

Article

Zero-Cost Enhancement of the Steffensen Method for Nonlinear Equations

Avram Sidi 

Computer Science Department, Technion-Israel Institute of Technology, Haifa 32000, Israel; asidi@cs.technion.ac.il

Abstract

The Steffensen method is a very effective root-finding procedure used for solving nonlinear equations of the form $f(x) = 0$ numerically. When the root of $f(x)$ is simple, this method converges with order 2. Following a critical review of its convergence analysis, in this work, we provide a modification with memory of the Steffensen method, which achieves an order substantially higher than 2, without increasing the number of function evaluations per iteration. Specifically, with an arbitrary fixed integer $k \geq 1$ and with initial approximations $x_0, x_1, \dots, x_k$, the modified method generates a sequence of approximations to the solution of $f(x) = 0$ via $x_{n+1} = x_n - f(x_n)/f[c_n, x_n]$, $n = k, k+1, \dots$, where $c_n = x_n - f(x_n)/p'_{n,k}(x_n)$ and $p_{n,k}(x)$ is the polynomial of interpolation to $f(x)$ at the points $x_n, x_{n-1}, \dots, x_{n-k}$. Just as in the original method, the modified method makes use of only $f(x)$; no derivatives of $f(x)$ are needed. We prove that the order of this method is $1 + \sqrt{2} \doteq 2.414$ for $k = 1$ and it can be increased towards $(3 + \sqrt{5})/2 \doteq 2.618$ by increasing k . (For example, the order of the method is $2.617 \dots$ already for $k = 6$.)

Keywords: nonlinear equations; Steffensen method; secant method; order of convergence

MSC: 65H04; 65H05

1. Introduction and Background

Let α be the solution to the equation

$$f(x) = 0, \quad (1)$$

and assume that $f(x)$ is twice continuously differentiable in a closed interval I containing α in its interior. A very effective iterative method used for solving (1) when α is a simple zero of $f(x)$ is the Steffensen method. Starting from an initial approximation x_0 , this method computes a sequence of approximations $\{x_n\}_{n=0}^{\infty}$ to α , such that x_{n+1} is determined from x_n via

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}}, \quad n = 0, 1, \dots \quad (2)$$

Thus, at each iteration, two evaluations of $f(x)$ are performed. No derivatives of $f(x)$ are computed. When α is a simple zero of $f(x)$, that is, when $f'(\alpha) \neq 0$, which we assume throughout this work, the sequence $\{x_n\}_{n=0}^{\infty}$ converges quadratically.

Starting with the Steffensen method, in this work, we aim to develop a new class of iterative methods with memory for solving nonlinear equations with a simple root. In



Academic Editors: Chuanjun Chen,
Jiansong Zhang and Carlo Bianca

Received: 28 February 2026

Revised: 15 June 2026

Accepted: 18 June 2026

Published: 7 July 2026

Copyright: © 2026 by the author.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Section 2, we give a brief review of the essentials of the Steffensen method, especially of its convergence analysis. Based on this review, in Section 3, we propose to modify the Steffensen method and construct a sequence of iterative methods with memory. These methods:

- require only two evaluations of $f(x)$ per iteration, just like the Steffensen method,
- require no evaluations of derivatives of $f(x)$, and
- have order $1 + \sqrt{2} \doteq 2.414$ at minimum, which increases quickly towards $(3 + \sqrt{5})/2 \doteq 2.618$. (This should be compared with the order of Steffensen's method, which is 2.)

In this sense, the cost per iteration of each of our methods is the same as that of the Steffensen method, while its order and efficiency index are higher than those of the Steffensen method.

We present a complete and detailed convergence theory for the new methods in Section 4 that is supplemented in Appendix A. In Section 5, we present three numerical examples that validate the convergence theory. We emphasize that we do not address the issues of multiple roots, global convergence, or related basins of attraction.

The Steffensen method and others are discussed in many books on numerical analysis—see, for example, Atkinson [1], Dahlquist and Björck [2], Henrici [3], Ralston and Rabinowitz [4], Stoer and Bulirsch [5], and Traub [6]; see also Sidi [7].

We also mention that there are quite a few methods with memory in the literature that generalize Steffensen's method and that achieve higher accuracies and efficiency indices—see, for example, Cordero and Torregrosa [8,9], Džunić [10], Džunić and Petković [11], Liu et al. [12], Ren and Bi [13], and Wang and Zhang [14]; see also the book [15], the survey [16] by Petković et al., and the recent work [17] by Cordero et al.

2. Review of the Steffensen Method

2.1. The Steffensen Method as a Special Case of a General Iterative Method

For future use, let us consider iterative methods for the equation $f(x) = 0$ that generate sequences of approximations $\{x_n\}_{n=0}^{\infty}$ to the solution α of this equation via

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(c_n) - f(x_n)}{c_n - x_n}}, \quad \text{for some appropriate } c_n. \quad (3)$$

[Comparing (2) with (3), it is easy to verify that x_{n+1} in the Steffensen method is generated by (3) by choosing $c_n = x_n + f(x_n)$ in the latter.] Subtracting α from both sides of (3), and making use of the facts that

$$\frac{f(c_n) - f(x_n)}{c_n - x_n} = f[c_n, x_n] \quad \text{and} \quad f(x_n) = f(x_n) - f(\alpha) = f[x_n, \alpha](x_n - \alpha), \quad (4)$$

where $f[z_0, z_1, \dots, z_m]$ denotes the m th order divided difference of $f(x)$ over the set of points $\{z_0, z_1, \dots, z_m\}$, we obtain

$$x_{n+1} - \alpha = \frac{f[c_n, x_n] - f[x_n, \alpha]}{f[c_n, x_n]}(x_n - \alpha) = \frac{f[c_n, x_n, \alpha]}{f[c_n, x_n]}(c_n - \alpha)(x_n - \alpha). \quad (5)$$

Here, we have used the fact that divided differences are symmetric functions of their arguments and that

$$f[u, v, w] = \frac{f[u, v] - f[u, w]}{v - w}. \quad (6)$$

The relation given in (5) is actually the starting point of a unified analytic treatment of several methods that are as described in this paragraph. For this, we make use of the properties of divided differences; these are summarized in [7].

2.2. Analysis of the Steffensen Method

Going back to the Steffensen method with $c_n = x_n + f(x_n)$, we have

$$c_n - \alpha = x_n - \alpha + f(x_n) = (1 + f[x_n, \alpha])(x_n - \alpha), \quad (7)$$

where we have invoked (4). Thus, (5) gives

$$x_{n+1} - \alpha = \frac{f[c_n, x_n, \alpha]}{f[c_n, x_n]}(1 + f[x_n, \alpha])(x_n - \alpha)^2. \quad (8)$$

The rest of the analysis of the Steffensen method follows by letting $n \rightarrow \infty$ and making use of the fact that $x_n, c_n \rightarrow \alpha$ provided x_0 is sufficiently close to α . We then have

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\alpha)}{2f'(\alpha)}[1 + f'(\alpha)] \quad \text{when } f \in C^2(\alpha - \rho, \alpha + \rho) \text{ for some } \rho > 0, \quad (9)$$

provided $f'(\alpha) \neq 0$, which we assume throughout this note. Here, we have made use of the facts that

$$f[z_0, z_1, \dots, z_m] = \frac{f^{(m)}(\hat{z})}{m!} \quad \text{for some } \hat{z} \in (\min\{z_0, z_1, \dots, z_m\}, \max\{z_0, z_1, \dots, z_m\}),$$

assuming $f \in C^m[a, b]$ and $\{z_0, z_1, \dots, z_m\} \in [a, b]$. (10)

From this, it is clear that the method has an order of convergence 2 when $f''(\alpha) \neq 0$ and $f'(\alpha) \neq -1$.

