i}\theta}$, $\cos\theta = x = \mathfrak{j}(z)$, $\beta_k = \mathfrak{j}^{inv}(\alpha_k)$, and $\tilde{\mu}'(z) = \mu'(\cos\theta)|\sin\theta|$, and ϕ_k the ORF for μ on I and poles $\{\alpha_k\}_k$, then if $\int_I \frac{\mu'(x)}{\sqrt{1-x^2}}\mathrm{d}x > -\infty$ (equivalent with $\tilde{\mu}\in SZ$) and if the α_k's stay away from I (i.e. the set of β_k's is in CI), then

$$\lim_{n\to\infty} \eta_n B_n(z) \frac{1-\overline{\beta}_n z}{\sqrt{1-|\beta_n|^2}} \phi_n(\mathfrak{j}(z)) = \frac{1}{\sqrt{2\pi}\sigma(z)}, \quad l.u. \quad z\in\mathbb{D}$$

for an appropriate unimodular constant η_n, $x = \mathfrak{j}(z)$, B_n the Blaschke product with zeros $\beta_1,\ldots,\beta_n$, and σ the outer spectral factor of $\tilde{\mu}$. Under the same conditions also

$$\lim_{n\to\infty} \eta_n \frac{z-\beta_{n+1}}{1-\overline{\beta}_n z} \sqrt{\frac{1-|\beta_n|^2}{1-|\beta_{n+1}|^2}} \frac{\phi_{n+1}(z)}{\phi_n(z)} = 1 \quad l.u. \quad z\in\hat{\mathbb{C}}^I.$$

For these and other related results see [194]. The same ratio asymptotics hold even for the weaker ET condition that $\mu' > 0$ a.e. in I if we restrict ourselves to $\mathbb{A}\subset\mathbb{R}^I$ and $z\in\mathbb{R}^I$ [533]. For least squares convergence and interpolation in a general situation for the poles which

are assumed to be bounded away from the interval $[-1,1]$ see [192]. For corresponding weak star convergence see also [531].

There are also estimates for the rate of convergence to $S_\mu(z) = \int \frac{\mathrm{d}\mu(x)}{x-z}$ of the rational interpolants $S_n = -\frac{\sigma_n}{\phi_n}$, ϕ_n is an ORF and $\sigma_n(z) = \int \frac{\phi_n(x)-\phi_n(z)}{x-z}\mathrm{d}\mu(x)$ and for the rational Szegő quadrature formula if the integrand f is analytic in a neighborhood G of I and $\mathbb{A} \subset \mathbb{C}\backslash G$. See for the general rational case [142, 143, 347]. For the two point situation of Laurent polynomials see [85–90, 126].

The whole real line $\mathbb{R}$

An example for the whole real line from [98, Thm. 4.3] that holds for the QORF $Q_n(z,\tau_n)$ that form a regular sequence. Recall that this means that each τ_n is chosen such that $Q_n(z,\tau_n)$ has n simple zeros none of which is equal to one of the α_k. An ORF ϕ_n is called regular if α_{n-1} is not a zero of its numerator.

Theorem 14.14 *Suppose the sequence of QORFs $Q_n(z,\tau_n)$ is regular. Then there exists a subsequence n_k such that*

$$\lim_{k\to\infty} \Omega_{n_k}(z) = -\lim_{k\to\infty} \frac{P_{n_k}(z,\tau)}{Q_{n_k}(z,\tau)} = \Omega_\mu(z) \quad l.u.\ z \in \mathbb{U}.$$

If $\phi_n(z)$ is a regular sequence and $\tau = 0$ is regular then

$$\lim_{n\to\infty} \Omega_{n_k}(z) = -\lim_{k\to\infty} \frac{\psi_{n_k}(z)}{\phi_{n_k}(z)} = \Omega_\mu(z) \quad l.u.\ z \in \mathbb{U}$$

The half line $\mathbb{R}^+ = [0,\infty)$

If $\operatorname{supp}\mu$ is on $[0,\infty)$, then the poles can be chosen on the negative real axis, so that there can be no problem with regularity. Then we can prove some monotone convergence for the even and the odd numbered approximants. Indeed, one may prove [114, Thm. 13.1-2], [95, Thm. 4.3-5]:

Theorem 14.15 *Suppose $-\infty \le \alpha_{2m} \le \beta < \alpha \le \alpha_{2m+1} \le 0$ then for the Stieltjes function*

$$S_\mu(z) = \int_0^\infty \frac{\mathrm{d}\mu(x)}{z-x} \quad \textit{and} \quad \sigma_n(z) = \int_0^\infty \frac{\phi_n(t)-\phi_n(z)}{t-z}\mathrm{d}\mu(t), \quad z \in \mathbb{C}\setminus[0,\infty)$$

where ϕ_n are the ORF we have with $S_n = \sigma_n/\phi_n$

$$\begin{aligned} \lim_{n\to\infty} S_{2n}(z) &= \frac{\sigma_{2n}(z)}{\phi_{2n}(z)} \searrow S^\infty(z), \quad z \in (\beta,\alpha) \\ \lim_{n\to\infty} S_{2n+1}(z) &= \frac{\sigma_{2n+1}(z)}{\phi_{2n+1}(z)} \nearrow S^0(z), \quad z \in (\beta,\alpha) \end{aligned}$$

and thus $S^0 \le S^\infty$ on (β,α). In $z\setminus[0,\infty)$, the rational functions $S_{2n}(z)$ converge l.u. to S^∞ and the $S_{2n+1}(z)$ converge l.u. to S^0. In fact the measures for S^∞ and S^0 are both the Stieltjes transforms of 'extreme' solutions for the rational Stieltjes moment problem on $\mathcal{L}\cdot\mathcal{L}$ in the sense that for any other solution μ, its Stieltjes transform S_μ will satisfy $S^0 \le S_\mu \le S^\infty$.

Another result holds for rational functions with complex poles. From Theorem 14.12 we know that the rationals will be dense if $\sum \frac{\sqrt{|\alpha_k|}}{1+|\alpha_k|} = \infty$ implying the denseness of the ORF. See [111, Thm. 5.1] and [111, Thm. 4.3] but also [398, 400].

Theorem 14.16 *Suppose $\mathbb{A}$ satisfies the denseness condition (14.8) then for S_μ and S_n as defined in Theorem 14.15*

$$\lim_{n\to\infty} S_n(z) = S_\mu(z) \quad l.u.\ z \in \mathbb{C} \setminus [0,\infty)$$

and the rational Szegő QF $\mathfrak{I}_n$ converges for any integrable function, i.e., $\lim_{n\to\infty} \mathfrak{I}_n[f] = \mathfrak{I}_\mu[f]$ for all $f \in L^1(\mu;\mathbb{R}^+)$.

14.3 Convergence of modified approximants and moment problem

The convergence of the PORF (for $\mathbb{T}$) and of the QORF (for $\mathbb{R}$) still depend on the parameter τ. This τ can vary over $\mathbb{T}$ for the PORF and over $\mathbb{R}$ for the QORF. By varying this τ, each approximant $\Omega_n(z,\tau) = \frac{P_n(z,\tau)}{Q_n(z,\tau)}$ that interpolates in the α_k, and hence matches a number of moments, will describe a circle with radius r_n (a so called *Weyl circle*) that is the boundary of a disk Δ_n. As n increases the disks form a nested shrinking sequence. If the radius r_n goes to 0, then the moment problem has a unique solution, and it will be indeterminate in the other case. See for example [93–96, 99, 100, 102, 103, 117].

Let us illustrate this in the case of the circle. See [105, Chap. 10] and [110]. Consider for fixed $z \in \mathbb{C} \setminus \mathbb{T}$ the approximants

$$K_n(z) = \Omega_n(z,\tau) = -\frac{P_n(z,\tau)}{Q_n(z,\tau)} = -\frac{\psi_n(z) - \tau\psi_n^*(z)}{\phi_n(z) + \tau\phi_n^*(z)}, \quad \tau \in \mathbb{T}$$

where $Q_n(z,\tau)$ is a PORF and $P_n(z,\tau)$ is the associated rational function of the second kind. Making use of Green's formula (7.4) and the CD relation, one may show that $K_n(z)$ will describe a circle with center $c_n(z)$ and radius $r_n(z)$ given by

$$c_n = \frac{\psi_n^*\overline{\phi_n^*} + \psi_n\overline{\phi_n}}{|\phi_n^*|^2 - |\phi_n|^2}, \quad r_n = \left|\frac{\psi_n^*\phi_n + \psi_n\phi_n^*}{|\phi_n^*|^2 - |\phi_n|^2}\right| = \frac{2|z|}{|1-|z|^2|}\frac{|B_{n-1}(z)|}{k_{n-1}(z,z)}$$

where B_{n-1} is the usual Blaschke product and k_{n-1} the reproducing kernel for $\mathcal{L}_{n-1}$. This shows that the circle $K_n(z)$ bounds a disk $\Delta_n(z)$ that will be in the right half plane if $z \in \mathbb{D}$ and in the left half plane if $z \in \mathbb{E}$. That is $\frac{\operatorname{Re} K_n(z)}{1-|z|^2} > 0$ for $z \in \mathbb{C} \setminus \mathbb{T}$. Moreover, the radius is monotonically shrinking with n, so that we have a sequence of nested disks. The limiting disk Δ_∞ will shrink to a point if and only if the Blaschke product diverges to 0, i.e., $\mathbb{A} \in BD$ thus $\sum_k (1-|\alpha_k|) = \infty$. Note that this conclusion is independent of z.

Let us explain this in some more detail. If we assume that we only know the rational moments of some unknown probability measure μ supported on $\mathbb{T}$:

$$m_k = \int B_k(t)\mathrm{d}\mu(t), \quad k = 0,1,\ldots$$

then this is sufficient to define an inner product in our space of rational functions $\mathcal{L}$. Indeed, since we assume μ real, $m_{-k} = \int \frac{\mathrm{d}\mu(t)}{B_k(t)} = \int B_{k*}(t)\mathrm{d}\mu(t) = \overline{m}_k$. Hence, by partial fraction expansion, we also know all the moments $m_{p,q} = \int B_{p*}(t)B_q(t)\mathrm{d}\mu(t)$, which allows to define the inner product. Thus, if we know the rational moments for a basis of $\mathcal{L}$, then we also know the rational moments for a basis of $\mathcal{L}\cdot\mathcal{L}$ (for a real measure on $\mathbb{T}$). Therefore, under these conditions, the rational moment on $\mathcal{L}$ and on $\mathcal{L}\cdot\mathcal{L}$ is the same (which in general is not

true for the real line – see below). Suppose all solutions of the moment problem forms a set $\mathcal{M}$, then it follows that for $z \in \mathbb{C} \setminus \mathbb{T}$ and $\mu \in \mathcal{M}$ we may define

$$\Omega_\mu(z) = \int D(t,z)\mathrm{d}\mu(t), \quad D(t,z) = \frac{t+z}{t-z}$$

and it turns out that $\{\Omega_\mu(z) : \mu \in \mathcal{M}\} = \Delta_\infty(z)$. See [105, Thm. 10.3.1].

Thus, from the above expression for the radius of the circles, it follows that the divergence of the Blaschke product is sufficient for the moment problem to be determinate. If the Blaschke product converges, then the $K_n(z)$ will converge to a circle with positive radius and the moment problem is indeterminate. The latter implies that both $\sum_{k=0}^n |\phi_k|^2$ and $\sum_{k=0}^n |\psi_k|^2$ converge locally uniformly in $\mathbb{C} \setminus \mathbb{T}$.

The case of the real line is analogous and yet there are some twists to the theory [105, Sec. 11.7-8]. Now we have to consider QORF and we can define for any regular index n

$$K_n(z) = \Omega_n(z,\tau) = -\frac{P_n(z,\tau)}{Q_n(z,\tau)} = -\frac{\psi_n(z) + \tau\frac{Z_n(z)}{Z_{n-1}(z)}\psi_{n-1}(z)}{\phi_n(z) + \tau\frac{Z_n(z)}{Z_{n-1}(z)}\phi_{n-1}(z)}, \quad \tau \in \hat{\mathbb{R}}.$$

The $Q_n(z,\tau)$ are the QORF and $P_n(z,\tau)$ are the associated functions of the second kind. Again, it can be shown that the $K_n(z)$ are the boundaries of nested disks with radius

$$r_n(z) = \frac{1}{k_{n-1}(z,z)}\left|\frac{2\zeta_{\mathrm{i}}(z)}{1-|\zeta_{\mathrm{i}}(z)|^2}\right|, \quad \zeta_{\mathrm{i}}(z) = \frac{z-\mathrm{i}}{1+\mathrm{i}z}.$$

lying in the right half plane if $z \in \mathbb{U}$ and in the left half plane if z is in the lower half plane. Moreover, since $\Omega_n(z,\infty) = \Omega_{n-1}(z,0)$, the circles $K_n(z)$ and $K_{n-1}(z)$ will touch.

The points $\pm\mathrm{i}$ are exceptional in the sense that in those points all the circles have radius $r_n(\pm\mathrm{i}) = 1$ and $K_n(\pm\mathrm{i}) = \pm 1$.

So as long as we stay away from $\operatorname{supp}\mu$ and from $\pm\mathrm{i}$, and if we select a subsequence of regular indices, we can draw analog conclusions as in the case of the circle.

For the moment problem we should consider moments

$$m_n = \int b_n(x)\mathrm{d}\mathring{\mu}(x), \quad b_n(z) = \prod_{k=1}^n Z_k(z), \quad Z_k(z) = \frac{z}{1-z/\alpha_k}, \ n = 0,1,2,\ldots.$$

It makes perfect sense to find a positive measure on $\mathbb{R}$ having these moments. However, that is not enough to define an inner product on $\mathcal{L}$ because $\mathcal{R} = \mathcal{L}\cdot\mathcal{L} \neq \mathcal{L}$, unless all poles are repeated an infinite number of times. Hence we need doubly indexed moments to be given

$$m_{p,q} = \int b_p(x)b_q(x)\mathrm{d}\mathring{\mu}(x), \quad p,q = 0,1,2,\ldots.$$

We can call the first moment problem with prescribed m_k the moment problem on $\mathcal{L}$ with solution set $\mathcal{M}^{\mathcal{L}}$ (possibly empty) while if the prescribed moments are all $m_{p,q}$ then we call it the moment problem on $\mathcal{R}$ with solution set $\mathcal{M}^{\mathcal{R}}$. Clearly $\mathcal{M}^{\mathcal{R}} \subset \mathcal{M}^{\mathcal{L}}$.

Define as before

$$\Omega_\mu(z) = \int D(x,z)\mathrm{d}\mathring{\mu}(x), \quad D(x,z) = \frac{1}{\mathrm{i}}\frac{1+xz}{x-z}$$

and the limiting disk $\Delta_\infty(z)$, $z \in \mathbb{C} \setminus (\mathbb{R} \cup \{\mathrm{i}\} \cup \{-\mathrm{i}\})$ then

$$\{\Omega_\mu(z) : \mu \in \mathcal{M}^{\mathcal{R}}\} \subset \Delta_\infty(z) \subset \{\Omega_\mu(z) : \mu \in \mathcal{M}^{\mathcal{L}}\}.$$

For further details consult [99, 114, 120].

Almendral in [10] gave a Nevanlinna parametrization of all solutions to the rational moment problem on $\mathcal{R}$ for a measure with $\operatorname{supp}\mu \subset \mathbb{R}$ and $\mathbb{A} \subset \hat{\mathbb{R}} \setminus \{0\}$. He proves

Theorem 14.17 *Let ϕ_n be the ORF corresponding to the moment problem in $\mathcal{R}$ and ψ_n the associated functions of the second kind. Set as before $\tilde{\phi}_n(x) = x\phi_n(x)/Z_n(x)$ and $\tilde{\psi}_n(x) = x\psi_n(x)/Z_n(x)$ and $x \in \mathbb{R} \setminus \mathbb{A}$. Then, assuming there is an infinite set Λ of regular indices and the limits are taken over this set then the functions A, B, C and D defined as*

$$\begin{aligned}
A(z) = A(z,x) &= \lim_{n\in\Lambda}[\tilde{\psi}_n(x)\tilde{\psi}_{n-1}(z) - \tilde{\psi}_n(z)\tilde{\psi}_{n-1}(z)] \\
B(z) = B(z,x) &= \lim_{n\in\Lambda}[\tilde{\psi}_n(x)\tilde{\phi}_{n-1}(z) - \tilde{\phi}_n(z)\tilde{\psi}_{n-1}(z)] \\
C(z) = C(z,x) &= \lim_{n\in\Lambda}[\tilde{\phi}_n(x)\tilde{\psi}_{n-1}(z) - \tilde{\psi}_n(z)\tilde{\phi}_{n-1}(z)] \\
D(z) = D(z,x) &= \lim_{n\in\Lambda}[\tilde{\phi}_n(x)\tilde{\phi}_{n-1}(z) - \tilde{\phi}_n(z)\tilde{\phi}_{n-1}(z)]
\end{aligned}$$

are analytic in $\mathbb{C} \setminus \mathbb{A}$. The convergence is locally uniform in $\mathbb{C} \setminus (\mathbb{R} \cup \{\pm\mathrm{i}\})$. Finally there is a one-to-one correspondence between the functions $\{\Omega_\mu : \mu \in \mathcal{M}^{\mathcal{R}}\}$ (and hence also between $\mu \in \mathcal{M}^{\mathcal{R}}$) and the functions ω in the Carathéodory Class $\mathcal{C}$. It is given by

$$\int D(x,z)\mathrm{d}\mathring{\mu}(x) = -\frac{A(z)\omega(z) - C(z)}{B(z)\omega(z) - D(z)}.$$

A similar result can be formulated for the rational moment problem on $\mathbb{T}$. A simultaneous treatment of both cases is given in [453]. See also [129]. For the limiting behaviour of the A, B, C, D functions in the case of a finite number of poles that each are repeated and infinite number of times, see [102, 127]. All this is related to the Nevanlinna-Pick algorithm and the recurrence for the reproducing kernels. Let us quickly review this. See also the brief introduction at the end of Section 6.

The Schur class $\mathcal{B}$ for the unit disk is the set of functions $\Gamma(z)$ that are analytic in $\mathbb{D}$ and are bounded by 1 in modulus: $\mathcal{B} = \{\Gamma \in H(\mathbb{D}) : \Gamma(\mathbb{D}) \subset \mathbb{D} \cup \mathbb{T}\}$. They are related to the Carathéodory functions because if $\Omega \in \mathcal{C}$ is a Carathéodory function (analytic and with positive real part in $\mathbb{D}$) with $\Omega(0) = 1$, then $\Gamma(z) = (1-\Omega(z))/(1+\Omega(z))$ is a Schur function with $\Gamma(0) = 0$ and vice versa, $\Omega(z) = (1-\Gamma(z))/(1+\Gamma(z))$ will be a Carathéodory function with $\Omega(0) = 1$. Furthermore, it is easily seen that if $\alpha_n \in \mathbb{D}$, then if $\Gamma_n \in \mathcal{B}$ also

$$\Gamma_{n+1}(z) = \frac{\Gamma_n(z) - \gamma_n}{1 - \overline{\gamma}_n\Gamma_n(z)}\frac{1}{\zeta_n(z)} \in \mathcal{B}, \quad \gamma_n = \Gamma_n(\alpha_n), \quad \zeta_n(z) = \frac{z - \alpha_n}{1 - \overline{\alpha}_n z}.$$

The Schur algorithm just repeats the previous transform with $\alpha_n = 0$ in every step. If $\Gamma_0 \in \mathcal{B}$, then $\Gamma_n(0)$ should be in $\mathbb{D}$ in all steps, unless Γ_0 is rational, in which case the algorithm stops after n steps. In the Nevanlinna-Pick algorithm, the α_k are selected arbitrarily from $\mathbb{D}$, but again if $\Gamma_0 \in \mathcal{B}$, then the $\Gamma_n(\alpha_n)$ should all be in $\mathbb{D}$. The original papers are [430–433, 477–479]. The parameters γ_n are sometimes called Schur parameters and because of their physical interpretation in scattering applications they were also called reflection coefficients [130].

If we invert the algorithm after n steps and write Γ_0 as a function of Γ_{n+1} then it will give something of the form

$$\Gamma_0(z) = \frac{R_n(z) + S_n^*(z)\Gamma_{n+1}(z)}{S_n(z) + R_n^*(z)\Gamma_{n+1}(z)}$$

with $R_n, S_n \in \mathcal{L}_n$ and the superstar has the usual meaning. It should be clear that if we choose Γ_{n+1} arbitrarily in $\mathcal{B}$ then the right-hand side of this relation gives all possible functions in $\mathcal{B}$ that interpolate Γ_0 in $\{\alpha_1, \ldots, \alpha_n\}$. But, unlike in the Schur algorithm (see Section 4.4) these are not so directly related to the recurrence of the ORFs, although they are connected as explained in [453]. However if we include an intermediate step and for some arbitrary w we set

$$\Gamma_n'(z) = \frac{\Gamma_n(z) - \gamma_n}{1 - \overline{\gamma}_n\Gamma_n(z)}\frac{1}{\zeta_n(z)}, \quad \Gamma_{n+1}(z) = \frac{\Gamma_n'(z) - \rho_n}{1 - \overline{\rho}_n\Gamma_n'(z)}\frac{1}{\zeta_w(z)}, \quad \zeta_w(z) = \frac{z - w}{1 - \overline{w}z}, \ \rho_n = \Gamma_n'(w),$$

then the duality between this recursion and the recursion of the reproducing kernels given in Section 7.3 is much of the same style as it was explained in Section 4.4 for the Schur algorithm. See [105, Sec. 6.4], but also [101, 212, 213]. Many properties can be derived from the theory of continued fractions (the R_n and S_n are called rational *Wall functions*), but it also follows in a straightforward way from the special properties of the recursion matrices $t_n(z)$ and their products used in the recurrences of Section 7.3. These matrices are some matrix valued inner functions for an indefinite metric, as it will be explained in Section 15.3. For some recent convergence results see [43, 44, 373].

15 Matrix valued versions

So far, we have only considered scalar versions, but there is a very extensive literature for the block case. Matrix valued polynomials was already considered by Wiener and Masani [408, 560, 561] and it was picked up in applications [39, 205, 212, 213, 215] and rational generalizations soon appeared [74, 204, 206, 302], while at the same time functional analysis and operator theory considered related problems [13, 35, 41, 115, 168, 226, 231, 233, 238, 264, 329, 330]. And it is still an active research area for both theory and applications

See also [180] and its many references for a recent survey on orthogonal matrix polynomials.

15.1 Matrix valued ORFs

The Hilbert space has to be replaced by a Hilbert module. The measure μ is now a Hermitian positive definite measure. Its entries are complex valued measures. The scalar trace measure is defined as $\tau = \mathrm{tr}(\mu)$ and with μ'_τ the Radon-Nikodym derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\tau}$, we define a left and a right matrix valued inner product for $N \times N$ valued functions f and g as

$$\langle\!\langle f, g \rangle\!\rangle_\ell = \int f \mathrm{d}\mu g^H = \int f \mu'_\tau g^H \mathrm{d}\tau \quad \text{and} \quad \langle\!\langle f, g \rangle\!\rangle_r = \int f^H \mathrm{d}\mu g = \int f^H \mu'_\tau g \mathrm{d}\tau.$$

Note that $\mathrm{tr}\,\langle\!\langle f, g \rangle\!\rangle_\ell$ and $\mathrm{tr}\,\langle\!\langle f, g \rangle\!\rangle_r$ are genuine scalar inner products.

This left/right duality and their interplay is the main difference that will formally show in the relations and formulas that will generalize the scalar ones that we had before. We give some examples.

For the case of the real line, block versions are used in applications such as model reduction. For this we refer to the applications of Section 16. Most of the papers on a more pure mathematical side are however dealing with the unit circle. So let us also here consider the circle case to formulate the main ideas.

Consider the module $\mathcal{L}_n$ of complex rational $N \times N$ functions that is spanned by $\{\mathrm{B}_k\}_{k=0}^n$ with $\mathrm{B}_k(z) = \prod_{j=0}^k \zeta_j(z) I_N = B_k(z) I_N$ with B_k the usual scalar Blaschke products for a sequence of points $\{\alpha_j\} \subset \mathbb{D}$, $\alpha_0 = 0$ as used before. This set can be orthogonalized by some Gram-Schmidt type algorithm using $\langle\!\langle \cdot, \cdot \rangle\!\rangle_\ell$ to give an orthonormal set $\{\phi_j^\ell\}$ satisfying $\langle\!\langle \phi_i^\ell, \phi_j^\ell \rangle\!\rangle_\ell = \delta_{ij} I_N$. Similarly there is also a set of right ORFs: $\{\phi_j^r\}$. So there are left and right reproducing kernels. For example $k_n^\ell(z, w) = \sum_{k=0}^n [\phi_k^\ell(w)]^H \phi_k^\ell(z)$ is left reproducing since writing $f(z) = \sum_{k=0}^n \phi_k^\ell(z) a_k$, directly implies that $\langle\!\langle k_n^\ell(z, w), f(w) \rangle\!\rangle_\ell = f(w)$. These reproducing kernels also satisfy Christoffel-Darboux relations. For example

$$k_n^\ell(z, w) = \sum_{k=0}^n [\phi_k^\ell(w)]^H \phi_k^\ell(z) = \frac{[\phi_n^{r*}(w)]^H \phi_n^{r*}(z) - \zeta_n(z)\overline{\zeta_n(w)}[\phi_n^\ell(w)]^H \phi_n^\ell(z)}{1 - \zeta_n(z)\overline{\zeta_n(w)}}.$$

Of course, there is a right version as well. Like in the scalar case we may derive a recurrence relation for the kernels, and from there, the recurrence for the ORF can be derived:

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

$$\tilde{t}_n(z) = \frac{\sqrt{1 - |\alpha_n|^2}}{\sqrt{1 - |\alpha_{n-1}|^2}} \frac{1 - \overline{\alpha}_{n-1} z}{1 - \overline{\alpha}_n z} H(\upsilon_n) Z_{n-1}(z), \tag{15.1}$$

where υ_n is a (symmetrized) matrix valued Verblunsky coefficient. It is contractive, meaning that $\upsilon\upsilon^H \prec I_N$, i.e., $I_N - \upsilon\upsilon^H$ is positive definite. Although there are left and right versions of these coefficients, one may use the freedom in normalization to get only one central version, which is the one we are using here. The block matrix $H(\upsilon)$ denotes the so called *Halmos extension* of a contractive matrix υ defined by

$$H(\upsilon_n) = \begin{bmatrix} (I_N - \upsilon_n\upsilon_n^H)^{-1/2} & 0 \\ 0 & (I_N - \upsilon_n^H\upsilon_n)^{-1/2} \end{bmatrix} \begin{bmatrix} I_N & \upsilon_n \\ \upsilon_n^H & I_N \end{bmatrix},$$

and $Z_{n-1}(z) = \zeta_{n-1}(z)I_N \oplus I_N$.

