Article

Application of a Generalized Secant Method to Nonlinear Equations with Complex Roots

Avram Sidi

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

Abstract: The secant method is a very effective numerical procedure used for solving nonlinear equations of the form $f(x) = 0$. In a recent work (A. Sidi, Generalization of the secant method for nonlinear equations. *Appl. Math. E-Notes*, 8:115–123, 2008), we presented a generalization of the secant method that uses only one evaluation of $f(x)$ per iteration, and we provided a local convergence theory for it that concerns real roots. For each integer k, this method generates a sequence $\{x_n\}$ of approximations to a real root of $f(x)$, where, for $n \geq k$, $x_{n+1} = x_n - f(x_n)/p'_{n,k}(x_n)$, $p_{n,k}(x)$ being the polynomial of degree k that interpolates $f(x)$ at $x_n, x_{n-1}, \ldots, x_{n-k}$, the order s_k of this method satisfying $1 < s_k < 2$. Clearly, when $k = 1$, this method reduces to the secant method with $s_1 = (1 + \sqrt{5})/2$. In addition, $s_1 < s_2 < s_3 < \cdots$, such that $\lim_{k \to \infty} s_k = 2$. In this note, we study the application of this method to simple complex roots of a function $f(z)$. We show that the local convergence theory developed for real roots can be extended almost as is to complex roots, provided suitable assumptions and justifications are made. We illustrate the theory with two numerical examples.

Keywords: secant method; generalized secant method; complex roots

MSC: 65H05

1. Introduction

Let α be the solution to the equation $f(x) = 0$. An effective iterative method used for solving this equation that makes direct use of $f(x)$ (but no derivatives of $f(x)$) is the *secant method* that is discussed in many books on numerical analysis. See, for example, Atkinson [1], Dahlquist and Björck [2], Henrici [3], Ralston and Rabinowitz [4], and Stoer and Bulirsch [5]. See also the recent note [6] by the author, in which the treatment of the secant method and those of the Newton–Raphson, regula falsi, and Steffensen methods are presented in a unified manner.

Recently, this method was generalized by the author in [7] as follows: Starting with $x_0, x_1, \ldots, x_k$, $k + 1$ initial approximations to α, we generate a sequence of approximations $\{x_n\}$, via the recursion

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

$p'_{n,k}(x)$ being the derivative of the polynomial $p_{n,k}(x)$ that interpolates $f(x)$ at the points $x_n, x_{n-1}, \ldots, x_{n-k}$. (Thus, $p_{n,k}(x)$ is of degree k.) Clearly, the case $k = 1$ is simply the secant method. In [7], we also showed that, provided $x_0, x_1, \ldots, x_k$ are sufficiently close to α, the method converges with order s_k, that is, $\lim_{n \to \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^{s_k}} = C \neq 0$ for some constant C, and that $1 < s_k < 2$. (We call s_k the *order of convergence* of the method or the *order* of the method for short.) Here s_k is the only positive root of the polynomial $s^{k+1} - \sum_{i=0}^{k} s^i$. We also have that

$$\frac{1 + \sqrt{5}}{2} = s_1 < s_2 < s_3 < \cdots < 2; \quad \lim_{k \to \infty} s_k = 2.$$

Actually, rounded to four significant figures,

$$s_1 \doteq 1.618, \ s_2 \doteq 1.839, \ s_3 \doteq 1.928, \ s_4 \doteq 1.966, \ s_5 \doteq 1.984, \ s_6 \doteq 1.992, \ s_7 \doteq 1.996, \text{ etc.}$$

Note that to compute x_{n+1} we need knowledge of only $f(x_n), f(x_{n-1}), \ldots, f(x_{n-k})$, and because $f(x_{n-1}), \ldots, f(x_{n-k})$ have already been computed, $f(x_n)$ is the only new quantity to be computed. Thus, each step of the method requires $f(x)$ to be computed only once. From this, it follows that the efficiency index of this method is simply s_k and that this index approaches 2 by increasing k even moderately, as can be concluded from the values of $s_1, \ldots, s_7$ given above.

In this work, we consider the application of this method to simple complex roots of a function $f(z)$, where z is the complex variable. Let us denote a real or complex root of $f(z)$ by α again; that is, $f(\alpha) = 0$ and $f'(\alpha) \neq 0$. Thus, starting with $z_0, z_1, \ldots, z_k$, $k + 1$ initial approximations to α, we generate a sequence of approximations $\{z_n\}$ via the recursion

$$z_{n+1} = z_n - \frac{f(z_n)}{p'_{n,k}(z_n)}, \quad n = k, k+1, \ldots, \tag{2}$$

$p'_{n,k}(z)$ being the derivative of the polynomial $p_{n,k}(z)$ that interpolates $f(z)$ at the points $z_n, z_{n-1}, \ldots, z_{n-k}$. As in [7], we can use Newton's interpolation formula to generate $p_{n,k}(z)$ and $p'_{n,k}(z)$. Thus

$$p_{n,k}(z) = f(z_n) + \sum_{i=1}^{k} f[z_n, z_{n-1}, \ldots, z_{n-i}] \prod_{j=0}^{i-1} (z - z_{n-j}) \tag{3}$$

and

$$p'_{n,k}(z_n) = f[z_n, z_{n-1}] + \sum_{i=2}^{k} f[z_n, z_{n-1}, \ldots, z_{n-i}] \prod_{j=1}^{i-1} (z_n - z_{n-j}). \tag{4}$$

Here, $g[\zeta_0, \zeta_1, \ldots, \zeta_m]$ is the divided difference of order m of the function $g(z)$ over the set of points $\{\zeta_0, \zeta_1, \ldots, \zeta_m\}$ and is a symmetric function of these points. For details, we refer the reader to [7].

