Journal of Online Mathematics and Its Applications

Unified Treatment of Regula Falsi, Newton–Raphson, Secant, and Steffensen Methods for Nonlinear Equations

Avram Sidi
Computer Science Department
Technion - Israel Institute of Technology
Haifa 32000, Israel
e-mail: asidi@cs.technion.ac.il
http://www.cs.technion.ac.il/ asidi

May 2006

Abstract

Regula falsi, Newton–Raphson, secant, and Steffensen methods are four very effective numerical procedures used for solving nonlinear equations of the form $f(x) = 0$. They are derived via linear interpolation procedures. Their analyses can be carried out by making use of interpolation theory through divided differences and Newton's interpolation formula. In this note, we unify these analyses. The analysis of the Steffensen method given here seems to be new and is especially simpler than the standard treatments. The contents of this note should also be a useful exercise/example in the application of polynomial interpolation and divided differences in introductory courses in numerical analysis.

1 Introduction

Let α be the solution to the equation

$$f(x) = 0, \tag{1}$$

and assume that $f(x)$ is twice continuously differentiable in a closed interval I containing α in its interior. Some iterative methods used for solving (1) and that make direct use of $f(x)$ are the *regula falsi method* (or *false position method*), the *secant method*, the *Newton–Raphson method*, and the *Steffensen method.* These methods are discussed in many books on numerical analysis. See, for example, Atkinson [1], Henrici [2], Ralston and Rabinowitz [3], and Stoer and Bulirsch [5].

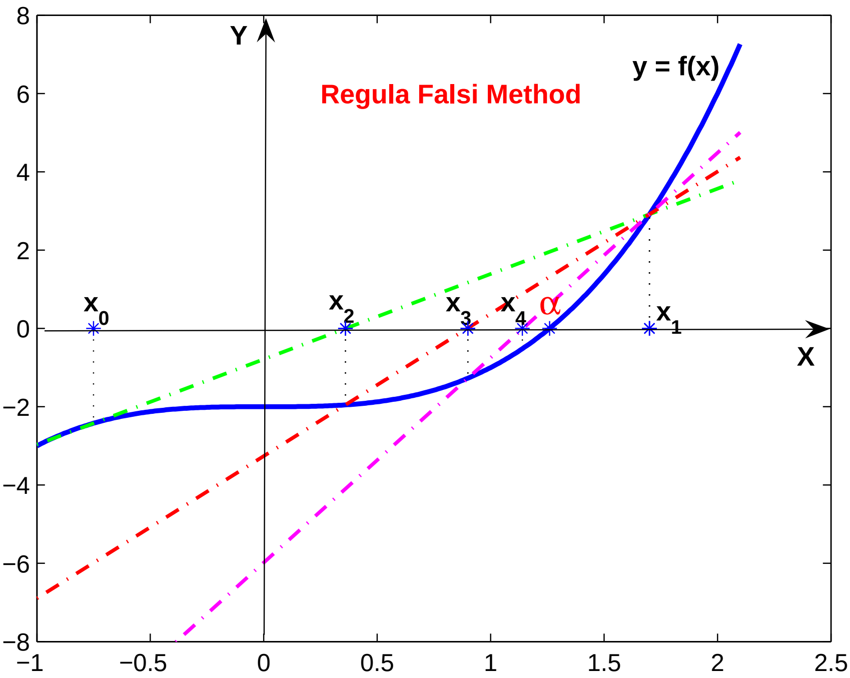

Figure 1: The regula falsi method.

All four methods are derived by a linear interpolation procedure as follows: Assuming that the approximations x_k to the solution α of (1) have been determined for all $k \leq n$, and another approximation c is available, the next approximation x_{n+1} is determined as the point of intersection (in the x-y plane) of the straight line through the points $(x_n, f(x_n))$ and $(c, f(c))$ with the x-axis. (See Figures 1–3 for the regula falsi, secant, and Newton–Raphson methods.) Since the equation of this straight line is

