A new quadrature method for locating the zeros of analytic functions with applications to engineering problems

Nikolaos I. Ioakimidis

Division of Applied Mathematics and Mechanics, Department of Engineering Sciences, School of Engineering, University of Patras, GR-265 04 Patras, Greece
e-mail: n.ioakimidis@upatras.gr

Abstract A new method for the computation of real or complex zeros of analytic functions and/or poles of meromorphic functions outside a fundamental interval \([a, b]\) of the real axis is proposed. This method is based on appropriately taking into account the error terms in the Gauss– and Lobatto–Chebyshev quadrature rules for ordinary integrals and it leads to a very simple non-iterative algorithm for the computation of these zeros. The results obtained by this algorithm with very few functional evaluations are of a very good accuracy and they can further be improved, if required, by local methods, which are generally inappropriate for the original localization of the zeros. The proposed method was tested in two engineering problems: a problem of neutron moderation in nuclear reactors and a problem of determining the critical buckling load of an elastic frame. The corresponding transcendental equations were solved by this method and numerical results for their zeros are presented. In all equations solved, numerical values for their zeros accurate to at least five significant digits were obtained by the present method with no more than thirteen functional evaluations.

Keywords Analytic functions · Meromorphic functions · Transcendental functions · Zeros · Poles · Chebyshev polynomials · Chebyshev functions of the second kind · Numerical integration · Gauss– and Lobatto–Chebyshev quadrature rules · Convergence · Error term · Quadrature error · Neutron moderation · Elastic buckling

1980 Mathematics Subject Classification numbers 65H05, 65D30, 65E05

Copyright © 1984 by the author

Presentation details2 The results of this technical report were presented to the “International Congress on Computational and Applied Mathematics” held to celebrate the 10th anniversary of

---

1 Both the internal and the external links (all appearing in blue) were added by the author on 21 January 2018 for the online publication of this technical report.

2 These details were also added by the author on 21 January 2018 for the online publication of this technical report.
1. Introduction

Several methods for the localization of zeros of analytic functions in the complex plane are available in the literature [10, 19]. Among these we mention the graphical methods [14, 15] and the semi-analytical methods of Delves and Lyness [8], Abd-Elall, Delves and Reid [1] and Burniston and Siewert [3, 17]. Improvements of the method of Delves and Lyness [8] were recently proposed by Carpentier and Dos Santos [4] and by Li [13]. Moreover, a generalization of the method of Burniston and Siewert [3, 17] was recently made by Anastasselou and Ioakimidis [2]. These methods are generally used for the approximate localization of zeros of analytic functions (or of poles of meromorphic functions), which is further used (if necessary) for the original approximations to these zeros for their use in local methods [10] like the classical Newton–Raphson method.

In a recent paper [11], Ioakimidis and Anastasselou proposed a new method for the computation of one simple zero of an analytic function inside a finite interval \([a, b]\) of the real axis. This method makes use of the Gauss– and Lobatto–Chebyshev quadrature rules for Cauchy-type principal value integrals on the interval \([-1, 1]\) after the error terms were ignored. In this technical report, we extend this method for locating zeros of analytic functions (or poles of meromorphic functions) outside a finite interval \([a, b]\) of the real axis. We are again based on the Gauss– and Lobatto–Chebyshev quadrature rules (but now for ordinary and not for Cauchy-type principal value integrals) [12, pp. 383, 396] and, particularly, on the error terms for these rules due to poles of the integrands. These terms were investigated in great detail by many researchers. Here we cite the papers by Chawla [5], Chawla and Jain [6] and Donaldson and Elliott [9] only since these are the most relevant to our developments.

After the development of the method in Section 2, we will apply it to two transcendental equations appearing in engineering problems (neutron moderation [16] and elastic buckling [18, p. 66]) in Sections 3 and 4, respectively, and we will present numerical results to illustrate the effectiveness and accuracy of the proposed method. Finally, in Section 5, we will proceed to a discussion of the method concerning its advantages and disadvantages when it is compared with the aforementioned competitive methods [1, 3, 4, 8, 10, 13–17, 19].