2.3. Can the Steffensen Method Be Enhanced?

Now, if $f'(\alpha) = -1$, the right-hand side of (9) vanishes and the method achieves order higher than 2. The question then is whether we can somehow achieve this *at zero additional cost*. Here, we describe the approach taken in Traub [6] (pp. 185–187) for dealing with this question.

The most direct way of achieving this is to apply the Steffensen method to the equation

$$F(x) = 0; \quad F(x) = \mu f(x), \quad \mu = -\frac{1}{f'(\alpha)}, \quad (11)$$

since $F'(\alpha) = -1$. With this, (2) becomes

$$x_{n+1} = x_n - \frac{F(x_n)}{\frac{F(x_n + F(x_n)) - F(x_n)}{F(x_n)}} = x_n - \frac{f(x_n)}{\frac{f(x_n + \mu f(x_n)) - f(x_n)}{\mu f(x_n)}}. \quad (12)$$

This, of course, is impractical simply because $f'(\alpha)$ is not available. We can, however, replace $f'(\alpha)$ by suitable *approximations*. In particular, we can achieve this by replacing μ in (11) and (12) by μ_n , which has the property that

$$\lim_{n \rightarrow \infty} \mu_n = \mu = -\frac{1}{f'(\alpha)}. \quad (13)$$

This seems to have a good chance of improving the order of the method. Let us also note that (12) (with μ replaced by μ_n) can be expressed as in (3) as follows:

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(c_n) - f(x_n)}{c_n - x_n}}, \quad c_n = x_n + \mu_n f(x_n). \quad (14)$$

Needless to say, the error expression given in (5) remains valid. Two preliminary such choices of $\{\mu_n\}$ are discussed in the next section. In Section 4, we describe a new choice of $\{\mu_n\}$ that generalizes one of two choices described in the next section; this results in a sequence of methods whose orders can be improved towards $(3 + \sqrt{5})/2 \doteq 2.618$, which is substantially larger than 2. The cost per iteration of these methods is the same as that of the Steffensen method, namely, two evaluations of $f(x)$ per iteration.

3. Preliminary Improvement of the Steffensen Method

3.1. An Expensive Choice: $\mu_n = -1/f'(x_n)$

To motivate the approach we take later, we first describe a method that is *not* zero-cost and was suggested in Traub [6] (p. 180, Equations (8)–(55)).

Taking into account the fact that the Steffensen method converges, we first choose $\mu_n = -1/f'(x_n)$, which clearly satisfies (13). This choice of μ_n is not what we are looking for since the resulting method requires, at each iteration step, one evaluation of $f'(x)$, in addition to the two evaluations of $f(x)$. Nevertheless, we show shortly that this will direct us to the right path. With this choice of μ_n ,

$$c_n = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (15)$$

thus, c_n is the result of applying one step of the Newton–Raphson method to the equation $f(x) = 0$, starting with the approximation x_n to α . Subtracting α from both sides of (15) and invoking (4) gives

$$c_n - \alpha = (x_n - \alpha) - \frac{f[x_n, \alpha](x_n - \alpha)}{f'(x_n)} = \left(1 - \frac{f[x_n, \alpha]}{f'(x_n)}\right)(x_n - \alpha). \quad (16)$$

Next, making use of the fact that $f'(z) = f[z, z]$ and of (6), we have

$$c_n - \alpha = \frac{f[x_n, x_n] - f[x_n, \alpha]}{f'(x_n)}(x_n - \alpha) = \frac{f[x_n, x_n, \alpha]}{f'(x_n)}(x_n - \alpha)^2. \quad (17)$$

Substituting this into (5), we obtain

$$x_{n+1} - \alpha = \frac{f[c_n, x_n, \alpha]}{f[c_n, x_n]} \frac{f[x_n, x_n, \alpha]}{f'(x_n)} (x_n - \alpha)^3. \quad (18)$$

Dividing both sides of (18) by $(x_n - \alpha)^3$ and letting $n \rightarrow \infty$ (and hence $x_n, c_n \rightarrow \alpha$), we finally obtain

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^3} = \left[\frac{f''(\alpha)}{2f'(\alpha)} \right]^2, \quad (19)$$

which shows that this method has order 3. It is clear that the cost per iteration of the method we have just described is two evaluations of $f(x)$ and one evaluation of $f'(x)$.

3.2. A Preliminary Zero-Cost Choice: $\mu_n = -1/f[x_n, x_{n-1}]$

The choice $\mu_n = -1/f'(x_n)$ that we made in the preceding subsection is actually very effective in that it increases the order of the Steffensen method from 2 to 3. It has the drawback that it requires the evaluation of $f'(x)$ once every iteration. Its success suggests to us that we could improve the order of the Steffensen method by using a *zero-cost approximation to $f'(x_n)$* instead of $f'(x_n)$ itself. We can achieve this *at zero additional cost*, for example, by letting

$$\mu_n = -\frac{1}{f[x_n, x_{n-1}]} = -\frac{1}{\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}}, \quad (20)$$

since $f(x_n)$ is already needed (hence computed) in (14) and $f(x_{n-1})$ is available from the previous iteration step, and $\lim_{n \rightarrow \infty} f[x_n, x_{n-1}] = f[\alpha, \alpha] = f'(\alpha)$, assuming that the sequence $\{x_n\}_{n=0}^{\infty}$ converges to α .

With this choice of μ_n ,

$$c_n = x_n - \frac{f(x_n)}{f[x_n, x_{n-1}]}, \quad (21)$$

thus, c_n is the result of applying one step of the secant method to the equation $f(x) = 0$, starting with the approximations x_{n-1} and x_n to α . Upon subtracting α from both sides of (21) and invoking (4), we obtain

$$c_n - \alpha = (x_n - \alpha) - \frac{f[x_n, \alpha](x_n - \alpha)}{f[x_n, x_{n-1}]} = \left(1 - \frac{f[x_n, \alpha]}{f[x_n, x_{n-1}]}\right)(x_n - \alpha). \quad (22)$$

Next, making use of (6), we have

$$c_n - \alpha = \frac{f[x_n, x_{n-1}] - f[x_n, \alpha]}{f[x_n, x_{n-1}]}(x_n - \alpha) = \frac{f[x_n, x_{n-1}, \alpha]}{f[x_n, x_{n-1}]}(x_n - \alpha)(x_{n-1} - \alpha). \quad (23)$$

Substituting this into (5), we obtain

$$x_{n+1} - \alpha = \frac{f[c_n, x_n, \alpha]}{f[c_n, x_n]} \frac{f[x_n, x_{n-1}, \alpha]}{f[x_n, x_{n-1}]} (x_n - \alpha)^2 (x_{n-1} - \alpha). \quad (24)$$

Dividing both sides of (24) by $(x_n - \alpha)^2 (x_{n-1} - \alpha)$ and letting $n \rightarrow \infty$ (and hence $x_n, c_n \rightarrow \alpha$), we finally obtain

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2 (x_{n-1} - \alpha)} = \left[\frac{f''(\alpha)}{2f'(\alpha)} \right]^2 \equiv L, \quad (25)$$

which shows that this method has order s_1 with $2 < s_1 < 3$. Indeed, letting

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^{s_1}} = Q, \quad (26)$$

it can be shown that s_1 is the positive solution of the polynomial equation $s^2 - 2s - 1 = 0$, namely, $s_1 = 1 + \sqrt{2} \doteq 2.414$, and that $Q = |L|^{\sqrt{2}}$.

