Functiones et Approximatio
XXXI (2003), 7–22

CONCERNING THE ONVERGENCE OF NEWTON-LIKE METHODS UNDER WEAK HÖLDER CONTINUITY CONDITIONS

IOANNIS K. ARGYROS

Abstract: The concept of majorizing sequences is employed to provide a convergence analysis for Newton-like methods in a Banach space. We use Hölder and center-Hölder continuity assumptions on the Fréchet-derivative of the operators involved. This way we show that our convergence conditions are weaker; error bounds on the distances involved finer and the location of the solution more precise than in earlier results.
Keywords: Newton-like method, Banach space, majorizing sequence, local-semilocal convergence, Fréchet-derivative, Hölder continuity, radius of convergence.

1. Introduction

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

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

where F is a Fréchet-differentiable operator defined on a convex subset D of a Banach space X with values in a Banach space Y.

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 $\dot{x} = Q(x)$ (for some suitable operator Q), 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 (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are

2001 Mathematics Subject Classification: 65B05, 65G99, 65J15, 47H17, 49M15, CR:1.5.

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.

Newton-like methods

$$x_{n+1} = x_n - A(x_n)^{-1}F(x_n) \quad (n \geqslant 0) \quad (x_0 \in D) \tag{2}$$

has been used extensively to generate a sequence approximating x^*. For a survey on local and semilocal convergence theorems for (2) we refer the reader to [2]–[5], [6], [10], [12]. Here $A(x) \in L(X,Y)$ the space of bounded linear operators from X into Y, and is an approximation to the Fréchet-derivative $F'(x)$ of operator F. Most results have used Lipschitz-type hypotheses. Here we use Hölder and center Hölder continuity conditions on F' and A instead of just Hölder assumptions, to provide local and semilocal convergence theorems for (2). This way the upper bounds on the inverses are finer than before. As far as we know and at this level of generality the most important results by others are given in [6], [10], [12].

The advantages of such an approach when compared with the ones mentioned above are:

(1) We cover a wider range of problems than [6], [10], [12]. since we use Hölder instead of Lipschitz conditions;

(2) our results compare favorably in the special case when our conditions are reduced to Lipschitz;

(3) in the local case we provide a larger convergence radius.

2. Semilocal Convergence Analysis of Newton's Method

We need the following result on majorizing sequences.

Theorem 1. *Assume: there exist parameters* $K \geqslant 0$, $M \geqslant 0$, $L \geqslant 0$, $\ell \geqslant 0$, $\mu \geqslant 0$, $\eta \geqslant 0$, λ_1, λ_2, $\lambda_3 \in [0,1]$, $\delta \in [0,2)$ *such that:*

$$h_q = K\eta^{\lambda_1} + (1+\lambda_1)\left[M\left(\frac{\eta}{1-q}\right)^{\lambda_2} + \mu\right] + \left[\ell + L\left(\frac{\eta}{1-q}\right)^{\lambda_3}\right]\delta \leqslant \delta, \tag{3}$$

and

$$\ell + L\left(\frac{\eta}{1-q}\right)^{\lambda_3} \leqslant 1, \tag{4}$$

where

$$q = \frac{\delta}{1+\lambda_1}. \tag{5}$$

Then, iteration $\{t_n\}$ $(n \geqslant 0)$ *given by*

$$t_0 = 0, \quad t_1 = \eta,$$
$$t_{n+2} = t_{n+1} + \frac{K(t_{n+1} - t_n)^{\lambda_1} + (1 + \lambda_1)[Mt_n^{\lambda_2} + \mu]}{(1 + \lambda_1)[1 - \ell - Lt_{n+1}^{\lambda_3}]} \cdot (t_{n+1} - t_n) \quad (n \geqslant 0) \tag{6}$$

is non-decreasing, bounded above by

$$t^{**} = \frac{\eta}{1 - q}, \tag{7}$$

and converges to some t^* *such that*

$$0 \leqslant t^* \leqslant t^{**}. \tag{8}$$

Moreover, the following error bounds hold for all $n \geqslant 0$

$$0 \leqslant t_{n+2} - t_{n+1} \leqslant q(t_{n+1} - t_n) \leqslant q^{n+1}\eta. \tag{9}$$

Proof. We must show:

$$K(t_{k+1} - t_k)^{\lambda_1} + (1 + \lambda_1)[Mt_k^{\lambda_2} + \mu] + (\ell + Lt_{k+1}^{\lambda_3})\delta \leqslant \delta, \tag{10}$$
$$1 - \ell - Lt_{k+1}^{\lambda_3} > 0, \tag{11}$$

and

$$0 \leqslant t_{k+1} - t_k \quad \text{for all} \quad k \geqslant 0. \tag{12}$$

Estimate (9) can then follow immediately from (10)–(12) and (6). For $k = 0$ (10), (12) hold by (3). We also get