With these ORFs $\phi_k^{\ell/r}$, we can again associate some matrix valued ORFs of the second kind $\psi_k^{\ell/r}$ and their ratios that will interpolate the Carathéodory function $\Omega(z) = \int D(t,z)\mathrm{d}\mu(t)$, with $D(t,z)$ again the Riesz-Herglotz kernel:

$$\Omega_n^\ell(z) = [\psi_n^{\ell *}(z)][\phi_n^{\ell *}(z)]^{-1} = \Omega_n^r(z) = [\phi_n^{r*}(z)]^{-1}[\psi_n^{r*}(z)]$$

and

$$\Omega(z) - \Omega_n^\ell(z) = \Omega(z) - \Omega_n^r(z) = zB_n(z)h(z),$$

with h analytic in $\mathbb{D}$.

These interpolation properties are related to quadrature as we have seen in the scalar case. However the quadrature aspects have not been explored much. Of course there are extra problems directly related to the fact that scalar values are replaced by matrices. A scalar value can be zero or not, while a matrix can be become singular without being zero.

All this and much more can be found in the papers mentioned above and in papers from the team in Leipzig [265, 268–272, 384, 387, 388] and others like [27, 257–261].

15.2 Nevanlinna-Pick interpolation

Much of the literature is lifting Schur (polynomial case) and Nevanlinna-Pick (rational case) interpolation problems to a matrix version. This inspired people to call this domain of research 'Schur analysis' which involves complex analysis, function and operator theory as well as systems theory and linear algebra. Much of this work appears in the book series *Operator Theory, Advances and Applications* by Birkhäuser.

The physical interpretation in scattering applications gave rise to the name reflection coefficients for the Verblunsky parameters υ_n [130] and the matrices $t_n(z)$ in the recurrence relation are called *chain scattering matrices* [211, 212].

The Nevanlinna-Pick interpolation problem is the inverse of what we described in the previous section. Now we are given a set of interpolation data $\{(z_k, y_k) : k = 0, 1, \ldots\}$ with the z_k points in $\mathbb{D}$ and the problem is to find $\Gamma \in \mathcal{B}$ such that $\Gamma(z_k) = y_k$, $k = 0, 1, \ldots$, or more precisely under what conditions, the problem has a solution, and characterize (i.e., parametrize) all solutions if they exist. This is equivalent with a rational moment problem and there is a solution if the Pick matrix for Schur functions with entries $\frac{1-\overline{y}_i y_j}{1-\overline{z}_i z_j}$ is positive definite. Constructing the Schur function from the interpolation data is somehow inverting the Nevanlinna-Pick algorithm, which is essentially the recurrence relation for the ORF or for the reproducing kernels. A matrix version is similar, except that the y_k are matrices. However, more general situations can be considered where also the scalar points $z_j \in \mathbb{D}$ are replaced by contractive matrices Z_j (i.e., their spectral radius is less than 1). If $\Gamma(z) =$

$\sum_{k=0}^{\infty} G_k z^k$, then the left value at Z is defined as $\Gamma(Z)_\ell = \sum_{k=0}^{\infty} Z^k G_k$ (and similarly for a right value). So the (left) matrix interpolation problem can be formulated as $\Gamma(Z_k)_\ell = Y_k$, $k = 0, 1, \ldots$.

A tangential version selects directions v_k and w_k, such that we require

$$v_k \Gamma(Z_k)_\ell = v_k Y_k \quad \text{or} \quad \Gamma(Z_k)_\ell w_k = Y_k w_k \quad \text{or} \quad v_k \Gamma(Z_k)_\ell w_k = v_k Y_k w_k, \quad k = 0, 1, 2, \ldots$$

or a combination of these, or even more generally, the v_k and w_k can be matrices as well. Original papers on the tangential problem are [245–248] and in a different version [455, 456]. For further generalizations and extensions we refer to the extensive literature. Some connection with a recent application using Krylov subspace techniques is given in the applications of Section 16.

15.3 J-inner functions and Nehari interpolation

Another direction of research is related to the recurrence of the ORFs. The matrix $\tilde{t}_n(z)$ that is used in the scalar recursion is special in the sense that it is analytic in $\mathbb{D}$ and it satisfies $\tilde{t}_n(z)^H J \tilde{t}_n(z) = J$ for $z \in \mathbb{T}$ and $\tilde{t}_{n*}(z) J \tilde{t}_n(z) \prec J$ for $z \in \mathbb{D}$ with the substar for a matrix function defined as $t_*(z) = [t(1/\overline{z})]^H$ and J the signature matrix $J = 1 \oplus -1 = \operatorname{diag}(1, -1)$. We say that the matrices $\tilde{t}_n(z)$ (and products of such matrices) are J-inner in $\mathbb{D}$. That means that they are J-contractive in $\mathbb{D}$ and J-unitary on $\mathbb{T}$. Such matrix functions were originally studied by Potapov [485] and they clearly play an essential role in the theory of ORF and rational interpolation problems in the complex plane. Also here matrix versions have been extensively studied and applied in interpolation problems related to ORFs. The generalization consists often in considering signature matrices J of the form $j_{pq} = \operatorname{diag}(I_p, -I_q)$. They are examples of matrix-valued inner functions. See e.g., [14, 29, 233, 368]. Since the recurrence for the ORF or for the Nevanlinna-Pick algorithm consists of a product of elementary J-inner factors, the product of the first n elementary factors, contains all the information used so far and one may in this way generate all solutions of the moment problem or of the Nevanlinna-Pick interpolation problem. They also show up in other rational approximation problems like Sarason interpolation [490] and Nehari interpolation [428]. Nehari interpolation means that one looks for solutions not necessary analytic but they allow a finite (usually small or minimal) number of poles in the domain of analyticity. Thereby one may minimize the number of poles for a given bound on the norm, or one may find the minimum norm with a given number of poles. If the Schur function is not analytic, the corresponding measure is indefinite and much of the previous theory may break down. The problem is important for finding Hankel-norm approximations (see also Section 16.2) as first proposed by Adamjan, Arov and Kreĭn [4, 5]. So there is an expensive bibliography on this topic that we will not include here since we are diverging from our original target of positive quadrature formulas.

Describing all the solutions of the moment problems and interpolation problems that we have mentioned is also related to completion problems (i.e., to find the missing elements in a partially prescribed j_{pq}-inner matrix) and extension problems (i.e., to extend a moment matrix with only a finite number of elements prescribed) to a larger, possibly infinite one with the same properties. Also these problems have a long history. Some recent publication is [163]. Many of the papers are still dealing with the polynomial and Laurent polynomial case like [40, 161, 162, 565, 566] and the papers in the book [15] where many more references can be found.

For these problems and the connection with Wold decomposition, the study of the shift operator, intertwining dilations, the commutant lifting theorem, and their applications in systems and control theory etc., we refer to the literature.

15.4 Refined versions

Starting with a space generated by the simple basis $\{\mathrm{B}_k\} = \{B_k I_N\}$ we basically get a 'non-commutative scalar version' leading to rational functions that can be written with one scalar denominator and a matrix valued numerator. It is a major challenge to unravel a much finer structure, in which an initial basis is used that consists of products of elementary factors of the form

$$Z_{i,k}(z) = \begin{bmatrix} I_{i-1} & & \\ & \zeta_k(z) & \\ & & I_{N-i} \end{bmatrix}, \quad \zeta_k(z) = s_k \frac{z-\alpha_k}{1-\overline{\alpha}_k z}, \quad s_k \in \mathbb{T}$$

or by even placing a factor ζ_k on an off-diagonal position. This is probably also related to a much finer structure that is possible for the recurrence relation matrices $\tilde{t}_n(z)$ that we described in (15.1). Indeed, it is shown in [233, Thm. 1.4] that any elementary J-inner matrix of size $N \times N$, introducing a pole in $\alpha \in \mathbb{C}$ must take one of the following forms (up to a constant unitary matrix factor)

$$I_N + \left(\frac{z-\alpha}{1-\overline{\alpha}z} - 1 \right) u(u^H J u)^{-1} u^H J, \quad u \in \mathbb{C}^N, \ s.t. \ \frac{u^H J u}{1-|\alpha|^2} > 0, \ \text{if } \alpha \in \mathbb{C} \setminus \mathbb{T},$$

or

$$I_N - \frac{c}{2} \frac{\alpha + z}{\alpha - z} u u^H J, \quad c > 0, \quad u \in \mathbb{C}^N, \ s.t. \ u^H J u = 0, \ \text{if } \alpha \in \mathbb{T},$$

or

$$I_N + (z^{-1} - 1) u (u^H J u)^{-1} u^H J, \quad u \in \mathbb{C}^N, \ s.t. \ u^H J u < 0, \ \text{if } \alpha = \infty.$$

Note that these are rank-one perturbations of I_N. Such recurrence matrices do not introduce an 'overall pole' but introduce a pole at a particular position in the rational matrix. It is not clear what this means in terms of ORF that are generated by this kind of recursion. These elementary matrices also appear is the Nevanlinna-Pick type pole-extraction algorithms of [30–34, 210, 211, 273, 274].

16 Some applications

To illustrate the use of ORF, and of rational quadrature we pick a couple of applications.

16.1 Rational Krylov methods

In several situations it is required to evaluate $f(A)v$ for a given vector v, some function f and a (usually very large but sparse) matrix A. This occurs when solving (time-dependent) differential equations, then $f(z)$ appears in the form of $\exp(z), z^{-1/2}, \cos(z), \ldots$. For example, take a simple diffusion problem $\dot{\boldsymbol{u}}(t) = -A\boldsymbol{u}(t)$, $\boldsymbol{u}(0) = \boldsymbol{u}_0$, then its solution is $\boldsymbol{u}(t) = \mathrm{e}^{-At}\boldsymbol{u}_0$. The matrix A is large if a fine space grid is used and sparse since the value at some point is only related to its immediate neighbors.

Another example can be found in the analysis of complex networks. The network is represented by a graph $G(V, E)$ with V the set of vertices and E the set of edges, the *adjacency matrix* A ($A_{i,j} = 1$ if there is an edge connecting vertex v_i to vertex v_j, and 0 otherwise). This matrix A is often huge and very sparse. Sparsity is caused by the so called *small world* property stating that it takes only a few edges to connect two arbitrary knots in the network. The number of triangles (3-cycles) that node v_i participates in is $\frac{1}{2}[A^3]_{ii} = \frac{1}{2}\mathbf{e}_i^T A^3 \mathbf{e}_i$. The *centrality* of a node is a way to measure its importance. Google's PageRank is a variant of eigenvector centrality for a directed graph consisting of links on web pages [385]. It is the eigenvector of the adjacency matrix corresponding to the (largest) eigenvalue 1. Its ith entry is the relative importance of node i in the network. This is an alternative for the ranking by *Katz centrality* [369] given by the entries of the vector $\mathbf{x}$ solving $(I - cA)\mathbf{x} = \mathbf{1}$ with $\mathbf{1}$ a vector of all ones and c some number, or one may rank *hubs* (many outgoing links pointing to authorities) and *authorities* (many incoming links from hubs) in which case we can solve $(I - cA)\mathbf{x} = \mathbf{1}$ and $(I - cA^T)\mathbf{y} = \mathbf{1}$. Note that this corresponds to taking row or column sums of $(I - cA)^{-1}$. But there are other measures of centrality ranking by taking row or column sums of other matrix functions like for example $\mathrm{e}^A = I + A + A^2/2! + \cdots$ or just take the diagonal elements. Hyperbolic functions play a role in *returnability* of digraphs [241]. *Communicability* measures how easy it is to send a message from node v_i to node v_j and it is given by the (i, j) entry of e^A [240] giving an average for v_i as the ith component of the vector $\mathrm{e}^A\mathbf{1}$. Also heath equations and even Schrödinger equations on graphs rely on the evaluation of matrix functions [242, 506].

Another approach that also applies to operators is when the function $f(A)$ has a Cauchy integral representation

$$f(A) = \int_\Gamma f(z)(zI - A)^{-1}\mathrm{d}\mu(z), \quad \Gamma = \partial\Omega$$

where $\Omega \subset \mathbb{C}$ contains the spectrum of A and f is supposed to be analytic in the closure $\hat{\Omega} = \Omega \cup \partial\Omega$. This makes the link with quadrature formulas and the rational approximation of Cauchy transforms of a function and moment problems even more explicit. But whatever the situation, the idea explained below remains essentially the same.

Thus in many situations it is important to evaluate $f(A)v$ or possibly $u^T f(A)v$ with u and v vectors and A a very large matrix that is often Hermitian or unitary.

Suppose for simplicity that $A \in \mathbb{C}^{N\times N}$ is Hermitian positive definite and we want to evaluate $v^H f(A)v$.

For example suppose $v = \mathbf{e}_i$, i.e., we want to compute the diagonal element $[f(A)]_{ii}$. Off-diagonal elements $[f(A)]_{ij}$ can be computed in a similar way using a nonsymmetric Lanczos algorithm.

Clearly, if the eigenvalue decomposition of A is $A = U_N \Lambda_N U_N^H$, $U_N^H U_N = I_N$ then

$$v^H f(A) v = v^H U_N f(\Lambda_N) U_N^H v = \mathfrak{I}_\mu[f] = \sum_{k=0}^{N-1} f(\lambda_{N,k}) w_{N,k}, \quad w_{N,k} = |u_k|^2$$

where $v^H U_N = [u_0, \ldots, u_{N-1}]$ and $\mu = \sum_{k=0}^{N-1} w_{N,k} \delta_{\lambda_{N,k}}$. Since N is assumed to be very large, we may consider $\mathfrak{I}_\mu[f]$ as an 'integral' with a discrete N-point measure that has almost infinitely many points in its support.

Since A is too large for evaluating $f(A)$, we project it onto a smaller dimensional subspace. Therefore define a space on which to project as the column space of the matrix $Q_n \in \mathbb{C}^{N \times n}$ with $Q_n^H Q_n = I_n$. Hence $f(A) \approx Q_n f(A_n) Q_n^H$ with $A_n \in \mathbb{C}^{n \times n}$ is a much smaller matrix given by $A_n = Q_n^H A Q_n$. Suppose the eigenvalue decomposition is $A_n = \tilde{U}_n \tilde{\Lambda}_n \tilde{U}_n^H$. Thus $f(A) \approx \hat{U}_n f(\tilde{\Lambda}_n) \hat{U}_n^H$ with $\hat{U}_n = Q_n \tilde{U}_n$ and $\hat{U}_n^H \hat{U}_n = I_n$. Thus the N-point 'integral' $v^H f(A) v$ is approximated by an n-point quadrature formula

$$\mathfrak{I}_n[f] = \sum_{k=0}^{n-1} f(\tilde{\lambda}_{nk}) \tilde{w}_{n,k}, \quad \tilde{w}_{n,k} = |\tilde{u}_k|^2$$

with $v^H \hat{U}_n = [\tilde{u}_0, \ldots, \tilde{u}_{n-1}]$. The $\tilde{\lambda}_{n,k}$ are called *Ritz values* that approximate the eigenvalues $\lambda_{N,k}$ of A.

That the QF is a (rational) Gauss formula is hidden in the orthogonality of the columns in the matrix Q_n that defines the projection. If we want to project onto the *Krylov space* $\mathcal{K}_n(A, v)$ spanned by $\{v, Av, \ldots, A^{n-1}v\}$, then the columns of Q_n form an orthogonal basis for this space and these are given by $\{\phi_k(A)v\}_{k=0}^{n-1}$ where the ϕ_k are orthogonal polynomials. Orthogonality is with respect to the linear functional defined by the moments $m_k = v^H A^k v$, $k = 0, 1, \ldots$. This corresponds to taking $f(x) = x^k$, $k = 0, 1, 2, \ldots$. Clearly $Q_n^H Q_n = I_n$, but the recurrence relation $x\boldsymbol{\phi}_n(x) = \boldsymbol{\phi}_n(x) J_n + [0, \ldots, 0, *]$ with J_n the truncated tridiagonal Jacobi matrix, translates as $AQ_n = Q_n J_n +$ rank one. Thus in the polynomial case, $A_n = J_n$ is the restriction of A to the Krylov subspace $\mathcal{K}_n(A, v)$ and the Ritz values are the eigenvalues of J_n, which are the zeros of the OP ϕ_n. Hence the QF $\mathfrak{I}_n$ is a Gaussian QF implying that it will be exact for all $f \in \mathcal{P}_{2n-1}$. In practice one may then use the symmetric Lanczos algorithm to compute the Jacobi matrix

$$J_n = \text{tridiag} \begin{pmatrix} a_1, a_2, \ldots, a_{n-1} \\ b_1, b_2, \ldots, b_{n-1}, b_n \\ a_1, a_2, \ldots, a_{n-1} \end{pmatrix}.$$

In its strip-down form, it computes the recursion

$$\begin{aligned} &a_j q_j = y_j = (A - b_j I) q_{j-1} - a_{j-1} q_{j-2}, \quad j = 1, 2, \ldots \\ &b_j = q_{j-1}^T A q_{j-1}, \\ &a_j = \|y_j\| \end{aligned}$$

with initial vectors $q_{-1} = 0$ and $q_0 = v$. For $n \ll N$, the spectrum of J_n is an approximation for the spectrum of A, and computing $f(A)$ boils down to essentially computing $f(J_n)$, which is a much smaller and simpler problem.

Off-diagonal elements $[f(A)]_{ij}$ can be computed in a similar way using a nonsymmetric Lanczos algorithm.

This illustrates that the quadrature formulas are indeed underlying these well known algorithms, but they are somewhat hidden in the iterative algorithms of linear algebra. So, all the variants (like Gauss-Lobatto and Gauss-Radau and rational versions) are applicable and can be translated in terms of linear algebra [150]. As we have seen, quadrature formulas are associated with interpolation and one may also consider the previous techniques as variants of standard Krylov methods that approximate the function f by a Padé approximant.

The n Ritz values, i.e., the eigenvalues of J_n approximate in a sense the N eigenvalues of A. The previous algorithm corresponds to the power method (multiply a vector v with powers of A) for eigenvalue computation and we will therefore have convergence to the dominant eigenvalue(s) of A. Indeed the Rayleigh quotient $\lim_{k\to\infty} v^H A^k v/\|v\|^2$ will converge to the dominant eigenvalue of A if there is a dominant one.

In the case of a non-Hermitian matrix, the technique is the same, except that, since we loose symmetry, the Jacobi matrix has to be replaced by a less sparse Hessenberg matrix. This is also known as the *Arnoldi* algorithm. Of course we restrict ourselves here to the basic facts, and for practical implementation much care has to be taken to keep complexity under control and at the same time deal with stability issues. Here we refer to the extensive literature in iterative methods in linear algebra.

To approximate a non-dominant eigenvalue (e.g., the one(s) near α), 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 α_k and diminish the influence of eigenvalues in the neighborhood of $1/\overline{\beta}_k$.

Obviously, an orthogonal basis can now be obtained by a set of ORFs ϕ_k instead of OPs. So, we define as before $b_0 = 1$ and

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

Note that if A is a square 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.

The rational Krylov space is spanned by $\{v, b_1(A)v, \ldots, b_{n-1}(A)v\}$. An orthogonal basis is given by the columns q_k of the matrix Q_n where $q_k = \phi_k(A)v$, $k = 0, \ldots, n-1$ with ϕ_k the kth ORF for $\mathcal{L}_n$. Clearly $Q_n^H Q_n = I_n$ and $A_n = Q_n^H A Q_n$.

Assuming regularity of the ORFs, they satisfy a recurrence relation corresponding to $x\boldsymbol{\phi}_n(x)(I + D_n^{-1}J_n) = \boldsymbol{\phi}_n(x)J_n + [0, \ldots, 0, *]$, which translates into (replace x by A and multiply by v):

$$A[\phi_0(A)v, \ldots, \phi_{n-1}(A)v](I_n + D_n^{-1}J_n) = [\phi_0(A)v, \ldots, \phi_{n-1}(A)v]J_n + \text{rank } 1$$

or

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

where $D_n = \text{diag}\,(\alpha_0, \ldots, \alpha_{n-1})$ and J_n is a tridiagonal matrix. Then $A_n = Q_n^H A Q_n$ is the restriction of A to the rational Krylov subspace spanned by the columns of Q_n. The previous relation implies that the eigenvalues of A_n are obtained by solving the generalized eigenvalue problem for the pencil $(I_n + D_n^{-1}J_n, J_n)$. If the eigenvalues are $\{\lambda_{n,k}\}$ then this corresponds to

approximating the discrete integral with respect to μ with N mass points at the eigenvalues of A by a rational Gauss quadrature formula $\mathfrak{I}_n$ whose nodes are the eigenvalues of A_n. The QF will be exact for all f in $\mathcal{L}_{n,n-1} = \mathcal{L}_n \cdot \mathcal{L}_{n-1}$ where the $\mathcal{L}_k$ are the rational function spaces with poles among $\alpha_1, \alpha_2, \ldots$. It might be needed to use a QORF if the zeros of the numerator of the ORF coincides with one of the α_k, or in practice to avoid numerical problems when they are close. In that case, the formula will be only exact for $f \in \mathcal{L}_{n-1,n-1} = \mathcal{L}_{n-1} \cdot \mathcal{L}_{n-1}$.

These rational Gauss QFs approximates the original integral by a discrete measure with nodes at the zeros of the ORFs, which are approximating eigenvalues of A that are 'near' the α_k. In the classical Gauss case, the ORFs are polynomials and all the α's are at ∞, hence the largest eigenvalues of A attract the eigenvalues of J_n, which are the nodes of the Gauss QF.

For further details about how to compute the coefficients c_k and d_k of the recurrence and about the choice of the α_k (one has to avoid 'forbidden poles') consult [193]. About the speed of convergence for the different eigenvalues, an analysis based on potential theory see [51, 52, 159]. Earlier references are [239, 425], and about the choice of the poles there are several proposals in the recent literature [53, 228, 229, 318, 320, 424]. As a general introduction to the subject one may consult [338]. An excellent survey with the state of the art is [319].

One more example is the *Hyperlink-Induced Topic Search* (HITS) algorithm introduced by Kleinberg in 1999 [376]. Given the directed graph of the network one wants to find a ranking of hubs and authorities. Finding the ranking vectors corresponds to computing the eigenvectors for the dominant eigenvalue $\lambda = 1$ for $A^T A$ and AA^T which are the authorities and the hubs matrices. This requires to iterate $u_k = \mathcal{A}^2 u_{k-1}$ where the ranking vectors $u_k = [y_k^T, x_k^T]^T$ are combining the ranking of the hubs y_k and the ranking vector of the authorities x_k while

$$\mathcal{A}^2 = \begin{bmatrix} AA^T & \\ & A^T A \end{bmatrix}.$$

See [55] for an alternative.

So far we only considered evaluating $v^H f(A)v$, but to evaluate the vector $f(A)v$ or $\mathrm{tr}[f(A)]$ one may find combinations of (rational) Gauss-type quadrature formulas and Lanczos, Arnoldi or other Krylov subspace methods to deal with these problems. There is a very flourishing literature on this subject for which we refer to the literature where new papers on the topic are appearing almost every day. The so called *extended Krylov* methods use the Krylov space $\mathcal{K} = A^{-m_2}\mathcal{K}_{m_1+m_2}(A, v)$ with $m_2 \ll m_1$. The reason is that then only m_2 inverts of A are needed and m_1 matrix vector multiplications. This corresponds to a Gauss quadrature based on Laurent orthogonal polynomials [218, 227, 348, 350]. Obviously a rational variant is possible where one considers a space of rational functions $\mathcal{L}_n$ as considered before but m_1 of the poles are chosen to be ∞ and only $m_2 \ll m_1$ are finite, requiring a shift and invert, but also in the positive powers, forming the polynomial part, one may replace a factor x by a more general factor $x - \beta$ where the parameter β can be used to diminish the influence of the eigenvectors for eigenvalues near β in the Krylov subspace that is constructed. Note that this corresponds to consider the sum of a polynomial space $\mathcal{P}_{m_1}$ and some proper rational space $\mathcal{L}_{m_2}$. Such quadrature formulas have been discussed by Gautschi [285, 288]. To solve nonlinear eigenvalue problems, Hermite interpolation has been used with basis functions of the general form

$$b_k(z) = \frac{1}{\beta_0} \prod_{j=1}^{k} \frac{z - \gamma_{j-1}}{\beta_j(1 - z/\xi_j)}, \quad k = 0, 1, 2, \ldots$$

have been used e.g., in [54].

In some applications, the vector v has to be replaced by a matrix $V \in \mathbb{C}^{N\times m}$. This gives a block version of rational Krylov method in which the subspace considered is the range of the matrix $[V, b_1(A)V, \ldots, b_n(A)V] \in \mathbb{C}^{N\times(n+1)m}$. Since the columns may not all be independent, and also to reduce the complexity, it might be useful to select in each iteration one vector in the column space of V, say Vd_i with $d_i \in \mathbb{C}^{m\times 1}$, so that the projection will be on the subspace range$[Vd_0, b_1(A)Vd_1, \ldots, b_n(A)Vd_n]$. This requires in each iteration not only the selection of a pole α_i, but also of a vector d_i. See [230].

What has been described refers to the real line case, and this is also dominantly represented in the literature. However, one can imagine that there is an analog theory that can be developed not for Hermitian but for unitary matrices, which would correspond to the case of the unit circle.

16.2 Model reduction

With the previous examples we are very near the methods applied in model reduction for (linear) dynamical systems.

Here the system can have a continuous time or a discrete time formulation corresponding to an analog or a digital system. In dynamical systems, the input and output depends on a time variable which takes values in the time domain and that can be either $\mathbb{R}$ in the analog case or $\mathbb{Z}$ in the digital case. The problem in the time domain can be transformed to the frequency domain. In the analog case this is obtained by a Laplace transform and the frequency domain is the imaginary axis. In the digital case one uses the z-transform (or Fourier transform) giving the unit circle as frequency domain. To stay on the same track as in the previous subsection, let us consider the continuous time systems.

The state space description of the system with m inputs and p outputs is given by

$$\begin{array}{l}\dot{x}(t) = Ax(t) + Bu(t)\\ y(t) = Cx(t) + Du(t)\end{array} \quad \text{or} \quad \begin{array}{l}zX(z) = AX(z) + BU(z)\\ Y(z) = CX(z) + DU(z)\end{array} \quad \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathbb{C}^{(N+p)\times(N+m)}$$

where $y \in \mathbb{C}^{p\times 1}$ is the output, $u \in \mathbb{C}^{m\times 1}$ is the input, and $x \in \mathbb{C}^{N\times 1}$ is the state, all depending on time $t \in \mathbb{R}$. The dot denotes the derivative (with respect to time). We consider the stationary case, i.e., when the matrices A, B, C, D are constant. In the time-varying case, these also depend on time. The vectors denoted with capitals are the Laplace transforms of the time dependent vectors.

