Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Vol. No. 01, 1997, pp. 23–29.
Submitted: March 15, 1997. Published: July 28, 1997.

A NEW CONVERGENCE THEOREM FOR STEFFENSEN'S METHOD ON BANACH SPACES AND APPLICATIONS

Ioannis K. Argyros

Abstract. We provide sufficient conditions for the convergence of Steffensen's method to a locally unique solution of a nonlinear operator equation on a Banach space. We use a fixed point theorem and a special choice of divided differences. Some applications of our results are also provided.

1991 A.M.S. (MOS) Subject Classification Codes. 65J15, 47H17, 49D15.

Key Words and Phrases. Steffensen's method, fixed point theorem, Banach space, divided differences.

I. INTRODUCTION

In this study we are concerned with the problem of approximating a locally unique solution x^* of the equation

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

where f is a continuous operator defined on a closed convex domain D of a Banach space E_1 with values in a Banach space E_2.

We use Steffensen's method

$$x_{n+1} = x_n - [x_n, g(x_n); f]^{-1} f(x_n) \quad (n \geq 0) \tag{2}$$

to generate a sequence converging to x^*.

Here $[x, y; f] \in L(E_1, E_2)$ is a divided difference of order one for the operator f on the points $x, y \in D$. Operator $g\colon D \subseteq E_1 \to E_1$ is given and continuous on D. For $g(x) = x$ $(x \in D)$, method (2) reduces to Newton's method [3], [4] provided that f is Fréchet-differentiable on D. If $g(x_n) = x_{n-1}$ $(n \geq 0)$ we obtain secant method [3], [5] from (2). Several other choices for g are possible, including the most popular given by $g(x) = x - f(x)$ $(x \in D)$.

Department of Mathematics, Cameron University, Lawton, OK 73505, U.S.A.
E-mail Address: ioannisa@cua.cameron.edu

Typeset by $\mathcal{AMS}$-TEX

We define the operator $P\colon D \subseteq E_1 \to E_2$ by

$$P(x) = x - [x, g(x); f]^{-1} f(x), \tag{3}$$

in which case (2) can also be written in the form

$$x_{n+1} = P(x_n) \quad (n \geq 0) \quad (x_0 \in D). \tag{4}$$

A solution x^* of equation (1) is said to be accessible from x_0 by Steffensen's method (2) if

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} P^n(x_0). \tag{5}$$

The region of accessibility of x^* by method (2) is defined to be the set of all x_0 such that (5) is true.

Let us define operator $L_f \in L(E_1, E_2)$ by

$$L_f(x) = P'(x) \quad (x \in D). \tag{6}$$

This operator is the degree of logarithmic convexity of f on x and is a measure of the convexity of the function.

Here we assume that the divided difference $[x, y; F]$ satisfies a certain Lipschitz condition that allows us to choose

$$[x, g(x); f] = \int_0^1 f'(x + t(g(x) - x))dt \quad \text{for all } x, g(x) \in D. \tag{7}$$

Steffensen's method (2) has been examined extensively under various assumptions (see [1]–[3], [6], [7] and the references there). In the above-mentioned references the authors have found a ball $U(x_0, p) = \{x \in E_1 \mid \|x - x_0\| \leq p\} \subseteq D$ such that $x^* \in U(x_0, p)$ and method (2) converges to x^* starting at $x_0 \in D$. Here using the fixed point theorem [3], [5] on (4), and (6) we provide sufficient convergence conditions for the convergence of Steffensen's method to $x^* \in U(x_0, q)$ for $q > p$. This way the region of convergence is enlarged.

II. CONVERGENCE ANALYSIS

Definition. The first order divided difference of f on the points $u, v \in E_1$ is an application $[u, v; f] \in L(E_1, E_2)$ which satisfies:

(i) $[u, v; f](v - u) = f(v) - f(u)$; (8)

(ii) if f is Fréchet differentiable at $u \in E_1$ then $[u, u; f] = f'(u)$.

