# Serdica

## Bulgariacae mathematicae publicationes

# Сердика

## Българско математическо списание

*SERDICA* — Bulgaricae
mathematicae publicationes
**20** (1994) 148-157

# ON ROOT FINDING METHODS WITH CORRECTIONS

MIODRAG S. PETKOVIĆ*, CARSTEN CARSTENSEN

ABSTRACT. The paper is concerned with very fast methods for the simultaneous computations of polynomial roots which are obtained by corrections. The R–orders of these methods are generally calculated and discussed from a practical point of view. As an example a class of iterative methods of Halley's type is given.

**1. Introduction.** The purpose of this paper is to present some theoretical as well as practical results on iteration methods for the simultaneous improvement of polynomial zeros. First we give a general theory concerning the construction of very fast methods and their $R$-order of convergence. A general approach which uses **one** and/or **two** corrections is presented. The computational efficiency of these algorithms is pretty high since a great increase of the order of convergence is obtained without additional calculations. From the point of view of the actual research and the current availability of parallel computers we notice that algorithms implemented in a total-step mode can be realized only. This indicates a disadvantage of the single-step methods which are sequential in nature.

In §2 we give a general frame for iteration functions with one and two corrections. The R–order of convergence is determined in §3. The results are discussed in §4. An example of Halley like methods is given in §5. Three new iterative formulas of higher order of convergence are presented.

We emphasize that our notation covers only complex point arithmetic and the ideas and results can be applied to interval arithmetic as well.

**2. Iteration methods with corrections.** Let $\hat{z} = z - \varphi(z), \varphi : \mathbb{C} \mapsto \mathbb{C}$, define an iteration method of the order of convergence $s$ for finding a complex zero $\zeta$ of a given polynomial. Here $\hat{z}$ is an improved approximation to the former approximation $z$ of the zero $\zeta$ and

$$|\hat{z} - \zeta| \leq \gamma |z - \zeta|^s$$

*This paper was partially supported by the Science Fund of Serbia under Grant No. 0401.

holds for some positive constant $\gamma$ (provided $|z-\zeta|$ is sufficiently small). The complex function $\varphi$ will be called the *correction function of order s*.

Assume that correction functions $\varphi_i$ $(i=1,\ldots,n)$ have order $p+1\geq 2$ and let us consider a class of iteration formulas

$$\text{(T)}\qquad \hat{z}_i = z_i - \varphi_i(z_1,\ldots,z_n) \quad (i=1,\ldots,n)$$

for the simultaneous improvement of approximations $z_1,\ldots,z_n$ to the zeros $\zeta_1,\ldots,\zeta_n$ of a given polynomial of degree $n$. The formula (T), where only former approximations are used in calculation of new approximations $\hat{z}_i$, defines the so-called *total step* method, or TS method.

Let $h_i$ and $\hat{h}_i$ be multiples of $|z_i-\zeta_i|$ and $|\hat{z}_i-\zeta_i|$ respectively, $i=1,\ldots,n$. Then for a rather wide class of simultaneous methods of the form (T) it is possible to derive the following *h-relations* [6]

$$\text{(1)}\qquad \hat{h}_i \leq \frac{1}{n-1} h_i^p \sum_{j=1,j\neq i}^{n} h_j \quad (i=1,\ldots,n).$$

Through the paper we will assume that all quantities $h_1,\ldots,h_n$ are of the same order of size. According to this fact the $p+1$ order of convergence of the TS method (T) follows from (1).

In some cases the convergence of the TS method (T) can be accelerated using a suitable correction function $\vartheta(z)$ of order $q>1$ (say) as follows

$$\text{(Tc)}\qquad \begin{aligned}\hat{z}_i &= z_i - \varphi_i(z_1-\vartheta(z_1),\ldots,z_{i-1}-\vartheta(z_{i-1}),z_i,\\ &\quad z_{i+1}-\vartheta(z_{i+1}),\ldots,z_n-\vartheta(z_n)) \quad (i=1,\ldots,n).\end{aligned}$$

The notation (Tc) means that the total step method uses *one* correction. In particular cases, for a wide class of methods of the form (Tc), we can derive the $h$-relations

$$\text{(2)}\qquad \hat{h}_i \leq \frac{1}{n-1} h_i^p \sum_{j=1,j\neq i}^{n} h_j^q \quad (i=1,\ldots,n)$$

associated to (Tc). This means that calculations with the improved approximations $z_j-\vartheta(z_j)$ instead of $z_j$ provide the increased order of convergence $p+q$.