As proposed in [7], we generate the $k + 1$ initial approximations as follows: We choose the approximations z_0, z_1 first. We then generate z_2 by applying our method with $k = 1$ (that is, with the secant method). Next, we apply our method to z_0, z_1, z_2 with $k = 2$ and obtain z_3, and so on, until we have generated all $k + 1$ initial approximations, via

$$z_{n+1} = z_n - \frac{f(z_n)}{p'_{n,n}(z_n)}, \quad n = 1, 2, \ldots, k - 1. \tag{5}$$

Remark 1.

1.　*Instead of choosing z_1 arbitrarily, we can generate it as $z_1 = z_0 + f(z_0)$ as suggested in Brin [8], which is quite sensible since $f(z)$ is small near the root α. We can also use the method of Steffensen—which uses only $f(z)$ and no derivatives of $f(z)$—to generate z_1 from z_0; thus,*

$$z_1 = z_0 - \frac{[f(z_0)]^2}{f(z_0 + f(z_0)) - f(z_0)}.$$

2.　*It is clear that, in case $f(z)$ takes on only real values along the $\mathrm{Re}\, z$ axis and we are looking for nonreal roots of $f(z)$, at least one of the initial approximations must be chosen to be nonreal.*

3.　*We would like to mention that Kogan, Sapir, and Sapir [9] have proposed another generalization of the secant method for simple real roots of nonlinear equations $f(x) = 0$ that*

resembles our method described in (1). *In the notation of* (1), *this method produces a sequence of approximations* $\{x_n\}$ *via*

$$x_{n+1} = x_n - \frac{f(x_n)}{p'_{n,n}(x_n)}, \quad n = 1, 2, \ldots, \tag{6}$$

starting with arbitrary x_0 *and* x_1, *and it is of order* 2. *Note that, in* (6), $p_{n,n}(x)$ *interpolates* $f(x)$ *at the points* $x_0, x_1, \ldots, x_n$, *hence is of degree* n, *which is tending to infinity. In* (1), $p_{n,k}(x)$ *is of degree* k, *which is fixed.*

4. *Yet another generalization of the secant method for finding simple real roots of* $f(x)$ *was recently given by Nijmeijer* [10]. *This method too requires no derivative information, requires one evaluation of* $f(x)$ *per iteration, and has the same order of convergence as our method. It follows an idea of applying a convergence acceleration method, such as Aitken's* Δ^2-*process, to approximations obtained from the secant method, as proposed by Han and Potra* [11]. *Because Nijmeijer's method is not based on polynomial interpolation, it is completely different from our method, however. For Aitken's* Δ^2-*process, see* [1–5]. *See also* [12] (*Chapter 15*) *by the author.*

In the next section, we analyze the local convergence properties of the method as it is applied to complex roots. We show that the analysis of [7] can be extended to the complex case following some clever manipulation. We prove that the order s_k of the method is the same as that we discovered in the real case. In Section 3, we provide two numerical examples to confirm the results of our convergence analysis.

2. Local Convergence Analysis

We now turn to the analysis of the sequence $\{z_n\}_{n=0}^{\infty}$ that is generated via (2). Our treatment covers all $k \geq 1$.

In our analysis, we will make use of the Hermite-Genocchi formula that provides an integral representation for divided differences (For a proof of this formula, see Atkinson [1], for example). Even though this formula is usually stated for functions defined on real intervals, it is easy to verify (see Filipsson [13], for example) that it also applies to functions defined in the complex plane under proper assumptions. Thus, provided $g(z)$ is analytic on E, a bounded closed convex set in the complex plane, and provided $\zeta_0, \zeta_1, \ldots, \zeta_m$ are in E, there holds

$$g[\zeta_0, \zeta_1, \ldots, \zeta_m] = \int \cdots \int_{S_m} g^{(m)}(t_0\zeta_0 + t_1\zeta_1 + \cdots + t_m\zeta_m)\, dt_1 \cdots dt_m, \quad t_0 = 1 - \sum_{i=1}^{m} t_i. \tag{7}$$

Here S_m is the m-dimensional simplex defined as

$$S_m = \left\{ (t_1, \ldots, t_m) \in \mathbb{R}^m : t_i \geq 0, \ i = 1, \ldots, m, \ \sum_{i=1}^{m} t_i \leq 1 \right\}. \tag{8}$$

We note that (7) holds whether the ζ_i are distinct or not. We also note that $g[\zeta_0, \zeta_1, \ldots, \zeta_m]$ is a symmetric and continuous function of its arguments.

