J Optim Theory Appl (2011) 149: 647–664
DOI 10.1007/s10957-011-9796-4

On the Higher-Order Method for the Solution of a Nonlinear Scalar Equation

Min Chen · Tsu-Shuan Chang

Published online: 10 February 2011

Abstract In this paper, a higher-order method for the solution of a nonlinear scalar equation is presented. It is proved that the new method is locally convergent with an order of $(m+2)$, where m is the highest order derivative used in the iterative formula. Some numerical examples are used to demonstrate the new method.

Keywords Nonlinear equation · Iterative method · Simple root · Convergence

1 Introduction

One of the most basic problems in numerical analysis is finding a simple root of a nonlinear equation in one variable:

$$\text{Find } x^* \in R \text{ such that } f(x^*) = 0.$$

The Newton method is the most popular solution for such a problem. Although it has local quadratic convergence under proper conditions, a higher convergence rate may be needed sometimes. Thus some algorithms with higher orders of convergence were developed, such as the Halley method [1–3] and the Chebyshev method [4].

Recently, by using only the first-order derivative, various rates of convergence can be obtained. In [5], the author presents a variant of the Newton method, which is locally convergent with an order of three. In [4], by adding extra terms in combination

Communicated by F.A. Potra.

M. Chen (✉) · T.-S. Chang
Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA
e-mail: ch_min16@hotmail.com

T.-S. Chang
e-mail: tschang@ucdavis.edu

with the Chebyshev-Halley method, the order of convergence of that algorithm is proved to be four, by proper choice of some parameters. In [1], the author presents a modified Halley method which has a quintic convergence by adding a right predictor-corrector in the updating formula.

For the general case, the order of convergence $(m+1)$ can be obtained, where m is the highest order of derivative used in the iterative formula. In [6], a sequence of updating formulas is presented to achieve the above up to $m=4$. For the general m, a hypothesis is proposed. A different method with higher order of convergence is given in [7], which is proved to be locally convergent with order of $(m+1)$. Based upon [7], we propose a new algorithm, which has the convergence rate with one order higher than that in [7]; that is, the order of convergence is $(m+2)$.

The paper is organized as follows. In Sect. 2, some results in [7] are reviewed, which will be used later. In Sect. 3, the new method is presented. In Sect. 4, the convergence analysis is provided. In Sect. 4.1, the case $m=1$ is analyzed in details to motivate the convergence analysis for the general case. In Sect. 4.2, the main result regarding the order of convergence is presented. In Sect. 5, some numerical examples are used to demonstrate the new method. In Sect. 6, a short discussion is given.

2 Review

Since our convergence analysis is closely related to the result in [7], we will first review the essential idea in [7]. Let us denote α as a solution of $f(x)$, i.e. $f(\alpha)=0$. In the rest of the paper, it is assumed that

$$f'(\alpha) \neq 0 \tag{1}$$

and the given function will be smooth. Moreover, it is assumed that the starting point x_0 will be close to α, when discussing the convergence rate of an algorithm.

Basically, its approach is motivated by solving the system of n equations.

$$f^i(\hat{x}) = 0, \quad i = 1, \dots, n, \tag{2}$$

where $f^i(\hat{x}) := (f(\hat{x}))^i$ is the i-th power of $f(\hat{x})$, which of course has the same solution as that of the original problem. The functions $f, f^2, \dots, f^n$ are then approximated by n-degree Taylor polynomials. In other words, we have the following equation

$$\begin{bmatrix} f(\hat{x}_k) \\ \vdots \\ f^n(\hat{x}_k) \end{bmatrix} + \begin{bmatrix} f^{(1)}(\hat{x}_k) & \cdots & f^{(n)}(\hat{x}_k)/n! \\ \vdots & \ddots & \vdots \\ (f^n)^{(1)}(\hat{x}_k) & \cdots & (f^n)^{(n)}(\hat{x}_k)/n! \end{bmatrix} \begin{bmatrix} (\hat{x}-\hat{x}_k) \\ \vdots \\ (\hat{x}-\hat{x}_k)^n \end{bmatrix} = 0, \tag{3}$$

where $(f)^{(j)}(\hat{x}_k)$ represents the j-th order derivative of the function $f(\hat{x})$ evaluated at the point $\hat{x}_k$, and

$$(f^i)^{(j)}(\hat{x}_k) := (f^i(\hat{x}))^{(j)}|_{\hat{x}=\hat{x}_k}, \quad i, j = 1, \dots, n. \tag{4}$$

Then, the iterative equation is determined by solving the following linear equation:

$$\overline{f_n}(\hat{x}_k) + F_n(\hat{x}_k)\begin{bmatrix}\hat{y}_{1,k}\\ \vdots\\ \hat{y}_{n,k}\end{bmatrix} = 0, \tag{5}$$

where

$$\overline{f_n}(\hat{x}_k) := \begin{bmatrix} f(\hat{x}_k)\\ \vdots \\ f^n(\hat{x}_k)\end{bmatrix}, \tag{6}$$

$$F_n(\hat{x}_k) = \begin{bmatrix} f^{(1)}(\hat{x}_k) & \cdots & f^{(n)}(\hat{x}_k)/n!\\ \vdots & \ddots & \vdots\\ (f^n)^{(1)}(\hat{x}_k) & \cdots & (f^n)^{(n)}(\hat{x}_k)/n!\end{bmatrix}. \tag{7}$$

Once the unique solution $\hat{y}_{i,k}$ of linear equation (5) is obtained in [7], the iterative formula takes the form

$$\hat{x}_{k+1} = \hat{x}_k + \hat{y}_{1,k}. \tag{8}$$

As presented in [7], $\hat{y}_{i,k}$ is obtained iteratively by the following steps:

$$\hat{y}_{n,k} = (-f(\hat{x}_k))^n/[U_n(f;\hat{x}_k)]_{n+1,n+1}, \tag{9}$$

$$\hat{y}_{i,k} = \left((-f(\hat{x}_k))^i - \sum_{j=i+2}^{n+1}[U_n(f;\hat{x}_k)]_{i+1,j}\cdot \hat{y}_{j-1,k}\right)/[U_n(f;\hat{x}_k)]_{i+1,i+1}, \tag{10}$$

where $U_n(f;\hat{x}_k)$ is a upper triangular matrix with dimension of $(n+1)\times(n+1)$, and the first row, except from the first component, is null. Mathematically, it is given as

$$[U_n(f;\hat{x}_k)]_{1,1} = 1, \tag{11}$$

$$[U_n(f;\hat{x}_k)]_{1,i} = 0, \quad 2\le i\le n+1, \tag{12}$$

$$[U_n(f;\hat{x}_k)]_{i,i} = (f^{(1)}(\hat{x}_k))^{i-1}, \quad 2\le i\le n+1, \tag{13}$$

$$[U_n(f;\hat{x}_k)]_{i,j} = \sum_{h=i-1}^{j-1}[U_n(f;\hat{x}_k)]_{i-1,h}\cdot f^{(j-h)}(\hat{x}_k)/(j-h)!,$$
$$i<j \text{ and } i,j = 2,3,\dots,n+1. \tag{14}$$