$$y = f(x_n) + \frac{f(x_n) - f(c)}{x_n - c}\,(x - x_n), \tag{2}$$

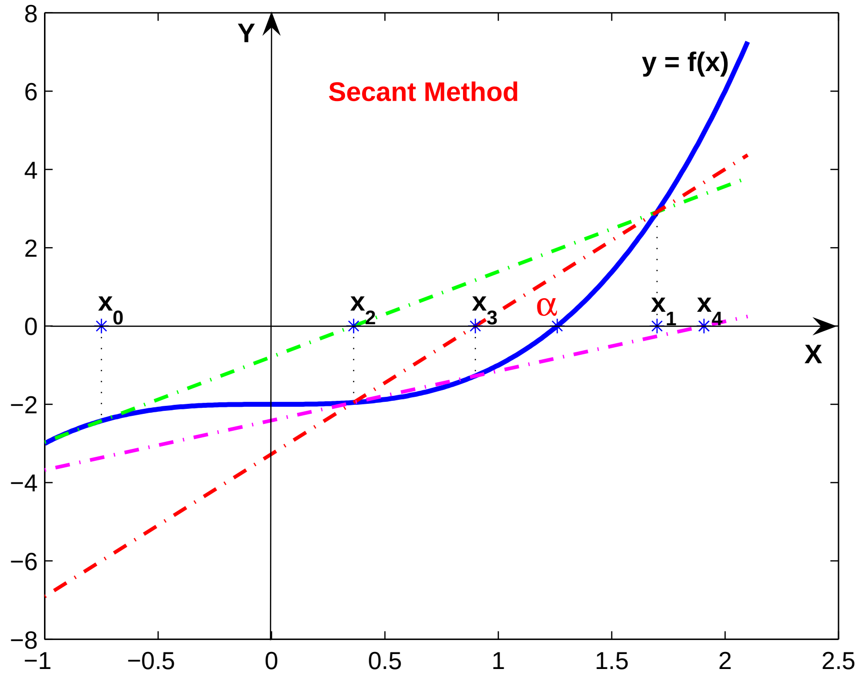

Figure 2: The secant method.

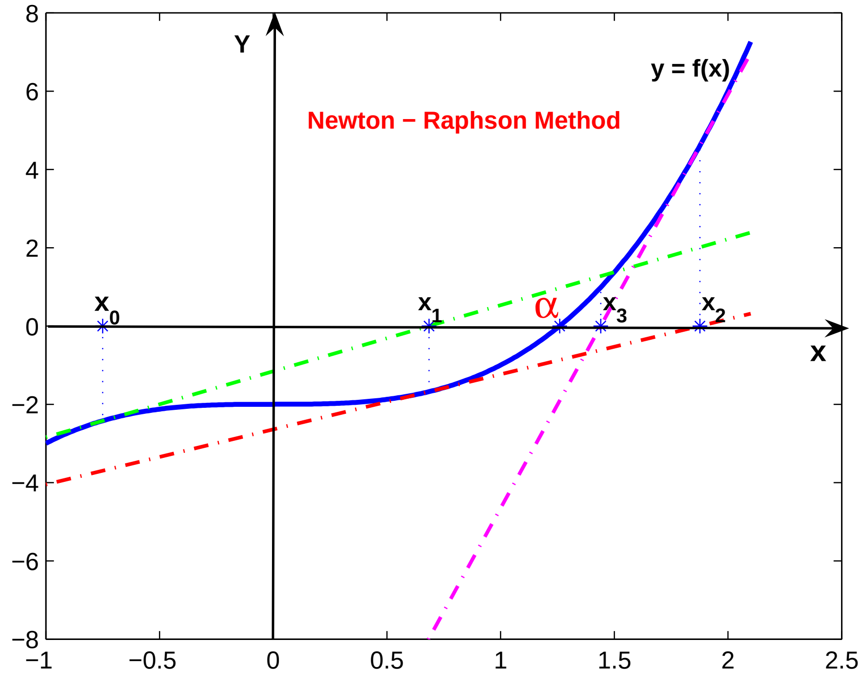

Figure 3: The Newton–Raphson method.

x_{n+1} is given as

$$x_{n+1} = x_n - \frac{f(x_n)}{\dfrac{f(x_n) - f(c)}{x_n - c}}. \tag{3}$$

Subtracting now α from both sides of (3), we obtain

$$x_{n+1} - \alpha = -\frac{f(x_n) + \dfrac{f(x_n) - f(c)}{x_n - c}(\alpha - x_n)}{\dfrac{f(x_n) - f(c)}{x_n - c}}. \tag{4}$$

It must be noted though that the above is valid also when $c = x_n$, as is the case in the Newton–Raphson method. In this case, the straight line in question, whose equation is $y = f(x_n) + f'(x_n)(x - x_n)$, is the tangent to the curve $y = f(x)$ at the point $(x_n, f(x_n))$, and can be obtained by letting $c \to x_n$ in (2). Having said this, we would like to emphasize nevertheless that the polynomial $p(x) = f(x_n) + f'(x_n)(x - x_n)$ is the linear interpolant to $f(x)$ (in the Hermite sense) that satisfies $p(x_n) = f(x_n)$ and $p'(x_n) = f'(x_n)$. It can be treated as any polynomial of interpolation via the Newton interpolation formula. We will discuss these issues in the next section in more generality and detail.

