The Computation of Orthogonal Rational Functions on an Interval

J. Van Deun and A. Bultheel [1]

Department of Computer Science, K.U.Leuven, Belgium

Abstract

A method is given to compute the recurrence coefficients for high degree orthogonal rational functions on an interval of the real line and some asymptotic results are used to derive computable error bounds. We also prove a density result for these functions. Several examples are given as illustrations.

Key words: orthogonal rational functions, quadrature

AMS Subject classification: 42C05, 65D05

1 Introduction

In [9] we presented an algorithm to compute the recurrence coefficients for orthogonal rational functions on (a subset of) the real line, based on certain interpolation properties of the Riesz-Herglotz-Nevanlinna transform of the orthogonality measure μ. An error analysis showed that this algorithm is very unstable and is only useful to produce the recurrence coefficients for orthogonal rational functions of low degree ($n < 10$), unless one is willing to use a multiprecision package. Furthermore, this algorithm only works if all poles are different from each other.

In this paper it is our aim to investigate the accurate computation of the recurrence coefficients for arbitrarily high degree working in ordinary (double)

[1] Email: {Joris.Vandeun|Adhemar.Bultheel}@cs.kuleuven.ac.be
This work is partially supported by the Fund for Scientific Research (FWO), project "CORFU: Constructive study of orthogonal functions", grant #G.0184.02 and the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the authors.

precision. We will derive expressions for the coefficients in terms of inner products. For a discrete measure these inner products reduce to finite sums whose computation is straightforward. If the measure is not discrete, the inner products can be approximated using Gaussian quadrature rules. We will limit our attention to the latter case and assume that the measure is supported on an interval (which can always be transformed to $[-1,1]$). Error bounds are derived to estimate the accuracy of the computed coefficients and several numerical examples are given as illustration.

2 Preliminaries

The complex plane is denoted by $\mathbb{C}$, the Riemann sphere by $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, the real line by $\mathbb{R}$ and the extended real line by $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. For the unit circle and its interior we introduce the following notation:

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

We will also use $I = [-1,1]$, $\mathbb{R}^I = \hat{\mathbb{R}} \setminus I$ and $\mathbb{C}^I = \hat{\mathbb{C}} \setminus I$. Given a positive bounded Borel measure μ on I whose support $\mathrm{supp}(\mu) \subset I$ is an infinite set, the inner product is defined as

$$\langle f\,,g\rangle = \int_{-1}^{1} f(x)\overline{g(x)}d\mu(x). \tag{1}$$

We will assume that the measure is normalized, i.e. $\mu(I) = 1$.

Next we will introduce the spaces of rational functions with real poles. Let a sequence $A = \{\alpha_1, \alpha_2, \ldots\} \subset \mathbb{R}^I$ be given. Define factors

$$Z_n(x) = \frac{x}{1 - x/\alpha_n}, \quad n = 1, 2, \ldots$$

and basis functions

$$b_0 = 1, \quad b_n(x) = b_{n-1}(x)Z_n(x), \quad n = 1, 2, \ldots$$

Then the space of rational functions with poles in A is defined as

$$\mathcal{L}_n = \mathrm{span}\{b_0, \ldots, b_n\}.$$

Let $\mathcal{P}_n$ denote the space of polynomials of degree at most n and define

$$\pi_n(x) = \prod_{k=1}^{n}(1 - x/\alpha_k),$$

then we may write equivalently

$$\mathcal{L}_n = \{p_n/\pi_n, p_n \in \mathcal{P}_n\}.$$

Note that $\mathcal{L}_n = \mathcal{P}_n$ if all poles are at infinity. Orthonormalizing the basis $\{b_0, \ldots, b_n\}$ with respect to the orthogonality measure μ we obtain orthogonal rational functions $\{\phi_0, \ldots, \phi_n\}$. Regarding these orthonormal functions we have the following important theorem from [3], completely analogous to the one for orthogonal polynomials. The regularity conditions mentioned in [3] are automatically satisfied because of the assumption on the location of the poles.

Theorem 1 *Put by convention $\alpha_{-1} = \alpha_0 = \infty$. Then for $n = 1, 2, \ldots$ the orthonormal rational functions ϕ_n satisfy the following three term recurrence relation:*

$$\phi_n(x) = \left(E_n Z_n(x) + F_n \frac{Z_n(x)}{Z_{n-1}(x)}\right) \phi_{n-1}(x) - \frac{E_n}{E_{n-1}} \frac{Z_n(x)}{Z_{n-2}(x)} \phi_{n-2}(x). \qquad (2)$$

The initial conditions are $\phi_{-1}(x) \equiv 0$, $\phi_0(x) \equiv 1$ and the coefficients E_n are nonzero.

Note that the coefficient E_0 is never used and can be arbitrarily chosen. We take it equal to $E_0 = 1$. If we take the coefficient E_n to be positive, then the functions ϕ_n will be uniquely determined. In this case we have the following lemma which can be deduced from [3].

Lemma 2 *The orthonormal functions ϕ_n normalized with $E_n > 0$ have real coefficients with respect to the basis $\{b_k\}$.*

It follows in particular that $\phi_n(x)$ is real for real x and for any inner product $\langle f\,, \phi_n \rangle$ we may omit the complex conjugate bar in (1).

The accurate computation of the recurrence coefficients is our main concern in this paper. In the next section we will derive expressions for these coefficients involving inner products.

3 Expressions for E_n and F_n

Taking the inner product on both sides of (2) with $\phi_k, k < n$ yields

$$E_n \langle Z_n \phi_{n-1}\,, \phi_k \rangle + F_n \left\langle \frac{Z_n}{Z_{n-1}} \phi_{n-1}\,, \phi_k \right\rangle - \frac{E_n}{E_{n-1}} \left\langle \frac{Z_n}{Z_{n-2}} \phi_{n-2}\,, \phi_k \right\rangle = 0$$

and solving for F_n we find

$$F_n = -E_n \frac{\langle Z_n \phi_{n-1}\,,\, \phi_k \rangle - \frac{1}{E_{n-1}} \left\langle \frac{Z_n}{Z_{n-2}} \phi_{n-2}\,,\, \phi_k \right\rangle}{\left\langle \frac{Z_n}{Z_{n-1}} \phi_{n-1}\,,\, \phi_k \right\rangle} = -E_n \hat{F}_n. \tag{3}$$