Note that we have not justified these convergence results rigorously here because they follow from the complete analysis given in the next section, where we generalize $\mu_n = -1/f[x_n, x_{n-1}]$ to something more sophisticated. That is, the convergence results stated here are special cases of those in the next section.

Before we go on, we mention that a zero-cost method that is similar (but not identical) to the method we have just described is treated in Traub [6] (pp. 185–187).

4. SSteff(n, k): General Enhancement of the Steffensen Method

4.1. Generalization of $\mu_n = -1/f[x_n, x_{n-1}]$

So far, we have observed that (i) by choosing $\mu_n = -1/f'(x_n)$ in (14), we can increase the order of the Steffensen method from 2 to 3, and (ii) by choosing $\mu_n = -1/f[x_n, x_{n-1}]$ in (14), we can increase the order of the Steffensen method from 2 to $1 + \sqrt{2} \doteq 2.414$. The reason the second choice achieves an inferior improvement compared to the first is that we are only approximating $f'(x_n)$ (in a crude fashion) by the derivative of the linear interpolant to $f(x)$ at x_n and x_{n-1} . Thus, it seems reasonable to conclude, if we can improve the approximation to $f'(x_n)$, we may be able to improve the order of the resulting method further, pushing it closer to 3. We do this by approximating $f'(x_n)$ by $p'_{n,k}(x_n)$, where $p_{n,k}(x)$ is the polynomial of interpolation (of degree at most k) to $f(x)$ at the points $x_n, x_{n-1}, \dots, x_{n-k}$, for some arbitrary integer k that is at our disposal. Thus, expressing $p_{n,k}(x)$ via Newton's interpolation formula, that is,

$$p_{n,k}(x) = f(x_n) + \sum_{i=1}^k f[x_n, x_{n-1}, \dots, x_{n-i}] \prod_{j=0}^{i-1} (x - x_{n-j}), \quad (27)$$

we have

$$\mu_n = -\frac{1}{p'_{n,k}(x_n)}; \quad p'_{n,k}(x_n) = f[x_n, x_{n-1}] + \sum_{i=2}^k f[x_n, x_{n-1}, \dots, x_{n-i}] \prod_{j=1}^{i-1} (x_n - x_{n-j}). \quad (28)$$

For the efficient computation of $p'_{n,k}(x_n)$, see [18], where the idea of approximating $f'(x_n)$ by $p'_{n,k}(x_n)$, where $p_{n,k}(x)$ is the polynomial of interpolation (of degree at most k) to $f(x)$ at the points $x_n, x_{n-1}, \dots, x_{n-k}$, was introduced by the author originally.

Observe that we need no extra computation of $f(x)$ for the present choice of μ_n since the values of $f(x)$ required for determining $p'_{n,k}(x_n)$ are (i) $f(x_n)$, which needs to be computed anyway, and (ii) $f(x_{n-i})$, $i = 1, \dots, k$, which are already available from previous iterations. Of course, we do not need to compute any derivatives of $f(x)$. It is clear that the computation of $p'_{n,k}(x_n)$ as presented in (28) can be achieved via a small number of arithmetic operations depending on k . Thus, the cost per iteration of the proposed method is practically two evaluations of $f(x)$, just as that of the Steffensen method itself. Clearly, for $k = 1$, we have $p'_{n,1}(x_n) = f[x_n, x_{n-1}]$ and we are back at the case described in Section 3.2.

With the choice $\mu_n = -1/p'_{n,k}(x_n)$, we have, for each fixed $k \in \{1, 2, \dots\}$,

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(c_n) - f(x_n)}{c_n - x_n}}, \quad c_n = x_n - \frac{f(x_n)}{p'_{n,k}(x_n)}, \quad (29)$$

and hence our method now is described via

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(x_n - f(x_n)/p'_{n,k}(x_n)) - f(x_n)}{-f(x_n)/p'_{n,k}(x_n)}}, \quad n = k, k+1, \dots, \quad (30)$$

with appropriate initial approximations $x_0, x_1, \dots, x_k$ to α . For convenience, let us denote this method (of computing x_{n+1} from $x_n, x_{n-1}, \dots, x_{n-k}$) by $SSteff(n, k)$ when $n \geq k$. As we show in Theorem 1, the method $SSteff(n, k)$ has an order that can be driven towards $(3 + \sqrt{5})/2 \doteq 2.618$ by increasing k .

A critical issue that arises immediately is that of generating systematically the initial approximations $x_1, x_2, \dots, x_k$, instead of choosing them arbitrarily. Choosing x_0 first, these can be generated one by one in the order listed as follows:

Determine x_1 via the Steffensen method in (2). (Note that x_1 can also be assigned any reasonable value. For example, we can choose x_1 close to x_0 but arbitrary otherwise.)

Determine x_2 applying SSteff(1,1) with x_0, x_1 as initial approximations.

Determine x_3 applying SSteff(2,2) with x_0, x_1, x_2 as initial approximations.

.....

Determine x_k applying SSteff($k-1, k-1$) with $x_0, x_1, x_2, \dots, x_{k-1}$ as initial approximations.

(From here on, continue with SSteff(n, k) and generate x_{n+1} , $n = k, k+1, k+2, \dots$)

Clearly, $x_1, x_2, \dots, x_k$ are *not* arbitrary initial values; they are true approximations to the root α as they are obtained by using valid root-finding methods, namely, Steffensen's method to determine x_1 , and SSteff(j, j) to determine x_{j+1} , $j = 1, 2, \dots, k-1$. In addition, it is clear that, for each $j = 0, 1, \dots, k-1$, x_{j+1} is likely to be a better approximation to α than x_j .

Remark 1. Precisely, this approach was developed originally in Sidi [18] to generalize the secant method to improve its order from $(1 + \sqrt{5})/2 \doteq 1.618$ towards 2 at zero additional cost. The application of this generalized secant method to complex roots has been described recently in Sidi [19].