By the conditions we have imposed on $g(z)$, it is easy to see that the integrand $g^{(m)}(\sum_{i=0}^{m} t_i\zeta_i)$ in (7) is always defined because $\sum_{i=0}^{m} t_i\zeta_i$ is in the set E and $g(z)$ is analytic on E. This is so because, by (7) and (8),

$$(t_1, \ldots, t_m) \in S_m \quad \Rightarrow \quad t_i \geq 0, \quad i = 0, 1, \ldots, m, \quad \text{and} \quad \sum_{i=0}^{m} t_i = 1,$$

which implies that $\sum_{i=0}^{m} t_i\zeta_i$ is a convex combination of $\zeta_0, \zeta_1, \ldots, \zeta_m$ hence is in the set $C = \text{conv}\{\zeta_0, \zeta_1, \ldots, \zeta_m\}$, the convex hull of the points $\zeta_0, \zeta_1, \ldots, \zeta_m$, and $C \subseteq E$. Consequently, taking moduli on both sides of (7), we obtain, for all ζ_i in E,

$$\left| g[\zeta_0, \zeta_1, \ldots, \zeta_m] \right| \le \int \cdots \int_{S_m} \left| g^{(m)} \left(\sum_{i=0}^{m} t_i \zeta_i \right) \right| dt_1 \cdots dt_m$$

$$\le \frac{\| g^{(m)} \|}{m!}, \quad \| g^{(m)} \| = \max_{z \in E} |g^{(m)}(z)|. \tag{9}$$

In addition, since $\sum_{i=0}^{m} t_i = 1$ in (7), as $\zeta_i \to \hat{\zeta}$ for all $i = 0, 1, \ldots, m$, there hold $\sum_{i=0}^{m} t_i \zeta_i \to \hat{\zeta}$ and $g^{(m)}(\sum_{i=0}^{m} t_i \zeta_i) \to g^{(m)}(\hat{\zeta})$, and hence

$$\lim_{\substack{\zeta_i \to \hat{\zeta} \\ i=0,1,\ldots,m}} g[\zeta_0, \zeta_1, \ldots, \zeta_m] = g[\underbrace{\hat{\zeta}, \hat{\zeta}, \ldots, \hat{\zeta}}_{m+1 \text{ times}}] = \frac{g^{(m)}(\hat{\zeta})}{m!}. \tag{10}$$

In (9) and (10), we have also invoked the fact that (see [14] (p. 346), for example)

$$\int \cdots \int_{S_m} dt_1 \cdots dt_m = \frac{1}{m!}.$$

We will make use of these in the proof of our main theorem that follows. This theorem and its proof are almost identical to that given in [7] once we take into account, where and when needed, the fact that we are now working in the complex plane. For convenience, we provide all the details of the proof.

Theorem 1. *Let α be a simple root of $f(z)$, that is, $f(\alpha) = 0$, but $f'(\alpha) \ne 0$. Let B_r be the closed disk of radius r containing α as its center, that is,*

$$B_r = \{ z \in \mathbb{C} : |z - \alpha| \le r \}. \tag{11}$$

Let $f(z)$ be analytic on B_r. Choose a positive integer k and let $z_0, z_1, \ldots, z_k$ be distinct initial approximations to α. Generate $z_{k+1}, z_{k+2}, \ldots$ via

$$z_{n+1} = z_n - \frac{f(z_n)}{p'_{n,k}(z_n)}, \quad n = k, k+1, \ldots, \tag{12}$$

where $p_{n,k}(z)$ is the polynomial of interpolation to $f(z)$ at the points $z_n, z_{n-1}, \ldots, z_{n-k}$. Then, provided $z_0, z_1, \ldots, z_k$ are in B_r and sufficiently close to α, we have the following cases:

1. *If $f^{(k+1)}(\alpha) \ne 0$, the sequence $\{z_n\}$ converges to α, and*

$$\lim_{n \to \infty} \frac{\epsilon_{n+1}}{\prod_{i=0}^{k} \epsilon_{n-i}} = \frac{(-1)^{k+1}}{(k+1)!} \frac{f^{(k+1)}(\alpha)}{f'(\alpha)} \equiv L; \quad \epsilon_n = z_n - \alpha \quad \forall n. \tag{13}$$

The order of convergence is s_k, $1 < s_k < 2$, where s_k is the only positive root of the polynomial $g_k(s) = s^{k+1} - \sum_{i=0}^{k} s^i$ and satisfies

$$2 - 2^{-k-1}e < s_k < 2 - 2^{-k-1} \quad \text{for } k \ge 2; \quad s_k < s_{k+1}; \quad \lim_{k \to \infty} s_k = 2, \tag{14}$$

e being the base of natural logarithms, and

$$\lim_{n \to \infty} \frac{|\epsilon_{n+1}|}{|\epsilon_n|^{s_k}} = |L|^{(s_k - 1)/k}, \tag{15}$$

which also implies that

$$s_k = \lim_{n \to \infty} \frac{\log |\epsilon_{n+1}/\epsilon_n|}{\log |\epsilon_n/\epsilon_{n-1}|}. \tag{16}$$

2. *If $f(z)$ is a polynomial of degree at most k, the sequence $\{z_n\}$ converges to α, and*

$$\lim_{n\to\infty} \frac{\epsilon_{n+1}}{\epsilon_n^2} = \frac{f''(\alpha)}{2f'(\alpha)}; \quad \epsilon_n = z_n - \alpha \quad \forall\, n. \tag{17}$$

Thus, $\{z_n\}$ converges with order 2 if $f''(\alpha) \neq 0$, and with order greater than 2 if $f''(\alpha) = 0$.

