APPLICATIONES MATHEMATICAE
28,1 (2001), pp. 1–15

IOANNIS K. ARGYROS (Lawton, OK)

ON A NEW METHOD FOR ENLARGING THE RADIUS OF CONVERGENCE FOR NEWTON'S METHOD

Abstract. We provide new local and semilocal convergence results for Newton's method. We introduce Lipschitz-type hypotheses on the mth Fréchet derivative. This way we manage to enlarge the radius of convergence of Newton's method. Numerical examples are also provided to show that our results guarantee convergence where others do not.

1. Introduction. Let $F : D \subseteq E_1 \to E_2$ be an m times continuously Fréchet-differentiable operator ($m \ge 2$ a positive integer) defined on an open convex subset D of a Banach space E_1 with values in a Banach space E_2. Suppose there exists $x^* \in D$ which is a solution of the equation

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

The most popular method for approximating such a point x^* is Newton's method:

$$x_{n+1} = G(x_n) \quad (n \ge 0,\ x_0 \in D), \tag{2}$$

where

$$G(x) \equiv x - F'(x)^{-1}F(x) \quad (x \in D). \tag{3}$$

Here $F'(x) \in L(E_1, E_2)$ $(x \in D)$, the space of bounded linear operators from E_1 into E_2. Sufficient convergence conditions for the convergence of Newton's method under Lipschitz hypotheses on the first Fréchet derivative have been given by many authors [1]–[8]. In particular, we refer the interested reader to [3] for a survey of such results. In the elegant paper [8] by Ypma, affine invariant results have been given concerning the radius of convergence

---

2000 *Mathematics Subject Classification*: 65B05, 47H17, 49D15.
Key words and phrases: Newton's method, Banach space, radius of convergence, affine invariant operator, mth Fréchet derivative.

of Newton's method. Ypma used Lipschitz conditions on the first Fréchet derivative as the basis for his analysis. In this study we use Lipschitz-like conditions on the mth Fréchet derivative $F^{(m)}(x) \in L(E_1^m, E_2)$ ($x \in D,\ m \geq 2$ a positive integer). This way we manage to enlarge the radius of convergence for Newton's method (2). Finally we provide numerical examples to show that our results guarantee convergence, where earlier ones do not [8]. This is important in numerical computations [3], [4], [6], [8].

2. Convergence analysis. We give an affine invariant form of the Banach lemma on invertible operators.

LEMMA 1. *Let $m \geq 2$ be a positive integer, $\alpha_2 > 0$, $\alpha_i \geq 0$ $(3 \leq i \leq m+1)$, $\eta \geq 0$, E_1, E_2 Banach spaces, D a convex subset of E_1 and $F : D \to E_2$ an m-times Fréchet-differentiable operator. Assume there exist $z \in D$ so that $F'(z)^{-1}$ exists, and some convex neighborhood $N(z)$ of z such that $N(z) \subseteq D$,*

$$\|F'(z)^{-1}F^{(i)}(z)\| \leq \alpha_i, \qquad i = 2, \ldots, m, \tag{4}$$

and

$$\|F'(z)^{-1}[F^{(m)}(x) - F^{(m)}(z)]\| \leq \alpha_{m+1}\|x - z\| \quad \text{for all } x \in N(z). \tag{5}$$

If $x \in N(z) \cap U(z, \delta)$, with δ the positive zero of the equation $f'(t) = 0$, where

$$f(t) = \frac{\alpha_{m+1}}{(m+1)!}t^{m+1} + \frac{\alpha_m}{m!}t^m + \ldots + \frac{\alpha_2}{2!}t^2 - t + d \tag{6}$$

then $F'(x)^{-1}$ exists and for $\|x - z\| < t \leq \delta$,

$$\|F'(z)^{-1}F''(x)\| \leq f''(t) \tag{7}$$

and

$$\|F'(x)^{-1}F'(z)\| \leq -f'(t)^{-1}. \tag{8}$$

Proof. It is convenient to define ε, b_1, b_i, $i = 2, \ldots, m$, by $\varepsilon = x - z_0$, $b_1 = z + \theta_1\varepsilon$, $b_i = z + \theta_i(b_{i-1} - z)$, $\theta_i \in [0, 1]$. We can have in turn

