Computing rational Gauss-Chebyshev quadrature formulas with complex poles*

K. Deckers, J. Van Deun, A. Bultheel

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

Abstract

We provide a fast algorithm to compute arbitrarily many nodes and weights for rational Gauss-Chebyshev quadrature formulas integrating exactly in spaces of rational functions with arbitrary *complex* poles outside $[-1, 1]$. This algorithm is based on the derivation of explicit expressions for the Chebyshev (para-) orthogonal rational functions.

Keywords: quadrature formulas, orthogonal rational functions, complex poles, algorithm, numerical integration, Gauss-Chebyshev.

1 Introduction

A class of Gaussian quadrature formulas on $[-1, 1]$, based on orthogonal rational functions, and their fast and efficient computation, has been dealt with in [2]. These formulas are rational generalisations of the Gauss-Chebyshev formulas and are exact in a maximal space of rational functions with arbitrary but prefixed *real* poles outside $[-1, 1]$. In contrast with most existing rational quadrature formulas, the computational effort of the algorithm in computing the nodes and weights for the Gauss-Chebyshev quadrature formulas is very low, and under certain conditions on the poles the complexity can be shown to be of order $O(n)$.

Furthermore, in [1] the expressions of the nodes and weights for the Gauss-Cheby-shev quadrature formulas, as well as the theorem on the asymptotic distribution of the nodes, are extended to the case of *complex* poles, arbitrary but fixed outside $[-1, 1]$. The main purpose of this paper is to extend the algorithm in [2] to this case. In Section 2 we resume the most important formulas and theorems from [1, 2]. Then, because the nonlinear equation defining the nodes cannot be solved analytically, we perform in Section 3 a thorough analysis of this equation based on the distribution of the poles

*The work is partially supported by the Fund for Scientific Research (FWO) project “RAM: Rational modelling: optimal conditioning and stable algorithms”, grant #G.0423.05 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.

in the complex plane. Afterwards, in Section 4, we discuss some numerical methods to compute the nodes. Once the nodes are known, the computation of the weights is straightforward, so we will not discuss this in detail. Finally in Section 5 we conclude the paper with some numerical examples.

2 Preliminaries

The field of complex numbers will be denoted by $\mathbb{C}$ and the Riemann sphere by $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. For the real line we use the symbol $\mathbb{R}$ and for the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. The unit circle and the open unit disc are denoted respectively by $\mathbb{T} = \{z : |z| = 1\}$ and $\mathbb{D} = \{z : |z| < 1\}$. The complement of the interval $I = [-1, 1]$ with respect to a set X will be given by X^I, e.g. $\overline{\mathbb{C}}^I = \overline{\mathbb{C}} \setminus I$. Furthermore, if $b = \lfloor a \rfloor$ with $a \in \mathbb{R}$, then b is the largest integer so that $b \leq a$.

Suppose a sequence of poles $A = \{\alpha_1, \alpha_2, \ldots\} \subset \overline{\mathbb{C}}^I$ is given and define the factors

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

and the basis functions

$$b_0 = 1, \quad b_k(x) = b_{k-1}(x) Z_k(x), \quad k = 1, 2, \ldots \quad . \tag{2}$$

Then the space of rational functions with poles in A is defined as $\mathcal{L}_n = \text{span}\{b_0, \ldots, b_n\}$. In the special case of all $\alpha_k = \infty$, the expression in (1) becomes $Z_k(x) = x$ and the expression in (2) becomes $b_k(x) = x^k$. Let $\mathcal{P}_n$ denote the space of polynomials of degree less than or equal to n and define $\pi_n(x) = \prod_{k=1}^n (1 - x/\alpha_k)$, then we may write equivalently $\mathcal{L}_n = \{p_n/\pi_n, p_n \in \mathcal{P}_n\}$.

We denote the Joukowski Transformation $x = \frac{1}{2}(z + z^{-1})$ by $x = J(z)$, mapping the open unit disc $\mathbb{D}$ onto the cut Riemann sphere $\overline{\mathbb{C}}^I$ and the unit circle $\mathbb{T}$ onto the interval I. The inverse mapping is denoted by $z = J^{-1}(x)$ and is chosen so that $z \in \mathbb{D}$ if $x \in \overline{\mathbb{C}}^I$. With the sequence $A = \{\alpha_1, \alpha_2, \ldots\} \subset \overline{\mathbb{C}}^I$ we associate a sequence $B = \{\beta_1, \beta_2, \ldots\} \subset \mathbb{D}$ so that $\beta_k = J^{-1}(\alpha_k)$.

Given this sequence of complex numbers $B = \{\beta_1, \beta_2, \ldots\} \subset \mathbb{D}$, we define the Blaschke factors

$$\zeta_k(z) = \frac{z - \beta_k}{1 - \overline{\beta}_k z}, \quad k = 1, 2, \ldots$$

and the Blaschke products

$$B_0 = 1, \quad B_k(z) = B_{k-1}(z)\zeta_k(z), \quad k = 1, 2, \ldots \quad .$$

We define the inner product of two functions $f(x)$ and $g(x)$ as

$$\langle f, g \rangle_w = \int_{-1}^{1} f(x)\overline{g(x)} w(x) dx,$$

where the weight function $w(x)$ can be one of

$$w(x) = \begin{cases} (1 - x^2)^{-1/2} & , \quad i = 1 \\ \left(\frac{1-x}{1+x}\right)^{1/2} & , \quad i = 2 \\ (1 - x^2)^{1/2} & , \quad i = 3 \end{cases} . \tag{3}$$

Whenever we want to refer to one of these specific weights, we shall use the index i that is mentioned in the corresponding notation.

We define the involution operation or substar conjugate of a function $f(z)$ as

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

and the superstar transformation as

$$f^*(z) = \frac{b_n(z)}{b_{n*}(z)} f_*(z).$$

Note that the factor $\frac{b_n(z)}{b_{n*}(z)}$ merely replaces the poles $\{\overline{\alpha}_k\}_{k=1}^n$ of $f_*(z)$ by the poles $\{\alpha_k\}_{k=1}^n$ so that $\mathcal{L}_n^* = \mathcal{L}_n$. Furthermore, we define the para-orthogonal rational functions $Q_n(x,\tau)$ with respect to the orthogonal rational function (ORF) $\varphi_n(x)$ as

$$Q_n(x,\tau) = \varphi_n(x) + \tau\varphi_n^*(x), \quad \tau \in \mathbb{T}, \quad n \geq 1.$$

The use of these para-orthogonal rational functions lies in the fact that their zeros are simple and real and can be used as nodes in the quadrature formulas. The quadrature formulas follow from the next theorem.

Theorem 2.1. *Assume that the para-orthogonal rational function $Q_n(x,\tau) = \frac{q_n(x,\tau)}{\pi_n(x)}$ is regular, i.e. none of the zeros $x_{nk}(\tau)$ of $q_n(x,\tau)$ coincides with any of the poles. Define*

$$\lambda_{nk} = \left(\sum_{j=0}^{n-1} \left[\varphi_j\left(x_{nk}(\tau)\right) \overline{\varphi_j\left(x_{nk}(\tau)\right)} \right] \right)^{-1}.$$

Then the quadrature formula

$$\int_{-1}^{1} w(x) f(x) dx \approx \sum_{k=1}^{n} \lambda_{nk} f\left(x_{nk}(\tau)\right)$$

is exact for $f \in \mathcal{L}_{n-1} \cdot \mathcal{L}_{n-1*}$. In the special case in which α_n is real, this quadrature formula is exact for $f \in \mathcal{L}_n \cdot \mathcal{L}_{n-1*}$ (see [3, p. 490]).

