APPLICATIONES MATHEMATICAE
40,4 (2013), pp. 459–481

IOANNIS K. ARGYROS (Lawton, OK)
SANJAY K. KHATTRI (Stord)

AN IMPROVED CONVERGENCE ANALYSIS OF NEWTON'S METHOD FOR TWICE FRÉCHET DIFFERENTIABLE OPERATORS

Abstract. We develop local and semilocal convergence results for Newton's method in order to solve nonlinear equations in a Banach space setting. The results compare favorably to earlier ones utilizing Lipschitz conditions on the second Fréchet derivative of the operators involved. Numerical examples where our new convergence conditions are satisfied but earlier convergence conditions are not satisfied are also reported.

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

$$\mathcal{F}(x) = 0, \tag{1.1}$$

where $\mathcal{F}$ is a twice Fréchet differentiable operator defined on a convex subset $\mathbf{D}$ of a Banach space $\mathbf{X}$ with values in a Banach space $\mathbf{Y}$. Numerous problems in science and engineering—such as optimization of chemical processes or multiphase, multicomponent flow—can be reduced to solving the above equation [7–9, 13–15]. For most problems, finding a closed form solution for the nonlinear equation (1.1) is not possible. Therefore, iterative solution techniques are employed. The study of convergence of iterative methods is usually divided into two categories: semilocal and local convergence analysis. The semilocal convergence analysis is based upon information around the initial point to give criteria ensuring the convergence of the iterative procedure, while the local convergence analysis is based on information around a solution to find estimates of the radii of convergence balls.

2010 *Mathematics Subject Classification*: Primary 65-XX, 47Hxx, 49Mxx; Secondary 65J20, 65B05, 65G99, 65J15, 65N30, 65N35, 65H10, 47H17, 49M15.
Key words and phrases: Newton's method, Banach space, local convergence, semilocal convergence, Fréchet derivative, majorizing sequence.

DOI: 10.4064/am40-4-6

The most popular iterative method for solving problem (1.1) is Newton's method

$$x_{n+1} = x_n - \mathcal{F}'(x_n)^{-1}\mathcal{F}(x_n) \quad \text{for } n = 0, 1, 2, \dots, \tag{1.2}$$

where $x_0 \in \mathbf{D}$ is an initial point. There exist extensive local as well as semilocal convergence analysis results under various Lipschitz type conditions for Newton's method (1.2) [1–17]. The following four conditions have been used to perform semilocal convergence analysis of Newton's method (1.2) [3, 5, 7–9, 13]:

$\mathbf{C}_1$. there exists $x_0 \in \mathbf{D}$ such that $\mathcal{F}'(x_0)^{-1} \in \mathbf{L}(\mathbf{Y}, \mathbf{X})$,
$\mathbf{C}_2$. $\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_0)\| \leq \eta$,
$\mathbf{C}_3$. $\|\mathcal{F}'(x_0)^{-1}\mathcal{F}''(x)\| \leq \mathcal{K}$ for each $x \in \mathbf{D}$,
$\mathbf{C}_4$. $\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}''(x) - \mathcal{F}''(y))\| \leq \mathcal{M}\|x - y\|$ for each $x, y \in \mathbf{D}$.

Let us also introduce the center-Lipschitz condition

$\mathbf{C}_5$. $\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x) - \mathcal{F}'(x_0))\| \leq \mathcal{L}_0\|x - x_0\|$ for each $x \in \mathbf{D}$.

We shall refer to ($\mathbf{C}_1$)–($\mathbf{C}_5$) as the ($\mathbf{C}$) *conditions*. The following conditions have also been employed [9–13, 16]:

$\mathbf{C}_6$. $\|\mathcal{F}'(x_0)^{-1}\mathcal{F}''(x_0)\| \leq \mathcal{K}_0$,
$\mathbf{C}_7$. $\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}''(x) - \mathcal{F}''(x_0))\| \leq \mathcal{M}_0\|x - x_0\|$ for each $x \in \mathbf{D}$.

Henceforth, the conditions ($\mathbf{C}_1$), ($\mathbf{C}_2$), ($\mathbf{C}_5$), ($\mathbf{C}_6$), ($\mathbf{C}_7$) are referred as the ($\mathbf{H}$) *conditions.*

For the semilocal convergence of Newton's method the conditions ($\mathbf{C}_1$), ($\mathbf{C}_2$), ($\mathbf{C}_3$) together with the following sufficient conditions are given [1–4, 9–17]:

$$\eta \leq \frac{4\mathcal{M} + \mathcal{K}^2 - \mathcal{K}\sqrt{\mathcal{K}^2 + 2\mathcal{M}}}{3\mathcal{M}(\mathcal{K} + \sqrt{\mathcal{K}^2 + 2\mathcal{M}})}, \tag{1.3}$$

$$\overline{U}(x_0, R_1) \subseteq \mathbf{D}, \tag{1.4}$$

where R_1 is the smallest positive root of

$$\mathscr{P}_1(t) = \frac{\mathcal{M}}{6}t^3 + \frac{\mathcal{K}}{2}t^2 - t + \eta. \tag{1.5}$$

The conditions ($\mathbf{C}_1$), ($\mathbf{C}_2$), ($\mathbf{C}_6$), ($\mathbf{C}_7$) together with

$$\eta \leq \frac{4\mathcal{M}_0 + \mathcal{K}_0^2 - \mathcal{K}_0\sqrt{\mathcal{K}_0^2 + 2\mathcal{M}_0}}{3\mathcal{M}_0(\mathcal{K}_0 + \sqrt{\mathcal{K}_0^2 + 2\mathcal{M}_0})}, \tag{1.6}$$

$$\overline{U}(x_0, R_2) \subseteq \mathbf{D}, \tag{1.7}$$

where R_2 is the smallest positive root of

$$\mathscr{P}_2(t) = \frac{\mathcal{M}_0}{6}t^3 + \frac{\mathcal{K}_0}{2}t^2 - t + \eta, \tag{1.8}$$

have also been used for the semilocal convergence of Newton's method. Conditions (1.3) and (1.6) cannot be directly compared with ours given in Sections 2 and 3, since we use $\mathcal{L}_0$ that does not appear in (1.3) and (1.6). However, comparisons can be made on concrete numerical examples. Let us consider $\mathbf{X} = \mathbf{Y} = \mathbb{R}$, $x_0 = 1$ and $\mathbf{D} = [\zeta, 2-\zeta]$ for $\zeta \in (0,1)$. Define a function $\mathcal{F}$ on $\mathbf{D}$ by

$$\mathcal{F}(x) = x^5 - \zeta. \tag{1.9}$$

Then, through some simple calculations, the conditions $(\mathbf{C}_2)$, $(\mathbf{C}_3)$, $(\mathbf{C}_4)$, $(\mathbf{C}_5)$, $(\mathbf{C}_6)$ and $(\mathbf{C}_7)$ yield

$$\eta = \frac{1-\zeta}{5}, \quad \mathcal{K} = 4(2-\zeta)^3, \quad \mathcal{M} = 12(2-\zeta)^2, \quad \mathcal{K}_0 = 4,$$
$$\mathcal{M}_0 = 4\zeta^2 - 20\zeta + 28, \quad \mathcal{L}_0 = 15 - 17\zeta + 7\zeta^2 - \zeta^3.$$

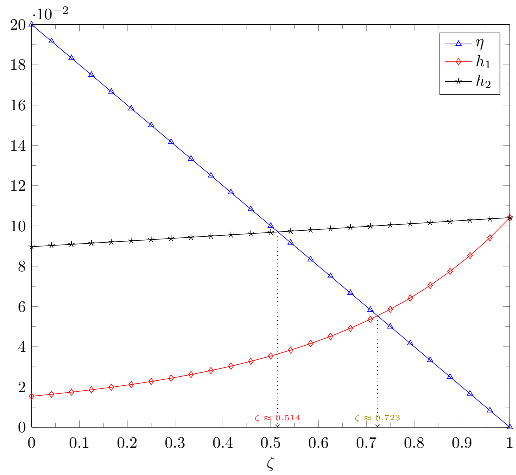

Fig. 1. Convergence criteria (1.3) and (1.6) for the equation (1.9). Here, h_1 and h_2 stand respectively for the right hand side of (1.3) and (1.6).

Figure 1 plots the criteria (1.3) and (1.6) for the problem (1.9). In Figure 1, h_1 stands for the right hand side of (1.3) and h_2 stands for the the right hand side of (1.6). In Figure 1, we observe that for $\zeta < 0.723$ the criterion (1.3) does not hold while for $\zeta < 0.514$ the criterion (1.6) does not hold. However, one can see that the method (1.2) is convergent.

In this work, we expand the applicability of Newton's method (1.2) first under the $(\mathbf{C})$ conditions and secondly under the $(\mathbf{H})$ conditions. The local convergence analysis of Newton's method (1.2) is also performed under similar conditions.