$$\begin{aligned} F''(x) &= F''(z) + [F''(x) - F''(z)] \\ &= F''(z) + \int_0^1 F'''[z + \theta_1(x - z)](x - z)\, d\theta_1 \\ &= F''(z) + \int_0^1 [F'''(z + \theta_1(x - z)) - F'''(z)](x - z)\, d\theta_1 \\ &\quad + \int_0^1 F'''(z)(x - z)\, d\theta_1 \end{aligned} \tag{9}$$

$$= F''(z) + \int_0^1 F'''(z)(x-z)\,d\theta_1$$

$$+ \int_0^1\int_0^1 F^{(4)}\{z+\theta_2[z+\theta_1(x-z)-z]\}[z+\theta_1(x-z)z](x-z)\,d\theta_2\,d\theta_1$$

$$= F''(z) + \int_0^1 F'''(z)\varepsilon\,d\theta_1 + \int_0^1\int_0^1 F^{(4)}(b_2)(b_1-z_0)\varepsilon\,d\theta_2\,d\theta_1 = \dots$$

$$= F''(z) + \int_0^1 F'''(z)\varepsilon\,d\theta_1 + \dots$$

$$+ \int_0^1\dots\int_0^1 F^{(m)}(b_{m-2})(b_{m-3}-z)\dots(b_1-z)\,d\theta_{m-2}\dots d\theta_1$$

$$= F''(z) + \int_0^1 F'''(z)\varepsilon d\theta_1 + \dots$$

$$+ \int_0^1\dots\int_0^1 F^{(m)}(z)(b_{m-3}-z)\dots(b_1-z)\varepsilon\,d\theta_{m-2}\dots d\theta_1$$

$$+ \int_0^1\dots\int_0^1 [F^{(m)}(b_{m-2})-F^{(m)}(z)](b_{m-3}-z)\dots(b_1-z)\varepsilon\,d\theta_{m-2}\dots d\theta_1.$$

Using the triangle inequality, (4), (5) and (6) in (9) after composing by $F'(z)^{-1}$ we obtain (7).

We also get

$$\begin{aligned}
(10)\qquad & -F'(z)^{-1}[F'(z)-F'(x)] \\
&= F'(z)^{-1}[F'(x)-F'(z)-F''(z)(x-z)+F''(z)(x-z)] \\
&= \int_0^1 F'(z)^{-1}\{F''[z+\theta_1\varepsilon]-F''(z)\}\,d\theta_1\varepsilon + F'(z)^{-1}\int_0^1 F''(z)\varepsilon\,d\theta_1 \\
&= \int_0^1\int_0^1 F'(z)^{-1}F'''(b_2)(b_1-z)\varepsilon\ d\theta_2\,d\theta_1 + F'(z)^{-1}\int_0^1 F''(z)\varepsilon\,d\theta_1 = \dots \\
&= \int_0^1\dots\int_0^1 F^{(m)}(b_{m-1})(b_{m-2}-z)\dots(b_1-z)\varepsilon\,d\theta_{m-1}\,d\theta_{m-2}\dots d\theta_2\,d\theta_1
\end{aligned}$$

$$+\int_0^1\dots\int_0^1 F^{(m-1)}(b_{m-2})(b_{m-3}-z)\dots(b_1-z)\varepsilon\,d\theta_{m-2}\dots d\theta_2\,d\theta_1$$

$$+\dots+\int F'(z)^{-1}F''(z)\varepsilon\,d\theta_1$$

$$=\int_0^1\dots\int_0^1 F'(z)^{-1}[F^{(m)}(b_{m-1})-F^{(m)}(z)](b_{m-2}-z)$$

$$\dots(b_1-z)\varepsilon\,d\theta_{m-2}\dots d\theta_1$$

$$+\int_0^1\dots\int_0^1 F'(z)^{-1}F^{(m)}(z)(b_{m-2}-z)\dots(b_1-z)\varepsilon\,d\theta_{m-1}\dots d\theta_1$$

$$+\int_0^1\dots\int_0^1 F'(z)^{-1}F^{(m-1)}(z)(b_{m-3}-z)\dots(b_1-z)\varepsilon d\theta_{m-2}\dots d\theta_1$$

$$+\dots+\int_0^1 F'(z)^{-1}F''(z)\varepsilon\,d\theta_1.$$

Since $f'(t)<0$ on $[0,\delta]$, using (4)–(6) in (10) we obtain, for $\|x-z\|<t$,