The order of convergence of the TS method (T) (without corrections) and (Tc) (with one correction) can be further increased if, calculating $\hat{z}_i$ $(i>1)$, we use the already found new approximations $\hat{z}_1,\ldots,\hat{z}_{i-1}$ (the so-called Gauss-Seidel approach). Thus we obtain the *single step* or shorter, SS method

$$\text{(S)}\qquad \hat{z}_i = z_i - \varphi_i(\hat{z}_1,\ldots,\hat{z}_{i-1},z_i,z_{i+1},\ldots,z_n) \quad (i=1,\ldots,n)$$

and

$$\text{(Sc)}\qquad \hat{z}_i = z_i - \varphi_i(\hat{z}_1, \ldots, \hat{z}_{i-1}, z_i, z_{i+1} - \vartheta(z_{i+1}), \ldots, z_n - \vartheta(z_n)) \quad (i = 1, \ldots, n).$$

The corresponding $h$-relations are of the form

$$\hat{h}_i \leq \frac{1}{n-1} h_i^p \left( \sum_{j<i} \hat{h}_j + \sum_{j>i} h_j \right) \quad (i = 1, \ldots, n) \tag{3}$$

and

$$\hat{h}_i \leq \frac{1}{n-1} h_i^p \left( \sum_{j<i} \hat{h}_j + \sum_{j>i} h_j^q \right) \quad (i = 1, \ldots, n). \tag{4}$$

The $R$-order of (S) and (Sc) will be discussed later. The iteration methods (Tc) and (Sc) will be referred to as the methods with one correction.

Further acceleration of convergence can be achieved if we improve the new approximations in the following way

$$\text{(Scc)}\qquad \begin{aligned} \hat{z}_i \;=\; & z_i - \varphi_i(\hat{z}_1 - u(z_1), \ldots, \hat{z}_{i-1} - u(z_{i-1}), z_i, \\ & z_{i+1} - \vartheta(z_{i+1}), \ldots, z_n - \vartheta(z_n)) \quad (i = 1, \ldots, n), \end{aligned}$$

where $u(z)$ is the correction function of order $r > 1$ (say). (Scc) is referred to as the single step method with *two* corrections. The subscript indices $c$ and $cc$ are attributed to the methods with one and two corrections, respectively. In particular cases, for the SS method (Scc) we have the $h$-relations

$$\hat{h}_i \leq \frac{1}{n-1} h_i^p \left( \sum_{j<i} \hat{h}_j^r + \sum_{j>i} h_j^q \right) \quad (i = 1, \ldots, n). \tag{5}$$

Considering the $h$-relations (1), (2), (3), (4), and (5) we notice that the absence of corrections, that is, $u(z) \equiv 0$ and $\vartheta(z) \equiv 0$, may be identified with the choice $r = 1$ and $q = 1$ in (1), (3) and (4).

Particular methods of the form (Tc), (Sc) and (Scc) were the subject of many papers, see e.g. the references in [6]. Clearly, a class of SS methods (Scc) includes (Sc) ($u(z) \equiv 0, r = 1$) and (S) ($u(z) \equiv 0$, $\vartheta(z) \equiv 0, r = q = 1$). Consequently, the $h$-relations (3) and (4) are special cases of (5) so that we shall study a generalized single step method (Scc) and the corresponding $h$-relations (5). In the sequel we shall often use the abbreviation (S) to denote any single step method. In particular, if we deal with the method (S) without corrections, that will be clearly emphasized.