It is proved in [7] that the order of convergence of the method is at least $(n+1)$, where n represents the highest order derivative of $f(\hat{x})$ in the iterative formula. Note that the method is reduced to the Newton method if $n=1$. When $n=2$, it is reduced to the Chebyshev method as below:

$$\hat{x}_{k+1} = \hat{x}_k - \frac{f(\hat{x}_k)}{f^{(1)}(\hat{x}_k)} - \frac{f^2(\hat{x}_k)f^{(2)}(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3}. \tag{15}$$

When $n=3$, it has the updating formula

$$\hat{x}_{k+1}=\hat{x}_k-\frac{f(\hat{x}_k)}{f^{(1)}(\hat{x}_k)}-\frac{f^{(2)}(\hat{x}_k)f^2(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3}-\frac{(f^{(2)}(\hat{x}_k))^2f^3(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^5}+\frac{f^{(3)}(\hat{x}_k)f^3(\hat{x}_k)}{6(f^{(1)}(\hat{x}_k))^4}. \tag{16}$$

For the notational simplicity, we denote the updating formula as follows.

$$\hat{x}_{k+1}=\hat{\Phi}_n(\hat{x}_k)=\hat{g}_{n-1}(\hat{x}_k)+\hat{h}_n(f^{(n)}(\hat{x}_k)), \tag{17}$$

where the index of $\hat{\Phi}_n(x_k)$ denotes the highest order n of the derivative used in the updating formula. The term $\hat{h}_n(f^{(n)}(\hat{x}_k))$ contains only the term with the highest order n of the derivative, and $\hat{g}_{n-1}(\hat{x}_k)$ is the collection of the remainders up to the order $(n-1)$ of the derivative.

3 A Higher-Order Method

The basic idea of the new method is to use the linear approximation of $f^{(n)}(x_k)$ as follows:

$$f^{(n)}(x_k)\approx\frac{f^{(n-1)}(x_k)-f^{(n-1)}(z_k)}{x_k-z_k}. \tag{18}$$

We will assume through this paper that

$$f^{(1)}(\alpha)\neq 0. \tag{19}$$

Due to this, we can always find a neighborhood of α such that $f^{(1)}(x_k)\neq 0$. Thus, we can choose in this neighborhood

$$z_k=x_k-f(x_k)/f^{(1)}(x_k). \tag{20}$$

By substituting (18) into (17), we obtain our updating formula

$$x_{k+1}=\hat{\Phi}_n(x_k)=\hat{g}_{n-1}(x_k)+\hat{h}_n\left(\frac{f^{(n-1)}(x_k)-f^{(n-1)}(z_k)}{x_k-z_k}\right) \tag{21}$$

or

$$x_{k+1}=\Phi_m(x_k)=g_m(x_k)+h_m(f^{(m)}(x_k)), \tag{22}$$

where $m=n-1$. Thus we have

$$g_m(.)=\hat{g}_{n-1}(.);\qquad h_m(f^{(m)}(x_k))=\hat{h}_n\left(\frac{f^{(n-1)}(x_k)-f^{(n-1)}(z_k)}{x_k-z_k}\right). \tag{23}$$

When $n=2$, the updating formula is

$$x_{k+1}=x_k-\frac{f(x_k)}{f^{(1)}(x_k)}-\frac{f^{(1)}(x_k)f(x_k)-f^{(1)}(z_k)f(x_k)}{2(f^{(1)}(x_k))^2}, \tag{24}$$

which is different from the formula in (15) only in its last term.

When $n = 3$, the updating formula is

$$x_{k+1} = x_k - \frac{f(x_k)}{f^{(1)}(x_k)} - \frac{f^{(2)}(x_k) f^2(x_k)}{2(f^{(1)}(x_k))^3} - \frac{(f^{(2)}(x_k))^2 f^3(x_k)}{2(f^{(1)}(x_k))^5} + \frac{(f^{(2)}(x_k) - f^{(2)}(z_k)) f^2(x_k)}{6(f^{(1)}(x_k))^3}, \tag{25}$$

which is also different from (16) only in its last term.

4 Convergence Analysis

In this section, we will present the convergence analysis of the new method, which has the order of convergence $(m+2)$, where m represents the highest order of derivative used in the iterative formula. In Sect. 4.1, two different proofs for the case of $m = 1$ are given to motivate the proof for the general case. In Sect. 4.2, the main result regarding the order of convergence is presented.

4.1 Motivation

As mentioned, we will compute the convergence coefficient explicitly in Lemma 4.1, and thus show that the order of the convergence rate is three when $m = 1$.

Lemma 4.1 *Let $I \in IR$ be an open interval and $\alpha \in I$ be a simple zero of sufficiently differentiable function $f : I \to IR$. If x_0 is sufficiently close to α, then the order of convergence of the method defined by* (20) *and* (24) *is three. It satisfies the error equation*:

$$e_{k+1} = \left(2c_2^2 + \frac{1}{2}c_3\right)e_k^3 + O(e_k^4), \tag{26}$$

where $e_i = x_i - \alpha$ and $c_i = \frac{f^{(i)}(\alpha)}{f^{(1)}(\alpha)\cdot i!}$, $i = 2, \dots, k+1$.

Proof From the definition, using Taylor series expansion, we have

$$\begin{aligned} f(x_k) &= f^{(1)}(\alpha)(x_k - \alpha) + \frac{1}{2!} f^{(2)}(\alpha)(x_k - \alpha)^2 \\ &\quad + \frac{1}{3!} f^{(3)}(\alpha)(x_k - \alpha)^3 + o(x_k - \alpha)^4 \\ &= f^{(1)}(\alpha)[e_k + c_2 e_k^2 + c_3 e_k^3 + o(e_k^4)], \end{aligned} \tag{27}$$

$$f^{(1)}(x_k) = f^{(1)}(\alpha)[1 + 2c_2 e_k + 3c_3 e_k^2 + o(e_k^3)]. \tag{28}$$

From (27) and (28), we can get

$$\frac{f(x_k)}{f^{(1)}(x_k)} = e_k - c_2 e_k^2 - (2c_3 - 2c_2^2)e_k^3 + O(e_k^4), \tag{29}$$

$$\begin{aligned} z_k &= x_k - \frac{f(x_k)}{f^{(1)}(x_k)} \\ &= \alpha + c_2 e_k^2 + (2c_3 - 2c_2^2)e_k^3 + o(e_k^4). \end{aligned} \tag{30}$$