$$\|-F'(z)^{-1}[F'(z)-F'(x)]\|\le 1+f'(\|x-z\|)<1+f'(t)<1. \tag{11}$$

It follows from the Banach lemma on invertible operators [3], [7], [8] and from (11) that $F'(x)^{-1}$ exists and

$$\|F'(x)^{-1}F'(z)\|\le[1-\|F'(z)^{-1}[F'(z)-F'(x)]]^{-1}\le -f'(t)^{-1},$$

which shows (8). ■

We need the following affine invariant form of the mean value theorem for m-Fréchet-differentiable operators.

LEMMA 2. *Let $m\ge 2$ be a positive integer, $\alpha_2>0$, $\alpha_i\ge 0$ $(3\le i\le m+1)$, E_1, E_2 Banach spaces, D a convex subset of E_1 and $F:D\to E_2$ an m-times Fréchet-differentiable operator. Assume there exist $z\in D$ so that $F'(z)^{-1}$ exists, and some convex neighborhood $N(z)$ of z such that $N(z)\subseteq D$,*

$$\|F'(z)^{-1}F^{(i)}(z)\|\le\alpha_i,\qquad i=2,\dots,m,$$

and

$$\|F'(z)^{-1}[F^{(m)}(x)-F^{(m)}(z)]\|\le\alpha_{m+1}\|x-z\|\quad \textit{for all } x\in N(z).$$

Then for all $x\in N(z)$,

$$\|F'(z)^{-1}[F(z)-F(x)-F'(x)(z-x)]\| \le\frac{m\alpha_{m+1}}{(m+1)!}\|x-z\|^{m+1}+\frac{(m-1)\alpha_m}{m!}\|x-z\|^m+\dots+\frac{\alpha_2}{2!}\|x-z\|^2. \tag{12}$$

Proof. We can write in turn:

$$
\begin{aligned}
(13)\quad & F(z)-F(x)-F'(x)(z-x)\\
&= \int_0^1 [F'(x+\theta_1(z-x))-F'(x)](z-x)\,d\theta_1\\
&= \int_0^1 [F''(z+\theta_1(x-z))-F''(z)]\theta_1\,d\theta_1(x-z)^2+\int_0^1 \theta_1 F''(z)(x-z)^2\,d\theta_1\\
&= \int_0^1\int_0^1 [F'''(z+\theta_2\theta_1(x-z))-F'''(z)]\theta_1(x-z)\,d\theta_2\,\theta_1\,d\theta_1(x-z)^2\\
&\quad+\int_0^1\int_0^1 F'''(z)\theta_1(x-z)\,d\theta_2\,\theta_1\,d\theta_1(x-z)^2+\int_0^1 \theta_1 F''(z)(x-z)^2\,d\theta_1=\ldots\\
&= \int_0^1\int_0^1\ldots\int_0^1 [F^{(m)}(z+\theta_{m-1}\theta_{m-2}\ldots\theta_1(x-z))-F^{(m)}(z)]\theta_{m-2}^1\\
&\quad\ldots\theta_3^{m-4}\theta_2^{m-3}\theta_1^{m-1}(x-z)^m\,d\theta_{m-1}\,d\theta_{m-2}\ldots d\theta_3\,d\theta_2\,d\theta_1\\
&\quad+\ldots+\int_0^1\int_0^1 F'''(z)\theta_1^2(x-z)^3\,d\theta_2\,d\theta_1+\int_0^1 \theta_1 F''(z)(x-z)^2\,d\theta_1.
\end{aligned}
$$

Composing both sides by $F'(z)^{-1}$, using the triangle inequality, (5) and (6) we obtain (12). ∎