Most textbooks in numerical analysis that treat the subject, provide the convergence analyses of the four methods mentioned above separately. They use Taylor series for the Newton–Raphson method, while they apply linear interpolation theory to the secant method. They use yet another approach for the Steffensen method.

By stressing the fact that all four methods are obtained via a linear interpolation procedure, in this note, we unify the analyses of all these methods by using the theory of polynomial interpolation via divided differences. As such, the treatment also provides a good example/exercise in the application of the subject of polynomial interpolation via Newton's interpolation formula and divided differences, whether the points of interpolation are distinct or not. Furthermore, the analysis of the Steffensen method presented here turns out to be especially simple and does not seem to have been given in the literature before.

In the next section, we recall the important properties of divided differences and Newton's interpolation formula. We refer the reader to the books mentioned above for detailed treatments of this subject. In Section 3, we express $x_{n+1} - \alpha$ via divided differences in a way that also reveals the order of each method. In Section 4, we complete the convergence proofs of the methods in a unified manner.

2 Divided differences and Newton's interpolation formula

In the sequel, we will denote by $f[z_0, z_1, \ldots, z_m]$ the divided difference of order m of $f(x)$ on the set of points $\{z_0, z_1, \ldots, z_m\}$ and will recall the following:

1. $f[z_i, z_{i+1}, \ldots, z_m]$ can be defined recursively via

$$f[z_i] = f(z_i); \quad f[z_i, z_j] = \frac{f[z_i] - f[z_j]}{z_i - z_j}, \quad z_i \neq z_j, \tag{5}$$

and, for $m > i+1$, via

$$f[z_i, z_{i+1}, \ldots, z_m] = \frac{f[z_i, z_{i+1}, \ldots, z_{m-1}] - f[z_{i+1}, z_{i+2}, \ldots, z_m]}{z_i - z_m}, \quad z_i \neq z_m. \tag{6}$$

In case the z_i are not distinct, the divided differences are defined as limits of the quotients above, provided the limits exist. For example, $f[a, a] = f'(a)$ provided $f(z)$ is differentiable at $z = a$.

2. $f[z_0, z_1, \ldots, z_m]$ is a symmetric function of its arguments, that is, it has the same value for every ordering of the points $z_0, z_1, \ldots, z_m$.

3. Provided $f \in C^m[a, b]$ and $z_i \in [a, b]$, $i = 0, 1, \ldots, m$, $f[z_0, z_1, \ldots, z_m]$ is a continuous function of its arguments $z_0, z_1, \ldots, z_m$; thus, $\lim_{y \to y_0} f[x, y, z] = f[x, y_0, z]$, for example. In addition, *whether the z_i are distinct or not,*

$$f[z_0, z_1, \ldots, z_m] = \frac{f^{(m)}(\xi)}{m!} \quad \text{for some } \xi \in \operatorname{int}(z_0, z_1, \ldots, z_m), \tag{7}$$

where

$$\operatorname{int}(z_0, z_1, \ldots, z_m) \equiv (\min\{z_i\}, \max\{z_i\}). \tag{8}$$

(This is a very convenient notation we use throughout this note.) Consequently,

$$f[\underbrace{z, z, \ldots, z}_{m+1 \text{ times}}] = \frac{f^{(m)}(z)}{m!}, \quad z \in [a, b]. \tag{9}$$

4. Newton's formula for the polynomial of interpolation $p(z)$ to the function $f(z)$ at the points $z_0, z_1, \ldots, z_m$, *whether these points are distinct or not,* is given by

$$p(z) = f(z_0) + \sum_{i=1}^{m} f[z_0, z_1, \ldots, z_i] \prod_{s=0}^{i-1} (z - z_s), \tag{10}$$

and the corresponding error formula is

$$f(z) - p(z) = f[z_0, z_1, \ldots, z_m, z] \prod_{s=0}^{m} (z - z_s). \tag{11}$$

Thus, unlike the Lagrange interpolation formula, the Newton interpolation formula allows points of interpolation to coincide, hence is much more flexible.

(i) When the z_i are distinct, $p(z)$ can be determined from $f(z_0), f(z_1), \ldots, f(z_m)$. In this case, the relevant divided difference table is computed via (5), (6) in a straightforward manner.

(ii) In case $z_0 = z_1 = \cdots = z_m$, by (7), $p(z)$ in (10) becomes the mth partial sum of the Taylor series of $f(z)$ about z_0, and the expression for $f(z) - p(z)$ given in (11) becomes the corresponding remainder.