4.2. Analysis of Convergence of SSteff(n, k)

Theorem 1. Let α be the solution to the equation $f(x) = 0$. Assume $f \in C^{k+1}(I)$, where I is a closed interval containing α in its interior, and assume also that α is simple, hence $f'(\alpha) \neq 0$, in addition to $f(\alpha) = 0$. Let $x_0, x_1, \dots, x_k$ be distinct initial approximations to α , and generate via (30) the approximations x_{n+1} , $n = k, k+1, \dots$, where $p_{n,k}(x)$ is the polynomial of interpolation to $f(x)$ at the points $x_n, x_{n-1}, \dots, x_{n-k}$, and hence $p'_{n,k}(x_n)$ is as given in (28). Then, provided $x_0, x_1, \dots, x_k$ are in I and sufficiently close to α , the sequence $\{x_n\}_{n=0}^\infty$ converges to α , and

$$\lim_{n \rightarrow \infty} \frac{\epsilon_{n+1}}{\epsilon_n^2 \prod_{i=1}^k \epsilon_{n-i}} = \frac{(-1)^{k+1}}{2! (k+1)!} \frac{f''(\alpha) f^{(k+1)}(\alpha)}{[f'(\alpha)]^2} \equiv L; \quad \epsilon_m \equiv x_m - \alpha \quad \forall m. \quad (31)$$

The order of convergence s_k is the only positive root of the polynomial

$$g_k(s) = s^{k+1} - 2s^k - \sum_{i=0}^{k-1} s^i, \quad (32)$$

and $2 < s_k < 3$. Actually,

$$1 + \sqrt{2} = s_1 < s_2 < s_3 < \dots; \quad \lim_{k \rightarrow \infty} s_k = \hat{s} = \frac{3 + \sqrt{5}}{2} = \left(\frac{1 + \sqrt{5}}{2} \right)^2 = 2.618 \dots, \quad (33)$$

and

$$\lim_{n \rightarrow \infty} \frac{|\epsilon_{n+1}|}{|\epsilon_n|^{s_k}} = |L|^{(k+1)/(s_k-1)} \Rightarrow \lim_{n \rightarrow \infty} \frac{\log |\epsilon_{n+1}/\epsilon_n|}{\log |\epsilon_n/\epsilon_{n-1}|} = s_k. \quad (34)$$

Remark 2. Note that, in Theorem 1,

$$s_1 \doteq 2.414, s_2 \doteq 2.547, s_3 \doteq 2.592, s_4 \doteq 2.608, s_5 \doteq 2.614, s_6 \doteq 2.617, \dots,$$

rounded to four significant figures. Thus, already with $k = 6$, we are almost achieving the highest order of convergence possible.

Proof. Below, we use the short-hand notation

$$\text{int}(a_1, \dots, a_m) = (\min\{a_1, \dots, a_m\}, \max\{a_1, \dots, a_m\}).$$

Analysis of c_n in (29):

We start by deriving a closed-form expression for $c_n - \alpha$. Subtracting α from both sides of (29), and noting that

$$f(x_n) = f(x_n) - f(\alpha) = f[x_n, \alpha](x_n - \alpha),$$

we have

$$c_n - \alpha = \left(1 - \frac{f[x_n, \alpha]}{p'_{n,k}(x_n)}\right)(x_n - \alpha) = \frac{p'_{n,k}(x_n) - f[x_n, \alpha]}{p'_{n,k}(x_n)}(x_n - \alpha). \quad (35)$$

We now split $p'_{n,k}(x_n) - f[x_n, \alpha]$ as in

$$p'_{n,k}(x_n) - f[x_n, \alpha] = \{p'_{n,k}(x_n) - f'(x_n)\} + \{f'(x_n) - f[x_n, \alpha]\}. \quad (36)$$

Invoking

$$f'(x_n) - f[x_n, \alpha] = f[x_n, x_n] - f[x_n, \alpha] = f[x_n, x_n, \alpha](x_n - \alpha) = \frac{f''(\eta_n)}{2!}(x_n - \alpha)$$

for some $\eta_n \in \text{int}(x_n, \alpha)$, and

$$\begin{aligned} f'(x_n) - p'_{n,k}(x_n) &= f[x_n, x_n, x_{n-1}, \dots, x_{n-k}] \prod_{i=1}^k (x_n - x_{n-i}) \\ &= \frac{f^{(k+1)}(\xi_n)}{(k+1)!} \prod_{i=1}^k (x_n - x_{n-i}) \end{aligned} \quad (37)$$

for some $\xi_n \in \text{int}(x_n, x_{n-1}, \dots, x_{n-k})$, (36) becomes

$$p'_{n,k}(x_n) - f[x_n, \alpha] = -\frac{f^{(k+1)}(\xi_n)}{(k+1)!} \prod_{i=1}^k (\epsilon_n - \epsilon_{n-i}) + \frac{f''(\eta_n)}{2!} \epsilon_n. \quad (38)$$

Substituting (37) and (38) into (35), and letting

$$\hat{D}_n = -\frac{f^{(k+1)}(\xi_n)}{(k+1)!} \quad \text{and} \quad \hat{E}_n = \frac{f''(\eta_n)}{2!}, \quad (39)$$

we finally obtain

$$c_n - \alpha = U_n \epsilon_n; \quad U_n \equiv \frac{p'_{n,k}(x_n) - f[x_n, \alpha]}{p'_{n,k}(x_n)} = \frac{\hat{D}_n \prod_{i=1}^k (\epsilon_n - \epsilon_{n-i}) + \hat{E}_n \epsilon_n}{f'(x_n) + \hat{D}_n \prod_{i=1}^k (\epsilon_n - \epsilon_{n-i})}. \quad (40)$$

We now prove that c_n is closer to α than x_n is, provided $x_n, x_{n-1}, \dots, x_{n-k}$ are all sufficiently close to α . Without loss of generality, we can assume that $I = [\alpha - T, \alpha + T]$ for some $T > 0$, and that $m_1 = \min_{x \in I} |f'(x)| > 0$. This is possible since $\alpha \in I$ and $f'(\alpha) \neq 0$. Let $M_s = \max_{x \in I} |f^{(s)}(x)|/s!$, $s = 1, 2, \dots$, and choose an interval $J = (\alpha - t/2, \alpha + t/2)$, $J \subset I$, sufficiently small to ensure that $m_1 > 2M_{k+1}t^k + M_2t/2$. It can now be shown that, provided x_{n-i} , $i = 0, 1, \dots, k$, are all in J , there holds

$$|U_n| \leq \frac{M_{k+1} \prod_{i=1}^k |\epsilon_n - \epsilon_{n-i}| + M_2 |\epsilon_n|}{m_1 - M_{k+1} \prod_{i=1}^k |\epsilon_n - \epsilon_{n-i}|} \leq \frac{M_{k+1} \prod_{i=1}^k (|\epsilon_n| + |\epsilon_{n-i}|) + M_2 |\epsilon_n|}{m_1 - M_{k+1} \prod_{i=1}^k (|\epsilon_n| + |\epsilon_{n-i}|)} \leq \bar{U}(t),$$

where

$$\bar{U}(t) \equiv \frac{M_{k+1}t^k + M_2t/2}{m_1 - M_{k+1}t^k} < 1.$$

Consequently, by (40), $|c_n - \alpha| < |\epsilon_n|$, which implies that $c_n \in J$, just like x_{n-i} , $i = 0, 1, \dots, k$, and is closer to α than x_n is.

Analysis of ϵ_{n+1} :

We now turn to x_{n+1} . Substituting (40) into (5), we obtain the closed-form expression

$$\epsilon_{n+1} = V_n \epsilon_n; \quad V_n = \frac{f[c_n, x_n, \alpha]}{f[c_n, x_n]} \cdot \frac{p'_{n,k}(x_n) - f[x_n, \alpha]}{p'_{n,k}(x_n)} (x_n - \alpha) = \frac{f[c_n, x_n, \alpha]}{f[c_n, x_n]} U_n \epsilon_n. \quad (41)$$

Since $c_n \in J$ just as x_{n-i} , $i = 0, 1, \dots, k$, we can write

$$V_n = \frac{f''(\tau_n)}{2f'(\rho_n)} U_n \epsilon_n, \quad \text{for some } \tau_n, \rho_n \in J.$$

Therefore, we also have

$$|V_n| \leq \frac{M_2}{m_1} \bar{U}(t) \frac{t}{2} = \bar{V}(t).$$

Clearly, by shortening the interval J (by making t smaller) if necessary, we can cause $\bar{V}(t) < 1$, and this implies that x_{n+1} is also in J and closer to α than x_n is. We have thus shown that, if $x_0, x_1, \dots, x_k$ are chosen in J and J is sufficiently small, then the whole sequence $\{x_n\}_{n=0}^\infty$ is in J .