Based on the above lemmas we derive affine invariant convergence results for the class $T \equiv T(\{\alpha_i\},\ 2 \le i \le m+1, \alpha)$ ($\alpha > 0$, $\alpha_2 > 0$, $\alpha_i \ge 0$, $3 \le i \le m+1$) of operators $F : D \subseteq E_1 \to E_2$ such that: D is an open convex set; F is m times continuously Fréchet-differentiable on D; there exists $x^* \in D$ such that $F(x^*) = 0$; $F'(x^*)^{-1}$ exists; $U(x^*, \alpha) \subseteq D$; x^* is the only solution of the equation $F(x) = 0$ in $U(x^*, \alpha)$; and for all $x \in U(x^*, \alpha)$,

$$\|F'(x^*)^{-1}[F^{(m)}(x^*)-F^{(m)}(x)]\| \le \alpha_{m+1}\|x^*-x\|, \tag{14}$$

and

$$\|F'(x^*)^{-1}F^{(i)}(x^*)\| \le \alpha_i, \qquad i=2,\ldots,m. \tag{15}$$

Let $F \in T$ and $x \in U(x^*, b)$ where $b \le \min\{\alpha, \delta\}$. By Lemma 1, $F'(x)^{-1}$ exists. Define

$$\mu(F,x) \equiv \sup\left\{\frac{\|F'(x)^{-1}[F^{(m)}(y)-F^{(m)}(x^*)]\|}{\|y-x^*\|}\ \middle|\ y\in U(x^*,b);\ y\neq x^*\right\}, \tag{16}$$

$$q_i = q_i(F,x) \equiv \|F'(x)^{-1}F^{(i)}(x^*)\|, \qquad 2 \le i \le m,\ x \in U(x^*, b). \tag{17}$$

It follows from (14)–(17) that

$$\mu(F,x^*) \le \alpha_{m+1}, \quad q_i(F,x^*) \le \alpha_i, \quad 2 \le i \le m, \tag{18}$$

$F \in T(\{q_i\}$, $2 \le i \le m$, $\mu(F,x^*),\alpha)$, and by Lemma 1,

$$\mu(F,x) \le \frac{\mu(F,x^*)}{1 - q_2\|x-x^*\| - \ldots - \frac{\mu(F,x^*)}{m!}\|x-x^*\|^m} \equiv \overline{\mu}(x). \tag{19}$$

We also have the estimates

$$\begin{aligned} \|F'(x)^{-1}F^{(i)}(x^*)\| &\le \|F'(x)^{-1}F'(x^*)\|\,\|F'(x^*)^{-1}F^{(i)}(x^*)\| \\ &\le q_i\|F'(x)^{-1}F'(x^*)\| \\ &\le \frac{q_i}{1-\alpha_2\|x-x^*\| - \ldots - \frac{\alpha_{m+1}}{m!}\|x-x^*\|^m} \equiv \overline{q}_i(x). \end{aligned} \tag{20}$$

The following lemma on fixed points is important.

Lemma 3. *Let F, x be as above. Then the Newton operator G defined in (3) satisfies*

$$\begin{aligned} \|G(x)-x^*\| \le{}& \frac{\mu(F,x)m}{(m+1)!}\|x-x^*\|^{m+1} \\ &+ \frac{q_m(m-1)}{m!}\|x-x^*\|^m + \ldots + \frac{q_2}{2!}\|x-x^*\|^2 \end{aligned} \tag{21}$$

and

$$\begin{aligned} &\|G(x)-x^*\| \\ &\le \frac{\frac{m\alpha_{m+1}}{(m+1)!}\|x-x^*\|^{m+1} + \frac{(m-1)\alpha_m}{m!}\|x-x^*\|^m + \ldots + \frac{\alpha_2}{2!}\|x-x^*\|^2}{1-\alpha_2\|x-x^*\| - \ldots - \frac{\alpha_{m+1}}{m!}\|x-x^*\|^m}. \end{aligned} \tag{22}$$

Proof. Using (3) we can write

$$\begin{aligned} G(x)-x^* &= x - F'(x)^{-1}F(x) - x^* \\ &= F'(x)^{-1}[F'(x)(x-x^*) - F(x)] \\ &= F'(x)^{-1}[F(x^*) - F(x) - F'(x)(x^*-x)] \\ &= [F'(x)^{-1}F'(x^*)]\{F'(x^*)^{-1}[F(x^*)-F(x)-F'(x)(x^*-x)]\}. \end{aligned} \tag{23}$$

As in Lemma 1 by taking norms in (23) and using (14), (15) we obtain (21). Moreover, using Lemma 2 and (12) we get (22). ■

