

IMPROVEMENT OF CONVERGENCE CONDITION OF THE SQUARE-ROOT INTERVAL METHOD FOR MULTIPLE ZEROS¹

Miodrag S. Petković², Dušan M. Milošević³

Abstract. A new theorem concerned with the convergence of the Ostrowski-like method for the simultaneous inclusion of multiple complex zeros in circular complex arithmetic is established. Computationally verifiable initial condition that guarantees the convergence of this parallel inclusion method is significantly relaxed compared with the classical theorem stated in [Z. Angew. Math. Mech. 62 (1982), 627–630].

AMS Mathematics Subject Classification (2000): 65H05, 65G20, 30C15

Key words and phrases: Polynomial zeros, convergence, simultaneous methods, inclusion methods, circular interval arithmetic

1. Introduction

Iterative methods for the simultaneous inclusion of polynomial zeros, realized in interval arithmetic, produce resulting real or complex intervals (disks or rectangles) that contain the sought zeros. In this manner, information about the upper error bounds of approximations to the zeros is provided. Besides, there exists the ability to incorporate rounding errors without altering the fundamental structure of the interval method. An extensive study and history of interval methods for solving algebraic equations may be found in the books [9] and [10].

The purpose of this paper is to present the improved initial convergence condition of the square-root inclusion method proposed in [8]. A similar problem was considered in the recent paper [11], where the initial condition for the convergence of the third-order Newton-like inclusion method, presented in [4], is relaxed.

The presentation of the paper is organized as follows. Some basic definitions and operations of circular complex interval arithmetic, necessary for the convergence analysis and the construction of inclusion methods, are given in section 1. The derivation of the square-root inclusion method and the criterion for the choice of a proper square root of a disk are presented in section 2. The convergence analysis of the considered method under the relaxed initial condition is given in section 3.

¹This research was supported by the Serbian Ministry of Science and Environmental Protection under grant number 144024

²Faculty of Electronic Engineering, University of Niš, A. Medvedeva 14, 18000 Niš, Serbia, msp@junis.ni.ac.yu

³Faculty of Electronic Engineering, University of Niš, A. Medvedeva 14, 18000 Niš, Serbia, dmilosev@elfak.ni.ac.yu

The construction of the inclusion method and its convergence analysis, presented in this paper, need the basic properties of the so-called circular complex arithmetic introduced by Gargantini and Henrici [4]. A circular closed region (disk) $Z := \{z : |z - c| \leq r\}$ with center $c := \text{mid } Z$ and radius $r := \text{rad } Z$ we denote by parametric notation $Z := \{c; r\}$. The following is valid:

$$\begin{aligned}\alpha\{c; r\} &= \{\alpha c; |\alpha|r\} \quad (\alpha \in \mathbb{C}), \\ \{c_1; r_1\} \pm \{c_2; r_2\} &= \{c_1 \pm c_2; r_1 + r_2\}.\end{aligned}$$

The inversion of a non-zero disk Z is defined by the Möbius transformation,

$$(1) \quad Z^{-1} = \left\{ \frac{\bar{c}}{|c|^2 - r^2}; \frac{r}{|c|^2 - r^2} \right\} \quad (|c| > r, \text{ i.e., } 0 \notin Z).$$

The addition, subtraction and inversion Z^{-1} are exact operations.

The set $\{z_1 z_2 : z_1 \in Z_1, z_2 \in Z_2\}$, in general, is not a disk. In order to remain within the realm of disks, Gargantini and Henrici [4] introduced the multiplication by

$$Z_1 \cdot Z_2 := \{c_1 c_2; |c_1|r_2 + |c_2|r_1 + r_1 r_2\} \supseteq \{z_1 z_2 : z_1 \in Z_1, z_2 \in Z_2\}.$$