The transfer function $H(z)$ in the relation $Y(z) = H(z)U(z)$ is given by $H(z) = C(zI - A)^{-1}B + D$. Even with N as small as possible, this representation is not unique. One could use a change of basis T in the state space resulting in an equivalent representation $(A', B', C', D') = (T^{-1}AT, T^{-1}B, CT, D)$. Obviously, the eigenvalues of A are the poles of the system. Note however that unlike our example in the previous section, the matrix A need not be symmetric. The idea is to use again projections to bring the large problem down to a smaller problem in a subspace. So for this projection operation we choose $P_n, Q_n \in \mathbb{C}^{N\times n}$ such that $P_n^H Q_n = I_n$ and setting $(A_n, B_n, C_n, D_n) = (P_n^H AQ_n, P_n^H B, CQ_n, D)$, this gives a 'small' system with transfer function $H_n(z) = C_n(zI_n - A_n)^{-1}B_n + D_n$ that approximates $H(z)$ in some sense. It seems a good idea to keep in this reduced system the dominant eigenvalues of A. Note that $H(z) = D + \sum_{k=1}^{\infty} CA^{k-1}Bz^{-k}$. The matrices $M_k = CA^kB \in \mathbb{C}^{p\times m}$ are moments, usually called *Markov parameters* in this context. The form of the Markov parameters (let us assume $D = 0$) suggests the use of Krylov spaces called *controllability*

and *observability* subspaces, which we shall denote as the column space of the matrix $\mathcal{C}$ and the row space of the matrix $\mathcal{O}^H$ respectively

$$\mathcal{C} = [B, AB, AB^2, \ldots], \quad \mathcal{O}^H = [C^H, A^H C^H, \ldots].$$

So

$$X(z) = \boldsymbol{C}(z)U(z), \quad \boldsymbol{C}(z) = (zI - A)^{-1}B = \textstyle\sum_{k\geq 1} A^{k-1}Bz^{-k}$$
$$Y(z) = \boldsymbol{O}(z)BU(z), \quad \boldsymbol{O}(z) = C(zI - A)^{-1} = \textstyle\sum_{k\geq 1} CA^{k-1}z^{-k}.$$

For simplicity we consider $p = m = 1$. The norm in which to approximate could be H^∞, or H^2, or some particular norm in between namely the *Hankel norm*. The moment matrix has a Hankel structure with entries M_k, $k = 0, 1, \ldots$, hence it is given by $\mathcal{OC}$. It is the matrix representation of the map from past inputs to future outputs. Its rank equals the degree N of the transfer function, i.e., the order of the system. The Hankel norm approximant will minimize the spectral norm of the Hankel matrix for $H - H_n$, thus we want to approximate a Hankel matrix of rank N by a Hankel matrix of lower rank n. This is an old idea that stems from Adamyan, Arov and Kreĭn [3, 4] and it is sometimes called AAK approximation. The latter is well approximated by *truncated balanced realization* [58, 299, 341]. The matrix $G_c = \mathcal{C}\mathcal{C}^H$ is called the *controllability Grammian* and $G_o = \mathcal{O}^H\mathcal{O}$ is called the *observability Grammian*. They satisfy the Lyapunov equations

$$AG_c + G_cA^H = BB^H \quad \text{and} \quad A^HG_o + G_oA = C^HC$$

for which efficient solvers exist. The Hankel matrix is $\mathcal{OC}$ and therefore the singular values of this matrix are the square roots of the eigenvalues of G_cG_o.

By an equivalence transformation T of the state space, it is possible to simultaneously diagonalize the two Grammians, such that $G_o = \Sigma = G_c$ with Σ diagonal. This is called a *balanced realization*. The Σ contains the singular values of the Hankel matrix. By keeping the largest n values in Σ, say Σ_n and reconstructing the reduced system gives the state space realization of the best Hankel norm approximant.

There exist many variations on this idea. The method will guarantee stability of the reduced system if the original system was stable, however the method is computationally intensive, there are still numerical issues, and the singular values of the system should decay sufficiently fast to be able to decide where to cut the matrix Σ. So there is still room for moment matching and rational interpolation.

The transfer function can be seen as a function $f(A)$ and thus computing the kth component of the output Y for some input U corresponds to evaluating a vector of the form $u^Hf(A)v$ where $f(A) = H(z)$, $u = e_k$ and $v = U(z)$. Recall from our previous section that evaluating $u^Hf(A)v$ corresponds to evaluating a (discrete) integral and interpolation corresponds to (rational) moment matching, and thus to evaluating a (rational) quadrature formula that approximates the integral. Thus all the variants and knowledge of quadrature formulas can and will be (implicitly) applied.

Interpolation techniques can be of Padé type, matching the initial Markov parameters (called the partial realization problem [339, 366, 367]). Also two-point Padé approximation has been used to approximate the transfer function both near ∞ and at the origin. This corresponds to matching simultaneously the transient behavior and the steady state of the system. See for example [131] for a survey of the literature at that time. Since then the emphasis of the literature shifted towards approximating large scale systems and the application of Krylov type methods. For example if the projections P_n, Q_n mentioned above

satisfy some (bi)orthogonality condition $P_n^H Q_n = I_n$, then if

$$\begin{aligned}&\text{colsp } Q_n = \text{colsp } [B, AB, A^2B, \ldots, A^{n-1}B] \text{ and}\\ &\text{colsp } P_n = \text{colsp } [C^H, (CA)^H, (CA^2)^H, \ldots, (CA^{n-1})^H],\end{aligned}$$

one gets a reduced system that matches the Markov parameters CA^kB for $k = 0, 1, \ldots, 2n-1$, which corresponds to a Gauss QF and the problem is best solved using Lanczos-type methods.

However, even if the given system is stable, the approximant need not be stable. So, also here, there are still some issues concerning convergence and stability of the resulting approximation that have to be solved. That observation holds for all interpolation based methods, also the ones below, but in most practical situations these methods are fast and work well.

To switch to the use of rational functions instead of polynomials, i.e., use rational Krylov methods, one needs moments of the form $C(\alpha_k I - A)^{-i}B$ for a set of points α_k and $i = 0, 1, \ldots, i_k$. See e.g. [38, 263, 276, 548]. Then, provided $(P_n^H Q_n)$ is invertible, the reduced system

$$\begin{bmatrix} \tilde{A}_n & \tilde{B}_n \\ \tilde{C}_n & \tilde{D}_n \end{bmatrix} = \begin{bmatrix} (P_n^H Q_n)^{-1} P_n^H & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} \begin{bmatrix} Q_n & 0 \\ 0 & I \end{bmatrix}$$

will match the (rational) moments used if the columns of these P_n and Q_n span the proper Krylov subspaces. Rational Lanczos techniques that are far more complicated than described above can be used.

There is also the tangential idea that can be exploited. Again this idea is around for a while already [12, 16, 42, 372, 480] in block versions of the scalar theory of the previous chapters. See also the papers in the book [235] and more recently [33, 234]. Consider the approximation error $E_n(z) = H(z) - H_n(z)$ for the reduced transfer function. If we interpolate, then E_n (and maybe some of its derivatives in case of confluent interpolation points) will be zero at the interpolation points α_k. This can be relaxed by selecting directions $v(z)$ and $w(z)$ at the point α_k and require that we have some conditions like

$$E_n(\alpha_k)v_k = O(z-\alpha_k)^{i_k}, \quad w_k^H E_n(\alpha_k) = O(z-\alpha_k)^{i_k}, \quad \text{or} \quad w_k^H E_n(\alpha_k)v_k = O(z-\alpha_k)^{i_k}.$$

On tangential approximation see [41] and in the case of interpolation using Krylov subspaces and generalized Sylvester equations see [277]. There is a compact notation that can be used for interpolation in the direction v at α. The condition $[H(z) - H_n(z)]v(z) = O(z-\alpha)^k$ can be realized using an inclusion of a Krylov space into colsp Q_n, i.e.,

$$\text{colsp } [(\alpha I - A)^{-1}B, \ldots, (\alpha I - A)^{-k}B] \begin{bmatrix} v_0 & \cdots & v_{k-1} \\ & \ddots & \vdots \\ & & v_0 \end{bmatrix} \subset \text{colsp } (Q_n).$$

which is equivalent to solving the equation $AQ_n - Q_n A' = BB'$ with

$$A' = \alpha I_k - \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}, \quad B' = [v_0, \ldots, v_{k-1}].$$

Finally we mention that there is also an H^2 error that can be minimized. To illustrate it for a discrete time system we shall more extensively discuss it in the next subsection, but it can be applied to continuous time systems as well.

16.3 Linear prediction and system identification

Another area where the ORFs are often used is signal processing and time series analysis.

Since the previous paragraph was mainly discussing the case of the real line, let us now take the discrete time as an example to illustrate the idea on the unit circle.

Suppose we observe a signal which may be the output of a (linear) system via measurements of its spectrum. Thus if $y = \{y_k\}$ is the signal, then its spectrum is $Y(z) = \sum_k y_k z^k$, which for $z = e^{\mathrm{i}\theta}$ can be considered to be a Fourier series. The power spectrum is $|Y(z)|^2$. If this y is the output of a linear system, then we may write $Y(z) = H(z)U(z)$ where U represents the spectrum of the input and H is the transfer function. This product in the frequency domain corresponds to a convolution in the original time domain.

One of the possible problems is *linear prediction* theory or *time series analysis*, of which the ideas go back to Wiener and Szegő [316, 559–561], 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 sample given the previous inputs, i.e., predict the near future from the (infinite) past:

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

where[11] $P(z) = \sum_{j=0}^{\infty} a_j z^j$ with $a_0 = 1$ so that z acts as a backward shift operator: $zy_k = y_{k-1}$ because $zY(z) = \sum_{k\in\mathbb{Z}} y_{k-1} z^k$. The prediction error is $e_k = y_k + \sum_{j=1}^{\infty} a_j y_{k-j}$ If $e = (e_k) \in \ell_2$, then minimizing $\|e\|^2$ is by Parseval equivalent with minimizing $\|E\|^2 = \int_{\mathbb{T}} |P(z)Y(z)|^2 \mathrm{d}\lambda(z)$, or with $\mathrm{d}\mu(z) = |Y(z)|^2 \mathrm{d}\lambda(z)$ (λ is the Lebesgue measure on $\mathbb{T}$), we have to minimize the μ-norm of $P(z) \in H^2$, given the fact that its constant term is 1.

When the signal Y is considered to be the transfer function of a system, i.e., $Y(z) = H(z)U(z)$ with $U(z) = 1$, then the ideal filter P is perfectly whitening, i.e., it extracts all the information from the signal Y, leaving a perfectly flat spectrum for the error. Thus ideally $H = 1/P$, which is the model that generates the signal Y. Thus we may consider this as a spectral factorization problem: given a positive measure μ, find an outer spectral factor H such that $\mu' = |H|^2$. Thus it is essentially the same as the prediction problem where one wants to find $P = 1/H$. The spectral factorization problem was considered by Szegő and supposing $\log \mu' \in L^1$ the solution is $H(z) = c \exp\{\frac{1}{2} \int D(t, z) \log \mu'(t) \mathrm{d}\lambda(t)\}$ with $D(t, z)$ the Riesz-Herglotz kernel and $c \in \mathbb{T}$. See Section 14.1.3. We have seen before that σ will be asymptotically approximated by a scaled version of ϕ_n^* (see Section 14.1.3 and Theorem 14.6).

The infinite dimensional problem is solved recursively and one considers finite dimensional subspaces of H^2 to solve the problem. If this subspace is $\mathcal{P}_n$, then the solution is known to be φ_n^*, the reciprocal of the monic orthogonal polynomial. It holds that (assuming $\int \mathrm{d}\mu = 1$) $\min_{p_n \in \mathcal{P}_n} \|1/H - p_n\|_\mu = \min_{p_n \in \mathcal{P}_n} \|1 - p_n H\|_\lambda$ has the solution $p_n = \varphi_n^*$ [73, Thm. 3.4].

To see under what conditions we will find P as $n \to \infty$ one has to analyze the asymptotics of the φ_n^*. Convergence can be rather slow though, especially when there is a pole close to $\mathbb{T}$. Suppose $P(z) = 1/(1 - \overline{\alpha}z) = \sum_{k=0}^{\infty} (\overline{\alpha}z)^k$. Then, if $|\alpha|$ is close to 1, this will converge very slowly. Therefore it may be much wiser to consider subspaces of rational functions $\mathcal{L}_n$ where even a rough estimate of the poles close to $\mathbb{T}$ (and these show prominently as sharp peaks in the power spectrum) can be fixed in advance, and the convergence will be much faster. In

[11] With this convention, stability means that the series has to converge inside $\mathbb{D}$, thus that P is analytic in $\mathbb{D}$.

the rational case, the solution to the problem $\min_{p_n \in \mathcal{L}_n} \|1/H - p_n\|_\mu = \min_{p_n \in \mathcal{L}_n} \|1 - p_n H\|_\lambda$ is given by $p_n(z) = k_n(z, 0)$, where $k_n(z, w)$ is the reproducing kernel for $\mathcal{L}_n$ [73, Thm. 3.5], [105, Thm. 6.3.1].

These ideas are closely related to Nevanlinna-Pick interpolation theory and there is an excessive literature on the topic. Early references are [212, 213, 233]. See also the books [6, 11, 15, 41]. The polynomial situation has even been generalized to the time-varying case [214].

A somewhat different approach is taken in system identification and model reduction. See for example the books [337, 481] and their references like [7, 8, 76, 187, 336, 434–437, 507, 551–553]. Suppose that we have observations in the frequency domain of input and output, possibly corrupted by noise for the system $Y(z) = H(z)U(z)$. This means that we have measurements of $(U(z_k), Y(z_k))$ thus also of $H(z_k) = Y(z_k)/U(z_k)$ for a number of points $\{z_1, \ldots, z_N\} \subset \mathbb{T}$. We want to approximate H by a rational function $H_n \in \mathcal{L}_n$ with $n \ll N$. Because the transfer function should be stable, then in our convention, it has to be analytic *inside* $\mathbb{D}$.

Suppose we have some weights $|\omega_k|^2$ associated with the measurements (e.g., based on the noise statistics). Then we want to model the transfer function H by a low degree rational function H_n by minimizing the discrete least squares error

$$\|H - H_n\|_v^2 = \sum_{k=1}^{N} |\omega_k|^2 |H(z_k) - H_n(z_k)|^2.$$

If the observed Y is corrupted by noise V and we actually measure $Y = HU + V$, then we just want to find H_n such that $Y_n = H_n U + V$ will minimize $\|Y - Y_n\|_\omega = \|(H - H_n)U\|_\omega = \|H - H_n\|_{\tilde{w}}$ where $\tilde{w}_k = \omega_k U(z_k)$.

To summarize, we have to solve a least squares problem, approximating H by $H_n \in \mathcal{L}_n$ in a Euclidean space with inner product $\langle f, g \rangle = \int \overline{f(e^{\mathrm{i}\theta})} g(e^{\mathrm{i}\theta}) \mathrm{d}\mu(e^{\mathrm{i}\theta})$ with $\mu(t) = \sum_{k=1}^{N} |\tilde{w}_k|^2 \delta_{z_k}$. We want to approximate it by a (rational) Szegő quadrature formula with measure $\mu_n(t) = \sum_{k=1}^{n} |w_k|^2 \delta_{t_k}$. The inner product, and hence the norm will be the same in $\mathcal{L}_{n-1}$ for the measure μ and the measure μ_n. Thus the least squares approximant is $H_n(z) = \sum_{k=0}^{n-1} a_k \phi_k(z) \in \mathcal{L}_{(n-1)}$ with $a_k = \langle \phi_k, H \rangle$ if the ϕ_k are the orthonormal rational functions. Thus the coefficients a_k will depend on the choice of the poles, and are easily computed once the poles are known. Thus minimizing the norm $\|H - H_n\|^2$ can be seen as a least squares problem with the α_k as the nonlinear parameters: $J(\alpha_1, \ldots, \alpha_n) = \|H - H_n(\alpha_1, \ldots, \alpha_n)\|^2$. Because the measure is discrete, and the points are all on $\mathbb{T}$ we can make use of Discrete Fourier Transforms (DFT) and it is therefore feasible to 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 (see [541]) to compute a lower order model. The nonlinear problem is however notoriously ill conditioned and one should also take care that the poles α_k do not leave the unit disk to guarantee stability. It is a drawback of these methods that even if the given system is stable, it can not be guaranteed that the best reduced approximant will be stable. So the latter may only work in a limited set of cases. A comparison with the linearized rational approach mentioned in Section 12 can be found in [135].

Another way to solve the problem is to tackle it as an inverse eigenvalue problem that we have explained before. We indeed know the N nodes and weights, but it might be interesting to use only $n \ll N$ of these which can be selected and for which we can recursively compute the best least squares solution.

In classical papers solving this problem, the poles were fixed in advance and 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 $\mathbb{T}$ [391]. Later one used the *Kautz basis* which involved not just one α, but a pair of complex conjugate poles $(\alpha, \overline{\alpha})$ [371]. A unifying theory was given in [435]. Then in [344] 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, Wahlberg, Gustafsson, Heuberger, and many more. See for example the books [337, 481] and their references. For example [7, 8, 76, 187, 336, 434–437, 507, 551–553].

All these use however orthogonality with respect to the continuous Lebesgue measure. Thus, 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.

Heuberger et al. also used state space realizations for the rational approximants in the case of cyclically repetition of a finite set of poles (they call it a *Hambo basis*, see also [551]). This kind of realization is particularly interesting when the input and output are vectors (MIMO systems). It is then describes in discrete time as

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

In this approach the state space representation grows in size as the degree of the approximant increases (see [540]). So in the Hambo basis it is assumed that there are only a finite number n of poles and that these are repeated over and over. This allows them to keep the dimension of the matrices under control. There is some link with the balanced realization. Indeed, the vector valued function $\boldsymbol{C}(z) = (zI - A_n)^{-1}B_n$ is constructed such that its entries are the first n ORF, and therefore the Grammian $G_c = \mathcal{C}\mathcal{C}^H$ will be diagonal, giving a balanced realization. This only works because in this case the ORFs are formed by the Takenaka-Malmquist basis and ϕ_k for $k \geq 1$ all have a zero at the origin.

If we want to minimize the L^2 norm w.r.t. the Lebesgue measure for a very large system, then one may have a look at the Iterative Rational Krylov Algorithm (IRKA) proposed in [318]. It has been mentioned in the previous subsection already. It was first proposed by Wilson in 1970 [563], and it has been revised by many authors since e.g. [346] and recently [138, 323, 537, 538].

In the multivariate case, the norm to minimize is $J = \operatorname{tr} \int E_n(z) E_n(z)^H \mathrm{d}\lambda(z)$ with $E_n = H - H_n$ the approximation error of the reduced transfer function H_n [24]. This is best formulated in a state space form for an extended system:

$$(\hat{A}, \hat{B}, \hat{C}) = \left(\begin{bmatrix} A & \\ & A_n \end{bmatrix}, \begin{bmatrix} B \\ B_n \end{bmatrix}, \begin{bmatrix} C & -C_n \end{bmatrix} \right).$$

It is easily checked that its transfer function is $H - H_n$. The extended Grammians are solutions of the equations[12]

$$\hat{A}\hat{G}_c\hat{A}^H + \hat{B}\hat{B}^H = \hat{G}_c \quad \text{and} \quad \hat{A}^H\hat{G}_o\hat{A} + \hat{C}^H\hat{B} = \hat{G}_o.$$

It is then possible to compute the gradients of J with respect to A_n, B_n and C_n. This allows IRKA to iterate, i.e., with initial conditions for A_n, B_n and C_n, one computes the elements of

[12]These Stein equations are the discrete time equivalent of the Lyapunov equations in continuous time.

blocks in the extended $\hat{G}_c$ and $\hat{G}_o$, from these new projections P_n and Q_n can be obtained to give new $(A_n, B_n, C_n) = (P_n^H A Q_n, P_n^H B, C Q_n)$ etc. which should converge to a fixed point of J. This is the large scale state space equivalent of what was mentioned before in the scalar case as the nonlinear least squares problem $J(\alpha_1, \ldots, \alpha_{n-1})$ to be solved with the poles as parameters.

Many of these ideas have been extended to higher order systems (involving second derivatives), coupled systems, time-varying systems (state space matrices depend on time), nonlinear systems, implicit systems ($(zI - A)$ is replaced by $(zE - A)$ with E possibly singular), etc. This is a very active research area in linear algebra.

17 What we did not discuss...

... mostly because these topics are still largely unexplored for ORFs although there is lots of material available in the polynomial case.

1. Associated rational functions: these are rational functions obtained by shifting the recurrence coefficients as well as the poles. This corresponds to a shift in the (rational) moments. These functions are again ORFS but for a different measure. See [188, 190, 199]. They were investigated originally in the context of positive interpolatory quadrature formulas by Peherstorfer [472, 474–476].

2. Other rational quadrature formulas. Gautschi also has quadrature formulas related to the integration of rational functions. This is related to the approach given here in the sense that the spaces of rational functions in which the quadrature formulas are exact are the sum of a polynomial and a rational function. This corresponds in our setting to choosing a number of the $\alpha_k = 0$ (which give the polynomial part). These also appear in the work of Guillermo Lopez. The approach is somewhat different, although the results could well be compared. See [285, 288, 290, 294, 309, 396, 398, 400].

   The work of Guillermo Lopez dealing with rational interpolation and quadrature, mostly uses orthogonality with respect to varying measures. See [144] for an excellent survey. That means that the denominators of the rational functions, say $\pi_n(z)$, are absorbed by the measure as $\mathrm{d}\tilde{\mu}(z) = \mathrm{d}\mu(z)/|\pi_n(z)|^2$. If $\tilde{p}_n$ is the nth orthogonal polynomial for $\tilde{\mu}$, then $\tilde{p}_n/\pi_n$ is orthogonal to $\tilde{\mathcal{L}}_{n-1} = \mathcal{P}_{n-1}/\pi_n$ while the ORF ϕ_n in this survey is orthogonal to $\mathcal{L}_{n-1} = \mathcal{P}_{n-1}/\pi_{n-1}$. Another difference with the presentation of this paper is that Lopez works with a triangular table of poles: $\pi_n(z) = \prod_{k=1}^{n}(1-\overline{\alpha}_{n,k}z)$, i.e., all the poles of $\mathcal{L}_n$ can change with n.

3. Gaussian interval quadrature. This is based on an idea of Kreĭn who generalized the idea of Gaussian quadrature formulas from polynomials to an arbitrary (extended) Chebyshev system (ET-system) [380]. If $\{u_k : k = 1, 2, \ldots\}$ is such a system then there exists a quadrature exact in $\mathrm{span}\{u_1, \ldots, u_{2n}\}$ namely

$$\mathfrak{I}_\mu[f] = \int_a^b f(x)\mathrm{d}\mu(x) \approx \mathfrak{I}_n[f] = \sum_{k=1}^{n} w_k f_k, \quad f_k = \frac{\int_{I_k} f(x)\mathrm{d}\mu(x)}{\int_{I_k} \mathrm{d}\mu(x)},$$

   where $I_k = [x_k, x_k + h_k]$ with $a < x_1 \le x_1 + h_1 \le \cdots \le x_n \le x_n + h_n < b$. Thus instead of matching function values, one matches averages over certain subintervals. It the length of the interval is zero, this means point evaluation, so that the classical Gaussian formula is a special case. This idea has been worked out in all its variations for the real line and for polynomials by many authors. See [37, 59–61, 412–415].

4. The relation to continued fractions. The recurrence relation for the orthogonal polynomials are closely related to the recurrence between the successive approximants of continued fractions. So, different continued fractions have been associated with orthogonal polynomials on the line, the half line and the unit circle. The theory of continued fractions also has a long history, which we have not pointed out here. For some connections with the different moment problems see [64, 65, 67, 158, 332, 342, 343, 354, 355, 357, 358, 360, 362, 373, 427, 516]. All these are for the polynomial case. For the

rational case, some indications are given in [105, Sec. 4.4], but the continued fraction formulation of the result is mostly a rephrasing of results that are already mentioned.

5. Convergence and strong asymptotics of the zeros. We have briefly looked at the asymptotics of the ORF. For the polynomials, a flourishing literature exists to also look at the asymptotic behavior of the zeros, thus of the points where the QFs are nonzero which should supposedly converge to the support of the measure or of the moment problem that is being solved. If we want to use n-point QFs with n becoming very large, then the $O(n^2)$ complexity of the Golub-Welsch algorithm or Newton iteration to find these nodes, may become unfeasible. In such cases, one may prefer to use the asymptotic approximations of the zeros in order to reduce the complexity to $O(n)$. Such asymptotic expansions can be obtained via Riemann-Hilbert problems. Some recent examples for Jacobi and other particular polynomials, sometimes with respect to particular varying measures, can be found in [322, 382, 406, 407, 409, 410]. For ORFs this has not been done so far. See [411] and also [159] and [57, 103, 127].

6. A measure supported on one or more intervals on $\mathbb{T}$ or multiple intervals on $\mathbb{R}$. This is another topic that has been investigated for polynomials, but has never been considered for ORFs. Similarly multiple or simultaneous orthogonality and the corresponding quadratures are investigated for polynomials, but could be explored also for rational functions. [25, 26, 179, 254, 255, 327, 365, 418, 517].

7. Sobolev type orthogonality. Again a topic with many publications when orthogonal polynomials are involved but still wide open for ORFs. Although we have seen that coinciding interpolation points did not give any problem, other instances where derivatives of the ORF (or rational functions in general) are much more complicated than for polynomials. For example answering the question if there is a close connection between differential equations and ORF like for example a Rodrigues type formula, that would probably be rather complicated.

8. ...

A Hardy, Lebesgue and other spaces

A.1 The unit circle

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 μ be a probability measure on $\mathbb{T}$ with infinite support and $L_2(\mu)$ the Hilbert space with inner product $\langle f, g\rangle_\mu = \int \overline{f(t)} g(t) \mathrm{d}\mu(t)$. The normalized Lebesgue measure is $\mathrm{d}\lambda = \mathrm{d}t/(2\pi)$. Since in the larger part of the text it is clear what the measure is, we shall only add the index in $\langle \cdot, \cdot \rangle_\mu$ when confusion is possible. The trigonometric moments are denoted by $m_k = \int t^{-k} \mathrm{d}\mu(t)$, $k \in \mathbb{Z}$.