Remark 1. Consider Newton's method (2)–(3) for some $x_0 \in U(x^*,b)$. Define a sequence $\{c_n\}$ $(n \ge 0)$ by

$$c_n \equiv \|x_n - x^*\| \quad (n \ge 0) \tag{24}$$

and a function g on $[0,\delta)$ by

$$(25)\qquad g(t) \equiv \frac{\dfrac{m\alpha_{m+1}}{(m+1)!}t^{m+1} + \dfrac{(m-1)\alpha_m}{m!}t^m + \ldots + \dfrac{\alpha_2}{2!}t^2}{1 - \alpha_2 t - \ldots - \dfrac{\alpha_{m+1}}{m!}t^m}.$$

By (24) and (25), estimate (22) becomes

$$(26)\qquad c_{n+1} \le g(c_n) \quad (n \ge 0).$$

It is simple algebra to show that $g(t) < t$ iff $t < \delta_0$, with δ_0 the positive zero of the equation

$$(27)\qquad h(t) = 0,$$

where

$$(28)\qquad h(t) = \frac{(2m+1)\alpha_{m+1}}{(m+1)!}t^m + \frac{(2m-1)\alpha_m}{m!}t^{m-1} + \ldots + \frac{3}{2!}\alpha_2 t - 1.$$

Note that for $m = 2$, using (28) we obtain

$$(29)\qquad \delta_0 = \frac{12}{9\alpha_2 + \sqrt{81\alpha_2^2 + 120\alpha_3}}.$$

Hence, we proved the following local convergence result for Newton's method (2)–(3).

THEOREM 1. *Newton's method* $\{x_n\}$ $(n \ge 0)$ *generated by* (2)–(3) *converges to the solution* x^* *of the equation* $F(x) = 0$, *for all* $F \in T$, *iff the initial guess* x_0 *satisfies*

$$(30)\qquad \|x_0 - x^*\| < \min\{\alpha, \delta_0\}.$$

We also have the following consequence of Theorem 1:

THEOREM 2. *Newton's method* $\{x_n\}$ $(n \ge 0)$ *generated by* (2)–(3) *converges to the solution* x^* *of the equation* $F(x) = 0$, *for all* $F \in T$, *if* $F'(x_0)^{-1}$ *exists at the initial guess* x_0, *and*

$$(31)\qquad \|x_0 - x^*\| < \min\{\alpha, \overline{\delta}_0\},$$

where $\overline{\delta}_0$ *is the positive zero of the equation resulting from* (28) *by replacing* α_{m+1} *with* $\mu(F, x_0)$ (*defined by* (10)) *and* α_i, $2 \le i \le m$, *with* $q_i(F, x_0)$ (*defined by* (17)).

Proof. By Lemma 1, since $F'(x_0)^{-1}$ exists and $\|x_0 - x^*\| < \overline{\delta}_0$, we get

$$(32)\qquad \mu(F, x^*) \le m_0 \equiv \frac{\mu(F, x_0)}{1 - q_2(F, x_0)\|x_0 - x^*\| - \ldots - \dfrac{\mu(F, x_0)}{m!}\|x - x_0\|^m}.$$

Moreover, we have

$$(33)\quad \begin{aligned} q_i(F,x^*) &= \|F'(x^*)^{-1}F^{(i)}(x^*)\| \\ &\le \|F'(x^*)^{-1}F'(x_0)\|\cdot\|F'(x_0)^{-1}F^{(i)}(x^*)\| \\ &\le q_0^i \equiv \frac{q_i(F,x_0)}{1-q_2(F,x_0)\|x_0-x^*\|-\ldots-\dfrac{\mu(F,x_0)}{m!}\|x_0-x^*\|^m}. \end{aligned}$$

Denote by $\overline{\overline{\delta}}_0$ the positive zero of the equation resulting from (28) by replacing α_{m+1} with $\mu(F,x^*)$ (defined by (16)) and α_i, $2\le i\le m$, with $q_i(F,x^*)$. Furthermore, denote by $\overline{\overline{\overline{\delta}}}_0$ the positive zero of the equation resulting from (28) by replacing α_{m+1} with m_0 and α_i, $2\le i\le m$, with q_0^i.