The paper is organized as follows. In Sections 2 and 3, we study majorizing sequences for the Newton's iterate $\{x_n\}$. Section 4 is devoted to the semilocal convergence analysis of Newton's method. The local convergence is analyzed in Section 5. Finally, numerical examples are given in Section 6.

2. Majorizing sequences—I. In this section, we give scalar sequences that are majorizing for Newton's method (1.2). We need the following convergence results for majorizing sequences under the (**C**) conditions.

LEMMA 2.1. *Let* $\mathcal{K}, \mathcal{L}_0, \mathcal{M}, \eta > 0$. *Define*

$$\alpha = \frac{2\mathcal{K}}{\mathcal{K} + \sqrt{\mathcal{K}^2 + 8\mathcal{L}_0\mathcal{K}}},$$

$$\eta_0 = \frac{2}{\mathcal{K}/2 + (1+\alpha)\mathcal{L}_0 + \sqrt{(\mathcal{K}/2 + (1+\alpha)\mathcal{L}_0)^2 + 2\mathcal{M}\alpha/3}},$$

$$\eta_1 = \frac{2\alpha}{\mathcal{K}/2 + \alpha\mathcal{L}_0 + \sqrt{(\mathcal{K}/2 + \alpha\mathcal{L}_0)^2 + 2\mathcal{M}\alpha/3}}. \tag{2.1}$$

Suppose that

$$\eta \leq \begin{cases} \eta_1 & \text{if } \mathcal{L}_0\eta \leq \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2}, \\ \eta_0 & \text{if } \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2} \leq \mathcal{L}_0\eta. \end{cases} \tag{2.2}$$

Then the sequence $\{t_n\}$ *generated by*

$$t_0 = 0, \quad t_1 = \eta, \quad t_{n+2} = t_{n+1} + \frac{\mathcal{K} + \frac{\mathcal{M}}{3}(t_{n+1} - t_n)}{2(1 - \mathcal{L}_0 t_{n+1})}(t_{n+1} - t_n)^2 \tag{2.3}$$

is well defined, increasing, bounded from above by

$$t^{\star\star} = \frac{\eta}{1-\alpha} \tag{2.4}$$

and converges to its unique least upper bound $t^\star$ *which satisfies* $t^\star \in [\eta, t^{\star\star}]$. *Moreover*

$$t_{n+1} - t_n \leq \alpha^n \eta, \tag{2.5}$$

$$t^\star - t_n \leq \frac{\alpha^n \eta}{1-\alpha}. \tag{2.6}$$

Proof. We use induction to prove (2.5). Set

$$\alpha_k = \frac{\mathcal{K} + \frac{\mathcal{M}}{3}(t_{k+1} - t_k)}{2(1 - \mathcal{L}_0 t_{k+1})}. \tag{2.7}$$

According to (2.3) and (2.7), we must prove that

$$\alpha_k \leq \alpha. \tag{2.8}$$

Estimate (2.8) holds for $k=0$ by (2.2) and the choice of η_1 given in (2.1). We also have $t_2-t_1\le\alpha(t_1-t_0)$ and so

$$t_2\le t_1+\alpha(t_1-t_0)=\eta+\alpha\eta=(1+\alpha)\eta=\frac{1-\alpha^2}{1-\alpha}\eta<\frac{\eta}{1-\alpha}=t^{\star\star}.$$

Let us assume that (2.7) holds for all $k\le n$. Then, by (2.3),

$$t_{k+1}-t_k\le\alpha^k\eta,\qquad t_{k+1}\le\frac{1-\alpha^{k+1}}{1-\alpha}\eta<t^{\star\star}.$$

Therefore, we must prove that

$$\left(\frac{\mathcal{K}}{2}+\frac{\mathcal{M}}{6}\alpha^k\eta\right)\alpha^k\eta+\alpha\mathcal{L}_0\frac{1-\alpha^{k+1}}{1-\alpha}\eta-\alpha\le 0. \tag{2.9}$$

Estimate (2.9) motivates us to define recurrent functions f_k on $[0,1)$ for each $k=1,2,\dots$ by

$$f_k(t)=\frac{1}{2}\left(\mathcal{K}+\frac{\mathcal{M}}{3}t^k\eta\right)t^{k-1}\eta+\mathcal{L}_0(1+t+\cdots+t^k)\eta-1. \tag{2.10}$$

We need a relationship between two consecutive functions f_k. Using (2.10) we get

$$f_{k+1}(t)=f_k(t)+g_k(t), \tag{2.11}$$

where

$$\begin{aligned} g_k(t)&=\left[\frac{1}{2}\left(\mathcal{K}+\frac{\mathcal{M}}{3}t^{k+1}\eta\right)t-\frac{1}{2}\left(\mathcal{K}+\frac{\mathcal{M}}{3}t^k\eta\right)+\mathcal{L}_0t^2\right]t^{k-1}\eta\\ &=\left[\frac{1}{2}(2\mathcal{L}_0t^2+\mathcal{K}t-\mathcal{K})+\frac{\mathcal{M}}{6}t^k\eta(t^2-1)\right]t^{k-1}\eta. \end{aligned} \tag{2.12}$$

In particular,

$$g_k(\alpha)\le 0, \tag{2.13}$$

since $\alpha\in(0,1)$ and

$$2\mathcal{L}_0\alpha^2+\mathcal{L}\alpha-\mathcal{K}=0$$

by the choice of α. Evidently (2.9) holds if

$$f_k(\alpha)\le 0\quad\text{for each } k=1,2,\dots. \tag{2.14}$$

But in view of (2.11)–(2.13) we have

$$f_k(\alpha)\le f_{k-1}(\alpha)\le\cdots\le f_1(\alpha).$$

Hence, (2.14) holds if $f_1(\alpha)\le 0$, which is true by the choice of η_0. The induction for (2.5) is complete. Hence, $\{t_n\}$ is increasing, bounded from above by $t^{\star\star}$ and as such it converges to some $t^\star$. Estimate (2.6) follows from (2.5) by standard majorization techniques [7, 8, 13–15, 17]. ∎

Let us denote by γ_0 and γ_1, respectively, the minimal positive zeros of the following equations with respect to η:

$$\begin{aligned}&\left[\frac{\mathcal{K}}{2}+\frac{\mathcal{M}}{6}\alpha(t_2-t_1)\right](t_2-t_1)+\mathcal{L}_0(1+\alpha)(t_2-t_1)+\mathcal{L}_0t_1-1=0,\\ &\left[\frac{\mathcal{K}}{2}+\frac{\mathcal{M}}{6}(t_2-t_1)\right](t_2-t_1)+\alpha\mathcal{L}_0t_2-\alpha=0.\end{aligned}\tag{2.15}$$

Let us set

$$\gamma=\min\{\gamma_0,\gamma_1,1/\mathcal{L}_0\}.\tag{2.16}$$

Then we can show the following result.

Lemma 2.2. *Suppose that*

$$\eta\begin{cases}\le\gamma & \text{if } \gamma\ne 1/\mathcal{L}_0,\\ <\gamma & \text{if } \gamma=1/\mathcal{L}_0,\end{cases}\tag{2.17}$$

Then the sequence $\{t_n\}$ *generated by* (2.3) *is well defined, increasing, bounded from above by*

$$t_1^{\star\star}=t_1+\frac{t_2-t_1}{1-\alpha}$$

and converges to its least upper bound $t_1^\star\in[0,t_1^{\star\star}]$*. Moreover, for each* $n=1,2,\dots,$

$$t_{n+2}-t_{n+1}\le\alpha^n(t_2-t_1).\tag{2.18}$$

Proof. As in Lemma 2.1 we shall prove (2.18) using induction. By the choice of γ_1,

$$\alpha_1=\frac{\mathcal{K}+\frac{\mathcal{M}}{3}(t_2-t_1)}{2(1-\mathcal{L}_0t_2)}(t_2-t_1)\le\alpha.\tag{2.19}$$

It follows from (2.19) and (2.15) that

$$\begin{gathered}0<t_3-t_2\le\alpha(t_2-t_1)\\ t_3\le t_2+\alpha(t_2-t_1)\\ t_3\le t_2+(1+\alpha)(t_2-t_1)-(t_2-t_1)\\ t_3\le t_1+\frac{1-\alpha^2}{1-\alpha}(t_2-t_1)<t^{\star\star}.\end{gathered}$$

Assume that

$$0<\alpha_k\le\alpha\tag{2.20}$$

for all $n\le k$. Then, by (2.3) and (2.20),

$$\begin{gathered}0<t_{k+2}-t_{k+1}\le\alpha^k(t_2-t_1),\\ t_{k+2}\le t_1+\frac{1-\alpha^{k+1}}{1-\alpha}(t_2-t_1)<t_1^{\star\star}.\end{gathered}$$

Estimate (2.20) is true with k replaced by $k+1$ provided that

$$\left[\frac{\mathcal{K}}{2}+\frac{\mathcal{M}}{6}(t_{k+2}-t_{k+1})\right](t_{k+2}-t_{k+1})\leq\alpha(1-\mathcal{L}_0t_{k+2})$$

or

(2.21)

$$\left[\frac{\mathcal{K}}{2}+\frac{\mathcal{M}}{6}\alpha^k(t_2-t_1)\right]\alpha^k(t_2-t_1)+\alpha\mathcal{L}_0\left[t_1+\frac{1-\alpha^{k+1}}{1-\alpha}(t_2-t_1)\right]-\alpha\leq 0.$$