Then the division is defined by

$$Z_1 : Z_2 = Z_1 \cdot Z_2^{-1}.$$

The square root of a disk $\{c; r\}$ that does not contain the origin, where $c = |c|e^{i\theta}$ and $|c| > r$, is defined as the union of two disjoint disks (see [2]):

$$(2) \quad \{c; r\}^{1/2} := \left\{ \sqrt{|c|}e^{i\theta/2}; \frac{r}{\sqrt{|c|} + \sqrt{|c| - r}} \right\} \cup \left\{ -\sqrt{|c|}e^{i\theta/2}; \frac{r}{\sqrt{|c|} + \sqrt{|c| - r}} \right\}.$$

In this paper we will use the following obvious properties:

$$(3) \quad z \in \{c; r\} \iff |z - c| \leq r,$$

$$(4) \quad \{c_1; r_1\} \cap \{c_2; r_2\} = \emptyset \iff |c_1 - c_2| > r_1 + r_2.$$

More details about circular arithmetic can be found in the books [1], [9] and [10].

2. Ostrowski-like method

Let P be a monic polynomial of degree $N \geq 3$

$$(5) \quad P(z) = \prod_{j=1}^n (z - \zeta_j)^{\mu_j}$$

with $n (\leq N)$ distinct real or complex zeros ζ_1, \dots, ζ_n of respective multiplicities μ_1, \dots, μ_n , where $\mu_1 + \dots + \mu_n = N$ and let

$$\delta_2(z) = \frac{P'(z)^2 - P(z)P''(z)}{P(z)^2}.$$

From the factorization of (5) we find

$$\delta_2(z) = -\frac{d^2}{dz^2}(\log P(z)) = \sum_{j=1}^n \frac{\mu_j}{(z - \zeta_j)^2} = \frac{\mu_i}{(z - \zeta_i)^2} + \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j}{(z - \zeta_j)^2}.$$

Solving the last equation in ζ_i we obtain the following fixed-point relation

$$(6) \quad \zeta_i = z - \frac{\sqrt{\mu_i}}{\left[\delta_2(z) - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j}{(z - \zeta_j)^2} \right]_*^{1/2}}.$$

It is assumed that only one complex value (of two) of the square root has to be taken in the last formula, which is indicated by the symbol *. This value is chosen in such a way that the right-hand side reduces to ζ_i .

Let $\mathcal{I}_n := \{1, \dots, n\}$ be the index set and suppose that n disjoint disks Z_1, \dots, Z_n such that $\zeta_j \in Z_j$ ($j \in \mathcal{I}_n$) have been found. Let us put $z = z_i = \text{mid } Z_i$ in (6). Since $\zeta_j \in Z_j$ ($j \in \mathcal{I}_n$), according to the inclusion isotonicity property we obtain

$$(7) \quad \zeta_i \in z_i - \frac{\sqrt{\mu_i}}{\left[\delta_2(z_i) - \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{z_i - Z_j} \right)^2 \right]_*^{1/2}} \quad (i \in \mathcal{I}_n).$$

Remark 1. We write $\left(\frac{1}{z_i - Z_j} \right)^2$ rather than $\frac{1}{(z_i - Z_j)^2}$ since $\text{rad} \left(\frac{1}{Z} \right)^2 \leq \text{rad} \frac{1}{Z^2}$ ($0 \notin Z$), see [6].

Let $Z_1^{(0)}, \dots, Z_n^{(0)}$ be initial disjoint disks containing the zeros ζ_1, \dots, ζ_n , that is, $\zeta_i \in Z_i^{(0)}$ for all $i \in \mathcal{I}_n$. The relation (7) suggests the following method for the simultaneous inclusion of all multiple zeros of P :

$$(8) \quad Z_i^{(m+1)} = z_i^{(m)} - \frac{\sqrt{\mu_i}}{\left[\delta_2(z_i^{(m)}) - \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{z_i^{(m)} - Z_j^{(m)}} \right)^2 \right]_*^{1/2}},$$

($i \in \mathcal{I}_n$; $m = 0, 1, \dots$). Assuming that the denominator does not contain the number 0, according to (3) there follows that the square root of a disk gives two disks. Since these disks are disjoint, only one of them gives a circular outer approximation that contains the exact zero ζ_i . The choice of this “proper” disk is indicated by the symbol *. The criterion for the choice of a proper disk is similar to that considered in [2] and reads:

Let

$$\left[\delta_2(z_i^{(m)}) - \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{z_i^{(m)} - Z_j^{(m)}} \right)^2 \right]_*^{1/2} = D_{1,i}^{(m)} \cup D_{2,i}^{(m)},$$

where $D_{1,i}^{(m)}$ and $D_{2,i}^{(m)}$ are determined according to (2). Among the disks $D_{1,i}^{(m)}$ and $D_{2,i}^{(m)}$ one has to choose that disk whose center minimizes

$$|P'(z_i^{(m)})/(\mu_i P(z_i^{(m)})) - \text{mid } D_{k,i}^{(m)}| \quad (k = 1, 2).$$

For the iteration index m let us introduce the abbreviations

$$\begin{aligned} r^{(m)} &= \max_{1 \leq i \leq n} r_i^{(m)}, & \mu &= \min_{1 \leq j \leq n} \mu_j \\ \rho^{(m)} &= \min_{\substack{1 \leq i, j \leq n \\ i \neq j}} \{|z_i^{(m)} - z_j^{(m)}| - r_j^{(m)}\}, \\ z_i^{(m)} &= \text{mid } Z_i^{(m)}, & r_i^{(m)} &= \text{rad } Z_i^{(m)}. \end{aligned}$$

For simplicity, we will omit sometimes the iteration index.

Remark 2. The iterative method (8) was proposed by M. Petković in [8]. It was proved that the order of convergence is equal four under the initial condition

$$(9) \quad \rho^{(0)} > 3(N - \mu)r^{(0)}.$$

The main goal of this paper is to improve this condition, that is, to find a multiplier significantly smaller than $3(N - \mu)$.

Remark 3. Omitting the sum in the iterative formula (8) we obtain the Ostrowski iterative formula

$$z^{(m+1)} = z^{(m)} - \frac{\sqrt{\mu_i}}{\left[\delta_2(z^{(m)}) \right]_*^{1/2}}$$

with the cubic convergence, extensively studied by Ostrowski [5]. For this reason, the inclusion method (8) is referred to as *Ostrowski-like* method.

3. Convergence analysis

In this section we give the convergence analysis of the interval method (8). In the sequel we will always assume that $N \geq 3$.

Lemma 1. *Let the inequality*

$$(10) \quad \rho > 2\sqrt{N - \mu} r$$

hold. Then

$$(i) \quad \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left| (z_i - \zeta_j)^{-2} - (z_i - z_j)^{-2} \right| \leq \frac{2(N - \mu)r}{\rho^3};$$

$$(ii) \quad \left| \delta_2(z_i) - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j}{(z_i - z_j)^2} \right| > \frac{4\mu}{5r^2}.$$

Proof. Of (i): Since

$$|z_i - \zeta_j| \geq |z_i - z_j| - |z_j - \zeta_j| \geq |z_i - z_j| - r_j \geq \rho,$$

we conclude that

$$(11) \quad \frac{1}{|z_i - \zeta_j|} \leq \frac{1}{\rho} \quad \text{and} \quad \frac{1}{|z_i - z_j|} \leq \frac{1}{\rho}.$$

Using (10) and (11) we estimate

$$\begin{aligned} \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left| (z_i - \zeta_j)^{-2} - (z_i - z_j)^{-2} \right| &= \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j |(z_i - z_j)^2 - (z_i - \zeta_j)^2|}{|z_i - \zeta_j|^2 |z_i - z_j|^2} \\ &= \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j |\zeta_j - z_j| (|z_i - z_j + z_i - \zeta_j|)}{|z_i - \zeta_j|^2 |z_i - z_j|^2} \leq r \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j (|z_i - z_j| + |z_i - \zeta_j|)}{|z_i - \zeta_j|^2 |z_i - z_j|^2} \\ &\leq r \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{|z_i - \zeta_j|^2 |z_i - z_j|} + \frac{1}{|z_i - \zeta_j| |z_i - z_j|^2} \right) \leq \frac{2(N - \mu)r}{\rho^3}. \end{aligned}$$