(iii) In case the z_i are not all distinct, we proceed as follows: We denote the distinct z_i by $a_1, a_2, \ldots, a_r$. For each $i = 1, \ldots, r$, we denote the multiplicity of a_i by s_i. Thus, $m + 1 = \sum_{i=1}^{r} s_i$. We then order the z_i as in

$$\begin{aligned}
&z_0 = z_1 = \cdots = z_{s_1-1} = a_1,\\
&z_{s_1} = z_{s_1+1} = \cdots = z_{s_1+s_2-1} = a_2,\\
&z_{s_1+s_2} = z_{s_1+s_2+1} = \cdots = z_{s_1+s_2+s_3-1} = a_3,\\
&\text{and so on.}
\end{aligned}$$

Then the polynomial $p(z)$ in (10), now called the *generalized Hermite interpolant*, interpolates $f(z)$ in the following sense:

$$p^{(j)}(a_i) = f^{(j)}(a_i), \quad j = 0, 1, \ldots, s_i - 1, \quad i = 1, \ldots, r.$$

Given $f^{(j)}(a_i)$, $j = 0, 1, \ldots, s_i - 1$, and $i = 1, \ldots, r$, the relevant divided difference table can be constructed via (5), (6), and (9), and $p(z)$ determined. See, for example, [5, pp. 51–59].

3 Divided difference formulas for $x_{n+1} - \alpha$

Going back to (3) and (4), we realize that they can be rewritten in terms of divided differences, respectively, as in

$$x_{n+1} = x_n - \frac{f(x_n)}{f[x_n, c]} \tag{12}$$

and

$$x_{n+1} - \alpha = -\frac{f(x_n) + f[x_n, c](\alpha - x_n)}{f[x_n, c]}. \tag{13}$$

Clearly, the numerator of the quotient on the right-hand side of (13) is related to the polynomial interpolating $f(x)$ at the points x_n and c, and from (10) and (11), we have

$$f(x) = f(x_n) + f[x_n, c](x - x_n) + f[x_n, c, x](x - x_n)(x - c). \tag{14}$$

Letting $x = \alpha$ in (14), and recalling that $f(\alpha) = 0$, we obtain

$$f(x_n) + f[x_n, c](\alpha - x_n) = -f[x_n, c, \alpha](\alpha - x_n)(\alpha - c), \tag{15}$$

and thus

$$x_{n+1} - \alpha = \frac{f[x_n, c, \alpha]}{f[x_n, c]}\,(x_n - \alpha)(c - \alpha). \tag{16}$$

Finally, we also recall that the arguments c and x_n above may coincide by the fact that divided differences are defined as limits in such a case.

We now specialize the result in (16) to the different methods mentioned above. We recall that $f \in C^2(I)$, where I is some closed interval containing α. No further differentiability properties are assumed for $f(x)$ in our treatment.

3.1 Regula falsi method

In the regula falsi method, we start with two initial points, $x_0 = c$ and x_1, such that $f(c)f(x_1) < 0$ so that $f(x) = 0$ has a solution α between c and x_1. We assume that α is the unique solution to $f(x) = 0$ between c and x_1. The point x_2 is determined as in (3), that is, x_2 is the point of intersection of the straight line passing through $(c, f(c))$ and $(x_1, f(x_1))$ with the x-axis. If $f(x_2) = 0$, then $\alpha = x_2$ and we stop. If $f(c)f(x_2) < 0$, then we leave c unchanged and continue to the next iteration; otherwise, we set $c = x_1$ and continue to the next iteration in the same way.

In case $f'(x)$ and $f''(x)$ have fixed signs in an interval containing α, which is the situation of interest to us here, the point c ultimately remains fixed. Therefore, in such a case, the regula falsi method becomes a fixed-point method at some point during the iteration process. Without loss of generality, we will assume that $c = x_0$ remains fixed.

Provided $c, x_n \in I$, by (7), the formula for x_{n+1} in (12) and the error formula (16) can be expressed as in

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(\eta_n)} \quad \text{and} \quad x_{n+1} - \alpha = \frac{f''(\xi_n)(c-\alpha)}{2f'(\eta_n)}\,(x_n - \alpha),$$
$$\xi_n \in \operatorname{int}(x_n, c, \alpha), \quad \eta_n \in \operatorname{int}(x_n, c). \tag{17}$$

If, in addition, $\lim_{n\to\infty} x_n = \alpha$ holds, then (16) gives

$$\lim_{n\to\infty} \frac{x_{n+1} - \alpha}{x_n - \alpha} = \frac{f[c, \alpha, \alpha]}{f[c, \alpha]}(c - \alpha) = \frac{f''(\overline{\xi})(c-\alpha)}{2f'(\overline{\eta})}, \quad \text{for some } \overline{\xi}, \overline{\eta} \in \operatorname{int}(c, \alpha),$$

and, as we show later, also

$$\lim_{n\to\infty} \frac{x_{n+1} - \alpha}{x_n - \alpha} = 1 - \frac{f'(\alpha)}{f[c, \alpha]},$$

which suggests that the convergence of $\{x_n\}$ may be linear. This needs to be proved rigorously, however.