Estimate (2.21) motivates us to define recurrent functions f_k on $[0,1)$ by

$$\begin{aligned}f_k(t)=&\left[\frac{\mathcal{K}}{2}+\frac{\mathcal{M}}{6}t^k(t_2-t_1)\right]t^k(t_2-t_1)+t\mathcal{L}_0(1+t+\cdots+t^k)(t_2-t_1)\\&-t(1-\mathcal{L}_0t_1).\end{aligned}$$

We have

$$f_{k+1}(t)=f_k(t)+\left[\frac{1}{2}(2\mathcal{L}_0t^2+\mathcal{K}t-\mathcal{K})+\frac{\mathcal{M}}{6}t^k(t^2-1)(t_2-t_1)\right]t^k(t_2-t_1).$$

In particular, by the choice of α,

$$f_{k+1}(\alpha)\leq f_k(\alpha)\leq\cdots\leq f_1(\alpha)\leq 0. \tag{2.22}$$

Evidently, estimate (2.21) holds if $f_k(\alpha)\leq 0$, or by (2.22) if $f_1(\alpha)\leq 0$, which is true by the choice of η_0. ∎

Lemmas 2.1 and 2.2 admit the following useful extensions. The proofs are omitted since they can simply be obtained by replacing $\eta=t_1-t_0$ with $t_{N+1}-t_N$ where $N=1,2,\dots$ for Lemma 2.3 and $N=2,3,\dots$ for Lemma 2.4.

LEMMA 2.3. *Suppose there exists $N=1,2,\dots$ such that*

$$t_0<t_1<\cdots<t_N<t_{N+1}<1/\mathcal{L}_0$$

and

$$t_{N+1}-t_N\leq\begin{cases}\eta_1 & \text{if } \mathcal{L}_0\eta\leq\dfrac{1-\alpha^2}{2+2\alpha-\alpha^2},\\[2mm] \eta_0 & \text{if } \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2}\leq\mathcal{L}_0\eta.\end{cases}$$

Then the conclusions of Lemma 2.1 for the sequence $\{t_n\}$ hold.

LEMMA 2.4. *Suppose there exists $N=2,3,\dots$ such that*

$$t_0<t_1<\cdots<t_N<t_{N+1}<1/\mathcal{L}_0$$

and

$$\eta\begin{cases}\leq\gamma & \text{if } \gamma\neq 1/\mathcal{L}_0,\\ <\gamma & \text{if } \gamma=1/\mathcal{L}_0,\end{cases}$$

where γ is defined by (2.16) *and where t_2-t_1, t_1, t_2 are replaced, respectively, by $t_{N+1}-t_N$, t_N, t_{N+1}. Then the conclusions of Lemma* 2.1 *for the sequence $\{t_n\}$ hold.*

REMARK 2.5. Another sequence related to Newton's method (1.2) is given by (see Theorem 4.1)

$$s_0 = 0, \quad s_1 = \eta, \quad s_2 = s_1 + \frac{\mathcal{K}_0 + \frac{\mathcal{M}_1}{3}(s_1 - s_0)}{2(1-\mathcal{L}_0 s_1)}(s_1 - s_0)^2, \tag{2.23}$$

$$s_{n+2} = s_{n+1} + \frac{\mathcal{K} + \frac{\mathcal{M}}{3}(s_{n+1} - s_n)}{2(1-\mathcal{L}_0 s_{n+1})}(s_{n+1} - s_n)^2,$$

for each $n = 1, 2, \ldots$ and some $\mathcal{K}_0 \in (0, \mathcal{K}]$, $\mathcal{M}_1 \in (0, \mathcal{M}]$. Then a simple inductive argument shows that

$$s_n \le t_n, \tag{2.24}$$

$$s_{n+1} - s_n \le t_{n+1} - t_n, \tag{2.25}$$

$$s^\star = \lim_{n\to\infty} s_n \le t^\star.$$

Moreover, if $\mathcal{K}_0 < \mathcal{K}$ or $\mathcal{M}_1 < \mathcal{M}$ then (2.24) and (2.25) hold as strict inequalities. Clearly, the sequence $\{s_n\}$ converges under the hypotheses of Lemma 2.1 or Lemma 2.2. However, $\{s_n\}$ can converge under weaker hypotheses than those of Lemma 2.2. Indeed, denote by γ_0^1 and γ_1^1, respectively, the minimal positive zeros of the equations

$$\left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}\alpha(s_2 - s_1)\right](s_2 - s_1) + \mathcal{L}_0(1+\alpha)(s_2 - s_1) + \mathcal{L}_0 s_1 - 1 = 0,$$

$$\left[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_1}{6}(s_2 - s_1)\right](s_2 - s_1) + \alpha\mathcal{L}_0 s_2 - \alpha = 0.$$

Set

$$\gamma^1 = \min\{\gamma_0^1, \gamma_1^1, 1/\mathcal{L}_0\}.$$

Then

$$\gamma \le \gamma^1. \tag{2.26}$$

Moreover, the conclusions of Lemma 2.2 hold for $\{s_n\}$ if (2.26) replaces (2.17).

Note also that strict inequality can hold in (2.26), which implies that the sequence $\{s_n\}$—which is tighter than $\{t_n\}$—converges under weaker conditions.

3. Majorizing sequences—II. We show convergence of sequences that are majorizing for Newton's method (1.2) under the (**H**) conditions.

LEMMA 3.1. *Let $\mathcal{K}_0, \mathcal{L}_0, \mathcal{M}_0, \eta > 0$ with $\mathcal{K}_0 \le \mathcal{L}_0$. Define*

$$a = \frac{2\mathcal{K}_0}{\mathcal{K}_0 + \sqrt{\mathcal{K}_0^2 + 8\mathcal{K}_0\mathcal{L}_0}}, \tag{3.1}$$

$$\theta_0 = \frac{2}{\frac{\mathcal{K}_0}{2} + (1+a)\mathcal{L}_0 + \sqrt{(\mathcal{K}_0/2 + (1+a)\mathcal{L}_0)^2 + 2\mathcal{M}_0(a+3)/3}},$$

$$\theta_1 = \frac{2a}{\mathcal{K}_0/2 + a\mathcal{L}_0 + \sqrt{(\mathcal{K}_0/2 + a\mathcal{L}_0)^2 + 2\mathcal{M}_0 a/3}}.$$

Suppose that

$$\eta \le \begin{cases} \theta_1 & \text{if } \mathcal{L}_0\eta \le \dfrac{1-a^2}{2+2a-a^2}, \\ \theta_0 & \text{if } \dfrac{1-a^2}{2+2a-a^2} \le \mathcal{L}_0\eta. \end{cases} \tag{3.2}$$

Then the sequence $\{v_n\}$ generated by

$$\begin{aligned} &v_0 = 0, \quad v_1 = \eta, \\ &v_{n+2} = v_{n+1} + \frac{\frac{\mathcal{M}_0}{6}(v_{n+1}-v_n) + \frac{\mathcal{M}_0}{2}v_n + \frac{\mathcal{K}_0}{2}}{1-\mathcal{L}_0 v_{n+1}}(v_{n+1}-v_n) \end{aligned} \tag{3.3}$$

is well defined, increasing, bounded from above by

$$v^{\star\star} = \frac{\eta}{1-a}$$

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

$$v_{n+1} - v_n \le a^n\eta, \tag{3.4}$$

$$v^\star - v_n \le \frac{a^n\eta}{1-a}. \tag{3.5}$$

Proof. As in Lemma 2.1 we use induction to prove that

$$\beta_k = \frac{\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}v_k + \frac{\mathcal{M}_0}{6}(v_{k+1}-v_k)}{1-\mathcal{L}_0 v_{k+1}}(v_{k+1}-v_k) \le a. \tag{3.6}$$

Estimate (3.6) holds for $k = 0$ by the choice of θ_1. Let us assume that it holds for all $k \le n$. Then we must prove that

$$\left(\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\frac{1-a^k}{1-a}\eta + \frac{\mathcal{M}_0}{6}a^k\eta\right)a^k\eta + a\mathcal{L}_0\frac{1-a^{k+1}}{1-a}\eta - a \le 0. \tag{3.7}$$

Define recurrent functions f_k on $[0,1)$ for $k = 1, 2, \dots$ by

$$\begin{aligned} f_k(t) = {}&\left(\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}(1 + t + \cdots + +t^{k-1})\eta + \frac{\mathcal{M}_0}{6}a^k\eta\right)t^{k-1}\eta \\ &+ \mathcal{L}_0(1 + t + \cdots + t^k)\eta - 1. \end{aligned} \tag{3.8}$$