By using Taylor series and (30), we can also write $f(z_k)$ and $f^{(1)}(z_k)$ in terms of α and z_k:

$$\begin{aligned} f(z_k) &= f^{(1)}(\alpha)(z_k - \alpha) + \frac{1}{2!} f^{(2)}(\alpha)(z_k - \alpha)^2 + \cdots \\ &= f^{(1)}(\alpha)[(z_k - \alpha) + c_2(z_k - \alpha)^2 + \cdots] \\ &= f^{(1)}(\alpha)[c_2 e_k^2 + (2c_3 - 2c_2^2)e_k^3 + o(e_k^4)], \end{aligned} \tag{31}$$

$$\begin{aligned} f^{(1)}(z_k) &= f^{(1)}(\alpha) + f^{(2)}(\alpha)(z_k - \alpha) + \cdots \\ &= f^{(1)}(\alpha)[1 + 2c_2(z_k - \alpha) + \cdots] \\ &= f^{(1)}(\alpha)[1 + 2c_2^2 e_k^2 + 4c_2(c_3 - c_2^2)e_k^3 + o(e_k^4)]. \end{aligned} \tag{32}$$

From (28) and (32), we have

$$\frac{f^{(1)}(z_k)}{f^{(1)}(x_k)} = 1 - 2c_2 e_k + 3(2c_2^2 - c_3)e_k^2 + o(e_k^3). \tag{33}$$

The last term in (24) equals

$$\begin{aligned} &\frac{(f^{(1)}(x_k) - f^{(1)}(z_k))f(x_k)}{2(f^{(1)}(x_k))^2} = \frac{f(x_k)}{2f^{(1)}(x_k)}\left(1 - \frac{f^{(1)}(z_k)}{f^{(1)}(x_k)}\right) \\ &= c_2 e_k^2 + \left(\frac{3}{2}c_3 - 4c_2^2\right)e_k^3 + o(e_k^4). \end{aligned} \tag{34}$$

Note that the minimum order of the first term in (34) is e_k; the second term only needs to contain order up to e_k^2. Finally,

$$x_{k+1} = \alpha + \left(2c_2^2 + \frac{1}{2}c_3\right)e_k^3 + o(e_k^4), \tag{35}$$

$$e_{k+1} = x_{k+1} - \alpha = \left(2c_2^2 + \frac{1}{2}c_3\right)e_k^3 + o(e_k^4). \tag{36}$$

□

Although it has been proved in [7] that the order of convergence, by using (15), is three, we will now compute explicitly the coefficient of convergence by following the same approach as in Lemma 4.1.

Lemma 4.2 *Let $I \in IR$ be an open interval and $\alpha \in I$ be a simple zero of sufficiently differentiable function $f : I \to IR$. If $\hat{x}_0$ is sufficiently close to α, then the order of* *convergence of the method defined by* (15) *is three. It satisfies the error equation*:

$$\hat{e}_{k+1} = (2c_2^2 - c_3)\hat{e}_k^3 + o(\hat{e}_k^4), \tag{37}$$

where $\hat{e}_i = \hat{x}_i - \alpha$ *and* $c_i = \frac{f^{(i)}(\alpha)}{f^{(1)}(\alpha)\cdot i!}$, $i = 2, \dots, k+1$.

Proof The proof is similar to that in Lemma 4.1. By using the Taylor series expansion, we have

$$\begin{aligned} f(\hat{x}_k) &= f^{(1)}(\alpha)(\hat{x}_k - \alpha) + \frac{1}{2!} f^{(2)}(\alpha)(\hat{x}_k - \alpha)^2 \\ &\quad + \frac{1}{3!} f^{(3)}(\alpha)(\hat{x}_k - \alpha)^3 + o(\hat{x}_k - \alpha)^4 \\ &= f^{(1)}(\alpha)[\hat{e}_k + c_2\hat{e}_k^2 + c_3\hat{e}_k^3 + o(\hat{e}_k^4)], \end{aligned} \tag{38}$$

$$f^{(1)}(\hat{x}_k) = f^{(1)}(\alpha)[1 + 2c_2\hat{e}_k + 3c_3\hat{e}_k^2 + o(\hat{e}_k^3)], \tag{39}$$

$$f^{(2)}(\hat{x}_k) = f^{(1)}(\alpha)[2c_2 + 6c_3\hat{e}_k + o(\hat{e}_k^2)]. \tag{40}$$

From (38) and (39), we can get

$$\frac{f(\hat{x}_k)}{f^{(1)}(\hat{x}_k)} = \hat{e}_k - c_2\hat{e}_k^2 - (2c_3 - 2c_2^2)\hat{e}_k^3 + o(\hat{e}_k^4). \tag{41}$$

Note that

$$\frac{f^2(\hat{x}_k)f^{(2)}(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3} = \left(\frac{f(\hat{x}_k)}{f^{(1)}(\hat{x}_k)}\right)^2 \cdot \frac{f^{(2)}(\hat{x}_k)}{2f^{(1)}(\hat{x}_k)}, \tag{42}$$

and the minimal order of the first term of (42) is $\hat{e}_k^2$. Thus, the second term only needs to contain order up to $\hat{e}_k$.

$$\frac{f^{(2)}(\hat{x}_k)}{2f^{(1)}(\hat{x}_k)} = c_2 + (3c_3 - 2c_2^2)\hat{e}_k + o(\hat{e}_k^2). \tag{43}$$

Substitute (41) and (43) to (42), we have

$$\frac{f^2(\hat{x}_k)f^{(2)}(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3} = c_2\hat{e}_k^2 + (3c_3 - 4c_2^2)\hat{e}_k^3 + o(\hat{e}_k^4). \tag{44}$$

Finally,

$$\begin{aligned} \hat{x}_{k+1} &= \alpha + (2c_2^2 - c_3)\hat{e}_k^3 + o(\hat{e}_k^4), \\ \hat{e}_{k+1} &= \hat{x}_{k+1} - \alpha = (2c_2^2 - c_3)\hat{e}_k^3 + o(\hat{e}_k^4). \end{aligned}$$

□

Remark Since all the equations remain the same during the derivation if we replace $\hat{x}_k$ by x_k and $\hat{e}_k$ by e_k, (44) is still true with such substitutions. That is, we have

$$\frac{f^2(x_k)f^{(2)}(x_k)}{2(f^{(1)}(x_k))^3} = c_2e_k^2 + (3c_3 - 4c_2^2)e_k^3 + o(e_k^4). \tag{45}$$

From the proofs of Lemma 4.1 and Lemma 4.2, we can see how they are related. We will give an alternative proof of Lemma 4.1 by observing the difference of the proofs of Lemmas 4.1 and 4.2. Now, we restate Lemma 4.1 in Lemma 4.3 and present the alternative proof.