Substituting this back into the recurrence relation we get $\phi_n(x) = E_n \hat{\phi}_n(x)$ where $\hat{\phi}_n(x)$ can be computed as

$$\hat{\phi}_n(x) = \left(Z_n(x) - \hat{F}_n \frac{Z_n(x)}{Z_{n-1}(x)} \right) \phi_{n-1}(x) - \frac{1}{E_{n-1}} \frac{Z_n(x)}{Z_{n-2}(x)} \phi_{n-2}(x)\ . \tag{4}$$

To obtain an expression for E_n we use the fact that the function ϕ_n is normalized

$$E_n = \frac{1}{\|\hat{\phi}_n\|}\ . \tag{5}$$

These formulas are straightforward to implement, once we have a procedure to compute the inner products. One way of doing this is by a partial fraction decomposition of the functions appearing in the inner products. Integrating term by term then gives the desired result. This procedure requires the knowledge of integrals of the form $\int_{-1}^{1} d\mu/(x-\alpha_i)^j$ (the *moments* of the measure μ). For non-trivial measures the exact evaluation of these integrals may be difficult or even impossible and even if we can evaluate them, the fact remains that the computation of the inner products using the moments is a highly ill-conditioned problem [2]. For the case of orthogonal polynomials, this is well known. For orthogonal rational functions the analysis in [6] may be repeated using the above moments in stead of $\int_{-1}^{1} x^i d\mu$, Hermite interpolation rational functions in stead of Hermite interpolation polynomials and the asymptotic behaviour of ϕ_n outside the interval of integration as described in [11].

Another possibility is to approximate the inner products using some kind of quadrature rule. This is the approach we will follow. We will use Gaussian quadrature rules based on orthogonal rational functions and provide algorithms to estimate the accuracy of the computed coefficients (without these error bounds there would be little practical use).

4 Quadrature formulas

The quadrature formulas from this section are very similar to the ones for the polynomial case. In fact if we take all poles at infinity, the formulas reduce to the classic Gauss quadrature formulas. The following theorem can be found in [3], chapter 11.6 and in more detail (for the case of the interval) in [10].

[2] This is basically what causes the algorithm in [9] to fail.

Theorem 3 *Let $\{\xi_k\}_{k=1}^n$ be the zeros of the n-th orthogonal rational function ϕ_n and let $\{\lambda_{nk}\}_{k=1}^n$ be defined by*

$$\lambda_{nk} = \left[\sum_{j=0}^{n-1} (\phi_j(\xi_k))^2\right]^{-1}$$

then the quadrature formula

$$I_\mu(f) = \int_{-1}^{1} f(x)d\mu(x) \approx I_n(f) = \sum_{k=1}^{n} \lambda_{nk} f(\xi_k)$$

is exact for $f \in \mathcal{L}_n \cdot \mathcal{L}_{n-1}$.

Also in [10] it is shown that the nodes are all distinct and inside the interval of integration $[-1, 1]$. It follows from [4] that these nodes can be found as the eigenvalues of a generalized eigenvalue problem and that the weights are defined by the first components of the corresponding eigenvectors. The matrices involved contain the recurrence coefficients and the poles. Essentially this means that, given the poles and the recurrence coefficients, the weights and nodes can be easily and accurately computed using linear algebra techniques.

Next we wish to study the convergence of these quadrature formulas. To this end we need some results about the density of the rational functions. First we give the following lemma, which is a special case of the more general result in [1, p. 250].

Lemma 4 *Let the polynomial*

$$w_n(x) = \prod_{k=1}^{n} \left(1 - \frac{x}{a_k}\right)$$

be given with real zeros a_k such that $|a_k| > 1$ (a_k may equal infinity) and take c_k such that $|c_k| < 1$ and

$$a_k = \frac{1}{2}\left(c_k + \frac{1}{c_k}\right), \quad k = 1, \ldots, n$$

then we have for $m > n$ and $p = 1$ or $p = \infty$

$$\min_{A_k} \left\| \frac{x^m + A_1 x^{m-1} + \ldots + A_m}{w_n(x)} \right\|_p = \frac{1}{2^{m-1}} \prod_{k=1}^{n} (1 + c_k^2).$$

Using this lemma it is not difficult to obtain conditions under which the rational functions are dense in the space $C(I)$ of continuous functions on I or in the $L^p(I)$ spaces.

Theorem 5 *The space $\mathcal{L} = \cup_0^\infty \mathcal{L}_n$ is dense in any $L^p(I)$ space with $p \geq 1$ as well as in the class $C(I)$ of continuous functions on I if and only if $\sum_{k=1}^\infty (1 - |c_k|) = \infty$, where c_k satisfies $\alpha_k = (c_k + 1/c_k)/2$ as in the previous lemma.*

PROOF. The proof is very much like the one of theorem 7.1.2 in [3], but we give it for the sake of completeness.

It suffices to show that divergence of the sum implies the completeness of the system $\{b_k\}_{k=0}^\infty$ in the space $C(I)$ and that it is necessary for completeness in the space $L^1(I)$.

First note that if there are infinitely many poles at infinity, then there is nothing to prove because in that case the system contains all the polynomials and these are known to be complete while the sum certainly diverges.

So assume (without loss of generality) that the first q poles are at infinity ($q < \infty$) and that the sum diverges. Then we have to show that all x^k with $k > q$ can be approximated arbitrarily close (according to the Chebyshev norm) by elements of $\mathcal{L}_n$ for n sufficiently large. This follows from the following observation (using the previous lemma):

$$\begin{aligned}
&\min_{f \in \mathcal{L}_n, A_i} \left\| x^{m+q} + A_1 x^{m+q-1} + \ldots + A_{m-1} x^{q+1} + f(x) \right\|_\infty \\
&= \frac{1}{|\alpha_{q+1}\alpha_{q+2}\ldots\alpha_n|} \min_{p \in \mathcal{P}_{n+m-1}} \left\| \frac{x^{n+m} + p(x)}{\pi_n(x)} \right\|_\infty = \frac{1}{2^{m+q-1}} \prod_{k=q+1}^{n} |c_k|
\end{aligned}$$

Since the sum diverges to infinity, we must have that $\prod_{q+1}^n |c_k| \to 0$. The rest of the proof is then by induction.

To prove that divergence of the sum is necessary for completeness in $L^1(I)$, we proceed as follows. By the previous lemma for $p = 1$ we have that

$$\min_{f \in \mathcal{L}_n} \|x^{q+1} - f(x)\|_1 = \frac{1}{|\alpha_{q+1}\alpha_{q+2}\ldots\alpha_n|} \min_{p \in \mathcal{P}_n} \left\| \frac{x^{n+1} + p(x)}{\pi_n(x)} \right\|_1 = \frac{1}{2^q} \prod_{k=q+1}^{n} |c_k|.$$

Since the system is supposed to be complete in $L^1(I)$, the previous expression should tend to zero as $n \to \infty$. Whence $\sum(1 - |c_k|) = \infty$. $\square$

As a direct consequence of this theorem and the fact that the weights in the quadrature formula are positive, we have the following theorem which we give without proof.

