REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Rev. Anal. Numér. Théor. Approx., vol. 38 (2009) no. 2, pp. 130–143
ictp.acad.ro/jnaat

ON THE CONVERGENCE OF STEFFENSEN–TYPE METHODS USING RECURRENT FUNCTIONS

IOANNIS K. ARGYROS* and SAÏD HILOUT†

Abstract. We introduce the new idea of recurrent functions to provide a new semilocal convergence analysis for Steffensen–type methods (STM) in a Banach space setting. It turns out that our sufficient convergence conditions are weaker, and the error bounds are tighter than in earlier studies in many interesting cases [1]–[5], [12], [14]–[17], [23], [24], [26]. Applications and numerical examples, involving a nonlinear integral equation of Chandrasekhar–type, and a differential equation are also provided in this study.

MSC 2000. 65H10, 65G99, 65J15, 47H17, 49M15.
Keywords. Steffensen–type method, recurrent functions, Banach space, semilocal convergence.

1. INTRODUCTION

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

$$F(x) + G(x) = 0, \tag{1}$$

where F is a Fréchet–differentiable operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in $\mathcal{X}$, and $G : \mathcal{D} \longrightarrow \mathcal{X}$ is a continuous operator.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time–invariant system is driven by the equation $\dot{x} = T(x)$, for some suitable operator T, where x is the state. Then the equilibrium states are determined by solving equation (1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers

* Cameron University, Department of Mathematics Sciences, Lawton, OK 73505, USA, e-mail: iargyros@cameron.edu.
† Poitiers University, Laboratoire de Mathématiques et Applications, Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France, e-mail: said.hilout@math.univ-poitiers.fr.

(single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative – when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

We introduce Steffensen's–type method (STM)

$$(2)\qquad \begin{aligned} &x_{n+1} = x_n - A_n^{-1}\,(F(x_n) + G(x_n)) \quad (n \geq 0), \quad (x_0 \in \mathcal{D}),\\ &A_n := A(x_n) = [x_n, q(x_n); F]^{-1}, \end{aligned}$$

has been used to generate a sequence $\{x_n\}$ approximating $x^\star$, when $G = 0$ on $\mathcal{D}$ [4], [12], [14]–[17], [23], [24], [26].

Here, $[x, y; F]$ denotes a divided difference of order one of F at points x, y, satisfying

$$(3)\qquad [x, y; F]\,(y - x) = F(y) - F(x), \quad x, y \in \mathcal{D}, \quad x \neq y,$$

and $q : \mathcal{X} \longrightarrow \mathcal{X}$ is a continuous operator.

Note that if $G = 0$, and $q(x) = x$, then, (2) reduces to Newton's method [3], [4], [12], [20]–[23], [24]–[28]. Moreover, if $q(x) = x$, (STM) reduces to Zincenko's method [4], [28].

Case $G = 0$. If $q(x) = x$ for $x \in \mathcal{D}$ and under Kantorovich–type assumptions, Rheinboldt [25] established a convergence theorem for Newton–type method (NTM), which includes the Kantorovich theorem for the Newton method ($A(x) = F'(x)$) as a special case [1], [4], [14]. A further generalization was given by Dennis in [10], Deuflhard and Heindl in [11], [12], Potra in [23].

Miel [20], [21] improved the error bounds for Rheinboldt [25]. Moret [22] obtained a convergence theorem as well as error bounds for (NTM) under stronger conditions than those of Rheinboldt.

Case $G \neq 0$. If $A(x) = F'(x)$, ($x \in \mathcal{D}$), Zabrejko and Nguen [27] established a convergence theorem for the Krasnoselskii–Zincenko–type iteration [28]. Later, Yamamoto and Chen [8] extended the results in [27], [28], when $A(x)$ is not necessarily equal to $F'(x)$. Under the same conditions, the results were specialized in [7], [9], [18], [19] for Broyden–like methods. More recently, Argyros [5] provided a unified convergence theory for even more general (NTM).

In this study, we provide a semilocal convergence analysis for (STM) by introducing recurrent functions. This new idea leads to an analysis with weaker sufficient convergence conditions than before [1]–[5], [12], [14]–[17], [23], [24], [26].

Numerical examples and special cases are also given in this study, to show: our results can apply to solve equations, where earlier ones cannot, and also provide tighter error bounds.

2. SEMILOCAL CONVERGENCE ANALYSIS OF (STM)

We need the following result on majorizing sequences for (NTM).

LEMMA 1. *Assume: there exist constants $K > 0$, $M > 0$, $\mu \geq 0$, $L > 0$, and $\eta > 0$, such that:*

$$2\ M < K; \tag{4}$$

Quadratic polynomial f_1 given by

$$f_1(s) = 2\ L\ \eta\ s^2 - \big(2\ (1 - L\ \eta) - K\ \eta\big)\ s + 2\ (M\ \eta + \mu), \tag{5}$$

has a root in $(0, 1)$, denoted by $\frac{\delta}{2}$, and for

$$\delta_0 = \frac{K\ \eta + 2\ \mu}{1 - L\ \eta}, \tag{6}$$

$$\alpha = \frac{2\ (K - 2\ M)}{K + \sqrt{K^2 + 8\ L\ (K - 2\ M)}}, \tag{7}$$

the following holds

$$\delta_0 \leq \delta \leq 2\ \alpha. \tag{8}$$

Then, scalar sequence $\{t_n\}$ $(n \geq 0)$ given by

$$t_0 = 0, \quad t_1 = \eta, \quad t_{n+2} = t_{n+1} + \frac{K\ (t_{n+1} - t_n) + 2\ (M\ t_n + \mu)}{2\ (1 - L\ t_{n+1})}\ (t_{n+1} - t_n) \tag{9}$$

is increasing, bounded above by

$$t^{\star\star} = \frac{2\ \eta}{2 - \delta}, \tag{10}$$

and converges to its unique least upper bound $t^\star \in [0, t^{\star\star}]$.