Lemma 4.3 *Let $I \in IR$ be an open interval and $\alpha \in I$ be a simple zero of sufficiently differentiable function $f : I \to IR$. If x_0 is sufficiently close to α, then the order of convergence of the method defined by (20) and (24) is three. It satisfies the error equation*:

$$e_{k+1} = \left(2c_2^2 + \frac{1}{2}c_3\right)e_k^3 + O(e_k^4), \tag{46}$$

where $e_i = x_i - \alpha$ and $c_i = \frac{f^{(i)}(\alpha)}{f^{(1)}(\alpha)\cdot i!}$, $i = 2, \dots, k+1$.

Proof Let us present the iteration formula (20) and (24) in terms of the format of the general updating formula in (22) by using $m = 1$:

$$x_{k+1} = \Phi_1(x_k) = g_1(x_k) + h_1(f^{(1)}(x_k)), \tag{47}$$

where $h_1(f^{(1)}(x_k)) = -\frac{f^{(1)}(x_k)f(x_k) - f^{(1)}(z_k)f(x_k)}{2(f^{(1)}(x_k))^2}$ and

$$g_1(x_k) = x_k - \frac{f(x_k)}{f^{(1)}(x_k)}.$$

The updating formula (15) in Lemma 4.2 can be expressed by using the general formula (17) with $n = 2$:

$$\hat{x}_{k+1} = \hat{\Phi}_2(\hat{x}_k) = \hat{g}_1(\hat{x}_k) + \hat{h}_2(f^{(2)}(\hat{x}_k)), \tag{48}$$

where $\hat{h}_2(f^{(2)}(\hat{x}_k)) = -\frac{f^2(\hat{x}_k)f^{(2)}(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3}$ and

$$\hat{g}_1(\hat{x}_k) = \hat{x}_k - \frac{f(\hat{x}_k)}{f^{(1)}(\hat{x}_k)}.$$

Note that both m and n represent the highest order of derivative in the iteration formula.

We know that the convergent order of the method defined by (48) (for $n = 2$) is three. So, we can assume

$$\hat{x}_{k+1} - \alpha = \hat{g}_1(\hat{x}_k) + \hat{h}_2(f^{(2)}(\hat{x}_k)) - \alpha = -\alpha + K_2\hat{e}_k^3 + o(\hat{e}_k^4), \tag{49}$$

where $K_2 \in IR$ is a constant.

By using Taylor expansion, after computation (see the proof of Lemma 4.2 for details), we know

$$\hat{h}_2(f^{(2)}(\hat{x}_k)) = -\frac{f^2(\hat{x}_k)f^{(2)}(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3} = -(c_2\hat{e}_k^2 + (3c_3 - 4c_2^2)\hat{e}_k^3 + o(\hat{e}_k^4)). \tag{50}$$

Substituting (50) to (49), we have

$$\hat{g}_1(\hat{x}_k) = -\alpha + c_2\hat{e}_k^2 + (K_2 + 3c_3 - 4c_2^2)\hat{e}_k^3 + o(\hat{e}_k^4).$$

Then, from the functional relationship $g_m(.) = \hat{g}_{n-1}(.)$ with $m = n - 1$, we get

$$g_1(x_k) = -\alpha + c_2e_k^2 + (K_2 + 3c_3 - 4c_2^2)e_k^3 + o(e_k^4). \tag{51}$$

Substituting (51) to (47), we have

$$x_{k+1} = -\alpha + c_2e_k^2 + (K_2 + 3c_3 - 4c_2^2)e_k^3 + h_1(f^{(1)}(x_k)) + o(e_k^4).$$

By using again Taylor expansion, after some steps of computation (see the proof of Lemma 4.1 for details), we know

$$h_1(f^{(1)}(x_k)) = -\frac{(f^{(1)}(x_k) - f^{(1)}(z_k))f(x_k)}{2(f^{(1)}(x_k))^2} = -\left(c_2e_k^2 + \left(\frac{3}{2}c_3 - 4c_2^2\right)e_k^3 + o(e_k^4)\right). \tag{52}$$

Finally, we can get

$$\begin{aligned}
x_{k+1} &= -\alpha + c_2e_k^2 + (K_2 + 3c_3 - 4c_2^2)e_k^3 - \left(c_2e_k^2 + \left(\frac{3}{2}c_3 - 4c_2^2\right)e_k^3\right) + o(e_k^4) \\
&= -\alpha + \left(K_2 + \frac{3}{2}c_3\right)e_k^3 + o(e_k^4), \\
e_{k+1} &= x_{k+1} - \alpha = \left(K_2 + \frac{3}{2}c_3\right)e_k^3 + o(e_k^4).
\end{aligned}$$

□

Thus, the method defined in (20) and (24) has the order of convergence three. Furthermore, we know from Lemma 4.2 that $K_2 = 2c_2^2 - c_3$. Thus, the result is the same as that in Lemma 4.1.

4.2 Main Result

From the proof of Lemma 4.3, we can see that it can be extended to general cases except that we need to prove: (1) There is only one term in the highest order of derivative in $\hat{\Phi}_n(\hat{x}_k)$; (2) The Taylor expansions of $h_m(f^{(m)}(x_k))$ and $\hat{h}_n(f^{(n)}(x_k))$ have the same coefficients up to e_k^n, where $m = n - 1$. In Lemma 4.4, we prove the former fact and in Lemma 4.5 the latter one.

Lemma 4.4 *For the method defined by*

$$\hat{x}_{k+1} = \hat{\Phi}_n(\hat{x}_k) = \hat{g}_{n-1}(\hat{x}_k) + \hat{h}_n(f^{(n)}(\hat{x}_k)), \tag{53}$$

it satisfies

$$\hat{h}_n(f^{(n)}(\hat{x}_k)) = (-1)^{n+1}\frac{f^{(n)}(\hat{x}_k)f^n(\hat{x}_k)}{n!(f^{(1)}(\hat{x}_k))^{n+1}}. \tag{54}$$

Proof According to [7], which is related to the unique solution of linear equation (5), can be expressed as

$$\hat{\Phi}_n(\hat{x}_k) = \hat{x}_k + \hat{y}_{1,k},$$

where $\hat{y}_{1,k}$ can be solved iteratively by (9), (10), (11), (12), (13) and (14).

From (14), we know

$$[U_n(f;\hat{x}_k)]_{2,j} = \sum_{h=1}^{j-1}[U_n(f;\hat{x}_k)]_{1,h} \cdot f^{(j-h)}(\hat{x}_k)/(j-h)!, \quad 3 \le j \le n+1.$$

From (11), (12) and (13), we can have

$$[U_n(f;\hat{x}_k)]_{2,j} = f^{(j-1)}(\hat{x}_k)/(j-1)!, \quad 2 \le j \le n+1. \tag{55}$$

From (55), we know that only $[U_n(f;\hat{x}_k)]_{2,n+1}$ has the highest n-th order derivative of $f(\hat{x})$. We can prove by induction that the other row components of $U_n(f;\hat{x}_k)$ do not contain the n-th order derivative of $f(\hat{x})$.