Theorem 6 *Under the same conditions as in theorem 5, it holds that*

$$\lim_{n\to\infty} I_n(f) = I_\mu(f) = \int_{-1}^{1} f(x)d\mu(x)$$

for any function f in $C(I)$.

If $f \notin \mathcal{L}_n \cdot \mathcal{L}_{n-1}$ then we need an expression for the error

$$e_n(f) = I_\mu(f) - I_n(f).$$

This will be derived in the following theorem.

Theorem 7 *Let f be analytic in a closed simply connected region Ω containing the interval $[-1,1]$. Let C be a Jordancurve that lies in Ω and surrounds the quadrature nodes $\{\xi_k\}_{k=1}^n$. Then*

$$e_n(f) = \frac{1}{2\pi\mathbf{i}} \int_C \frac{f(t)}{(1-t/\alpha_n)\phi_n^2(t)} \left\{ \int_{-1}^{1} \frac{\phi_n^2(x)(1-x/\alpha_n)}{t-x} d\mu(x) \right\} dt. \qquad (6)$$

PROOF. Let $R_{2n-1}(f,x)$ denote the (unique) element of $\mathcal{L}_n \cdot \mathcal{L}_{n-1}$ satisfying the interpolation conditions

$$R_{2n-1}(f,\xi_j) = f(\xi_j)$$
$$R'_{2n-1}(f,\xi_j) = f'(\xi_j)$$

for $j = 1,\ldots,n$. Then it can easily be shown that $e_n(f) = I_\mu(f - R_{2n-1}(f,\cdot))$. For a detailed explanation we refer to [2] where the (similar) case of quadrature on the real half line is treated.

Writing $R_{2n-1}(f,x) = p_{2n-1}(f,x)/(\pi_n(x)\pi_{n-1}(x))$ it is not difficult to see that $p_{2n-1}(f,x)$ is the unique Hermite interpolating polynomial for the function $\pi_n(x)\pi_{n-1}(x)f(x)$. From theorem 3.6.1 in [5] the interpolation error is given by

$$\pi_n(x)\pi_{n-1}(x)f(x) - p_{2n-1}(f,x) =$$
$$\frac{1}{2\pi\mathbf{i}} \int_C \frac{(x-\xi_1)^2 \ldots (x-\xi_n)^2}{(t-\xi_1)^2 \ldots (t-\xi_n)^2} \frac{\pi_n(t)\pi_{n-1}(t)f(t)}{t-x} dt$$

where C is as in the statement of this theorem. Dividing by $\pi_n(x)\pi_{n-1}(x)$ this yields

$$f(x) - R_{2n-1}(f,x) = \frac{1}{2\pi\mathbf{i}} \int_C \frac{(1-x/\alpha_n)\phi_n^2(x)}{(1-t/\alpha_n)\phi_n^2(t)} \frac{f(t)}{t-x} dt.$$

Integrating over I with respect to $d\mu(x)$ now proves the theorem. $\square$

The main concern of this paper is to compute the recurrence coefficients for rational functions with poles in A orthogonal with respect to μ, using the formulas derived in section 3. To approximate the inner products in these formulas, we will use a quadrature rule based on a different set of rational functions (poles in $\tilde{A}$), which we assume known. Before we can give an accurate estimate for the error in the quadrature formulas, we need some asymptotic results for orthogonal rational functions as n tends to ∞.

5 Asymptotic behaviour

As in the polynomial case, most convergence results for orthogonal rational functions on the interval $[-1,1]$ were obtained relating them to orthogonal rational functions on the complex unit circle using the Joukowski transformation. We denote this transformation by $x = J(z) = \frac{1}{2}(z + z^{-1})$, mapping the open unit disc $\mathbb{D}$ to $\mathbb{C}^I$ and the unit circle $\mathbb{T}$ to I. The inverse mapping is denoted by $z = J^{-1}(x)$ and is chosen such that $z \in \mathbb{D}$ if $x \in \mathbb{C}^I$. Then associate a sequence $B = \{\beta_1, \beta_2, \ldots\} \subset I$ to the sequence $A = \{\alpha_1, \alpha_2, \ldots\}$ such that $\beta_k = J^{-1}(\alpha_k)$ and define Blaschke products as follows,

$$B_0 = 1, \quad B_n(z) = \frac{z - \beta_n}{1 - \beta_n z} B_{n-1}(z), \quad n = 1, 2, \ldots$$

In the rest of this paper we will assume that the measure μ satisfies Szegő's condition

$$\int_{-1}^{1} \frac{\log \mu'(x)}{\sqrt{1-x^2}} dx > -\infty.$$

Next define the measure ν on $[0, 2\pi)$ as

$$\nu(E) = \mu(\{\cos\theta, \theta \in E \cap [0, \pi)\}) + \mu(\{\cos\theta, \theta \in E \cap [\pi, 2\pi)\})$$

and the associated Szegő function $\sigma(\nu, z)$ as

$$\sigma(\nu, z) = \exp\left\{ \frac{1}{4\pi} \int_0^{2\pi} \frac{e^{\mathbf{i}\theta} + z}{e^{\mathbf{i}\theta} - z} \log \nu'(\theta) d\theta \right\}, \quad |z| < 1.$$

The first theorem deals with strong asymptotics for orthogonal rational functions and the asymptotic behaviour of the recurrence coefficients. It is proved in [11].