The expression for the Chebyshev ORF $\varphi_n(x)$ related to the i^{th} weight in (3), as well as expressions for the computation of the nodes and weights in the quadrature formula are given in the next theorem. For the proof we refer to [1].

Theorem 2.2. *Let $x = J(z) \in \overline{\mathbb{C}}$ and $\alpha_k = J(\beta_k) \in \overline{\mathbb{C}}^I$. Suppose we define the numbers c, d, p and q for $i = 1, 2, 3$ according to Table 1. Then the orthonormal rational functions $\varphi_n(x)$ related to the i^{th} weight in* (3)*, with $n \geq 1$ are given by*

$$\varphi_n(x) = \sqrt{\frac{2^i}{\pi}} \sqrt{1 - |\beta_n|^2} \frac{q}{2z^{i-1} + q - 3} \left(\frac{z^i B_{n-1*}(z)}{1 - \beta_n z} - \frac{q}{(z - \beta_n) B_{n-1}(z)} \right)$$

and for $n = 0$ by

$$\varphi_0 = \sqrt{\frac{p}{\pi}}.$$

Furthermore, the nodes $x_{nk}(\tau) = \cos\theta_{nk}(\tau) \in [-1, 1]$ of the rational Gauss-Chebyshev quadrature formula are the zeros of the para-orthogonal rational function $Q_n(x, \tau)$. They satisfy

$$F_n\left(\theta_{nk}(\tau)\right) = \pi k - d\frac{\pi}{2}, \quad k = 1, 2, \ldots, n, \tag{4}$$

where

$$\begin{aligned} F_n(\theta) &= \sum_{j=1}^{n-1} f_{\beta_j}(\theta) + \frac{1}{2} f_{\beta_{n,\tau}}(\theta) - (n - c)\,\theta, \\ f_\beta(\theta) &= \arctan\frac{\sin\theta - \Im(\beta)}{\cos\theta - \Re(\beta)} + \arctan\frac{\sin\theta + \Im(\beta)}{\cos\theta - \Re(\beta)} \end{aligned}$$

and

$$\beta_{n,\tau} = \frac{\beta_n + \tau\overline{\beta}_n}{1 + \tau}$$

supposing $\tau \in \mathbb{T} \setminus \{-1\}$ is chosen so that $\beta_{n,\tau} \in]-1, 1[$.

Finally, the weights $\lambda_{nk}(\tau)$ of the rational Gauss-Chebyshev quadrature formula are given by

$$\lambda_{nk}(\tau) = 2\pi\frac{1 - (1 - d)[x_{nk}(\tau)]^{i-1}}{i + g_n(x_{nk}(\tau))}, \quad k = 1, 2, \ldots, n,$$

where

$$\begin{aligned} g_n(x) &= \sum_{j=1}^{n-1} \left\{P(z, \beta_j) + P(z, \overline{\beta}_j)\right\} + P(z, \beta_{n,\tau}), \\ P(z, \beta) &= \frac{1 - |\beta|^2}{|z - \beta|^2} \quad \textit{and} \quad x = J(z). \end{aligned}$$

Note that for all poles equal to infinity, $\varphi_n(x)$ becomes the Chebyshev polynomial of the first (respectively second) kind for $i = 1$ (respectively $i = 3$). Furthermore, we have in this case that $\mathcal{L}_n = \mathcal{P}_n$ and $\mathcal{L}_n \cdot \mathcal{L}_{n-1} = \mathcal{P}_{2n-1}$.

The distribution of the points $x_{nk}(\tau)$ as $n \to \infty$ depends on the asymptotic distribution of the poles, as shown below.

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

Table 1: Definition of c, d, p and q in function of i.

Theorem 2.3. *Assume that the sequence of poles* $A = \{\alpha_1, \alpha_2, \ldots\}$ *is bounded away from* I *and that the asymptotic distribution of the poles is given by a measure* ν *on (a subset of)* $\overline{\mathbb{C}}^I$*, i.e. for any continuous function* f *with compact support,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\alpha_k) = \int f(z) d\nu(z).$$

If $\nu = p\delta_\infty + (1-p)\nu_0$ *with* $0 \leq p \leq 1$ *(where* δ_z *is the unit measure whose support is the point* z*) and*

$$\int \log |t| d\nu_0(t) < \infty,$$

then the asymptotic distribution of the zeros of $Q_n(x,\tau)$ *is given by an absolutely continuous measure* λ *with weight function*

$$\lambda'(x) = \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}} \int \Re \left\{ \frac{\sqrt{t^2-1}}{t-x} \right\} d\nu(t)$$

where the square root is positive for $t > 1$ *and the branch cut is* $[-1, 1]$ *(see [1, p. 11–13]).*

In general, Equation (4) cannot be solved analytically. Assume for the remainder of this section that $A = \{\alpha_1, \alpha_2, \ldots\} \subset \overline{\mathbb{R}}^I$ (and thus $B = \{\beta_1, \beta_2, \ldots\} \subset I$, and the nodes and weights independent of τ). Then it is possible, however, to numerically calculate the nodes in the quadrature formulas using Newton's method. In this respect, the following properties are particularly interesting (see [2, p. 316]).

Lemma 2.4. *Suppose* $A = \{\alpha_1, \alpha_2, \ldots\} \subset \overline{\mathbb{R}}^I$. *Then the functions* $F_n(\theta)$ *from Theorem 2.2, with c given by Table 1, are strictly increasing for* $0 \leq \theta \leq \pi$. *If all poles have equal sign, these functions are concave (positive poles) or convex (negative poles) on* $]0, \pi[$. *In general there can be only one interior inflection point.*

Newton's method for finding zeros works particularly well for monotonic functions, especially if the initial values are not too far from the exact solutions. We discuss two different methods for determining these initial values.

The first method is based on linear extrapolation (LE). Let $\{\theta_{nk}\}_{k=1}^n$ denote the n exact zeros, then the initial values can be determined using one of the following equations

$$\theta_{n,k+1}^{(0)} = \theta_{n,k} + (\theta_{n,k} - \theta_{n,k-1}), \quad \theta_{n,1}^{(0)} = \theta_{n,0} = 0 \tag{5}$$

or

$$\theta_{n,n-k}^{(0)} = \theta_{n,n-k+1} + (\theta_{n,n-k+1} - \theta_{n,n-k+2}), \quad \theta_{n,n}^{(0)} = \theta_{n,n+1} = \pi \tag{6}$$

for $k = 1, 2, \ldots, n-1$.

The advantage of LE is that it can be used for any sequence of arbitrary poles $\{\alpha_k\}_{k=1}^n \subset \overline{\mathbb{R}}^I$. The disadvantage, however, is that the initial values cannot be determined all at the same time.

The second method for determining the initial values is based on the asymptotic zero distribution (AZD). First we assume that the poles tend to a fixed limit with increasing n, i.e. $\lim_{n\to\infty} \alpha_n = \alpha$, so that the zero distribution is given by a measure λ whose derivative is equal to

$$\lambda'(x) = \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}} \frac{\sqrt{1-1/\alpha^2}}{1-x/\alpha}.$$

The zero density on the interval $[-1, x]$ equals

$$t(x) = \int_{-1}^{x} \lambda'(u) du = \frac{1}{\pi} \arcsin \frac{\alpha x - 1}{\alpha - x} + \frac{1}{2}.$$

Solving for x gives

$$x = \frac{1 - \alpha \cos(\pi t)}{\alpha - \cos(\pi t)}, \qquad t \in [0, 1], \tag{7}$$

so if we evaluate this in n equidistant points

$$t_{nk} = \frac{2k-1}{2n} \in [0, 1], \qquad k = 1, 2, \ldots, n, \tag{8}$$

then we get an estimation for the zeros x_{nk}.