As already mentioned, the situation described here happens when $f'(x)$ and $f''(x)$ have fixed signs on I, for example, when $f'(x) > 0$ and $f''(x) > 0$ on I and $f(c) > 0$, $f(x_1) < 0$. We come back to this in Section 4, where we show that the regula falsi method converges strictly linearly in this case.

3.2 Secant method

In the secant method, we start with two initial points x_0 and x_1, and for computing x_{n+1} via (3), we set $c = x_{n-1}$; that is, (3) becomes

$$x_{n+1} = x_n - \frac{f(x_n)}{\dfrac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}} = x_n - \frac{f(x_n)}{f[x_n, x_{n-1}]}.$$

The error formula (16) now is

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

and, provided $x_{n-1}, x_n \in I$, by (7), it becomes

$$x_{n+1} - \alpha = \frac{f''(\xi_n)}{2f'(\eta_n)}\,(x_n - \alpha)(x_{n-1} - \alpha),$$
$$\xi_n \in \operatorname{int}(x_n, x_{n-1}, \alpha), \quad \eta_n \in \operatorname{int}(x_n, x_{n-1}). \tag{18}$$

In case, $f'(\alpha) \neq 0$ and x_0 and x_1 are sufficiently close to α, we have $\lim_{n\to\infty} x_n = \alpha$, by Section 4 and hence

$$\lim_{n\to\infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)(x_{n-1} - \alpha)} = \frac{f''(\alpha)}{2f'(\alpha)}.$$

From this, one derives the conclusion that the order of convergence of the secant method is at least $(1 + \sqrt{5})/2$.

3.3 Newton–Raphson method

In the Newton–Raphson method, we start with one initial point x_0, and for computing x_{n+1} via (3), we set $c = x_n$; that is, x_{n+1} is the point at which the tangent line to the function $f(x)$ at x_n intersects the x-axis, and (3) becomes

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Hence, the error formula (16) now is

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

which, provided $x_n \in I$, by (7), becomes

$$x_{n+1} - \alpha = \frac{f''(\xi_n)}{2f'(x_n)}\,(x_n - \alpha)^2, \quad \xi_n \in \operatorname{int}(x_n, \alpha). \tag{19}$$

In case $f'(\alpha) \neq 0$ and x_0 is sufficiently close to α, we have, by Section 4, $\lim_{n\to\infty} x_n = \alpha$, and hence

$$\lim_{n\to\infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\alpha)}{2f'(\alpha)}.$$

From this, one derives the conclusion that the order of convergence of the Newton–Raphson method is at least two.

3.4 Steffensen method

In the Steffensen method, we start with one initial point x_0, and for computing x_{n+1} via (3), we set $c = x_n + f(x_n)$; that is, (3) becomes

$$x_{n+1} = x_n - \frac{f(x_n)}{\dfrac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}} = x_n - \frac{f(x_n)}{f[x_n, x_n + f(x_n)]}.$$

In this case, an error formula is a little more tricky to obtain. First, we note that

$$c - \alpha = x_n - \alpha + f(x_n) - f(\alpha) = \big(1 + f[x_n, \alpha]\big)(x_n - \alpha).$$

Thus, the error formula in (16) first becomes

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

Next, provided $x_n, x_n + f(x_n) \in I$, by (7), this gives

$$x_{n+1} - \alpha = \frac{f''(\xi_n)}{2f'(\eta_n)}\big[1 + f'(\theta_n)\big]\,(x_n - \alpha)^2,$$
$$\xi_n \in \operatorname{int}(x_n, x_n + f(x_n), \alpha), \quad \eta_n \in \operatorname{int}(x_n, x_n + f(x_n)), \quad \theta_n \in \operatorname{int}(x_n, \alpha). \tag{20}$$

In case, $f'(\alpha) \neq 0$ and x_0 is sufficiently close to α, by Section 4, we have $\lim_{n\to\infty} x_n = \alpha$, and hence

$$\lim_{n\to\infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^2} = \frac{f''(\alpha)}{2f'(\alpha)}\big[1 + f'(\alpha)\big].$$

From this, one derives the conclusion that the order of convergence of the Steffensen method is at least two.