Theorem 8 *Assume that the sequence $A = \{\alpha_1, \alpha_2, \ldots\} \subset \mathbb{R}^I$ is bounded away from I and let μ be a positive bounded Borel measure with* $\mathrm{supp}(\mu) \subset I$ *an infinite set, which satisfies the Szegő condition*

$$\int_{-1}^{1} \frac{\log \mu'(x)}{\sqrt{1-x^2}} dx > -\infty.$$

Let ν and $\sigma(\nu, z)$ be as defined above. If $\{\phi_n\}$ are the orthonormal rational functions on I associated with A and μ, then locally uniformly in $\mathbb{C}^I$ we have

$$\lim_{n\to\infty} c_n B_n(z) \frac{1-\beta_n z}{\sqrt{1-\beta_n^2}} \phi_n(x) = \frac{1}{\sqrt{2\pi}\sigma(\nu, z)}$$

where $z = J^{-1}(x)$, $\beta_k = J^{-1}(\alpha_k)$ and $c_n = \pm 1$ according to the normalization $E_n > 0$. Under the same assumptions we have

$$E_n \sim 2\frac{\sqrt{(1-\beta_{n-1}^2)(1-\beta_n^2)}(1-\beta_{n-1}\beta_n)}{(1+\beta_{n-1}^2)(1+\beta_n^2)},$$
$$F_n \sim -\sqrt{\frac{1-\beta_n^2}{1-\beta_{n-1}^2}}\frac{(1-\beta_{n-1}^2)(\beta_n+\beta_{n-2})+2\beta_{n-1}(1-\beta_n\beta_{n-2})}{(1+\beta_n^2)(1-\beta_{n-1}\beta_{n-2})},$$

in the sense that the ratio of the left hand side and the right hand side tends to 1 as n tends to infinity.

We will also use the following theorem from [12]. It gives a weak convergence result similar to the one for the polynomial case, see e.g. [7].

Theorem 9 *Under the same assumptions as in theorem 8 we have*

$$\frac{1-x/\alpha_n}{\sqrt{1-1/\alpha_n^2}}\phi_n^2(x)d\mu(x) \xrightarrow{*} \frac{1}{\pi}\frac{dx}{\sqrt{1-x^2}}.$$

where $\xrightarrow{*}$ means convergence in the weak star topology for measures.

This result also holds onder the weaker condition $\mu' > 0$ a.e., but we do not need that here.

Now we are ready to derive computable error bounds for the inner products from section 3 when they are calculated using the quadrature formulas from section 4.

6 Error bounds

Let us recapitulate the problem under consideration. Given a set of poles A and a measure μ we want to compute the recurrence coefficients associated with them (up to arbitrarily high degree) and estimate how accurate the computed coefficients are, i.e. estimate how much the numerical values differ from the theoretical value. To compute the coefficients we use the formulas (3)-(5) where

we approximate the inner products using a quadrature rule based on a different set of orthogonal rational functions, which we assume known.

In this section we will estimate the accuracy of the computed recurrence coefficients, based on certain assumptions. First let us recall some notation. The functions whose recurrence coefficients we wish to compute are $\{\phi_k\}$. They have poles in A and are orthogonal with respect to μ. The functions used in the quadrature formulas are $\{\tilde{\phi}_k\}$, with poles in $\tilde{A}$. The number of coefficients we wish to compute is m, the number of nodes in the quadrature formula is n. The quadrature error for the function f is denoted by $e_n(f)$ and the Blaschke products associated with A and $\tilde{A}$ are $\{B_k(z)\}$ and $\{\tilde{B}_k(z)\}$ respectively. The assumptions we make in our analysis are the following:

- m and n are large (theoretically infinite),
- A and $\tilde{A}$ are bounded away from I,
- the quadrature error for each of the inner products in (3) - (5) is roughly the same (i.e. of the same order of magnitude); to justify this assumption note that for any p, q we have $Z_p(x)/Z_q(x) = 1 + Z_p(x)/Z_q(\alpha_p)$ and if $\alpha_p \neq \alpha_q$ and they are not both equal to infinity, then $Z_p(x)Z_q(x) = Z_q(\alpha_p)Z_p(x) + Z_p(\alpha_q)Z_q(x)$,
- the absolute errors on E_m and F_m are of the same order of magnitude and approximately equal to $e_n(Z_m\phi_{m-1}^2)$; this is justified by taking $k = m-1$ in (3) and by the previous assumption; we denote this error by Δ_{nm}.

Now we can prove the following theorem.

Theorem 10 *Put $\rho_1 = \min_{1\leq k\leq m} |\alpha_k|$. Then under the assumptions made in the introduction to this section we have the following error bound,*

$$\Delta_{nm} \lesssim \min_{1\leq r<\rho_1} J^{-1}(r) \max_{0\leq\theta\leq\pi} \frac{(1-\beta_{m-1}^2)|1+z^2|}{|1-\beta_m z|^2|1-\beta_{m-1}z|^2} \frac{|\tilde{B}_n(z)\tilde{B}_{n-1}(z)|}{|B_m(z)B_{m-1}(z)|} \tag{7}$$

where $z = J^{-1}(t)$ and $t = re^{\mathbf{i}\theta}$. If $A = \{\alpha, \alpha, \ldots\}$ however, then put $\rho_2 = min_{1\leq k\leq n}|1/\tilde{\beta}_k - \beta|$. In this case an alternative estimate is given by

$$\Delta_{nm} \lesssim \min_{0<r<\rho_2} \frac{1-\beta^2}{r^{2m-2}} \max_{0\leq\theta\leq\pi} |1+z^2||\tilde{B}_n(z)\tilde{B}_{n-1}(z)||1-\beta z|^{2m-5} \tag{8}$$

where $z = \beta + re^{\mathbf{i}\theta}$. We use the symbol $\lesssim$ to denote that the inequality only holds if n and m tend to infinity. The rest of the notation is as in the previous section.

PROOF. We start by showing that with $z = J^{-1}(t) \in \mathbb{D}$,

$$H(t) = H(J(z)) = \int_{-1}^{1} \frac{\tilde{\phi}_n^2(x)(1-x/\tilde{\alpha}_n)}{t-x} d\mu(x) \approx \frac{2z}{1-z^2} \frac{1-\tilde{\beta}_n^2}{1+\tilde{\beta}_n^2} \tag{9}$$

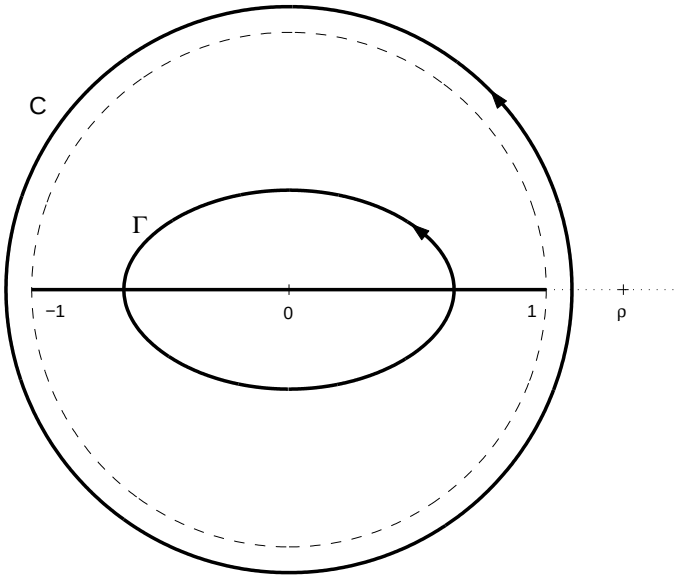

Fig. 1. Curves used in the proof of theorem 10.

as n tends to infinity (in the sense that the ratio tends to one). First note that

$$\int_{-1}^{1} \frac{dx}{(t-x)\sqrt{1-x^2}} = \frac{1}{2}\int_0^{2\pi} \frac{d\theta}{t-\cos\theta} = \frac{z}{\mathbf{i}}\oint_{\mathbb{T}} \frac{du}{(u-z)(1-zu)} = \frac{2\pi z}{1-z^2}$$

where the last equality follows from the residue theorem. Using this formula, theorem 9 and some algebra then proves (9).