$$0 \leqslant t_2 - t_1 \leqslant q(t_1 - t_0). \tag{13}$$

Let us assume (10)–(12) hold for all $k \leqslant n + 1$. We can have in turn

$$\begin{aligned}
&K(t_{k+2} - t_{k+1})^{\lambda_1} + (1 + \lambda_1)[Mt_{k+1}^{\lambda_2} + \mu] + (\ell + Lt_{k+2}^{\lambda_3})\delta \\
\leqslant& K\eta^{\lambda_1} q^{k+1} + (1 + \lambda_1)\left[M\left(\frac{1 - q^{k+1}}{1 - q}\eta\right)^{\lambda_2} + \mu\right] + \left[\ell + L\left(\frac{1 - q^{k+2}}{1 - q}\eta\right)^{\lambda_3}\right]\delta \\
\leqslant& K\eta^{\lambda_1} + (1 + \lambda_1)\left[M\left(\frac{\eta}{1 - q}\right)^{\lambda_2} + \mu\right] + \left[\ell + L\left(\frac{\eta}{1 - q}\right)^{\lambda_3}\right]\delta
\end{aligned} \tag{14}$$

which is smaller or equal to δ by (3). We used:

$$t_{k+2} - t_{k+1} \leqslant q^{k+1}\eta$$

and

$$t_{k+1} \leqslant t_1 + q\eta + q^2\eta + \cdots + q^k\eta = \frac{1-q^{k+1}}{1-q}\eta.$$

Hence the first inequality in (10) holds for all $k \geqslant 0$.

We must also show:

$$t_k \leqslant t^{**} \quad (k \geqslant 0). \tag{15}$$

For $k = 0, 1, 2$ we get

$$t_0 = 0 \leqslant t^{**}, \quad t_1 = \eta \leqslant t^{**} \quad \text{and} \quad t_2 \leqslant \eta + q\eta = (1+q)\eta \leqslant t^{**}.$$

Assume (15) holds for all $k \leqslant n+1$. It follows from (9)

$$\begin{aligned} t_{k+2} &\leqslant t_{k+1} + q(t_{k+1} - t_k) \leqslant t_k + q(t_k - t_{k-1}) + q(t_{k+1} - t_k) \\ &\leqslant \cdots \leqslant t_1 + q(t_1 - t_0) + \cdots + q(t_k - t_{k-1}) + q(t_{k+1} - t_k) \\ &\leqslant \eta + q\eta + q^2\eta + \cdots + q^{k+1}\eta = \frac{1-q^{k+2}}{1-q}\eta < \frac{\eta}{1-q} = t^{**}. \end{aligned}$$

Moreover inequality (11) holds since

$$Lt_{k+1}^{\lambda_3} + \ell < L\left(\frac{\eta}{1-q}\right)^{\lambda_3} + \ell \leqslant 1 \quad \text{by (3)}.$$

Furthermore, (12) also holds by (6), (11) and (12). The induction for (11)–(13) is completed.

Iteration $\{t_n\}$ $(n \geqslant 0)$ is now non-decreasing and bounded above by t^{**} and as such it converges to some t^* satisfying (8).

That completes the proof of Theorem 1. ■

In the Lipschitz case we provide the following result on majorizing sequences which seems to be weaker than the corresponding result in Theorem 1.

Proposition 1. *Let* $\lambda_1 = \lambda_2 = \lambda_3 = 1$. *Assume: there exist parameters* $K \geqslant 0$, $M \geqslant 0$, $L \geqslant 0$, $\ell \geqslant 0$, $\mu \geqslant 0$, $\eta \geqslant 0$, $\delta \in [0,1]$ *such that:*

$$h_\delta = \left(K + L\delta + \frac{4M}{2-\delta}\right)\eta + \delta\ell + 2\mu \leqslant \delta, \tag{16}$$

$$\ell + \frac{2L\eta}{2-\delta} \leqslant 1, \tag{17}$$

$$L \leqslant K, \tag{18}$$

and

$$\ell + 2\mu < 1, \tag{19}$$

then, iteration $\{t_n\}$ $(n \geqslant 0)$ *given by (6) is non-decreasing, bounded above*

$$t^{**} = \frac{2\eta}{2-\delta}$$

and converges to some t^* such that

$$0 \leqslant t^* \leqslant t^{**}.$$

Moreover the following error bounds hold for all $n \geqslant 0$

$$0 \leqslant t_{n+2} - t_{n+1} \leqslant \frac{\delta}{2}(t_{n+1} - t_n) \leqslant \left(\frac{\delta}{2}\right)^{n+1} \eta. \tag{20}$$

Proof. As in Theorem 1 but to show (14) it suffices

$$\delta L \left[\frac{2}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+2}\right) - 1\right] \leqslant K \left[1 - \left(\frac{\delta}{2}\right)^{k+1}\right] \tag{21}$$

and

$$\frac{4M}{2-\delta}\left[1 - \left(\frac{\delta}{2}\right)^{k+1}\right] \leqslant \frac{4M}{2-\delta}. \tag{22}$$

But (22) certainly hold by the choice of δ. Instead of (21) we can show

$$\left[1 - \left(\frac{\delta}{2}\right)^{k+1}\right] \frac{\delta^2 L}{2-\delta} \leqslant K,$$

or (since $L \leqslant K$)

$$\frac{\delta^2 L}{2-\delta} \leqslant K,$$

which is also true by the choice of δ.