Moreover the following estimates hold for all $n \geq 1$:

$$t_{n+1} - t_n \leq \frac{\delta}{2}\ (t_n - t_{n-1}) \leq \left(\frac{\delta}{2}\right)^n\ \eta, \tag{11}$$

and

$$t^\star - t_n \leq \frac{2\ \eta}{2 - \delta}\ \left(\frac{\delta}{2}\right)^n. \tag{12}$$

Proof. We shall show using induction on the integer m:

$$\begin{aligned} 0 \ &< \ t_{m+2} - t_{m+1} = \frac{K\ (t_{m+1} - t_m) + 2\ (M\ t_m + \mu)}{2\ (1 - L\ t_{m+1})}\ (t_{m+1} - t_m) \\ &\leq \ \frac{\delta}{2}\ (t_{m+1} - t_m), \end{aligned} \tag{13}$$

and

$$L\ t_{m+1} < 1. \tag{14}$$

If (13) and (14) hold, we have (11) holds, and

$$\begin{array}{rcl} t_{m+2} & \leq & t_{m+1} + \frac{\delta}{2}\ (t_{m+1} - t_m) \\ & \leq & t_m + \frac{\delta}{2}\ (t_m - t_{m-1}) + \frac{\delta}{2}\ (t_{m+1} - t_m) \\ & \leq & \eta + \left(\frac{\delta}{2}\right)\ \eta + \cdots + \left(\frac{\delta}{2}\right)^{m+1}\ \eta \\ & = & \frac{1-\left(\frac{\delta}{2}\right)^{m+2}}{1-\frac{\delta}{2}}\ \eta < \frac{2\ \eta}{2-\delta} = t^{\star\star} \qquad \text{by (10).} \end{array}$$

Estimates (13) and (14) hold by the initial conditions for $m = 0$. Indeed (13) and (14) become:

$$\begin{array}{rcl} 0 < t_2 - t_1 & = & \frac{K\ (t_1 - t_0) + 2\ (M\ t_0 + \mu)}{2\ (1 - L\ t_1)}\ (t_1 - t_0) \\ & = & \frac{K\ \eta + 2\ \mu}{2\ (1 - L\ \eta)}\ (t_1 - t_0) = \frac{\delta_0}{2}\ (t_1 - t_0) \leq \frac{\delta}{2}\ (t_1 - t_0), \\ L\ \eta < 1, & & \end{array}$$

which are true by the choices of δ_0, and δ, (8), (9), and the initial conditions. Let us assume (11), (13) and (14) hold for all $m \leq n + 1$.

Estimate (13) can be re-written as

$$K\ (t_{m+1} - t_m) + 2\ (M\ t_m + \mu) \leq (1 - L\ t_{m+1})\ \delta$$

or

$$K\ (t_{m+1} - t_m) + 2\ (M\ t_m + \mu) + \delta\ L\ t_{m+1} \leq 0, \tag{15}$$

or

$$K\ \left(\frac{\delta}{2}\right)^m\ \eta + 2\ \left(M\ \frac{1-\left(\frac{\delta}{2}\right)^m}{1-\frac{\delta}{2}}\ \eta + \mu\right) + \delta\ L\ \frac{1-\left(\frac{\delta}{2}\right)^{m+1}}{1-\frac{\delta}{2}}\ \eta - \delta \leq 0. \tag{16}$$

Replace $\frac{\delta}{2}$ by s, and define functions f_m on $[0, +\infty)$ $(m \geq 1)$:

$$\begin{array}{rcl} f_m(s) & = & K\ s^m\ \eta + 2\ \left[M\ (1 + s + s^2 + \cdots + s^{m-1})\ \eta + \mu\right] \\ & & +2\ s\ L\ (1 + s + \cdots + s^m)\ \eta - 2\ s. \end{array} \tag{17}$$

We need to find a relationship between two consecutive f_m:

(18)

$$\begin{array}{rcl} f_{m+1}(s) & = & K\ s^{m+1}\ \eta + 2\ \left(M\ (1 + s + s^2 + \cdots + s^{m-1} + s^m)\ \eta + \mu\right) \\ & & \quad +2\ s\ L\ (1 + s + \cdots + s^m + s^{m+1})\ \eta - 2\ s \\ & = & K\ s^{m+1}\ \eta - K\ s^m\ \eta + K\ s^m\ \eta \\ & & \quad +2\ \left(M\ (1 + s + s^2 + \cdots + s^{m-1})\ \eta + \mu\right) \\ & & \quad +2\ M\ s^m\ \eta + 2\ s\ L\ (1 + s + \cdots + s^m)\ \eta + 2\ s\ L\ s^{m+1}\ \eta - 2\ s \\ & = & f_m(s) + K\ s^{m+1}\ \eta - K\ s^m\ \eta + 2\ M\ s^m\ \eta + 2\ s\ L\ s^{m+1}\ \eta \\ & = & f_m(s) + g(s)\ s^m\ \eta, \end{array}$$

where

$$g(s) = 2\ L\ s^2 + K\ s + 2\ M - K. \tag{19}$$

Note that in view of (4), function g has a unique positive zero α given by (7), and

$$g(s) < 0 \qquad s \in (0, \alpha). \tag{20}$$

Estimate (16) certainly holds, if

$$f_m(\tfrac{\delta}{2}) \leq 0 \qquad (m \geq 1). \tag{21}$$

Clearly, (21) holds for $m = 1$ as equality. We then get by (17), (18)–(20):

$$f_2(\tfrac{\delta}{2}) = f_1(\tfrac{\delta}{2}) + g(\tfrac{\delta}{2})\,\tfrac{\delta}{2}\,\eta = g(\tfrac{\delta}{2})\,\tfrac{\delta}{2}\,\eta \leq 0,$$

Assume (21) holds for all $k \leq m$. We shall show (21) for m replaced by $m + 1$. Indeed, we have

$$f_{m+1}(\tfrac{\delta}{2}) = f_m(\tfrac{\delta}{2}) + g(\tfrac{\delta}{2})\ \left(\tfrac{\delta}{2}\right)^m \eta \leq 0,$$

which shows (21) for all m.

Moreover, we obtain

$$f_\infty(\tfrac{\delta}{2}) = \lim_{m \to \infty} f_m(\tfrac{\delta}{2}) \leq 0.$$

That completes the induction.

Moreover, estimate (12) follows from (11) by using standard majorization techniques [2], [4], [14].

Finally, sequence $\{t_n\}$ is increasing, bounded above by $t^{\star\star}$, and as such it converges to its unique least upper bound $t^\star$.

That completes the proof of Lemma 1. □

REMARK 2. The hypotheses of Lemma 1 have been left as uncluttered as possible. Note that these hypotheses involve only computations only at the initial point x_0. Below, we shall provide some simpler but stronger hypotheses under which the hypotheses of Lemma 1 hold. □

We can also show the following alternative to Lemma 1.

LEMMA 3. *Let $K > 0$, $M > 0$, $\mu > 0$, with $L > 0$, and $\eta > 0$, be such that:*

$$\mu < \alpha, \quad 2\ M < K,$$

and

$$0 < h_A = a\ \eta \leq \tfrac{1}{2}, \tag{22}$$

where

$$a = \tfrac{1}{4\ (\alpha - \mu)}\ \max\{2\ L\ \alpha^2 + 2\ L\ \alpha + K\ \alpha + 2\ M,\ K + 2\ \alpha\ L\}. \tag{23}$$

Then, the following hold:

f_1 has a positive root $\frac{\delta}{2}$, $\max\{\delta_0, \delta\} \leq 2\ \alpha$,

and

the conclusions of Lemma 1 hold, with α replacing $\frac{\delta}{2}$.

Proof. It follows from (18), (22) that

$$f_m(\alpha) = f_1(\alpha) \leq 0, \quad m \geq 1, \tag{24}$$

which together with $f_1(0) = 2\ (M\ \eta + \mu) > 0$, imply that there exists a positive root $\frac{\delta}{2}$ of polynomial f_1, satisfying

$$\delta \leq 2\ \alpha. \tag{25}$$

It also follows from (6), and (22) that

$$\delta_0 \leq 2\ \alpha, \tag{26}$$

and (21) holds, with α replacing $\frac{\delta}{2}$ (by (24)). Note also $\alpha \in (0,1)$ by (4).
That completes the proof of Lemma 3 $\square$

We shall show the following semilocal convergence theorem for (STM).

THEOREM 4. *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{X}$ be a Fréchet–differentiable operator, $G : \mathcal{D} \longrightarrow \mathcal{X}$ be a continuous operator, $[x, y; F]$ be a divided difference of order one of F on $\mathcal{D}$, satisfying* (3)*, $q : \mathcal{D} \longrightarrow \mathcal{X}$ a continuous operator, and let $A(x) \in \mathcal{L}(\mathcal{X}, \mathcal{X})$ given in* (2) *be an approximation of $F'(x)$. Assume that there exist an open convex subset $\mathcal{D}$ of $\mathcal{X}$, $x_0 \in \mathcal{D}$, a bounded inverse A_0^{-1} of A_0, and constants $K > 0$, $M_0 > 0$, $\beta \geq 0$, $b \geq 0$, $\mu_0 \geq 0$, $\mu_1 \geq 0$, $\eta > 0$, parameter $\alpha \in [0, 1)$, such that for all $x, y \in \mathcal{D}$:*

$$\|A_0^{-1}\ [F(x_0) + G(x_0)]\| \leq \eta, \tag{27}$$

$$\|A_0^{-1}\ [F'(x) - F'(y)]\| \leq K\ \|x - y\|, \tag{28}$$

$$\|A_0^{-1}\ [F'(x) - A(x)]\| \leq M_0\ \|x - q(x)\| + \mu_0, \tag{29}$$

$$\|A_0^{-1}\ [A(x) - A_0]\| \leq \beta\ (\|x - x_0\| + \|q(x) - q(x_0)\|), \tag{30}$$