To prove (7) write

$$\Delta_{nm} = |e_n(Z_m\phi_{m-1}^2)| = \left|\frac{1}{2\pi\mathbf{i}}\int_C \frac{Z_m(t)\phi_{m-1}^2(t)}{(1-t/\tilde{\alpha}_n)\tilde{\phi}_n^2(t)}H(t)dt\right|$$

using (6), where C is a circle with center 0 and radius $1 \le r < \rho_1$ (so that the integrand is analytic inside C and I is contained in C). It follows from [10] that all the quadrature nodes are on I (and therefore inside C), so we may indeed write down the above formula. Use the transformation $t = J(z)$ to convert this into

$$\Delta_{nm} = \left|\frac{1}{4\pi\mathbf{i}}\int_\Gamma \frac{Z_m(J(z))\phi_{m-1}^2(J(z))}{(1-J(z)/J(\tilde{\beta}_n))\tilde{\phi}_n^2(J(z))}H(J(z))\left(\frac{1}{z^2}-1\right)dz\right| \qquad (10)$$

where $\Gamma = J^{-1}(C)$ is as shown in figure 1. Next use theorem 8 to obtain that

$$\frac{\phi_{m-1}^2(J(z))}{\tilde{\phi}_n^2(J(z))} \approx \frac{\tilde{B}_n^2(z)}{B_{m-1}^2(z)}\frac{(1-\tilde{\beta}_n z)^2}{(1-\beta_{m-1}z)^2}\frac{(1-\beta_{m-1}^2)}{(1-\tilde{\beta}_n^2)} \qquad (11)$$

if n and m tend to infinity and use the definition of $Z_m(x)$ to get after some

calculations

$$Z_m(J(z)) = \frac{1}{2}\frac{(1+\beta_m^2)(1+z^2)}{(z-\beta_m)(1-\beta_m z)}, \quad 1 - \frac{J(z)}{J(\tilde{\beta}_n)} = \frac{(z-\tilde{\beta}_n)(1-\tilde{\beta}_n z)}{z(1+\tilde{\beta}_n^2)} \ . \quad (12)$$

Combining formulas (9) - (12) and canceling terms yields

$$\Delta_{nm} \approx \left| \frac{1}{4\pi \mathbf{i}} \int_\Gamma \frac{(1+\beta_m^2)(1-\beta_{m-1}^2)(1+z^2)}{(1-\beta_m z)^2(1-\beta_{m-1}z)^2} \frac{\tilde{B}_n(z)\tilde{B}_{n-1}(z)}{B_m(z)B_{m-1}(z)} dz \right| . \quad (13)$$

To estimate an upper bound for the integral we take the length of Γ times the maximum of the absolute value of the integrand. An upper bound for length(Γ) is given by $2\pi J^{-1}(r)$ (the circumference of the smallest circle which contains Γ in its interior) [3] . Because all poles are real the absolute value of the integrand is symmetric with respect to $\theta = \pi$ so that we may limit our search for the maximum value to the range $0 \le \theta \le \pi$. Also note that $1 \le 1+\beta_m^2 < 2$. This gives

$$\Delta_{nm} \lesssim J^{-1}(r) \max_{0\le\theta\le\pi} \frac{(1-\beta_{m-1}^2)|1+z^2|}{|1-\beta_m z|^2|1-\beta_{m-1}z|^2} \frac{|\tilde{B}_n(z)\tilde{B}_{n-1}(z)|}{|B_m(z)B_{m-1}(z)|}$$

Taking the minimum over all possible curves gives (7).

In the case where $A = \{\alpha, \alpha, \dots, \}$ we can obtain a more accurate approximation to the integral in (13). Now the integrand is a meromorphic function with a pole of order $2m-1$ at $z = \beta$. The poles in $\tilde{B}_n(z)\tilde{B}_{n-1}(z)$ are not important because they are outside of Γ. Using the residue theorem we get that

$$\Delta_{nm} \approx \frac{1}{2}(1+\beta^2)(1-\beta^2)|\operatorname{Res}(F;\beta)|$$

where the function F equals

$$F(z) = \frac{(1+z^2)(1-\beta z)^{2m-5}}{(z-\beta)^{2m-1}} \tilde{B}_n(z)\tilde{B}_{n-1}(z).$$

If we put $G(z) = F(z)(z-\beta)^{2m-1}$ then we have that

$$\operatorname{Res}(F;\beta) = \frac{G^{(2m-2)}(\beta)}{(2m-2)!} \ .$$

The accurate numerical calculation of the high order derivative is very difficult, but we can approximate it using a Cauchy estimate, as given by theorem 10.26

[3] In practice we will approximate $J^{-1}(r) = r - \sqrt{r^2-1}$ by $1/(2r)$ (even for r close to 1). Since we are only interested in the number of correct digits of the recurrence coefficients (i.e. the order of magnitude of the error), this approximation is good enough and the formulas become a bit simpler (one could even drop the factor 1/2).

in [8, p. 213]. This then yields

$$|G^{(2m-2)}(\beta)| \leq \frac{(2m-2)!\max_C |G(z)|}{r^{2m-2}}$$

where C is a circle with center at β and radius r and such that $G(z)$ is analytic inside C. Obviously, the maximum radius for C is $\min_{1\leq k\leq n} |1/\tilde{\beta}_k - \beta|$. Putting everything together and taking the minimum over all possible circles then finally gives (8). $\square$

Remark 11 *Theoretically we could also have used the residue theorem to estimate the error in the first case. However, in practice this will usually be impossible. The numerical computation of the residues is very difficult and even if we can compute them, they tend to be very large (especially for large m and n) and it is a well known fact that summing large numbers (not all positive or negative) to obtain a very small number is a numerical disaster.*

The previous theorem provides us with an easy way of estimating the accuracy of the computed recurrence coefficients. However, it would be more useful if we could use the formulas in the opposite direction: suppose we want to compute the reccurence coefficients with a given precision, then how many nodes would we need in the quadrature formula? If we regard Δ_{nm} as a function of n, then this amounts to finding the inverse function $n = \Delta_{nm}^{-1}(\epsilon)$, where ϵ is the required precision. In general this will be very difficult, but if we make some additional assumptions, then we can obtain useful results. First we need the following lemma.