That completes the proof of Proposition 1. ■

Remark 1. The range for δ in Proposition 1 can be extended to $\delta \in [0, 2)$, if (16) is replaced by

$$h_\delta \leqslant \delta \quad \text{and} \quad \frac{\delta^2 L}{2-\delta} \leqslant K, \tag{16$'$}$$

as it can easily be seen from the proof. Replace (16) by (16)$'$ and call the corresponding proposition, Proposition 1$'$. Clearly the hypotheses of Proposition 1$'$ are weaker than Proposition 1.

Remark 2. Note that conditions (3), (4) (or (16)–(19)) are of the Newton–Kantorovich-type hypotheses (see also (57)), which are always present in the study of Newton-like methods [5], [9].

We can show the main semilocal convergence theorem for Newton-like methods.

Theorem 2. *Let $F\colon D \subseteq X \to Y$ be a Fréchet-differentiable operator. Assume:*

(a) *there exist an approximation $A(x) \in L(X,Y)$ of $F'(x)$, an open convex subset D_0 of D, $x_0 \in D_0$, parameters $\eta \geqslant 0$, $K \geqslant 0$, $M \geqslant 0$, $L \geqslant 0$, $\mu \geqslant 0$, $\ell \geqslant 0$, $\lambda_1 \in [0,1]$, $\lambda_2 \in [0,1]$, $\lambda_3 \in [0,1]$ such that:*

$$A(x_0)^{-1} \in L(Y,X), \quad \|A(x_0)^{-1}F(x_0)\| \leqslant \eta, \tag{23}$$

$$\|A(x_0)^{-1}[F'(x) - F'(y)]\| \leqslant K\|x-y\|^{\lambda_1}, \tag{24}$$

$$\|A(x_0)^{-1}[F'(x) - A(x)]\| \leqslant M\|x-x_0\|^{\lambda_2} + \mu, \tag{25}$$

and

$$\|A(x_0)^{-1}[A(x) - A(x_0)]\| \leqslant L\|x-x_0\|^{\lambda_3} + \ell \quad \text{for all } x, y \in D_0; \tag{26}$$

(b) *hypotheses of Theorem 1 or Proposition 1 or Proposition 1′ hold;*

(c)

$$\overline{U}(x_0, t^*) = \{x \in X \mid \|x - x_0\| \leqslant t^*\} \subseteq D_0. \tag{27}$$

Then, sequence $\{x_n\}$ $(n \geqslant 0)$ generated by Newton's method (2) is well defined, remains in $\overline{U}(x_0,t^*)$ for all $n \geqslant 0$ and converges to a solution $x^* \in \overline{U}(x_0,t^*)$ of equation $F(x) = 0$.

Moreover the following error bounds hold for all $n \geqslant 0$:

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n \tag{28}$$

and

$$\|x_n - x^*\| \leqslant t^* - t_n. \tag{29}$$

Futhermore the solution x^* is unique in $\overline{U}(x_0,t^*)$ if

$$\frac{1}{1 - \ell - L(t^*)^{\lambda_3}}\left[\frac{K}{1+\lambda_1}(t^*)^{1+\lambda_1} + M(t^*)^{\lambda_2} + \mu\right] < 1, \tag{30}$$

or in $U(x_0,R_0)$ if $R_0 > t^*$, $U(x_0,R_0) \subseteq D_0$, and

$$\frac{1}{1 - \ell - L(t^*)^{\lambda_3}}\left[\frac{K}{1+\lambda_1}(R + t^*)^{1+\lambda_1} + M(t^*)^{\lambda_2} + \mu\right] \leqslant 1. \tag{31}$$

Proof. By induction on n we must show (28). Since,

$$\|x_1 - x_0\| = \|A(x_0)^{-1}F(x_0)\| \leqslant \eta \leqslant t_1 - t_0$$

(28) holds for $n = 0$, and we have by (26) for $x \in U(x_0,t^*)$

$$\|A(x_0)^{-1}[A(x) - A(x_0)]\| \leqslant L\|x - x_0\|^{\lambda_3} + \ell \leqslant L(t^*)^{\lambda_3} + \ell < 1.$$

It follows from the Banch Lemma on invertible operators [9] $A(x_1)^{-1}$ exists and

$$\|A(x)^{-1}A(x_0)\| \leqslant (1 - L\|x-x_0\|^{\lambda_3} - \ell)^{-1} \leqslant (1 - L(t^*)^{\lambda_3} - \ell)^{-1}. \tag{32}$$