2. The proposed method

We consider the Gauss– and Lobatto–Chebyshev quadrature rules for ordinary integrals [12, pp. 383, 396] concerning the weight function \(w(t) = 1/\sqrt{1-t^2}\), i.e.

\[
\frac{1}{\pi} \int_{-1}^{1} \frac{g(t)}{\sqrt{1-t^2}} \, dt = \frac{1}{n} \sum_{i=1}^{n} g(t_{iG}) + R_{nG},
\]

\[
\frac{1}{\pi} \int_{-1}^{1} \frac{g(t)}{\sqrt{1-t^2}} \, dt = \frac{1}{n} \sum_{i=0}^{n} g(t_{iL}) + R_{nL},
\]

respectively. Here the nodes \(t_{iG}\) and \(t_{iL}\) are determined by

\[
t_{iG} = \cos \left( \frac{(2i-1)\pi}{2n} \right), \quad i = 1, 2, \ldots, n,
\]

\[
t_{iL} = \cos \left( \frac{i\pi}{n} \right), \quad i = 0, 1, \ldots, n.
\]
Moreover, $R_{nG}$ and $R_{nL}$ denote the corresponding quadrature errors and the double prime in the sum of Eq. (2) means (as usual) that the terms corresponding to $i = 0$ and $i = n$ should be halved.

Here the integrand $g(t)$ in Eqs. (1) and (2) is assumed to be a meromorphic function inside and on an ellipse $E_\rho$ with foci the points $\pm 1$ and semi-axes $[(\rho \pm (1/\rho))/2]$ with just one simple pole at a point $a$ inside $E_\rho$ with residue $A$. Then the major contribution to the error terms in Eqs. (1) and (2) is due to this pole and we can write [7, 9]

$$R_{nG} = A \frac{U_{n-1}^*(a)}{T_n(a)} + E_{nG},$$

$$R_{nL} = -A \frac{T_n^*(a)}{(1-a^2)U_{n-1}(a)} + E_{nL},$$

where $T_n(z)$ and $U_n(z)$ denote the Chebyshev polynomials of the first and the second kind, respectively, and degree $n$ and $T_n^*(z)$ and $U_n^*(z)$ denote the corresponding functions of the second kind. The explicit formulae for these functions are (see, e.g., [7])

$$T_n(z) = \frac{w^n+w^{-n}}{2},$$

$$U_n(z) = \frac{w^{n+1}-w^{-n-1}}{2\sqrt{z^2-1}},$$

$$T_n^*(z) = w^{-n},$$

$$U_n^*(z) = -w^{-n-1}\sqrt{z^2-1}$$

with

$$w = z + \sqrt{z^2-1}.$$  

Furthermore, $E_{nG}$ and $E_{nL}$ in Eqs. (5) and (6), respectively, denote the remaining parts of the error terms $R_{nG}$ and $R_{nL}$, respectively. These are given by [5, 6]

$$E_{nG} = -\frac{1}{2\pi i} \int_{E_\rho} \frac{U_{n-1}^*(z)}{T_n(z)} g(z) \, dz,$$

$$E_{nL} = \frac{1}{2\pi i} \int_{E_\rho} \frac{T_n^*(z)}{(1-z^2)U_{n-1}(z)} g(z) \, dz.$$  

For sufficiently large values of $n$, it is well known that [5, 6]

$$|E_{nG}|, |E_{nL}| \leq \frac{2M(\rho)}{\rho^{2n-1}},$$

where $M(\rho) = \max |g(z)|$ for $z \in E_\rho$. From the inequality (14) it is clear that $E_{nG}$ and $E_{nL}$ tend to zero very rapidly for sufficiently large values of $\rho$ as $n \to \infty$. From now on we will ignore these terms in Eqs. (1), (2), (5) and (6).

We define now the quantity

$$S_n = \sum_{j=0}^{2n} \sum (-1)^j g(x_j)$$

with

$$x_j = \cos \theta_j$$

and

$$\theta_j = \frac{j\pi}{2n}.$$
Then, by subtracting Eq. (1) from Eq. (2) and taking into account Eqs. (5) and (6), we obtain

\[ S_n \approx nA \left[ \frac{U_{n-1}(a)}{T_n(a)} + \frac{T_n(a)}{(1-a^2)U_{n-1}(a)} \right]. \]  

(17)

By taking into account Eqs. (7) to (10), we can rewrite Eq. (17) in the form

\[ S_n \approx -\frac{4nA}{(p^{2n} - p^{-2n})\sqrt{a^2 - 1}}, \]  

(18)

where \( p = a + \sqrt{a^2 - 1} \).  