Note. The definition of the Steffensen method we have given here can be found in Atkinson [1], for example. There is another completely different approach to the Steffensen method that goes through fixed-point iteration and the Aitken Δ^2-process that is used to accelerate linear convergence. In this approach, we start with the equation $x = \phi(x)$, whose solution we denote α again. Starting with x_0, an initial approximation to α, we perform the following steps:

Step 0. Set $z_0 = x_0$ and $n = 0$.

Step 1. Compute z_1 and z_2 via $z_1 = \phi(z_0)$ and $z_2 = \phi(z_1)$.

Step 2. Apply the Aitken Δ^2-process to $\{z_0, z_1, z_2\}$ to obtain x_{n+1}:

$$x_{n+1} = \frac{z_0 z_2 - z_1^2}{z_0 - 2z_1 + z_2} = z_0 - \frac{(z_1 - z_0)^2}{z_0 - 2z_1 + z_2}.$$

Step 3. Set $z_0 = x_{n+1}$ and let $n \leftarrow n + 1$, and go to Step 1.

Let us set $f(x) = \phi(x) - x$. Then α is the solution to $f(x) = 0$. Consequently, in terms of $f(z)$, we have

$$z_1 = z_0 + f(z_0) \quad \text{hence also} \quad z_1 - z_0 = f(z_0),$$

and also

$$z_2 - z_1 = f(z_1) = f(z_0 + f(z_0)).$$

Therefore,

$$z_0 - 2z_1 + z_2 = (z_2 - z_1) - (z_1 - z_0) = f(z_1) - f(z_0) = f(z_0 + f(z_0)) - f(z_0).$$

Combining all this in Step 2 above, we have

$$x_{n+1} = z_0 - \frac{[f(z_0)]^2}{f(z_0 + f(z_0)) - f(z_0)} = z_0 - \frac{f(z_0)}{\dfrac{f(z_0 + f(z_0)) - f(z_0)}{f(z_0)}}.$$

Recalling that $z_0 = x_n$ in Step 2, we finally have

$$x_{n+1} = x_n - \frac{f(x_n)}{\dfrac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}},$$

which is what we had in the beginning of this subsection. The approach we have just described can be found in [5], for example. For the derivation and convergence acceleration properties of the Aitken Δ^2-process, we refer the reader to Sidi [4, Chapter 15].

Note that, with $f(x) = \phi(x) - x$, the condition that $f'(\alpha) \neq 0$ is the same as $\phi'(\alpha) \neq 1$.

4 Completion of proofs of convergence

In this section, we show how the convergence proofs of the four methods above can be completed. The regula falsi method, with fixed c, has a special proof of its own. The remaining three methods can be shown to converge, in a unified manner, via the following simple and well-known result:

Lemma 4.1 *Let the sequence* $\{x_n\}_{n=0}^{\infty}$ *be such that*

$$x_{n+1} - \alpha = C_n(x_n - \alpha), \quad |C_n| \leq \overline{C} < 1 \quad \forall n.$$

Then, (i) $|x_{n+1} - \alpha| < |x_n - \alpha|$, (ii) $x_n \in [\alpha - \delta, \alpha + \delta]$, *where* $\delta = |x_0 - \alpha|$, (iii) $|x_n - \alpha| \leq \overline{C}^n |x_0 - \alpha|$, *and* (iv) $\lim_{n\to\infty} x_n = \alpha$.

In relation to this lemma, we recall that if $0 < \lim_{n\to\infty} |C_n| < 1$, the sequence $\{x_n\}$ converges linearly, whereas if $\lim_{n\to\infty} C_n = 0$, it is said to converge superlinearly.

Let us recall our assumption that $f \in C^2(I)$, where I is a closed interval containing α in its interior. In the treatment of the secant, Newton–Raphson, and Steffensen methods, we assume further that $f'(\alpha) \neq 0$, and choose the interval I as $I = [\alpha - \rho, \alpha + \rho]$ for some $\rho > 0$ such that $f'(x) \neq 0$ on I. Therefore, for all $x \in I$, $0 < K \leq |f'(x)| \leq L$ and $|f''(x)| \leq M$ for some positive constants K, L, and M.