Assume that for $0 \leqslant k \leqslant n+1$

$$\|x_{k+1} - x_k\| \leqslant t_{k+1} - t_k. \tag{33}$$

Then, we get

$$\|x_{k+1} - x_0\| \leqslant t_{k+1} - t_0 = t_{k+1} \tag{34}$$

and

$$\begin{aligned} x_{k+2} - x_{k+1} &= -\, A(x_{k+1})^{-1}F(x_{k+1}) \\ &= -\, A(x_{k+1})^{-1}[F(x_{k+1}) - A(x_k)(x_{k+1} - x_k) - F(x_k)] \\ &= -\, [A(x_{k+1})^{-1}A(x_0)]A(x_0)^{-1} \\ &\quad \cdot \left\{ \int_0^1 [F'(x_{k+1} + t(x_k - x_{k+1})) - F'(x_k)](x_{k+1} - x_k)dt \right. \\ &\quad \left. + (F'(x_k) - A(x_k))(x_{k+1} - x_k) \right\}. \end{aligned} \tag{35}$$

Using (6), (24)–(26), (32), (33) and (35) we obtain in turn

$$\begin{aligned} \|x_{k+2} - x_{k+1}\| &\leqslant \|A(x_{k+1})^{-1}A(x_0)\| \\ &\quad \cdot \left\{ \int_0^1 \|A(x_0)^{-1}[F'(x_{k+1} + t(x_k - x_{k+1})) - F'(x_k)]\|dt \right. \\ &\quad \left. + \|A(x_0)^{-1}[F'(x_k) - A(x_k)]\| \right\} \|x_{k+1} - x_k\| \\ &\leqslant \frac{1}{1 - \ell - L\|x_{k+1} - x_0\|^{\lambda_3}} \\ &\quad \times \left\{ \frac{K}{1+\lambda_1}\|x_{k+1} - x_k\|^{1+\lambda_1} + (M\|x_k - x_0\|^{\lambda_2} + \mu)\|x_{k+1} - x_k\| \right\} \\ &\leqslant \frac{1}{(1 - \ell - Lt_{k+1}^{\lambda_3})(1+\lambda_1)} \{K(t_{k+1} - t_k)^{\lambda_1} \\ &\quad + (Mt_k^{\lambda_2} + \mu)(1+\lambda_1)\}(t_{k+1} - t_k) = t_{k+2} - t_{k+1}, \end{aligned} \tag{36}$$

which completes the induction. Hence we have:

$$\begin{aligned} \|x_{k+2} - x_{k+1}\| &\leqslant t_{k+2} - t_{k+1}, \\ \|A(x_0)^{-1}[A(x_{k+1}) - A(x_0)]\| &\leqslant L\|x_{k+1} - x_0\|^{\lambda_3} + \ell \\ &\leqslant Lt_{k+1}^{\lambda_3} + \ell \leqslant L(t^*)^{\lambda_3} + \ell < 1, \end{aligned}$$

and

$$\|x_{k+2} - x_0\| \leqslant t_{k+2} \leqslant t^*.$$

It follows $x_k \in U(x_0, t^*)$ for all $k \geqslant 0$. Iteration $\{x_n\}$ $(n \geqslant 0)$ is Cauchy by (28) in a Banach space X and as such it converges to some $x^* \in \overline{U}(x_0, t^*)$ (since

$\overline{U}(x_0, t^*)$ is a closed set). By letting $k \to \infty$ in (36) we obtain $F(x^*) = 0$. Estimate (29) follows from (28) by using standard majorization techniques [9].

Finally to show uniqueness let $y^* \in U(x_0, t^*)$ be a solution of $F(x) = 0$. Using the approximation

$$y^* - x_{k+1} = y^* - x_{k+1} + A(x_{k+1})^{-1}F(x_{k+1}) - A(x_{k+1})^{-1}F(y^*), \tag{37}$$

(24)–(26) we get as in (36) for x_{k+2} replaced by y^*

$$\begin{aligned}
\|y^* - x_{k+1}\| &\leqslant \|A(x_k)^{-1}A(x_0)\| \\
&\quad \cdot \left\{ \int_0^1 \|A(x_0)^{-1}[F'(x_k + t(y^* - x_k)) - F'(x_k)\| dt \right. \\
&\quad \left. + \|A(x_0)^{-1}[F'(x_k) - A(x_k)\| \right\} \|y^* - x_k\| \\
&\leqslant \frac{1}{1 - \ell - Lt_k^{\lambda_3}} \left\{ \frac{K}{1+\lambda_1} \|y^* - x_k\|^{1+\lambda_1} + (M\|x_k - x_0\|^{\lambda_2} + \mu)\|y^* - x_k\| \right\} \\
&\leqslant \frac{1}{1 - L(t^*)^{\lambda_3} - \ell} \left\{ \frac{K}{1+\lambda_1}(t^*)^{1+\lambda_1} + (M(t^*)^{\lambda_2} + \mu) \right\} \|y^* - x_k\|.
\end{aligned} \tag{38}$$

By (30) and (38)

$$\|y^* - x_{k+1}\| < \|y^* - x_k\|. \tag{39}$$

By letting $k \to \infty$ in (39) we obtain $\lim_{k\to\infty} x_k = y^*$. But we already showed $\lim_{k\to\infty} x_k = x^*$. Hence, we deduce $x^* = y^*$. The second case of uniqueness uses (31) instead of (30).