Proof. We start by deriving a closed-form expression for the error in z_{n+1}. Subtracting α from both sides of (12), and noting that

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

we have

$$z_{n+1} - \alpha = \left(1 - \frac{f[z_n, \alpha]}{p'_{n,k}(z_n)}\right)(z_n - \alpha) = \frac{p'_{n,k}(z_n) - f[z_n, \alpha]}{p'_{n,k}(z_n)}(z_n - \alpha). \tag{18}$$

We now note that

$$p'_{n,k}(z_n) - f[z_n, \alpha] = \{p'_{n,k}(z_n) - f'(z_n)\} + \{f'(z_n) - f[z_n, \alpha]\}, \tag{19}$$

and that

$$f'(z_n) - p'_{n,k}(z_n) = f[z_n, z_n, z_{n-1}, \dots, z_{n-k}]\prod_{i=1}^{k}(z_n - z_{n-i}) \tag{20}$$

and

$$f'(z_n) - f[z_n, \alpha] = f[z_n, z_n] - f[z_n, \alpha] = f[z_n, z_n, \alpha](z_n - \alpha). \tag{21}$$

Note that (20) can be obtained by starting with the divided difference representation of $f(z) - p_{n,k}(z)$, namely, $f(z) - p_{n,k}(z) = f[z, z_n, z_{n-1}, \dots, z_{n-k}]\prod_{i=0}^{k}(z - z_{n-i})$, and by computing $\lim_{z\to z_n}[f(z) - p_{n,k}(z)]/\prod_{i=0}^{k}(z - z_{n-i})$ via L'Hôpital's rule.

For simplicity of notation, let

$$-f[z_n, z_n, z_{n-1}, \dots, z_{n-k}] = \widehat{D}_n \quad \text{and} \quad f[z_n, z_n, \alpha] = \widehat{E}_n, \tag{22}$$

and rewrite (19) and (20) as

$$p'_{n,k}(z_n) - f[z_n, \alpha] = \widehat{D}_n \prod_{i=1}^{k}(\epsilon_n - \epsilon_{n-i}) + \widehat{E}_n \epsilon_n, \tag{23}$$

$$p'_{n,k}(z_n) = f'(z_n) + \widehat{D}_n \prod_{i=1}^{k}(\epsilon_n - \epsilon_{n-i}). \tag{24}$$

Substituting these into (18), we finally obtain

$$\epsilon_{n+1} = C_n \epsilon_n; \quad C_n \equiv \frac{p'_{n,k}(z_n) - f[z_n, \alpha]}{p'_{n,k}(z_n)} = \frac{\widehat{D}_n \prod_{i=1}^{k}(\epsilon_n - \epsilon_{n-i}) + \widehat{E}_n \epsilon_n}{f'(z_n) + \widehat{D}_n \prod_{i=1}^{k}(\epsilon_n - \epsilon_{n-i})}. \tag{25}$$

We now prove that convergence takes place. First, let us assume without loss of generality that $f'(z) \neq 0$ for all $z \in B_r$, and set $m_1 = \min_{z\in B_r}|f'(z)| > 0$. (This is possible since $\alpha \in B_r$ and $f'(\alpha) \neq 0$, and we can choose r as small as we wish to also guarantee $m_1 > 0$.) Next, let $M_s = \max_{z\in B_r}|f^{(s)}(z)|/s!, s = 1, 2, \dots$. Thus, assuming that $\{z_n, z_{n-1}, \dots, z_{n-k}\} \subset B_r$ and noting that B_r is a convex set, we have by (9) that

$$|\widehat{D}_n| \le M_{k+1}, \quad |\widehat{E}_n| \le M_2, \quad \text{because } \{\alpha, z_n, z_{n-1}, \dots, z_{n-k}\} \subset B_r.$$