Let a, b, c, α be nonnegative numbers, $x_0 \in D$ be such that the inverse of $[x_0, g(x_0); f]$ exists on D, and define parameters d, δ, β, ε by

$$d = a(1 + b), \ \delta = \|[x_0, g(x_0); f]^{-1} f(x_0)\|, \ \beta = a\|x_0 - g(x_0)\| \tag{9}$$

and

$$\varepsilon = \frac{\|x_0 - g(x_0)\|}{1 - b} \quad (\text{for } 0 \leq b < 1). \tag{10}$$

Moreover define real functions h_1, h_2, h_3, h_4 on $[0, +\infty)$ by

$$h_1(r) = r - \frac{1}{d}, \tag{11}$$

$$h_2(r) = 2d^2r^2 - d(3 + c - \beta)r + 1 - \beta - d\delta, \tag{12}$$

$$h_3(r) = 2d^2r^3 - d(c + \delta d - \beta + 3)r^2 + (1 - \beta + d\delta)r - \delta \tag{13}$$

and

$$h_4(r) = r - \varepsilon. \tag{14}$$

For our convergence theorem that follows we would like to be able to find an interval I which is a subset of the nonnegative axis such that

$$h_1(r) < 0, \tag{15}$$

$$h_2(r) > 0, \tag{16}$$

$$h_3(r) \geq 0, \tag{17}$$

$$h_4(r) \geq 0 \tag{18}$$

and

$$U(x_0, r) \subseteq D \tag{19}$$

for all $r \in I$.

The example at the end of this study shows that this can happen.

We state and prove the following semilocal convergence theorem for method (2):

Theorem 1. *Let $f: D \subseteq E_1 \to E_2$, $g: D \subseteq E_1 \to E_1$ be continuous operators, D a closed convex subset of E_1, and E_1, E_2 Banach spaces. Assume:*

(a) *f is twice and g once Fréchet-differentiable on D;*

(b) *there exist $a, b, c > 0$, $x_0 \in D$ such that*

$$\|[x_0, g(x_0); f]^{-1}([x, y; f] - [v, w; f])\| \leq a(\|x - v\| + \|y - w\|), \tag{20}$$

$$\|g(x) - g(y)\| \leq b\|x - y\|, \quad (b < 1), \tag{21}$$

$$\|[x_0, g(x_0); f]^{-1}(f(x) - f(y))\| \leq c\|x - y\| \tag{22}$$

and

$$\|[x_0, g(x_0); f]^{-1}([x, x, g(x); f] + [x, g(x), g(x); f]g'(x))\| \leq \alpha \tag{23}$$

for all $x, y, v, w \in D$;

(c) *conditions (15)–(19) hold on some interval $I \subseteq R^+ \cup \{0\}$;*

(d) *hypotheses (c) and (d) of the Proposition hold.*

Then the following hold:

(i) $$\|P'(x)\| \leq c(r) < 1 \tag{24}$$

and

(25) $$r \geq \frac{\delta}{1-c(r)}$$

for all $r \in I$, where

(26) $$c(r) = \frac{1}{(1-dr)^2}[(1-dr)(dr+\beta)+\alpha(cr+\delta)].$$

(ii) *P has a unique fixed point x^* in $U(x_0, r)$.*

(iii) *Iteration $x_{n+1} = P(x_n)$ $(n \geq 0)$ converges to x^*. Moreover the following error bound holds for all $n \geq 0$*

(27) $$\|x_n - x^*\| \leq c^n(r)r, \quad r \in I.$$

Proof. (i) Let $f(x), g(x) \in U(x_0, r)$ $(r \in I)$ whenever $x \in U(x_0, r)$. Using (3) and (7), we obtain

(28)
$$\begin{aligned}
P'(x) &= I - ([x, g(x); f]^{-1})' f(x) - [x, g(x); f]^{-1} f'(x) \\
&= I - [x, g(x); f]^{-1}[x, g(x); f]'[x, g(x); f]^{-1} f(x) \\
&\quad - \ [x, g(x); f]^{-1} f'(x) \\
&= [x, g(x); f]^{-1}[x_0, g(x_0); f][x_0, g(x_0); f]^{-1}\Big\{([x, g(x); f] - f'(x)]) \\
&\quad - \int_0^1 f''(x + t(g(x) - x))(I + t(g'(x) - I))dt \cdot \\
&\quad \cdot [x, g(x); f]^{-1}[x_0, g(x_0); f][x_0, g(x_0); f]^{-1} f(x)\Big\}.
\end{aligned}$$

By (15), (20), (21) and the estimate

$$\begin{aligned}
\|[x_0, g(x_0); f]^{-1}([x, g(x); f] - [x_0, g(x_0); f])\| &\leq a(\|x - x_0\| + \|g(x) - g(x_0)\|) \\
&\leq a(r + br) = dr < 1,
\end{aligned}$$

and from the Banach lemma on invertible operators [3], [5] it follows that $[x, g(x); f]^{-1}$ exists $(x \in U(x_0, r)$ $(r \in I))$ and