That completes the proof of Theorem 2. ∎

In order for us to compare Theorem 2 with a relevant one already in the literature [12] we state:

Theorem 3. *Assume:*

(a) *hypotheses (a) of Theorem 2 hold for* $\lambda_1 = \lambda_2 = \lambda_3 = 1$;

(b)

$$\ell + \mu < 1, \quad \sigma = \max\left\{1, \frac{L+M}{K}\right\}, \quad K \neq 0, \tag{40}$$

$$h = \frac{2\sigma K\eta}{(1 - \ell - \mu)^2} \leqslant 1 \tag{41}$$

and set

$$s^* = \frac{(1 - \ell - \mu)(1 - \sqrt{1-h})}{\sigma K}, \tag{42}$$

$$s^{**} = \frac{1 - \mu + \sqrt{(1-\mu)^2 - 2K\eta}}{K}; \tag{43}$$

(c)

$$\overline{U}(x_1, s^* - \eta) \subseteq D_0. \tag{44}$$

Then, sequence $\{x_n\}$ $(n \geqslant 0)$ generated by Newton's method (2) is well defined, remains in $U(x_1, x^* - \eta)$ for all $n \geqslant 1$ and converges to a solution $x^* \in \overline{U}(x_1, s^* - \eta)$ of equation $F(x) = 0$.

The solution x^* is unique in

$$\widetilde{U} = \begin{cases} U(x_0, s^{**}) \cap D_0 & \text{if } 2K\eta < (1-\mu)^2 \\ \overline{U}(x_0, s^{**}) \cap D_0 & \text{if } 2K\eta = (1-\mu)^2 . \end{cases} \tag{45}$$

Moreover the following error bounds hold for all $n \geqslant 0$

$$\|x_{n+1} - x_n\| \leqslant s_{n+1} - s_n \tag{46}$$

$$\|x_n - x^*\| \leqslant s^* - s_n, \tag{47}$$

where,

$$s_0 = 0, \quad s_{n+1} = s_n +. \frac{f(s_n)}{g(s_n)} \quad (n \geqslant 0), \tag{48}$$

$$f(s) = \frac{1}{2}\sigma K s^2 - (1 - \ell - \mu)s + \eta \tag{49}$$

and

$$g(s) = 1 - \ell - Ls. \tag{50}$$

We now show that the error bounds obtained in Theorem 2 are more precise than the corresponding ones in Theorem 3.

Theorem 4. *Under the hypotheses of Theorem 2 (for $\lambda_1 = \lambda_2 = \lambda_3 = 1$), 3 the following error bounds hold:*

$$t_{n+1} \leqslant s_{n+1} \quad (n \geqslant 1) \tag{51}$$

$$t_{n+1} - t_n \leqslant s_{n+1} - s_n \quad (n \geqslant 1) \tag{52}$$

$$t^* - t_n \leqslant r^* - s_n \quad (n \geqslant 0) \tag{53}$$

and

$$t^* \leqslant s^*. \tag{54}$$

Moreover strict inequality holds in (51) and (52) if $K < M + L$.

Proof. We use mathematical induction on the integer i to first show (46) and (47). For $n = 0$ in (6) and (44) we obtain

$$\begin{aligned} t_2 - \eta &= \frac{\frac{K}{2}\eta^2 + (M0 + \mu)\eta}{1 - \ell - L\eta} \leqslant \frac{\frac{\sigma}{2}n^2 + (M \cdot 0 + \mu)\eta}{1 - \ell - L\eta} \\ &\leqslant \frac{\frac{\sigma}{2}\eta^2 + M(\eta - 0)0 + \mu(\eta - 0) - g(0)(\eta - 0) + f(0)}{g(\eta)} \\ &\leqslant \frac{\frac{\sigma}{2}s_1^2 - (1 - \mu - \ell)s_1 + \eta - (\sigma - M - L)s_0(s_1 - s_0)}{g(s_1)} \\ &\leqslant \frac{f(s_1)}{g(s_2)} = s_2 - s_1, \end{aligned}$$

and

$$t_2 \leqslant s_2.$$

Assume:

$$t_{i+1} \leqslant s_{i+1}, \quad t_{i+1} - t_i \leqslant s_{i+1} - s_i. \tag{55}$$

Using (6), (46) and (54) we obtain in turn

$$\begin{aligned}
&t_{i+2} - t_{i+1} \\
&= \frac{\frac{K}{2}(t_{i+1} - t_i)^2 + (Mt_i + \mu)(t_{i+1} - t_i)}{1 - \ell - Lt_{i+1}} \\
&\leqslant \frac{\frac{\sigma}{2}(s_{i+1} - s_i)^2 + (Ms_i + \mu)(s_{i+1} - s_i)}{g(s_{i+1})} \\
&= \frac{\frac{\sigma}{2}(s_{i+1} - s_i)^2 + M(s_{i+1} - s_i)s_i + \mu(s_{i+1} - s_i) - g(s_i)(s_{i+1} - s_i) + f(s_i)}{g(s_{i+1})} \\
&= \frac{\frac{\sigma}{2}s_{i+1}^2 - (1 - \mu - \ell)s_{i+1} + \eta - (\sigma - M - L)s_i(s_{i+1} - s_i)}{g(s_{i+1})} \\
&\leqslant \frac{f(s_{i+1})}{g(s_{i+1})} = s_{i+2} - s_{i+2},
\end{aligned}$$

which show (47) and (48) for all $(n \geqslant 1)$.