$$\|A_0^{-1}\ [G(x) - G(y)]\| \leq \mu_1\ \|x - y\|, \tag{31}$$

$$\|q(x) - q(x_0)\| \leq b\ \|x - x_0\|, \tag{32}$$

$$t^\star \geq \tfrac{\|x_0 - q(x_0)\|}{1 - b}, \tag{33}$$

$$\overline{U}(x_0, t^\star) = \{x \in \mathcal{X},\ \|x - x_0\| \leq t^\star\} \subseteq \mathcal{D}, \tag{34}$$

and the hypotheses of Lemmas 1*, or* 3 *hold with*

$$\mu = \mu_0 + \mu_1, \qquad M = M_0\ (1 + b),$$

$$\mu_0 = M_0\ \|x_0 - q(x_0)\ \|, \quad \text{and} \quad L = (1 + b)\ \beta.$$

Then, sequence $\{x_n\}$ $(n \geq 0)$ generated by (STM) *is well defined, remains in $\overline{U}(x_0, t^\star)$ for all $n \geq 0$, and converges to a solution $x^\star$ of equation $F(x) + G(x) = 0$ in $\overline{U}(x_0, t^\star)$.*

Moreover, the following estimates hold for all $n \geq 0$:

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \tag{35}$$

and

$$\|x_n - x^\star\| \leq t^\star - t_n, \tag{36}$$

where sequence $\{t_n\}$ $(n \geq 0)$*, and* $t^\star$ *are given in Lemma* 1.

Furthermore, the solution $x^\star$ *of equation* (1) *is unique in* $\overline{U}(x_0, t^\star)$ *provided that:*

$$\left(\tfrac{K}{2} + M + L\right) t^\star + \mu < 1.$$

Proof. We shall show using induction on $m \geq 0$:

$$\|x_{m+1} - x_m\| \leq t_{m+1} - t_m, \tag{37}$$

and

$$\overline{U}(x_{m+1}, t^\star - t_{m+1}) \subseteq \overline{U}(x_m, t^\star - t_m). \tag{38}$$

For every $z \in \overline{U}(x_1, t^\star - t_1)$,

$$\begin{array}{rcl} \|z - x_0\| & \leq & \|z - x_1\| + \|x_1 - x_0\| \\ & \leq & t^\star - t_1 + t_1 = t^\star - t_0, \end{array}$$

implies $z \in \overline{U}(x_0, t^\star - t_0)$. We also have

$$\|x_1 - x_0\| = \|A_0^{-1}\,[F(x_0) + G(x_0)]\| \leq \eta = t_1 - t_0.$$

That is (37) and (38) hold for $m = 0$. Given they hold for $n \leq m$, then:

$$\begin{array}{rcl} \|x_{m+1} - x_0\| & \leq & \displaystyle\sum_{i=1}^{m+1} \|x_i - x_{i-1}\| \\ & \leq & \displaystyle\sum_{i=1}^{m+1} (t_i - t_{i-1}) = t_{m+1} - t_0 = t_{m+1}, \end{array}$$

and

$$\begin{array}{rcl} \|x_m + \theta\,(x_{m+1} - x_m) - x_0\| & \leq & t_m + \theta\,(t_{m+1} - t_m) \\ & \leq & t^\star, \end{array}$$

for all $\theta \in (0, 1)$.

We have:

$$\begin{array}{rcl} \|q(x_{m+1}) - x_0\| & \leq & \|q(x_{m+1}) - q(x_0)\| + \|q(x_0) - x_0\| \\ & \leq & b\ \|x_{m+1} - x_0\| + \|q(x_0) - x_0\| \\ & \leq & b\ (t_{m+1} - t_0) + \|q(x_0) - x_0\| \\ & \leq & b\ t^\star + \|q(x_0) - x_0\| \leq t^\star < 1 \end{array}$$

by (33).

Using (30), (32), and the induction hypotheses, we get:

$$\begin{array}{rcl} \|A_0^{-1}\ [A_{m+1} - A_0]\| & \leq & \beta\ (\|x_{m+1} - x_0\| + \|q(x_{m+1}) - q(x_0)\|) \\ & \leq & \beta\ (\|x_{m+1} - x_0\| + b\|x_{m+1} - x_0\|) \\ & = & L\ \|x_{m+1} - x_0\| \\ & \leq & L\ (t_{m+1} - t_0) = L\ t_{m+1} < 1. \end{array} \tag{39}$$

It follows from (39), and the Banach lemma on invertible operators [4], [14] that A_{m+1}^{-1} exists, and

$$\|A_{m+1}^{-1}\ A_0\| \leq (1 - L\ t_{m+1})^{-1}. \tag{40}$$

We have for $x \in \overline{U}(x_0, t^\star)$:

(41)
$$\begin{aligned} \|A_0^{-1}\ [F'(x) - A(x)]\| &\le M_0\ \|x - q(x)\| \\ &\le M_0\ \|(x-x_0)+(x_0-q(x_0))+(q(x_0)-q(x))\| \\ &\le M_0\ (\|x-x_0\|+S\|x_0-q(x_0)\| + b\ \|x-x_0\|) \\ &\le M_0\ ((1+b)\ \|x-x_0\| + \|x_0 - q(x_0)\|) \\ &= M\ \|x - x_0\| + \mu_0. \end{aligned}$$