Using (3.8), we get

$$f_{k+1}(a) = f_k(a) + \left[\frac{1}{2}(2\mathcal{L}_0 a^2 + \mathcal{K}_0 a - \mathcal{K}_0) + \frac{\mathcal{M}_0}{6}(a^{k+2} + 3a^{k+1} + 2a^k - 3)\eta\right] a^{k-1}\eta \leq f_k(a), \tag{3.9}$$

since a given by (3.1) solves the equation $2\mathcal{L}_0 a^2 + \mathcal{K}_0 a - \mathcal{K}_0 = 0$, and $a^{k+2} + 3a^{k+1} + 2a^k - 3 \leq 0$ for each $k = 1, 2, \dots$ if $a \in [0, 1/2]$. Evidently, it follows from (3.9) that (3.7) holds by the choice of θ_0. ∎

Denote by δ_0 and δ_1, respectively, the minimal positive zeros of the equations

$$\left[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\left(v_2 + \frac{\mathcal{M}_0}{6}a(v_2 - v_1)\right)\right](v_2 - v_1) + \mathcal{L}_0(v_1 + (1+a)(v_2 - v_1)) - 1 = 0,$$

$$\left[\frac{\mathcal{M}_0}{6}(v_2 - v_1) + \frac{\mathcal{M}_0}{2}v_1 + \frac{\mathcal{K}_0}{2}\right](v_2 - v_1) + a\mathcal{L}_0 v_2 - a = 0.$$

Set

$$\delta = \min\{\delta_0, \delta_1, 1/\mathcal{L}_0\}.$$

Then we can show:

Lemma 3.2. *Suppose that*

$$\eta \begin{cases} \leq \delta & \text{if } \delta \neq 1/\mathcal{L}_0, \\ < \delta & \text{if } \delta = 1/\mathcal{L}_0. \end{cases} \tag{3.10}$$

Then the sequence $\{v_n\}$ generated by equation (3.3) is well defined, increasing, bounded from above by

$$v_1^{\star\star} = v_1 + \frac{v_2 - v_1}{1 - a} \tag{3.11}$$

and converges to its least upper bound $v_1^\star \in [0, v_1^{\star\star}]$. Moreover, for each $n = 1, 2, \dots,$

$$v_{n+2} - v_{n+1} \leq a^n(v_2 - v_1).$$

Proof. We have $\beta_1 \leq a$ by the choice of δ_1. This time we must have

$$\left[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\left(v_1 + \frac{1 - a^k}{1 - a}(v_2 - v_1)\right) + \frac{\mathcal{M}_0}{6}a^k(v_2 - v_1)\right]a^k(v_2 - v_1) + a\mathcal{L}_0\left[v_1 + \frac{1 - a^{k+1}}{1 - a}(v_2 - v_1)\right] - a \leq 0. \tag{3.12}$$

Define functions f_k on $[0,1)$ by

(3.13)
$$f_k(t) = \left[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\left(v_1 + \frac{1-t^k}{1-t}(v_2 - v_1)\right) + \frac{\mathcal{M}_0}{6}a^k(t_2 - t_1)\right]t^k(v_2 - v_1) + t\mathcal{L}_0\left[v_1 + \frac{1-t^{k+1}}{1-t}(v_2 - v_1)\right] - t.$$

We have

$$f_{k+1}(a) = f_k(a) + \left[\frac{1}{2}(2\mathcal{L}_0 a^2 + \mathcal{K}_0 a - \mathcal{K}_0) + \frac{\mathcal{M}_0}{6}(v_2 - v_1)\big(3(a-1)a + (a^{k+2} + 3a^{k+1} + 2a^k - 3)\big)\right]a^k(v_2 - v_1).$$

Thus $f_{k+1}(a) \le f_k(a) \le \cdots \le f_1(a)$. But by the choice of η_0 we have $f_1(a) \le 0$. ∎

Remark 3.3. A sequence related to Newton's method (1.2) under the (**H**) conditions is defined by

$$(3.14)\qquad \begin{aligned} &u_0 = 0, \quad u_1 = \eta, \quad u_2 = u_1 + \frac{\mathcal{K}_0 + \frac{\mathcal{M}_1}{3}(u_1 - u_0)}{2(1 - \mathcal{L}_0 u_1)}(u_1 - u_0)^2, \\ &u_{n+2} = u_{n+1} + \frac{\mathcal{K}_0 + \frac{\mathcal{M}_0}{3}(u_{n+1} - u_n)}{2(1 - \mathcal{L}_0 u_{n+1})}(u_{n+1} - u_n)^2 \end{aligned}$$

for $n = 1, 2, \ldots$ and $\mathcal{M}_1 \in (0, \mathcal{M}]$. Then a simple inductive argument shows that for each $n = 2, 3, \ldots,$

$$(3.15)\qquad u_n \le v_n,$$

$$(3.16)\qquad u_{n+1} - u_n \le v_{n+1} - v_n,$$

$$(3.17)\qquad u^\star = \lim_{n\to\infty} \le v^\star.$$

Moreover, if $\mathcal{K}_0 < \mathcal{K}$ or $\mathcal{M}_1 < \mathcal{M}_0$ then (3.15) and (3.16) hold as strict inequalities. The sequence $\{u_n\}$ converges under the hypotheses of Lemma 3.1 or 3.2. However, $\{u_n\}$ can converge under weaker hypotheses than those of Lemma 3.2. Indeed, denote by δ_0^1 and δ_1^1, respectively, the minimal positive zeros of the equations

$$\left[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\left(u_2 + \frac{\mathcal{M}_0}{6}(u_2 - u_1)\right)\right](u_2 - u_1) + \mathcal{L}_0(u_1 + (1+a)(u_2 - u_1)) - 1 = 0,$$

$$\left[\frac{\mathcal{M}_0}{6}(u_2 - u_1) + \frac{\mathcal{M}_0}{2}u_1 + \frac{\mathcal{K}_0}{2}\right](u_2 - u_1) + a\mathcal{L}_0 u_2 - a = 0.$$

Set

$$\delta^1 = \min\{\delta_0^1, \delta_1^1, 1/\mathcal{L}_0\}. \tag{3.18}$$

Then $\delta \le \delta^1$. Moreover, the conclusions of Lemma 3.2 hold for the sequence $\{u_n\}$ if (3.18) replaces (3.10). Note also that strict inequality may hold in (3.18), which implies that, tighter than $\{v_n\}$, the sequence $\{u_n\}$ converges under weaker conditions. Finally note that $\{t_n\}$ is tighter than $\{v_n\}$ although the sufficient convergence conditions for $\{v_n\}$ are weaker than those for $\{t_n\}$.

Lemmas similar to Lemmas 2.3 and 2.4 for $\{v_n\}$ can be obtained in an analogous way.

4. Semilocal convergence. We present the semilocal convergence of Newton's method (1.2) first under the (**C**) and then under the (**H**) conditions. Let $U(x, R)$ and $\overline{U}(x, R)$ stand, respectively, for the open and closed balls in $\mathbf{X}$ centered at $x \in \mathbf{X}$ and of radius $R > 0$.

THEOREM 4.1. *Let $\mathcal{F} : \mathbf{D} \subseteq \mathbf{X} \to \mathbf{Y}$ be twice Fréchet differentiable. Suppose that the* (**C**) *conditions and the hypotheses of Lemma* 2.1 *hold and*

$$\overline{U}(x_0, t^\star) \subseteq \mathbf{D}. \tag{4.1}$$

Then the sequence $\{x_n\}$ defined by Newton's method (1.2) *is well defined, remains in $\overline{U}(x_0, t^\star)$ for all $n \ge 0$ and converges to a unique solution $x^\star \in \overline{U}(x_0, t^\star)$ of the equation $\mathcal{F}(x) = 0$. Moreover, for all $n \ge 0$,*

$$\|x_{n+2} - x_{n+1}\| \le t_{n+2} - t_{n+1}, \tag{4.2}$$

$$\|x_n - x^\star\| \le t^\star - t_n, \tag{4.3}$$

where $\{t_n\}$ $(n \ge 0)$ is given by (2.3). *Furthermore, if there exists $R \ge t^\star$ such that*

$$U(x_0, R) \subseteq \mathbf{D} \quad \text{and} \quad \mathcal{L}_0(t^\star + R) \le 2,$$

then the solution $x^\star$ is unique in $U(x_0, R)$.