**3. $R$-order of convergence.** Let $O_R(IP, n)$ denote the $R$-order of convergence of an iteration process $IP$, defined as in [5], and let $r \le q \le p$. The explanation for such ordering will be given later (see Remarks 2 and 4 in §4).

Let $h_i^{(k)} = |z_i^{(k)} - \zeta_i|$ $(i = 1, ..., n)$ for $k = 0, 1, ...$ and let $A_n(p, q, r)$ be the $n \times n$ matrix given by

$$A_n(p,q,r) = \begin{bmatrix} p & q & & & \\ & p & q & & \\ & & \ddots & & \\ & & & p & q \\ rp & rq & & & p \end{bmatrix}.$$

**Theorem 1.** *Assume that the initial approximations $z_1^{(0)}, \dots, z_n^{(0)}$ are such that*

$$h_i^{(0)} \le h = \max_{1 \le i \le n} h_i^{(0)} < 1 \tag{6}$$

*Then the R-order of convergence of the single step method (Scc) is given by*

$$O_R((\mathrm{Scc}), n) \ge p + t_n(p, q, r),$$

*where $r \le q \le p$ and $t_n(p, q, r)$ is the unique positive root of the equation*

$$t^n - trq^{n-1} - rpq^{n-1} = 0.$$

Proof. Taking into consideration the iteration index $k = 0, 1, \dots$, the $h$-relations (5) become

$$h_i^{(k+1)} \le \frac{1}{n-1} \left(h_i^{(k)}\right)^p \left[\sum_{j<i} \left(h_j^{(k+1)}\right)^r + \sum_{j>i} \left(h_j^{(k)}\right)^q\right] \tag{7}$$

$(i = 1, \dots, n; k = 0, 1 \dots)$. In the similar way as in [1], starting from (6) and (7) we derive the following relations

$$h_i^{(k+1)} \le h^{s_i^{(k+1)}} \quad (i = 1, \dots, n; k = 0, 1, \dots), \tag{8}$$

where the column vector $s^{(k)} = [s_1^{(k)} \cdots s_n^{(k)}]^T$ is successively computed by

$$s^{(k+1)} = A_n(p, q, r) s^{(k)} \tag{8a}$$

starting with $s^{(0)} = [1 \cdots 1]^T$. For more details see [1] and [6] as well as the references cited there. If $r = 1$ (the absence of corrections at the new approximations), then we have a less general case considered in [6].

The characteristic polynomial of the matrix $A_n(p,q,r)$ (with a leading coefficient 1) is

$$f_n(\lambda; p,q,r) = (\lambda - p)^n - (\lambda - p)rq^{n-1} - rpq^{n-1}.$$

Putting $t = \lambda - p$ we get

$$g_n(t; p,q,r) = f_n(t+p; p,q,r) = t^n - trq^{n-1} - rpq^{n-1}.$$

The graph of the function $y_1(t) = t^n$ and the straight line $y_2(t) = trq^{n-1} + rpq^{n-1}$ intersect only at one point for $t > 0$. Hence the equation

$$t^n - trq^{n-1} - rpq^{n-1} = 0$$

has the unique positive root $t_n(p,q,r) > q$. The corresponding (positive) eigenvalue of the matrix $A_n(p,q,r)$ is $p + t_n(p,q,r)$ and it is equal to the spectral radius $\rho(A_n(p,q,r))$ (since this matrix is nonnegative and irreducible). Following the analysis presented by Alefeld and Herzberger [1] it can be proved that the lower bound of the $R$-order of the iteration method (Scc) for which the $h$-relation (7) holds, is given by the spectral radius $\rho(A_n(p,q,r))$. Hence

$$O_R((\mathrm{Scc}), n) \geq \rho(A_n(p,q,r)) = p + t_n(p,q,r). \quad \square$$

We note that the assertion of Theorem 1 can also be derived from the general results presented in [2]; the details are left to the reader.

Since $t_n(p,q,r) > q$ we see that $O_R((\mathrm{Scc}), n) > p + q$. A more precise lower bound of the $R$-order can be established using Deutsch's result [3]. As in [6] we are able to prove the following assertion.