Introducing the Riesz-Herglotz kernel

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

we can associate with μ its Riesz-Herglotz transform

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

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

$$\mathcal{C} = \{f \in H(\mathbb{D}) : \operatorname{Re} f(z) \geq 0, \forall z \in \mathbb{D}\},$$

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

Define the *Poisson kernel*[13]

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

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

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

Then it follows that

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

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

$$\Omega_\mu(z) = m_0 + 2 \sum_{k=1}^{\infty} m_k z^k, \quad (m_0 = 1)$$

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

[13] Note that for $t = e^{\mathrm{i}\theta}$ and $z = r e^{\mathrm{i}\sigma}$, then $P(t, z) = \frac{1 - r^2}{1 - 2r\cos(\theta - \sigma) + r^2} = P_r(\theta - \sigma)$ with $P_r(\eta) = \frac{1 - r^2}{1 - 2r\cos\eta + r^2}$.

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

$$\mathcal{B}(\mathbb{D}) = \{f \in H(\mathbb{D}) : f(\mathbb{D}) \subseteq \mathbb{D} \cup \mathbb{T}\}$$

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

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

and $s_n = 1$ 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

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

with ν 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

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

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 μ

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

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., $t \in \mathbb{T}$. One may prove that

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

More about the material in this section can be found in many textbooks a.o. [232, 278, 340, 378, 489, 555].

Let λ 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 $t = \mathrm{e}^{i\theta}$, 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. 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* $s(z,w) = s_w(z) = \lim_{n\to\infty} \sum_{k=0}^n (\overline{w}z)^k = \frac{1}{1-z\overline{w}}$. It 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 $f \in H^2(\mathbb{D}, \mathrm{d}\lambda)$.

A.2 The real line

Let $\mathbb{U} = \{z \in \mathbb{C} : \mathrm{Im}\, z > 0\}$ be the upper half plane and $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Re}\, z > 0\}$ the right half plane.

The Cayley transform

$$z = \mathfrak{c}[w] = \mathrm{i}\frac{1-w}{1+w}, \quad w \in \mathbb{D} \cup \mathbb{T}, \quad z \in \mathbb{U} \cup \mathbb{R}$$

is a one-to-one relation mapping the unit disk $\mathbb{D}$ onto the upper half plane $\mathbb{U}$, of $\mathbb{T}$ onto $\hat{\mathbb{R}}$, and it maps -1 onto ∞. This implies that

$$t = e^{\mathrm{i}\theta} = \mathfrak{c}^{inv}[x] = \frac{\mathrm{i}-x}{\mathrm{i}+x}, \quad x \in \mathbb{R}, \quad t \in \mathbb{T}, \quad \theta \in (-\pi, \pi)$$

maps the unit circle to the real line and the Lebesgue measure is mapped onto

$$\mathrm{d}\lambda(t, \mathbb{T}) = \frac{\mathrm{d}\theta}{2\pi} = \frac{\mathrm{d}x}{\pi(1+x^2)} = \frac{\mathrm{d}\lambda(x, \mathbb{R})}{1+x^2} = \mathrm{d}\mathring{\lambda}(x, \mathbb{R}).$$

Similarly we can define $\mathrm{d}\mathring{\mu}(x) = \frac{\mathrm{d}\mu(x)}{1+x^2}$.

As before we consider the space $L^2(\mu, \mathbb{R})$ with inner product $\langle f, g\rangle_\mu = \int \overline{f(x)} g(x) \mathrm{d}\mu(x)$.

The Hardy spaces $H^p(\mathbb{U})$ can be defined as the set of functions analytic in $\mathbb{U}$ such that [278, p. 51], [232, Chap. 10], [378]

$$\sup_{y>0} \int_{-\infty}^{\infty} |f(x+\mathrm{i}y)|^p \mathrm{d}x = \|f\|_p^p < \infty \quad \text{and} \quad \sup_{z\in\mathbb{U}} |f(z)| = \|f\|_\infty < \infty.$$

This means that $H^\infty(\mathbb{U}) = \{f : f \circ \mathfrak{c} \in H^\infty(\mathbb{D})\}$ and

$$H^p(\mathbb{U}) = \{(z+\mathrm{i})^{-2/p} f(z) : f \circ \mathfrak{c} \in H^p(\mathbb{D})\}, \quad 0 < p < \infty.$$

Besides that we also have spaces $\mathring{H}^p(\mathbb{U})$ defined by

$$f \in \mathring{H}^p(\mathbb{U}) \leftrightarrow (z+\mathrm{i})^{-2/p} f(z) \in H^p(\mathbb{U}) \leftrightarrow f \circ \mathfrak{c} \in H^p(\mathbb{D}).$$

Yet another way to define Hardy spaces is via the Fourier transform, which is defined as

$$\hat{f}(x) = \int e^{\mathrm{i}xt} f(t) \mathrm{d}\lambda(t)$$

and $H^p(\mathbb{U})$ are those functions in $L^p(\mathbb{R})$ for which the Fourier transform vanishes for $x < 0$.

The Carathéodory class is defined as

$$\mathcal{C} = \{f \in H(\mathbb{U}) : \mathrm{Re}\ f(\mathbb{U}) \geq 0\}$$

and the bounded analytic functions are

$$\mathcal{B} = \{f \in H(\mathbb{U}) : |f(\mathbb{U})| \leq 1\}.$$

The Nevanlinna and Poisson kernel are[14] is

$$D(x,z) = \frac{1}{\mathrm{i}} \frac{1+xz}{x-z} \quad \text{and} \quad P(x,z) = \frac{1}{2\mathrm{i}} \frac{z-\overline{z}}{(x-z)(x-\overline{z})} = \frac{\mathrm{Im}\, z}{|x-z|^2}, \quad x \in \mathbb{R}$$

and it holds that

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

and [334, p. 588], [378, p. 144] the Nevanlinna transform of the measure is defined as

$$\Omega_\mu(z) = \mathrm{i}c + \int D(x,z) \mathrm{d}\mathring{\mu}(x), \ c \in \mathbb{R}, \quad z \in \mathbb{U}.$$

It is a function in the class $\mathcal{C}$ and any function in $\mathcal{C}$ has such a Nevanlinna representation with μ nonnegative. If $\sup |y\Omega(\mathrm{i}y)|$ is bounded for $y > 1$, then μ is finite and we can use the simpler Hamburger (or Markov or Stieltjes or Cauchy) representation [334, p. 590], [2, p. 92]

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

Note that [278, p. 29], [378, p. 146]

$$\mathrm{Re}\, \Omega_\mu(z) = \int P(x,z) \mathrm{d}\mu(z) > 0 \quad \text{and} \quad \lim_{y \downarrow 0} \mathrm{Re}\, \Omega_\mu(x+\mathrm{i}y) = \mu'(x).$$

[14] Originally a Nevanlinna function refers to a function holomorphic in the upper half plane and having there a nonnegative imaginary part. Here we have added the factor $1/\mathrm{i}$ to make the real part nonnegative. In systems theory, the frequency axis is the imaginary axis, in which case the integration variable is of the form $\mathrm{i}s$ with $s \in \mathbb{R}$ which cancels the i factor in the definition of the kernel. Stability requires that the functions are *positive real*, i.e., have a positive real part in the right half plane. We kept this idea and therefore we call it Carathéodory rather than Nevanlinna functions.

With the normalization $\int \mathrm{d}\mathring{\mu} = 1$ and $\Omega(0) = 1$, the relation between them is one-to-one.

An inner function is again a function in $H(\mathbb{U})$ with boundary value equal to 1 a.e. and an example of a meromorphic inner function is a Blaschke product $B(z) = \prod_n \zeta_n(z)$ where now

$$\zeta_n(z) = s_n \frac{z - \alpha_n}{z - \overline{\alpha}_n}, \quad s_n = \frac{|1 + \alpha_n^2|}{1 + \alpha_n^2}, \quad \alpha_n \in \mathbb{U}.$$

If $\alpha_n = \alpha_0$ ($\alpha_0 = \mathrm{i}$), then we set $s_n = 1$. A Blaschke product converges if

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

Inner functions (defined in $\mathbb{U}$) have a pseudo-analytic extension to the lower half plane by setting $f(z) = f_*(z)$ where the latter is by definition $f_*(z) = \overline{f(\overline{z})}$ which means that for $x \in \mathbb{R}$, $f_*(x) = f(x)$.

Suppose Szegő's condition $\int \log \mu'(x) \mathrm{d}\mathring{\lambda}(x) > -\infty$ is satisfied, then we can again define an outer spectral factor for μ as

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

We shall assume that we choose $c = 1$. We have

$$|\sigma(z)|^2 = \exp\left\{ P(x, z) \log \mu'(x) \mathrm{d}\lambda(x) \right\}, \quad \mu'(x) = |\sigma(x)|^2 \quad \text{a.e. } x \in \mathbb{R}.$$

Or more generally (with appropriate conditions)

$$\sigma(z)\sigma_*(z) = \frac{1}{2}[\Omega(z) + \Omega_*(z)].$$

We could say that Ω_μ is the analytic term in an additive split and σ the analytic factor in a multiplicative split (i.e., an additive split of the logarithm) of the (density of the) measure.

References

This list gives historical as well as recent references. It is far from being complete in covering all areas that were discussed in this survey. It reflects the personal preferences and the limited knowledge of the author. None of the many omissions is intentional.

[1] N.I. Achieser. (1956) *Theory of approximation*. Frederick Ungar Publ. Co., New York (cited on pp. 81, 89).

[2] N.I. Akhiezer [Achieser]. (1969) *The classical moment problem and some related questions in analysis*. Originally published Moscow, 1961. Oliver and Boyd, Edinburgh (cited on pp. 12, 16, 17, 19, 21, 22, 115).

[3] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. (1968) *Infinite Hankel matrices and generalized Carathéodory–Fejér and F. Riesz problems*. Funk. Anal. i Prilozen **2(1)**. (In Russian). English translation: *Functional Anal. Appl.* **2** (1968) 1–18, pp. 1–19 (cited on p. 104).

[4] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. (1968) *Infinite Hankel matrices and generalized Carathéodory-Fejér and Riesz problems*. Functional Anal. Appl. **2**, pp. 1–18 (cited on pp. 97, 104).

[5] V.M. Adamjan, D.Z. Arov, and M.G. Kreĭn. (1968) *Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and I. Schur*. Functional Anal. Appl. **2**, pp. 269–281 (cited on p. 97).

[6] J. Agler and J.E. McCarthy. (2002) *Pick interpolation and Hilbert function spaces*. Vol. 44. of Graduate Studies in Mathematics. AMS (cited on p. 107).

[7] H. Akçay and B. Ninness. (1999) *Orthonormal basis functions for continuous-time systems and L_p convergence*. Math. Control Signals Systems **12**, pp. 295–305 (cited on pp. 107, 108).

[8] H. Akçay and B. Ninness. (1998) *Rational basis functions for robust identification from frequency and time domain measurements*. Automatica **34(9)**, pp. 1101–1117 (cited on pp. 107, 108).

[9] E. Aldén. (1985) *A survey of weak and strong moment problems with generalizations*. Preprint Series 2. Dept. Math. Umeå, pp. 1–38 (cited on p. 19).

[10] A. Almendral Vazquez. (2003) *Nevanlinna parametrization of the solutions of some rational moment problems*. Analysis **23(2)**, pp. 107–124 (cited on p. 93).

[11] D. Alpay. (2001) *The Schur algorithm, reproducing kernel spaces and system theory*. SMF/AMS Texts and Monographs 5. AMS (cited on p. 107).

[12] D. Alpay, V. Bolotnikov, and Y. Peretz. (1995) *On the tangential interpolation problem for H_2 functions*. Trans. Amer. Math. Soc. **347**, pp. 675–686 (cited on p. 105).

[13] D. Alpay, A. Dijksma, J. Rovyak, and H. de Snoo. (1997) *Schur functions, operator colligations and reproducing kernel Pontryagin spaces*. Vol. 96. of Oper. Theory: Adv. Appl. Birkhäuser Verlag (cited on p. 95).

[14] D. Alpay and H. Dym. (1986) *On applications of reproducing kernel spaces to the Schur algorithm and relational J unitary factorization*. In: *I. Schur methods in operator theory and signal processing*. Ed. by I. Gohberg. Vol. 18. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel, pp. 89–159 (cited on p. 97).

[15] D. Alpay and B. Kirstein, eds. (2012) *Interpolation, schur functions and moment problems II*. Vol. 226. of Oper. Theory: Adv. Appl. Birkhäuser Verlag (cited on pp. 97, 107).

[16] D. Alpay and P. Loubaton. (1993) *The tangential trigonometric moment problem on an interval and related topics*. In: *New aspects in interpolation and completion theories*. Ed. by I. Gohberg. Vol. 64. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 1–21 (cited on p. 105).

[17] G. Ammar, D. Calvetti, and L. Reichel. (1999) *Computation of Gauss-Kronrod quadrature rules with non-positive weights*. Electron. Trans. Numer. Anal. **9**, pp. 26–38 (cited on p. 80).

[18] G. Ammar, G.H. Golub, and L. Reichel. (1993) *Direct and inverse unitary eigenproblems in signal processing: an overview*. In: *Linear Algebra for Large Scale and Real-Time Applications*. Ed. by G.H. Golub M.S. Moonen and B.L.R. De Moor. NATO Adv. Sci. Inst. Ser. E Appl. Sci. Kluwer Academic Publisher, pp. 341–343 (cited on p. 74).

[19] G. Ammar, W.B. Gragg, and L. Reichel. (1986) *On the eigenproblem for orthogonal matrices*. In: *Proceedings of the 25th Conference on Decision and Control Athens*. IEEE, pp. 1963–1966 (cited on p. 25).

[20] G. Ammar and C. He. (1995) *On an inverse eigenvalue problem for unitary Hessenberg matrices*. Linear Algebra Appl. **218**, pp. 263–272 (cited on p. 74).

[21] G.S. Ammar, W.B. Gragg, and L. Reichel. (1991) *Constructing a unitary Hessenberg matrix from spectral data*. In: *Numerical linear algebra, digital signal processing and parallel algorithms*. Ed. by G. Golub and P. Van Dooren. Vol. 70. of NATO-ASI Series, F: Computer and Systems Sciences. Springer, Berlin, pp. 385–395 (cited on p. 74).

[22] G.S. Ammar, L. Reichel, and D. Sorenson. (1992) *Algorithm 730. An implementation of a divide and conquer algorithm for the unitary eigenproblem*. ACM Trans. Math. Software **18(3)**, pp. 292–307 (cited on p. 25).

[23] G.S. Ammar, L. Reichel, and D. Sorenson. (1994) *Corrigendum: Algorithm 730. An implementation of a divide and conquer algorithm for the unitary eigenproblem*. ACM Trans. Math. Software **20(1)**, p. 161 (cited on p. 25).

[24] A.C. Antoulas. (2005) *Approximation of large-scale dynamical systems*. SIAM (cited on p. 108).

[25] A.I. Aptekarev. (1999) *Strong asymptotics of multiply orthogonal polynomials for Nikishin systems*. Math. Sb. **190(5)**. (Translation from Mat. Sb. 190, No.5, 3-44 (1999).), pp. 631–669 (cited on p. 111).

[26] A.I. Aptekarev and J. Kuijlaars A.B. (2011) *Hermite-Padé approximations and multiple orthogonal polynomial ensembles*. Russ. Math. Surv. **66(6)**, pp. 1133–1199 (cited on p. 111).

[27] D.Z. Arov. (1995) *Computation of the resolvent matrix for the generalized bitangential Schur and Carathéodory-Nevanlinna-Pick interpolation problem in the strictly completely indeterminate case*. Integral Equations Operator Theory **22(3)**, pp. 253–272 (cited on p. 96).

[28] D.Z. Arov. (1973) *Darlington realization of matrix-valued functions*. Math. USSR-Izv. **7**, pp. 1295–1326 (cited on p. 57).

[29] D.Z. Arov. (1988) *On regular and singular J-inner matrix functions and related extrapolation problems*. Functional Anal. Appl. **22**. (Russian), pp. 57–59 (cited on p. 97).

[30] D.Z. Arov and H. Dym. (1997) *J-inner matrix functions, interpolation and inverse problems for canonical systems, I: Foundations*. Integral Equations Operator Theory **29(4)**, pp. 373–454 (cited on p. 98).

[31] D.Z. Arov and H. Dym. (2000) *J-inner matrix functions, interpolation and inverse problems for canonical systems, II: The inverse monodromy problem*. Integral Equations Operator Theory **36(1)**, pp. 11–70 (cited on p. 98).

[32] D.Z. Arov and H. Dym. (2000) *J-inner matrix functions, interpolation and inverse problems for canonical systems, III: More on the inverse monodromy problem*. Integral Equations Operator Theory **36(2)**, pp. 127–181 (cited on p. 98).

[33] D.Z. Arov and H. Dym. (2002) *J-inner matrix functions, interpolation and inverse problems for canonical systems, IV: Direct and inverse bitangential input scattering problems*. Integral Equations Operator Theory **43(1)**, pp. 1–67 (cited on pp. 98, 105).

[34] D.Z. Arov, B. Fritzsche, and B. Kirstein. (2000) *On a parametrization formula for the solution set of a completely indeterminate generalized matricial Carathéodory-Fejér problem*. Math. Nachr. **219**, pp. 5–43 (cited on p. 98).

[35] G. Arsene, and T. Constantinescu. (1987) *The structure of positive block matrices and nonstationary prediction*. J. Functional Anal. **70**, pp. 402–425 (cited on p. 95).

[36] N.M. Atkishiyev and S.M. Nagiyev. (1994) *On Rogers-Szegő polynomials*. J. Phys. A **27**, pp. L611–L615 (cited on p. 63).

[37] V.F. Babenko and D.S. Skorokhodov. (2007) *On the best interval quadrature formulae for classes of differentiable periodic functions*. J. Complexity **23(4-6)**, pp. 890–917 (cited on p. 110).

[38] Z. Bai, P.M. Dewilde, and R.W. Freund. (2005) *Reduced order modeling*. In: *Numerical Methods in Electromagnetics*. Ed. by W.H.A. Schilders and E.J.W. ter Maten. Vol. 13. of Handbook of Numerical Analysis, pp. 825–895 (cited on p. 105).

[39] M. Bakonyi and T. Constantinescu. (1992) *Schur's algorithm and several applications*. Vol. 261. of Pitman Research Notes in Mathematics. Longman, Harlow, UK (cited on p. 95).

[40] M. Bakonyi and H.J. Woerdeman. (2011) *Matrix completions, moments, and sums of Hermitian squares*. Princeton Series in Applied Mathematics. Princeton University Press (cited on p. 97).

[41] J.A. Ball, I. Gohberg, and L. Rodman. (1990) *Interpolation of rational matrix functions*. Vol. 45. of Oper. Theory: Adv. Appl. Birkhäuser Verlag (cited on pp. 95, 105, 107).

[42] J.A. Ball and J. Kim. (1996) *Bitangential interpolation problems for symmetric rational functions*. Linear Algebra Appl. **241-243**, pp. 113–152 (cited on p. 105).

[43] L. Baratchart, L. Golinskii, and S. Kupin. (2010) *Multipount Schur algorithm, II: generalized moment problems, Gaussian processes and prediction*. ArXiv Mathematics e-prints. eprint: `arxiv:1007.1363` (cited on p. 94).

[44] L. Baratchart, S. Kupin, V. Lunot, and M. Olivi. (2011) *Multipount Schur algorithm and orthogonal rational functions: I convergence properties*. J. Anal. Math. **14(1)**, pp. 207–253 (cited on p. 94).

[45] N. Bebiano, C.M. da Fonseca, and J. da Providência. (2011) *An inverse eigenvalue problem for periodic Jacobi matrices in Minkowski spaces*. Linear Algebra Appl. **435(8)**, pp. 2033–2045 (cited on p. 74).

[46] N. Bebiano, S. Furtado, and J. da Providência. (2012) *An algorithm for constructing a pseudo-Jacobi matrix from given spectral data*. Numer. Linear Algebra Appl. **20**, pp. 185–197 (cited on p. 74).

[47] N. Bebiano, S. Furtado, and J. da Providência. (2011) *On the eigenvalues of principal submatrices of J-normal matrices*. Linear Algebra Appl. **435**, pp. 3101–3114 (cited on p. 74).

[48] N. Bebiano, S. Furtado, and J. da Providência. (2011) *Quasi-Jacobi form for J-normal matrices and the inverse spectral problem*. Linear Algebra Appl. **436**, pp. 1739–1753 (cited on p. 74).

[49] N. Bebiano and J. da Providência. (2011) *Inverse problems for pseudo-Jacobi matrices, existence and uniqueness results*. Inverese Problems **27**, p. 025005 (cited on p. 74).

[50] B. Beckermann, J. Bustamante, R. Martínez-Cruz, and J.M. Quesada. (2013) *Gaussian, Lobatto and Radau positive quadrature rules with a prescribed abscissa*. Calcolo. To appear. Online with doi:10.1007/s10092-013-0087-3 (cited on p. 67).

[51] B. Beckermann and S. Güttel. (2012) *Superlinear convergence of the rational Arnoldi method for the approximation of matrix functions*. Numer. Math. **121(2)**, pp. 205–236 (cited on p. 102).

[52] B. Beckermann, S. Güttel, and R. Vandebril. (2010) *On the convergence of rational Ritz values*. SIAM J. Matrix Anal. Appl. **31**, pp. 1740–1774 (cited on p. 102).

[53] B. Beckermann and L. Reichel. (2009) *Error estimation and evaluation of matrix functions via the Faber transform*. SIAM J. Numer. Anal. **47(5)**, pp. 3849–3883 (cited on p. 102).

[54] R. van Beeumen, K. Meerbergen, and W. Michiels. (2013) *A rational Krylov method based on Hermite interpolation for nonlinear eigenvalue problems*. SIAM J. Sci. Comput. **35(1)** (cited on p. 103).

[55] M. Benzi, E. Estrada, and C. Klymko. (2013) *Ranking hubs and authorities using matrix functions*. Linear Algebra Appl. **438**, pp. 2447–2474 (cited on p. 102).

[56] J.-P. Berrut. (1997) *The barycentric weights of rational interpolation with prescribed poles*. J. Comput. Appl. Math. **86(1)**, pp. 45–52 (cited on p. 63).

[57] T. Bloom, D.S. Lubinsky, and H. Stahl. (1992) *Interpolatory integration rules and orthogonal polynomials with varying weight*. Numer. Algorithms **3(1-4)**, pp. 55–65 (cited on p. 111).

[58] F. Bo-Yeh, C.D. Yang, and M.C. Tsai. (1988) *Unified approach to H^∞ optimization, Hankel approximation and balanced realization problems*. Internat. J. Control **47(4)**, pp. 967–978 (cited on p. 104).

[59] B. Bojanov and P. Petrov. (2006) *Gaussian interval quadrature formulae for Tchebycheff systems*. SIAM J. Numer. Anal. **43(2)**, pp. 787–795 (cited on p. 110).

[60] B. Bojanov and P. Petrov. (2003) *Uniqueness of the Gaussian interval quadrature formula*. Numer. Math. **95**, pp. 53–62 (cited on p. 110).

[61] B. Bojanov and Petrov P.P. (2001) *Gaussian interval quadrature formula*. Numer. Math. **87**, pp. 625–643 (cited on p. 110).

[62] D.L. Boley and G.H. Golub. (1987) *A survey of matrix inverse eigenvalue problems*. In: *Inverse problems*. Vol. 3. Physics Trust Publications, Bristol, England, pp. 595–622 (cited on p. 74).

[63] C. Bonan-Hamada. (1994) *Orthogonal Laurent polynomials and indeterminate strong Stieltjes moment problems*. PhD thesis. University of Colorado, Boulder (cited on p. 17).

[64] C. Bonan-Hamada and W.B. Jones. (2004) *Stieltjes continued fractions and special functions: a survey*. Comm. Anal. Th. Continued Fractions **12**, pp. 5–68 (cited on p. 110).

[65] C. Bonan-Hamada, W.B. Jones, and O. Njåstad. (1998) *Natural solutions of strong Stieltjes moment problems derived from PC-fractions*. In: *Orthogonal functions, moment theory and continued fractions*. Ed. by W.B. Jones and A.S. Ranga. Vol. 199. of Lecture Notes in Pure and Appl. Math. Marcel Dekker, pp. 15–30 (cited on pp. 17, 19, 110).

[66] C. Bonan-Hamada, W.B. Jones, and O. Njåstad. (2010) *Zeros of orthogonal Laurent polynomials and solutions of the strong Stieltjes moment problem*. J. Comput. Appl. Math. **235(4)**, pp. 895–903 (cited on pp. 17, 19).

[67] C. Bonan-Hamada, W.B. Jones, O. Njåstad, and W.J. Thron. (1998) *A class of indeterminate strong Stieltjes moment problems with discrete distributions*. In: *Orthogonal functions, moment theory and continued fractions*. Ed. by W.B. Jones and A.S. Ranga. Vol. 199. of Lecture Notes in Pure and Appl. Math. Marcel Dekker, pp. 31–55 (cited on pp. 17, 19, 110).

[68] C.F. Borges, R. Frezza, and W.B. Gragg. (1995) *Some inverse eigenproblems for Jacobi and arrow matrices*. Numer. Linear Algebra Appl. **2(3)**, pp. 195–203 (cited on p. 74).

[69] P. Borwein and T. Erdélyi. (1995) *Polynomials and polynomial inequalities*. Springer Verlag, New York (cited on p. 89).

[70] P. Borwein, T. Erdélyi, and J. Zhang. (1994) *Chebyshev polynomials and Markov-Bernstein type inequalities for the rational spaces*. J. London Math. Soc. (2) **50**, pp. 501–519 (cited on p. 57).

[71] C. Brezinski. (1990) *A direct proof of the Christoffel-Darboux identity and its equivalence to the recurrence relationship*. J. Comput. Appl. Math. **32**, pp. 17–25 (cited on p. 13).

[72] C. Brezinski and A. Sameh, eds. (2013) *Walter Gautschi. Selected works with commentaries*. Contemporary Mathematicians. (3 volumes). Birkhäuser Verlag (cited on p. 12).

[73] A. Bultheel. (1984) *On the ill-conditioning of locating the transmission zeros in least squares ARMA filtering*. J. Comput. Appl. Math. **11(1)**, pp. 103–118 (cited on pp. 106, 107).

[74] A. Bultheel. (1980) *Orthogonal matrix functions related to the multivariable Nevanlinna-Pick problem*. Bull. Soc. Math. Belg. Sér. B **32(2)**, pp. 149–170 (cited on p. 95).

[75] A. Bultheel and M.J. Cantero. (2009) *A matricial computation of rational quadrature formulas on the unit circle*. Numer. Algorithms **52(1)**, pp. 47–68 (cited on pp. 53, 54).

[76] A. Bultheel and P. Carrette. (2003) *Algebraic and spectral properties of general Toeplitz matrices*. SIAM J. Control Optim. **41(5)**, pp. 1413–1439 (cited on pp. 57, 107, 108).