Let $j \geqslant 0$ we can get

$$\begin{aligned}
t_{i+j} - t_i &\leqslant (t_{i+j} - t_{i+j-1}) + (t_{i+j-1} - t_{i+j-2}) + \cdots + (t_{i+1} - t_i) \\
&\leqslant (s_{i+j} - s_{i+j-1}) + (s_{i+j-1} - s_{i+j-2}) + \cdots + (s_{i+1} - s_i) \\
&\leqslant s_{i+j} - s_i.
\end{aligned} \tag{56}$$

By letting $j \to \infty$ in (55) we obtain (52).

Finally (52) implies (53) (since $t_1 = s_1 = 0$). It can easily be seen from (6) and (47) that strict inequality holds in (51) and (52) if $K < M + L$.

That completes the proof of Theorem 4. ■

Remark 3. Due to (50) our Theorem 2 provides at least as precise information on the location of the solution x^* than Theorem 3. Note also that $t^* \in [\eta, t^{**}]$, and under the hypotheses of Theorems 2 and 3 $t^* \in [\eta, s^*]$ where s^* is given by (42).

Remark 4. Let us compare condition (3) with condition (41). For simplicity let us consider only Newton's method. We set $A(x) = F'(x)$ and choose $M = \mu = \ell = 0$, $K = L$, $\lambda_1 = \lambda_2 = \lambda_3 = 1$ to obtain:

$$h = 2K\eta \leqslant 1, \tag{57}$$

which is the famous Newton–Kantorovich hypothesis for the convergence of Newton's method [9]. However in general $L \leqslant K$. Hence, our condition becomes

$$h_1 = (K + L)\eta \leqslant 1. \tag{58}$$

Note that

$$h \leqslant 1 \Longrightarrow h_1 \leqslant 1 \tag{59}$$

but not vice versa (unless if $K = L$).

Condition (57) is widely used in the literature and for a long time whereas (58) should be used instead. In the following example we show $\frac{K}{L}$ can be arbitrarily large.

In the three examples that follow we use the choices of operators and parameters given in Remark 4.

Example 1. Let $X = Y = \mathbf{R}$, $x_0 = 0$ and define function F on $\mathbf{R}$ by

$$F(x) = c_0 x + c_1 + c_2 \sin e^{c_3 x}, \tag{60}$$

where c_i, $i = 0, 1, 2, 3$ are given parameters. Using (60) it can easily be seen that for c_3 large and c_2 sufficiently small, $\frac{K}{L}$ may be arbitrarily large. That is (58) may be satisfied but not (57). In the next example we show that (58) holds where (57) is violated.

Example 2. Let $X = Y = \mathbf{R}$, $D = [\sqrt{2}-1, \sqrt{2}+1]$, $x_0 = \sqrt{2}$ and define function F on D by

$$F(x) = \frac{1}{6}x^3 - \left(\frac{2^{3/2}}{6} + .23\right). \tag{61}$$

Using (61) we obtain

$$\eta = .23, \quad K = 2.4142136, \quad L = 1.914213562,$$
$$h = 1.1105383 > 1 \quad \text{and} \quad h_1 = .995538247 < 1.$$

That is there is no guarantee that Newton's method $\{x_n\}$ $(n \geqslant 0)$ starting at x_0 converges to a solution x^* of equation $F(x) = 0$, since (57) is violated. However since (58) holds our Theorem 2 shows convergence of Newton's method to $x^* = 1.614507018$ since (58) holds.

Example 3. Let $X = Y = \mathbf{R}$, $D = [a, 2-a]$, $a \in \left[0, \frac{1}{2}\right)$, $x_0 = 1$ and define function F on D by

$$F(x) = x^3 - a.$$

Using (23), (24) and (26) we obtain

$$\eta = \frac{1}{3}(1-a), \quad K = 2(2-a), \quad L = 3 - a.$$

Then (57) becomes

$$h = \frac{4}{3}(1-a)(2-a) > 1 \quad \text{for all } a \in \left[0, \frac{1}{2}\right).$$

However (16) for $\delta = 1$ gives

$$h_1 = \frac{1}{3}(1-a)[(3-a)+2(2-a)] \leqslant 1$$

for all

$$a \in \left[\frac{5-\sqrt{13}}{3}, \frac{1}{2}\right).$$

Remark 5. The results obtained here can be used to solve equations of the form

$$A^{\#}(x_0)F(x_0) = 0. \tag{62}$$

Define

$$x_{n+1} = x_n - A(x_n)^{\#}F(x_n) \quad (n \geqslant 0) \quad (x_0 \in D), \tag{63}$$

where, $A(x_n)^{\#}$ denotes an outer inverse of $A(x_n)$, i.e.,

$$A(x_n)^{\#}A(x_n)A(x_n)^{\#} = A(x_n)^{\#} \quad (n \geqslant 0). \tag{64}$$

As in [10] using the Banach-type Lemma on outer inverses [10, p. 240] instead of the Banch Lemma on invertible operators and setting

$$A(x_n)^{\#} = [I + A^{\#}(x_0)(A(x_n) - A(x_0))]^{-1}A^{\#}(x_0) \tag{65}$$

exactly as in [10] and Theorem 2 we show the conclusions of Theorem 2 hold. Note that $A(x_0)^{-1}$ must be replaced by $A^{\#}(x_0)$ in (23)–(26).