Next, choose the ball $B_{t/2}$ sufficiently small (with $t/2 \leq r$) to ensure that $m_1 > 2M_{k+1}t^k + M_2t/2$. It can now be verified that, provided $z_n, z_{n-1}, \ldots, z_{n-k}$ are all in $B_{t/2}$, there holds

$$
\begin{aligned}
|C_n| &\leq \frac{M_{k+1}\prod_{i=1}^{k}|\epsilon_n - \epsilon_{n-i}| + M_2|\epsilon_n|}{m_1 - M_{k+1}\prod_{i=1}^{k}|\epsilon_n - \epsilon_{n-i}|} \\
&\leq \frac{M_{k+1}\prod_{i=1}^{k}(|\epsilon_n| + |\epsilon_{n-i})| + M_2|\epsilon_n|}{m_1 - M_{k+1}\prod_{i=1}^{k}(|\epsilon_n| + |\epsilon_{n-i}|)} \leq \overline{C},
\end{aligned}
$$

where

$$
\overline{C} \equiv \frac{M_{k+1}t^k + M_2t/2}{m_1 - M_{k+1}t^k} < 1.
$$

Consequently, by (25), $|\epsilon_{n+1}| \leq \overline{C}|\epsilon_n| < |\epsilon_n|$, which implies that $z_{n+1} \in B_{t/2}$, just like $z_n, z_{n-1}, \ldots, z_{n-k}$. Therefore, if $z_0, z_1, \ldots, z_k$ are chosen in $B_{t/2}$, then $|C_n| \leq \overline{C} < 1$ for all $n \geq k$, hence $\{z_n\} \subset B_{t/2}$ and $\lim_{n\to\infty} z_n = \alpha$.

As for (13) when $f^{(k+1)}(\alpha) \neq 0$, we proceed as follows: By the fact that $\lim_{n\to\infty} z_n = \alpha$, we first note that, by (20) and (21),

$$
\lim_{n\to\infty} p'_{n,k}(z_n) = f'(\alpha) = \lim_{n\to\infty} f[z_n, \alpha], \tag{26}
$$

and thus $\lim_{n\to\infty} C_n = 0$. This means that $\lim_{n\to\infty}(\epsilon_{n+1}/\epsilon_n) = 0$ and, equivalently, that $\{z_n\}$ converges with order greater than 1. As a result,

$$
\lim_{n\to\infty}(\epsilon_n/\epsilon_{n-i}) = 0 \quad \text{for all } i \geq 1,
$$

and

$$
\epsilon_n/\epsilon_{n-i} = o(\epsilon_n/\epsilon_{n-j}) \quad \text{as } n \to \infty, \quad \text{for } j < i.
$$

Consequently, expanding in (25) the product $\prod_{i=1}^{k}(\epsilon_n - \epsilon_{n-i})$, we have

$$
\begin{aligned}
\prod_{i=1}^{k}(\epsilon_n - \epsilon_{n-i}) &= \prod_{i=1}^{k}\left(-\epsilon_{n-i}[1 - \epsilon_n/\epsilon_{n-i}]\right) \\
&= (-1)^k\left(\prod_{i=1}^{k}\epsilon_{n-i}\right)[1 + O(\epsilon_n/\epsilon_{n-1})] \quad \text{as } n \to \infty. \tag{27}
\end{aligned}
$$

Substituting (27) into (25), and defining

$$
D_n = \frac{\widehat{D}_n}{p'_{n,k}(z_n)}, \quad E_n = \frac{\widehat{E}_n}{p'_{n,k}(z_n)}, \tag{28}
$$

we obtain

$$
\epsilon_{n+1} = (-1)^k D_n\left(\prod_{i=0}^{k}\epsilon_{n-i}\right)[1 + O(\epsilon_n/\epsilon_{n-1})] + E_n\epsilon_n^2 \quad \text{as } n \to \infty. \tag{29}
$$

Dividing both sides of (29) by $\prod_{i=0}^{k}\epsilon_{n-i}$, and defining

$$
\sigma_n = \frac{\epsilon_{n+1}}{\prod_{i=0}^{k}\epsilon_{n-i}}, \tag{30}
$$

we have

$$
\sigma_n = (-1)^k D_n[1 + O(\epsilon_n/\epsilon_{n-1})] + E_n\sigma_{n-1}\epsilon_{n-k-1} \quad \text{as } n \to \infty. \tag{31}
$$

Now, by (10), (22), and (26),

$$\lim_{n\to\infty} D_n = -\frac{1}{(k+1)!}\frac{f^{(k+1)}(\alpha)}{f'(\alpha)}, \quad \lim_{n\to\infty} E_n = \frac{f^{(2)}(\alpha)}{2f'(\alpha)}. \tag{32}$$

Because $\lim_{n\to\infty} D_n$ and $\lim_{n\to\infty} E_n$ are finite, and because $\lim_{n\to\infty}(\epsilon_n/\epsilon_{n-1}) = 0$ and $\lim_{n\to\infty}\epsilon_{n-k-1} = 0$, it follows that there exist a positive integer N and positive constants $\beta < 1$ and D, with $|E_n\epsilon_{n-k-1}| \le \beta$ when $n > N$, for which (31) gives

$$|\sigma_n| \le D + \beta|\sigma_{n-1}| \quad \text{for all } n > N. \tag{33}$$

Using (33), it is easy to show that

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

which, by the fact that $\beta < 1$, implies that $\{\sigma_n\}$ is a bounded sequence. Making use of this fact, we have $\lim_{n\to\infty} E_n\sigma_{n-1}\epsilon_{n-k-1} = 0$. Substituting this into (31), and invoking (32), we next obtain $\lim_{n\to\infty}\sigma_n = (-1)^k\lim_{n\to\infty}D_n = L$, which is precisely (13).