From (14), we also know

$$[U_n(f;\hat{x}_k)]_{3,j} = \sum_{h=2}^{j-1}[U_n(f;\hat{x}_k)]_{2,h} \cdot f^{(j-h)}(\hat{x}_k)/(j-h)!, \quad 4 \le j \le n+1. \tag{56}$$

Since $2 \le h \le j-1 \le n$ and $j-h \le n-1$, the term of $[U_n(f;\hat{x}_k)]_{2,h}$ and $f^{(j-h)}(\hat{x}_k)$ will not contain the n-th order derivative of $f(\hat{x})$. Thus, the components of the third row of $U_n(f;\hat{x}_k)$ do not contain the n-th order derivative of $f(\hat{x})$.

Now, by assuming that the components of the s-th row of $U_n(f;\hat{x}_k)$ do not contain the n-th order derivative of $f(\hat{x})$ is true for $3 \le s \le n$, we will prove that the same is also true for $(s+1)$. From (14), we can get

$$[U_n(f;\hat{x}_k)]_{s+1,j} = \sum_{h=s}^{j-1}[U_n(f;\hat{x}_k)]_{s,h} \cdot f^{(j-h)}(\hat{x}_k)/(j-h)!, \quad s+1 \le j \le n+1. \tag{57}$$

Since $s \le h \le j-1 \le n$, we have $j-h \le n+1-s \le n-2$. Thus, the components of the $(s+1)$-th row of $U_n(f;\hat{x}_k)$ do not contain the n-th order derivative of $f(\hat{x})$. Therefore, the components of $U_n(f;\hat{x}_k)$ do not contain the highest n-th order derivative of $f(\hat{x})$ except $[U_n(f;\hat{x}_k)]_{2,n+1}$.

From (10), we have

$$\hat{y}_{1,k} = \left(-f(\hat{x}_k) - \sum_{j=3}^{n+1}[U_n(f;\hat{x}_k)]_{2,j} \cdot \hat{y}_{j-1,k}\right)/[U_n(f;\hat{x}_k)]_{2,2}. \tag{58}$$

From (58), by considering the property of $U_n(f;\hat{x}_k)$, we know the n-th derivative of $f(\hat{x}_k)$ in $\hat{y}_{1,k}$ must come from the term $[U_n(f;\hat{x}_k)]_{2,n+1} \cdot \hat{y}_{n,k}/[U_n(f;\hat{x}_k)]_{2,2}$. Thus we have

$$\hat{h}_n(f^{(n)}(\hat{x}_k)) = [U_n(f;\hat{x}_k)]_{2,n+1} \cdot \hat{y}_{n,k}/[U_n(f;\hat{x}_k)]_{2,2}. \tag{59}$$

From (9) and (13), we can get

$$\hat{y}_{n,k} = \frac{(-f(x_k))^n}{(f^{(1)}(x_k))^n}. \tag{60}$$

Substitute (55), (60) to (59), we have

$$\hat{h}_n(f^{(n)}(\hat{x}_k)) = -\frac{f^{(n)}(\hat{x}_k)}{n!} \cdot \frac{(-f(\hat{x}_k))^n}{(f^{(1)}(\hat{x}_k))^n} / f^{(1)}(\hat{x}_k) = (-1)^{n+1} \frac{f^{(n)}(\hat{x}_k) f^n(\hat{x}_k)}{n!(f^{(1)}(\hat{x}_k))^{n+1}}.$$

□

Lemma 4.5 *Let $I \in IR$ be an open interval and $\alpha \in I$ be a simple zero of sufficiently differentiable function $f : I \to IR$. If x_k is generated by the method defined by (20) and (22), $\hat{x}_k$ is generated by the method defined by (17) (presented in [7]). Then, for an arbitrary integer n, it satisfies*

$$\frac{f^{(n)}(\hat{x}_k) f^n(\hat{x}_k)}{n!(f^{(1)}(\hat{x}_k))^{n+1}} = A_n \hat{e}_k^n + \hat{B}_n \hat{e}_k^{n+1} + o(\hat{e}_k^{n+2}), \tag{61}$$

$$\frac{f^{(n)}(x_k) f^n(x_k)}{n!(f^{(1)}(x_k))^{n+1}} = A_n e_k^n + \hat{B}_n e_k^{n+1} + o(e_k^{n+2}), \tag{62}$$

$$\frac{f^{n-1}(x_k)(f^{(n-1)}(x_k) - f^{(n-1)}(z_k))}{n!(f^{(1)}(x_k))^n} = A_n e_k^n + B_n e_k^{n+1} + o(e_k^{n+2}), \tag{63}$$

where $e_k = x_k - \alpha$, $\hat{e}_k = \hat{x}_k - \alpha$, $A_n \in IR$, $B_n \in IR$ and $\hat{B}_n \in IR$.

Proof We will prove this Lemma by induction. From the proof of Lemma 4.1 and Lemma 4.2, we know

$$\frac{f^2(\hat{x}_k) f^{(2)}(\hat{x}_k)}{2(f^{(1)}(\hat{x}_k))^3} = c_2 \hat{e}_k^2 + (3c_3 - 4c_2^2)\hat{e}_k^3 + o(\hat{e}_k^4),$$

$$\frac{f^2(x_k) f^{(2)}(x_k)}{2(f^{(1)}(x_k))^3} = c_2 e_k^2 + (3c_3 - 4c_2^2) e_k^3 + o(e_k^4),$$

$$\frac{(f^{(1)}(x_k) - f^{(1)}(z_k)) f(x_k)}{2(f^{(1)}(x_k))^2} = c_2 e_k^2 + \left(\frac{3}{2}c_3 - 4c_2^2\right) e_k^3 + o(e_k^4).$$

It shows that (61), (62) and (63) are true for $n = 2$. Now, assume (61), (62) and (63) are true for a given $1 \le i \le n$. From (61), we have

$$f^{(i)}(\hat{x}_k) = [A_i \hat{e}_k^i + \hat{B}_i \hat{e}_k^{i+1} + M_i \hat{e}_k^{i+2} + o(\hat{e}_k^{i+3})] \frac{i!(f^{(1)}(\hat{x}_k))^{i+1}}{f^i(\hat{x}_k)}, \tag{64}$$

where $M_i \in IR$ is the coefficient of the first component of its last term. By taking the derivative of (64), we can get

$$f^{(i+1)}(\hat{x}_k) = [iA_i \hat{e}_k^{i-1} + (i+1)\hat{B}_i \hat{e}_k^i + (i+2)M_i \hat{e}_k^{i+1} + o(\hat{e}_k^{i+2})] \frac{i!(f^{(1)}(\hat{x}_k))^{i+1}}{f^i(\hat{x}_k)}$$

$$+ [A_i \hat{e}_k^i + \hat{B}_i \hat{e}_k^{i+1} + M_i \hat{e}_k^{i+2} + o(\hat{e}_k^{i+3})]$$
$$\times \left[\frac{(i+1)(f^{(1)}(\hat{x}_k))^i f^{(2)}(\hat{x}_k)}{(f(\hat{x}_k))^i} - \frac{i(f^{(1)}(\hat{x}_k))^{i+2}}{(f(\hat{x}_k))^{i+1}} \right] i!.$$