By (18), (19), and (20), the $C_n = (x_{n+1} - \alpha)/(x_n - \alpha)$ in the lemma relevant to the secant, Newton–Raphson, and Steffensen methods are then

$$C_n = \frac{f''(\xi_n)}{2f'(\eta_n)}(x_{n-1} - \alpha), \quad \xi_n, \eta_n \in I, \quad \text{if } x_{n-1}, x_n \in I. \text{ (secant)}$$

$$C_n = \frac{f''(\xi_n)}{2f'(x_n)}(x_n - \alpha), \quad \xi_n \in I, \quad \text{if } x_n \in I. \text{ (Newton–Raphson)}$$

$$C_n = \frac{f''(\xi_n)}{2f'(\eta_n)}[1 + f'(\theta_n)]\,(x_n - \alpha), \quad \xi_n, \eta_n, \theta_n \in I, \quad \text{if } x_n, x_n + f(x_n) \in I. \text{ (Steffensen)}$$

Thus, by the fact that $|f''(x)/f'(y)| \leq M/K$ when $x, y \in I$, and letting $Q = M/(2K)$, there holds

$$
\begin{aligned}
&|C_n| \leq Q\,|x_{n-1} - \alpha|, \quad \text{if } x_{n-1}, x_n \in I. \quad \text{(secant)} \\
&|C_n| \leq Q\,|x_n - \alpha|, \quad \text{if } x_n \in I. \quad \text{(Newton–Raphson)} \\
&|C_n| \leq Q\,(1 + L)\,|x_n - \alpha|, \quad \text{if } x_n, x_n + f(x_n) \in I. \quad \text{(Steffensen)}
\end{aligned}
$$

We make use of these in the sequel. It is important to realize that, in order to be able to make use of these bounds on the $|C_n|$, we must show that $x_{n-1}, x_n, x_n + f(x_n) \in I$ for the relevant methods.

1. *Regula falsi method.* Let us assume that $c, x_1 \in I$, $f(c) > 0$, $f(x_1) < 0$, $f'(x) > 0$ and $f''(x) > 0$ on I. From these and from (17), it follows that c remains fixed and $x_1 < x_2 < \cdots < \alpha < c$. That is, $\{x_n\}$ is an increasing sequence bounded above by α, thus has a limit $\leq \alpha$. By the continuity of $f(x)$ and $f'(x)$ and by the assumption that $f'(x) > 0$ on I, it follows from (12) that $\lim_{n\to\infty} x_n = \alpha$. Now, by the fact that $x_n < x_{n+1} < \alpha$, we already have $0 < C_n = (x_{n+1} - \alpha)/(x_n - \alpha) < 1$ for every n. To show that the convergence is linear, we must show that $0 < \lim_{n\to\infty} C_n < 1$. Let us recall that
$$C_n = \frac{x_{n+1} - \alpha}{x_n - \alpha} = \frac{f[x_n, c, \alpha]}{f[x_n, c]}(c - \alpha).$$
First, because $x_n \in I$ for all n and $\lim_{n\to\infty} x_n = \alpha \in I$ and because of our assumptions on $f(x)$, we have that
$$\lim_{n\to\infty} C_n = \frac{f[\alpha, \alpha, c]}{f[\alpha, c]}\,(c - \alpha) = \frac{f''(\overline{\xi})(c - \alpha)}{2f'(\overline{\eta})} > 0, \quad \text{for some } \overline{\xi}, \overline{\eta} \in \operatorname{int}(c, \alpha).$$
Next, by the recursion relations among the divided differences, there holds
$$f[\alpha, \alpha, c] = \frac{f[\alpha, c] - f[\alpha, \alpha]}{c - \alpha}.$$
Therefore,
$$\lim_{n\to\infty} C_n = \frac{f[\alpha, c] - f[\alpha, \alpha]}{f[\alpha, c]} = 1 - \frac{f[\alpha, \alpha]}{f[\alpha, c]} = 1 - \frac{f'(\alpha)}{f[\alpha, c]}.$$
It is now easy to see that the slope $f[\alpha, c]$ of the straight line through $(\alpha, f(\alpha))$ and $(c, f(c))$ and the slope $f[\alpha, \alpha] = f'(\alpha)$ of the tangent to $f(x)$ at $(\alpha, f(\alpha))$ satisfy $f[\alpha, \alpha] < f[\alpha, c]$. From this, we conclude that $\lim_{n\to\infty} C_n < 1$. This completes the proof.

   We have assumed in the treatment above that $f'(x) > 0$ and $f''(x) > 0$ on I. It is easy to see that the same technique applies to all cases in which $f'(x)$ and $f''(x)$ have fixed signs on I.