For a more general case of a finite number of arbitrary poles $\{\alpha_k\}_{k=1}^{n} \subset \overline{\mathbb{R}}^I$ for which the distribution is not known, we can use the cubic interpolating spline $s(t)$ through the points $(t_n(\xi_k), \arccos \xi_k)$, with

$$t_n(x) = \frac{1}{n\pi} \sum_{j=1}^{n} \arcsin \frac{\alpha_j x - 1}{\alpha_j - x} + \frac{1}{2} \tag{9}$$

and with

$$\xi_k = \cos\left(\pi \frac{2k-1}{2m}\right), \qquad k = 1, 2, \ldots, m, \tag{10}$$

the zeros of the Chebyshev polynomial $T_m(x)$ for a suitable value of m. Note that the spline is an approximation for the inverse of (9) and that (9) converges pointwise to $t(x)$ for $n \to \infty$. The initial values for the zeros θ_{nk} are then given by

$$\theta_{nk}^{(0)} = s(t_{nk}) \tag{11}$$

with t_{nk} the points from Equation (8).

The advantage of AZD is that all the initial values can be computed at the same time using vector or matrix operations (depending on whether the sequence contains all equal poles). Of course the estimates will be better for larger n because the method is based on the asymptotic behaviour of the poles.

3 Analysis of the equation for the nodes in the case of complex poles

As mentioned in Section 2, Equation (4) cannot be solved analytically, so the nodes have to be computed numerically using an iterative method. An analysis of the equation for the nodes has been done in the case of all real poles in Lemma 2.4. The equation had some properties that justified the use of Newton's method in computing the nodes. In this section we will perform a thorough analysis of this equation in the case of complex poles, to investigate whether these properties still hold.

The first property of Lemma 2.4 concerns the first derivative of $F_n(\theta)$ with respect to θ. It says that these functions are strictly increasing for $0 \leq \theta \leq \pi$. We will prove in the following lemma that this property holds in general for complex poles.

Lemma 3.1. *The functions $F_n(\theta)$ are strictly increasing for $0 \leq \theta \leq \pi$.*

Proof. Define K,L and M as

$$K = 1 - \Re(\beta)\cos\theta, \quad L = 1 - 2\Re(\beta)\cos\theta + |\beta|^2 \quad \text{and} \quad M = \Im(\beta)\sin\theta.$$

Then we have that

$$\begin{aligned} \frac{df_\beta(\theta)}{d\theta} &= \frac{2}{L}\left(K + \frac{2M^2(2K-L)}{L^2-4M^2}\right) \\ &\geq \frac{2K}{L} = \frac{2}{1+\frac{|\beta|^2-\Re(\beta)\cos\theta}{1-\Re(\beta)\cos\theta}} > 1, \end{aligned} \tag{12}$$

where the first inequality follows from the fact that $2K - L = 1 - |\beta|^2 > 0$ and that $L^2 - 4M^2$ is strictly positive on $]0, \pi[$. So we have that $\frac{dF_n(\theta)}{d\theta} > c - 1/2 > 0$, which proves the statement. □

The second property of Lemma 2.4 concerns the second derivative of $F_n(\theta)$ with respect to θ, stating that if all poles have equal sign, these functions are concave (positive poles) or convex (negative poles) on $]0, \pi[$. As will be proven in the next lemma, this property cannot be extended to the entire complex plane when simply divided into two parts, but it can be extended to a part of the complex plane.

Lemma 3.2. *Suppose that all the poles satisfy the condition given by*

$$|\Im(\beta)| \leq \sqrt{\frac{|\Re(\beta)|}{4-|\Re(\beta)|}(1-|\Re(\beta)|)} \tag{13}$$

(see Figure 1). If the real parts of all the poles have equal sign, then the functions $F_n(\theta)$ are concave (positive real part) or convex (negative real part) on $]0, \pi[$.

Proof. First note that $\frac{d^2 f_\beta(\theta)}{d\theta^2} = 2\sin(\theta)\widetilde{f}(\theta)$ with

$$\widetilde{f}(\theta) = (|\beta|^2 - 1)\times \frac{4\Re(\beta)|\beta|^2\cos^2(\theta) - 4|\beta|^2(1+|\beta|^2)\cos(\theta) + \Re(\beta)[(1+|\beta|^2)^2 + 4\Im(\beta)^2]}{(L^2-4M^2)^2},$$

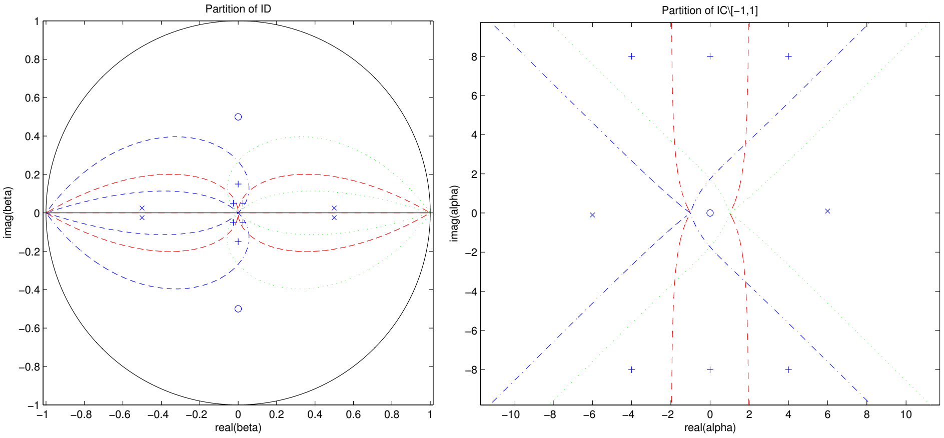

Figure 1: Partition of $\mathbb{D}$ (figure on the left), respectively the complex plane (figure on the right). Condition (13) holds for poles inside the red curve (figure on the left), respectively outside of the red curve (figure on the right). Furthermore, the poles within the areas marked with an 'x' and an 'o' satisfy condition (17), while the poles within the areas marked with a '+' satisfy condition (16).

and L and M as in Lemma 3.1. The denominator of $\widetilde{f}(\theta)$ is strictly positive on $]0,\pi[$. If $\beta = 0$, then the numerator equals zero. If on the other hand $\beta \neq 0$, then the zeros of the numerator are given by

$$\cos(\theta) = \begin{cases} \operatorname{sign}(\Re(\beta)) \left[\frac{1+|\beta|^2}{2|\Re(\beta)|} \pm \sqrt{\frac{\left(\frac{1+|\beta|^2}{2\Re(\beta)}\right)^2 - 1}{\left(\frac{\Re(\beta)}{\Im(\beta)}\right)^2+1}} \right] & , \quad \Re(\beta) \neq 0 \\ 0 & , \quad \Re(\beta) = 0 \end{cases} . \tag{14}$$

Note that $1 + |\beta|^2 > 2|\Re(\beta)|$ so that the numerator has no zeros iff (14) has no solution for $\theta \in]0,\pi[$, which means that

$$\Theta(\beta) := \frac{1+|\beta|^2}{2|\Re(\beta)|} - \sqrt{\frac{\left(\frac{1+|\beta|^2}{2\Re(\beta)}\right)^2 - 1}{\left(\frac{\Re(\beta)}{\Im(\beta)}\right)^2 + 1}} \geq 1. \tag{15}$$

Some computations lead to the condition given by (13). Under this condition we get that $\operatorname{sign}(\widetilde{f}(\theta)) = -\operatorname{sign}(\Re(\beta))$ on $]0,\pi[$, which ends the proof. □

Finally the third property of Lemma 2.4 concerns the second derivative of $F_n(\theta)$ with respect to θ again, stating that there can be only one interior inflection point. As before, this property cannot be extended to the entire complex plane, but it can be extended to a part of it, as will be proven in the next lemma.