Proof. Let us prove that

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \tag{4.4}$$

$$\overline{U}(x_{k+1}, t^\star - t_{k+1}) \subseteq \overline{U}(x_k, t^\star - t_k), \tag{4.5}$$

for all $k \ge 0$. For every $z \in \overline{U}(x_1, t^\star - t_1)$,

$$\begin{aligned}\|z - x_0\| &\le \|z - x_1\| + \|x_1 - x_0\| \\ &\le (t^\star - t_1) + (t_1 - t_0) = t^\star - t_0\end{aligned}$$

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

$$\|x_1 - x_0\| = \|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_0)\| \le \eta = t_1 - t_0,$$

(4.4) and (4.5) hold for $k = 0$. Given they hold for $n = 0, 1 \dots, k$, we have

$$\|x_{k+1} - x_0\| \leq \sum_{i=1}^{k+1} \|x_i - x_{i-1}\| \leq \sum_{i=1}^{k+1} (t_i - t_{i-1}) = t_{k+1} - t_0 = t_{k+1},$$
$$\|x_k + \theta(x_{k+1} - x_k) - x_0\| \leq t_k + \theta(t_{k+1} - t_k) \leq t^\star,$$

for all $\theta \in [0, 1]$. Using (1.2), we obtain the approximation

$$\begin{aligned}
\mathcal{F}(x_{k+1}) &= \mathcal{F}(x_{k+1}) - \mathcal{F}(x_k) - \mathcal{F}'(x_k)(x_{k+1} - x_k) \\
&= \int_0^1 [\mathcal{F}'(x_k + \theta(x_{k+1} - x_k)) - \mathcal{F}'(x_k)](x_{k+1} - x_k)\, d\theta \\
&= \int_0^1 \mathcal{F}''(x_k + \theta(x_{k+1} - x_k))(1 - \theta)(x_{k+1} - x_k)^2\, d\theta.
\end{aligned}$$

Then we get, by $(\mathbf{C}_3)$, $(\mathbf{C}_4)$ and (4.1),

$$\begin{aligned}
(4.6) \quad & \|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1})\| \\
&\leq \int_0^1 \big(\|\mathcal{F}'(x_0)^{-1}[\mathcal{F}''(x_k + \theta(x_{k+1} - x_k)) - \mathcal{F}''(x^\star)]\| \\
&\quad + \|\mathcal{F}'(x_0)^{-1}\mathcal{F}''(x^\star)\|\big)\|x_{k+1} - x_k\|^2(1 - \theta)\, d\theta \\
&\leq \left[\mathcal{M}\left(\int_0^1 \|x_{k+1} - x_k\|(1 - \theta)\, d\theta\right) + \frac{\mathcal{K}}{2}\right]\|x_{k+1} - x_k\|^2 \\
&\leq \frac{\mathcal{M}}{6}\|x_{k+1} - x_k\|^3 + \frac{\mathcal{K}}{2}\|x_{k+1} - x_k\|^2 \\
&\leq \left[\overline{\mathcal{M}}\left(\frac{1}{6}(t_{k+1} - t_k)\right) + \frac{\overline{\mathcal{K}}}{2}\right](t_{k+1} - t_k)^2,
\end{aligned}$$

where

$$\overline{\mathcal{K}} = \begin{cases} \mathcal{K}_0, & \mathcal{K} = 0, \\ \mathcal{K}, & \mathcal{K} > 0, \end{cases} \quad \text{and} \quad \overline{M} = \begin{cases} \mathcal{M}_0, & \mathcal{K} = 0, \\ \mathcal{M}, & \mathcal{K} > 0. \end{cases}$$

Using $(\mathbf{C}_5)$, we obtain

$$\begin{aligned}
(4.7) \qquad \|\mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x_{k+1}) - \mathcal{F}'(x_0))\| &\leq \mathcal{L}_0\|x_{k+1} - x_0\| \\
&\leq \mathcal{L}_0 t_{k+1} \leq \mathcal{L}_0 t^\star < 1.
\end{aligned}$$

It follows from the Banach lemma on invertible operators [7, 8, 13–15] and (4.7) that $\mathcal{F}'(x_{k+1})^{-1}$ exists and

$$(4.8) \quad \|\mathcal{F}'(x_{k+1})^{-1}\mathcal{F}'(x_0)\| \leq (1 - \mathcal{L}_0\|x_{k+1} - x_0\|)^{-1} \leq (1 - \mathcal{L}_0 t_{k+1})^{-1}.$$

Therefore by (1.2), (4.6) and (4.8), we obtain in turn

$$\begin{aligned}\|x_{k+2}-x_{k+1}\| &\le \|\mathcal{F}'(x_{k+1})^{-1}\mathcal{F}'(x_{k+1})\| \\ &\le \|\mathcal{F}'(x_{k+1})^{-1}\mathcal{F}'(x_0)\|\,\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1})\| \le t_{k+2}-t_{k+1}.\end{aligned}$$

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

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

That is,

$$z \in \overline{U}(x_{k+1}, t^\star - t_{k+1}). \tag{4.9}$$

Estimates (4.8) and (4.9) imply that (4.4) and (4.5) hold for $n = k+1$. The proof of (4.4) and (4.5) is now complete by induction.

Lemma 2.1 implies that $\{t_n\}$ is a Cauchy sequence. From (4.4) and (4.5), $\{x_n\}$ is a Cauchy sequence too and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star)$. Estimate (4.3) follows from (4.2) by using standard majorization techniques [7, 8, 13–15, 17]. Moreover, by letting $k \to \infty$ in (4.6), we obtain $\mathcal{F}(x^\star) = 0$. Finally, to show uniqueness let $y^\star$ be a solution of $\mathcal{F}(x) = 0$ in $U(x_0, R)$. It follows from $(\mathbf{C}_5)$ for $x = y^\star + \theta(x^\star - y^\star)$, $\theta \in [0,1]$, that

$$\begin{aligned}&\Big\|\mathcal{F}'(x_0)^{-1}\int_0^1 (\mathcal{F}'(y^\star+\theta(x^\star-y^\star)) - \mathcal{F}'(x_0))\Big\|\,d\theta \\ &\qquad \le \mathcal{L}_0\int_0^1 \|y^\star+\theta(x^\star-y^\star)-x_0\|\,d\theta \\ &\qquad \le \mathcal{L}_0\int_0^1 (\theta\|x^\star-x_0\| + (1-\theta)\|y^\star-x_0\|)\,d\theta \\ &\qquad \le \frac{\mathcal{L}_0}{2}(t^\star+R) \le 1 \quad \text{(by (4.1))},\end{aligned}$$

and the Banach lemma on invertible operators implies that the linear operator $T^{\star\star} = \int_0^1 \mathcal{F}'(y^\star+\theta(x^\star-y^\star))\,d\theta$ is invertible. Using the identity $0 = \mathcal{F}(x^\star) - \mathcal{F}'(y^\star) = T^{\star\star}(x^\star - y^\star)$, we deduce that $x^\star = y^\star$.

Similarly, we show the uniqueness in $\overline{U}(x_0, t^\star)$ by setting $t^\star = R$. ∎

REMARK 4.2. The conclusions of Theorem 4.1 hold if $\{t_n\}$, $t^\star$ are replaced by $\{r_n\}$, $r^\star$, respectively.

Using the approximation

$$\begin{aligned}&\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1}) \\ &\quad = \int_0^1 \mathcal{F}'(x_0)^{-1}[\mathcal{F}''(x_k+\theta(x_{k+1}-x_k)) - \mathcal{F}''(x_0)](x_{k+1}-x_k)^2(1-\theta)\,d\theta \\ &\qquad + \int_0^1 \mathcal{F}'(x_0)^{-1}\mathcal{F}''(x_0)(1-\theta)\,d\theta\,\|x_{k+1}-x_k\|^2\end{aligned}$$

instead of (4.6) and ($\mathbf{C}_6$), ($\mathbf{C}_7$) instead of, respectively, ($\mathbf{C}_3$), ($\mathbf{C}_4$), we arrive at the following semilocal convergence result under the ($\mathbf{H}$) conditions [7, Theorem 6.3.7 p. 210 for proof].

THEOREM 4.3. *Let $\mathcal{F} : \mathbf{D} \subseteq \mathbf{X} \to \mathbf{Y}$ be twice Fréchet differentiable. Furthermore suppose that the* ($\mathbf{H}$) *conditions, $\overline{U}(x_0, v^\star) \subseteq \mathbf{D}$, and the hypotheses of Lemma* 3.1 *hold. Then the sequence $\{x_n\}$ generated by Newton's method* (1.2) *is well defined, remains in $\overline{U}(x_0, t^\star)$ for all $n \geq 0$ and converges to a unique solution $x^\star \in \overline{U}(x_0, t^\star)$ of the equation $\mathcal{F}(x) = 0$. Moreover, for all $n \geq 0$,*

$$\|x_{n+2} - x_{n+1}\| \leq v_{n+2} - v_{n+1}, \qquad \|x_n - x^\star\| \leq v^\star - v_n \tag{4.10}$$

where $\{v_n\}$ $(n \geq 0)$ is given by (3.3). *Furthermore, if there exists $R \geq t^\star$ such that*

$$U(x_0, R) \subseteq \mathbf{D} \quad \textit{and} \quad \mathcal{L}_0(t^\star + R) \leq 2,$$

then the solution $x^\star$ is unique in $U(x_0, R)$.