Then,

$$\begin{aligned} &\frac{f^{(i+1)}(\hat{x}_k)(f(\hat{x}_k))^{i+1}}{(i+1)!(f^{(1)}(\hat{x}_k))^{i+2}} \\ &\quad = (iA_i \hat{e}_k^{i-1} + (i+1)\hat{B}_i \hat{e}_k^i + (i+2)M_i \hat{e}_k^{i+1} + o(\hat{e}_k^{i+2})) \frac{f(\hat{x}_k)}{(i+1)f^{(i)}(\hat{x}_k)} \\ &\qquad + (A_i \hat{e}_k^i + \hat{B}_i \hat{e}_k^{i+1} + M_i \hat{e}_k^{i+2} + o(\hat{e}_k^{i+3})) \left(\frac{f(\hat{x}_k) f^{(2)}(\hat{x}_k)}{(f^{(1)}(\hat{x}_k))^2} - \frac{i}{(i+1)} \right). \end{aligned} \tag{65}$$

From (41) and (43), we know

$$\frac{f(\hat{x}_k)}{f^{(1)}(\hat{x}_k)} = \hat{e}_k - c_2 \hat{e}_k^2 - (2c_3 - 2c_2^2)\hat{e}_k^3 + o(\hat{e}_k^4), \tag{66}$$

$$\frac{f(\hat{x}_k) f^{(2)}(\hat{x}_k)}{(f^{(1)}(\hat{x}_k))^2} = 2c_2 \hat{e}_k + (-6c_2^2 + 6c_3)\hat{e}_k^2 + o(\hat{e}_k^3). \tag{67}$$

Substituting (66) and (67) into (65), we get

$$\begin{aligned} &\frac{f^{(i+1)}(\hat{x}_k)(f(\hat{x}_k))^{i+1}}{(i+1)!(f^{(1)}(\hat{x}_k))^{i+2}} \\ &\quad = \left(\hat{B}_i - \frac{i}{i+1} c_2 A_i - \frac{i}{i+1} \hat{B}_i + 2c_2 A_i \right) \hat{e}_k^{i+1} \\ &\qquad + \left(\frac{i+2}{i+1} M_i - c_2 \hat{B}_i - \frac{i}{i+1} M_i + 2c_2 \hat{B}_i + A_i(-6c_2^2 + 6c_3) \right. \\ &\qquad \left. - \frac{i}{i+1} A_i (2c_3 - 2c_2^2) \right) \hat{e}_k^{i+2} + o(\hat{e}_k^{i+3}) \\ &\quad = \left(\frac{i+2}{i+1} c_2 A_i + \frac{1}{i+1} \hat{B}_i \right) \hat{e}_k^{i+1} \\ &\qquad + \left(\frac{2}{i+1} M_i + c_2 \hat{B}_i + \frac{2i+3}{i+1} A_i (2c_3 - 2c_2^2) \right) \hat{e}_k^{i+2} + o(\hat{e}_k^{i+3}). \end{aligned}$$

That is,

$$\frac{f^{(i+1)}(\hat{x}_k)(f(\hat{x}_k))^{i+1}}{(i+1)!(f^{(1)}(\hat{x}_k))^{i+2}} = A_{i+1} \hat{e}_k^{i+1} + \hat{B}_{i+1} \hat{e}_k^{i+2} + o(\hat{e}_k^{i+3}), \tag{68}$$

where $A_{i+1} = \frac{i+2}{i+1} c_2 A_i + \frac{1}{i+1} \hat{B}_i$ and

$$\hat{B}_{i+1} = \frac{2}{i+1} M_i + c_2 \hat{B}_i + \frac{2i+3}{i+1} A_i (2c_3 - 2c_2^2).$$

Thus, (61) is true for $i+1$.

For the same reason as in (45), if we replace $\hat{x}_k$ by x_k and $\hat{e}_k$ by e_k, (68) is still true. That is, we also have

$$\frac{f^{(i+1)}(x_k)(f(x_k))^{i+1}}{(i+1)!(f^{(1)}(x_k))^{i+2}} = A_{i+1}e_k^{i+1} + \hat{B}_{i+1}e_k^{i+2} + o(e_k^{i+3}). \tag{69}$$

Equation (62) is true for $i+1$.

Recalling the Taylor expansion, we have

$$f^{(i)}(z_k) = f^{(i)}(x_k) + f^{(i+1)}(x_k)(z_k - x_k) + \frac{1}{2}f^{(i+2)}(x_k)(z_k - x_k)^2 + o(z_k - x_k)^3.$$

Then,

$$\begin{aligned} f^{(i)}(x_k) - f^{(i)}(z_k) &= f^{(i+1)}(x_k)(x_k - z_k) - \frac{1}{2}f^{(i+2)}(x_k)(x_k - z_k)^2 + o(x_k - z_k)^3 \\ &= f^{(i+1)}(x_k)\frac{f(x_k)}{f^{(1)}(x_k)} - \frac{1}{2}f^{(i+2)}(x_k)\left(\frac{f(x_k)}{f^{(1)}(x_k)}\right)^2 \\ &\quad + o\left(\frac{f(x_k)}{f^{(1)}(x_k)}\right)^3. \end{aligned}$$

Thus,

$$\begin{aligned} &\frac{(f^{(i)}(x_k) - f^{(i)}(z_k))f^i(x_k)}{(i+1)!(f^{(i)}(x_k))^{(i+1)}} \\ &\quad = \frac{f^{(i+1)}(x_k)f^{i+1}(x_k)}{(i+1)!(f^{(i)}(x_k))^{(i+2)}} - \frac{f^{(i+2)}(x_k)}{2(i+1)!f^{(1)}(x_k)}\left(\frac{f(x_k)}{f^{(1)}(x_k)}\right)^{i+2} \\ &\qquad + o\left(\frac{f(x_k)}{f^{(1)}(x_k)}\right)^{i+3}. \end{aligned} \tag{70}$$

Since the minimal order of $\frac{f(x_k)}{f^{(1)}(x_k)}$ is e_k, $o(\frac{f(x_k)}{f^{(1)}(x_k)})^{i+3}$ is equivalent to $o(e_k^{i+3})$. Substituting (66) and (69) into (70), after some steps we get the following:

$$\begin{aligned} \frac{(f^{(i)}(x_k) - f^{(i)}(z_k))f^i(x_k)}{(i+1)!(f^{(i)}(x_k))^{(i+1)}} &= A_{i+1}e_k^{i+1} + \hat{B}_{i+1}e_k^{i+2} \\ &\quad - \frac{f^{(i+2)}(x_k)}{2(i+1)!f^{(1)}(x_k)}e_k^{i+2} + o(e_k^{i+3}) \\ &= A_{i+1}e_k^{i+1} + B_{i+1}e_k^{i+2} + o(e_k^{i+3}), \end{aligned}$$

where $B_{i+1} = \hat{B}_{i+1} - \frac{f^{(i+2)}(x_k)}{2(i+1)!f^{(1)}(x_k)}$. In other words, (63) is true for $i+1$. □

Theorem 4.1 *Let $I \in IR$ be an open interval and $\alpha \in I$ be a simple zero of sufficiently differentiable function $f: I \to IR$. Assume $f^{(1)}(\alpha) \neq 0$ and x_0 be sufficiently* *close to* α. *Then the order of convergence of the algorithm defined by* (20) *and* (22) *is* $(m+2)$.