(29) $$\|[x, g(x); f]^{-1}[x_0, g(x_0); f]\| \leq (1 - dr)^{-1}.$$

Condition (21) implies that $f'(x) = [x, x; f]$ $(x \in D)$ [3], [5], and

(30)
$$\begin{aligned}
&\|[x_0, g(x_0); f]^{-1}([x, g(x); f] - [x, x; f])\| \leq a\|x - g(x)\| \\
&\leq \ a(\|x - x_0\| + \|x_0 - g(x_0)\| + \|g(x) - g(x_0)\|) \leq a(r + \|x_0 - g(x_0)\| + br).
\end{aligned}$$

Moreover by (20)–(23) and (29), the second member inside the braces in (28) is bounded above by

$$\alpha(1-dr)^{-1}(cr+\delta). \tag{31}$$

Finally using majorizations (29), (30), and (31), (28) gives

$$\|P'(x)\| \leq c(r) \quad (r \in I),$$

where function $c(r)$ is given by (26). It now follows easily from (15) and (16) that $c(r) \in [0,1)$ which shows (24). Estimate (25) reduces to showing (17) which is true by hypothesis.

(ii) It follows from (18), (25) and (2) for $n = 0$ that $\|x_1 - x_0\| \leq (1-c(r))r \leq r$, and

$$\|g(x_0) - x_0\| \leq (1-b)r < r,$$

which imply $x_1,\ g(x_0) \in U(x_0, r)$. For all positive integers smaller or equal to n assume that

$$x_n \in U(x_0, r) \quad \text{and} \quad \|x_n - x_0\| \leq (1-c^n(r))r \leq r.$$

Then we can get

$$\begin{aligned}\|g(x_n) - x_0\| &\leq \|g(x_n) - g(x_0)\| + \|g(x_0) - x_0\| \leq b\|x_n - x_0\| + \|g(x_0) - x_0\| \\ &\leq br + \|g(x_0) - x_0\| \leq r \quad \text{(by (18))}.\end{aligned}$$

That is $g(x_n) \in U(x_0, r)$. Using (4) and (24) we obtain

$$\begin{aligned}\|x_{n+1} - x_n\| &= \|P(x_n) - P(x_{n-1})\| \leq \sup_{y \in [x_{n-1}, x_n]} \|P'(y)\|\,\|x_n - x_{n-1}\| \\ &\leq c(r)\|x_n - x_{n-1}\| \quad (r \in I).\end{aligned} \tag{32}$$

Therefore we can obtain in turn

$$\|x_{n+1} - x_n\| \leq c(r)\|x_n - x_{n-1}\| \leq \cdots \leq c^n(r)\|x_1 - x_0\| = c^n(r)(1-c(r))r,$$

and

$$\begin{aligned}\|x_{n+1} - x_0\| &\leq \|x_{n+1} - x_n\| + \|x_n - x_0\| \leq c^n(r)(1-c(r))r + (1-c^n(r))r \\ &= (1-c^{n+1}(r))r \leq r.\end{aligned}$$

That is we have shown $x_n \in U(x_0, r)$ and $\|x_n - x_0\| \leq (1-c^n(r))r \leq r$ for all $n \in N$. From (32) it now follows that for $m, n \in N$

$$\|x_{n+m} - x_n\| \leq (1-c^m(r))c^n(r)r. \tag{33}$$

It follows from (33) that $\{x_n\}$ $(n \geq 0)$ is a Cauchy sequence in a Banach space E_1 and as such it converges to some limit x^*. Using the continuity of f and g, (4) gives

$P(x^*) = x^*$. To show uniqueness, assume $y^* \in U(x_0, r)$ and $P(y^*) = y^*$. Then by (24) we get

$$\begin{aligned}\|x^* - y^*\| = \|P(x^*) - P(y^*)\| &\leq \sup_{z \in [x^*, y^*]} \|P'(z)\| \, \|x^* - y^*\| \\ &\leq c(r)\|x^* - y^*\|\end{aligned}$$

which shows $x^* = y^*$.

Finally by letting $m \to \infty$ in (33) we obtain (27).

That completes the proof of Theorem 1.

The region of accessibility to x^* can be extended also to a ball around x_0 in a different way.