Lemma 3.3. *If all the poles satisfy the condition given by*

$$\begin{aligned} \Im(\beta)^6 - \{14 + 8|\Re(\beta)| - \Re(\beta)^2\}\Im(\beta)^4 + \{(1+|\Re(\beta)|)^2(1 - 10|\Re(\beta)| - \Re(\beta)^2)\}\Im(\beta)^2 \\ - \{\Re(\beta)^2(1+|\Re(\beta)|)^4\} \geq 0, \end{aligned} \tag{16}$$

or if they all satisfy the condition given by (13) *together with*

$$\Im(\beta)^6 - \{14 - 8|\Re(\beta)| - \Re(\beta)^2\}\Im(\beta)^4 + \{(1 - |\Re(\beta)|)^2(1 + 10|\Re(\beta)| - \Re(\beta)^2)\}\Im(\beta)^2 \\ - \{\Re(\beta)^2(1 - |\Re(\beta)|)^4\} \leq 0 \quad (17)$$

(see Figure 1), then there can be only one inflection point $\theta_b \in]0, \pi[$*. Furthermore,* $\frac{dF_n(\theta)}{d\theta}$ *reaches a maximum (respectively minimum) for* $\theta = \theta_b$ *in the case that condition* (16) *(respectively the conditions* (13) *and* (17)*) is (are) satisfied .*

Proof. Differentiating $\widetilde{f}(\theta)$ with respect to θ gives us

$$\frac{d\widetilde{f}(\theta)}{d\theta} = u(\theta)v(\theta),$$

with

$$u(\theta) = \frac{4(1 - |\beta|^2)\sin\theta}{(L^2 - 4M^2)^3}, \quad v(\theta) = -d_3\cos^3(\theta) + d_2\cos^2(\theta) - d_1\cos(\theta) + d_0,$$
$$d_3 = 8\Re(\beta)|\beta|^4, \quad d_2 = 12|\beta|^4(1 + |\beta|^2), \quad d_1 = 6\Re(\beta)|\beta|^2[(1 + |\beta|^2)^2 + 4\Im(\beta)^2],$$
$$\text{and} \quad d_0 = (1 + |\beta|^2)[(\Re(\beta)^2 - \Im(\beta)^2)(1 + |\beta|^4) + 6|\beta|^4 - 4(\Re(\beta)^4 + \Im(\beta)^4)].$$

Note that $u(\theta) > 0$ for $\theta \in]0, \pi[$, that $d_2 \geq 0$, that $\text{sign}(d_3) = \text{sign}(d_1) = \text{sign}(\Re(\beta))$ and that $v(\theta)$ has a global minimum in $\cos(\theta) = \text{sign}(\Re(\beta))\Theta(\beta)$ if $\Theta(\beta) \leq 1$. Evaluating $v(\theta)$ in $\theta = 0$ and $\theta = \pi$ gives us

$$v(0) = -d_3 + d_2 - d_1 + d_0 = s_0(\beta)t_0(\beta) \quad \text{and} \quad v(\pi) = d_3 + d_2 + d_1 + d_0 = s_\pi(\beta)t_\pi(\beta),$$

where

$$s_0(\beta) = -(\Im^2(\beta) + (1 - \Re(\beta))^2), \quad s_\pi(\beta) = -(\Im^2(\beta) + (1 + \Re(\beta))^2),$$

and $t_0(\beta)$ (respectively $t_\pi(\beta)$) is given by the left part of (17) (respectively (16)) omitting the absolute values. If $\widetilde{f}(\theta)$ decreases with increasing $\theta \in]0, \pi[$ for every pole in the sequence, then there can be at most one interior inflection point in which $\frac{dF_n(\theta)}{d\theta}$ reaches a maximum. This will be the case if $v(0)$ and $v(\pi)$ are both negative (see the areas marked with a '+' on Figure 1), which is equivalent with the condition given by (16). If, on the other hand, $\widetilde{f}(\theta)$ increases with increasing $\theta \in]0, \pi[$ for every pole in the sequence, then there can be only one interior inflection point in which $\frac{dF_n(\theta)}{d\theta}$ reaches a minimum. This will certainly be the case if $v(0)$ and $v(\pi)$ are both positive and $\Theta(\beta) \geq 1$ for every pole in the sequence (see the areas marked with an 'x' on Figure 1), which is equivalent with the conditions given by (13) and (17). □

Remark 3.4. If $v(0)$ and $v(\pi)$ are both positive and $\theta_b = \arccos(\text{sign}(\Re(\beta))\Theta(\beta)) \in]0, \pi[$ with $v(\theta_b) \geq 0$, then $\widetilde{f}(\theta)$ will increase with increasing $\theta \in]0, \pi[$ so that there can be only one interior inflection point in which $\frac{dF_n(\theta)}{d\theta}$ reaches a minimum. Although there is no analytical proof at the moment of writing, extensive observations confirm that $v(\theta_b) < 0$ for every $\beta \in \mathbb{D}$ not satisfying the condition given by (13). From here on we will assume that this hypothesis is correct.

Remark 3.5. If the sequence of poles does not satisfy any of the conditions mentioned in Lemma 3.2 and 3.3, then there can be more than one interior inflection point. Based on many tests, the maximum number of interior inflection points found was $2n-3$ ($n \geq 2$), for a sequence of poles containing $n-2$ poles among the first $n-1$ poles in the areas marked with an ’o’, all having different real parts, and two real poles with different sign close to one in absolute value. One can indeed expect that the shape of $F_n(\theta)$ becomes worse when the poles are closer to the boundary $[-1,1]$. In the remainder of the text, we will assume that this kind of distribution of the poles is the worst situation possible.

Based on the previous remark, we will investigate the behaviour of $F_n(\theta)$ if one or more poles in the sequence tends to a value in $[-1,1]$.

Lemma 3.6. *Consider a finite sequence of n poles and suppose that we let one of the poles α_k in the sequence tend to a value $\alpha \in [-1,1]$. Then $\frac{dF_n(\theta)}{d\theta}$ has a local maximum in $\theta = \theta_{b_k}$ and $\lim_{\alpha_k \to \alpha} \theta_{b_k} = \arccos(\alpha)$.*

Proof. First note that

$$L^2 - 4M^2 = 4|\beta_k|^2\{(\cos(\theta) - \alpha_k)(\cos(\theta) - \overline{\alpha}_k)\} = 4|\beta_k|^2\{(\cos(\theta) - \Re(\alpha_k))^2 + \Im(\alpha_k)^2\},$$

so that for $\theta \in \{z \in \overline{\mathbb{C}} : \Re(z) \in [0,\pi] \quad \text{and} \quad \cos(z) \in \overline{\mathbb{R}}\}$, $L^2 - 4M^2$ has a minimal value for $\theta_{b_k^o} = \arccos(\Re(\alpha_k))$. Plug this into (12) and simplify to give

$$\frac{d}{d\theta} f_{\beta_k}(\theta_{b_k^o}) = \frac{2}{1 - |\beta_k|^2}.$$

So, if $\alpha_k \to \alpha$, then $\theta_{b_k^o} \to \arccos(\alpha) \in [0,\pi]$ and $\frac{d}{d\theta} f_{\beta_k}(\theta_{b_k^o}) \to \infty$. If $\theta \neq \theta_{b_k^o}$, then $\frac{d}{d\theta} f_{\beta_k}(\theta) \to 1$ for $|\beta_k| \to 1$. Assume now that the other poles α_l in the sequence are not too close to the boundary $[-1,1]$. Then $\sum_{l=1,l\neq k}^{n} \frac{d}{d\theta} f_{\beta_l}(\theta) < \infty$ so that $\frac{dF_n(\theta)}{d\theta}$ will have a local maximum in $\theta_{b_k} \to \arccos(\alpha)$. Clearly the same holds if the sequence does have a pole α_l that tends to the same value $\alpha_2 = \alpha \in [-1,1]$. If on the other hand $\alpha_2 \neq \alpha$, then $\frac{d}{d\theta} f_{\beta_l}(\theta) \to 1$ for each $\theta \in]\theta_{b_k} - \delta, \theta_{b_k} + \delta[$ with $\delta > 0$. In this case $\frac{dF_n(\theta)}{d\theta}$ will have another local maximum in $\theta_{b_l} \to \arccos(\alpha_2)$. □