**Theorem 2.** *For the single step method (Scc) we have*

$$O_R((\mathrm{Scc}), n) > p + q + \frac{rpq + (r-1)q^2}{rp + q + r(n-2)(p+q)}. \tag{9}$$

**Remark 1.** We note that the roots $t_n(p,q,r)$ are strictly decreasing functions of the degree $n$ and strictly increasing functions of the parameters $p, q$ and $r$. The proof can be easily derived as in [4]. Since $\rho(A_n) = p + t_n(p,q,r)$, these assertions also hold for $\rho(A_n)$.

**4. Some practical remarks.** In this section we give some comments and remarks concerning a practical application of single step methods, the choice of correction functions and the convergence speed. When we speak about the computational efficiency of iteration methods, we assume a measure of efficiency discussed in [6, Ch. 6] which takes into account a total CPU (central processor unit) time for a required accuracy of approximations.

**Remark 2.** Iteration methods of the form (Scc) will possess a great efficiency if the correction functions $u$ and $\vartheta$ do not require additional calculations. In other words, these functions should use either 1) already calculated values necessary for the calculation of the basic functions $\varphi_i$ or 2) the values that will be applied in the next iteration steps. This means that $u$ and $\vartheta$ have to use polynomial derivatives of order not higher than that necessary for $\varphi_i$ in the basic formula (T). This implies $r \le p$ and $q \le p$. In the second case it is important to note that the computation of the corrections $u(\hat{z}_1), \dots, u(\hat{z}_{n-1})$ in the last iteration step is inefficient because those values are not applied further. For this reason, one recommends to combine (Scc) and (Sc) so that (Sc) is employed only in the last iteration step.

**Remark 3.** From Theorem 2 we notice that $O_R((\text{Scc}), n) \to p + q$ as $n \to \infty$, thus the $R$-order approaches to the order of convergence of the corresponding TS method for great degrees $n$ of polynomials. Moreover, the correction terms $u(\hat{z}_j)$ become pointless.

**Remark 4.** The conditions $r \le p$ and $q \le p$ are discussed in Remark 2 with respect to observing the greatest efficiency of the applied method (Scc). From Remark 1 we see that the order of convergence is greater if $q$ and $r$ are higher. A simple analysis of the behaviour of the function $y_2(t) = rq^{n-1}(t + p)$ (introduced in the proof of Theorem 1) which takes into account a practical choice $r, q \le 3$ shows that, in order to accelerate the convergence, it is preferable to increase $q$ rather than $r$. According to this and Remark 2 we should choose $r \le q \le p$.

**Remark 5.** Note that

$$t_n(p, q + 1, r) - t_n(p, q, r) > 1$$

for every $n$. Thus, the increase of the convergence order is *at least 1* taking a correction $\vartheta(z)$ of order $q + 1$ instead of that with order $q$. Such a large acceleration cannot be attained increasing order $r$ to $r + 1$ concerning corrections $u(z)$ (see (9)). Moreover, the effect of increasing $r$ for higher $n$ is negligible, which can be noticed from Table 1 for the presented particular (but sufficiently illustrative) example.

**5. An example: Halley's class of iteration methods.** To illustrate the above consideration and comments, we present a class of simultaneous root-finding methods of Halley's type. The methods with two corrections appear to be new ones.

Let $z \in \mathbb{C}$ and let

$$f_k(z) = \frac{P^{(k)}(z)}{P'(z)} \quad (k = 0, 1, 2),$$

where $P$ is a polynomial of degree $n$ having $n$ distinct real or complex zeros $\zeta_1, \ldots, \zeta_n$. Let us introduce

$$g(z) = \frac{1}{f_0(z)} - \frac{1}{2} f_2(z),$$

$$N(z) = f_0(z) = \frac{P(z)}{P'(z)} \qquad \text{(Newton's correction)},$$

$$H(z) = \frac{1}{g(z)} = \frac{P(z)/P'(z)}{1 - P''(z)P(z)/P'(z)^2} \qquad \text{(Halley' correction)}.$$