Using (2), we obtain the approximation:

(42)
$$\begin{aligned} x_{m+2} - x_{m+1} &= -A_{m+1}^{-1}\ (F(x_{m+1} + G(x_{m+1})) \\ &= -A_{m+1}^{-1}\ A_0\ A_0^{-1} \\ &\quad \Big(\int_0^1 [F'(x_{m+1}+\theta\,(x_m - x_{m+1})) - F'(x_m)]\ (x_{m+1}-x_m)\ \mathrm{d}\theta \\ &\quad +(F'(x_m) - A_m)\ (x_{m+1} - x_m) + G(x_{m+1}) - G(x_m)\Big) \end{aligned}$$

Using (28), (29), (31), (40)–(42), and the induction hypotheses, we obtain in turn:

(43)
$$\begin{aligned} \|x_{m+2} - x_{m+1}\| &\le (1 - L\ t_{m+1})^{-1}\ \big(\tfrac{K}{2}\ \|x_{m+1} - x_m\|^2 \\ &\quad +(M\|x_m-x_0\|+\mu_0)\|x_{m+1}-x_m\| + \mu_1\|x_{m+1}-x_m\ \|\big) \\ &\le (1-L\ t_{m+1})^{-1}\big(\tfrac{K}{2}(t_{m+1}-t_m)+Mt_m+\mu\big)(t_{m+1}-t_m) \\ &= t_{m+2} - t_{m+1}, \end{aligned}$$

which shows (37) for all $m \ge 0$.

Thus, for every $z \in \overline{U}(x_{m+2}, t^\star - t_{m+2})$, we have:

$$\begin{aligned} \|z - x_{m+1}\| &\le \|z - x_{m+2}\| + \|x_{m+2} - x_{m+1}\| \\ &\le t^\star - t_{m+2} + t_{m+2} - t_{m+1} = t^\star - t_{m+1}, \end{aligned}$$

which shows (38) for all $m \ge 0$.

Lemmas 1 or 3 imply that sequence $\{t_n\}$ is Cauchy. Moreover, it follows from (37) and (38) that $\{x_n\}$ $(n \ge 0)$ is also a Cauchy sequence in a Banach space $\mathcal{X}$, and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star)$ (since $\overline{U}(x_0, t^\star)$ is a closed set).

We also have by (42) and (43):

$$\begin{aligned} &\big\|A_0^{-1}(F(x_{m+1} + G(x_{m+1}))\big\| \\ &\le \left[\frac{K}{2}(t_{m+1} - t_m) + Mt_m + \mu\right](t_{m+1} - t_m) \\ &\le \left(\frac{K}{2}(t_{m+1} - t_m) + Mt^* + \mu\right)(t_{m+1} - t_m) \to 0 \end{aligned}$$

as $m \to \infty$.

Hence, by the continuity of operators F and G we obtain

$$F(x^*) + G(x^*) = 0.$$

Furthermore estimate (36) is obtained from (35) by using standard majorization techniques [2], [4], [14]. Finally to show that $x^\star$ is the unique solution of equation (1) in $\overline{U}(x_0, t^\star)$, as in (42) and (43), we get in turn for $y^\star \in \overline{U}(x_0, t^\star)$, with $F(y^\star) + G(y^\star) = 0$, the estimation:

(44)
$$\begin{aligned}
\|y^\star - x_{m+1}\| \leq\ & \|A_m^{-1}\ A_0\| \\
& \Big\{ \Big(\int_0^1 \|A_0^{-1}\ (F'(x_m + \theta\,(y^\star - x_m)) - F'(x_m))\|\ \mathrm{d}\theta \\
& + \|A_0^{-1}\ [F'(x_m) - A_m]\|\Big)\ \|y^\star - x_m\| + \\
& \|A_0^{-1}\ [G(x_m) - G(y^\star)]\| \Big\} \\
\leq\ & (1 - L\ t_{m+1})^{-1}\ \big(\tfrac{K}{2}\ \|y^\star - x_m\|^2 + \\
& (M\ \|x_m - x_0\| + \mu)\ \|y^\star - x_m\|\big) \\
\leq\ & (1 - L\ t_{m+1})^{-1} \left(\tfrac{K}{2}\ (t^\star - t_m) + M\ t_m + \mu\right) \|y^\star - x_m\| \\
\leq\ & (1 - L\ t^\star)^{-1} \left(\tfrac{K}{2}\ (t^\star - t_0) + M\ t^\star + \mu\right) \|x^\star - x_m\| \\
<\ & \|y^\star - x_m\|,
\end{aligned}$$

by the uniqueness hypothesis.

It follows by (44) that $\lim\limits_{m \longrightarrow \infty} x_m = y^\star$. But we showed $\lim\limits_{m \longrightarrow \infty} x_m = x^\star$. Hence, we deduce $x^\star = y^\star$.

That completes the proof of Theorem 4. $\square$

Note that $t^\star$ can be replaced by $t^{\star\star}$ given by (10) in the uniqueness hypothesis provided that $\overline{U}(x_0, t^{\star\star}) \subseteq \mathcal{D}$, or in all hypotheses of the theorem.