Two things are of importance here. In the first place, note that the previous lemma explains the amount of $2n-3$ interior inflection points mentioned in Remark 3.5. If we have $n-2$ complex poles all tending to a different value in $]-1,1[$ and two real poles tending to 1 and -1, then according to the previous lemma this will result in $\frac{dF_n(\theta)}{d\theta}$ having n local maxima. Between two maxima there has to be a minimum, so we get $n + (n-1) = 2n-1$ inflection points of which $2n-3$ are in the interval $]0,\pi[$. Secondly, note that $\text{sign}(\Re(\beta))\Theta(\beta)$ for $|\beta| \to 1$ tends to $\Re(\beta)$ while this in turn will tend to $\Re(\alpha) \in [-1,1]$. Furthermore, because of the assumption in Remark 3.4 that $v(\theta) < 0$ in the only possible real zero of the numerator of $\widetilde{f}(\theta)$, this zero is a local maximum for $\frac{df_\beta(\theta)}{d\theta}$, but not necessarily a local maximum for $\frac{dF_n(\theta)}{d\theta}$. However, based on the previous lemma and the asymptotic behaviour of this zero we can conclude that it will approach the exact local maximum when the related pole or all the other poles come closer to the boundary, or when the multiplicity of the related pole increases.

4 Computing the nodes for complex poles

We will now discuss three methods for determining the initial values for Newton's method. Two of them have already been briefly discussed in Section 2 for the case of all real poles, namely LE and AZD.

4.1 Asymptotic zero distribution

Consider first the case of one multiple pole $\alpha \in \overline{\mathbb{R}}^I$. We can determine β in function of α as

$$\beta = \alpha - \operatorname{sign}(\alpha)\sqrt{\alpha^2 - 1} = \frac{1}{\alpha + \operatorname{sign}(\alpha)\sqrt{\alpha^2 - 1}}, \tag{18}$$

where the second expression, from a numerical point of view, is better because it avoids computing the difference between two almost equal values. For $\alpha = \infty$, (7) gives us

$$x = -\cos(\pi t) = \cos(\pi(1 - t)). \tag{19}$$

From (18) we can deduce that $\beta = 0$ for the given value of α, which means that we can determine the exact solution of (4) for a given multiplicity n of the pole. This gives us

$$\theta_{nk} = \pi t_{nk}, \tag{20}$$

where

$$t_{nk} = \frac{k - d/2}{n + c - 1} \tag{21}$$

and

$$x_{nk} = \cos\left(\pi t_{nk}\right). \tag{22}$$

So when we evaluate (7) in the points

$$t = 1 - t_{nk}, \tag{23}$$

with t_{nk} given by (21) instead of by (8), the initial values will be the exact solution. When using the same points for an arbitrary pole $\alpha \in \mathbb{R}^I$, this results in the following initial values

$$x_{nk}^{(0)} = \frac{1 + \alpha\cos\left(\pi t_{nk}\right)}{\alpha + \cos\left(\pi t_{nk}\right)} = \frac{2\beta + (\beta^2 + 1)\cos\left(\pi t_{nk}\right)}{\beta^2 + 1 + 2\beta\cos\left(\pi t_{nk}\right)}. \tag{24}$$

In the case of one real pole with multiplicity n, $F_n(\theta)$ can be simplified to $F_n(\theta) = f_{n,\beta}(\theta) - (n - c)\theta$, where

$$f_{n,\beta}(\theta) = (2n - 1)\arctan\frac{\sin(\theta)}{\cos(\theta) - \beta}.$$

If $F_n(\theta)$ is considered a function of β instead of θ, we will use the notation $H_n(\beta)$. So $H_n(\beta) = F_n(\theta_{nk}^{(0)})$, with $x_{nk}^{(0)} = \cos(\theta_{nk}^{(0)})$ given by (24).

As explained in Section 2, the advantage of AZD for one multiple pole is that we can compute all the initial values at the same time using vector operations. This makes it the fastest method of all.

A disadvantage, however, is that it can only be used if all poles are equal to each other. Clearly this has to be a real value, because in the case of complex poles the last pole α_n (if not real already) will be replaced by $\alpha_{n,\tau} = (\beta_{n,\tau} + \beta_{n,\tau}^{-1})/2 \in \overline{\mathbb{R}}^I$ during the computations. We will now prove another advantage of AZD for one multiple pole, but first we need the following two lemmas.

Lemma 4.1. *For each $\beta \in]-1,1[$ we have that*

$$f_{n,\beta}\left(\theta_{nk}^{(0)}\right) = f_{n,-\beta}\left(\pi t_{nk}\right), \tag{25}$$

with t_{nk} as in (21).

Proof. Using $x_{nk}^{(0)} = \cos\theta_{nk}^{(0)}$ gives us

$$f_{n,\beta}\left(\theta_{nk}^{(0)}\right) = (2n-1)\arctan\frac{\sqrt{1-\left(x_{nk}^{(0)}\right)^2}}{x_{nk}^{(0)}-\beta}.$$

Then replacing $x_{nk}^{(0)}$ with the expression in (24) and doing some computations proves the lemma. □

Lemma 4.2. *Function $H_n(\beta)$ is a non-increasing function with increasing β.*

Proof. From (24) and (25) we deduce that

$$H_n(\beta) = f_{n,-\beta}\left(\pi t_{nk}\right) - (n-c)\arccos\left(\frac{2\beta + (\beta^2+1)\cos\left(\pi t_{nk}\right)}{\beta^2+1+2\beta\cos\left(\pi t_{nk}\right)}\right).$$

Differentiating with respect to β gives us

$$\frac{dH_n(\beta)}{d\beta} = (1-2c)\frac{\sin\left(\pi t_{nk}\right)}{\left(\cos\left(\pi t_{nk}\right)-\beta\right)^2 + \left(\sin\left(\pi t_{nk}\right)\right)^2}.$$

The first part is strictly negative because it follows from Table 1 that $c > 1/2$, while the second factor is not negative because $t_{nk} \in [0,1]$, which means that the derivative is not positive. □

Where $F_n(\theta)$ is strictly increasing concave (respectively convex), Newton's method converges monotonically if $\theta_{nk}^{(0)} \leq \theta_{nk}$ (respectively $\theta_{nk}^{(0)} \geq \theta_{nk}$). The following theorem will prove that the initial values $\theta_{nk}^{(0)} = \arccos\left(x_{nk}^{(0)}\right)$ with $x_{nk}^{(0)}$ given by (24) satisfy this condition.