[77] A. Bultheel, R. Cruz Barroso, K. Deckers, and P. González-Vera. (2009) *Rational Szegő quadratures associated with Chebyshev weight functions*. Math. Comp. **78**, pp. 1031–1059 (cited on pp. 59, 60).

[78] A. Bultheel, R. Cruz Barroso, P. González-Vera, and F. Perdomo Pío. (2010) *On the computation of symmetric Szegő-type quadrature formulas*. In: *Proceedings of the MCalv6(5) Conference, Zaragoza, 7-9 September 2009*. Ed. by A. Elipe, J.I. Montijano, and L. Rández. Vol. 33. of Monografías de la Real Academia de Ciencias Exactas Físicas Químicas y Naturales de Zaragoza, pp. 177–196 (cited on p. 55).

[79] A. Bultheel, R. Cruz Barroso, and M. Van Barel. (2010) *On Gauss-type quadrature formulas with prescribed nodes anywhere on the real line*. Calcolo **47(1)**, pp. 21–48 (cited on p. 67).

[80] A. Bultheel, R. Cruz Barroso, and Marc Van Barel. (2013) *Erratum to : on Gauss-type quadrature formulas with prescribed nodes anywhere on the real line*. Calcolo **50(2)**, pp. 163–164 (cited on p. 67).

[81] A. Bultheel, L. Daruis, and P. González-Vera. (2001) *A connection between quadrature formulas on the unit circle and the interval* $[-1, 1]$. J. Comput. Appl. Math. **132(1)**, pp. 1–14 (cited on pp. 51, 70).

[82] A. Bultheel, L. Daruis, and P. González-Vera. (2005) *Positive interpolatory quadrature formulas and para-orthogonal polynomials*. J. Comput. Appl. Math. **179(1-2)**, pp. 97–119 (cited on p. 70).

[83] A. Bultheel, L. Daruis, and P. González-Vera. (2009) *Quadrature formulas on the unit circle with prescribed nodes and maximal domain of validity*. J. Comput. Appl. Math. **231(2)**, pp. 948–963 (cited on p. 70).

[84] A. Bultheel and P. Dewilde. (1979) *Orthogonal functions related to the Nevanlinna-Pick problem*. In: *Proc. 4th Int. Conf. on Math. Theory of Networks and Systems at Delft*. Ed. by P. Dewilde. Western Periodicals, North-Hollywood, pp. 207–212 (cited on pp. 30–32).

[85] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. (1999) *Estimates of the rate of convergence for certain quadrature formulas on the half line*. In: *Continued fractions: from number theory to constructive approximation*. Ed. by B. Berndt and F. Gesztesy. Vol. 236. of Contemp. Math. Amer. Math. Soc., pp. 85–100 (cited on p. 90).

[86] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. (2000) *On the convergence of certain Gauss-type quadrature formulas for unbounded intervals*. Math. Comp. **69**, pp. 721–747 (cited on p. 90).

[87] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. (2003) *Orthogonal Laurent polynomials and quadrature formulas for unbounded intervals: Part I. Gauss-type formulas*. Rocky Mountain J. Math. **33(2)**, pp. 585–608 (cited on p. 90).

[88] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. (2005) *Orthogonal Laurent polynomials and quadrature formulas for unbounded intervals: Part II. Interpolatory rules*. Appl. Numer. Math. **54(1)**, pp. 39–63 (cited on p. 90).

[89] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. (1997) *Quadrature on the half line and two-point Padé approximants to Stieltjes functions. Part II: Convergence*. J. Comput. Appl. Math. **77**, pp. 53–76 (cited on pp. 20, 44, 90).

[90] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. (1997) *Quadrature on the half line and two-point Padé approximants to Stieltjes functions. Part III: The unbounded case*. J. Comput. Appl. Math. **87**, pp. 95–117 (cited on pp. 20, 44, 90).

[91] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1992) *A Favard theorem for orthogonal rational functions on the unit circle*. Numer. Algorithms **3**, pp. 81–89 (cited on p. 34).

[92] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1997) *A Favard theorem for rational functions with poles on the unit circle*. East J. Approx. **3**, pp. 21–37 (cited on pp. 34, 46).

[93] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1992) *A moment problem associated to rational Szegő functions*. Numer. Algorithms **3**, pp. 91–104 (cited on p. 91).

[94] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1997) *A rational moment problem on the unit circle*. Methods Appl. Anal. **4(3)**, pp. 283–310 (cited on p. 91).

[95] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2002) *A rational Stieltjes moment problem*. Appl. Math. Comput. **128(2-3)**, pp. 217–235 (cited on pp. 90, 91).

[96] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2008) *An indeterminate rational moment problem and Carathéodory functions*. J. Comput. Appl. Math. **219(2)**, pp. 359–369 (cited on p. 91).

[97] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2010) *Computation of rational Szegő-Lobatto quadrature formulas*. Appl. Numer. Math. **60(12)**, pp. 1251–1263 (cited on p. 65).

[98] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1995) *Convergence of modified approximants associated with orthogonal rational functions*. J. Comput. Appl. Math. **57(1/2)**, pp. 77–86 (cited on pp. 37, 87, 90).

[99] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2001) *Determinacy of a rational moment problem*. J. Comput. Appl. Math. **133(1-2)**, pp. 241–252 (cited on pp. 91, 93).

[100] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1999) *Elements of a theory of orthogonal rational functions*. Rev. Acad. Canaria Cienc. **9(1-2)**, pp. 127–152 (cited on pp. 29, 91).

[101] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1995) *Favard theorem for reproducing kernels*. J. Comput. Appl. Math. **57(1/2)**, pp. 57–76 (cited on pp. 35, 94).

[102] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2010) *Growth properties of Nevanlinna matrices for rational moment problems*. J. Approx. Theory **162(12)**, pp. 2184–2201 (cited on pp. 91, 93).

[103] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2011) *Natural solutions of rational Stieltjes moment problems*. J. Math. Anal. Appl. **377(2)**, pp. 571–583 (cited on pp. 91, 111).

[104] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1996) *On the convergence of multipoint Padé-type approximants and quadrature formulas associated with the unit circle*. Numer. Algorithms **13**, pp. 321–344 (cited on pp. 77, 88).

[105] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1999) *Orthogonal rational functions*. Vol. 5. of Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press (cited on pp. 29–34, 39, 43, 45–50, 82–87, 91, 92, 94, 107, 111).

[106] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1998) *Orthogonal rational functions and interpolatory product rules on the unit circle. I. Recurrence and interpolation*. Analysis **18**, pp. 167–183 (cited on p. 40).

[107] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1998) *Orthogonal rational functions and interpolatory product rules on the unit circle. II. Quadrature and convergence*. Analysis **18**, pp. 185–200 (cited on p. 40).

[108] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2000) *Orthogonal rational functions and interpolatory product rules on the unit circle. III. Convergence of general sequences*. Analysis **20**, pp. 99–120 (cited on p. 40).

[109] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1996) *Orthogonal rational functions and modified approximants*. Numer. Algorithms **11**, pp. 57–69 (cited on pp. 37, 38).

[110] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1997) *Orthogonal rational functions and nested disks*. J. Approx. Theory **89**, pp. 344–371 (cited on pp. 34, 91).

[111] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2003) *Orthogonal rational functions and quadrature on the real half line*. J. Complexity **19(3)**, pp. 212–230 (cited on pp. 76, 90).

[112] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1992) *Orthogonal rational functions and quadrature on the unit circle*. Numer. Algorithms **3**, pp. 105–116 (cited on p. 36).

[113] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2003) *Orthogonal rational functions and tridiagonal matrices*. J. Comput. Appl. Math. **153(1-2)**, pp. 89–97 (cited on p. 56).

[114] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2005) *Orthogonal rational functions on the real half line with poles in* $[-\infty, 0]$. J. Comput. Appl. Math. **179(1-2)**, pp. 121–155 (cited on pp. 44, 51, 68, 90, 93).

[115] 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: *Orthogonal Polynomials and Special Functions*. Ed. by W. Van Assche and F. Marcellán. Vol. 1883. of Lecture Notes in Math. Springer Verlag, pp. 187–228 (cited on p. 95).

[116] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1994) *Orthogonal rational functions on the unit circle: theory, applications and generalizations*. In: *Orthogonal Polynomials on the Unit Circle: Theory and Applications*. Ed. by M. Alfaro, A. García, C. Jagels, and F. Marcellan. Universidad Carlos III de Madrid, pp. 21–42 (cited on p. 29).

[117] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2003) *Positivity of continued fractions associated with rational Stieltjes moment problems*. Rocky Mountain J. Math. **33(2)**, pp. 609–627 (cited on p. 91).

[118] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2001) *Quadrature and orthogonal rational functions*. J. Comput. Appl. Math. **127(1-2)**. Invited paper, pp. 67–91 (cited on pp. 29, 76, 77).

[119] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1994) *Quadrature formulas on the unit circle based on rational functions*. J. Comput. Appl. Math. **50**, pp. 159–170 (cited on pp. 61, 87).

[120] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2013) *Quadratures associated with pseudo-orthogonal rational functions on the real half line with poles in* $[-\infty, 0]$. J. Comput. Appl. Math. **237(1)**, pp. 589–602 (cited on pp. 50, 51, 68, 93).

[121] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1997) *Rates of convergence of multipoint rational approximants and quadrature formulas on the unit circle*. J. Comput. Appl. Math. **77(1-2)**, pp. 77–102 (cited on pp. 87, 88).

[122] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2010) *Rational quadrature formulas on the unit circle with prescribed nodes and maximal domain of validity*. IMA J. Numer. Anal. **30(4)**, pp. 940–963 (cited on pp. 64, 65, 71, 72).

[123] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1994) *Recurrence relations for orthogonal functions*. In: *Continued Fractions and Orthogonal Functions*. Ed. by S.C. Cooper and W.J. Thron. Vol. 154. of Lecture Notes in Pure and Appl. Math. Marcel Dekker, pp. 23–46 (cited on p. 40).

[124] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (2004) *Separation of zeros of para-orthogonal rational functions*. Rev. Acad. Canaria Cienc. **16(1-2)**, pp. 9–14 (cited on p. 37).

[125] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. (1992) *The computation of orthogonal rational functions and their interpolating properties*. Numer. Algorithms **2(1)**, pp. 85–114 (cited on p. 38).

[126] A. Bultheel, P. González-Vera, and R. Orive. (1995) *Quadrature on the half line and two-point Padé approximants to Stieltjes functions. Part I: Algebraic aspects*. J. Comput. Appl. Math. **65**, pp. 57–72 (cited on pp. 20, 44, 90).

[127] A. Bultheel, E. Hendriksen, and O. Njåstad. (2014) *Holomorphic functions associated with indeterminate rational moment problems*. J. Comput. Appl. Math. submitted (cited on pp. 93, 111).

[128] A. Bultheel and A. Lasarow. (2010) *On the existence of para-orthogonal rational functions on the unit circle*. Analysis **30(3)**, pp. 301–316 (cited on p. 37).

[129] A. Bultheel and A. Lasarow. (2008) *Solution of the multiple Schur-Nevanlinna problem for Schur functions via orthogonal rational functions*. Analysis **28**, pp. 95–123 (cited on p. 93).

[130] A. Bultheel and M. Van Barel. (1994) *Linear prediction: mathematics and engineering*. Bull. Belgian Math. Soc. Simon Stevin **1**, pp. 1–58 (cited on pp. 14, 93, 96).

[131] A. Bultheel and M. Van Barel. (1986) *Padé techniques for model reduction in linear system theory: a survey*. J. Comput. Appl. Math. **14(3)**, pp. 401–438 (cited on pp. 27, 104).

[132] A. Bultheel and M. Van Barel. (1995) *Vector orthogonal polynomials and least squares approximation*. SIAM J. Matrix Anal. Appl. **16(3)**, pp. 863–885 (cited on p. 75).

[133] A. Bultheel, M. Van Barel, Y. Rolain, and R. Pintelon. (2005) *Numerically robust transfer function modelling from noisy frequency domain data*. IEEE Trans. Automat. Control **50(11)**, pp. 1835–1839 (cited on p. 75).

[134] A. Bultheel, M. Van Barel, Y. Rolain, and J. Schoukens. (2001) *Numerically robust rational approximation for identification*. In: *40th IEEE Conference on Decision and Control, Orlando FL, December 4-7, 2001*. Ed. by IEEE. Vol. 5. IEEE, pp. 4770–4775 (cited on p. 75).

[135] A. Bultheel, M. Van Barel, and P. Van gucht. (2004) *Orthogonal bases in discrete least squares rational approximation*. J. Comput. Appl. Math. **164-165**, pp. 175–194 (cited on pp. 75, 107).

[136] A. Bultheel and P. Van gucht. (2000) *Boundary asymptotics for orthogonal rational functions on the unit circle*. Acta Applicandae Mathematicae **61**, pp. 333–349 (cited on p. 85).

[137] A. Bultheel, P. Van gucht, and M. Van Barel. (2010) *Orthogonal rational functions with real coefficients and semiseparable matrices*. J. Comput. Appl. Math. **233(5)**, pp. 1192–1201 (cited on pp. 55, 75).

[138] A. Bunse-Gerstner, D. Kubalińska, G. Vossen, and D. Wilczek. (2010) *H^2-norm optimal model reduction for large scale discrete dynamical MIMO systems*. J. Comput. Appl. Math. **233(5)**, pp. 1202–1216 (cited on p. 108).

[139] A. Bunse-Gestner and L. Elsner. (1991) *Schur parameters pencils for the solution of the unitary eigenproblem*. Linear Algebra Appl. **154-156**, pp. 741–778 (cited on p. 25).

[140] R.B. Burckel. (1971) *An introduction to classical complex analysis*. Birkhäuser Verlag, Basel (cited on p. 21).

[141] F. Cala Rodríguez, P. González-Vera, and M. Jiménez Páiz. (1999) *Quadrature formulas for rational functions*. Electron. Trans. Numer. Anal. **9**, pp. 39–52 (cited on p. 51).

[142] F. Cala Rodríguez and G. López Lagomasino. (1998) *Multipoint Padé-type approximants. Exact rate of convergence*. Constr. Approx. **14**, pp. 259–272 (cited on p. 90).

[143] F. Cala Rodríguez and G. López Lagomasino. (2001) *Multipoint rational approximants with preassigned poles*. J. Math. Anal. Appl. **256(1)**, pp. 142–161 (cited on p. 90).

[144] B. de la Calle Ysern. (2010) *A walk though approximation theory*. In: *Recent Trends in Orthogonal Polynomials and Approximation Theory*. Ed. by J. Arvesú, F. Marcellán, and A. Martínez-Finkelshtein. Vol. 507. of Contemporary Mathematics. AMS, pp. 25–86 (cited on pp. 83, 110).

[145] B. de la Calle Ysern. (2005) *Error bounds for rational quadrature formulae of analytic functions*. Numer. Math. **101(2)**, pp. 251–271 (cited on pp. 77, 78).

[146] B. de la Calle Ysern. (2013) *Optimal extension of the Szegő quadrature*. IMA J. Numer. Anal. Submitted (cited on p. 80).

[147] B. de la Calle Ysern and P. González-Vera. (2007) *Rational quadrature formulas on the unit circle with arbitrary poles*. Numer. Math. **107(4)**, pp. 559–587 (cited on pp. 77, 78).

[148] B. de la Calle Ysern, G. Lopez Lagomasino, and L. Reichel. (2009) *Stieltjes-type polynomials on the unit circle*. Math. Comp. **78(266)**, pp. 969–997 (cited on p. 80).

[149] D. Calvetti, G.H. Golub, W.B. Gragg, and L. Reichel. (2000) *Computation of Gauss-Kronrod quadrature rules*. Math. Comp. **69(231)**, pp. 1035–1052 (cited on p. 80).

[150] D. Calvetti, S.-M. Kim, and L. Reichel. (2005) *Quadrature rules based on the Arnoldi process*. SIAM J. Matrix Anal. Appl. **26(3)**, pp. 765–781 (cited on p. 101).

[151] M. Camacho and P. González-Vera. (1992) *A note on para-orthogonality and biorthogonality*. In: *Det Kongelige Norske Videnskabers Selskab Skrifter* (No. 3), pp. 1–16 (cited on p. 55).

[152] M.J. Cantero, R. Cruz Barroso, and P. González-Vera. (2008) *A matrix approach to the computation of quadrature formulas on the interval*. Appl. Numer. Math. **58(3)**, pp. 296–318 (cited on p. 20).

[153] M.J. Cantero, L. Moral, and L. Velázquez. (2003) *Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle*. Linear Algebra Appl. **362**, pp. 29–56 (cited on p. 25).

[154] M.J. Cantero, L. Moral, and L. Velázquez. (2002) *Measures and para-orthogonal polynomials on the unit circle*. East J. Approx. **8**, pp. 447–464 (cited on p. 25).

[155] M.J. Cantero, L. Moral, and L. Velázquez. (2006) *Measures on the unit circle and unitary truncations of unitary operators*. J. Approx. Theory **139(1-2)**, pp. 430–468 (cited on p. 25).

[156] M.J. Cantero, L. Moral, and L. Velázquez. (2005) *Minimal representations of unitary operators and orthogonal polynomials on the unit circle*. Linear Algebra Appl. **408**, pp. 40–65 (cited on p. 25).

[157] K. Castillo, L. Garza, and F. Marcellán. (2010) *Linear spectral transformations, Hessenberg matrices, and orthogonal polynomials*. Rend. Circ. Mat. Palermo Serie II Suppl. 82, pp. 1–23 (cited on p. 61).

[158] G.N. Chen and Y.J. Hu. (1999) *The truncated Hamburger matrix moment problem in the nondegenerate and degenerate cases, and matrix continued fractions*. Linear Algebra Appl. **177**, pp. 199–236 (cited on p. 110).

[159] A. Chesnokov, K. Deckers, and M. Van Barel. (2010) *A numerical solution of the constrained weighted energy problem*. J. Comput. Appl. Math. **235(4)**, pp. 950–965 (cited on pp. 102, 111).

[160] T. Chihara. (1978) *An introduction to orthogonal polynomials*. Gordon and Breach Science Publishers, New York (cited on p. 12).

[161] A.E. Choque Rivero, Yu.M. Dyakarev, B. Fritzsche, and B. Kirstein. (2007) *A truncated matricial moment problem on a finite interval, the case of an odd number of given moments*. In: *System Theory, the Schur Algorithm and Multidimensional Analysis*. Ed. by D. Alpay and V. Vinnikov. Vol. 176. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 99–164 (cited on p. 97).

[162] A.E. Choque Rivero, A. Lasarow, and A. Rahn. (2011) *On ranges and Moore-Penrose inverses related to matrix Schur and Carathéodory functions*. Complex Anal. Oper. Theory **5(2)**, pp. 513–541 (cited on p. 97).

[163] A.E. Choque Rivero and S.M. Zagorodnyuk. (2012) *The strong matrix Stieltjes moment problem*. Bol. Soc. Mat. Mexicana **18(2)**, pp. 151–170 (cited on p. 97).

[164] M. Chu and G.H. Golub. (2005) *Inverse eigenvalue problems: Theory, algorithms, and applications*. Oxford University Press (cited on pp. 73, 74).

[165] M. Chu and G.H. Golub. *Structured inverse eigenvalue problems*. In: *Acta Numerica*. Vol. 11. Cambridge University Press, pp. 1–71 (cited on p. 74).

[166] S. Chung-Tsun. (2004) *Some inverse problems on Jacobi matrices*. Inverse Problems **20**, pp. 589–600 (cited on p. 74).

[167] L. Cochran and S.C. Cooper. (1994) *Orthogonal Laurent polynomials on the real line*. In: *Continued fractions and orthogonal functions*. Ed. by S.C. Cooper and W.J. Thron. Marcel Dekker, New York, pp. 47–100 (cited on p. 44).

[168] T. Constantinescu. (1996) *Schur analysis, factorization and dilation problems*. Vol. 82. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel (cited on p. 95).

[169] R. Cruz Barroso, L. Daruis, P. González-Vera, and O. Njåstad. (2005) *Quadrature rules for periodic integrands. Biorthogonality and para-orthogonality*. Ann. Math. Inf. **32**, pp. 5–44 (cited on p. 55).

[170] R. Cruz Barroso, L. Daruis, P. González-Vera, and O. Njåstad. (2007) *Sequences of orthogonal Laurent polynomials, bi-orthogonality and quadrature formulas on the unit circle*. J. Comput. Appl. Math. **200**, pp. 424–440 (cited on p. 55).

[171] R. Cruz Barroso and S. Delvaux. (2009) *Orthogonal Laurent polynomials on the unit circle and snake-shaped matrix factorizations*. J. Approx. Theory **161(1)**, pp. 65–87 (cited on p. 25).

[172] R. Cruz Barroso, C. Díaz Mendoza, and R. Orive. (2013) *Orthogonal Laurent polynomials. A new algebraic approach*. J. Math. Anal. Appl. **408(1)**, pp. 40–54 (cited on pp. 20, 25, 44).

[173] R. Cruz Barroso, P. González-Vera, and O. Njåstad. (2007) *On bi-orthogonal systems of trigonometric functions and quadrature formulas for periodic integrands*. Numer. Algorithms **44**, pp. 309–333 (cited on p. 55).

[174] R. Cruz Barroso, P. González-Vera, and F. Perdomo Pío. (2009) *An application of Szegő polynomials to the computation of certain weighted integrals on the real line*. Numer. Algorithms **52(3)**, pp. 273–293 (cited on p. 63).

[175] R. Cruz Barroso, P. González-Vera, and F. Perdomo Pío. (2011) *On the computation of the Fourier transform under the presence of nearby polar singularities*. New York J. Math. **17a**, pp. 127–146 (cited on p. 63).

[176] R. Cruz Barroso, P. González-Vera, and F. Perdomo Pío. (2010) *Orthogonality, interpolation and quadratures on the unit circle and the interval* $[-1, 1]$. J. Comput. Appl. Math. **235(4)**, pp. 966–981 (cited on p. 70).

[177] R. Cruz Barroso, P. González-Vera, and F. Perdomo Pío. (2009) *Quadrature formulas associated with Rogers-Szegő polynomials*. Comput. Math. Appl. **57(2)**, pp. 308–323 (cited on p. 63).

[178] R.E. Curto and L.A. Fialkow. (1998) *Flat extensions of positive moment matrices: recursively generated relations*. Vol. 648. of Trans. Amer. Math. Soc. AMS (cited on p. 27).

[179] E. Daems and A.B.J. Kuijlaars. (2004) *A Christoffel-Darboux formula for multiple orthogonal polynomials*. J. Approx. Theory **130(2)**, pp. 190–202 (cited on p. 111).

[180] D. Damanik, A. Pushnitski, and B. Simon. (2008) *The analytic theory of matrix orthogonal polynomials*. Surveys in Approximation Theory **4**, pp. 1–85 (cited on p. 95).

[181] S. Darlington. (1970) *Analytic approximations to approximations in the Chebyshev sense*. Bell Syst. Techn. J. **49**, pp. 1–32 (cited on p. 57).

[182] L. Daruis and P. González-Vera. (2002) *Interpolatory quadrature formulas on the unit circle for Chebyshev weight functions*. Numer. Math. **90(4)**, pp. 641–664 (cited on p. 60).

[183] L. Daruis, P. González-Vera, and M. Jiménez Páiz. (2006) *Quadrature formulas associated with rational modifications of the Chebyshev weight functions*. Comput. Math. Appl. **51**, pp. 419–430 (cited on p. 61).

[184] L. Daruis, P. González-Vera, and O. Njåstad. (2002) *Szegő quadrature formulas for certain Jacobi-type weight functions*. Math. Comp. **71(238)**, pp. 683–701 (cited on p. 60).

[185] L. Daruis, J. Hernández, and F. Marcellán. (2007) *Spectral transformations for Hermitian Toeplitz matrices*. J. Comput. Appl. Math. **202(2)**, pp. 155–176 (cited on p. 61).

[186] P.J. Davis and P. Rabinowitz. (1984) *Methods of numerical integration*. 2nd ed. First edition 1967. Academic Press (cited on pp. 8, 23).

[187] T.J. De Hoog. (2001) *Rational orthonormal bases and related transforms in linear system modeling*. PhD thesis. T.U. Delft (cited on pp. 107, 108).

[188] K. Deckers and A. Bultheel. (2008) *Associated rational functions based on a three-term recurrence relation for orthogonal rational functions*. IAENG Int. J. Appl. Math. **38(4)**, pp. 214–222 (cited on p. 110).

[189] K. Deckers and A. Bultheel. (2009) *Orthogonal rational functions and rational modifications of a measure on the unit circle*. J. Approx. Theory **157**, pp. 1–18 (cited on p. 61).

[190] K. Deckers and A. Bultheel. (2008) *Orthogonal rational functions, associated rational functions and functions of the second kind*. In: *Proceedings of the World Congress on Engineering 2008, vol. 2*. Ed. by S.I. Ao, L. Gelman, D.W.L. Hukins, A. Hunter, and A.M. Korsunsky. Lecture Notes in Engineering and Computer Science. CDROM, isbn 978-988-17012-3-7; Print: isbn: 978-988-98671-9-5. Newswood Limited, International Association of Engineers, London, United Kingdom, pp. 838–843 (cited on p. 110).

[191] K. Deckers and A. Bultheel. (2009) *Orthogonal rational functions with complex poles: The Favard theorem*. J. Math. Anal. Appl. **356(2)**, pp. 764–768 (cited on pp. 46, 52).

[192] K. Deckers and A. Bultheel. (2012) *Rational interpolation I: Least squares convergence*. J. Math. Anal. Appl. **395(2)**, pp. 455–464 (cited on p. 90).

[193] K. Deckers and A. Bultheel. (2007) *Rational Krylov sequences and orthogonal rational functions*. Technical Report TW499. Department of Computer Science, K.U.Leuven (cited on p. 102).

[194] K. Deckers and A. Bultheel. (2009) *Recurrence and asymptotics for orthogonal rational functions on an interval*. IMA J. Numer. Anal. **29**, pp. 1–23 (cited on pp. 51, 52, 89).

[195] K. Deckers and A. Bultheel. (2012) *The existence and construction of rational Gauss-type quadrature rules*. Appl. Math. Comput. **218(10)**, pp. 10299–10320 (cited on pp. 56, 69).

[196] K. Deckers, A. Bultheel, R. Cruz Barroso, and F. Perdomo Pío. (2010) *Positive rational interpolatory quadrature formulas on the unit circle and the interval*. Appl. Numer. Math. **60(12)**, pp. 1286–1299 (cited on pp. 52, 70).