Lemma 12 *Let $F(r, \theta, n)$ be a real continuous function defined on the box $[r_1, r_2] \times [\theta_1, \theta_2] \times [0, \infty)$, which has continuous partial derivatives up to second order with respect to r and θ and such that $\partial F/\partial n < 0$ on the domain of definition. We will call a point (r^*, θ^*)* stationary *if it satisfies $\frac{\partial F}{\partial r}(r^*, \theta^*, n) = \frac{\partial F}{\partial \theta}(r^*, \theta^*, n) = 0$. Then put*

$$\epsilon = f(n) = \min_{r_1\leq r\leq r_2} \max_{\theta_1\leq\theta\leq\theta_2} F(r, \theta, n)$$

and assume that for each n there is a unique stationary point (r_n, θ_n) such that $f(n) = F(r_n, \theta_n, n)$. Then it holds that

$$n = f^{-1}(\epsilon) = \min_{r_1\leq r\leq r_2} \max_{\theta_1\leq\theta\leq\theta_2} F^{-1}(r, \theta, \epsilon) = F^{-1}(r_n, \theta_n, \epsilon)$$

where F^{-1} is the inverse function with respect to n, i.e. $F^{-1}(r, \theta, \epsilon) = n$ implies $F(r, \theta, n) = \epsilon$.

PROOF. First define the function $G(r, n) = \max_{\theta_1\leq\theta\leq\theta_2} F(r, \theta, n)$. Since we assumed that the saddle point (r_n, θ_n) is stationary, we may write down the following equations,

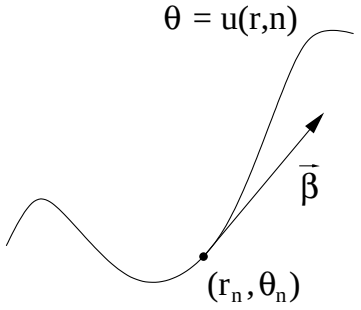

Fig. 2. Illustration for the proof of lemma 12.

$$\frac{\partial F}{\partial \theta}(r_n, \theta_n, n) = 0,\ \frac{\partial^2 F}{\partial \theta^2}(r_n, \theta_n, n) \leq 0, \tag{14}$$

$$\frac{\partial G}{\partial r}(r_n, n) = 0,\ \frac{\partial^2 G}{\partial r^2}(r_n, n) \geq 0. \tag{15}$$

Because F is continuous with continuous derivatives and (r_n, θ_n) is unique for each n, the equation $\partial F/\partial\theta = 0$ defines $\theta = u(r, n)$ as a function of r and n in a neighbourhood of (r_n, θ_n), such that $G(r, n) = F(r, u(r, n), n)$. The function u corresponds to a ridge on the surface $\epsilon = F(r, \theta, n)$. Figure 2 may serve as illustration. Using the chain rule and (14), we may rewrite the last equation of (15) as

$$\begin{aligned}&\frac{\partial^2 F}{\partial r^2}(r_n, \theta_n, n) + 2\frac{\partial^2 F}{\partial r \partial \theta}(r_n, \theta_n, n)\frac{\partial u}{\partial r}(r_n, n)\\ &+ \frac{\partial^2 F}{\partial \theta^2}(r_n, \theta_n, n)\left(\frac{\partial u}{\partial r}(r_n, n)\right)^2 \geq 0\end{aligned} \tag{16}$$

which is the same as saying that the second order directional derivative of F in the direction $\vec{\beta} = (1, \frac{\partial u}{\partial r}(r_n, n))$ is nonnegative, i.e.

$$\frac{\partial^2 F}{\partial \vec{\beta}^2}(r_n, \theta_n, n) \geq 0.$$

The relation $F(r, \theta, n) = \epsilon$ defines $n = F^{-1}(r, \theta, \epsilon)$ as a function of r, θ and ϵ. Differentiating this expression for fixed ϵ with respect to r gives

$$\frac{\partial F}{\partial r}(r, \theta, n) + \frac{\partial F}{\partial n}(r, \theta, n)\frac{\partial F^{-1}}{\partial r}(r, \theta, \epsilon) = 0$$

and substituting (r_n, θ_n) for (r, θ) we get

$$\frac{\partial F^{-1}}{\partial r}(r_n, \theta_n, \epsilon) = 0.$$

In the same way we obtain the relations

$$\begin{aligned}
\frac{\partial^2 F}{\partial r^2}(r_n,\theta_n,n)+\frac{\partial F}{\partial n}(r_n,\theta_n,n)\frac{\partial^2 F^{-1}}{\partial r^2}(r_n,\theta_n,\epsilon)&=0,\\
\frac{\partial^2 F}{\partial r\partial\theta}(r_n,\theta_n,n)+\frac{\partial F}{\partial n}(r_n,\theta_n,n)\frac{\partial^2 F^{-1}}{\partial r\partial\theta}(r_n,\theta_n,\epsilon)&=0,\\
\frac{\partial^2 F}{\partial \theta^2}(r_n,\theta_n,n)+\frac{\partial F}{\partial n}(r_n,\theta_n,n)\frac{\partial^2 F^{-1}}{\partial \theta^2}(r_n,\theta_n,\epsilon)&=0.
\end{aligned}$$

Substituting this into equation (16) yields

$$-\frac{\partial F}{\partial n}(r_n,\theta_n,n)\frac{\partial^2 F^{-1}}{\partial \vec{\beta}^2}(r_n,\theta_n,\epsilon)\geq 0$$

and because $\partial F/\partial n<0$ we get

$$\frac{\partial^2 F^{-1}}{\partial \vec{\beta}^2}(r_n,\theta_n,\epsilon)\geq 0.$$

Using the same procedure we obtain the conditions

$$\frac{\partial F^{-1}}{\partial \theta}(r_n,\theta_n,\epsilon)=0,\quad \frac{\partial^2 F^{-1}}{\partial \theta^2}(r_n,\theta_n,\epsilon)\leq 0.$$

Together these relations show that (r_n,θ_n) is a stationary saddle point for the function F^{-1}, thus proving the theorem. □

Using this lemma it is not difficult to invert the formulas from theorem 10, if we make the additional assumptions that all the poles in the quadrature formula are equal to each other as given in the following theorem.