Proof From Lemma 4.4, we know,

$$\hat{x}_{k+1} = \hat{\Phi}_n(\hat{x}_k) = \hat{g}_{n-1}(\hat{x}_k) + \hat{h}_n(f^{(n)}(\hat{x}_k)), \tag{71}$$

where $\hat{h}_n(f^{(n)}(\hat{x}_k)) = (-1)^{n+1}\frac{f^{(n)}(\hat{x}_k)f^n(\hat{x}_k)}{n!(f^{(1)}(\hat{x}_k))^{n+1}}$.

From $m = n-1$ and the definition of the algorithm, we have

$$x_{k+1} = \Phi_m(x_k) = g_m(x_k) + h_m(f^{(m)}(x_k)), \tag{72}$$

where

$$\begin{aligned} h_m(f^{(m)}(x_k)) &= \hat{h}_n\left(\frac{f^{(n-1)}(x_k) - f^{(n-1)}(z_k)}{x_k - z_k}\right) \\ &= (-1)^{n+1}\frac{f^{n-1}(x_k)(f^{(n-1)}(x_k) - f^{(n-1)}(y_k))}{n!(f^{(1)}(x_k))^n}. \end{aligned}$$

From [7], we know the order of convergence of $\hat{\Phi}_n(\hat{x}_k)$ is at least $(n+1)$. From Lemma 4.5 and (71), we have

$$\begin{aligned} \hat{\Phi}_n(\hat{x}_k) - \alpha &= (-1)^{n+1}\frac{f^{(n)}(\hat{x}_k)f^n(\hat{x}_k)}{n!(f^{(1)}(\hat{x}_k))^{n+1}} + g_{n-1}(\hat{x}_k) - \alpha \\ &= (-1)^{n+1}(A_n\hat{e}_k^n + \hat{B}_n\hat{e}_k^{n+1}) + o(\hat{e}_k^{n+2}) + g_{n-1}(\hat{x}_k) - \alpha \\ &= K_n\hat{e}_k^{n+1} + o(\hat{e}_k^{n+2}), \end{aligned}$$

where $K_n \in IR$ is a constant. Thus,

$$\hat{g}_{n-1}(\hat{x}_k) - \alpha = -(-1)^{n+1}A_n\hat{e}_k^n + (K_n - (-1)^{n+1}\hat{B}_n)\hat{e}_k^{n+1} + o(\hat{e}_k^{n+2}). \tag{73}$$

Note that $g_m(.) = \hat{g}_{n-1}(.)$, we have

$$g_m(x_k) - \alpha = -(-1)^{n+1}A_n e_k^n + (K_n - (-1)^{n+1}\hat{B}_n)e_k^{n+1} + o(e_k^{n+2}). \tag{74}$$

Substituting (74) and (63) into (72), we get

$$\begin{aligned} \Phi_m(x_k) - \alpha &= (-1)^{n+1}\frac{f^{n-1}(x_k)(f^{(n-1)}(x_k) - f^{(n-1)}(y_k))}{n!(f^{(1)}(x_k))^n} + g_m(x_k) - \alpha \\ &= (-1)^{n+1}(A_n e_k^n + B_n e_k^{n+1}) - (-1)^{n+1}A_n e_k^n \\ &\quad + (K_n - (-1)^{n+1}\hat{B}_n)e_k^{n+1} + o(e_k^{n+2}) \\ &= [K_n + (-1)^{n+1}(B_n - \hat{B}_n)]e_k^{n+1} + o(e_k^{n+2}). \end{aligned} \tag{75}$$

Since $m = n-1$, we finally have

$$\Phi_m(x_k) - \alpha = [K_{m+1} + (-1)^{m+2}(B_{m+1} - \hat{B}_{m+1})]e_k^{m+2} + o(e_k^{m+3}).$$

Thus, the order of convergence of $\Phi_m(x_k)$ is $(m+2)$. □

5 Numerical Results

We implemented the new method in Matlab and tested it on five numerical examples. We also compared the results obtained by the method presented in [2] and [7], using the same stopping criterion $|f(x^k)| \le 10^{-10}$.

For each example, we chose three different starting points, which ensured that the solution stayed the same for the same example for whatever method used. Therefore, we compared these methods by evaluating the number of iterations required to perform them. The following are the test functions, and the solution α found by those methods:

Example 1: $f(x) = xe^x + 2e^x - 1$, $\alpha = -0.4428544009$
Example 2: $f(x) = x^7 + 2x^5 + 3x^3 + x^2 + x + 1$, $\alpha = -0.5841144224$
Example 3: $f(x) = xe^{x^2} - \sin^2 x + 3\cos x + 5$, $\alpha = -1.2076478271$
Example 4: $f(x) = x^2 - e^x - 3x + 2$, $\alpha = +0.2575302854$
Example 5: $f(x) = -0.5x^7 + 0.1x^5 + 10x^3 - 10x^2 - 70x - 7$, $\alpha = -0.1016253384$

It is known that not only the number of iterations counts, but also the cost of each iteration counts. In the sense of Ostrowski [8], its asymptotic efficiency index is equal to $(m+2)^{1/c}$, where c is the number of arithmetic operations per iteration. Since for complicated functions the number p of function evaluations per iteration dominates the other arithmetic operations, it is reasonable to take $c = p$. The efficiency index of the new method equals $(m+2)^{1/(m+2)}$.

The iteration formulae of the Traub's method [2] for given x_k are as follows.

$$x_k^0 = x_k; \qquad x_k^{j+1} = x_k^j - f(x_k^j)/f^{(1)}(x_k), \quad j = 0, \dots, m-1; \qquad x_{k+1} = x_k^m.$$

It has an order of $(m+1)$ and requires $(m+1)$ functions per iteration. Hence its efficiency index is $(m+1)^{1/(m+1)}$. Note that our method uses higher derivatives, which in some cases may be difficult to compute.