[197] K. Deckers, A. Bultheel, and F. Perdomo Pío. (2011) *Rational Gauss-Radau and rational Szegő-Lobatto quadrature on the interval and the unit circle respectively*. Jaen J. Approx. **3(1)**, pp. 15–66 (cited on pp. 51, 70).

[198] K. Deckers, A. Bultheel, and J. Van Deun. (2010) *A generalized eigenvalue problem for quasi-orthogonal rational functions*. Numer. Math. **117(3)**, pp. 463–506 (cited on pp. 52, 56, 68).

[199] K. Deckers, M.J. Cantero, L. Moral, and L Velázquez. (2011) *An extension of the associated rational functions on the unit circle*. J. Approx. Theory **163(4)**, pp. 524–546 (cited on p. 110).

[200] K. Deckers, J. Van Deun, and A. Bultheel. (2007) *An extended relation between orthogonal rational functions on the unit circle and the interval* $[-1, 1]$. J. Math. Anal. Appl. **334(2)**, pp. 1260–1275 (cited on pp. 51, 59).

[201] K. Deckers, J. Van Deun, and A. Bultheel. (2006) *Computing rational Gauss-Chebyshev quadrature formulas with complex poles*. In: *Proceedings of the Fifth International Conference on Engineering Computational Technology*. Ed. by B. H. V. Topping, G. Montero, and R. Montenegro. Paper 30. Civil-Comp Press, Kippen, Stirlingshire, United Kingdom (cited on p. 59).

[202] K. Deckers, J. Van Deun, and A. Bultheel. (2009) *Computing rational Gauss-Chebyshev quadrature formulas with complex poles: the algorithm*. Adv. Eng. Softw. **40(8)**, pp. 707–717 (cited on p. 59).

[203] K. Deckers, J. Van Deun, and A. Bultheel. (2008) *Rational Gauss-Chebyshev quadrature formulas for complex poles outside* $[-1, 1]$. Math. Comp. **77(262)**, pp. 967–983 (cited on pp. 58, 59).

[204] Ph. Delsarte, Y. Genin, and Y. Kamp. (1981) *On the role of the Nevanlinna-Pick problem in circuit and system theory*. Internat. J. Circuit Th. Appl. **9**, pp. 177–187 (cited on pp. 27, 95).

[205] Ph. Delsarte, Y. Genin, and Y. Kamp. (1978) *Orthogonal polynomial matrices on the unit circle*. IEEE Trans. Circuits and Systems **25**, pp. 149–160 (cited on p. 95).

[206] Ph. Delsarte, Y. Genin, and Y. Kamp. (1979) *The Nevanlinna-Pick problem for matrix-valued functions*. SIAM J. Appl. Math. **36**, pp. 47–61 (cited on p. 95).

[207] S. Delvaux. (2007) *Rank structured matrices*. PhD. KU Leuven (cited on p. 75).

[208] S. Delvaux and M. Van Barel. (2005) *Orthonormal rational function vectors*. Numer. Math. **100(3)**, pp. 409 –440 (cited on p. 75).

[209] P. Dewilde. (1991) *A course on algebraic Schur and Nevanlinna-Pick interpolation problems*. In: *Algorithms and parallel VLSI architectures, volume A: tutorials*. Ed. by E.F. Deprettere and A.-J. van der Veen. Elsevier Science Publ., pp. 13–69 (cited on p. 27).

[210] P. Dewilde and E.F.A. Deprettere. (1988) *The generalized Schur algorithm: approximation and hierarchy*. In: *Topics in operator theory and interpolation*. Ed. by I. Gohberg. Vol. 29. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel, pp. 97–116 (cited on p. 98).

[211] P. Dewilde and H. Dym. (1981) *Lossless chain scattering matrices and optimum linear prediction : the vector case*. Internat. J. Circuit Th. Appl. **9**, pp. 135–175 (cited on pp. 96, 98).

[212] P. Dewilde and H. Dym. (1984) *Lossless inverse scattering, digital filters, and estimation theory*. IEEE Trans. Inf. Th. **IT-30**, pp. 644–662 (cited on pp. 31, 94–96, 107).

[213] P. Dewilde and H. Dym. (1981) *Schur recursions, error formulas, and convergence of rational estimators for stationary stochastic sequences*. IEEE Trans. Inf. Th. **IT-27**, pp. 446–461 (cited on pp. 14, 31, 57, 94, 95, 107).

[214] P. Dewilde and A.J. Van Der Veen. (1998) *Time-varying systems and computations*. Kluwer, Dordrecht (cited on p. 107).

[215] P. Dewilde, A. Viera, and T. Kailath. (1978) *On a generalized Szegő-Levinson realization algorithm for optimal linear predictors based on a network synthesis approach*. IEEE Trans. Circuits and Systems **CAS-25**, pp. 663–675 (cited on p. 95).

[216] C. Díaz Mendoza, P. González-Vera, and M. Jiménez Páiz. (2005) *Strong Stieltjes distributions and orthogonal Laurent polynomials with applications to quadratures and Padé approximation*. Math. Comp. **74(252)**, pp. 1843–1870 (cited on pp. 19, 20, 44).

[217] C. Díaz Mendoza, P. González-Vera, M. Jiménez Páiz, and F. Cala Rodríguez. (2005) *Orthogonal Laurent polynomials corresponding to certain strong Stieltjes distributions with applications to numerical quadratures*. Math. Comp. **75**, pp. 281–305 (cited on pp. 17, 19).

[218] C. Díaz Mendoza, P. González-Vera, M. Jiménez Páiz, and O. Njåstad. (2010) *Orthogonality and recurrence for ordered Laurent polynomial sequences*. J. Comput. Appl. Math. **235(4)**, pp. 982–997 (cited on pp. 20, 44, 102).

[219] C. Díaz Mendoza, P. González-Vera, M. Jiménez Páiz, and F. Cala Rodríguez. (2006) *Orthogonal Laurent polynomials corresponding to certain strong Stieltjes distributions with applications to numerical quadratures*. Math. Comp. **75(253)**, pp. 281–305 (cited on pp. 17, 19, 20, 44).

[220] M.M. Djrbashian. (1990) *A survey on the theory of orthogonal systems and some open problems*. In: *Orthogonal polynomials: Theory and practice*. Ed. by P. Nevai. Vol. 294. of Series C: Mathematical and Physical Sciences. NATO-ASI. Kluwer Academic Publishers, Boston, pp. 135–146 (cited on p. 57).

[221] M.M. Djrbashian. (1967) *Expansions by systems of rational functions with fixed poles*. Izv. Akad. Nauk Armyan. SSR **2**. (In Russian), pp. 3–51 (cited on p. 57).

[222] M.M. Djrbashian. (1962) *Expansions in systems of rational functions on a circle with a given set of poles*. Doklady Akademii Nauk SSSR **143**. (In Russian. Translation in Soviet Mathematics Doklady, vol. 3, 315–319, 1962), pp. 17–20 (cited on p. 57).

[223] M.M. Djrbashian. (1966) *Orthogonal systems of rational functions on the circle*. Izv. Akad. Nauk Armyan. SSR **1**. (In Russian), pp. 3–24 (cited on p. 57).

[224] M.M. Djrbashian. (1966) *Orthogonal systems of rational functions on the unit circle*. Izv. Akad. Nauk Armyan. SSR **1**. (In Russian), pp. 106–125 (cited on p. 57).

[225] M.M. Djrbashian. (1962) *Orthogonal systems of rational functions on the unit circle with given set of poles*. Doklady Akademii Nauk SSSR **147**. (In Russian. Translation in Soviet Mathematics Doklady, vol. 3, 1794–1798, 1962), pp. 1278–1281 (cited on p. 57).

[226] W.F. Donoghue Jr. (1974) *Monotone matrix functions and analytic continuation*. Springer, Berlin (cited on p. 95).

[227] V. Druskin and L. Knizhnerman. (1998) *Extended Krylov subspaces: approximation of the matrix square root and related functions*. SIAM J. Matrix Anal. Appl. **19(2)**, pp. 755–771 (cited on p. 102).

[228] V. Druskin, C. Lieberman, and M. Zaslavsky. (2010) *On adaptive choice of shifts in rational Krylov subspace reduction of evolutionary problems*. SIAM J. Sci. Comput. **32(5)**, pp. 2485–2496 (cited on p. 102).

[229] V. Druskin and V. Simoncini. (2011) *Adaptive rational Krylov subspaces for large-scale dynamical systems*. Systems Control Lett. **60(8)**, pp. 546–560 (cited on p. 102).

[230] V. Druskin, V. Simoncini, and M. Zaslavsky. (2013) *Adaptive tangential interpolation in rational Krylov subspaces for MIMO model reduction data*. SIAM J. Matrix Anal. Appl. To appear (cited on p. 103).

[231] V.K. Dubovoj, B. Fritzsche, and B. Kirstein. (1992) *Matricial version of the classical Schur problem*. Vol. 129. of Teubner-Texte zur Mathematik. Teubner Verlagsgesellschaft, Stuttgart, Leipzig (cited on p. 95).

[232] P.L. Duren. (1970) *The theory of H^p spaces*. Vol. 38. of Pure and Applied Mathematics. Academic Press, New York (cited on pp. 113, 114).

[233] H. Dym. (1989) *J-contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*. Vol. 71. of CBMS Regional Conf. Ser. in Math. Amer. Math. Soc., Providence, Rhode Island (cited on pp. 95, 97, 98, 107).

[234] H. Dym. (2004) *Linear fractional transformations, Riccati equations and bitangential interpolation, revisited*. In: *Reproducing kernel spaces and applications*. Ed. by D. Alpay. Vol. 143. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 171–212 (cited on p. 105).

[235] H. Dym, B. Fritzsche, V.E. Katsnelson, and B. Kirstein, eds. (1997) *Topics in interpolation theory*. Vol. 85. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel (cited on p. 105).

[236] S. Elhay, G.H. Golub, and J. Kautsky. (1992) *Jacobi matrices for sums of weight functions*. BIT **32**, pp. 143–166 (cited on p. 74).

[237] S. Elhay, G.H. Golub, and J. Kautsky. (1991) *Updating and downdating of orthogonal polynomials with data fitting applications*. SIAM J. Matrix Anal. Appl. **12**, pp. 327–353 (cited on p. 74).

[238] R.L. Ellis and I. Gohberg. (2002) *Orthogonal systems and convolution operators*. Vol. 140. of Oper. Theory: Adv. Appl. Birkhäuser Verlag (cited on p. 95).

[239] J. van den Eshof and M. Hochbruck. (2006) *Preconditioning Lanczos approximations to the matrix exponential*. SIAM J. Sci. Comput. **27(4)**, pp. 1438–1457 (cited on p. 102).

[240] E. Estrada and N. Hatano. (2008) *Communicability in complex networks*. Phys. Rev. E **77**, p. 036111 (cited on p. 99).

[241] E. Estrada and N. Hatano. (2009) *Returnability in complex directed networks (digraphs)*. Linear Algebra Appl. **490(8-9)**, pp. 1886–1896 (cited on p. 99).

[242] E. Estrada, N. Hatano, and M. Benzi. (2012) *The physics of communicability in complex networks*. Physics reports **514(3)**, pp. 89–119 (cited on p. 99).

[243] D. Fasino, N. Mastronardi, and M. Van Barel. (2002) *Fast and stable algorithms for reducing diagonal plus semiseparable matrices to tridiagonal and bidiagonal form*. Tech. rep. IRMA 3/2002. I.R.M.A. - C.N.R., Bari (It.) (cited on p. 75).

[244] H. Fassbender. (1997) *Inverse unitary eigenproblems and related orthogonal functions*. Numer. Math. **77**, pp. 323–345 (cited on p. 74).

[245] I.P. Fedčina. (1972) *A criterion for the solvability of the Nevanlinna-Pick tangent problem*. Mat. Issled. (Kishinev) **7**. (In Russian), (See MR 50 # 4960), pp. 213–227 (cited on p. 97).

[246] I.P. Fedčina. (1975) *A description of the solutions of the Nevanlinna-Pick tangent problem*. Dokl. Akad. Nauk Armjan SSR, Ser. Mat. **60**. (In Russian), (See MR 52 # 5974), pp. 37–42 (cited on p. 97).

[247] I.P. Fedčina. (1978) *The Schur problem for vector valued functions*. Ukrain Mat. Z. **30(6)**, pp. 797–805,861 (cited on p. 97).

[248] I.P. Fedčina. (1975) *The tangential Nevanlinna-Pick problems with multiple points*. Dokl. Akad. Nauk Armjan SSR, Ser. Mat. **61**. (MR 53 # 13576), pp. 214–218 (cited on p. 97).

[249] L. Fejér. (1933) *Mechanische Quadraturen mit positiven Cotesschen Zahlen*. Math. Z. **37**, pp. 287–309 (cited on p. 7).

[250] L. Fialkow and J. Nie. (2012) *The truncated moment problem via homogenization and flat extensions*. J. Funct. Anal. **263(6)**, pp. 1682–1700 (cited on p. 27).

[251] L.A. Fialkow. (2011) *Solution of the truncated moment problem with variety $y = x^3$*. Trans. Amer. Math. Soc. **363**, pp. 3133–3165 (cited on p. 27).

[252] L.A. Fialkow. (2008) *Truncated multivariable moment problems with finite variety*. J. Operator Theory **60(2)**, pp. 343–377 (cited on p. 27).

[253] L.A. Fialkow and J. Nie. (2013) *On the closure of positive flat moment matrices*. J. Operator Theory **69(1)**, pp. 257–277 (cited on p. 27).

[254] U. Fidalgo Prieto, J.R. Illán González, and G. López Lagomasino. (2007) *Convergence and computation of simultaneous rational quadrature formulas*. Numer. Math. **106(1)**, pp. 99–128 (cited on p. 111).

[255] G. Filipuk, W. Van Assche, and L. Zhang. (2013) *Ladder operators and differential equations for multiple orthogonal polynomials*. J. Phys. A **46(20)**, id:205204 (cited on p. 111).

[256] B. Fischer and G.H. Golub. (1992) *How to generalte unknown orthogonal polynomials out of known orthogonal polynomials*. J. Comput. Appl. Math. **43(1-2)**, pp. 99–115 (cited on p. 61).

[257] C. Foiaş and A. Frazho. (1990) *The commutant lifting approach to interpolation problems*. Vol. 44. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel (cited on p. 96).

[258] C. Foiaş, A. Frazho, and W.S. Li. (1995) *On H^2 minimization for the Carathéodory-Schur interpolation problem*. Integral Equations Operator Theory **21**, pp. 24–32 (cited on p. 96).

[259] C. Foiaş and A.E. Frazho. (1986) *On the Schur representation in the commutant lifting theorem I*. In: *I. Schur methods in operator theory and signal processing*. Ed. by I. Gohberg. Vol. 18. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel, pp. 207–218 (cited on p. 96).

[260] C. Foiaş, A.E. Frazho, I. Gohberg, and M.A. Kaashoek. (1998) *Metric constrained interpolation, commutant lifting and systems*. Vol. 100. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel (cited on p. 96).

[261] A.E. Frazho. (1986) *Three inverse scattering algorithms for the lifting theorem*. In: *I. Schur methods in operator theory and signal processing*. Ed. by I. Gohberg. Vol. 18. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel, pp. 219–248 (cited on p. 96).

[262] G. Freud. (1971) *Orthogonal polynomials*. Pergamon Press, Oxford (cited on pp. 12, 82).

[263] R.W. Freund. (2003) *Model reduction methods based on Krylov subspaces*. In: *Acta Numerica*, 267319 (cited on p. 105).

[264] B. Fritzsche and B. Kirstein, eds. (1991) *Ausgewählte Arbeiten zu den Ursprüngen der Schur-Analysis*. Vol. 16. of Teubner-Archiv zur Mathematik. Teubner, Leipzig (cited on p. 95).

[265] B. Fritzsche, B. Kirstein, and A. Lasarow. (2003) *On a moment problem for rational matrix-valued functions*. Linear Algebra Appl. **372**, pp. 1–31 (cited on pp. 28, 96).

[266] B. Fritzsche, B. Kirstein, and A. Lasarow. (2012) *On maximal weight solutions of a moment problem for rational matrix-valued functions*. In: *Interpolation, Schur Functions and Moment Problems II*. Ed. by D. Alpay and B. Kirstein. Vol. 226. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 251–300 (cited on p. 28).

[267] B. Fritzsche, B. Kirstein, and A. Lasarow. (2004) *On rank invariance of moment matrices of nonnegative Hermitian-valued Borel measures on the unit circle*. Math. Nachr. **263-264(1)**, pp. 103–132 (cited on p. 28).

[268] B. Fritzsche, B. Kirstein, and A. Lasarow. (2005) *Orthogonal rational matrix-valued functions on the unit circle*. Math. Nachr. **278(5)**, pp. 525–553 (cited on pp. 28, 96).

[269] B. Fritzsche, B. Kirstein, and A. Lasarow. (2005) *Orthogonal rational matrix-valued functions on the unit circle*. Math. Nachr. **278(5)**, pp. 525–553 (cited on p. 96).

[270] B. Fritzsche, B. Kirstein, and A. Lasarow. (2006) *Orthogonal rational matrix-valued functions on the unit circle: Recurrence relations and a Favard-type theorem*. Math. Nachr. **279**, pp. 513–542 (cited on pp. 28, 96).

[271] B. Fritzsche, B. Kirstein, and A. Lasarow. (2012) *Para-orthogonal rational matrix-valued functions on the unit circle*. Oper. Matrices **6(4)**, pp. 631–680 (cited on pp. 28, 96).

[272] B. Fritzsche, B. Kirstein, and A. Lasarow. (2005) *Szegő pairs of orthogonal rational matrix-valued functions on the unit circle*. In: *Operator theory and indefinite inner product spaces*. Ed. by M. Langer, A. Luger, and H. Woracek. Vol. 163. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 163–189 (cited on p. 96).

[273] B. Fritzsche, B. Kirstein, and K. Müller. (1998) *An analysis of the block structure of certain subclasses of j_{qq}-inner functions*. Z. Anal. Anwendungen **17(2)**, 459478 (cited on p. 98).

[274] B. Fritzsche, B. Kirstein, and K. Müller. (1998) *An analysis of the block structure of j_{qq}-inner functions*. In: *Contributions to Operator Theory in Spaces with an Indefinite Metric*. Ed. by A. Dijksma, I. Gohberg, M.A. Kaashoek, and Mennicken R. Vol. 106. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 157–185 (cited on p. 98).

[275] D. Galant. (1971) *An implementation of Christoffel's theorem in the theory of orthogonal polynomials*. Math. Comp. **25(113)**, pp. 111–113 (cited on p. 61).

[276] K. Gallivan, E.J. Grimme, and P. Van Dooren. (1996) *A rational Lanczos algorithm for model reduction*. Numer. Algorithms **12**, pp. 33–64 (cited on p. 105).

[277] K. Gallivan, A. Vandendorpe, and P. Van Dooren. (2004) *Model reduction of MIMO systems via tangential interpolation*. SIAM J. Matrix Anal. Appl. **26(2)**, pp. 328–349 (cited on p. 105).

[278] J. B. Garnett. (1981) *Bounded analytic functions*. Academic Press, New York (cited on pp. 21, 113–115).

[279] L. Garza, F. Hernández, and F. Marcellán. (2010) *Orthogonal polynomials and measures on the unit circle. The Geronimus transformations*. J. Comput. Appl. Math. **233(5)**, pp. 1220–1231 (cited on p. 61).

[280] L. Garza and F. Marcellán. (2010) *Quadrature rules on the unit circle. a survey*. In: *Recent trend is orthogonal polynomials and approximation theory*. Ed. by J. Arvesú, F. Marcellán, and A. Martínez-Finkelstein. Vol. 507. of Contemporary Mathematics. AMS, pp. 113–139 (cited on p. 61).

[281] L. Garza and F. Marcellán. (2009) *Szegő transformations and Nth order associated polynomials on the unit circle*. Comput. Math. Appl. **57(10)**, pp. 1659–1671 (cited on p. 61).

[282] L. Garza and F. Marcellán. (2009) *Verblunski parameters and linear spectral transformations*. J. Math. Anal. Appl. **16**, pp. 69–89 (cited on p. 61).

[283] L.E. Garza Gaone. (2009) *Transformatciones espectrales, funciones de Carathéodory y polynomios ortogonales en la circunferencia unidad*. PhD thesis. Univ. Carlos III de Madrid (cited on p. 61).

[284] W. Gautschi. (1981) *A survey of Gauss-Christoffel quadrature formulae*. In: *E.B. Christoffel. The influence of his work on mathematical and physical sciences*. Ed. by P.L. Butzer and F. Fehér. Birkhäuser Verlag, Basel, pp. 72–147 (cited on p. 9).

[285] W. Gautschi. (1999) *Algorithm 793: GQRAT - Gauss quadrature for rational functions*. ACM Trans. Math. Software **25**, pp. 213–139 (cited on pp. 102, 110).

[286] W. Gautschi. (1968) *Construction of Gauss–Christoffel quadrature formulas*. Math. Comp. **22**, pp. 251–270 (cited on p. 23).

[287] W. Gautschi. (1988) *Gauss-Kronrod quadrature - a survey*. In: *Numerical Methods and Approximation Theory III*. Ed. by G.V. Milivanović. University Nič, pp. 39–66 (cited on p. 79).

[288] W. Gautschi. (1993) *Gauss-type quadrature rules for rational functions*. In: *Numerical integration IV*. Ed. by H. Brass and G. Hämmerlin. Vol. 112. of Internat. Ser. of Numer. Math. Birkhäuser Verlag, Basel, pp. 111–130 (cited on pp. 102, 110).

[289] W. Gautschi. (2004) *Generalized Gauss-Radau and Gauss-Lobatto formulae*. BIT **44(4)**, pp. 711–720 (cited on p. 70).

[290] W. Gautschi. (1993) *On the computation of generalized Fermi-Dirac and Bose-Einstein integrals*. Computer Physics Comm. **74**, pp. 233–238 (cited on p. 110).

[291] W. Gautschi. (1999) *Orthogonal polynomials and quadrature*. Electron. Trans. Numer. Anal. **9**, pp. 65–76 (cited on pp. 12, 23).

[292] W. Gautschi. (2004) *Orthogonal polynomials computation and approximation*. Oxford University Press, Oxford (cited on pp. 12, 23).

[293] W. Gautschi. (2002) *The interplay between classical analysis and (numerical) linear algebra: a tribute to Gene H. Golub*. Electron. Trans. Numer. Anal. **i13**, pp. 119–147 (cited on pp. 23, 61).

[294] W. Gautschi, L. Gori, and M.L. Lo Cascio. (2000) *Quadrature rules for rational functions*. Numer. Math. **86(4)**, pp. 617–633 (cited on p. 110).

[295] W. Gautschi, H.J. Landau, and G.V. Milovanović. (1987) *Polynomials orthogonal on the semicircle, II*. Constr. Approx. **3(1)**, pp. 389–404 (cited on p. 17).

[296] W. Gautschi and G.V. Milovanović. (1986) *Polynomials orthogonal on the semicircle*. J. Approx. Theory **46**, pp. 230–250 (cited on p. 17).

[297] Ya. Geronimus. (1961) *Orthogonal polynomials*. Consultants Bureau, New York (cited on p. 12).

[298] Ya. Geronimus. *Polynomials orthogonal on a circle and their applications*. In: vol. 3. of Transl. Math. Monographs. Amer. Math. Soc., pp. 1–78 (cited on p. 12).

[299] K. Glover. (1984) *All optimal Hankel norm appoximations of linear multivariable systems and their L_∞ error bounds*. Internat. J. Control **39(6)**, pp. 1115–1193 (cited on p. 104).

[300] E. Godoy and F. Marcellan. (1993) *Orthogonal polynomials and rational modification of measures*. Canad. J. Math. **45**, pp. 930–943 (cited on p. 61).

[301] E. Godoy and F. Marcellan. (1991) *Orthogonal polynomials on the unit circle: distribution of zeros*. J. Comput. Appl. Math. **37(1-3)**, pp. 195–208 (cited on p. 61).

[302] I. Gohberg, ed. (1986) *I. Schur methods in operator theory and signal processing*. Vol. 18. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, Basel (cited on pp. 21, 95, 148).

[303] L. Golinskii. (2002) *Quadrature formula and zeros of para-orthogonal polynomials on the unit circle*. Acta Math. Hungar. **96**, pp. 169–186 (cited on p. 37).

[304] G.H. Golub and J. Kautsky. (1983) *Calculation of Gauss quadratures with multiple free and fixed knots*. Numer. Math. **41**, pp. 147–163 (cited on pp. 61, 71).

[305] G.H. Golub and G. Meurant. (1994) *Matrices, moments and quadrature*. In: *Numerical analysis 1993*. Vol. 303. of Pitman Research Notes in Mathematics. Longman Scientific & Technical (cited on pp. 23, 74).

[306] G.H. Golub and G. Meurant. (2009) *Matrices, moments and quadrature with applications*. Princeton Series in Applied Mathematics. Princeton University Press (cited on pp. 23, 74).

[307] G.H. Golub and C.F. Van Loan. (1996) *Matrix computations*. 3rd edition. John Hopkins University Press (cited on p. 23).

[308] G.H. Golub and J.H. Welsch. (1969) *Calculation of Gauss quadrature rules*. Math. Comp. **23(106)**, pp. 221–230 (cited on pp. 23, 53).

[309] A. Gonchar and G. López. (1978) *On Markov's theorem for multipoint Padé approximants for functions of Stieltjes type*. Math. USSR-Sb. **105**. English translation. *Math. USSR-Sb*, 34 (1978) 449-459, pp. 512–524 (cited on p. 110).

[310] P. González-Vera, H. Martínez, and J.J. Trujillo. (2007) *An application of Szegő quadratures to the computation of the Fourier transform*. Appl. Math. Comput. **187(1)**, pp. 183–194 (cited on p. 63).

[311] P. González-Vera and O. Nåstad. (1990) *Szegő functions and multipoint Padé approximation*. J. Comput. Appl. Math. **32(1-2)**, pp. 107–116 (cited on p. 26).

[312] W. B. Gragg and W. J. Harrod. (1984) *The numerically stable reconstruction of Jacobi matrices from spectral data*. Numer. Math. **44**, pp. 317–335 (cited on p. 74).

[313] W.B. Gragg. (1982) *Positive definite Toeplitz matrices, the Arnoldi process for isometric operators and Gaussian quadrature on the unit circle*. In: *Numerical methods in linear algebra*. Ed. by E.S. Nikolaev. (Russian, English translation in J. Comput. Appl. Math., 46 (1-2) (1993) 183-198.) Moscow university press, pp. 16–32 (cited on p. 25).

[314] W.B. Gragg. (1886) *The QR algorithm for unitary Hessenberg matrices*. J. Comput. Appl. Math. **16**, pp. 1–8 (cited on p. 25).