Convergence of $\{\epsilon_n\}_{n=0}^\infty$:

By the fact that $\epsilon_{n+1} = V_n \epsilon_n$ and $|V_n| \leq \bar{V}(t) < 1$ for all n , it follows that $\lim_{n \rightarrow \infty} \epsilon_n = 0$. Therefore, the sequence $\{x_n\}_{n=0}^\infty$ converges and $\lim_{n \rightarrow \infty} x_n = \alpha$. By (29), this implies that $\lim_{n \rightarrow \infty} c_n = \alpha$ as well. Next, by the fact that $\lim_{n \rightarrow \infty} x_n = \alpha$, we note that

$$\lim_{n \rightarrow \infty} p'_{n,k}(x_n) = f'(\alpha) = \lim_{n \rightarrow \infty} f[x_n, \alpha],$$

and this implies $\lim_{n \rightarrow \infty} U_n = 0$, which in turn forces $\lim_{n \rightarrow \infty} V_n = 0$. Consequently, $\lim_{n \rightarrow \infty} \epsilon_{n+1}/\epsilon_n = 0$, and this implies that $\{x_n\}_{n=0}^\infty$ converges superlinearly, that is, of order greater than 1. As a result, we also have $\lim_{n \rightarrow \infty} \epsilon_n/\epsilon_{n-i} = 0$ for every $i \geq 1$ and, consequently,

$$\epsilon_n/\epsilon_{n-i} = o(\epsilon_n/\epsilon_{n-j}) \quad \text{as } n \rightarrow \infty \quad \forall j < i. \quad (42)$$

We make use of this in the next paragraph.

Proof of (31):

Expanding in (40) the product $\prod_{i=1}^k (\epsilon_n - \epsilon_{n-i})$, we have

$$\begin{aligned} \prod_{i=1}^k (\epsilon_n - \epsilon_{n-i}) &= \prod_{i=1}^k (-\epsilon_{n-i} [1 - \epsilon_n/\epsilon_{n-i}]) \\ &= (-1)^k \left(\prod_{i=1}^k \epsilon_{n-i} \right) [1 + O(\epsilon_n/\epsilon_{n-1})] \quad \text{as } n \rightarrow \infty \quad \text{by (42)}. \end{aligned} \quad (43)$$

Substituting (43) into (40)–(41), and defining

$$D_n = \frac{f''(\tau_n)}{2f'(\rho_n)} \frac{\hat{D}_n}{p'_{n,k}(x_n)}, \quad E_n = \frac{f''(\tau_n)}{2f'(\rho_n)} \frac{\hat{E}_n}{p'_{n,k}(x_n)}, \quad (44)$$

we obtain

$$\epsilon_{n+1} = (-1)^k D_n \left(\epsilon_n^2 \prod_{i=1}^k \epsilon_{n-i} \right) [1 + O(\epsilon_n/\epsilon_{n-1})] + E_n \epsilon_n^3 \quad \text{as } n \rightarrow \infty. \quad (45)$$

Dividing both sides of (45) by $\epsilon_n^2 \prod_{i=1}^k \epsilon_{n-i}$, and defining

$$\sigma_n = \frac{\epsilon_{n+1}}{\epsilon_n^2 \prod_{i=1}^k \epsilon_{n-i}}, \quad (46)$$

we have

$$\sigma_n = (-1)^k D_n [1 + O(\epsilon_n / \epsilon_{n-1})] + E_n \sigma_{n-1} \epsilon_{n-1} \epsilon_{n-k-1} \quad \text{as } n \rightarrow \infty. \quad (47)$$

Now,

$$\lim_{n \rightarrow \infty} D_n = -\frac{1}{(k+1)!} \frac{f''(\alpha)}{2f'(\alpha)} \frac{f^{(k+1)}(\alpha)}{f'(\alpha)}, \quad \lim_{n \rightarrow \infty} E_n = \left[\frac{f''(\alpha)}{2f'(\alpha)} \right]^2. \quad (48)$$

Because (i) $\lim_{n \rightarrow \infty} D_n$ and $\lim_{n \rightarrow \infty} E_n$ exist and are finite, (ii) $\lim_{n \rightarrow \infty} (\epsilon_n / \epsilon_{n-1}) = 0$ and (iii) $\lim_{n \rightarrow \infty} \epsilon_{n-1} \epsilon_{n-k-1} = 0$, it follows that there exist a positive integer N and positive constants $\beta < 1$ and D , with $|E_n \epsilon_{n-1} \epsilon_{n-k-1}| \leq \beta$ when $n \geq N$, for which (47) gives

$$|\sigma_n| \leq D + \beta |\sigma_{n-1}| \quad \text{for all } n \geq N. \quad (49)$$

Using (49), it is easy to show that

$$|\sigma_{N+j}| \leq D \frac{1 - \beta^j}{1 - \beta} + \beta^j |\sigma_N|, \quad j = 1, 2, \dots,$$

which, by the fact that $\beta < 1$, implies that $\{\sigma_n\}_{n=N}^\infty$ is a bounded sequence. Making use of this fact, we have $\lim_{n \rightarrow \infty} E_n \sigma_{n-1} \epsilon_{n-1} \epsilon_{n-k-1} = 0$. Substituting this into (47), and invoking (48), we next obtain $\lim_{n \rightarrow \infty} \sigma_n = (-1)^k \lim_{n \rightarrow \infty} D_n = L$, which is precisely (31).

Proof of (33) (Order of method):

That the order s_k of the method is greater than 2 can be ascertained from the fact that $\lim_{n \rightarrow \infty} \epsilon_{n+1} / \epsilon_n^2 = 0$, which follows from (31). In order to prove the rest of the theorem, it suffices to consider the case $\epsilon_{n+1} = L \epsilon_n^2 \prod_{i=1}^k \epsilon_{n-i}$ for all n . First, assuming that the method converges with some order s , we must show that $\lim_{n \rightarrow \infty} (|\epsilon_{n+1}| / |\epsilon_n|^s) = Q$ indeed holds for some constants $s > 2$ and $Q > 0$. To prove the validity of this ansatz, it is sufficient to consider the simpler case of $|\epsilon_{j+1}| = Q |\epsilon_j|^s$ for all $j = 0, 1, 2, \dots$ (See Traub [6] (pp. 57–59).) For this, we express all $|\epsilon_{n-i}|$ in terms of $|\epsilon_n|$, $i = 1, 2, \dots$. Starting with

$$|\epsilon_n| = Q |\epsilon_{n-1}|^s \Rightarrow |\epsilon_{n-1}| = \frac{|\epsilon_n|^{1/s}}{Q^{1/s}},$$

$$|\epsilon_{n-1}| = Q |\epsilon_{n-2}|^s \Rightarrow |\epsilon_{n-2}| = \frac{|\epsilon_{n-1}|^{1/s}}{Q^{1/s}} = \frac{|\epsilon_n|^{1/s^2}}{Q^{1/s + 1/s^2}},$$

we obtain

$$|\epsilon_{n-i}| = \frac{|\epsilon_n|^{1/s^i}}{Q^{m_i}}, \quad m_i = \frac{1}{s} + \frac{1}{s^2} + \dots + \frac{1}{s^i}, \quad i = 1, 2, \dots$$