Of (ii): Using the inequality (10) and the assertion (i) of Lemma 1 we obtain

$$\begin{aligned} \left| \delta_2(z_i) - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j}{(z_i - z_j)^2} \right| &\geq \frac{\mu_i}{|z_i - \zeta_i|^2} - \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left| (z_i - \zeta_j)^{-2} - (z_i - z_j)^{-2} \right| \\ &\geq \frac{\mu}{r^2} - \frac{2(N - \mu)r}{\rho^3} > \frac{\mu}{r^2} \left(1 - \frac{1}{4\mu\sqrt{N - \mu}} \right) \\ &> \frac{4\mu}{5r^2}. \quad \square \end{aligned}$$

Now we state the convergence theorem of the Ostrowski-like method (8) under the relaxed initial condition of the form (10).

Theorem 1. *Let be given the initial disjoint disks $Z_1^{(0)}, \dots, Z_n^{(0)}$ containing the zeros ζ_1, \dots, ζ_n of the polynomial P and let the interval sequences $\{Z_i^{(m)}\}$ ($i \in \mathcal{I}_n$) be defined by the iterative formula (8). Then, under the condition*

$$(12) \quad \rho^{(0)} > 2\sqrt{N - \mu} r^{(0)},$$

for each $i \in \mathcal{I}_n$ and $m = 0, 1, \dots$ we have

$$\begin{aligned}
1^\circ \quad & \zeta_i \in Z_i^{(m)}; \\
2^\circ \quad & r^{(m+1)} < \frac{8(N-\mu)(r^{(m)})^4}{5\mu\left(\rho^{(0)} - \frac{5}{3}r^{(0)}\right)^3}.
\end{aligned}$$

Proof. Of (i): We will prove the assertion 1° by induction. Suppose that $\zeta_i \in Z_i^{(m)}$ for $i \in \mathcal{I}_n$ and $m \geq 1$. Then

$$z_i^{(m)} - \frac{\sqrt{\mu_i}}{\left[\delta_2(z_i^{(m)}) - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j}{(z_i^{(m)} - \zeta_j)^2}\right]^*} \equiv \zeta_i,$$

where the symbol $*$ denotes the (complex) value of the square root equal to $\sqrt{\mu_i}/(z_i^{(m)} - \zeta_i)$. Since

$$\sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{z_i^{(m)} - \zeta_j}\right)^2 \in \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{z_i^{(m)} - Z_j^{(m)}}\right)^2,$$

from (8) one obtains $\zeta_i \in Z_i^{(m+1)}$. Since $\zeta_i \in Z_i^{(0)}$, the assertion 1° follows by mathematical induction.

Let us prove now that the interval method (8) has the order of convergence equal to four (assertion 2°). We use induction and start with $m = 0$. For simplicity, all indices are omitted and all quantities in the first iteration are denoted by $\hat{}$.

We use the following inclusion derived in [7]

$$(13) \quad \left(\frac{1}{z_i - Z_j}\right)^k \subset \left\{ \frac{1}{(z_i - z_j)^k}; \frac{kr}{\rho^{k+1}} \right\} \quad (k = 1, 2, \dots).$$

According to the inclusion (13) (for $k = 2$) we obtain

$$(14) \quad \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j \left(\frac{1}{z_i - Z_j}\right)^2 \subset \left\{ \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mu_j}{(z_i - z_j)^2}; \frac{2(N-\mu)r}{\rho^3} \right\} =: \{c_i; \eta\}.$$