It is well known that the correction functions $N(z)$ and $H(z)$ appear in the iteration formulas

$$\hat{z} = z - N(z)$$

and

$$\hat{z} = z - H(z)$$

with the convergence orders 2 and 3. Therefore, the corrections $N(z)$ and $H(z)$ have orders 2 and 3, respectively. Furthermore, let

$$\begin{aligned} \Sigma_i(\boldsymbol{a}, \boldsymbol{b}) &= \left[ \sum_{j=1}^{i-1} (z_i - a_j)^{-1} + \sum_{j=i+1}^{n} (z_i - b_j)^{-1} \right]^2 \\ &\quad + \sum_{j=1}^{i-1} (z_i - a_j)^{-2} + \sum_{j=i+1}^{n} (z_i - b_j)^{-2}, \end{aligned}$$

where $\boldsymbol{a} = (a_1, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, \ldots, b_n)$ are some vectors and suppose that $z_1, \ldots, z_n$ are reasonably good approximations to the zeros. We introduce the following vectors of approximations of various type:

$$\begin{aligned}
\boldsymbol{z} &= (z_1, \ldots, z_n) && \text{(the former approximations)}, \\
\hat{\boldsymbol{z}} &= (\hat{z}_1, \ldots, \hat{z}_n) && \text{(the new approximations)} \\
\boldsymbol{z}_N &= (z_{N,1}, \ldots, z_{N,n}), && z_{N,i} = z_i - N(z_i) \\
& && \text{(the Newton approximations)} \\
\boldsymbol{z}_H &= (z_{H,1}, \ldots, z_{H,n}), && z_{H,i} = z_i - H(z_i) \\
& && \text{(the Halley approximations)} \\
\hat{\boldsymbol{z}}_N &= (\hat{z}_{N,1}, \ldots, \hat{z}_{N,n}), && \hat{z}_{N,i} = \hat{z}_i - N(\hat{z}_i) \\
& && \text{(the improved Newton's approximations)} \\
\hat{\boldsymbol{z}}_H &= (\hat{z}_{H,1}, \ldots, \hat{z}_{H,n}), && \hat{z}_{H,i} = \hat{z}_i - H(\hat{z}_i) \\
& && \text{(the improved Halley's approximations)}
\end{aligned}$$

Starting from the Wang–Zheng fixed point relation (see [7]) we present several iteration formulas for simultaneous computation of polynomial zeros:

(T) $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\boldsymbol{z}, \boldsymbol{z})\right]^{-1};$$

(S) $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\hat{\boldsymbol{z}}, \boldsymbol{z})\right]^{-1};$$

$(T_N)$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\boldsymbol{z}_N, \boldsymbol{z}_N)\right]^{-1};$$

$(S_N)$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\hat{\boldsymbol{z}}, \boldsymbol{z}_N)\right]^{-1};$$

$(T_H)$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\boldsymbol{z}_H, \boldsymbol{z}_H)\right]^{-1};$$

$(S_H)$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\hat{\boldsymbol{z}}, \boldsymbol{z}_H)\right]^{-1};$$

$(S_{NN})$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\hat{\boldsymbol{z}}_N, \boldsymbol{z}_N)\right]^{-1};$$

$(S_{NH})$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\hat{\boldsymbol{z}}_N, \boldsymbol{z}_H)\right]^{-1};$$

$(S_{HH})$ $$\hat{z}_i = z_i - \left[g(z_i) - \frac{N(z_i)}{2}\Sigma_i(\hat{\boldsymbol{z}}_H, \boldsymbol{z}_H)\right]^{-1};$$

for $i = 1, \ldots, n$. The first six formulas have already been listed in the book [6, p.219]. The remaining three formulas are new. The subscript indices denote the type of corrections (Newton or Halley's). Since the iteration formula without sums

$$\hat{z}_i = z_i - \frac{1}{g(z_i)} = z_i - H(z_i)$$

defines the well–known Halley's methods, all iteration methods presented above will be referred to as Halley-like methods.