Substituting this into $|\epsilon_{n+1}| = |L| \epsilon_n^2 \prod_{i=1}^k |\epsilon_{n-i}|$, we obtain

$$Q |\epsilon_n|^s = |L| \epsilon_n^2 \prod_{i=1}^k \frac{|\epsilon_n|^{1/s^i}}{Q^{m_i}} = \frac{|L|}{Q^M} |\epsilon_n|^r; \quad r = 2 + \sum_{i=1}^k \frac{1}{s^i}, \quad M = \sum_{i=1}^k m_i.$$

Of course, this is possible when s satisfies $s = r$ and $Q^{M+1} = |L|$. Now, the requirement that $s = r$ is the same as $g_k(s) = 0$; hence, we have shown that the order s should be a root of the polynomial $g_k(s)$. By Theorem A1 in Appendix A to this work (with $a = 2$ there), we have that $g_k(s)$ has only one positive root s_k in the interval $(2, 3)$, with all the other roots

having moduli smaller than 1, which implies that this positive root is indeed the order of the method. Again, by Theorem A1, the sequence $\{s_k\}_{k=1}^{\infty}$ is monotonically increasing and $\hat{s}$ given in (33) is its limit and least upper bound as well. The expression given for M can be simplified considerably, as we show next.

Proof of (34):

What remains to be shown is that

$$Q = |L|^{1/\beta}, \quad \beta = M + 1 = \frac{s_k - 1}{k + 1}.$$

First, it is easy to verify that

$$M = \sum_{i=1}^k \frac{k-i+1}{s_k^i} = \frac{1}{s_k^k} \sum_{i=1}^k i s_k^{i-1}.$$

Next,

$$s_k^k M = \left(\frac{d}{ds} \sum_{i=0}^k s^i \right) \Big|_{s=s_k} = \left(\frac{d}{ds} \frac{s^{k+1} - 1}{s - 1} \right) \Big|_{s=s_k} = \frac{(k+1)s^k(s-1) - (s^{k+1} - 1)}{(s-1)^2} \Big|_{s=s_k}.$$

By $s^{k+1} - 1 = (s-1) \sum_{i=0}^k s^i$, this becomes

$$s_k^k M = \frac{(k+1)s^k - \sum_{i=0}^k s^i}{s-1} \Big|_{s=s_k} = \frac{ks_k^k - \sum_{i=0}^{k-1} s_k^i}{s_k - 1}.$$

Now, by the fact that $g_k(s_k) = 0$, we have $\sum_{i=0}^{k-1} s_k^i = s_k^{k+1} - 2s_k^k$. Consequently,

$$M = \frac{k+2-s_k}{s_k-1} \Rightarrow \beta = M+1 = \frac{k+1}{s_k-1},$$

which is the required result. This completes the proof of the theorem. $\square$

Remark 3. We recall that, for methods that converge superlinearly, that is, with order greater than 1, a measure of their effectiveness is the so-called efficiency index. If an iterative method for solving $f(x) = 0$ that requires p evaluations of $f(x)$ and its derivatives (here, we are assuming that the cost of computing any derivative of $f(x)$ is about the same as the cost of computing $f(x)$) has order $s > 1$, the efficiency index of the method is defined as $EI = s^{1/p}$. Figuratively speaking, EI is the order of the method per function evaluation. (See Traub [6] (pp. 11–13) for more on this subject.) Thus, the larger EI , the more effective the iterative method. In comparing methods, we should look at their performance after they have done a fixed (and the same) number of function evaluations. In view of this, let us compare the efficiency indices of the Newton–Raphson and the Steffensen methods, and the Steffensen–Traub method in Section 3.1, with the efficiency indices of the new enhanced methods $SSteff(n, k)$, $k = 1, 2, \dots, 6$:

Newton–Raphson	$s = 2$	$p = 2$	$EI \doteq 1.414$
Steffensen	$s = 2$	$p = 2$	$EI \doteq 1.414$
Steffensen–Traub (Section 3.1)	$s = 3$	$p = 3$	$EI \doteq 1.442$
$SSteff(n, 1)$	$s \doteq 2.414$	$p = 2$	$EI \doteq 1.554$
$SSteff(n, 2)$	$s \doteq 2.547$	$p = 2$	$EI \doteq 1.596$
$SSteff(n, 3)$	$s \doteq 2.592$	$p = 2$	$EI \doteq 1.600$
$SSteff(n, 4)$	$s \doteq 2.608$	$p = 2$	$EI \doteq 1.615$
$SSteff(n, 5)$	$s \doteq 2.614$	$p = 2$	$EI \doteq 1.617$
$SSteff(n, 6)$	$s \doteq 2.617$	$p = 2$	$EI \doteq 1.618$

5. Numerical Examples

We have applied the method in Section 4, with $k = 1, 2, 3$, to three nonlinear equations $f(x) = 0$ of different types with simple roots that we denote α . We recall that, in this method,

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(x_n - f(x_n)/p'_{n,k}(x_n)) - f(x_n)}{-f(x_n)/p'_{n,k}(x_n)}}, \quad n = k, k+1, k+2, \dots,$$

with

$$p'_{n,k}(x_n) = f[x_n, x_{n-1}] + \sum_{i=2}^k f[x_n, x_{n-1}, \dots, x_{n-i}] \prod_{j=1}^{i-1} (x_n - x_{n-j}).$$

We also recall that, for a fixed k , $k \in \{1, 2, 3, \dots\}$, and with x_0 as the initial approximation that we choose, we determine x_1 from x_0 , via the Steffensen method. (Again, x_1 can also be assigned any reasonable value. For example, we can choose x_1 close to x_0 but arbitrary otherwise.) Following that, we compute x_2 from x_0 and x_1 via the method $\text{SSteff}(1, 1)$. Next, we compute x_3 from x_0, x_1, x_2 via $\text{SSteff}(2, 2)$, and continue this way until we obtain x_k from $x_0, x_1, \dots, x_{k-1}$ via $\text{SSteff}(k-1, k-1)$. Following that, we continue with $\text{SSteff}(n, k)$ to compute x_{n+1} , $n = k, k+1, \dots$, as described in Section 4.1.

Let us write, for each $k \in \{1, 2, \dots\}$,

$$\epsilon_n = x_n - \alpha, \quad \rho_n = \frac{\log |\epsilon_{n+1}/\epsilon_n|}{\log |\epsilon_n/\epsilon_{n-1}|}.$$

By Theorem 1, we have, for each k ,

$$\lim_{n \rightarrow \infty} \rho_n = s_k.$$

This, of course, means that the ρ_n approximate s_k for each fixed k . (Recall that $s_1 \doteq 2.414$, $s_2 \doteq 2.547$, and $s_3 \doteq 2.592$.)

In the examples that follow, we have done all our computations using 1000-digit arithmetic to also demonstrate numerically the validity of our theory. The entries denoted “— — —” in the accompanying tables are not absolute zeros; they only mean that errors are below the working precision of the arithmetic used in the computations.