Let $u_i = \delta_2(z_i) - c_i$. Then, using (1), (2) and (14), from (8) we obtain

$$\begin{aligned}
\hat{r}_i &= \text{rad } \hat{Z}_i \leq \text{rad} \left(\frac{\sqrt{\mu_i}}{\{u_i; \eta\}^{1/2}} \right) \leq \sqrt{\mu_i} \text{rad} \left(\frac{1}{\{u_i; \eta\}} \right)^{1/2} \\
&= \sqrt{\mu_i} \text{rad} \left\{ \frac{\bar{u}_i}{|u_i|^2 - \eta^2}; \frac{\eta}{|u_i|^2 - \eta^2} \right\}^{1/2} \\
(15) \quad &= \frac{\sqrt{\mu_i} \eta}{(|u_i|^2 - \eta^2)^{1/2} (|u_i|^{1/2} + (|u_i| - \eta)^{1/2})}.
\end{aligned}$$

Here we have used the inequality $\text{rad} \frac{1}{Z^{1/2}} \leq \text{rad} \left(\frac{1}{Z} \right)^{1/2}$ ($0 \notin Z$) proved in [6].

By virtue of Lemma 1 and the inequality (12) we estimate

$$\eta = \frac{2(N - \mu)r}{\rho^3} < \frac{1}{4\sqrt{N - \mu}} \cdot \frac{1}{r^2} < \frac{\mu}{5r^2}$$

and

$$|u_i| - \eta > \frac{4\mu}{5r^2} - \frac{\mu}{5r^2} = \frac{3\mu}{5r^2}.$$

Using the last two inequalities and Lemma 1 (ii), we obtain from (15)

$$\hat{r}_i < \frac{\frac{2\sqrt{\mu}(N - \mu)r^4}{\rho^3}}{\mu^{3/2} \left(\left(\frac{4}{5} \right)^2 - \left(\frac{1}{5} \right)^2 \right)^{1/2} \left(\sqrt{\frac{4}{5}} + \sqrt{\frac{3}{5}} \right)} < \frac{8}{5} \frac{N - \mu}{\mu} \frac{r^4}{\rho^3}$$

and, using (12),

$$(16) \quad \hat{r} < \frac{1}{7}r.$$

According to a geometric construction and the fact that the disks $Z_i^{(m)}$ and $Z_i^{(m+1)}$ must have at least one common point (the zero ζ_i), the following relation can be derived (see [3]):

$$(17) \quad \rho^{(m+1)} \geq \rho^{(m)} - r^{(m)} - 3r^{(m+1)}.$$

Using the inequalities (16) and (17) (for $m = 0$), we find

$$\begin{aligned} \rho^{(1)} &\geq \rho^{(0)} - r^{(0)} - 3r^{(1)} > 2\sqrt{N - \mu} r^{(0)} - r^{(0)} - \frac{3}{7}r^{(0)} \\ &> 7r^{(1)} \left(2\sqrt{N - \mu} - 1 - \frac{3}{7} \right), \end{aligned}$$

wherefrom it follows

$$(18) \quad \rho^{(1)} > 2\sqrt{N - \mu} r^{(1)}.$$

This is the condition (10) for the index $m = 1$, which means that all assertions of Lemma 1 are valid for $m = 1$.

Using the definition of ρ and (18), for arbitrary pair of indices $i, j \in \mathcal{I}_n$ ($i \neq j$) we have

$$(19) \quad |z_i^{(1)} - z_j^{(1)}| \geq \rho^{(1)} > 2\sqrt{N - \mu} r^{(1)} \geq 2r^{(1)} \geq r_i^{(1)} + r_j^{(1)}.$$

Therefore, in regard to (4), the disks $Z_1^{(1)}, \dots, Z_n^{(1)}$, produced by (8), are disjoint.

Applying mathematical induction with the argumentation used for the derivation of (16), (18) and (19) (which makes the part of the proof with respect to

$m = 1$), we prove that the disks $Z_1^{(m)}, \dots, Z_n^{(m)}$ are disjoint for each $m = 0, 1, \dots$, and the following relations are true:

$$(20) \quad r^{(m+1)} < \frac{8(N-\mu)(r^{(m)})^4}{5\mu(\rho^{(m)})^3},$$

$$(21) \quad r^{(m+1)} < \frac{1}{7}r^{(m)},$$

$$(22) \quad \rho^{(m)} > 2\sqrt{N-\mu}r^{(m)}.$$

In addition, we note that the last inequality (22) means that the assertions of Lemma 1 hold for each $m = 0, 1, 2, \dots$. Finally, from (21) we conclude that the sequence $\{r^{(m)}\}$ monotonically converges to 0, in other words, the Ostrowski-like inclusion method (8) is convergent under the initial condition (12).