3. Local Analysis of Newton-like Methods

We can show the following local result for Newton-like methods:

Theorem 5. *Let $F: D \subseteq X \to Y$ be a Fréchet-differentiable operator. Assume:*

(a) *there exist an approximation $A(x) \in L(X,Y)$ of $F'(x)$, a simple solution $x^* \in D$ of equation $F(x) = 0$, a bounded inverse $A(x^*)$ and parameters $\overline{K}$, $\overline{L}$, $\overline{M}$, $\overline{\mu}$, $\overline{\ell} \geqslant 0$, λ_4, λ_5, $\lambda_6 \in [0,1]$ such that:*

$$\|A(x^*)^{-1}[F'(x) - F'(y)]\| \leqslant \overline{K}\|x-y\|^{\lambda_4}, \tag{66}$$

$$\|A(x^*)^{-1}[F'(x) - A(x)]\| \leqslant \overline{M}\|x-x^*\|^{\lambda_5} + \overline{\mu}, \tag{67}$$

and

$$\|A(x^*)^{-1}[A(x) - A(x^*)]\| \leqslant \overline{L}\|x-x^*\|^{\lambda_6} + \overline{\ell} \tag{68}$$

for all $x, y \in D$;

(b) *equation*

$$\frac{\overline{K}}{1+\lambda_4} r^{\lambda_4} + \overline{L} r^{\lambda_6} + \overline{M} r^{\lambda_5} + \overline{\mu} + \overline{\ell} - 1 = 0 \tag{69}$$

has a minimal positive zero r_0 *which also satisfies:*

$$\overline{L} r_0^{\lambda_6} + \overline{\ell} < 1 \tag{70}$$

and

$$U(x^*, r_0) \subseteq D. \tag{71}$$

Then, sequence $\{x_n\}$ $(n \geqslant 0)$ *generated by Newton's method (2) is well defined, remains in* $U(x^*, r_0)$ *for all* $n \geqslant 0$, *and converges to* x^* *provided that* $x_0 \in U(x^*, r_0)$. *Moreover the following error bounds hold for all* $n \geqslant 0$:

$$\begin{aligned} &\|x_{n+1} - x^*\| \\ &\leqslant \frac{1}{1 - \overline{L}\|x_n - x^*\|^{\lambda_6} - \overline{\ell}} \left[\frac{\overline{K}}{1+\lambda_4} \|x_n - x^*\|^{\lambda_4} + \overline{M}\|x_n - x^*\|^{\lambda_5} + \overline{\mu} \right] \|x_n - x^*\|. \end{aligned} \tag{72}$$

Proof. Let $x \in U(x^*, r_0)$. Using (67) and (69) we get

$$\|A(x^*)^{-1}[A(x) - A(x^*)]\| \leqslant \overline{L}\|x - x^*\|^{\lambda_6} + \overline{\ell} \leqslant \overline{L} r_0^{\lambda_6} + \overline{\ell} < 1. \tag{73}$$

It follows from (73) and the Banach Lemma on invertible operators $A(x)^{-1}$ exists and

$$\begin{aligned} \|A(x)^{-1} A(x^*)\| &\leqslant (1 - \overline{\ell} - \overline{L}\|x - x^*\|^{\lambda_6})^{-1} \\ &\leqslant (1 - \overline{\ell} - \overline{L} r_0^{\lambda_6})^{-1}. \end{aligned} \tag{74}$$

By hypothesis $x_0 \in U(x^*, r_0)$. Assume $x_k \in U(x^*, r_0)$. Using (2), (65)–(67) we get in turn

$$\begin{aligned} &\|x_{k+1} - x^*\| \\ &= \|A(x_k)^{-1}[A(x_k)(x^* - x_k) + F(x_k) - F(x^*)]\| \\ &\leqslant \|A(x_k)^{-1} A(x^*)\| \left\{ \left\| A(x^*)^{-1} \int_0^1 [F'(x_k + t(x^* - x_k)) - F'(x_k)] \right\| dt \right. \\ &\quad \left. + \|A(x^*)^{-1}[F'(x_k) - A(x_k)]\| \right\} \|x^* - x_k\| \\ &\leqslant \frac{1}{1 - \overline{L}\|x_k - x^*\|^{\lambda_6} - \overline{\ell}} \left[\frac{\overline{K}}{1+\lambda_4} \|x_k - x^*\|^{\lambda_4} + \overline{M}\|x^* - x_k\|^{\lambda_5} + \overline{\mu} \right] \|x_k - x^*\| \\ &< \frac{1}{1 - \overline{\ell} - \overline{L} r_0^{\lambda_6}} \left[\frac{\overline{K}}{1+\lambda_4} r_0^{\lambda_4} + \overline{M} r_0^{\lambda_5} + \overline{\mu} \right] \|x_k - x^*\|, \end{aligned} \tag{75}$$

which shows $x_k \in U(x^*, r_0)$ $(n \geqslant 0)$. Moreover by the choice of r_0 and (75)

$$\|x_{k+1} - x^*\| < \|x_k - x^*\| < r_0. \tag{76}$$

Hence, we deduce $\lim_{k\to\infty} x_k = x^*$.