The order of convergence of the total step methods (T), ($T_N$) and ($T_H$) is *four, five* and *six*, respectively. The $R$-order of the single step methods (S), ($S_N$), ($S_H$), ($S_{NN}$), ($S_{NH}$) and ($S_{HH}$) is given by Theorem 1 (with $p = 3$) and listed in Table 1 for several values of the polynomial degree $n$.

**Table 1.**

| n | $r = 1, q = 1$ (S) | $r = 1, q = 2$ ($S_N$) | $r = 1, q = 3$ ($S_H$) | $r = 2, q = 2$ ($S_{NN}$) | $r = 2, q = 3$ ($S_{NH}$) | $r = 3, q = 3$ ($S_{HH}$) |
|---|---|---|---|---|---|---|
| 3 | 4.649 | 5.862 | 6.974 | 6.787 | 8.308 | 9.311 |
| 4 | 4.441 | 5.585 | 6.662 | 6.149 | 7.484 | 8.054 |
| 5 | 4.335 | 5.443 | 6.502 | 5.847 | 7.093 | 7.487 |
| 6 | 4.269 | 5.357 | 6.404 | 5.671 | 6.866 | 7.166 |
| 7 | 4.226 | 5.299 | 6.338 | 5.555 | 6.716 | 6.958 |
| 8 | 4.194 | 5.257 | 6.291 | 5.473 | 6.611 | 6.813 |
| 9 | 4.170 | 5.225 | 6.255 | 5.413 | 6.523 | 6.706 |
| 10 | 4.152 | 5.200 | 6.227 | 5.366 | 6.472 | 6.624 |
| 15 | 4.097 | 5.130 | 6.147 | 5.233 | 6.301 | 6.395 |
| 20 | 4.072 | 5.096 | 6.109 | 5.171 | 6.221 | 6.289 |
| $\infty$ | 4. | 5. | 6. | 5. | 6. | 6. |

The concept of the computational efficiency of an iteration method, (see [6, Ch. 6]), takes into account the R-order of convergence and the number of basic arithmetic operations per iteration, taken with ceratin *weight* depending on the processor time. Accordingly, an iteration method possesses a greater computational efficiency if is needed in order less CPU time to provide the required accuracy of the produced approximations. Since the presented methods have a very high order of convergence (see Table 1) and do not require additional numerical operations in reference to the basic method (of the fourth order), we can conclude that their computational efficiency is great.

## REFERENCES

[1] G. Alefeld, J. Herzberger. On the convergence speed of some algorithms for the simultaneous approximation of the polynomial roots. *SIAM J. Numer. Anal.*, **11** (1974), 237–243.

[2] W. Burmeister, J.W. Schmidt. On the R–order of coupled sequences arising in single–step type methods, *Numer. Math.*, **53** (1988), 653–661.

[3] E. DEUTSCH. Lower bounds for the Perron root of a non-negative irreducible matrix. *Math. Proc. Cambridge Philos. Soc.*, **92** (1982), 49-54.

[4] J. HERZBERGER. Bounds for the R-order of certain iterative numerical processes. *BIT*, **26** (1986), 259-262.

[5] J.M. ORTEGA, W.C. RHEINBOLDT. Iterative solution of nonlinear equations in several variables. Academic Press, New-York, 1970.

[6] M.S. PETKOVIĆ. Iterative methods for simultaneous inclusion of polynomial zeros. Springer-Verlag, Berlin, 1989.

[7] X. WANG, S. ZHENG. A family of parallel and interval iterations for finding all roots of a polynomial simultaneously with rapid convergence (I). *J. Comput. Math.*, **1** (1984), 70-76.

*Faculty of Electronic Engineering*
*University of Niš*
*Beogradska 14, 18000 Niš*
*Yugoslavia*

*Institut für Angewandte Mathematik*
*Universität Hannover*
*Welfengarten 1, D-3000 Hannover*
*FR Germany*

*Received 17.05.1993*