(19)

By replacing \( n \) by \( 2n \) in Eqs. (1) and (2), quite similarly, we find that

\[ S_{2n} \approx -\frac{8nA}{(p^{4n} - p^{-4n})\sqrt{a^2 - 1}}. \]  

(20)

Equations (18) and (20) can be used for the approximate computation of the pole \( a \) of \( g(z) \) inside \( E_\rho \) and of the corresponding residue \( A \). Thus, by dividing Eqs. (18) and (20), we find that

\[ \frac{S_n}{S_{2n}} \approx \frac{1}{2} \left( \frac{p^{2n} + p^{-2n}}{p^{2n} - p^{-2n}} \right). \]  

(21)

We denote approximations to \( a, p \) and \( A \) by \( a_n, p_n \) and \( A_n \), respectively. Then, by solving Eq. (21) with respect to \( p_n \approx p \), we obtain

\[ p_n = \left[ s_n + \sqrt{s_n^2 - 1} \right]^{1/(2n)}, \]  

(22)

where

\[ s_n = \frac{S_n}{S_{2n}}. \]  

(23)

Then Eq. (19) yields

\[ a \approx a_n = \frac{1}{2} \left( p_n + \frac{1}{p_n} \right). \]  

(24)

If we are interested in \( A \) too, we determine it approximately, \( A_n \), from either Eq. (18) or Eq. (20). But generally, we are interested just in the pole \( a \) of \( g(z) \), determined approximately from Eq. (24), where \( p_n \) is computed from Eq. (22) and \( s_n \) from Eq. (23). Finally, we note that Eq. (24) can alternatively be written as

\[ a \approx a_n = \cosh \left( \frac{1}{2n} \cosh^{-1} s_n \right) \]  

(25)

as can easily be verified.

We have thus constructed a non-iterative algorithm for the approximate computation of the pole \( a \) of \( g(z) \) inside \( E_\rho \). Clearly, this method is also applicable to the approximate computation of zeros of analytic functions \( f(z) \) too provided that we use \( g(z) = 1/f(z) \) in our method. The computation of both a pole and a zero of a meromorphic function \( g(z) \) inside \( E_\rho \) is also possible by a double application of the previous method, i.e. both for \( g(z) \) and for \( f(z) = 1/g(z) \).

Now let us consider the convergence of the method for increasing values of \( n \). Clearly, the fundamental quadrature rules (1) and (2) converge as \( n \to \infty \) as is well known [5–7, 9] even if the contributions due to the pole \( a \) are ignored. This is clear from Eqs. (5), (6), (14), (17) and (18). More explicitly, the error terms \( E_nG \) and \( E_nL \), which were the only terms ignored and to which the approximate character of the present method is exclusively due, tend to zero almost like \( 1/\rho^{2n} \) for large values of \( n \). On the other hand, the right-hand side of Eq. (18), due to the pole \( a \) of \( g(z) \), tends to zero almost like \( 1/|p|^{2n} \) for large values of \( n \) and since \( |p| < \rho \) (because the pole \( a \) of \( g(z) \) was
assumed to lie inside the ellipse \(E_\rho\), the contribution of the error terms \(E_{nG}\) and \(E_{nL}\) to the accuracy of Eq. (18) tends to zero as \(n \to \infty\). Therefore, at least in theory, the previous method converges for \(n \to \infty\), that is, \(a_n \to a\) for \(n \to \infty\).

In practice, the situation is worse. This is exclusively due to round-off errors always present when using Eq. (15) for the computation of \(S_n\) and \(S_{2n}\) or, in another way of thinking, to the fact that the first terms of the right-hand sides of Eqs. (5) and (6), i.e. the contributions of the pole \(a\) to the error terms \(R_{nG}\) and \(R_{nL}\), tend to zero for \(n \to \infty\).