Remark 5. A direct comparison between earlier results and ours is not possible at this generality. Let us then set $G = 0$ on $\mathcal{D}$. Păvăloiu [15]–[17] has made an extensive research on Steffensen's method in this case, and under various conditions. In particular, in [17], Păvăloiu extended some results of ours [1] from the Secant method to Steffensen's method using the set of hypotheses:

$$\|[y, u; F] - [x, y; F]\| \leq c_1 \|x - u\| + c_2 \|x - y\| + c_3 \|u - y\|, \tag{45}$$

$$\|[x, y; F]^{-1}\| \leq c_4, \tag{46}$$

$$\|F(q(x))\| \leq c_5 \|F(x)\|^{c_6}, \tag{47}$$

$$\|x - q(x)\| \leq c_7 \|F(x)\|, \tag{48}$$

to provide the semilocal convergence theorem for Steffensen's method.

Here, some advantages of our approach:

(a) Our results are given in affine invariant form. The advantages of affine invariant over non–affine invariant results have been explained in [4], [12].
(b) Hypotheses (28)–(30), and (32) are simpler, and weaker than (45)–(48).

In a future paper, we hope to find a result, similar to Lemma 1, but where conditions (45)–(48) are used, so we can have a more direct comparison between our results and the corresponding ones in [17]. $\square$

3. SPECIAL CASES AND APPLICATIONS

APPLICATION 6. *Let $q(x)=x$, $G(x)=0$ $(x\in\mathcal{D})$. We can now compare our Theorem* 4 *with the Newton–Kantorovich theorem for solving equations in the interesting case of Newton's method* [4], [14].

The famous for its simplicity and clarity Newton–Kantorovich hypothesis for solving nonlinear equations

$$h_K = K\ \eta \le \tfrac{1}{2}. \tag{49}$$

Note that in this case, functions f_m $(m\ge 1)$ should be defined by

$$f_m(s) = \big(K\ s^{m-1} + 2\ L\ (1+s+s^2+\cdots+s^m)\big)\ \eta - 2, \tag{50}$$

and

$$f_{m+1}(s) = f_m(s) + g(s)\ s^{m-1}\ \eta. \tag{51}$$

However, it is simple algebra to show that conditions (8) reduces to:

$$h_A = \overline{L}\ \eta \le \tfrac{1}{2}, \tag{52}$$

where

$$\overline{L} = \tfrac{1}{8}\ \big(K + 4\ L + \sqrt{K^2 + 8\ K\ L}\big).$$

Note also that

$$L \le K \tag{53}$$

holds in general, and $\frac{K}{L}$ can be arbitrarily large [4].

In view of (49), (52) and (53), we get

$$h_K \le \tfrac{1}{2} \Longrightarrow h_A \le \tfrac{1}{2}, \tag{54}$$

but not necessarily vice versa unless if $L=K$.

In the example that follows, we show that $\frac{K}{L}$ can arbitrarily large. Indeed:

EXAMPLE 7. Let $\mathcal{X}=\mathcal{Y}=\mathbb{R}$, $x_0=1$, and define scalar functions F and G by

$$F(x) = c_0\ x + c_1 + c_2\ \sin \mathrm{e}^{c_3\ x}, \qquad G(x)=0, \tag{55}$$

where c_i, $i=0,1,2,3$ are given parameters. Using (55), it can easily be seen that for c_3 large and c_2 sufficiently small, $\frac{K}{L}$ can be arbitrarily large. $\square$

In the next examples, we show (49) is violated but (52) holds.

Example 8. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$, equipped with the max–norm, and

$$x_0 = (1,1)^{\mathrm{T}}, \quad U_0 = \{x \,:\, \|x - x_0\| \le 1 - c\}, \quad c \in \left[0, \tfrac{1}{2}\right).$$

Define function F on U_0 by

$$F(x) = (\xi_1^3 - c, \xi_2^3 - c), \qquad x = (\xi_1, \xi_2)^{\mathrm{T}}. \tag{56}$$

The Fréchet–derivative of operator F is given by

$$F'(x) = \begin{bmatrix} 3\,\xi_1^2 & 0 \\ 0 & 3\,\xi_2^2 \end{bmatrix}.$$

Using hypotheses of Theorem 4, we get:

$$\eta = \tfrac{1}{3}\ (1 - c), \quad L = 3 - c, \quad \text{and} \quad K = 2\ (2 - c).$$

The Newton–Kantorovich condition (49) is violated, since

$$\tfrac{4}{3}\ (1 - c)\ (2 - c) > 1 \quad \text{for all} \quad c \in \left[0, \tfrac{1}{2}\right).$$

Hence, there is no guarantee that Newton's method (2) converges to $x^\star = (\sqrt[3]{c}, \sqrt[3]{c})^{\mathrm{T}}$, starting at x_0.