Theorem 2. *Under the hypotheses of Theorem* 1 *iteration* $y_{n+1} = P(y_n)$, $y_0 \in U(x_0, r)$ $(r \in I)$ *converges to* x^*. *Moreover the following error bounds hold for all* $n \geq 0$

$$\|x^* - y_n\| \leq \frac{c^n(r)}{1 - c(r)} \|y_1 - y_0\|$$

and

$$\|x^* - y_n\| \leq c^n(r)\|x^* - y_0\|.$$

Proof. For $x \in U(x_0, r)$ we obtain by (3) and (24) that

$$\|x_1 - p(x)\| = \|p(x_0) - p(x)\| \leq c(r)\|x - x_0\| \leq c(r)r.$$

That is P maps $U(x_0, r)$ into $U(x_1, c(r)r)$ with $U(x_1, c(r)r) \subseteq U(x_0, r)$. Hence $P: U(x_0, r) \to U(x_0, r)$ is a contraction, with contraction factor $c(r) < 1$.

The result follows from the contraction mapping principle [3], [5].

We complete our study with a simple example that shows that our convergence ball includes the ones found in earlier results [6], [7].

Example. Let $E_1 = E_2 = R$, $D = [-.1, .1]$, $x_0 = .01$ and define $f(x): D \to E_2$ by

$$f(x) = x - x^2. \tag{34}$$

Then we have in turn

$$g(x) = x - f(x) = x - (x - x^2) = x^2,$$

$$[x, y; f] = \frac{f(x) - f(y)}{x - y} = 1 - (x + y), \ [x, x, g(x); f] = -1,$$

$$[x, g(x), g(x); f]g'(x) = \frac{2x(x+3)}{x - 1}, \quad (x \neq 0), \quad \alpha = 1.7061206,$$

$$a = \|[x_0, g(x_0); f]^{-1}\| = \left| \frac{1 - x_0}{1 - 2x_0 + x_0^3} \right| = 1.010203,$$

$$|g(x) - g(y)| = |x + y| \cdot |x - y| \Rightarrow b = .2,$$

$$|f(x) - f(y)| = |1 - (x + y)|\,|x - y| \Rightarrow c = d = 1.4546923,$$
$$\delta = \beta = .010001 \quad \text{and} \quad \varepsilon = .012375.$$

The hypotheses of Theorem 1 are satisfied if $h_1(r) = r - .6874306 < 0$, $h_2(r) = 4.2322594r^2 - 6.465451r + .9754506 > 0$, $h_3(r) = 4.2322594r^3 - 6.4866216r^2 + 1.0045605r - .010001 \geq 0$ and $h_4(r) = r - .012375 \geq 0$ for r belonging in a suitable subset I of the nonnegative axis, and such that $U(x_0, r) \subseteq D$. It is simple algebra to show that these conditions are satisfied if we set

$$I = [.012375,\ .09]. \tag{35}$$

It follows from Theorem 2 and (35) that iteration $y_{n+1} = P(y_n)$ converges to a solution $x^* = 0$ of equation $f(x) = 0$, where f is given by (24) provided that

$$y_0 \in [-.08,\ .1] = I_A. \tag{36}$$

To compare our results with the ones obtained in [6, p. 858] (similarly for [7]), we first note that no uniqueness result was given there. Moreover the convergence ball is $U(.01,\ .039802)$. That means convergence is assured for

$$y_0 \in [-.029802,\ .049802] = I_P$$

and by (36)

$$I_P \subseteq I_A.$$

Hence we showed that there exist infinitely many new starting points $y_0 \in I_A - I_P$ for which the convergence of iteration $y_{n+1} = P(y_n)$ is assured via Theorem 2, whereas this is not the case for Theorem 1 in [6, p. 858].

REFERENCES

[1] Argyros, I.K. An error analysis for Steffensen's method, Revue d'analyse Numerique et de theorie de l'approximation (to appear).

[2] Argyros, I.K. An error analysis for the Steffensen method under generalized Zabrejko-Nguen-type assumptions, Revue d'analyse Numerique et de theorie de l'approximation, 25, 1–2, (1996), 11-22.

[3] Argyros, I.K. and Szidarovszky, F. The Theory and Application of Iteration Methods, C.R.C. Press, Inc., Boca Raton, Florida, 1993.

[4] Cătinaş, E. On some iterative methods for solving nonlinear equations, Revue d'analyse Numerique et de theorie de l'approximation, 23, 1, (1994), 47–53.

[5] Kantorovich, L.V. and Akilov, G.P. Functional Analysis in Normed Spaces, Academic Press, New York, 1978.

[6] Păvăloiu, I. Sur la methode de Steffensen pour la resolution des equations operationnelles nonlinearies, Revue Roumaine de Mathematiques pures et appliquees, XIII, 1, (1968), 149–158.

[7] Păvăloiu, I. Sur une generalisation de la methode de Steffensen, Revue d'analyse Numerique et de theorie de l'approximation, 21, 1, (1992), 59–65.