5. Local convergence. We study the local convergence of Newton's method under the ($\mathbf{A}$) *conditions*:

$\mathbf{A}_1$. there exists $x^\star \in \mathbf{D}$ such that $\mathcal{F}(x^\star) = 0$ and $\mathcal{F}'(x^\star)^{-1} \in \mathbf{L}(\mathbf{Y}, \mathbf{X})$,
$\mathbf{A}_2$. $\|\mathcal{F}'(x^\star)^{-1}\mathcal{F}''(x^\star)\| \leq b$,
$\mathbf{A}_3$. $\|\mathcal{F}'(x^\star)^{-1}[\mathcal{F}''(x) - \mathcal{F}''(x^\star)]\| \leq c\|x - x^\star\|$ for each $x \in \mathbf{D}$,
$\mathbf{A}_4$. $\|\mathcal{F}'(x^\star)^{-1}[\mathcal{F}'(x) - \mathcal{F}'(x^\star)]\| \leq d\|x - x^\star\|$ for each $x \in \mathbf{D}$.

Note also that in view of ($\mathbf{A}_3$) and ($\mathbf{A}_4$), respectively, there exist $c_0 \in (0, c]$ and $d_0 \in (0, d]$ such that for each $\theta \in [0, 1]$:

$\mathbf{A}'_3$. $\|\mathcal{F}'(x^\star)^{-1}(\mathcal{F}''(x_0 + \theta(x^\star - x_0)) - \mathcal{F}''(x^\star))\| \leq c_0(1 - \theta)\|x_0 - x^\star\|$,
$\mathbf{A}'_4$. $\|\mathcal{F}'(x^\star)^{-1}(\mathcal{F}'(x_0) - \mathcal{F}'(x^\star))\| \leq d_0(1 - \theta)\|x_0 - x^\star\|$.

Then we can show:

THEOREM 5.1. *Suppose that the* ($\mathbf{A}$) *conditions hold and $\overline{U}(x^\star, r) \subseteq \mathbf{D}$, where*

$$r = \frac{2}{b/2 + d + \sqrt{(b/2 + d)^2 + 4c/3}}. \tag{5.1}$$

Then the sequence $\{x_n\}$ (starting from $x_0 \in U(x^\star, r)$) generated by Newton's method (1.2) *is well defined, remains in $U(x^\star, r)$ for all $n \geq 0$ and converges to $x^\star$. Moreover,*

$$\|x_{n+1} - x^\star\| \leq e_n\|x_n - x^\star\|^2, \tag{5.2}$$

$$e_n = \frac{\frac{\overline{c}}{3}\|x_n - x^\star\| + \frac{b}{2}}{1 - \overline{d}\|x_n - x^\star\|} \quad \textit{and} \quad q(t) = \frac{\frac{ct}{3} + \frac{b}{2}}{1 - dt}t$$

where

$$\bar{c} = \begin{cases} c_0 & \textit{if } n = 0, \\ c & \textit{if } n > 0, \end{cases} \qquad \bar{d} = \begin{cases} d_0 & \textit{if } n = 0, \\ d & \textit{if } n > 0. \end{cases}$$

Proof. The starting point x_0 is in $U(x^\star, r)$. Suppose that $x_k \in U(x^\star, r)$ for all $k \le n$. Using $(\mathbf{A}_4)$ and the definition of r we get

$$\|\mathcal{F}'(x^\star)^{-1}(\mathcal{F}'(x_k) - \mathcal{F}'(x^\star))\| \le d\|x_k - x^\star\| < dr < 1. \tag{5.3}$$

It follows from (5.3) and the Banach lemma on invertible operators that $\mathcal{F}'(x_k)^{-1}$ exists and

$$\|\mathcal{F}'(x_k)^{-1}\mathcal{F}'(x^\star)\| \le \frac{1}{1 - d\|x_k - x^\star\|}. \tag{5.4}$$

Hence, x_{k+1} exists. Using (1.2), we obtain the approximation

$$\begin{aligned} &x^\star - x_{k+1} \\ &= -\mathcal{F}'(x_k)^{-1}\mathcal{F}'(x^\star)\int_0^1 \mathcal{F}'(x^\star)^{-1}\big[\big(\mathcal{F}''(x_k + \theta(x^\star - x_k)) - \mathcal{F}''(x^\star)\big) + \mathcal{F}''(x^\star)\big] \\ &\qquad\qquad \times (x^\star - x_k)^2(1-\theta)\,d\theta. \end{aligned} \tag{5.5}$$

In view of $(\mathbf{A}_2)$, $(\mathbf{A}_3)$, $(\mathbf{A}_4)$, (5.4), (5.5) and the choice of r we have

$$\begin{aligned} \|x_{k+1} - x^\star\| &\le \frac{c\int_0^1(1-\theta)^2\|x_k - x^\star\|^3\,d\theta + b\int_0^1(1-\theta)\,d\theta\,\|x_k - x^\star\|^2}{1 - d\|x_k - x^\star\|} \\ &\le e_k\|x_k - x^\star\|^2 < q(r)\|x_k - x^\star\| = \|x_k - x^\star\|, \end{aligned}$$

which implies that $x_{k+1} \in U(x^\star, r)$ and $\lim_{k\to\infty} x_k = x^\star$. ∎

Remark 5.2. The local results or projection methods such as the Arnolds method, the generalized minimum residual method (GMRES), the generalized conjugate method (GCR) for combined Newton/finite projection methods and in connection with the mesh independence principle can be used to develop the cheapest and most efficient mesh refinement strategies [4, 7, 8, 14, 15]. These results can also be used to solve equations of the form (1.1), where $\mathcal{F}'$, $\mathcal{F}''$ satisfy differential equations of the form

$$\mathcal{F}'(x) = \mathcal{P}(\mathcal{F}(x)) \quad \text{and} \quad \mathcal{F}''(x) = \mathcal{Q}(\mathcal{F}(x)).$$

where $\mathcal{P}$ and $\mathcal{Q}$ are known operators. Since $\mathcal{F}'(x^\star) = \mathcal{P}(\mathcal{F}(x^\star)) = \mathcal{P}(0)$ and $\mathcal{F}''(x^\star) = \mathcal{Q}(\mathcal{F}(x^\star)) = \mathcal{Q}(0)$ we can apply our results without actually knowing the solution $x^\star$ of equation (1.1).

6. Numerical examples

Example 6.1. Let $\mathbf{X} = \mathbf{Y} = \mathbb{R}$ be equipped with the max-norm, $x_0 = \omega$, $\mathbf{D} = [-\exp(1), \exp(1)]$. Let us define $\mathcal{F}$ on $\mathbf{D}$ by

$$\mathcal{F}(x) = x^3 - \exp(1). \tag{6.1}$$

Here, $\omega \in \mathbf{D}$. Through some algebraic manipulations, we obtain

$$\begin{cases} \eta = \dfrac{|\omega^3 - \exp(1)|}{3\omega^2}, \quad \mathcal{K} = \dfrac{4\exp(1)}{\omega^2}, \quad \mathcal{L}_0 = \dfrac{2\exp(1) + \omega}{\omega^2}, \quad \mathcal{K}_0 = \dfrac{2}{\omega}, \\ \mathcal{M} = \dfrac{2}{\omega^2}, \quad \mathcal{M}_0 = \dfrac{2}{\omega^2}. \end{cases}$$

For $\omega = 0.48\exp(1)$, the criteria (1.3) and (1.6) read

$$0.09730789545 \leq 0.07755074734 \quad \text{and} \quad 0.09730789545 \leq 0.2856823952$$

respectively. Thus we observe that the criterion (1.3) fails while the criterion (1.6) holds. The hypothesis of Lemma 2.1 reads

$$0.09730789545 \leq \begin{cases} 0.2017739733 & \text{if } 0.08268226632 \leq 0.2499999999, \\ 0.2036729480 & \text{if } 0.2499999999 \leq 0.08268226632. \end{cases}$$

Table 1. Newton's method applied to (6.1)

n	x_n	$\|x_{n+2} - x_{n+1}\|$	$\|x_n - x^\star\|$
0	$1.30 \times 10^{+00}$	6.44×10^{-03}	9.08×10^{-02}
1	$1.40 \times 10^{+00}$	2.98×10^{-05}	6.47×10^{-03}
2	$1.40 \times 10^{+00}$	6.37×10^{-10}	2.98×10^{-05}
3	$1.40 \times 10^{+00}$	2.91×10^{-19}	6.37×10^{-10}
4	$1.40 \times 10^{+00}$	6.06×10^{-38}	2.91×10^{-19}
5	$1.40 \times 10^{+00}$	2.63×10^{-75}	6.06×10^{-38}
6	$1.40 \times 10^{+00}$	4.95×10^{-150}	2.63×10^{-75}
7	$1.40 \times 10^{+00}$	1.76×10^{-299}	4.95×10^{-150}
8	$1.40 \times 10^{+00}$	2.22×10^{-598}	1.76×10^{-299}
9	$1.40 \times 10^{+00}$	3.52×10^{-1196}	2.22×10^{-598}

Table 2. Sequences $\{t_n\}$ (2.3)

n	t_n	$t_{n+2} - t_{n+1}$	$t^\star - t_n$
0	$0.00 \times 10^{+00}$	4.95×10^{-02}	1.69×10^{-01}
1	9.73×10^{-02}	1.87×10^{-02}	7.16×10^{-02}
2	1.47×10^{-01}	3.26×10^{-03}	2.21×10^{-02}
3	1.66×10^{-01}	1.02×10^{-04}	3.36×10^{-03}
4	1.69×10^{-01}	1.01×10^{-07}	1.02×10^{-04}
5	1.69×10^{-01}	9.75×10^{-14}	1.01×10^{-07}
6	1.69×10^{-01}	9.16×10^{-26}	9.75×10^{-14}
7	1.69×10^{-01}	8.08×10^{-50}	9.16×10^{-26}
8	1.69×10^{-01}	6.30×10^{-98}	8.08×10^{-50}
9	1.69×10^{-01}	3.82×10^{-194}	6.30×10^{-98}

Thus the hypothesis of Lemma 2.1 holds. As a consequence, we can apply Theorem 4.1. Table 1 reports the convergence behavior of Newton's method (1.2) applied to (4.6) with $x_0 = 1$ and $\psi = 0.55$. Numerical computations are performed to the decimal point accuracy of 2005 by employing the high-precision library ARPREC.