Theorem 4.3. *For each $\beta \in]-1,0]$ (respectively $\beta \in [0,1[$) we have $\theta_{nk}^{(0)} \geq \theta_{nk}$ (respectively $\theta_{nk}^{(0)} \leq \theta_{nk}$).*

Proof. We will only prove the theorem for $\beta \in [0,1[$. Proving the theorem for $\beta \in]-1,0]$ will be similar but with the inequality signs reversed. With $F_n(\theta)$ a strictly increasing concave function (see Lemma 2.4), it suffices to prove that $F_n\left(\theta_{nk}^{(0)}\right) \leq F_n\left(\theta_{nk}\right)$. Because θ_{nk} is the k^{th} exact solution of (4), we know that

$$F_n\left(\theta_{nk}\right) = k\pi - d\pi/2 = \text{constant}$$

for each $\beta \in [0,1[$. For $\beta = 0$ we also have that

$$F_n\left(\theta_{nk}^{(0)}\right) = F_n\left(\theta_{nk}\right) = k\pi - d\pi/2$$

because of (19)–(23). Using Lemma 4.2 then proves the inequality. □

Finally, consider the case of different poles $\{\alpha_k\}_{k=1}^n$ with $\alpha_k \in \overline{\mathbb{C}}^I$. Like before, we can determine the initial values as

$$\theta_{nk}^{(0)} = s\left(1 - t_{nk}\right),$$

with t_{nk} given by (21), instead of using (8) and (11), replacing (9) and (10) with

$$t_n(x) = \frac{1}{n\pi}\sum_{j=1}^{n-1}\Re\left(\arcsin\frac{\alpha_j x - 1}{\alpha_j - x}\right) + \arcsin\frac{\alpha_{n,\tau}x - 1}{\alpha_{n,\tau} - x} + \frac{1}{2} \tag{26}$$

and

$$\xi_k = \cos\left(\pi\frac{k - d/2}{m + c - 1}\right), \qquad k = 1, 2, \ldots, m.$$

This way, when all the poles are equal to infinity with multiplicity $n = m$, the initial values will again be the exact solution.

Again, as mentioned in Section 2, the advantage of AZD for different poles is that we can compute all the initial values at the same time using matrix operations. Theoretically, this makes it faster than LE, but slower than AZD for one multiple pole because we first have to form the cubic interpolating spline for the inverse of (26). In practice, however, when the matrices become too large, a bigger but slower memory can become necessary for the computations, which means that LE would become faster. Nevertheless, this can be solved by still determining the initial values using AZD, but computing the nodes in a for-loop using vector operations instead of matrix operations. A disadvantage of AZD for different poles is that the initial values do not necessarily converge monotonically to the exact solution. Another disadvantage of AZD in general is that it does not work well when too many poles in the given sequence are too close to the boundary. In this case, however, it is not excluded that still some of the nodes can be computed using AZD.

4.2 Linear extrapolation

As we have seen in Section 2, the initial values cannot be determined simultaneously when using LE. Not only is LE slower than AZD because of this, but also convergence while computing the previous node is essential so that it can continue with the computation of the next node. Because of this, LE is more interesting in combination with AZD when some but not all nodes can be computed with the latter, rather than as a method on its own. Nevertheless, as we will prove in the next theorem, LE has an advantage over AZD for sequences of different poles because under certain conditions it converges monotonically to the exact solution.

Theorem 4.4. *If all poles have positive (respectively negative) real parts and satisfy condition* (13)*, Newton's method converges monotonically if the initial values are determined by* (5) *(respectively* (6)*). (This statement is already mentioned for all real poles without proof in [2, p. 317]).*

Proof. Under the conditions given by the theorem, $F_n(\theta)$ is concave (poles with positive real parts) or convex (poles with negative real parts) on $]0,\pi[$ (see Lemma 3.2). Therefore we only need to prove that with (5) for poles with positive real parts (respectively (6) for poles with negative real parts), $\theta_{n,k}^{(0)} \leq \theta_{n,k}$ (respectively $\theta_{n,k}^{(0)} \geq \theta_{n,k}$) is satisfied for $k = 1, 2, \ldots, n$. We will only prove the theorem for poles with positive real parts. The proof for poles with negative real parts will be similar but with the inequality signs reversed. For $k = 1$ the inequality is trivial because $\theta_{n,1}^{(0)} = 0$ and $\theta_{n,1} \in \,]0,\pi[$. For $k > 1$ we base ourselves on the mean value theorem on the intervals $[\theta_{n,k-2}, \theta_{n,k-1}]$ and $[\theta_{n,k-1}, \theta_{n,k}]$:

$$\left\{ \begin{array}{l} \exists \theta_1 \in [\theta_{n,k-2}, \theta_{n,k-1}] : \frac{dF_n}{d\theta}(\theta_1) = \frac{F_n(\theta_{n,k-1}) - F_n(\theta_{n,k-2})}{\theta_{n,k-1} - \theta_{n,k-2}} \\ \exists \theta_2 \in [\theta_{n,k-1}, \theta_{n,k}] : \frac{dF_n}{d\theta}(\theta_2) = \frac{F_n(\theta_{n,k}) - F_n(\theta_{n,k-1})}{\theta_{n,k} - \theta_{n,k-1}} \end{array} \right. .$$

Because $F_n(\theta)$ on the right-hand side is evaluated in two successive exact solutions of (4), we can rewrite this as

$$\left\{ \begin{array}{l} \theta_{n,k-1} - \theta_{n,k-2} = \pi \Big/ \frac{dF_n}{d\theta}(\theta_1) \\ \theta_{n,k} - \theta_{n,k-1} = \pi \Big/ \frac{dF_n}{d\theta}(\theta_2) \end{array} \right. . \tag{27}$$

Further, because $\theta_1 \leq \theta_2$ and $F_n(\theta)$ is strictly increasing concave, we have that

$$\frac{dF_n}{d\theta}(\theta_1) \geq \frac{dF_n}{d\theta}(\theta_2),$$

or

$$\theta_{n,k-1} - \theta_{n,k-2} \leq \theta_{n,k} - \theta_{n,k-1},$$

so

$$\theta_{n,k}^{(0)} = \theta_{n,k-1} + (\theta_{n,k-1} - \theta_{n,k-2}) \leq \theta_{n,k-1} + (\theta_{n,k} - \theta_{n,k-1}) = \theta_{n,k}.$$

Finally, we remark that for $k = 2$ the value $\theta_{n,k-2} = \theta_{n,0} = 0$ is not a solution of (4). We have that $F_n(\theta_{n,0}) = F_n(0) = 0$ which means that here the first expression in (27) has to be replaced with

$$\theta_{n,k-1} - \theta_{n,k-2} = \left(\pi - d\frac{\pi}{2}\right) \Big/ \frac{dF_n}{d\theta}(\theta_1).$$

When $d = 0$ this does not make any difference, but when $d = 1$ this becomes

$$\theta_{n,k-1} - \theta_{n,k-2} = \pi \Big/ \left(2\frac{dF_n}{d\theta}(\theta_1)\right).$$

□

In the case of poles with positive and negative real parts, we need the extra condition (17) to assure that $F_n(\theta)$ has at most one interior inflection point $\theta_b \in \,]0,\pi[$, for which $\frac{dF_n(\theta)}{d\theta}$ has a local minimum. In this way, $F_n(\theta)$ will be strictly increasing concave on $]0,\theta_b[$ while the other part will be strictly increasing convex (see Lemma 3.3). So, when applying (5) (respectively (6)) to the interval $[0,\theta_b]$ (respectively $[\theta_b,\pi]$), Newton's method also converges monotonically for these kinds of sequences of poles when using LE.

In practice it is not easy to compute the interior inflection point. Nevertheless, because convergence is not excluded when an initial value lies on the wrong side of the exact solution, computing this is not necessary at all. Instead, we can start with one of both formulas (i.e. (5)) until it does not converge anymore and then continue with the other one (i.e. (6)) if necessary. For other kinds of sequences of poles, convergence cannot be assured with LE but is not necessarily excluded. However, for sequences containing poles in the area marked with an 'o' close to the boundary, it is very unlikely that all the nodes can be computed using LE because of peak shaped interior local maxima of $\frac{dF_n(\theta)}{d\theta}$.