That s_k, the order of the method, as defined in the statement of the theorem, satisfies (14) and (15) follows from Traub [15] (Chapter 3). We provide a simplified treatment of this topic in Appendix A.

This completes the proof of part 1 of the theorem.

When $f(z)$ is a polynomial of degree at most k, we first observe that $f^{(k+1)}(z) = 0$ for all z, which implies that $p_{n,k}(z) = f(z)$ for all z, hence also $p'_{n,k}(z) = f'(z)$ for all z. Therefore, we have that $p'_{n,k}(z_n) = f'(z_n)$ in the recursion of (12). Consequently, (12) becomes

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}, \quad n = k, k+1, \dots,$$

which is the recursion for the Newton–Raphson method. Thus, (17) follows. This completes the proof of part 2 of the theorem. $\square$

3. Numerical Examples

In this section, we present two numerical examples that we treated with our method. Our computations were done in quadruple-precision arithmetic (approximately 35-decimal-digit accuracy). Note that in order to verify the theoretical results concerning iterative methods with order greater than unity, we need to use computer arithmetic of high precision (preferably, of variable precision, if available) because the number of correct significant decimal digits in the z_n increases dramatically from one iteration to the next as we are approaching the solution.

In both examples below, we take $k = 2$. We choose z_0 and z_1 and compute z_2 using one step of the secant method, namely,

$$z_2 = z_1 - \frac{f(z_1)}{f[z_0, z_1]}. \tag{34}$$

Following that, we compute $z_3, z_4, \dots$, via

$$z_{n+1} = z_n - \frac{f(z_n)}{f[z_n, z_{n-1}] + f[z_n, z_{n-1}, z_{n-2}](z_n - z_{n-1})}, \quad n = 2, 3, \dots. \tag{35}$$

In our examples, we have carried out our computations for several sets of z_0, z_1, and we have observed essentially the same behavior that we observe in Tables 1 and 2.

Example 1. *Consider* $f(z) = 0$, *where* $f(z) = z^3 - 8$, *whose solutions are* $\alpha_r = 2e^{i2\pi r/3}$, $r = 0, 1, 2$. *We would like to obtain the root* $\alpha_1 = 2e^{i2\pi/3} = -1 + i\sqrt{3}$. *We chose* $z_0 = 2i$ *and* $z_1 = -2 + 2i$. *The results of our computations are given in Table 1.*

Table 1. Results obtained by applying the generalized secant method with $k = 2$, as shown in (34) and (35), to the equation $z^3 - 8 = 0$, to compute the root $\alpha_1 = -1 + i\sqrt{3}$. The entries denoted "**" mean that the limit of the extended-precision arithmetic has been reached.

n	$\lvert\epsilon_n\rvert$	$\dfrac{\epsilon_{n+1}}{\epsilon_n\epsilon_{n-1}\epsilon_{n-2}}$	$\dfrac{\log\lvert\epsilon_{n+1}/\epsilon_n\rvert}{\log\lvert\epsilon_n/\epsilon_{n-1}\rvert}$
0	$1.035D + 00$	-	-
1	$1.035D + 00$	-	-
2	$4.808D - 01$	$-8.972D - 02 + i\,1.015D - 01$	2.516
3	$6.979D - 02$	$1.224D - 01 - i\,2.727D - 02$	1.437
4	$4.355D - 03$	$1.009D - 01 - i\,4.079D - 02$	2.023
5	$1.591D - 05$	$4.561D - 02 - i\,9.794D - 02$	1.839
6	$5.223D - 10$	$3.793D - 02 - i\,7.268D - 02$	1.839
7	$2.967D - 18$	$3.741D - 02 - i\,7.579D - 02$	1.838
8	$2.083D - 33$	**	**
9	$0.000D + 00$	**	**

From (13) and (16) in Theorem 1, we should have

$$\lim_{n\to\infty} \frac{\epsilon_{n+1}}{\epsilon_n\epsilon_{n-1}\epsilon_{n-2}} = \frac{(-1)^3}{3!}\frac{f'''(\alpha_1)}{f'(\alpha_1)} = \frac{1}{24}(1 - i\sqrt{3}) = 0.04166\cdots - i\,0.07216\cdots$$

and

$$\lim_{n\to\infty} \frac{\log\lvert\epsilon_{n+1}/\epsilon_n\rvert}{\log\lvert\epsilon_n/\epsilon_{n-1}\rvert} = s_2 = 1.83928\cdots,$$

and these seem to be confirmed in Table 1. Furthermore, in infinite-precision arithmetic, z_9 should have close to 60 correct significant figures; we do not see this in Table 1 due to the fact that the arithmetic we have used to generate Table 1 can provide an accuracy of at most 35 digits.