Table 2 reports the behavior of $\{t_n\}$ (2.3).

Comparing Tables 1 and 2, we observe that the estimates of Theorem 4.1 hold.

Example 6.2. In this example, we provide an application of our results to a special nonlinear Hammerstein integral equation of the second kind. Consider the integral equation

$$x(s) = 1 + \frac{4}{5}\int_0^1 G(s,t)x(t)^3\,dt, \quad s \in [0,1], \tag{6.2}$$

where G is the Green kernel on $[0,1]\times[0,1]$ defined by

$$G(s,t) = \begin{cases} t(1-s), & t \le s, \\ s(1-t), & s \le t. \end{cases} \tag{6.3}$$

Let $\mathbf{X} = \mathbf{Y} = \mathcal{C}[0,1]$ and let $\mathbf{D}$ be a suitable open convex subset of $\mathbf{X}_1 := \{x \in \mathbf{X} : x(s) > 0,\ s \in [0,1]\}$, which will be given below. Define $\mathcal{F} : \mathbf{D} \to \mathbf{Y}$ by

$$[\mathcal{F}(x)](s) = x(s) - 1 - \frac{4}{5}\int_0^1 G(s,t)x(t)^3\,dt, \quad s \in [0,1]. \tag{6.4}$$

The first and second derivatives of $\mathcal{F}$ are given by

$$[\mathcal{F}(x)'y](s) = y(s) - \frac{12}{5}\int_0^1 G(s,t)x(t)^2y(t)\,dt, \quad s \in [0,1],$$

$$[\mathcal{F}(x)''yz](s) = \frac{24}{5}\int_0^1 G(s,t)x(t)y(t)z(t)\,dt, \quad s \in [0,1].$$

We use the max-norm. Let $x_0(s) = 1$ for all $s \in [0,1]$. Then, for any $y \in \mathbf{D}$, we have

$$[(I - \mathcal{F}'(x_0))(y)](s) = \frac{12}{5}\int_0^1 G(s,t)y(t)\,dt, \quad s \in [0,1],$$

which means

$$\|I - \mathcal{F}'(x_0)\| \le \frac{12}{5}\max_{s\in[0,1]}\int_0^1 G(s,t)\,dt = \frac{12}{5\times 8} = \frac{3}{10} < 1.$$

It follows from the Banach theorem that $\mathcal{F}'(x_0)^{-1}$ exists and

$$\|\mathcal{F}'(x_0)^{-1}\| \leq \frac{1}{1-\frac{3}{10}} = \frac{10}{7}. \tag{6.5}$$

On the other hand, from (6.4) we have

$$\|\mathcal{F}(x_0)\| = \frac{4}{5}\max_{s\in[0,1]}\int_0^1 G(s,t)\,dt = \frac{1}{10}.$$

Thus, we get $\eta = 1/7$. Note that $\mathcal{F}''(x)$ is not bounded in $\mathbf{X}$ or its subset $\mathbf{X}_1$. Take into account that a solution $x^\star$ of (1.1) with $\mathcal{F}$ given by (6.4) must satisfy

$$\|x^\star\| - 1 - \frac{1}{10}\|x^\star\|^3 \leq 0,$$

i.e., $\|x^\star\| \leq \rho_1 = 1.153467305$ and $\|x^\star\| \geq \rho_2 = 2.423622140$, where ρ_1 and ρ_2 are the positive roots of the real equation $z-1-z^3/10=0$. Consequently, if we look for a solution such that $x^\star < \rho_1 \in \mathbf{X}_1$, we can consider $\mathbf{D} := \{x : x\in\mathbf{X}_1 \text{ and } \|x\|<r\}$, with $r\in(\rho_1,\rho_2)$, as a nonempty open convex subset of $\mathbf{X}$. For example, choose $r=1.7$. Using (3.4) and (3.5), we obtain, for any $x,y,z\in\mathbf{D}$,

$$\begin{aligned}
&\|[(\mathcal{F}'(x)-\mathcal{F}'(x_0))y](s)\|\\
&\qquad = \frac{12}{5}\Big\|\int_0^1 G(s,t)(x(t)^2-x_0(t)^2)y(t)\,dt\Big\|\\
&\qquad \leq \frac{12}{5}\int_0^1 G(s,t)\|x(t)-x_0(t)\|\,\|x(t)+x_0(t)\|y(t)\,dt\\
&\qquad \leq \frac{12}{5}\int_0^1 G(s,t)(r+1)\|x(t)-x_0(t)\|y(t)\,dt, \quad s\in[0,1],
\end{aligned}$$

and

$$\|(F''(x)yz)(s)\| = \frac{24}{5}\int_0^1 G(s,t)x(t)y(t)z(t)\,dt, \quad s\in[0,1].$$

Then we get

$$\|\mathcal{F}'(x)-\mathcal{F}'(x_0)\| \leq \frac{12}{5}\,\frac{1}{8}(r+1)\|x-x_0\| = \frac{81}{100}\|x-x_0\|,$$
$$\|F''(x)\| \leq \frac{24}{5}\times\frac{r}{8} = \frac{51}{50}$$

and

$$\|[[F''(x) - \mathcal{F}''(\overline{x})]yz](s)\| = \frac{24}{5}\left\|\int_0^1 G(s,t)(x(t) - \overline{x}(t)))y(t)z(t)\,dt\right\|$$
$$\leq \frac{24}{5}\frac{1}{8}\|x - \overline{x}\| = \frac{3}{5}\|x - \overline{x}\|.$$

Now we can choose constants as follows:

$$\eta = \frac{1}{7}, \quad \mathcal{M} = \frac{6}{7}, \quad \mathcal{M}_0 = \frac{6}{7}, \quad \mathcal{K} = \frac{51}{35}, \quad \mathcal{L}_0 = \frac{49}{70}, \quad \mathcal{K}_0 = \frac{11}{15}.$$

Conditions (1.3) and (1.5) read

$$0.1428571429 < 0.3070646192 \quad \text{and} \quad R_1 = 0.1627780248.$$

Conditions (1.6) and (1.8) read

$$0.1428571429 < 0.4988741112 \quad \text{and} \quad R_2 = 0.1518068730.$$

The hypotheses (2.2) and (3.2) read

$$\frac{1}{7} \leq \begin{cases} 0.5047037049 & \text{if } 0.1000000000 \leq 0.2131833880, \\ 0.5228360736 & \text{if } 0.2131833880 \leq 0.1000000000, \end{cases}$$

and

$$\frac{1}{7} \leq \begin{cases} 0.6257238049 & \text{if } 0.1000000000 \leq 0.2691240473, \\ 0.5832936968 & \text{if } 0.2691240473 \leq 0.1000000000, \end{cases}$$

respectively. Thus hypotheses (2.2) and (3.2) hold. Comparison of sequences (2.3), (2.23), (3.3) and (3.14) is reported in Table 3. From Table 3, we observe that the estimates (2.24) and (3.15) hold.

Table 3. Comparison of the sequences (2.3), (2.23), (3.3) and (3.14)

n	t_n	s_n	v_n	u_n
0	$0.000000 \times 10^{+00}$	$0.000000 \times 10^{+00}$	$0.000000 \times 10^{+00}$	$0.000000 \times 10^{+00}$
1	1.428571×10^{-01}	1.428571×10^{-01}	1.428571×10^{-01}	1.428571×10^{-01}
2	1.598408×10^{-01}	1.514801×10^{-01}	2.042976×10^{-01}	1.516343×10^{-01}
3	1.600782×10^{-01}	1.515408×10^{-01}	2.356037×10^{-01}	1.516661×10^{-01}
4	1.600783×10^{-01}	1.515408×10^{-01}	2.527997×10^{-01}	1.516661×10^{-01}
5	1.600783×10^{-01}	1.515408×10^{-01}	2.626215×10^{-01}	1.516661×10^{-01}
6	1.600783×10^{-01}	1.515408×10^{-01}	2.683548×10^{-01}	1.516661×10^{-01}
7	1.600783×10^{-01}	1.515408×10^{-01}	2.717435×10^{-01}	1.516661×10^{-01}
8	1.600783×10^{-01}	1.515408×10^{-01}	2.737612×10^{-01}	1.516661×10^{-01}
9	1.600783×10^{-01}	1.515408×10^{-01}	2.749678×10^{-01}	1.516661×10^{-01}

Concerning the uniqueness balls, from equation (1.5), we get $R_1 = 0.1627780248$ and from equation (1.8), we get $R_2 = 0.1518068730$, whereas from Theorem 4.1, we get $R \leq 1.257142857$. Therefore, the new approach provides the largest uniqueness ball.