However, our condition (52) is true for all $c \in I = \left[.450339002, \tfrac{1}{2}\right)$. Hence, the conclusions of our Theorem 4 can apply to solve equation (56) for all $c \in I$. □

Example 9. Let $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$ be the space of real–valued continuous functions defined on the interval $[0,1]$ with norm

$$\|x\| = \max_{0 \le s \le 1} |x(s)|.$$

Let $\theta \in [0,1]$ be a given parameter. Consider the "Cubic" integral equation

$$u(s) = u^3(s) + \lambda\, u(s) \int_0^1 q(s,t)\, u(t)\, \mathrm{d}t + y(s) - \theta. \tag{57}$$

Here the kernel $q(s,t)$ is a continuous function of two variables defined on $[0,1] \times [0,1]$; the parameter λ is a real number called the "albedo" for scattering; $y(s)$ is a given continuous function defined on $[0,1]$ and $x(s)$ is the unknown function sought in $\mathcal{C}[0,1]$. Equations of the form (57) arise in the kinetic theory of gasses [4], [6]. For simplicity, we choose $u_0(s) = y(s) = 1$, and $q(s,t) = \frac{s}{s+t}$, for all $s \in [0,1]$, and $t \in [0,1]$, with $s + t \ne 0$. If we let $\mathcal{D} = U(u_0, 1 - \theta)$, and define the operator F on $\mathcal{D}$ by

$$F(x)(s) = x^3(s) - x(s) + \lambda\, x(s) \int_0^1 q(s,t)\, x(t)\, \mathrm{d}t + y(s) - \theta, \tag{58}$$

for all $s \in [0,1]$, then every zero of F satisfies equation (57).

We have the estimates:

$$\max_{0\le s\le 1}\left|\int \tfrac{s}{s+t}\,\mathrm{d}t\right| = \ln 2.$$

Therefore, if we set $\xi = \|F'(u_0)^{-1}\|$, then it follows from hypotheses of Theorem 4 that

$$\eta = \xi\ (|\lambda|\ \ln 2 + 1 - \theta),$$

$$K = 2\ \xi\ (|\lambda|\ \ln 2 + 3\,(2-\theta)) \quad \text{and} \quad L = \xi\ (2\ |\lambda|\ \ln 2 + 3\,(3-\theta)).$$

It follows from Theorem 4 that if condition (52) holds, then problem (57) has a unique solution near u_0. This assumption is weaker than the one given before using the Newton–Kantorovich hypothesis (49).

Note also that $L < K$ for all $\theta \in [0,1]$. □

Example 10. Consider the following nonlinear boundary value problem [4]

$$\begin{cases} u'' = -u^3 - \gamma\ u^2 \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 Q(s,t)\ (u^3(t) + \gamma\ u^2(t))\ \mathrm{d}t \tag{59}$$

where Q is the Green function:

$$Q(s,t) = \begin{cases} t\ (1-s), & t \le s \\ s\ (1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0\le s\le 1}\int_0^1 |Q(s,t)| = \tfrac{1}{8}.$$

Let $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$, with norm

$$\|x\| = \max_{0\le s\le 1}|x(s)|.$$

Then problem (59) is in the form (1), where $F : \mathcal{D} \longrightarrow \mathcal{Y}$ is defined as

$$[F(x)]\,(s) = x(s) - s - \int_0^1 Q(s,t)\ (x^3(t) + \gamma\ x^2(t))\ \mathrm{d}t,$$

and

$$G(x)\,(s) = 0.$$

It is easy to verify that the Fréchet derivative of F is defined in the form

$$[F'(x)v]\,(s) = v(s) - \int_0^1 Q(s,t)\ (3\ x^2(t) + 2\ \gamma\ x(t))\ v(t)\ \mathrm{d}t.$$

If we set $u_0(s) = s$, and $\mathcal{D} = U(u_0, R)$, then since $\|u_0\| = 1$, it is easy to verify that $U(u_0, R) \subset U(0, R+1)$. It follows that $2\gamma < 5$, then

$$\begin{aligned} \|I - F'(u_0)\| &\leq \tfrac{3\|u_0\|^2 + 2\gamma\|u_0\|}{8} = \tfrac{3+2\gamma}{8}, \\ \|F'(u_0)^{-1}\| &\leq \tfrac{1}{1 - \frac{3+2\gamma}{8}} = \tfrac{8}{5-2\gamma}, \\ \|F(u_0)\| &\leq \tfrac{\|u_0\|^3 + \gamma\|u_0\|^2}{8} = \tfrac{1+\gamma}{8}, \\ \|F(u_0)^{-1}\ F(u_0)\| &\leq \tfrac{1+\gamma}{5-2\ \gamma}. \end{aligned}$$

On the other hand, for $x, y \in \mathcal{D}$, we have

$$[(F'(x) - F'(y))v]\,(s) = -\int_0^1 Q(s,t)(3x^2(t) - 3\ y^2(t) + 2\gamma\ (x(t) - y(t)))v(t)\mathrm{d}t.$$

Consequently,

$$\begin{aligned} \|F'(x) - F'(y)\| &\leq \tfrac{\|x-y\|(2\gamma + 3(\|x\| + \|y\|))}{8} \\ &\leq \tfrac{\|x-y\|(2\gamma + 6R + 6\|u_0\|)}{8} \\ &= \tfrac{\gamma + 6R + 3}{4}\|x - y\|, \\ \|F'(x) - F'(u_0)\| &\leq \tfrac{\|x-u_0\|(2\gamma + 3(\|x\| + \|u_0\|))}{8} \\ &\leq \tfrac{\|x-u_0\|(2\gamma + 3R + 6\|u_0\|)}{8} \\ &= \tfrac{2\gamma + 3R + 6}{8}\|x - u_0\|. \end{aligned}$$

Therefore, conditions of Theorem 4 hold with

$$\eta = \tfrac{1+\gamma}{5-2\gamma}, \quad K = \tfrac{\gamma + 6R + 3}{4}, \quad L = \tfrac{2\gamma + 3R + 6}{8}.$$

Note also that $L < K$. □

REFERENCES

[1] Argyros, I.K., *The secant method and fixed points of nonlinear operators*, Monatshefte für Math., **106**, pp. 85–94, 1988.