Example 2. *Consider* $f(z) = 0$, *where* $f(z) = \sin(iz) - \cos z$. $f(z)$ *has infinitely many roots* $\alpha_r = (1 - i)(\pi/4 + r\pi)$, $r = 0, \pm1, \pm2, \ldots$. *We would like to obtain the root* $\alpha_0 = (1 - i)\pi/4$. *We chose* $z_0 = 1.5 - i1.3$ *and* $z_1 = 0.6 - i0.5$. *The results of our computations are given in Table 2.*

Table 2. Results obtained by applying the generalized secant method with $k = 2$, as shown in (34) and (35), to the equation $\sin(iz) - \cos z = 0$, to compute the root $\alpha_0 = (1 - i)\pi/4$. The entries denoted "**" mean that the limit of the extended-precision arithmetic has been reached.

n	$\lvert\epsilon_n\rvert$	$\dfrac{\epsilon_{n+1}}{\epsilon_n\epsilon_{n-1}\epsilon_{n-2}}$	$\dfrac{\log\lvert\epsilon_{n+1}/\epsilon_n\rvert}{\log\lvert\epsilon_n/\epsilon_{n-1}\rvert}$
0	$6.608D - 01$	-	-
1	$3.403D - 01$	-	-
2	$1.341D - 01$	$3.163D - 01 + i\,1.397D - 01$	2.743
3	$1.043D - 02$	$1.466D - 01 - i\,1.846D - 01$	1.774
4	$1.122D - 04$	$-2.943D - 03 - i\,1.117D - 01$	1.934
5	$1.755D - 08$	$9.223D - 03 - i\,1.614D - 01$	1.766
6	$3.320D - 15$	$-7.686D - 04 - i\,1.658D - 01$	1.857
7	$1.084D - 27$	**	**
8	$9.630D - 35$	**	**

From (13) and (16) in Theorem 1, we should have

$$\lim_{n\to\infty} \frac{\epsilon_{n+1}}{\epsilon_n\epsilon_{n-1}\epsilon_{n-2}} = \frac{(-1)^3}{3!}\frac{f'''(\alpha_1)}{f'(\alpha_1)} = -\frac{i}{6} = -i\,0.16666\cdots$$

and

$$\lim_{n\to\infty} \frac{\log|\epsilon_{n+1}/\epsilon_n|}{\log|\epsilon_n/\epsilon_{n-1}|} = s_2 = 1.83928\cdots,$$

and these seem to be confirmed in Table 2. Furthermore, in infinite-precision arithmetic, z_8 should have close to 50 correct significant figures; we do not see this in Table 2 due to the fact that the arithmetic we have used to generate Table 2 can provide an accuracy of at most 35 digits.

Remark 2. *In relation to the examples we have just presented, we would like to discuss the issue of estimating the relative errors $|\epsilon_n/\alpha|$ in the z_n. This should help the reader when studying the numerical results included in Tables 1 and 2. Starting with (13) and (15), we first note that, for all large n,*

$$|\epsilon_{n+1}| \approx |L|^{(s_k-1)/k}|\epsilon_n|^{s_k}.$$

Therefore, assuming also that $\alpha \neq 0$, we have

$$|\epsilon_{n+1}/\alpha| \approx D|\epsilon_n/\alpha|^{s_k}, \quad D = \left(|L|^{1/k}|\alpha|\right)^{s_k-1}.$$

Now, if z_n has $q > 0$ correct significant figures, we have $|\epsilon_n/\alpha| = O(10^{-q})$. If, in addition, $D = O(10^r)$ for some r, then we will have

$$|\epsilon_{n+1}/\alpha| \approx O(10^{r-qs_k}).$$

For simplicity, let us consider the case $r = 0$, which is practically what we have in the two examples we have treated. Then z_{n+1} has approximately qs_k correct significant decimal digits. That is, if z_n has q correct significant decimal digits, then, due to the fact that $s_k > 1$, z_{n+1} will have s_k times as many correct significant decimal digits as z_n.

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

Appendix A

Before ending, we would like to provide a brief treatment of the order of convergence of our method stated in (14) and (15) by considering

$$\frac{\epsilon_{n+1}}{\prod_{i=0}^{k}\epsilon_{n-i}} = L \quad \forall n \quad \Leftrightarrow \quad \epsilon_{n+1} = L\prod_{i=0}^{k}\epsilon_{n-i} \quad \forall n,$$

instead of (13). We will show that $|\epsilon_{n+1}| = Q|\epsilon_n|^{s_k}$ is possible if s_k is a solution to the polynomial equation $s^{k+1} = \sum_{i=0}^{k} s^i$ and $Q = |L|^{(s_k-1)/k}$. (For a more detailed treatment, we refer the reader to [15] (Section 3.3)).

We start by expressing all $|\epsilon_{n-i}|$ in terms of $|\epsilon_n|$. We have

$$|\epsilon_{n-i}| = \frac{|\epsilon_n|^{1/s_k^i}}{Q^{m_i}}, \quad m_i = \sum_{j=1}^{i}\frac{1}{s_k^j}, \quad i = 1, 2, \ldots.$$

Substituting this into $|\epsilon_{n+1}| = |L|\prod_{i=0}^{k}|\epsilon_{n-i}|$, we obtain

$$Q|\epsilon_n|^{s_k} = |L|\,|\epsilon_n|\prod_{i=1}^{k}\frac{|\epsilon_n|^{1/s_k^i}}{Q^{m_i}} = \frac{|L|}{Q^M}|\epsilon_n|^{\rho}; \quad \rho = \sum_{i=0}^{k}\frac{1}{s_k^i}, \quad M = \sum_{i=1}^{k}m_i.$$

Of course, this is possible when $s_k = \rho$ and $Q^{M+1} = |L|$.