Theorem 13 *Assume that $\tilde{A}=\{\tilde{\alpha},\tilde{\alpha},\ldots\}$ and let ϵ denote the required precision. If for each n the point (r_n,θ_n) where the minimum and maximum in (7) occur, satisfies the conditions of the previous lemma, then under the assumptions and with the notation of theorem 10 we have the following estimate for the number of nodes n needed in the quadrature formula,*

$$n\gtrsim \min_{1\leq r<\rho_1}\max_{0\leq\theta\leq\pi}\frac{1}{2}\frac{\log\left(\frac{J^{-1}(r)(1-\beta_{m-1}^2)|1+z^2|}{\epsilon|1-\beta_m z|^2|1-\beta_{m-1}z|^2|B_m(z)B_{m-1}(z)|}\right)}{\log\left|\frac{1-\tilde{\beta}z}{z-\tilde{\beta}}\right|}+\frac{1}{2}. \tag{17}$$

If furthermore $A=\{\alpha,\alpha,\ldots\}$ then under the same assumptions we have

$$n\gtrsim \min_{1\leq r<\rho_2}\max_{0\leq\theta\leq\pi}\frac{1}{2}\frac{\log\left(\frac{(1-\beta^2)|1+z^2||1-\beta z|^{2m-5}}{\epsilon r^{2m-2}}\right)}{\log\left|\frac{1-\tilde{\beta}z}{z-\tilde{\beta}}\right|}+\frac{1}{2}. \tag{18}$$

PROOF. The inequality (7) may be written

$$\Delta_{nm} \lesssim \min_{1\leq r<\rho_1} \max_{0\leq\theta\leq\pi} F(r,\theta,n)$$

where it is not difficult to see that F is certainly twice continuously differentiable with respect to r and θ and also $\partial F/\partial n < 0$. The rest of the proof is a simple application of the previous lemma. Since $F(r_n,\theta_n,n)$ is an upper bound for the actual error, $F^{-1}(r_n,\theta_n,\epsilon)$ will of course be a lower bound for the required number of nodes. This gives (17). The proof of (18) is completely analogous. □

Note that the proof of this theorem depends on the assumption that for each n the minimum and maximum occur at a point (r_n,θ_n) which satisfies the conditions (14) and (15). This assumption is very difficult to analyze theoretically, but the examples in the next section show that in practice it seems to be satisfied, in the sense that the formulas in this theorem provide reasonable estimates for the required number of nodes.

7 Numerical examples

The examples in this section serve as illustration for theorems 10 and 13. The figures show the number of correct digits for the coefficient E_m, computed using the procedure which was explained at the end of section 4. Computations were done in double precision, while the exact values of the coefficients were calculated in Fortran using a multiprecision package. In solid line is the exact number of correct digits, in dashed line the estimated number of correct digits. To estimate this number for each coefficient, we used the formulas from theorem 10 (implemented in Matlab with a nested call to the function `fminbnd`). These formulas yield an estimate for the absolute error on E_m. To obtain an estimate for the relative error (and thus for the number of correct digits), we used the asymptotic formula from theorem 8 as an approximation to the real value of E_m.

In all four examples μ is equal to the normalized Lebesgue measure on $[-1,1]$. The number of nodes in the quadrature formula equals $n = 150$. Let us start with the case where

$$A = \{2,4,6,\ldots\},$$
$$\tilde{A} = \{\omega,-\omega,2\omega,-2\omega,\ldots\}, \quad \omega = 1.1\ .$$

The results are as shown in figure 3. The oscillatory behaviour in the beginning

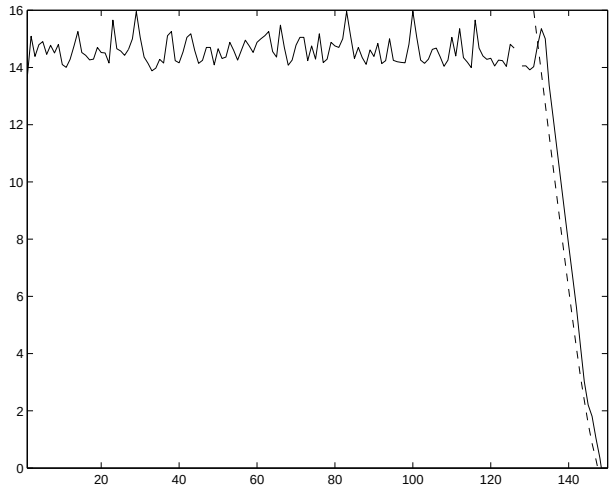

Fig. 3. Example 1. Estimated and actual number of correct digits for the recurrence coefficient E_m.

is of course due to machine limitations: working in IEEE double precision, no more than approximately 16 correct digits can be reached. Using 150 nodes in the quadrature formula, it is possible to compute over 130 coefficients up to machine precision. It is also clear from the figure that the estimated number of correct digits is very close to the actual number of correct digits.

In the previous example the poles in the computed functions are at a reasonable distance from the boundary. In that case, one might as well use Gauss-Legendre quadrature formulas to compute the inner products. This is shown in the following example, where we thus have

$$A = \{2, 4, 6, \ldots\},$$
$$\tilde{A} = \{\infty, \infty, \infty, \ldots\}.$$

Figure 4 gives the results. It is clear that the use of a quadrature formula based on rational functions was not necessary in example 1; ordinary Gauss-Legendre quadrature gives the same results. In this case all poles in $\tilde{A}$ are equal and we can use theorem 13 to predict the number of nodes needed to compute $m = 132$ coefficients up to a precision of $\epsilon = 10^{-15}$. This yields $n = 152$, which is a good approximation to the actual number of nodes used.