EXAMPLE 6.3. Let us consider the case when $\mathbf{X} = \mathbf{Y} = \mathbb{R}$, $\mathbf{D} = U(0,1)$ and define $\mathcal{F}$ on $\mathbf{D}$ by

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

Then we can define $\mathcal{P}(x) = x + 1$ and $\mathcal{Q}(x) = x + 1$. To compare our radius of convergence with earlier ones, we introduce the Lipschitz condition

$$\|\mathcal{F}'(x^\star)^{-1}(\mathcal{F}'(x) - \mathcal{F}'(y))\| \le \mathcal{L}\|x - y\| \quad \text{for each } x, y \in \mathbf{D}. \tag{6.6}$$

The radius of convergence given by Traub–Woźniakowski [7, 8, 15] is

$$r_0 = \frac{2}{3\mathcal{L}}. \tag{6.7}$$

The radius of convergence given by us in [5–7] is

$$r_1 = \frac{2}{2d + \mathcal{L}}. \tag{6.8}$$

Using $(\mathbf{A}_2)$, $(\mathbf{A}_3)$, $(\mathbf{A}_4)$ and (6.6), we get $b = 1$, $c = d = e - 1$ and $\mathcal{L} = e$. Then, using (5.1), (6.7) and (6.8), we obtain

$$r = 0.4078499356 > r_1 = 0.324947231 > r_0 = 0.245252961.$$

EXAMPLE 6.4. Let $X = Y = \mathbb{R}^3$, $D = U(0,1)$, $x^* = (0,0,0)$ and define $\mathcal{F}$ on $\mathbf{D}$ by

$$\mathcal{F}(x,y,z) = \left(e^x - 1, \frac{e-1}{2}y^2 + y, z\right)^{\mathrm{T}}. \tag{6.9}$$

For $u = (x,y,z)$ we have

$$\mathcal{F}'(u) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & (e-1)y + 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathcal{F}''(u) = \left(\begin{array}{ccc|ccc|ccc} e^x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & e-1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

and

$$\mathcal{F}'''(u) = \left(\begin{array}{c|c|c|c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \begin{matrix} e^x & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \end{array}\right).$$

Using the $(\mathbf{A})$ and $(\mathbf{A}')$ conditions—and $\mathcal{F}'(x^*) = \operatorname{diag}\{1,1,1\}$—we set

$$b = 1.0, \quad \bar{c} = c_0 = c = \bar{d} = d_0 = d = e - 1, \quad \mathcal{L} = e, \quad \text{and} \quad \mathcal{L}_0 = e - 1.$$

We obtain

$$r = 0.4078499356.$$

Thus, $r_0 < r$.

The following iterations have been used before:

$$\|x_{n+1} - x^\star\| \le p_n \|x_n - x^\star\|^2 \quad [6\text{–}8,\ 12],$$
$$\|x_{n+1} - x^\star\| \le \lambda_n \|x_n - x^\star\|^2 \quad [6\text{–}8],$$
$$\|x_{n+1} - x^\star\| \le \mu_n \|x_n - x^\star\|^2 \quad [15],$$
$$\|x_{n+1} - x^\star\| \le \xi_n \|x_n - x^\star\|^2 \quad [6\text{–}8,\ 15],$$

where

$$p_n = \frac{(\mathcal{L}/3)\|x_n - x^\star\| + b/2}{1 - d\|x_n - x^\star\|}, \quad \lambda_n = \frac{\mathcal{L}/2}{1 - \mathcal{L}_0\|x_n - x^\star\|},$$
$$\mu_n = \frac{\mathcal{L}/2}{1 - \mathcal{L}\|x_n - x^\star\|} \quad \text{and} \quad \xi_n = \frac{(\mathcal{L}/3)\|x_n - x^\star\| + b/2}{1 - \big((\mathcal{L}/2)\|x_n - x^\star\| + b\big)\|x_n - x^\star\|}.$$

To compare the above iterations with the iteration (5.2), we produce the comparison tables 4 and 5; we apply Newton's method (1.2) to the equation (6.9) with $x_0 = (0.21, 0.21, 0.21)^{\mathrm{T}}$. In Table 4, we note that the estimate (5.2)—of Theorem 5.1—holds.

Table 4. Comparison of various iterative procedures

n	$\|x_{n+1} - x^\star\|$	$e_n\|x_n - x^\star\|^2$	$\lambda_n\|x_n - x^\star\|^2$
1	0.034624745433299	0.292667362771974	0.479494429606589
2	0.000669491177317	0.000677347930013	0.001732513344520
3	0.000000347374133	0.000000224639537	0.000000609893622
4	0.000000000000103	0.000000000000060	0.000000000000164
5	0.000000000000000	0.000000000000000	0.000000000000000

Table 5. Comparison of various iterative procedures

n	$\mu_n\|x_n - x^\star\|^2$	$\xi_n\|x_n - x^\star\|^2$
1	15.944478671072201	0.240445748047369
2	0.001798733838791	0.000661013573819
3	0.000000610302684	0.000000224531576
4	0.000000000000164	0.000000000000060
5	0.000000000000000	0.000000000000000

Acknowledgements. We thank the reviewers for useful and constructive remarks which have improved the quality of our work.

References

[1] S. Amat, C. Bermúdez, S. Busquier, M. J. Legaz, and S. Plaza, *On a family of high-order iterative methods under Kantorovich conditions and some applications*, Abstract Appl. Anal. 2012, art. ID 782170, 14 pp.

[2] S. Amat, C. Bermúdez, S. Busquier and S. Plaza, *On a third-order Newton-type method free of bilinear operators*, Numerical Linear Algebra Appl. 17 (2010), 639–653.

[3] S. Amat and S. Busquier, *Third-order iterative methods under Kantorovich conditions*, J. Math. Anal. Appl. 336 (2007), 243–261.

[4] S. Amat, S. Busquier and J. M. Gutiérrez, *Third-order iterative methods with applications to Hammerstein equations: A unified approach*, J. Comput. Appl. Math. 235 (2011), 2936–2943.

[5] I. K. Argyros, *A Newton–Kantorovich theorem for equations involving m-Fréchet differentiable operators and applications in radiative transfer*, J. Comput. Appl. Math. 131 (2001), 149–159.

[6] I. K. Argyros, *On the Newton–Kantorovich hypothesis for solving equations*, J. Comput. Appl. Math. 169 (2004), 315–332.

[7] I. K. Argyros, *Computational Theory of Iterative Methods*, Elsevier, 2007.

[8] I. K. Argyros, Y. J. Cho and S. Hilout, *Numerical Methods for Equations and its Applications*, CRC Press/Taylor and Francis, New York, 2012.

[9] J. A. Ezquerro, D. González and M. A. Hernández, *Majorizing sequences for Newton's method from initial value problems*, J. Comput. Appl. Math. 236 (2012), 2246–2258.

[10] J. A. Ezquerro, D. González and M. A. Hernández, *On the local convergence of Newton's method under generalized conditions of Kantorovich*, Appl. Math. Lett. 26 (2013), 566–570.

[11] J. A. Ezquerro and M. A. Hernández, *Generalized differentiability conditions for Newton's method*, IMA J. Numer. Anal. 22 (2002), 187–205.

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

[13] L. V. Kantorovich, *The majorant principle and Newton's method*, Dokl. Akad. Nauk SSSR 76 (1951), 17–20 (in Russian).

[14] A. M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, Academic Press, New York, 1973.

[15] J. F. Traub, *Iterative Methods for the Solution of Equations*, Amer. Math. Soc., 1982.

[16] W. Werner, *Some improvements of classical iterative methods for the solution of nonlinear equations*, in: Numerical Solution of Nonlinear Equations, Lecture Notes in Math. 878, Springer, 1981, 426–440.

[17] T. Yamamoto, *On the method of tangent hyperbolas in Banach spaces*, J. Comput. Appl. Math. 21 (1988), 75–86.

Ioannis K. Argyros
Department of Mathematical Sciences
Cameron University
Lawton, OK 73505-6377, U.S.A.
E-mail: iargyros@cameron.edu

Sanjay K. Khattri
Department of Engineering
Stord/Haugesund University College
N-414 Stord, Norway
E-mail: sanjay.khattri@hsh.no

Received on 22.3.2013;
revised version on 17.11.2013 (2174)