Example 1. Our first example is $f(x) = x^5 - 243$, whose root is $\alpha = 3$. Note that $f'(\alpha) = 405$. The results of our computations, starting with $x_0 = 4$, are given in Table 1. With this choice of the initial approximation, the method converges smoothly to the root $\alpha = 3$ for all values of k . The computed values of ρ_n clearly approach the theoretical orders of convergence, and the asymptotic behavior is well-observed before the onset of finite-precision saturation.

Example 2. Our second example is $f(x) = x + \log(x+1) - (3 + \log 4)$, whose root is $\alpha = 3$. Note that $f'(\alpha) = 5/4$. The results of our computations, starting with $x_0 = 4$, are given in Table 2. As observed, the method exhibits rapid convergence for all values of k . Moreover, the estimated orders ρ_n approach the theoretical values associated with each choice of k . Due to the high convergence speed, the iterates reach the exact solution within the working precision after only a few steps, and the subsequent entries are affected by finite-precision saturation.

Table 1. Errors $|\epsilon_n| = |x_n - \alpha|$ and observed orders ρ_n for $f(x) = x^5 - 243 = 0$ with $x_0 = 4$.

n	$k = 1$		$k = 2$		$k = 3$	
	$ \epsilon_n $	ρ_n	$ \epsilon_n $	ρ_n	$ \epsilon_n $	ρ_n
1	1.00D+000		1.00D+000		1.00D+000	
2	1.73D−001	1.7152	1.73D−001	2.6655	1.73D−001	2.6655
3	8.57D−003	2.4568	1.62D−003	2.1004	1.62D−003	2.5454
4	5.30D−006	2.3988	8.85D−008	2.6798	1.11D−008	2.5328
5	1.07D−013	2.4167	3.35D−019	2.5519	9.14D−022	2.6102
6	2.69D−032	2.4138	2.38D−048	2.5381	6.47D−056	2.5939
7	3.43D−077	2.4143	2.47D−122	2.5484	1.69D−144	2.5915
8	1.40D−185	2.4142	7.20D−311	2.5470	4.61D−374	2.5916
9	3.00D−447		4.52D−791		5.25D−969	
10	---		---		---	

Table 2. Errors $|\epsilon_n| = |x_n - \alpha|$ and observed orders ρ_n for $f(x) = x + \log(x + 1) - (3 + \log 4) = 0$ with $x_0 = 4$.

n	$k = 1$		$k = 2$		$k = 3$	
	$ \epsilon_n $	ρ_n	$ \epsilon_n $	ρ_n	$ \epsilon_n $	ρ_n
1	3.75D−002		3.75D−002		3.75D−002	
2	7.81D−007	2.2930	7.81D−007	2.4749	7.81D−007	2.4749
3	1.44D−017	2.4363	2.02D−018	2.5207	2.02D−018	2.5903
4	1.01D−043	2.4104	1.26D−047	2.5571	1.97D−048	2.5869
5	9.13D−107	2.4149	2.61D−122	2.5462	4.52D−126	2.5961
6	5.25D−259	2.4141	1.80D−312	2.5463	1.24D−327	2.5916
7	1.57D−626		1.11D−796		5.45D−850	
8	---		---		---	
9	---		---		---	
10	---		---		---	

Example 3. Our third example is $f(x) = (x - 2)e^x - e^3$, whose root is $\alpha = 3$. Note that $f'(\alpha) = 2e^3$. The results of our computations, starting with $x_0 = 4$, are given in Table 3. With this choice of the initial approximation, the method converges smoothly to the root $\alpha = 3$ for all values of k . The computed values of ρ_n clearly approach the theoretical orders of convergence, and the asymptotic behavior is well-observed before the onset of finite-precision saturation.

Table 3. Errors $|\epsilon_n| = |x_n - \alpha|$ and observed orders ρ_n for $f(x) = (x - 2)e^x - e^3 = 0$ with $x_0 = 4$.

n	$k = 1$		$k = 2$		$k = 3$	
	$ \epsilon_n $	ρ_n	$ \epsilon_n $	ρ_n	$ \epsilon_n $	ρ_n
1	1.00D+000		1.00D+000		1.00D+000	
2	2.31D−001	1.7036	2.31D−001	2.2639	2.31D−001	2.2639
3	1.91D−002	2.4379	8.40D−003	2.1077	8.40D−003	2.5246
4	4.35D−005	2.3975	7.75D−006	2.7672	1.95D−006	2.5851
5	2.02D−011	2.4167	3.10D−014	2.5361	7.79D−016	2.6193
6	1.00D−026	2.4138	1.56D−035	2.5367	1.89D−040	2.5889
7	1.14D−063	2.4143	1.45D−089	2.5496	3.54D−104	2.5907
8	7.34D−153	2.4142	2.54D−227	2.5468	2.81D−269	2.5920
9	3.46D−368	2.4142	3.65D−578		3.22D−697	
10	4.95D−888		---		---	

Interestingly, the equation we have treated is related to the Lambert equation. This can be shown as follows: letting $x = y + 2$, $f(x) = 0$ becomes

$$ye^{y+2} - e^3 = 0 \quad \Rightarrow \quad ye^y = e,$$

which is a Lambert equation. Clearly, the solution of this equation for y is simply $y = W(e) = 1$, where $W(e)$, the solution of $ye^y = e$, is the principal branch of the Lambert W -function. This results in $\alpha = 3$ for the solution of $f(x) = 0$.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The author would like to thank the anonymous referees for their critical comments and valuable suggestions that helped to improve the presentation of the topic of this paper.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

Theorem A1. *The real polynomial*

$$g_k(z) = z^{k+1} - az^k - \sum_{i=0}^{k-1} z^i, \quad a > 1, \quad (\text{A1})$$

has only one positive simple zero z_k in the interval $(a, a+1)$, and all the other zeros have moduli smaller than 1. In addition, z_k form a monotonically increasing sequence, and we have

$$\frac{a + \sqrt{a^2 + 4}}{2} = z_1 < z_2 < z_3 < \dots; \quad \lim_{k \rightarrow \infty} z_k = \hat{z} = \frac{a + 1 + \sqrt{(a-1)^2 + 4}}{2} < a + 1. \quad (\text{A2})$$

Proof. Define the polynomial $G_k(z)$ as in

$$G_k(z) = (z-1)g_k(z) = z^{k+2} - (a+1)z^{k+1} + (a-1)z^k + 1.$$

Then,

$$G_k(a) = -a^k + 1 < 0 \quad \text{and} \quad G_k(a+1) = (a-1)(a+1)^k + 1 > 0.$$

Therefore, $G_k(z)$ has a zero z_k in the interval $(a, a+1)$. Clearly, z_k is also a zero of $g_k(z)$ since $a > 1$. By Descartes' rule of signs, $G_k(z)$ has either two positive zeros or none. But, since $z = 1$ is such a zero, $G_k(z)$ must have one more, and this is z_k . Thus, z_k is the only positive root of $g_k(z)$. Of course, this also implies that

$$G_k(z) > 0 \quad \text{if } z > z_k, \quad G_k(z) < 0 \quad \text{if } 1 < z < z_k. \quad (\text{A3})$$