In any case, for low values of \(n\) (\(n = 1, 2\) or 3) the proposed method leads to sufficiently accurate numerical results with very few functional evaluations (5, 9 or 13, respectively). This will be verified in the numerical applications of the next two sections. Of course, if the pole \(a\) of \(g(z)\) lies too close to the integration interval \([-1,1]\) or if the corresponding residue \(A\) is large, then the accuracy of the method increases. Moreover, evidently, the numerical results obtained by the present method can be improved by the Newton–Raphson method or by some other appropriate local method [10].

We conclude this section with four remarks: (i) At first, it is possible by a simple change of variable of the form

\[x = \frac{1}{2}[(b-a)t + b + a]\]

(26)

to transform our fundamental interval \([-1,1]\) to a more general interval \([a,b]\). (ii) Secondly, if \(g(z)\) has more than one pole \(a\) (or \(f(z) = 1/g(z)\) has more than one zero \(a\)) inside \(E_\rho\), then the algorithm will converge to the pole (or zero) \(a_1\), which corresponds to the absolutely least value of \(p\) determined from Eq. (19). This is clear through an argumentation quite analogous to that used for the proof of the convergence of the method. Next, after the determination of \(a_1\), the algorithm can be used again with \(g(z)\) replaced by \((z-a_1)g(z)\) for the determination of the next pole \(a_2\) inside \(E_\rho\) and so on. Of course, it is assumed that no poles of \(g(z)\) have equal or almost equal (in absolute value) corresponding values of \(p\). (iii) Thirdly, as is clear from Eqs. (22) and (24), the present method gives \(n\) values \(a_n\) for each \(n\) and we have to select one of them. For small values of \(n\), this is easy by a direct substitution to \(f(z) = 1/g(z)\). For larger values of \(n\), the information about \(a\) available from previous values of \(n\) permits us to evaluate \(f(z)\) usually only two or three times in order to select the appropriate value of \(a_n\). (iv) Finally, as was already mentioned, round-off errors do not permit us to use large values of \(n\). Therefore, the accuracy of the computer used is of fundamental importance for the accuracy of the method and double-precision arithmetic is recommended.

### 3. Application to an equation of the theory of neutron moderation

As a first application of the method proposed in the previous section, we consider the transcendental equation

\[xe^x - be^b = 0,\]

(27)

where \(b\) is a parameter, appearing in the theory of neutron moderation in nuclear reactors. This equation was solved by Siewert and Burkart [16] as an extension of previous results by Siwerst and Burniston [17]. Further references are also reported in References [16] and [17].

For \(b \in (-1,0)\), Eq. (27) possesses not only the obvious root \(x = b\), but also a second root \(a \in (-\infty,-1)\). We will determine this root by using our method. Clearly, the function \(g(z)\) should be defined as

\[g(z) = \frac{z-b}{ze^z - be^b},\]

(28)

where the numerator aims simply at reducing the number of poles of \(g(z)\) to one: a single pole \(a\).

In Table 1, we display the numerical results for the root \(a\) of Eq. (27) for several values of the parameter \(b\) (as these numerical results were obtained by a computer working in ten significant
Table 1: Numerical results for the nontrivial root \( a \) of the transcendental equation \( xe^x - be^b = 0 \) for several values of the parameter \( b \), resulted by the method of Section 2 for \( n = 1,2,\ldots,5 \) as well as by the Newton–Raphson method: \( a_{ex} \) with this symbol denoting the “exact” value of the root \( a \). (The accuracy of the computer was ten significant digits.)

|   | \( n \) | \( b = -0.4 \) | \( b = -0.6 \) | \( b = -0.8 \) | \( b = -0.9 \) | \( b = -0.95 \) | \( b = -0.99 \) | \( b = -0.999 \) |
|---|---|---|---|---|---|---|---|
| 1 | -1.9357942 | -1.5233929 | -1.2253530 | -1.1053129 | -1.0509826 | -1.0099429 | -1.0009887 |   |
| 2 | -2.0188109 | -1.5474246 | -1.2308449 | -1.1071471 | -1.0517245 | -1.0100672 | -1.0010006 |   |
| 3 | -2.0187817 | -1.5474089 | -1.2308421 | -1.1071465 | -1.0517243 | -1.0100672 | -1.0010006 |   |
| 4 | -1.9951696 | -1.5472311 | -1.2308407 | -1.1071459 | -1.0517244 | -1.0100672 | -1.0010006 |   |
| 5 | -2.0428886 | -1.5432273 | -1.2308093 | -1.1071469 | -1.0517240 | -1.0100673 | -1.0010006 |   |