Tables 1 through 5 show that for each example, respectively, the number of iterations (listed under column "I") is performed by the methods starting from three initial points x_0. Column "F" indicates the total number of function and derivative evaluations. The methods in the tables are denoted as follows:

"Previous method with $m = 1$" is the Newton method;
"Previous method with $m = 2$" is the Chebyshev method (defined by (15));
"Previous method with $m = 3$" is defined by (16);
"New method with $m = 1$" is defined by (20) and (24);
"New method with $m = 2$" is defined by (20) and (25);
"New method with $m = 3$" is defined by (20) and (22) when $m = 3$.

The column "Traub" is the Traub method described.

Tables 1 to 5 indicate that the new method has a higher convergent rate than the previous method, when using the same highest order derivative. Table 5 suggests

Table 1 Results for Example 1

Initial	$x_0 = 2$						$x_0 = 4$						$x_0 = 6$					
Meth	Prev		New		Traub		Prev		New		Traub		Prev		New		Traub	
	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
$m = 1$	8	16	6	18	8	16	10	20	7	21	10	20	12	24	9	27	12	24
$m = 2$	5	15	4	16	5	15	7	21	6	24	7	21	8	24	7	28	9	27
$m = 3$	4	16	4	20	5	20	6	24	5	25	6	24	7	28	6	30	7	28

Table 2 Results for Example 2

Initial	$x_0 = -5$						$x_0 = -2$						$x_0 = -1$					
Meth	Prev		New		Traub		Prev		New		Traub		Prev		New		Traub	
	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
$m = 1$	15	30	11	33	15	30	10	20	7	21	10	20	6	12	4	12	6	12
$m = 2$	10	30	10	40	11	33	7	21	6	24	7	21	4	12	4	16	4	12
$m = 3$	9	36	8	40	9	36	5	20	5	25	6	24	4	16	3	15	4	16

Table 3 Results for Example 3

Initial	$x_0 = -5$						$x_0 = -3$						$x_0 = -1$					
Meth	Prev		New		Traub		Prev		New		Traub		Prev		New		Traub	
	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
$m = 1$	30	60	23	69	30	60	13	26	10	30	13	26	5	10	4	12	5	10
$m = 2$	20	60	17	68	22	66	9	27	8	32	10	30	3	9	3	12	4	12
$m = 3$	16	64	14	70	18	72	7	28	6	30	8	32	3	12	3	15	3	12

Table 4 Results for Example 4

Initial	$x_0 = 0$						$x_0 = 2$						$x_0 = 7$					
Meth	Prev		New		Traub		Prev		New		Traub		Prev		New		Traub	
	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
$m = 1$	3	6	2	6	3	6	4	8	4	12	4	8	9	18	7	21	9	18
$m = 2$	2	6	2	8	2	6	3	9	3	12	3	9	7	21	6	24	7	21
$m = 3$	2	8	2	10	2	8	3	12	3	15	3	12	6	24	5	25	6	24

that the new method, when using a derivative with one order lower, could have a faster convergence rate for some problems. This can be seen from the convergence coefficients in Lemmas 4.1 and 4.2.

However, if we take the total number of function evaluation as our criterion, the new method is comparable to the Traub's method. The comparison of the efficiency indices reveals that the maximum is achieved at $m = 2$ for the Traub's method and

Table 5 Results for Example 5

Initial	$x_0 = -9$						$x_0 = 2.5$						$x_0 = 4.5$					
Meth	Prev		New		Traub		Prev		New		Traub		Prev		New		Traub	
	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
$m = 1$	20	40	14	42	20	40	7	14	5	15	7	14	10	20	7	21	10	20
$m = 2$	73	219	32	128	132	396	23	69	5	20	5	15	8	24	6	24	7	21
$m = 3$	61	244	19	95	10	40	86	344	63	315	4	16	7	28	5	25	6	24

at $m = 1$ for our method. We can also compare the efficiency indices for some other methods. There are two methods in [9]. The first one has the order of convergence 3 and requires 3 function evaluations per step, and its efficiency index is $3^{1/3} = 1.4422$. The second one with the order $1 + \sqrt{2}$ requires only two function evaluations per step, and its efficiency index is $(1 + \sqrt{2})^{1/2} = 1.5538$. In [10], the two methods have the efficiency indices 1.5538 and 1.67, respectively. In [11], the order is 1.839 and only one function estimation per step is required, and its efficiency index is 1.839. In [12], the efficiency index of several iterative procedures for solving nonlinear equations in Banach spaces is discussed.

6 Concluding Remarks

We have proposed in this paper a new algorithm for finding the simple root of a nonlinear equation. Not only have we given its motivation, but also we have proved that the order of convergence of the new method is $(m + 2)$, where m is the highest order derivative used in the iteration formula. Its proofs are explained in stages for better understanding. Before the main result is presented, the case $m = 1$ is analyzed in details to motivate the convergence analysis. It also uses the results of a previous paper, where the order of convergence is $(m + 1)$. Numerical examples are used to demonstrate the new algorithm with the same stopping criteria as some previous papers. The limited numerical experience suggests that the new method would be a valuable alternative for solving such a problem.

References

1. Noor, M.A., Khan, W.A., Hussain, A.: A new modified Halley method without second derivatives for nonlinear equation. Appl. Math. Comput. (2007)
2. Traub, J.: Iterative Methods for the Solution of Equations. Prentice-Hall, Englewood Cliffs (1964)
3. Bateman, H.: Haley's method for solving equations. Am. Math. Mon. **45**(1), 11–17 (1938)
4. Chun, C.: Certain improvement of Chebyshev-Halley methods with accelerated fourth-order convergence. Appl. Math. Comput. (2007)
5. Frontini, M., Sormani, E.: Some variant of Newton's method with third-order convergence. Appl. Math. Comput. **140**, 419–426 (2003)

6. Chun, C.: Iterative methods improving Newton's method by the decomposition method. Comput. Math. Appl. **50**, 1559–1568 (2005)
7. Germani, A., Manes, C., Palumbo, P.: Higher-Order method for the solution of a nonlinear scalar equation. J. Optim. Theory Appl. **131**, 347–364 (2006)
8. Ostrowski, A.M.: Solution of Equations in Euclidean and Banach Spaces. Academic Press, New York (1973)
9. Alefeld, G., Potra, F.A.: On two higher order enclosing methods of J.W. Schmidt. Z. Angew. Math. Mech. **68**(8), 331–337 (1988)
10. Alefeld, G., Potra, F.A., Shi, Y.: On enclosing simple roots of nonlinear equations. Math. Comput. **61**(204), 733–744 (1993)
11. Potra, F.A.: On an iterative algorithm of order 1.839... for solving nonlinear operator equations. Numer. Funct. Anal. Optim. **7**(1), 75–106 (1985)
12. Potra, F.A., Ptak, V.: Nondiscrete Induction and Iterative Processes. Research Notes in Mathematics, vol. 103. Wiley, Boston (1984)