Next, by the fact that

$$zG_{k-1}(z) = G_k(z) + z - 1$$

and because $G_{k-1}(z_{k-1}) = 0$, we also have that

$$G_k(z_{k-1}) = -z_{k-1} + 1 < 0, \quad \text{since } z_{k-1} > a > 1,$$

which, upon invoking (A3), implies that $z_{k-1} < z_k$. Thus, the sequence $\{z_k\}_{k=1}^{\infty}$ is monotonically increasing and is bounded from above by $a+1$. Therefore, $\lim_{k \rightarrow \infty} z_k = \hat{z}$ exists. To determine $\hat{z}$, we rewrite the equality $G_k(z_k) = 0$ as in

$$z_k^2 - (a+1)z_k + (a-1) + z_k^{-k} = 0.$$

Letting now $k \rightarrow \infty$ and recalling that $z_k > a > 1$ so that $\lim_{k \rightarrow \infty} z_k^{-k} = 0$, we obtain

$$\hat{z}^2 - (a+1)\hat{z} + (a-1) = 0.$$

That is, $\hat{z}$ is a solution to the quadratic equation $z^2 - (a+1)z + (a-1) = 0$. This equation has two positive solutions, one being in $(0, a)$ while the other is in $(a, a+1)$. Clearly, $\hat{z}$ is the solution that is in the interval $(a, a+1)$, as given in (A2).

We now turn to the remaining k roots of the polynomial $g_k(z)$, which are the roots of the polynomial $h_k(z) = g_k(z)/(z - z_k)$ that has degree k , that is,

$$h_k(z) = \sum_{j=0}^k c_j z^j, \quad c_k = 1.$$

Then, $g_k(z) = (z - z_k)h_k(z)$, which gives the identity

$$z^{k+1} - az^k - \sum_{j=0}^{k-1} z^j = -z_k c_0 + \sum_{j=1}^k (c_{j-1} - z_k c_j) z^j + c_k z^{k+1},$$

hence leading to the following $k + 1$ equations:

$$-z_k c_0 = -1; \quad c_{j-1} - z_k c_j = -1, \quad 1 \leq j \leq k-1; \quad c_{k-1} - z_k c_k = -a.$$

Solving the first k of these equations for $c_0, c_1, \dots, c_{k-1}$, in this order, we obtain

$$c_j = \frac{1}{z_k} \sum_{i=0}^j \frac{1}{z_k^i}, \quad 0 \leq j \leq k-1 \quad \Rightarrow \quad 0 < c_0 < c_1 < \dots < c_{k-1}.$$

The last equation, with $c_k = 1$ and $z_k < a + 1$, gives

$$c_{k-1} = z_k c_k - a = z_k - a < 1 = c_k \quad \Rightarrow \quad c_{k-1} < c_k.$$

Therefore,

$$0 < c_0 < c_1 < \dots < c_{k-1} < c_k,$$

and, by the Eneström–Kakeya theorem, the polynomial $\sum_{j=0}^k c_j z^j$ has all its zeros in the interior of the unit disk $\{z : |z| \leq 1\}$. In fact, these zeros are in the annulus D ,

$$D = \{z : \min_{0 \leq j \leq k-1} \{c_j / c_{j+1}\} \leq |z| \leq \max_{0 \leq j \leq k-1} \{c_j / c_{j+1}\}\}.$$

This completes the proof. $\square$

Note that the problem treated in Theorem A1 is analogous to, yet different from, a similar problem treated in Traub [6] (pp. 46–51).

For Descartes' rule of signs, see, for example, Henrici [20] (p. 442). For the Eneström–Kakeya theorem, see, for example, Henrici [20] (p. 462), Marden [21] (p. 136), and Pólya and Szegő [22] (p. 107).

References

1. Atkinson, K.E. *An Introduction to Numerical Analysis*, 2nd ed.; John Wiley & Sons Inc.: New York, NY, USA, 1989.
2. Dahlquist, G.; Björck, Å. *Numerical Methods in Scientific Computing: Volume I*; SIAM: Philadelphia, PA, USA, 2008.
3. Henrici, P. *Elements of Numerical Analysis*; Wiley: New York, NY, USA, 1964.
4. Ralston, A.; Rabinowitz, P. *A First Course in Numerical Analysis*, 2nd ed.; McGraw-Hill: New York, NY, USA, 1978.
5. Stoer, J.; Bulirsch, R. *Introduction to Numerical Analysis*, 3rd ed.; Springer-Verlag: New York, NY, USA, 2002.
6. Traub, J.F. *Iterative Methods for the Solution of Equations*; Prentice Hall: Englewood Cliffs, NJ, USA, 1964.
7. Sidi, A. Unified treatment of regula falsi, Newton–Raphson, secant, and Steffensen methods for nonlinear equations. *J. Online Math. Appl.* **2006**, *6*.
8. Cordero, A.; Torregrosa, J.R. A class of Steffensen type methods with optimal order of convergence. *Appl. Math. Comput.* **2011**, *217*, 7653–7659. [[CrossRef](#)]
9. Cordero, A.; Torregrosa, J.R. Low-complexity root-finding iteration functions with no derivatives of any order of convergence. *J. Comp. Appl. Math.* **2015**, *275*, 502–515.
10. Džunić, J. On efficient two-parameter methods for solving nonlinear equations. *Numer. Algorithms* **2013**, *63*, 549–569.

11. Džunić, J.; Petković, M.S. A cubically convergent Steffensen-like method for solving nonlinear equations. *Appl. Math. Lett.* **2012**, *25*, 1881–1886. [[CrossRef](#)]
12. Liu, Z.; Zheng, Q.; Zhao, P. A variant of Steffensen's method of fourth-order convergence and its applications. *Appl. Math. Comput.* **2010**, *216*, 1978–1983.
13. Ren, H.; Wu, Q.; Bi, W. A class of two-step Steffensen type methods with fourth-order convergence. *Appl. Math. Comput.* **2009**, *209*, 206–210.
14. Wang, X.; Zhang, T. A family of Steffensen type methods with seventh-order convergence. *Numer. Algorithms* **2013**, *62*, 429–444.
15. Petković, M.S.; Neta, B.; Petković, L.D.; Džunić, J. *Multipoint Methods for Solving Nonlinear Equations*; Academic Press: Waltham, MA, USA, 2012.
16. Petković, M.S.; Neta, B.; Petković, L.D.; Džunić, J. Multipoint methods for solving nonlinear equations: A survey. *Appl. Math. Comput.* **2014**, *226*, 635–660. [[CrossRef](#)]
17. Cordero, A.; Garrido, N.; Torregrosa, J.R.; Triguero-Navarro, P. Three-step iterative weight function scheme with memory for solving nonlinear problems. *Math. Meth. Appl. Sci.* **2023**, *48*, 8024–8036.
18. Sidi, A. Generalization of the secant method for nonlinear equations. *Appl. Math. E-Notes* **2008**, *8*, 115–123.
19. Sidi, A. Application of a generalized secant method for nonlinear equations with complex roots. *Axioms* **2021**, *10*, 169. [[CrossRef](#)]
20. Henrici, P. *Applied and Computational Complex Analysis*; Wiley: New York, NY, USA, 1974; Volume 1.
21. Marden, M. Geometry of Polynomials. In *Mathematical Surveys Number 3*; American Mathematical Society: Providence, RI, USA, 1966.
22. Pólya, G.; Szegő, G. *Problems and Theorems in Analysis I*; Springer: Berlin/Heidelberg, Germany, 1972.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.