[315] W.B. Gragg and L. Reichel. (1990) *A divide and conquer method for unitary orthogonal eigenproblems*. Numer. Math. **57**, pp. 695–718 (cited on p. 25).

[316] U. Grenander and G. Szegő. (1958) *Toeplitz forms and their applications*. University of California Press, Berkley (cited on pp. 12, 82, 106).

[317] M. Gu, R. Guzzo, X.B. Chi, and X.Q. Cao. (2003) *A stable divide and conquer algorithm for the unitary eigenproblem*. SIAM J. Matrix Anal. Appl. **25(2)**, pp. 385–404 (cited on p. 25).

[318] S. Gugercin, A.C. Antoulas, and C.A. Beattie. (2008) H^2 *model reduction for large-scale linear dynamical systems*. SIAM J. Numer. Anal. **30(2)**, pp. 609–638 (cited on pp. 102, 108).

[319] S. Güttel. (2013) *Rational Krylov approximation of matrix functions: numerical methods and optimal pole selection*. GAMM Mitteilungen **36(1)**, pp. 8–31 (cited on p. 102).

[320] S. Güttel and L. Knizhnerman. (2013) *A black-box rational Arnoldi variant for Cauchy-Stieltjes matrix functions*. BIT **53(3)**, pp. 595–616 (cited on p. 102).

[321] O. Hald. (1976) *Inverse eigenvalue problems for Jacobi matrices*. Linear Algebra Appl. **14(1)**, pp. 63–85 (cited on p. 74).

[322] N. Hale and A. Townsend. (2013) *Fast and accurate computation of Gauss-Legendre and gauss-jacobi quadrature nodes and weights*. SIAM J. Sci. Comput. **35(1)**, A652–A672 (cited on p. 111).

[323] Y. Halevi. (2006) *Projection properties of the* L_2 *optimal reduced order model*. Internat. J. Control **79(4)**, pp. 298–310 (cited on p. 108).

[324] H. Hamburger. (1920) *Ueber eine Erweiterung des Stieltjesschen Moment Problems I*. Math. Ann. **81**, pp. 235–319 (cited on p. 17).

[325] H. Hamburger. (1921) *Ueber eine Erweiterung des Stieltjesschen Moment Problems II*. Math. Ann. **82**, pp. 120–164 (cited on p. 17).

[326] H. Hamburger. (1921) *Ueber eine Erweiterung des Stieltjesschen Moment Problems III*. Math. Ann. **82**, pp. 168–187 (cited on p. 17).

[327] M. Haneczok and A. Van Assche. (2012) *Interlacing properties of zeros of multiple orthogonal polynomials*. J. Math. Anal. Appl. **389(1)**, pp. 529–438 (cited on p. 111).

[328] F. Hausdorff. (1923) *Momentprobleme für ein endliches Interval*. Math. Z. **16**, pp. 220–248 (cited on p. 16).

[329] J.W. Helton. (1987) *Operator theory, analytic functions, matrices and electrical engineering*. Vol. 68. of CBMS Regional Conf. Ser. in Math. Amer. Math. Soc., Rhode Island, N.Y. (cited on p. 95).

[330] J.W. Helton, J.A. Ball, C.R. Johnson, and J.N. Palmer. (1991) *Operator theory, analytic functions, matrices , and electrical engineering*. Vol. 68. of Regional Conference Series Math. AMS, Providence, R.I. (cited on p. 95).

[331] E. Hendriksen and O. Njåstad. (1989) *A Favard theorem for rational functions*. J. Math. Anal. Appl. **142(2)**, pp. 508–520 (cited on p. 46).

[332] E. Hendriksen and O. Njåstad. (1989) *Positive multipoint Padé continued fractions*. Proc. Edinburgh Math. Soc. **32**, pp. 261–269 (cited on pp. 26, 110).

[333] E. Hendriksen and H. van Rossum. (1986) *Orthogonal Laurent polynomials*. Proc. of the Kon. Nederl. Akad. Wetensch, Proceedings A **89(1)**, pp. 17–36 (cited on p. 44).

[334] P. Henrici. (1988) *Applied and computational complex analysis: Special functions, integral transforms, asymptotics, continued fractions*. Vol. II. of Wiley Classics Library. John Wiley & Sons, New York - London - Sydney - Toronto (cited on pp. 27, 115).

[335] M. Bello Hernández and G. López Lagomasino. (1998) *Ratio and relative asymptotics of polynomials orthogonal on an arc of the unit circle*. J. Approx. Theory **92**, pp. 216–244 (cited on p. 18).

[336] P.S.C. Heuberger. (1991) *On approximate system identification with system based orthonormal functions*. PhD thesis. T.U.Delft (cited on pp. 107, 108).

[337] P.S.C. Heuberger, P.M.J. Van den Hof, and B. Wahlberg, eds. (2005) *Modelling and identification with rational orthogonal basis functions*. Springer Verlag (cited on pp. 107, 108).

[338] N. Higham. (2008) *Functions of matrices - Theory and computation*. SIAM (cited on p. 102).

[339] B.L. Ho and R.E. Kalman. (1966) *Effective construction of linear state-variable models from input/ output functions*. Regelungstechnik **14(12)**, pp. 545–548 (cited on p. 104).

[340] K. Hoffman. (1962) *Banach spaces of analytic functions*. Prentice-Hall, Englewood Cliffs (cited on p. 113).

[341] J.W. Hoffmann and P.A. Fuhrmann. (1996) *On balanced realizations of bounded real and positive real functions*. Linear Algebra Appl. **245**, pp. 107–146 (cited on p. 104).

[342] O. Holtz and M. Tyaglov. (2012) *Structured matrices, continued fractions, and root localization of polynomials*. SIAM Rev. **54(3)** (cited on p. 110).

[343] Y.J. Hu and G.N. Chen. (1998) *A unified treatment of the Stieltjes moment problem in both nondegenerate and degenerate cases, and matrix continued fractions*. Linear Algebra Appl. **2770**, pp. 199–236 (cited on p. 110).

[344] P. Hueberger, P. Van den Hof, and O.H. Bosgra. (1995) *A generalized orthonormal basis for linear dynamical systems*. IEEE Trans. Automat. Control **40(3)**, pp. 451–465 (cited on p. 108).

[345] M. Humet and M. Van Barel. (2014) *Algorithms for the Geronimus transformation for orthogonal polynomials on the unit circle*. J. Comput. Appl. Math. **267**, pp. 195–217 (cited on p. 61).

[346] D.C. Hyland and D.S. Bernstein. (1985) *The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore*. IEEE Trans. Automat. Control **30(12)**, pp. 1201–1211 (cited on p. 108).

[347] J. Illán González and G. López Lagomasino. (1984) *A note on generalized quadrature formulas of Gauss-Jacobi type*. In: *Constructive Theory of Functions '84, Sofia*. Publishing House of the Bulgarian Academy of Science, Sofia, pp. 513–518 (cited on p. 90).

[348] C. Jagels and L. Reichel. (2012) *Recursion relations for the extended Krylov subspace method*. Linear Algebra Appl. **434**, pp. 1716–1732 (cited on p. 102).

[349] C. Jagels and L. Reichel. (2007) *Szegő-Lobatto quadrature rules*. J. Comput. Appl. Math. **200(1)**, pp. 116–126 (cited on p. 65).

[350] C. Jagels and L. Reichel. (2013) *The structure of matrices in rational Gauss quadrature*. Math. Comp. **82**, pp. 2035–2060 (cited on p. 102).

[351] E.X. Jiang. (2003) *An inverse eigenvalue problem for Jacobi matrices*. J. Comput. Appl. Math. **21(5)**, pp. 569–584 (cited on p. 74).

[352] W.B. Jones and O. Njåstad. (1999) *Orthogonal Laurent polynomials and strong moment theory*. J. Comput. Appl. Math. **105**, pp. 51–91 (cited on p. 20).

[353] W.B. Jones, O. Njåstad, and W.J. Thron. (1983) *Continued fractions and strong Hamburger moment problems*. Proc. London Math. Soc., Ser III **47(3)**, pp. 363–384 (cited on pp. 17, 20).

[354] W.B. Jones, O. Njåstad, and W.J. Thron. (1986) *Continued fractions associated with the trigonometric moment problem and other strong moment problems*. Constr. Approx. **2**, pp. 197–211 (cited on p. 110).

[355] W.B. Jones, O. Njåstad, and W.J. Thron. (1989) *Moment theory, orthogonal polynomials, quadrature and continued fractions associated with the unit circle*. Bull. London Math. Soc. **21**, pp. 113–152 (cited on pp. 9, 10, 12, 110).

[356] W.B. Jones, O. Njåstad, and W.J. Thron. (1984) *Orthogonal Laurent polynomials and the strong Hamburger moment problem*. J. Math. Anal. Appl. **98**, pp. 528–554 (cited on pp. 20, 44).

[357] W.B. Jones, O. Njåstad, and W.J. Thron. (1987) *Perron-Carathéodory continued fractions*. In: *Rational approximation and its applications in mathematics and physics*. Ed. by J. Gilewicz, M. Pindor, and W. Siemaszko. Vol. 1237. of Lecture Notes in Math. Springer, Berlin, pp. 188–206 (cited on p. 110).

[358] W.B. Jones, O. Njåstad, and W.J. Thron. (1986) *Schur fractions, Perron-Carathéodory fractions and Szegő polynomials, a survey*. In: *Analytic theory of continued fractions II*. Ed. by W.J. Thron. Vol. 1199. of Lecture Notes in Math. Springer, pp. 127–158 (cited on p. 110).

[359] W.B. Jones, O. Njåstad, and W.J. Thron. (1983) *Two-point Padé expansions for a family of analytic functions*. J. Comput. Appl. Math. **9**, pp. 105–124 (cited on p. 44).

[360] W.B. Jones, V. Petersen, and H. Waadeland. (2003) *Convergence of PPC-continued fraction approximants in frequency analysis*. Rocky Mountain J. Math. **33(2)**, pp. 525–542 (cited on p. 110).

[361] W.B. Jones and W.J. Thron. (1984) *Orthogonal Laurent polynomials and Gaussian quadrature*. In: *Quantum mechanics in mathematics, chemistry and physics*. Ed. by K. Gustafson and W.P. Reinhardt. Plenum, New York, pp. 449–455 (cited on p. 17).

[362] W.B. Jones and W.J. Thron. (1982) *Survey of continued fraction methods of solving moment problems and related topics*. In: *Analytic Theory of Continued Fractions*. Ed. by W.B. Jones, W.J. Thron, and H. Waadeland. Vol. 932. of Lecture Notes in Math. Springer, Berlin, pp. 4–37 (cited on p. 110).

[363] W.B. Jones, W.J. Thron, and H. Waadeland. (1980) *A strong Stieltjes moment problem*. Trans. Amer. Math. Soc. **206**, pp. 503–528 (cited on pp. 17, 19, 20).

[364] H. Joulak and B. Beckermann. (2009) *On Gautschi's conjecture for generalized Gauss-Radau and Gauss-Lobatto formulae*. J. Comput. Appl. Math. **233(3)**, pp. 768–774 (cited on p. 71).

[365] V. Kaliaguine and A. Ronveaux. (1996) *On a system of 'classical' polynomials of simultaneous orthogonality*. J. Comput. Appl. Math. **67(2)**, pp. 207–217 (cited on p. 111).

[366] R.E. Kalman. (1963) *Mathematical description of linear dynamical systems*. SIAM J. Control **1**, pp. 152–192 (cited on p. 104).

[367] R.E. Kalman. (1979) *On partial realizations, transfer functions and canonical forms*. Acta Polytechnica Scandinavica **31**, pp. 9–32 (cited on p. 104).

[368] V.E. Katsnelson. (1994) *Left and right Blaschke-Potapov products and Arov-singular J-inner functions* (cited on p. 97).

[369] L. Katz. (1953) *A new status index derived from sociometric index*. Psychometrika **18(1)**, pp. 39–43 (cited on p. 99).

[370] J. Kautsky and G.H. Golub. (1983) *On the calculation of Jacobi matrices*. Linear Algebra Appl. **52-53**, pp. 439–455 (cited on p. 61).

[371] W.H. Kautz. (1954) *Transient synthesis in the time domain*. IRE Trans. on Circuits Theory **1**, pp. 29–39 (cited on pp. 57, 108).

[372] A.Ya. Kheifets and P.M. Yuditskii. (1994) *An analysis and extension of V.P. Potapov's approach to interpolation problems with applications to the generalized bi-tangential Schur-Nevanlinna-Pick problem and J-inner-outer factorization*. In: *Matrix and operator valued functions*. Ed. by I. Gohberg and L.A. Sakhnovich. Vol. 72. of Oper. Theory: Adv. Appl. Birkhäuser Verlag, pp. 133–161 (cited on p. 105).

[373] S. Khrushchev. (2001) *Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fraction in $L^2(T)$*. J. Approx. Theory **108(2)**, pp. 161–248 (cited on pp. 94, 110).

[374] S.-M. Kim and L. Reichel. (2007) *Anti Szegő quadrature rules*. Math. Comp. **76**, pp. 795–810 (cited on p. 79).

[375] G. Klein and J.-P. Berrut. (2010) *Linear barycentric rational quadrature*. BIT **52**, pp. 407–424 (cited on p. 63).

[376] J. Kleinberg. (1999) *Authoritative sources in a hyperlinked environment*. J. ACM **46(5)**, pp. 604–632 (cited on p. 102).

[377] L. Klotz and A. Lasarow. (2003) *Extremal problems for matrix-valued polynomials on the unit circle and applications to multivariate stationary sequences*. J. Approx. Theory **125**, pp. 42–62 (cited on p. 28).

[378] P. Koosis. (1998) *Introduction to H_p spaces*. 2nd. Vol. 115. of Cambridge Tracts in Mathematics. Cambridge University Press (cited on pp. 113–115).

[379] M.A. Kovačević. (2000) *Construction of generalized Gauss-Turán quadrature rules for systems of arbitrary functions*. J. Math. Anal. Appl. **40(8-9)**, pp. 895–909 (cited on p. 71).

[380] M.G. Krein. (1951) *The ideas of P.L. Tchebycheff and A.A. Markov in the theory of limiting values of integrals and their further developments*. Uspehi Matem. Nauk **6**. (Russian) Transl. Amer. Math. Soc. Transl. (2), 12 (1959), pp. 1-121, pp. 3–120 (cited on p. 110).

[381] A.S. Kronrod. (1964) *Nodes and weights for quadrature formulas (sixteen-places tables)*. Engl. transl. Consultants Bureau, N.Y., 1965. Nauka (cited on p. 79).

[382] A. Kuijlaars, K. McLaughlin, W. Van Assche, and M. Vanlessen. (2010) *The Riemann-Hilbert approach to strong asymptotics for orthogonal polynomials on* $[-1, 1]$. Adv. Math. **188(2)**, pp. 337–398 (cited on p. 111).

[383] H.J. Landau, ed. (1987) *Moments in mathematics*. Vol. 37. of Proc. Sympos. Appl. Math. Amer. Math. Soc., Providence, R.I. (cited on p. 16).

[384] H. Langer and A. Lasarow. (2004) *Solution of a multiple Nevanlinna-Pick problem via orthogonal rational functions*. J. Math. Anal. Appl. **293(2)**, pp. 605–632 (cited on pp. 28, 96).

[385] A.N. Langville and C.D. Meyer. (2006) *Google's PageRank and beyond: the science of search engine ranking*. Princeton University Press (cited on p. 99).

[386] A. Lasarow. (2000) *Aufbau einer Szegő-Theorie für rationale Matrixfunktionen*. PhD thesis. Universität Leipzig, Fak. Mathematik Informatik (cited on p. 28).

[387] A. Lasarow. (2006) *Dual pairs of orthogonal systems of rational matrix-valued functions on the unit circle*. Analysis **26**, pp. 209–244 (cited on p. 96).

[388] A. Lasarow. (2008) *Solution of a multiple Nevanlinna-Pick problem for Carathéodory functions via orthogonal rational functions in the matrix case*. Math. Nachr. **281(1)**, pp. 75–108 (cited on p. 96).

[389] D.P. Laurie. (1996) *Anti-Gaussian quadrature formulas*. Math. Comp. **76**, pp. 739–747 (cited on p. 78).

[390] D.P. Laurie. (2007) *Calculation of Radau-Kronrod and Lobatto-Kronrod quadrature formulas*. Numer. Algorithms **45(1-4)**, pp. 139–152 (cited on p. 80).

[391] Y.W. Lee. (1933) *Synthesis of electrical networks by means of the Fourier transforms of Laguerre functions*. J. Math. Phys. **11**, pp. 83–113 (cited on pp. 57, 108).

[392] X. Li and K. Pan. (1996) *Strong and weak convergence of rational functions orthogonal on the unit circle*. J. London Math. Soc. **53**, pp. 289–301 (cited on p. 57).

[393] Z. Li and X. Zhao. (2009) *Inverse eigenvalue problem for generalized periodic Jacobi matrices with linear relation*. In: *Third International Symposium on Intelligent Information Technology Application, 2009. IITA 2009*. Vol. 1, pp. 764–766 (cited on p. 74).

[394] A. Liu, D. Qu, and L. Lu. (2010) *An inverse eigenvalue problem for symmetric arrow-plus-Jacobi matrices*. In: *2010 International Conference on Computational and Information Sciences (ICCIS)*, pp. 764–766 (cited on p. 74).

[395] R. Lobatto. (1852) *Lessen over de differentiaal- en integraalrekening, deel 2: integraalrekening*. Gebroeders Van Cleef, 's Gravenhage (cited on p. 64).

[396] G.L. Lopes [López Lagomasino]. (1979) *Conditions for convergence of multipoint Padé approximants for functions of Stieltjes type*. Math. USSR-Sb. **35**, pp. 363–376 (cited on p. 110).

[397] G.L. Lopes [López Lagomasino]. (1987) *On the asymptotics of the ratio of orthogonal polynomials and convergence of multipoint Padé approximants*. Math. USSR-Sb. **56**, pp. 207–219 (cited on p. 83).

[398] G. López Lagomasino and J. Illán González. (1982) *Fórmulas de cuadratura para intervalos no acotados*. Cienc. Mat. (Havana) **3(3)**, pp. 29–47 (cited on pp. 90, 110).

[399] G. López Lagomasino and J. Illán González. (1982) *Quadrature formulas for unbounded intervals*. Rev. Cienc. Mat. **3**. (In Spanish), pp. 29–47 (cited on p. 88).

[400] G. López Lagomasino and J. Illán González. (1987) *Sobre los métodos interpolatorios de integración numéroca y su conexión con la aproximación racional*. Cienc. Mat. (Havana) **8(2)**, pp. 31–44 (cited on pp. 90, 110).

[401] G.L. López [López Lagomasino]. (1989) *Asymptotics of polynomials orthogonal with respect to varying measures*. Constr. Approx. **5**, pp. 199–219 (cited on p. 83).

[402] G.L. López [López Lagomasino]. (1988) *Szegő's theorem for orthogonal polynomials with respect to varying measures*. In: *Orthogonal polynomials and their applications*. Ed. by M. Alfaro et al. Vol. 1329. of Lecture Notes in Math. Springer, pp. 255–260 (cited on p. 83).

[403] T. Mach, M. Van Barel, and R. Vandebril. (2014) *Inverse eigenvalue problems linked to rational Arnoldi, and rational, (non)symmetric Lanczos*. J. Comput. Appl. Math. To appear (cited on p. 75).

[404] F. Malmquist. (1926) *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)*. Gjellerups, Copenhagen, pp. 253–259 (cited on p. 57).

[405] F. Marcellan. (2006) *Polinomios ortogonales no estandar. Aplicaciones en analisis numerico y teoria de aproximacion*. Rev. Acad. Colombiana Cienc. Exact., Fis. y Natur. **30(117)**. (In Spanish), pp. 563–579 (cited on p. 61).

[406] A. Martinez-Finkelshtein, K.T.-R. McLaughlin, and E.B. Saff. (2006) *Asymptotics of orthogonal polynomials with respect to an analytic weight with algebraic singularities on the circle*. ArXiv Mathematics e-prints, p. 36. eprint: `arxiv:math/0605715` (cited on p. 111).

[407] A. Martinez-Finkelshtein, K.T.-R. McLaughlin, and E.B. Saff. (2006) *Szegő orthogonal polynomials with respect to an analytic weight: canonical representation and strong asymptotics*. Constr. Approx. **24(3)**, pp. 319–363 (cited on p. 111).

[408] P. Masani. (1960) *The prediction theory of multivariate stochastic processes, III.* Acta Math. **104**, pp. 141–162 (cited on p. 95).

[409] K.T.-R. McLaughlin and P. Miller. (2006) *The $\bar{\partial}$ steepest descent method and the asymptotic behavior of polynomials orthogonal on the unit circle with fixed and exponentially varying nonanalytic weights*. Internat. Math. Res. Papers, id:48673, p. 78 (cited on p. 111).

[410] K.T.-R. McLaughlin and P. Miller. (2008) *The $\bar{\partial}$ steepest descent method for orthogonal polynomials on the real line with varying weights*. Internat. Math. Res. Notices, id:rnn075, p. 66 (cited on p. 111).

[411] K.T.-R. McLaughlin, A.H. Vartanian, and X. Zhou. (2008) *Rational functions with a general distribution of poles on the real line orthogonal with respect to varying exponential weights: I*. Mathematical Physiscs, Analysis and Geometry **11(3-4)**, pp. 187–364 (cited on p. 111).

[412] G.V. Milovanović and A.S. Cvetković. (2007) *Gauss-Hermite interval quadrature rule*. Comput. Math. Appl. **54(4)**, pp. 544–555 (cited on p. 110).

[413] G.V. Milovanović and A.S. Cvetković. (2005) *Gauss-Laguerre interval quadrature rule*. J. Comput. Appl. Math. **182(2)**, pp. 433–446 (cited on p. 110).

[414] G.V. Milovanović and A.S. Cvetković. (2006) *Gauss-Radau and Gauss-Lobatto interval quadrature for Jacobi weight function*. Numer. Math. **102(3)**, pp. 523–542 (cited on p. 110).

[415] G.V. Milovanović and A.S. Cvetković. (2004) *Uniqueness and computation of Gaussian interval quadrature for Jacobi weight function*. Numer. Math. **99(1)**, pp. 141–162 (cited on p. 110).

[416] G.V. Milovanović and M.M. Spalević. (2006) *Quadrature rules with multiple nodes for evaluating integrals with strong singularities*. J. Comput. Appl. Math. **189**, pp. 689–702 (cited on p. 71).

[417] G.V. Milovanović, N.M. Spalević, and A.S. Cvetković. (2004) *Calculation of Gaussian type quadratures with multiple nodes*. Math. Comput. Modelling **39(2-3)**, pp. 325–347 (cited on p. 71).

[418] G.V. Milovanović and M.P. Stanić. (2012) *Multiple orthogonality and applications in numerical integration*. In: *Nonlinear analysis. Stability, approximation, and inequalities*. Ed. by P. Pardalos. Springer verlag, pp. 431–455 (cited on p. 111).

[419] G. Min. (1998) *Hermite-Fejér interpolation for rational systems*. Constr. Approx. **14**, pp. 517–529 (cited on p. 57).

[420] G. Min. (1998) *Lagrange interpolation and quadrature formula in rational systems*. J. Approx. Theory **95**, pp. 123–145 (cited on p. 57).

[421] G. Min. (1998) *Lobatto type quadrature formulas in rational spaces*. J. Comput. Appl. Math. **94**, pp. 1–12 (cited on p. 69).

[422] G. Min. (1999) *On the denseness of rational systems*. J. Approx. Theory **98**, pp. 197–202 (cited on p. 57).

[423] G. Monegato. (2001) *An overview of the computational aspects of Kronrod quadrature rules*. Numer. Algorithms **26(2)**, pp. 173–196 (cited on p. 80).

[424] I. Moret. (2009) *Rational Lanczos approximations to the matrix square root and related functions*. Numer. Linear Algebra Appl. **16(6)**, pp. 431–445 (cited on p. 102).

[425] I. Moret and P. Novati. (2005) *Interpolating functions of matrices on zeros of quasi-kernel polynomials*. Numer. Linear Algebra Appl. **12(4)**, pp. 337–253 (cited on p. 102).

[426] K. Müller and A. Bultheel. (1997) *Translation of the Russian paper "Orthogonal systems of rational functions on the unit circle" by M.M. Džrbašian*. Technical Report TW253. Department of Computer Science, K.U. Leuven (cited on p. 57).

[427] O. Nåstad. (1988) *Laurent continued fractions corresponding to pairs of power series*. J. Approx. Theory **55(3)**, pp. 119–139 (cited on p. 110).

[428] Z. Nehari. (1957) *On bounded bilinear forms*. Ann. of Math. **65**, pp. 153–162 (cited on p. 97).

[429] P. Nevai. (1986) *Géza Freud, orthogonal polynomials and Christoffel functions. A case study*. J. Approx. Theory **48**, pp. 3–167 (cited on p. 12).

[430] R. Nevanlinna. (1922) *Asymptotische Entwickelungen beschränkter Funktionen und das Stieltjessche Momentenproblem*. Ann. Acad. Sci. Fenn. Ser. A. **18(5)**, 53pp. (Cited on pp. 26, 93).

[431] R. Nevanlinna. (1922) *Kriterien für die Randwerte beschränkter Funktionen*. Math. Z. **13**, pp. 1–9 (cited on pp. 26, 93).

[432] R. Nevanlinna. (1929) *Über beschränkte analytische Funktionen*. Ann. Acad. Sci. Fenn. Ser. A. **32(7)**, 75pp. (Cited on pp. 26, 93).

[433] R. Nevanlinna. (1919) *Über beschränkte Funktionen die in gegebenen Punkten vorgeschriebene Werte annehmen*. Ann. Acad. Sci. Fenn. Ser. A. **13(1)**, 71pp. (Cited on pp. 26, 93).

[434] B. Ninness and F. Gustafsson. (1997) *A unified construction of orthogonal bases for system identification*. IEEE Trans. Automat. Control **42(2)**, pp. 515–522 (cited on pp. 57, 107, 108).

[435] B. Ninness and F. Gustafsson. (1994) *A unifying construction of orthonormal bases for system identification*. In: *Proceedings of the IEEE Conference on Decision and Control Orlando Florida December 1994*, pp. 3388–3393 (cited on pp. 57, 107, 108).

[436] B. Ninness, H. Hjalmarsson, and F. Gustafsson. (1999) *Generalised Fourier and Toeplitz results for rational orthonormal bases*. SIAM J. Control Optim. **37**, pp. 429–460 (cited on pp. 57, 107, 108).

[437] B. Ninness, H. Hjalmarsson, and F. Gustafsson. (1999) *The fundamental role of general orthogonal bases in system identification*. IEEE Trans. Automat. Control **44(7)**, pp. 1384–1407 (cited on pp. 57, 107, 108).