4.3 An alternative way to determine initial values

The goal is to find a method that combines the advantages of AZD (that all nodes can be computed at the same time) and LE (assuring monotonic convergence), but works well for sequences of poles close to the boundary. To realise this, we will base ourselves on the distribution of the inflection points of $F_n(\theta)$.

Consider the list $\{\theta_{b_j}\}_{j=1}^m$ of inflection points of $F_n(\theta)$ with $\theta_{b_j} < \theta_{b_{j+1}}$. Certainly this list is not empty because $F_n(\theta)$ always has an inflection point in $\theta_{b_1} = 0$ and $\theta_{b_m} = \pi$, so that $m \geq 2$. Note that if $\frac{dF_n(\theta)}{d\theta}$ has a local maximum in θ_{b_j}, then it has a local minimum in $\theta_{b_{j-1}}$ and $\theta_{b_{j+1}}$, and vice versa. Assume now that we know the local maxima $\theta_{b_{2j-r}}$, with $r = 0$ or $r = 1$ and $j = 1, \ldots, \lfloor (m+r)/2 \rfloor$. Then we can compute the nodes $\{\theta_{k_0}, \ldots, \theta_{k_1}\} \subset [\theta_{b_{2j-r-1}}, \theta_{b_{2j-r+1}}]$ starting from $\theta_{b_{2j-r}}$ as initial value, assuming we can determine the indices k_0 and k_1 exactly. Note that this way we try to assure monotonic convergence of the initial values rather than trying to determine initial values close to the exact solution.

In practice we do not know the local maxima nor the indices k_0 and k_1. However, we can try to estimate the local maxima and the indices using Lemma 3.6. Suppose we have a sequence of $N = \lfloor (m+r)/2 \rfloor$ different poles α_j, each with multiplicity n_j so that $\sum_{j=1}^N n_j = n$. Then the local maxima can be estimated using $\cos(\theta_{b_{2j-r}}) = \text{sign}(\Re(\beta_j)) \min(\Theta(\beta_j), 1)$ where $\Theta(\beta)$ is given by (15). Furthermore, if $\alpha_j \in [-1, 1]$, $\frac{df_{\beta_j}(\theta_{b_{2j-r}})}{d\theta} = \infty$, while for every other θ we have that $\frac{df_{\beta_j}(\theta)}{d\theta} = 1$. So, with $\delta > 0$ and $\epsilon > 0$ we get that

$$\begin{aligned} f_{\beta_j}(\theta_{b_{2j-r}} + \delta) - f_{\beta_j}(\theta_{b_{2j-r}} - \epsilon) &= \big[f_{\beta_j}(\pi) - (\pi - \theta_{b_{2j-r}} - \delta)\big] \\ &\quad - \big[f_{\beta_j}(0) + (\theta_{b_{2j-r}} - \epsilon)\big] = f_{\beta_j}(\pi) - f_{\beta_j}(0) - \pi + \delta + \epsilon = \pi + \delta + \epsilon. \end{aligned}$$

Consequently, $F_n(\theta)$ shows a jump in $\theta_{b_{2j-r}}$ equal to $n_j\pi$ (if α_j is not the last pole in the sequence, otherwise the jump will be equal to $(n_j - 1/2)\pi$), while, if θ_k and θ_{k+1} are two successive exact solutions of (4), then $F_n(\theta_{k+1}) - F_n(\theta_k) = \pi$. Because of this, we can use as a rule that the multiplicity of $\theta_{b_{2j-r}}$ as initial value equals n_j.

Clearly this method improves the closer the poles in the sequence are to the boundary. For this reason, a disadvantage of this method is that it does not work well if the sequence of poles contains too many poles away from the boundary. Like for AZD, we can combine this method with LE if not all the nodes can be computed using this method. In the remainder of this text we will refer to this method as the method of asymptotic inflection point distribution (AIPD).

4.4 Some additional remarks

The main advantage of Newton's method is that, if the iterations converge, they converge quadratically to the exact nodes when starting from initial values close to the exact solution. Only LE and AIPD can theoretically assure (monotonic) convergence in a part of the complex plane. But none of the methods discussed before can assure convergence of the initial values to the exact solution for any sequence of complex poles. In practice, however, Newton's method always seems to converge for sequences only containing poles away from the boundary when using AZD, as well as for sequences only containing poles close to the boundary when using AIPD.

For sequences containing poles close to the boundary as well as poles away from the boundary, Newton's method does not always converge for all the nodes when using one of both methods for determining the initial values. But in many cases, improvement is possible through combining AZD or AIPD with LE, as mentioned before. The computations for such a combination are not difficult to organise. However, it cannot always solve the problem of convergence so that other combinations (like for instance combining AZD with AIPD) need to be considered, for which the computations are more difficult to organise. Many observations strongly indicate that, when considering more combinations of methods, Newton's method will converge for any sequence of poles.

However, for sequences containing poles extremely close to the boundary, the results for the nodes need to be very precise to get accurate results for the weights as well. When using a less severe criterium of accuracy during the iterations, Newton's method will converge but the results will be very inaccurate. On the other hand, when using a more severe criterium of convergence, rounding errors can cause the desirable accuracy to be unreachable. For this reason, some additional computations using the method of bisection can be necessary to get full precision. Nevertheless, this additional number of iterations will be negligible compared with the number of iterations needed when only using the method of bisection.

Finally note that, when the last pole α_n in the sequence is not real, it can be replaced by any $\alpha \in \overline{\mathbb{R}}^I$ during the computations, because one can always find a $\tau \in \mathbb{T}$ so that $\alpha_{n,\tau} = \alpha$. In such cases a proper choice of α can sometimes improve the results as well.

5 Numerical results

We shall now look at some examples of sequences of poles and compare the results for the nodes and weights for the different methods discussed in the previous section with the results for the nodes and weights when only using the method of bisection. With Δx and $\Delta\lambda$ we denote the maximal difference in absolute value between the resulting nodes and weights. We will also compare the largest number of iterations, denoted by p_{max}, needed for the initial value that converged the slowliest, as well as the total number of iterations, denoted by p_{total}, assuming the iterations for each node were done serially like for LE. All the computations in the examples that follow are done with $\tau = 1$ for the first weight function ($i = 1$). Furthermore, the Newton iterations are done with an absolute accuracy of $|\theta_k^{(p-1)} - \theta_k^{(p)}| \leq 10^{-10}$, while for the method of bisection the nodes are computed with an absolute accuracy of $|\theta_k^{(p-1)} - \theta_k^{(p)}| \leq 5 \times 10^{-16}$. To check the accuracy of the results for each method on its own, the absolute difference $|\pi - \sum_{k=1}^{n} \lambda_k|$, which theoretically has to equal zero, is checked as well.

Example 5.1. Let us first consider the sequence of poles given by

$$\begin{aligned} A = \{\alpha_{k+6} = 2.005 + 1.905\mathbf{i} + 0.001k(1+\mathbf{i}), \quad k = -5, \dots, 5\} \cup \\ \{\alpha_{k+17} = -2.000 - 1.900\mathbf{i} - 0.001k(1+\mathbf{i}), \quad k = -5, \dots, 5\}. \end{aligned} \tag{28}$$

Then Figure 2 shows the graph of $F_n(\theta)$, $\frac{dF_n(\theta)}{d\theta}$ and $\frac{d^2F_n(\theta)}{d\theta^2}$. Note that, although the graph of $F_n(\theta)$ looks like a straight line, it is not as can be deduced from the graph of the first and second derivative. The given sequence does not satisfy any of the conditions given by Lemma 3.2 and 3.3. However, none of the poles are close to the boundary, thus we can expect that Newton's method works well for the given sequence. Table 2 summarises the results for Newton's method and for the method of bisection.