\( a_{ex} = -2.0187881 \) -1.5474049 -1.2308422 -1.1071465 -1.0517243 -1.0100671 -1.0010006

digits) for \( n = 1,2,\ldots,5 \). In the same table, we also present the “exact” values of the root \( a \) as these values were obtained by the Newton–Raphson method with an original approximation to \( a \) the value resulted by the method of Section 2 for \( n = 2 \). For the transcendental Eq. (25), the Newton–Raphson method takes the following form

\[
a_{k+1} = a_k - \frac{a_k - be^b - a_k}{a_k + 1}.
\] (29)

From the numerical results of Table 1 we observe that the most accurate results are obtained for \( n = 3 \). In this case, with thirteen functional evaluations, at least six correct significant digits were obtained. This is a very satisfactory result. For larger values of \( n \), the contribution of round-off errors in the computer does not permit any increase of the accuracy of the numerical results. Of course, this would be possible if a computer working with more than ten significant digits were used.

On the other hand, from Table 1 we also observe that the numerical results for the root \( a \) become more accurate as \( b \rightarrow -1 \). This is due to the fact that in this case \( a \rightarrow -1 \) too, that is the pole \( a \) of \( g(z) \), defined by Eq. (28), approaches the integration interval \([-1,1]\) and, therefore, the error term due to this pole becomes significant tending to infinity as \( a \rightarrow -1 \). For this reason, the influence of round-off errors becomes also negligible (even for \( n = 5 \)) as is clear from Table 1 for \( |b| \geq 0.9 \). In any case, it is understood that the accuracy of the operations inside the computer strongly influences the numerical results. Now we proceed to a second application of the method.

### 4. Application to an equation of the theory of elastic buckling

As a second application of the method proposed in Section 2, we consider the transcendental equation

\[
xtan x - b = 0
\] (30)

appearing in the theory of elastic buckling for a frame consisting of three bars (two vertical and one horizontal) and loaded by two loads. The details of the problem and the derivation of Eq. (30) can be found in the classical book of Timoshenko and Gere [18, p. 66] and we omit them here. On the
Table 2: Analogous results to those of Table 1, but now for the smallest positive root of the transcendental equation \( x \tan x - b = 0 \).

<table>
<thead>
<tr>
<th></th>
<th>( b = 10 )</th>
<th>( b = 5 )</th>
<th>( b = 3 )</th>
<th>( b = 2 )</th>
<th>( b = 1.7 )</th>
<th>( b = 1.6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.4667638</td>
<td>1.3321021</td>
<td>1.1993695</td>
<td>1.0785047</td>
<td>1.0276950</td>
<td>1.0085457</td>
</tr>
<tr>
<td>2</td>
<td>1.4301260</td>
<td>1.3144184</td>
<td>1.1926387</td>
<td>1.0769022</td>
<td>1.0272535</td>
<td>1.0084225</td>
</tr>
<tr>
<td>3</td>
<td>1.4289347</td>
<td>1.3138651</td>
<td>1.1924658</td>
<td>1.0768747</td>
<td>1.0272400</td>
<td>1.0084212</td>
</tr>
<tr>
<td>4</td>
<td>1.4288741</td>
<td>1.3138392</td>
<td>1.1924591</td>
<td>1.0768740</td>
<td>1.0272478</td>
<td>1.0084212</td>
</tr>
<tr>
<td>5</td>
<td>1.4287624</td>
<td>1.3138279</td>
<td>1.1924583</td>
<td>1.0768740</td>
<td>1.0272478</td>
<td>1.0084212</td>
</tr>
<tr>
<td></td>
<td>( a_{ex} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.4288700</td>
<td>1.3138377</td>
<td>1.1924588</td>
<td>1.0768740</td>
<td>1.0272478</td>
<td>1.0084212</td>
</tr>
</tbody>
</table>

other hand, Eq. (30) is a more or less classical transcendental equation. Burniston and Siewert [3] solved this equation by their method and tables of the roots of the same equation are also available.