2. *Secant method.* We choose x_0 and x_1 in I sufficiently close to α to ensure that the inequality
$$\overline{C} \equiv Q \max\{|x_0 - \alpha|, |x_1 - \alpha|\} < 1$$

holds. Consequently, $|C_1| \leq \overline{C}$. By $|x_2 - \alpha| = |C_1|\,|x_1 - \alpha|$, this implies that $x_2 \in I$ and $|x_2 - \alpha| < |x_1 - \alpha|$. In addition,

$$|C_2| \leq Q\,|x_2 - \alpha| < Q\,|x_1 - \alpha| \leq \overline{C},$$

that is, $|C_2| \leq \overline{C}$ as well. We can now continue by induction and show that

$$x_n \in I, \quad |x_n - \alpha| < |x_{n-1} - \alpha|, \quad \text{and} \quad |C_n| \leq \overline{C} \quad \forall n.$$

Lemma 4.1 now applies. This completes the proof of convergence for the secant method.

3. *Newton–Raphson method.* We choose x_0 sufficiently close to α to ensure that the inequality

$$\overline{C} \equiv Q\,|x_0 - \alpha| < 1$$

holds. Consequently, $|C_0| \leq \overline{C}$. By $|x_1 - \alpha| = |C_0|\,|x_0 - \alpha|$, this implies that $x_1 \in I$ and $|x_1 - \alpha| < |x_0 - \alpha|$. In addition,

$$|C_1| \leq Q\,|x_1 - \alpha| < Q\,|x_0 - \alpha| = \overline{C},$$

that is, $|C_1| \leq \overline{C}$ too. We can now continue by induction and show that

$$x_n \in I, \quad |x_n - \alpha| < |x_{n-1} - \alpha|, \quad \text{and} \quad |C_n| \leq \overline{C} \quad \forall n.$$

Lemma 4.1 now applies. This completes the proof of convergence for the the Newton–Raphson method.

4. *Steffensen method.* First, let us observe that

$$x + f(x) - \alpha = x - \alpha + f(x) - f(\alpha) = [1 + f'(\theta(x))]\,(x - \alpha)$$
$$\text{for some } \theta(x) \in \operatorname{int}(x, \alpha), \quad \text{provided } x \in I.$$

From this, it is clear that

$$|x + f(x) - \alpha| \leq (1 + L)\,|x - \alpha| \quad \text{provided } x \in I.$$

Now choose $x_0 \in I$ such that $|x_0 - \alpha| \leq \rho/(1+L)$, which guarantees that $x_0 + f(x_0) \in I$ because $|x_0 + f(x_0) - \alpha| \leq (1 + L)\,|x_0 - \alpha| \leq \rho$. Next, let us restrict x_0 further and choose it sufficiently close to α to ensure that the inequality

$$Q\,(1 + L)^2\,|x_0 - \alpha| < 1$$

holds. From this, we also have that the inequalities

$$\overline{C} \equiv Q\,(1 + L)\,|x_0 - \alpha| < 1 \quad \text{and} \quad \overline{C}\,(1 + L) < 1$$

hold as well. Now, $|C_0| \leq \overline{C}$, which, by $|x_1 - \alpha| = |C_0|\,|x_0 - \alpha|$, implies that $|x_1 - \alpha| \leq \overline{C}\,|x_0 - \alpha|$, hence $x_1 \in I$ and $|x_1 - \alpha| < |x_0 - \alpha|$. Therefore,

$$|x_1 + f(x_1) - \alpha| \leq (1 + L)\,|x_1 - \alpha| \leq \overline{C}\,(1 + L)|x_0 - \alpha| < \rho,$$

that is, $x_1 + f(x_1) \in I$ too. In addition,

$$|C_1| \leq Q\,(1+L)\,|x_1 - \alpha| < Q\,(1+L)\,|x_0 - \alpha| = \overline{C},$$

that is, $|C_1| \leq \overline{C}$ too. Continuing by induction, we can now show that

$$x_n, x_n + f(x_n) \in I, \quad |x_n - \alpha| < |x_{n-1} - \alpha|, \quad \text{and} \quad |C_n| \leq \overline{C} \quad \forall n.$$

Lemma 4.1 now applies. This completes the proof of convergence for the Steffensen method.

Acknowledgement

The author wishes to thank Miss Enav Alcobi for producing the graphs in Section 1.

References

[1] K.E. Atkinson. *An Introduction to Numerical Analysis.* Wiley, New York, second edition, 1989.

[2] P. Henrici. *Elements of Numerical Analysis.* Wiley, New York, 1964.

[3] A. Ralston and P. Rabinowitz. *A First Course in Numerical Analysis.* McGraw-Hill, New York, second edition, 1978.

[4] A. Sidi. *Practical Extrapolation Methods: Theory and Applications.* Number 10 in Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 2003.

[5] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer-Verlag, New York, third edition, 2002.