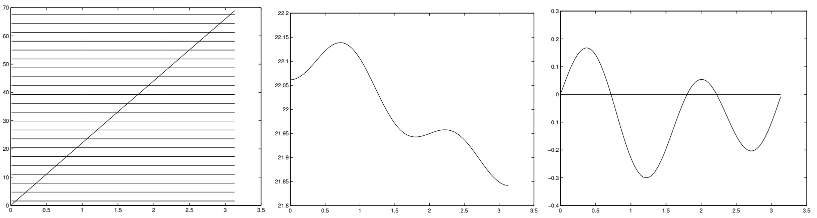

Figure 2: The graph of $F_n(\theta)$, $\frac{dF_n(\theta)}{d\theta}$ and $\frac{d^2F_n(\theta)}{d\theta^2}$ for the sequence of poles given by (28).

method	Δx	$\Delta\lambda$	$\lvert\pi - \sum_{k=1}^{n}\lambda_k\rvert$	p_{max}	p_{total}
bisection	0	0	8.9×10^{-16}	52	1133
AZD	5.6×10^{-16}	2.8×10^{-17}	8.9×10^{-16}	3	66
LE	7.8×10^{-16}	2.8×10^{-17}	8.9×10^{-16}	3	52
AIPD	7.8×10^{-16}	2.8×10^{-17}	8.9×10^{-16}	4	84

Table 2: The results for Newton's method and for the method of bisection for the sequence given by (28).

Example 5.2. Next, consider the sequence of poles given by

$$B = \{\alpha_k = 0.75 + 0.01\mathbf{i}, \quad k = 1, \dots, 4\} \cup \{\alpha_k = 2, \quad k = 5, 6\}. \tag{29}$$

Then Figure 3 shows the graph of $F_n(\theta)$, $\frac{dF_n(\theta)}{d\theta}$ and $\frac{d^2F_n(\theta)}{d\theta^2}$. The given sequence contains poles close to the boundary as well as poles away from the boundary. Because of this, not all the nodes can be computed with Newton's method when using only one method for determining the initial values. But all the nodes can be computed when LE is combined with one of the other two methods. Table 3 summarises the results for the two combinations and for the method of bisection. With n_i we denote the number of initial values that converged for the specific method.

Example 5.3. Finally consider the sequence of poles given by

$$C = \{\alpha_k = -\alpha_{k+5} = 0.75 + 0.01\mathbf{i}, \quad k = 1, \dots, 4\} \cup \{\alpha_5 = -\alpha_{10} = 2\}. \tag{30}$$

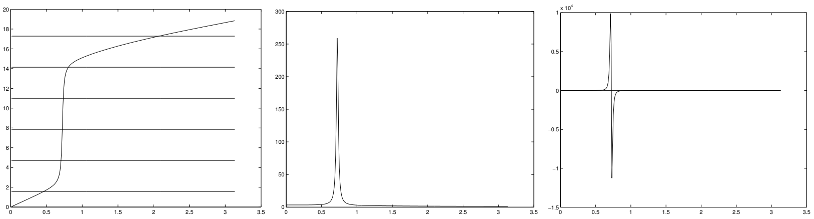

Figure 3: The graph of $F_n(\theta)$, $\frac{dF_n(\theta)}{d\theta}$ and $\frac{d^2F_n(\theta)}{d\theta^2}$ for the sequence of poles given by (29).

method	Δx	$\Delta\lambda$	$\lvert\pi - \sum_{k=1}^{n}\lambda_k\rvert$	p_{max}	p_{total}	n_i
bisection	0	0	1.8×10^{-15}	52	309	6
AZD and LE	2.3×10^{-15}	1.3×10^{-15}	1.3×10^{-15}	6 5	26 5	5 1
AIPD and LE	3.9×10^{-16}	1.9×10^{-15}	3.1×10^{-15}	8 6	32 6	5 1

Table 3: The results for Newton's method and for the method of bisection for the sequence given by (29).

Then Figure 4 shows the graph of $F_n(\theta)$, $\frac{dF_n(\theta)}{d\theta}$ and $\frac{d^2F_n(\theta)}{d\theta^2}$. The given sequence now contains much more poles close to the boundary than poles away from the boundary. In this case not all the nodes can be computed when combining AZD with LE. But they can now be computed when only using AIPD. Table 4 summarises the results for Newton's method and for the method of bisection.

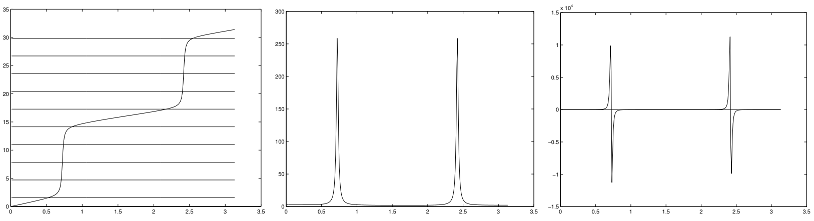

Figure 4: The graph of $F_n(\theta)$, $\frac{dF_n(\theta)}{d\theta}$ and $\frac{d^2F_n(\theta)}{d\theta^2}$ for the sequence of poles given by (30).

method	Δx	$\Delta\lambda$	$\lvert\pi - \sum_{k=1}^{n}\lambda_k\rvert$	p_{max}	p_{total}
bisection	0	0	5.8×10^{-15}	52	516
AIPD	1.8×10^{-15}	6.7×10^{-15}	0	9	60

Table 4: The results for Newton's method and for the method of bisection for the sequence given by (30).

6 Conclusion

We have provided a fast algorithm to compute arbitrarily many nodes and weights for rational Gauss-Chebyshev quadrature formulas integrating exactly in spaces of rational functions with arbitrary *complex* poles outside $[-1, 1]$. This algorithm was based on the derivation of explicit expressions for the Chebyshev (para-)orthogonal rational functions.

Once the nodes were known, the computation of the weights was straightforward. These nodes could not be found analytically but had to be solved numerically (using Newton's method) from the equation $f(x_k) = C_k$, $k = 1, \ldots, n$, where C_k is a constant only depending on the node x_k to be computed and f is a smooth, monotonically increasing function that depends on the first n poles. The success of our algorithm depended on obtaining accurate initial values for the exact solutions and on a thorough analysis of the graph of $f(x)$, with $x \in [-1, 1]$, based on the distribution of the poles in the complex plane.

Some characteristics of this function were already given for the case of all *real* poles in [2] and two methods were derived for determining a sequence of initial values for the nodes, assuring convergence of Newton's method. For *complex* poles, however, these characteristics did not hold in general and these two methods were not sufficient to assure convergence to the exact values in every possible case.

We derived a new method to obtain more accurate initial values for the case of arbitrary complex poles. In all cases under consideration, these values converged to the exact solutions. In some exceptional cases, a few additional iterations using the method of bisection were needed to obtain full precision.

When combining different methods for determining the initial values for Newton's method, we restricted ourselves in this paper to the combinations which were easy to organise in practice. Other combinations to improve the results (accuracy versus total number of iterations) are open for further research. Also determining the value for the parameter τ (in case the last pole in the sequence is not real) so that the computations of the nodes are optimal is an unresolved problem.

References

[1] Deckers, K. and J. Van Deun and Bultheel, A., *"Rational Gauss–Chebyshev quadrature formulas for complex poles outside* $[-1, 1]$*"*, Report TW 448, Department of Computer Science, Catholic University of Leuven, 2006.

[2] Van Deun, J. and A. Bultheel and González Vera, P., *"On computing rational Gauss–Chebyshev quadrature formulas"*, Mathematics of Computation, 75, 307-326, 2006.

[3] Van Deun, J. and A. Bultheel, *"Orthogonal rational functions and quadrature on an interval"*, Journal of Computational and Applied Mathematics, 153(1-2), 487-495, 2003.