[438] O. Njåstad. (1991) *A modified Schur algorithm and an extended Hamburger moment problem*. Trans. Amer. Math. Soc. **327(1)**, pp. 283–311 (cited on pp. 20, 26).

[439] O. Njåstad. (1985) *An extended Hamburger moment problem*. Proc. Edinburgh Math. Soc. **28**, pp. 167–183 (cited on pp. 17, 26).

[440] O. Njåstad. (2009) *Analytic functions associated with strong Hamburger moment problems*. Proc. Edinburgh Math. Soc. **52**, pp. 181–187 (cited on p. 20).

[441] O. Njåstad. (1995) *Classical and strong moment problems*. Comm. Anal. Th. Continued Fractions **4**, pp. 4–38 (cited on p. 19).

[442] O. Njåstad. (1993) *Convergence properties related to p-point Padé approximants of Stieltjes functions*. J. Approx. Theory **73**, pp. 149–161 (cited on p. 26).

[443] O. Njåstad. (1995) *Extremal solutions of the strong Stieltjes moment problem*. J. Comput. Appl. Math. **65**, pp. 309–318 (cited on pp. 17, 19).

[444] O. Njåstad. (1994) *Solutions of the Stieltjes moment problem*. Comm. Anal. Th. Continued Fractions **3**, pp. 32–51 (cited on p. 17).

[445] O. Njåstad. (1996) *Solutions of the strong Hamburger moment problem*. J. Math. Anal. Appl. **197**, pp. 227–248 (cited on pp. 17, 19, 20).

[446] O. Njåstad. (1995) *Solutions of the strong Stieltjes moment problem*. Methods Appl. Anal. **2**, pp. 320–347 (cited on pp. 17, 19).

[447] O. Njåstad. (1999) *Strong Stieltjes moment problems*. In: *Continued fractions: from number theory to constructive approximation*. Ed. by B. Berndt and F. Gesztesy. Vol. 236. of Contemp. Math. Amer. Math. Soc., pp. 257–275 (cited on p. 19).

[448] O. Njåstad. (1987) *Unique solvability of an extended Hamburger moment problem*. J. Math. Anal. Appl. **124**, pp. 502–519 (cited on pp. 17, 20, 26).

[449] O. Njåstad. (1988) *Unique solvability of an extended Stieltjes moment problem*. Proc. Amer. Math. Soc. **102(1)**, pp. 78–82 (cited on pp. 19, 26).

[450] O. Njåstad and J.C. Santos León. (2007) *Domain of validity of Szegő quadrature formulas*. J. Comput. Appl. Math. **202(2)**, pp. 440–449 (cited on p. 71).

[451] O. Njåstad and W.J. Thron. (1983) *The theory of sequences of orthogonal L-polynomials*. In: *Padé Approximants and Continued Fractions*. Ed. by H. Waadeland and H. Wallin. Det Kongelige Norske Videnskabers Selskab Skrifter (No. 1), pp. 54–91 (cited on p. 44).

[452] O. Njåstad and W.J. Thron. (1986) *Unique solvability of the strong Hamburger moment problem*. J. Austral. Math. Soc. (Series A) **40**, pp. 5–19 (cited on pp. 17, 20).

[453] O. Njåstad and L. Velázquez. (2009) *Wall rational functions and Khrushchev's formula for orthogonal rational functions*. Constr. Approx. **30**, pp. 277–297 (cited on pp. 93, 94).

[454] S.E. Notaris. (1994) *An overview on the results of existence or nonexistence and the error term of Gauss-Kronrod quadrature formulae*. In: *Approximation and Computation: A Festschrift in Honor of W. Gautschi*. Ed. by R.V.M. Zahar. Vol. 119. of Internat. Ser. of Numer. Math. Birkhäuser Verlag, pp. 485–496 (cited on p. 79).

[455] A.A. Nudel'man. (1981) *On a generalization of classical interpolation problems*. Soviet Math. Dokl. **23(1)**, pp. 125–128 (cited on p. 97).

[456] A.A. Nudel'man. (1977) *On a new problem of moment type*. Soviet Math. Dokl. **18**, pp. 507–510 (cited on p. 97).

[457] V. Olshevsky. (1998) *Eigenvector computation for almost unitary Hessenberg matrices and inversion of Szegő-Vandermonde matrices via discrete transmission lines*. Linear Algebra Appl. **285**, pp. 37–67 (cited on p. 74).

[458] V. Olshevsky, ed. (2003) *Fast algorithms for structured matrices: theory and applications*. Vol. 323. of Contemp. Math. AMS (cited on p. 75).

[459] V. Olshevsky, ed. (2001) *Structured matrices in mathematics, computer science, and engineering, vol. 1*. Vol. 280. of Contemp. Math. AMS (cited on p. 75).

[460] V. Olshevsky, ed. (2001) *Structured matrices in mathematics, computer science, and engineering, vol. 2*. Vol. 281. of Contemp. Math. AMS (cited on p. 75).

[461] K. Pan. (1996) *Asymptotic behavior of orthogonal rational functions corresponding to a measure with discrete part off the unit circle*. Rocky Mountain J. Math. **26**, pp. 1529–1540 (cited on p. 57).

[462] K. Pan. (1996) *Convergence of orthogonal rational functions outside of the unit disk*. J. Comput. Appl. Math. **69**, pp. 345–351 (cited on p. 57).

[463] K. Pan. (1995) *Extensions of Szegő's theory of rational functions orthogonal on the unit circle*. J. Comput. Appl. Math. **62**, pp. 321–331 (cited on p. 57).

[464] K. Pan. (1997) *Moment problem for rational orthogonal functions on the unit circle*. Rocky Mountain J. Math. **27**, pp. 257–269 (cited on p. 57).

[465] K. Pan. (1992) *On characerization theorems for measures associated with orthogonal systems of rational functions on the unit circle*. J. Approx. Theory **70**, pp. 265–272 (cited on p. 57).

[466] K. Pan. (1993) *On orthogonal systems of rational functions on the unit circle and polynomials orthogonal with respect to varying measures*. J. Comput. Appl. Math. **47(3)**, pp. 313–322 (cited on p. 57).

[467] K. Pan. (1994) *On the convergence of rational interpolation approximant of Carathéodory functions*. J. Comput. Appl. Math. **54**, pp. 371–376 (cited on p. 57).

[468] K. Pan. (1995) *On the orthogonal rational functions with arbitrary poles and interpolation properties*. J. Comput. Appl. Math. **60**, pp. 347–355 (cited on p. 57).

[469] K. Pan. (1993) *Strong and weak convergence of orthogonal systems of rational functions on the unit circle*. J. Comput. Appl. Math. **46**, pp. 427–436 (cited on p. 57).

[470] K. Pan. (1997) *The zero distribution of orthogonal rational functions on the unit circle*. Canad. J. Math. **49**, pp. 810–819 (cited on p. 57).

[471] B.N. Parlett. (1998) *The symmetric eigenvalue problem*. Vol. 20. of Classics in Applied Mathematics. Originally published Prentice-Hall, 1980. SIAM (cited on p. 23).

[472] F. Peherstorfer. (1996) *A special class of polynomials orthogonal on the unit circle including the associated polynomials*. Constr. Approx. **12**, pp. 161–185 (cited on p. 110).

[473] F. Peherstorfer. (2009) *Gauss-Turán quadrature formulas: asymptotics of weights*. SIAM J. Numer. Anal. **47**, pp. 2638–2659 (cited on p. 71).

[474] F. Peherstorfer. (1993) *On positive quadrature formulas*. In: *International Series on Numerical Mathematics*. Vol. 112. Birkhäuser Verlag, pp. 297–313 (cited on p. 110).

[475] F. Peherstorfer. (2008) *Positive quadrature formulas III: asympotics of weights*. Math. Comp. **77**, pp. 2241–2259 (cited on p. 110).

[476] F. Peherstorfer and R. Steinbauer. (1994) *Perturbation of orthogonal polynomials on the unit circle – a survey*. In: *Orthogonal polynomials on the unit circle: theory and applications*. Ed. by M. Alfaro, A. García, C. Jagels, and F. Marcellan. Universidad Carlos III, Madrid, pp. 97–119 (cited on p. 110).

[477] G. Pick. (1920) *Über beschränkte Funktionen mit vorgeschriebenen Wertzuordnungen*. Ann. Acad. Sci. Fenn. Ser. A **15(3)**, 17pp. (Cited on pp. 26, 93).

[478] G. Pick. (1918) *Über die Beschränkungen analytischen Funktionen durch vorgegebene Funktionswerte*. Math. Ann. **78**, pp. 270–275 (cited on pp. 26, 93).

[479] G. Pick. (1916) *Über die Beschränkungen analytischen Funktionen welche durch vorgegebene Funktionswerte bewirkt werden*. Math. Ann. **77**, pp. 7–23 (cited on pp. 26, 93).

[480] M.S. Piekarski and M. Uruski. (1996) *A state space approach to the tangential interpolation problem with bounded real matrices*. Internat. J. Circuit Th. Appl. **23(2)**, pp. 83–90 (cited on p. 105).

[481] R. Pintelon and I. Kollár. (2005) *On the frequency scaling in continuous-time modeling*. IEEE Trans. Instrum. and Meas. **53(5)**, pp. 318–321 (cited on pp. 107, 108).

[482] R. Pintelon, Y. Rolain, A. Bultheel, and M. Van Barel. (2004) *Frequency domain identification of multivariabe systems using vector orthogonal polynomials*. In: *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems*. Ed. by B. De Moor, B. Motmans, J. Willems, P. Van Dooren, and V. Blondel. Cdrom. K.U.Leuven, Dept. ESAT (cited on p. 75).

[483] R. Pintelon, Y. Rolain, A. Bultheel, and M. Van Barel. (2002) *Numerically robust frequency domain identification of multivariable systems*. In: *Proceedings of the 2002 International Conference on Modal Analysis, Noise and Vibration Engineering, Leuven 2002, September 16-18*. Ed. by P. Sas and B. Van Hal, pp. 1316–1321 (cited on p. 75).

[484] G. Polya. (1933) *Über die Convergenz von Quadraturverfahren*. Math. Z. **37**, pp. 264–286 (cited on p. 7).

[485] V.P. Potapov. *The multiplicative structure of J-contractive matrix functions*. In: vol. 15. of Amer. Math. Soc. Transl. Ser. 2. Amer. Math. Soc., Providence, Rhode Island, pp. 131–243 (cited on p. 97).

[486] M.S. Pranić and L. Reichel. (2012) *Recurrence relations for orthogonal rational functions*. Numer. Math. **123(4)**, pp. 629–642 (cited on p. 56).

[487] R. Radau. (1880) *Étude sur les formules d'approximation qui servent à calculer la valeur numérique d'une intégrale définie*. J. Math. Pures Appl. **6**, pp. 283–336 (cited on p. 64).

[488] L. Reichel and M.S. Pranić. (2013) *Rational Gauss quadrature*. SIAM J. Numer. Anal. (cited on p. 56).

[489] W. Rudin. (1987) *Real and complex analysis*. 3rd. McGraw-Hill, New York (cited on p. 113).

[490] D. Sarason. (1967) *Generalized interpolation in H^∞*. Trans. Amer. Math. Soc. **127**, pp. 179–203 (cited on p. 97).

[491] I. Schur. (1912) *Über ein Satz von C. Carathéodory*. S.-B. Preuss. Akad. Wiss. (Berlin), pp. 4–15 (cited on p. 21).

[492] I. Schur. (1917) *Über Potenzreihen die im Innern des Einheitskreises Beschränkt sind I*. J. Reine Angew. Math. **147**. See also [302, p.31-59], pp. 205–232 (cited on p. 21).

[493] I. Schur. (1918) *Über Potenzreihen die im Innern des Einheitskreises Beschränkt sind II*. J. Reine Angew. Math. **148**. See also [302, p.36-88], pp. 122–145 (cited on p. 21).

[494] L.-C. Shen. (2012) *On polynomial approximation of square integrable functions on a subarc of the unit circle*. Comput. Methods Funct. Th. **144(2)**, pp. 152–163 (cited on p. 17).

[495] Y.G. Shi and G. Xu. (2007) *Construction of σ-orthogonal polynomials and Gaussian quadrature formulas*. Adv. Comput. Math. **27(1)**, pp. 57–88 (cited on p. 71).

[496] C.T. Shieh. (2004) *Some inverse problems on Jacobi matrices*. Inverse Problems **20(2)**, pp. 589–600 (cited on p. 74).

[497] J.A. Shohat and J.D. Tamarkin. (1943) *The problem of moments*. Vol. 1. of Math. Surveys. Amer. Math. Soc., Providence, R.I. (cited on pp. 12, 16, 17).

[498] B. Simon. (2007) *CMV matrices: five years after*. J. Comput. Appl. Math. **208**, pp. 120–154 (cited on p. 14).

[499] B. Simon. (2005) *OPUC on one foot*. Bull. Amer. Math. Soc. **42**, pp. 431–460 (cited on p. 14).

[500] B. Simon. (2005) *Orthogonal polynomials on the unit circle*. Vol. 54. of Colloquium Publications. AMS (cited on pp. 12, 14, 25).

[501] H. Stahl and V. Totik. (1992) *General orthogonal polynomials*. Encyclopedia of Mathematics and its Applications. Cambridge University Press (cited on pp. 12, 86).

[502] W. Steklov. (1916) *On approximate evaluation of definite integrals by mens of formulas of mechanical quadrature. I Convergence of formulas of mechanical quadrature*. Izv. Imper. Akad. Nauk **10(6)**. (In Russian), pp. 169–186 (cited on p. 7).

[503] T.J. Stieltjes. (1894) *Recherches sur les fractions continues*. Ann. Fac. Sci. Toulouse **8**. Oeuvres Complèts, Collected Papers, Springer Verlag, 1993, Vol. 2, 402-566; English transl.: ibid, 609–745, J.1–122, 9:A.1–47 (cited on p. 17).

[504] A.H. Stroud and D. Secrest. (1966) *Gaussian quadrature formulas*. Prentice-Hall series in automatic computation. Prentice-Hall, Englewood Cliffs, N.J. (cited on p. 23).

[505] A.H. Stroud and D.D. Stancu. (1965) *Quadrature formulas with multiple nodes*. SIAM J. Numer. Anal. **2(1)**, pp. 129–143 (cited on p. 71).

[506] P. Suau, E.R. Hancock, and F. Escolano. (2013) *Graph characteristics from the schrödinger operator*. In: *Graph-based representitions in pattern recognition*. Ed. by W.G. Kropatsch, R.M. Artner, Y. Haxhimusa, and X. Jiang. Vol. 7877. of Lecture Notes in Computer Science, pp. 172–181 (cited on p. 99).

[507] Z. Szabó, J. Bokor, and F. Schipp. (1999) *Identification of rational approximate models in H_∞ using generalized orthogonal basis*. IEEE Trans. Automat. Control **44(1)**, pp. 153–158 (cited on pp. 107, 108).

[508] G. Szegő. (1963) *On bi-orthogonal systems of trigonometric polynomials*. Magyar Tud. Alcad. Kutato Int. Közl. **8**, pp. 255–273 (cited on p. 55).

[509] G. Szegő. (1975) *Orthogonal polynomials*. 4th. Vol. 33. of Amer. Math. Soc. Colloq. Publ. First edition 1939. Amer. Math. Soc., Providence, Rhode Island (cited on pp. 7, 10, 12, 51, 58).

[510] S. Takenaka. (1925) *On the orthogonal functions and a new formula of interpolation*. Japanese J. Math. **2**, pp. 129–145 (cited on p. 57).

[511] L. Tchakaloff. (1934) *Sur la structure des ensembles linéaires définis par une certaine propriété minimale*. Acta Math. **63(1)**. (In French), pp. 77–97 (cited on p. 7).

[512] S. Tokarzewski. (1996) *Two-point Padé approximants for the expansion of Stieltjes functions in the real domain*. J. Comput. Appl. Math. **67**, pp. 59–72 (cited on p. 19).

[513] S. Tokarzewski. *Two-point Padé approximants to Laurent expansions of Stieltjes functions*. Preprint, Inst. Fund. Technological Research, Warsaw. 1994 (cited on p. 19).

[514] S. Tokarzewski, J. Bławzdziewicz, and I. Adrianov. (1994) *Two-point Padé approximants for formal Stieltjes series*. Numer. Algorithms **8**, pp. 313–328 (cited on p. 19).

[515] P. Turán. (1950) *On the theory of mechanical quadrature*. Acta Sci. Math. (Szeged) **12**, pp. 30–37 (cited on p. 71).

[516] W. Van Assche. (1993) *The impact of Stieltjes' work on continued fractions and orthogonal polynomials*. In: *T.J. Stieltjes, Collected Papers*. Stringer Verlag, pp. 5–37 (cited on pp. 16, 110).

[517] W. Van Assche and E. Coussement. (2001) *Some classical multiple orthogonal polynomials*. Numerical analysis 2000, Vol. V, Quadrature and orthogonal polynomials, J. Comput. Appl. Math. **127(1-2)**, pp. 317–347 (cited on p. 111).

[518] W. Van Assche and I. Vanherwegen. (1993) *Quadrature formulas based on rational interpolation*. Math. Comp. **16**, pp. 765–783 (cited on p. 51).

[519] M. Van Barel and A. Bultheel. (1990) *A new approach to the rational interpolation problem*. J. Comput. Appl. Math. **32(1-2)**, pp. 281–289 (cited on p. 74).

[520] M. Van Barel and A. Bultheel. (1990) *A new approach to the rational interpolation problem: the vector case*. J. Comput. Appl. Math. **33(3)**, pp. 331–346 (cited on p. 74).

[521] M. Van Barel and A. Bultheel. (1992) *A new formal approach to the rational interpolation problem*. Numer. Math. **62**, pp. 87–122 (cited on p. 74).

[522] M. Van Barel and A. Bultheel. (1992) *A parallel algorithm for discrete least squares rational approximation*. Numer. Math. **63**, pp. 99–121 (cited on p. 74).

[523] M. Van Barel and A. Bultheel. (1994) *Discrete linearized least squares approximation on the unit circle*. J. Comput. Appl. Math. **50**, pp. 545–563 (cited on p. 75).

[524] M. Van Barel and A. Bultheel. (1995) *Orthogonal polynomial vectors and least squares approximation for a discrete inner product*. Electron. Trans. Numer. Anal. **3**, pp. 1–23 (cited on p. 74).

[525] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. (2005) *Orthogonal rational functions and structured matrices*. SIAM J. Matrix Anal. Appl. **26(3)**, pp. 810–829 (cited on p. 75).

[526] M. Van Barel and M. Humet. *Orthogonal functions and inverse eigenvalue problems*. 4èmes Journées Approximation, International conference on constructive complex approximation, Lille, France, 28-29 June 2012. 2012 (cited on p. 75).

[527] E. Van Camp. (2005) *Diagonal-plus-semiseparable matrices and their use in numerical linear algebra*. PhD thesis. KU Leuven (cited on p. 75).

[528] J. Van Deun. (2010) *Computing near-best fixed pole rational interpolants*. J. Comput. Appl. Math. **235(4)**, pp. 1077–1084 (cited on pp. 62, 63).

[529] J. Van Deun. (2007) *Eigenvalue problems to compute almost optimal points for rational interpolation with prescribed poles*. Numer. Algorithms **45(1-4)**, pp. 89–99 (cited on p. 63).

[530] J. Van Deun. (2007) *Electrostatics and ghost poles in near best fixed poles rational interpolation*. Electron. Trans. Numer. Anal. **26**, pp. 439–452 (cited on p. 56).

[531] J. Van Deun and A. Bultheel. (2005) *A weak-star convergence result for orthogonal rational functions*. J. Comput. Appl. Math. **178(1-2)**, pp. 453–464 (cited on p. 90).

[532] J. Van Deun and A. Bultheel. (2003) *Orthogonal rational functions and quadrature on an interval*. J. Comput. Appl. Math. **153(1-2)**, pp. 487–495 (cited on p. 68).

[533] J. Van Deun and A. Bultheel. (2003) *Ratio asymptotics for orthogonal rational functions on an interval*. J. Approx. Theory **123(2)**, pp. 162–172 (cited on p. 89).

[534] J. Van Deun and A. Bultheel. (2005) *The computation of orthogonal rational functions on an interval*. J. Comput. Appl. Math. **179(1-2)**, pp. 355–373 (cited on p. 51).

[535] J. Van Deun, A. Bultheel, and P. González-Vera. (2006) *On computing rational Gauss-Chebyshev quadrature formulas*. Math. Comp. **75(253)**, pp. 307–326 (cited on p. 59).

[536] J. Van Deun, K. Deckers, A. Bultheel, and J.A.C. Weideman. (2008) *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 (cited on pp. 62, 63).

[537] P. Van Dooren, K.A. Gallivan, and P.-A. Absil. (2008) *H^2-optimal model reduction of MIMO systems*. Appl. Math. Lett. **21**, pp. 1276–1273 (cited on p. 108).

[538] P. Van Dooren, K.A. Gallivan, and P.-A. Absil. (2010) *H^2-optimal model reduction of with higher order poles*. SIAM J. Matrix Anal. Appl. **31(5)**, pp. 2738–2753 (cited on p. 108).

[539] P. Van gucht and A. Bultheel. (2000) *A relation between orthogonal rational functions on the unit circle and the interval* $[-1, 1]$. Comm. Anal. Th. Continued Fractions **8**, pp. 170–182 (cited on pp. 51, 59).

[540] P. Van gucht and A. Bultheel. (2003) *State space representation for arbitrary orthogonal rational functions*. Systems Control Lett. **49(2)**, pp. 91–98 (cited on p. 108).

[541] P. Van gucht and A. Bultheel. (2003) *Using orthogonal rational functions for system identification: numerical aspects*. IEEE Trans. Automat. Control **48(4)**, pp. 705–709 (cited on p. 107).

[542] R. Vandebril. (2004) *Semiseparable matrices and the symmetric eigenvalue problem*. PhD thesis. Dept. Computer Science, K.U.Leuven (cited on p. 75).

[543] R. Vandebril, M. Van Barel, and N. Mastronardi. (2005) *A note on the representation and definition of semiseparable matrices*. Numer. Linear Algebra Appl. **12(8)**, pp. 839–858 (cited on p. 75).

[544] R. Vandebril, M. Van Barel, and N. Mastronardi. (2007) *Matrix computations and semiseparable matrices, volume 1. linear systems*. The Johns Hopkins University Press, Baltimore (cited on p. 75).

[545] R. Vandebril, M. Van Barel, and N. Mastronardi. (2008) *Matrix computations and semiseparable matrices, volume 2: eigenvalue and singular value methods*. The Johns Hopkins University Press, Baltimore (cited on pp. 25, 75).

[546] L. Velázquez. (2008) *Spectral methods for orthogonal rational functions*. J. Funct. Anal. **254(4)**, pp. 954–986 (cited on p. 53).

[547] S. Verblunsky. (1935) *On positive harmonic functions: A contribution to the algebra of Fourier series*. Proc. London Math. Soc. **38(2)**, pp. 125–157 (cited on p. 14).

[548] C.D. Villemagne and R.E. Skelton. (1987) *Model reduction using a projection formulation*. Internat. J. Control **46**, pp. 2141–2169 (cited on p. 105).

[549] Gautschi W and M.M. Spalević. (1997) *S-orthogonality and construction of Gauss-Turán-type quadrature formulae*. J. Comput. Appl. Math. **86(1)**, pp. 205–216 (cited on p. 71).

[550] H. Waadeland. (1992) *A Szegő quadrature formula for the Poisson integral*. In: *Computational and Applied Mathematics I*. Ed. by C. Brezinski and U. Kulish. Elsevier, pp. 479–486 (cited on pp. 57, 61).

[551] B. Wahlberg. (2003) *Orthogonal rational functions: a transformation analysis*. SIAM Rev. **45(4)**, pp. 689–705 (cited on pp. 107, 108).

[552] B. Wahlberg. (1994) *System identification using Kautz models*. IEEE Trans. Automat. Control **AC-39(6)**, pp. 1276–1282 (cited on pp. 107, 108).

[553] B. Wahlberg. (1991) *System identification using Laguerre models*. IEEE Trans. Automat. Control **AC-36(5)**, pp. 551–562 (cited on pp. 107, 108).

[554] J.L. Walsh. (1960) *Interpolation and approximation*. 3rd. Vol. 20. of Amer. Math. Soc. Colloq. Publ. First edition 1935. Amer. Math. Soc., Providence, Rhode Island (cited on pp. 57, 82).

[555] J.L. Walsh. (1932) *Interpolation and functions analytic interior to the unit circle*. Trans. Amer. Math. Soc. **34**, pp. 523–556 (cited on pp. 26, 113).

[556] T.L. Wang and W.B. Gragg. (2002) *Convergence of the shifted QR algorithm for unitary Hessenberg matrices*. Math. Comp. **71**, pp. 1473–1496 (cited on p. 25).

[557] T.L. Wang and W.B. Gragg. (2003) *Convergence of the unitary QR algorithm with a unimodular Wilkinson shift*. Math. Comp. **72**, pp. 375–385 (cited on p. 25).

[558] D.S. Watkins. (1993) *Some perspectives on the eigenvalues problem*. SIAM Rev. **35**, pp. 430–471 (cited on pp. 25, 61).

[559] N. Wiener. (1949) *Extrapolation, interpolation and smoothing of time series*. Wiley, New York (cited on p. 106).

[560] N. Wiener and P. Masani. (1957) *The prediction theory of multivariate stochastic processes, I. The regularity condition*. Acta Math. **98**, pp. 111–150 (cited on pp. 95, 106).

[561] N. Wiener and P. Masani. (1958) *The prediction theory of multivariate stochastic processes, II. The linear predictor*. Acta Math. **99**, pp. 93–139 (cited on pp. 95, 106).

[562] H. Wilf. (1962) *Mathematics for the physical sciences*. Wiley (cited on p. 23).

[563] D.A. Wilson. (1970) *Optimum solution of model-reduction problem*. Proc. IEEE **117**, pp. 1161–1165 (cited on p. 108).

[564] Y.H. Xu and E.X. Jiang. (2008) *A new algorithm for an inverse eigenvalue problem on Jacobi matrices*. J. Shanghai Univ. **12(4)**, pp. 289–293 (cited on p. 74).

[565] S.M. Zagorodnyuk. (2013) *Nevanlinna formula for the truncated matrix trigonometric moment problem*. Ukrainian Math. J. **64(8)**, pp. 1199–1214 (cited on p. 97).

[566] S.M. Zagorodnyuk. (2011) *Truncated matrix trigonometric problem of moments: operator approach*. Ukrainian Math. J. **63(6)**, pp. 914–926 (cited on p. 97).