The advantage of using rational quadrature formulas only becomes apparent in the case where the poles in the computed functions are close to the boundary. In example 3 we take

$$A = \{1.01, 1.01, 1.01, \ldots\},$$
$$\tilde{A} = \{\infty, \infty, \infty, \ldots\},$$

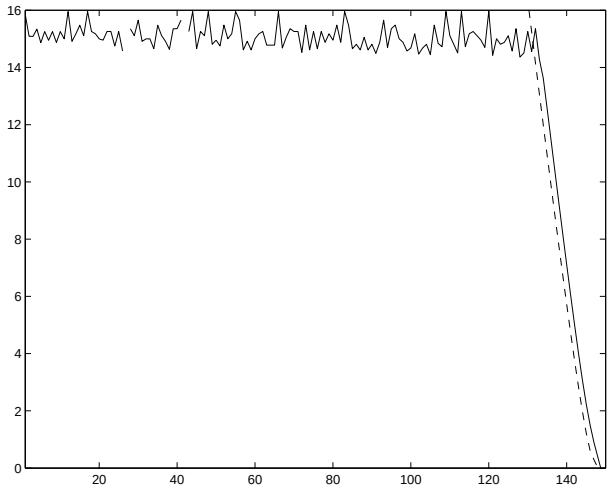

Fig. 4. Example 2. Estimated and actual number of correct digits for the recurrence coefficient E_m.

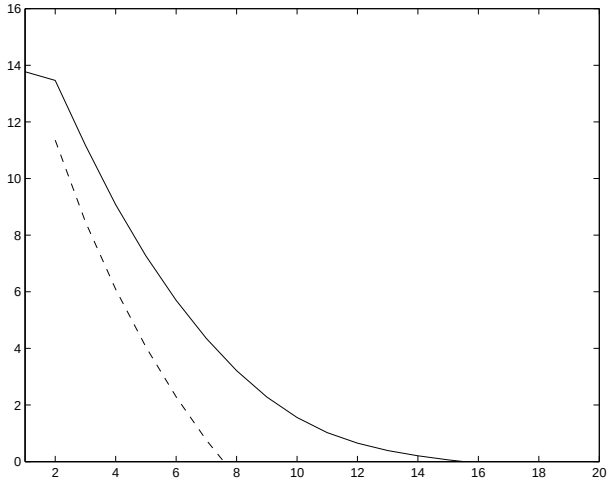

Fig. 5. Example 3. Estimated and actual number of correct digits for the recurrence coefficient E_m.

so there is a multiple pole very close to 1. The Gauss-Legendre quadrature formulas fail completely, as is shown in figure 5. Only 2 or 3 coefficients can be computed 'exactly'. In this case we have used the second formula in theorem 10 to estimate the number of correct digits. If we use the first formula, we get an estimate of only 1 correct digit even when $m = 2$, which is far from the actual number of correct digits. According to theorem 13, we would need approximately $n = 2600$ nodes in the quadrature formula to compute $m = 150$ coefficients up to machine precision.

However, if we take

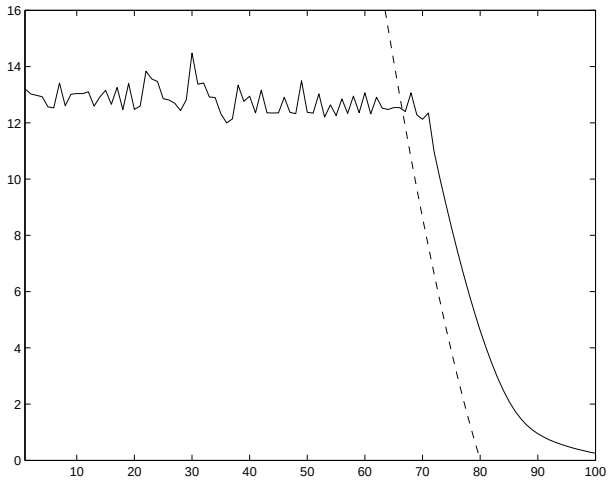

Fig. 6. Example 4. Estimated and actual number of correct digits for the recurrence coefficient E_m.

$$A = \{1.01, 1.01, 1.01, \ldots\},$$
$$\tilde{A} = \{1.03, 1.03, 1.03, \ldots\},$$

then we are able to compute approximately 70 coefficients up to machine precision, as shown in figure 6. The second formula in theorem 13 predicts that we would need $n = 157$ nodes to compute $m = 71$ coefficients up to a precision of $\epsilon = 10^{-13}$. This is a reasonable estimate.

8 Conclusion

Although we have only looked at two different situations, similar experiments give similar results. The main conclusions to be drawn are that the computation of orthogonal rational functions on an interval can be done using ordinary Gauss quadrature as long as the poles stay at a reasonable distance from the boundary. The number of nodes needed depends of course on the number of coefficients one wishes to compute and on the required precision. The formulas of theorem 13 provide reasonable estimates.

When the poles in the computed functions are close to the boundary, we can no longer use Gauss quadrature and we have to resort to rational quadrature formulas. This means that the computation of one set of rational functions requires the knowledge of a different but similar set of rational functions, which is of course a bit problematic. This problem is still open to investigation.

As a final remark we mention that it is also possible to compute the recurrence coefficients using a quadrature formula based on functions $\tilde{\phi}_n$ orthogonal with respect to a *different* measure $\tilde{\mu}$. Several experiments indicate however that

this is not a good idea, although this does not follow immediately from the error analysis.

References

[1] N.I. Achieser. *Theory of Approximation.* Frederick Ungar Publishing Co., New York, 1956.

[2] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions and quadrature on the real half line. *J. Complexity*, 2002. Accepted.

[3] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. *Orthogonal Rational Functions*, volume 5 of *Cambridge Monographs on Applied and Computational Mathematics.* Cambridge University Press, 1999.

[4] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonal rational functions and tridiagonal matrices. *J. Comput. Appl. Math.*, 153(1-2):89–97, 2003.

[5] P.J. Davis. *Interpolation and Approximation.* Blaisdell, New York, 1963.

[6] W. Gautschi. On the Construction of Gaussian Quadrature Rules from Modified Moments. *Math. Comp.*, 24:245–260, 1970.

[7] E. A. Rakhmanov. On the asymptotics of the ratio of orthogonal polynomials. *Math. USSR Sbornik*, 32:199–213, 1977.

[8] W. Rudin. *Real and Complex Analysis.* McGraw-Hill, New York, 1987. 3rd ed.

[9] J. Van Deun and A. Bultheel. An interpolation algorithm for orthogonal rational functions. *J. Comput. Appl. Math.*, 2003. Accepted.

[10] J. Van Deun and A. Bultheel. Orthogonal rational functions and quadrature on an interval. *J. Comput. Appl. Math.*, 153(1-2):487–495, 2003.

[11] J. Van Deun and A. Bultheel. Ratio asymptotics for orthogonal rational functions on an interval. *J. Approx. Theory*, 123:162–172, 2003.

[12] J. Van Deun and A. Bultheel. A weak-star convergence result for orthogonal rational functions. 2003. Submitted.