Here we will test our method by solving Eq. (30). The peculiarity of this equation is that it has an infinity of positive and negative roots \( \pm a_i \) with \( i = 0, 1, \ldots \). But, as was already mentioned in Section 2, our method will converge to the root with the smallest corresponding absolute value of \( p \), defined by Eq. (19). In our numerical application, this means that our method will give the absolutely smallest root \( a_0 \) of Eq. (30). The next roots of this equation, \( a_1, a_2, \ldots \), can be found by successive further applications of the same method. But, as is well known [18], the critical buckling loads depend only on the smallest positive root of the corresponding transcendental equation, that is, we are interested only in \( a_0 \). Moreover, it is not difficult to see that, although we are essentially seeking simultaneously two roots of Eq. (30), the roots \( a_0 \) and \( -a_0 \), the algorithm of Section 2 remains unchanged since these two roots are opposite numbers.

In Table 2 we present the numerical results for the smallest positive root \( a_0 \) of Eq. (30) for several values of the parameter \( b \) exactly as has been already the case in the application of the previous section, but now, evidently, for the equation

\[
 g(z) = \frac{1}{z \tan z - b}, \tag{31}
\]

Moreover, the iterative formula for the Newton–Raphson method was

\[
a_{0,k+1} = a_{0,k} - \frac{a_{0,k} \tan a_{0,k} - b}{a_{0,k} (1 + \tan^2 a_{0,k}) + \tan a_{0,k}}. \tag{32}
\]

The best numerical results were obtained for \( n = 4 \) and \( n = 5 \). For \( n = 4 \) (seventeen functional evaluations) at least six correct significant digits were obtained in all cases, whereas for \( n = 3 \) (thirteen functional evaluations) at least five correct significant digits were obtained in all cases. Several further remarks analogous to those made in the previous section can also be made for the numerical results of Table 2.

5. Discussion–conclusions

The method proposed in this technical report has several advantages and disadvantages when it is compared with competitive methods. We mention some of them.
Among its advantages over several (but not all) competitive methods, we can mention: (i) The fact that it is able to find real or complex zeros of analytic functions in the complex plane (and not only on a finite interval) without previous information on these zeros as is necessary in other methods, e.g. in the Newton–Raphson method. (ii) The very rapid convergence of the method for very low values of \( n \) \((n \leq 4)\) and the sufficiently accurate determination of zeros of analytic functions with very few functional evaluations. (iii) The simplicity of the method as an algorithm and the fact that it does not require computations of derivatives of the original analytic function. (iv) The very simple theoretical background of the method.

Among the disadvantages of the method over some (but not all) competitive methods, we can mention: (i) The fact that it is strongly influenced by round-off errors in the computer and, therefore, double-precision arithmetic is several times necessary. (ii) The fact that the zeros of an analytic function are not determined simultaneously but one after the other. (iii) The fact that the appropriate root of the quantity \( p_n \) in Eq. (22) should be selected essentially by inspection or by testing these roots. (iv) The lack of an elegant theory for the derivation of the method contrary to what happens in other quadrature methods [1, 3, 8], where the results of the theory of analytic functions are used.

Concluding, we believe that the proposed method can be considered as a good method for the original localization of zeros of analytic functions and in several cases without the need to improve the obtained numerical results by local methods. We also believe that it is the simplest method based on numerical integration rules (quadrature method) for the computation of zeros of analytic functions both as its theory and its corresponding algorithm are concerned. Of course, modifications and generalizations of the method are also possible. One such modification is the change of the fundamental weight function \( 1/\sqrt{1-t^2} \) used in the quadrature rules (1) and (2).

Acknowledgements The results reported here belong to a research project supported by the National Hellenic Research Foundation. The author gratefully acknowledges the financial support of this Foundation.

References


3All the links (external links in blue) in this section were added by the author on 21 January 2018 for the online publication of this technical report. Moreover, in Reference [11] final publication details were also added.