[2] Argyros, I.K., *The theory and application of abstract polynomial equations*, St.Lucie/CRC/Lewis Publ. Mathematics series, 1998, Boca Raton, Florida, U.S.A.

[3] Argyros, I.K., *A unifying local–semilocal convergence analysis and applications for two–point Newton–like methods in Banach space*, J. Math. Anal. Appl., **298**, pp. 374–397, 2004.

[4] Argyros, I.K., *Computational theory of iterative methods, Series: Studies in Computational Mathematics*, **15**, Editors: C.K. Chui and L. Wuytack, Elsevier Publ. Co., New York, USA, 2007.

[5] Argyros, I.K., *On a class of Newton–like methods for solving nonlinear equations*, J. Comput. Appl. Math., **228**, pp. 115–122, 2009.

[6] Chandrasekhar, S., *Radiative transfer*, Dover Publ., New York, 1960.

[7] Chen, X., *On the convergence of Broyden–like methods for nonlinear equations with nondifferentiable terms*, Ann. Inst. Statist. Math., **42**, pp. 387–401, 1990.

[8] Chen, X. and Yamamoto, T., *Convergence domains of certain iterative methods for solving nonlinear equations*, Numer. Funct. Anal. Optim., **10**, pp. 37–48, 1989.

[9] Chen, Y. and Cai, D., *Inexact overlapped block Broyden methods for solving nonlinear equations*, Appl. Math. Comput., **136**, pp. 215–228, 2003.

[10] Dennis, J.E., *Toward a unified convergence theory for Newton–like methods*, Nonlinear Functional Analysis and Applications (L.B. Rall, ed.), Academic Press, New York, pp. 425–472, 1971.

[11] Deuflhard, P., *Newton methods for nonlinear problems. Affine invariance and adaptive algorithms*, Springer Series in Computational Mathematics, **35**, Springer–Verlag, Berlin, 2004.

[12] Deuflhard, P. and Heindl, G., *Affine invariant convergence theorems for Newton's method and extensions to related methods*, SIAM J. Numer. Anal., **16**, pp. 1–10, 1979.

[13] Huang, Z., *A note of Kantorovich theorem for Newton iteration*, J. Comput. Appl. Math., **47**, pp. 211–217, 1993.

[14] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[15] Păvăloiu, I., *Sur la méthode de Steffensen pour la résolution des équations opérationnelles non linéaires*, Rev. Roumaine Math. Pures Appl., **13**(6), pp. 857–861, 1968.

[16] Păvăloiu, I., *Introduction in the theory of approximation of equations solutions*, Dacia Ed., Cluj–Napoca, 1976.

[17] Păvăloiu, I., On the convergency of a Steffensen–type method, Seminar on Mathematical Analysis, pp. 121–126, Preprint, pp. 91–7, "Babeş–Bolyai" Univ., Cluj–Napoca, 1991.

[18] Li, D. and Fukushima, M., *Globally convergent Broyden–like methods for semismooth equations and applications to VIP, NCP and MCP*, Optimization and numerical algebra (Nanjing, 1999), Ann. Oper. Res., **103**, pp. 71–97, 2001.

[19] Ma, C., *A smoothing Broyden–like method for the mixed complementarity problems*, Math. Comput. Modelling, **41**, pp. 523–538, 2005.

[20] Miel, G.J., *Unified error analysis for Newton–type methods*, Numer. Math., **33**, pp. 391–396, 1979.

[21] Miel, G.J., *Majorizing sequences and error bounds for iterative methods*, Math. Comp., **34**, pp. 185–202, 1980.

[22] Moret, I., *A note on Newton type iterative methods*, Computing, **33**, pp. 65–73, 1984.

[23] Potra, F.A., *Sharp error bounds for a class of Newton–like methods, Libertas Mathematica*, **5**, pp. 71–84, 1985.

[24] Ulm, S. Ju., *A generalization of Steffensen's method for solving non–linear operator equations*, (Russian), Ž. Vyčisl. Mat. i Mat. Fiz., **4**, pp. 1093–1097, 1964.

[25] Rheinboldt, W.C., *A unified convergence theory for a class of iterative processes*, SIAM J. Numer. Anal., **5**, pp. 42–63, 1968.

[26] Yamamoto, T., *A convergence theorem for Newton–like methods in Banach spaces*, Numer. Math., **51**, pp. 545–557, 1987.

[27] Zabrejko, P.P. and Nguen, D.F., *The majorant method in the theory of Newton–Kantorovich approximations and the Pták error estimates*, Numer. Funct. Anal. Optim., **9**, pp. 671–684, 1987.

[28] Zinčenko, A.I., *Some approximate methods of solving equations with non–differentiable operators*, (Ukrainian), Dopovidi Akad. Nauk Ukraïn. RSR, pp. 156–161, 1963.

Received by the editors: November 18, 2008.