Now, the requirement that $s_k = \rho$ is the same as $s_k^{k+1} = \sum_{i=0}^{k} s_k^i$, which implies that the order s_k should be a root of the polynomial

$$g_k(s) = s^{k+1} - \sum_{i=0}^{k} s^i = \frac{s^{k+2} - 2s^{k+1} + 1}{s - 1}.$$

By Descartes' rule of signs, $g_k(s)$ has only one positive root, which we denote by $\tilde{s}$. Since $g_k(1) = -k < 0$ and $g_k(2) = 1 > 0$, we have that $1 < \tilde{s} < 2$. The remaining k roots of $g_k(s)$ are the zeroes of the polynomial $\tilde{g}(s) = g_k(s)/(s - \tilde{s}) = \sum_{j=0}^{k} c_j s^j$, the c_j satisfying $\tilde{s}c_0 = 1$ and $\tilde{s}c_j - c_{j-1} = 1$, $j = 1, \ldots, k$, hence

$$c_j = \frac{1}{\tilde{s}} \sum_{i=0}^{j} \frac{1}{\tilde{s}^i}, \quad j = 0, 1, \ldots, k \quad \Rightarrow \quad 0 < c_0 < c_1 \cdots < c_{k-1} < c_k = 1.$$

Therefore, by the Eneström–Kakeya theorem, all k roots of $\tilde{g}(s)$ are in the unit disk. We thus conclude that $\tilde{s} = s_k$ since we already know that the order of our method is greater than 1. (For Descartes' rule of signs and the Eneström–Kakeya theorem, see, for example, Henrici [16] (pp. 442, 462)).

Next, we note that $g_k(s) = sg_{k-1}(s) - 1$. Therefore, $g_{k-1}(s_{k-1}) = 0$ implies $g_k(s_{k-1}) = -1 < 0$, which, along with $g_k(2) = 1 > 0$, implies that $s_{k-1} < s_k < 2$. Therefore, the sequence $\{s_k\}_{k=1}^{\infty}$ is monotonically increasing and is bounded from above by 2. Consequently, $\lim_{k \to \infty} s_k = \hat{s}$ exists and $\hat{s} \leq 2$. Now,

$$g_k(s_k) = 0 \quad \Rightarrow \quad s_k^{k+2} - 2s_k^{k+1} + 1 = 0 \quad \Rightarrow \quad s_k^2 - 2s_k = -\frac{1}{s_k^k}.$$

Upon letting $k \to \infty$ on both sides, we obtain $\hat{s}^2 - 2\hat{s} = 0$, which gives $\hat{s} = 2$.

The expression given for M can be simplified considerably as we show next. 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,

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

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

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

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

$$M = \frac{k - (s_k - 1)}{s_k - 1} \quad \Rightarrow \quad M + 1 = \frac{k}{s_k - 1},$$

which is the required result.

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: New York, NY, USA, 2002.

6. Sidi, A. Unified treatment of regula falsi, Newton–Raphson, secant, and Steffensen methods for nonlinear equations. *J. Online Math. Appl.* **2006**, *6*, 1–13.
7. Sidi, A. Generalization of the secant method for nonlinear equations. *Appl. Math. E-Notes* **2008**, *8*, 115–123.
8. Brin, L.Q. *Tea Time Numerical Analysis*; Southern Connecticut State University: New Haven, CT, USA, 2016.
9. Kogan, T.; Sapir, L.; Sapir, A. A nonstationary iterative second-order method for solving nonlinear equations. *Appl. Math. Comput.* **2007**, *188*, 75–82. [CrossRef]
10. Nijmeijer, M.J.P. A method to accelerate the convergence of the secant algorithm. *Adv. Numer. Anal.* **2014**, *2014*, 321592. [CrossRef]
11. Han, W.; Potra, F.A. Convergence acceleration for some root finding methods. *Comput. Suppl.* **1993**, *9*, 67–78.
12. Sidi, A. *Practical Extrapolation Methods: Theory and Applications*; Number 10 in Cambridge Monographs on Applied and Computational Mathematics; Cambridge University Press: Cambridge, UK, 2003.
13. Filipsson, L. Complex mean-value interpolation and approximation of holomorphic functions. *J. Approx. Theory* **1997**, *91*, 244–278. [CrossRef]
14. Davis, P.J.; Rabinowitz, P. *Methods of Numerical Integration*, 2nd ed.; Academic Press: New York, NY, USA, 1984.
15. Traub, J.F. *Iterative Methods for the Solution of Equations*; Prentice Hall: Englewood Cliffs, NJ, USA, 1964.
16. Henrici, P. *Applied and Computational Complex Analysis*; Wiley: New York, NY, USA, 1974; Volume 1.