That completes the proof of Theorem 5. ∎

Remark 6. As noted in [1]–[7], [8], [10] and [13] the local results obtained here can be used for projection methods such as Arnoldi's, the generalized minimum residual method (GMRES), the generalized conjugate residual method (GCR), for combined Newton/finite-difference projection methods and in connection with the mesh independence principle in order to develop the cheapest mesh refinement strategies.

From now on for simplicity we refer only to Newton's method. That is we set $A(x) = F'(x)$, $(x \in D)$.

Remark 7. The local results obtained here can also be used to solve equations of the form $F(x) = 0$, where F' satisfies the autonomous differential equation [5], [13]:

$$F'(x) = T(F(x)), \tag{77}$$

where, $T: Y \to X$ is a known continuous operator. Since $F'(x^*) = T(F(x^*)) = T(0)$, we can apply the results obtained here without actually knowing the solution x^* of equation (1).

We complete this section with a numerical example.

Example 4. Let $X = Y = \mathbf{R}$, $D = U(0,1)$ and define function F on D by

$$F(x) = e^x - 1. \tag{78}$$

Then it can easily be seen that we can set $T(x) = x + 1$ in (77). Since $F'(x^*) = 1$, we get $\|F'(x) - F'(y)\| \leqslant e\|x - y\|$. Hence we set $\overline{K} = e$. Moreover since $x^* = 0$ we obtain in turn

$$\begin{aligned} F'(x) - F'(x^*) &= \mathrm{e}^x - 1 = x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots \\ &= \left(1 + \frac{x}{2!} + \cdots + \frac{x^{n-1}}{n!} + \cdots\right)(x - x^*) \end{aligned}$$

and for $x \in U(0,1)$,

$$\|F'(x) - F'(x^*)\| \leqslant (\mathrm{e} - 1)\|x - x^*\|.$$

That is, $\overline{L} = e - 1$. We obtain for $\lambda_1 = \lambda_2 = \lambda_3 = 1$,

$$r_0 = .254028662$$

Rheinboldt in [11] used

$$r_1 = \frac{2}{3\overline{K}}.$$

In this case we get

$$r_1 = .245252961.$$

That is our convergence radius r_0 is larger than the corresponding one r_1 due to Rheinboldt [11]. This observation is very important in computational mathematics, since it allows a wider choice of initial guesses x_0 (see also Remark 6).

References

[1] Allgower, F.L., Böhmer, K., Potra, F.A., and Rheinboldt, W.C., A mesh independence principle for operator equations and their discretizations, *SIAM J. Numer. Anal.* **23**, 1 (1986), 160–169.

[2] Argyros, I.K., Convergence domains for some iterative processes in Banach spaces using outer and generalized inverses, *J. Comput. Anal. and Applic.* **1**, 1 (1999), 87–104.

[3] Argyros, I.K., Local convergence of inexact Newton-like iterative methods and applications, *Computers and Mathematics with Application* **39** (2000), 69–75.

[4] Argyros, I.K., *Advances in the Efficiency of Computational Methods and Applications*, World Scientific Publ. Co., River Edge, NJ, 2000.

[5] Argyros, I.K. and Szidarovszky, F., *The Theory and Applications of Iteration Methods*, C.R.C. Press, Boca Raton, Florida, 1993.

[6] Dennis, J.E., Jr., On Newton-like methods, *Numer. Math.* **11** (1968), 324–330.

[7] Gutiérrez, J.M., A new semilocal convergence theorem for Newton's method, *J. Comp. Appl. Math.* **79** (1997), 131–145.

[8] Gutiérrez, J.M., Hernandez, M.A., and Salanova, M.A., Accessibility of solutions by Newton's method, *Internat. J. Comput. Math.* **57** (1995), 239–247.

[9] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1964.

[10] Nashed, M.Z. and Chen, X., Convergence of Newton-like methods for singular operator equations using outer inverses, *Numer. Math.* **66** (1993), 235–257.

[11] Rheinboldt, W.C., An adaptive continuation process for solving systems of nonlinear equations, *Polish Academy of Sciences, Banach Ctr. Publ.* **3** (1977), 129–142.

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

[13] Ypma, T.J., Local convergence of inexact Newton's method, *SIAM J. Numer. Anal.* **21**, 3 (1984), 583–590.

Address: Cameron University, Department of Mathematical Sciences, Lawton, OK 73505, U.S.A.

E-mail: ioannisa@cameron.edu

Received: 10 September 2002