Setting $\omega = 1/7$ we find

$$(23) \quad 1 + 4(\omega + \omega^2 + \dots + \omega^m) - \omega^m < 1 + \frac{4\omega}{1-\omega} = \frac{5}{3}.$$

By the successive application of (17) and (21) we obtain

$$\begin{aligned} \rho^{(m)} &> \rho^{(m-1)} - r^{(m-1)} - 3\omega r^{(m-1)} = \rho^{(m-1)} - r^{(m-1)}(1 + 3\omega) \\ &> \rho^{(m-2)} - r^{(m-2)} - 3\omega r^{(m-2)} - \omega r^{(m-2)}(1 + 3\omega) \\ &= \rho^{(m-2)} - r^{(m-2)}(1 + 4\omega + 4\omega^2 - \omega^2) \\ &\vdots \\ &> \rho^{(0)} - r^{(0)}(1 + 4\omega + 4\omega^2 + \dots + 4\omega^m - \omega^m) \\ &> \rho^{(0)} - \frac{5}{3}r^{(0)}, \end{aligned}$$

where we used (23). According to the last inequality and (20) we find

$$r^{(m+1)} < \frac{8(N-\mu)(r^{(m)})^4}{5\mu\left(\rho^{(0)} - \frac{5}{3}r^{(0)}\right)^3}.$$

Therefore, the assertion 2° of Theorem 1 holds. The last relation shows that the order of convergence of the inclusion method (8) is four. \square

We conclude this paper with the remark that the initial condition (12) is significantly weakened compared to (9), see Remark 2. The ratio $R(N, \mu) = 3(N-\mu)/(2\sqrt{N-\mu}) = 1.5\sqrt{N-\mu}$ of multipliers appearing in (9) and (12), for $\mu = 1, 2, 3$ and $N \geq 4$, is given in Fig. 1.

Fig. 1 Ratio of multipliers

References

- [1] Alefeld, G., Herzberger, J., Introduction to Interval Computations. New York: Academic Press 1983.
- [2] Gargantini, I., Parallel Laguerre iterations: The complex case. Numer. Math. 26 (1976), 317–323.
- [3] Gargantini, I., Further application of circular arithmetic: Schröder-like algorithms with error bound for finding zeros of polynomials. SIAM J. Numer. Anal. 15 (1978), 497–510.
- [4] Gargantini, I., Henrici, P., Circular arithmetic and the determination of polynomial zeros. Numer. Math. 18 (1972), 305–320.
- [5] Ostrowski, A., Solution of Equations and Systems of Equations. New York: Academic Press 1966.
- [6] Petković, Lj. D., A note on the evaluation in circular arithmetic. Z. Angew. Math. Mech. 66 (1986), 371–373.
- [7] Petković, M. S., On a generalization of the root iterations for polynomial complex zeros in circular interval arithmetic. Computing 27 (1981), 37–55.
- [8] Petković, M. S., Generalized root iterations for simultaneous determination of multiple complex zeros. Z. Angew. Math. Mech. 62 (1982), 627–630.
- [9] Petković, M. S., Iterative Methods for Simultaneous Inclusion of Polynomial Zeros. Berlin-Heidelberg-New York: Springer-Verlag 1989.
- [10] Petković, M. S., Petković, Lj. D., Complex Interval Arithmetic and its Applications. Berlin-Weinheim-New York: Wiley-VCH 1998.
- [11] Zhu, L., On the convergent condition of Newton-like method in parallel circular iteration for simultaneously finding all multiple zeros of a polynomial. Appl. Math. Comp. 168 (2005), 677–685.

Received by the editors May 17, 2006