Using the above definitions we get

$$(34)\quad \begin{aligned} \overline{\overline{\delta}}_0 \ge \overline{\overline{\overline{\delta}}}_0 \ge{}& \frac{\mu(F,x_0)m}{(m+1)!}\|x_0-x^*\|^{m+1} \\ &+\frac{q_m(F,x_0)(m-1)}{m!}\|x_0-x^*\|^m \\ &+\ldots+\frac{q_2(F,x_0)}{2!}\|x_0-x^*\| \\ \ge{}& \|G(x_0)-x^*\|. \end{aligned}$$

The result now follows from (34) and Theorem 1. ■

Remark 2. Let us assume equality in (26) and consider the iteration $c_{n+1}=g(c_n)$ $(n\ge 0)$. Denote the numerator of g by g_1 and the denominator by g_2. By Ostrowski's theorem for convex functions [1], [3] the iteration $\{c_n\}$ $(n\ge 0)$ converges to 0 if $c_0\in[0,\overline{\overline{\overline{\delta}}}_0)$, $g'(c_0)<1$. Define the real function h_0 by

$$(35)\quad h_0(t)=g_2(t)^2-g_1'(t)g_2(t)+g_2'(t)g_1(t),$$

where $\overline{\alpha}_{m+1}=\mu(F,x^*)$ and $\overline{\alpha}_i=q_i(F,x^*)$, $2\le i\le m$, replace α_{m+1} and α_i in the definition of g respectively. Note that h is a polynomial of degree $2m$ and can be written in the form

$$(36)\quad \begin{aligned} h_0(t)={}&\frac{(m+1)!(m-1)!+(m!)^2}{(m!)^2(m+1)!(m-1)!}\overline{\alpha}_{m+1}^2t^{2m} \\ &+(\text{other lower order terms})+1. \end{aligned}$$

For example, in case $m=2$,

$$(37)\quad h_0(t)=\frac{5}{12}\overline{\alpha}_3^2t^4+\frac{7}{6}\overline{\alpha}_3\overline{\alpha}_2t^3+\left(\frac{3\overline{\alpha}_2^2}{2}-2\overline{\alpha}_3\right)t^2-3\overline{\alpha}_2t+1.$$

Since h_0 is continuous and $h_0(0)=1>0$, we deduce that there exists $t_0>0$ such that $h_0(t)>0$ for all $t\in[0,t_0)$.

Set

$$\overline{c}_0 = \min\{t_0, \overline{\overline{\delta}}_0\}. \tag{38}$$

It is simple algebra to show that $g'(c_0) < 1$ iff $h_0(c_0) > 0$. Hence, Newton's method converges to x^* for all $F \in T$ if the initial guess x_0 satisfies

$$\|x_0 - x^*\| \le \min\{\alpha, \overline{c}_0\}. \tag{39}$$

Condition (39) is weaker than (31).

Although Theorem 1 gives an optimal domain of convergence for Newton's method, the rate of convergence may be slow for x_0 near the boundaries of that domain. However, it is known that if the conditions of the Newton–Kantorovich theorem [3], [7] are satisfied at x_0 then convergence is rapid. The proof of this theorem can essentially be found in [3].

THEOREM 3. *Let $m \ge 2$ be a positive integer, E_1, E_2 Banach spaces, D an open convex subset of E_1, and $F : D \to E_2$ an m-times Fréchet-differentiable operator. Let $x_0 \in D$ be such that $F'(x_0)^{-1}$ exists, and suppose the positive numbers δ^*, $d(F, x_0)$, $\alpha_i(F, x_0)$, $2 \le i \le m+1$, satisfy*

$$\|F'(x_0)^{-1}F(x_0)\| \le d(F, x_0), \tag{40}$$

$$\|F'(x_0)^{-1}F^{(i)}(x_0)\| \le \alpha_i(F, x_0), \qquad i = 2, \ldots, m, \tag{41}$$

and

$$\|F'(x_0)^{-1}[F^{(m)}(x) - F^{(m)}(x_0)]\| \le \alpha_{m+1}(F, x_0)\|x - x_0\| \tag{42}$$

for all $x \in U(x_0, \delta^*) \subseteq D$. Denote by s the positive zero of the scalar equation

$$p'(t) = 0, \tag{43}$$

where

$$p(t) = \frac{\alpha_{m+1}(F, x_0)}{(m+1)!}t^{m+1} + \frac{\alpha_m(F, x_0)}{m!}t^m + \ldots + \frac{\alpha_2(F, x_0)}{2!}t^2 - t + d(F, x_0). \tag{44}$$

If

$$p(s) \le 0, \tag{45}$$

and

$$\delta^* \ge r_1, \tag{46}$$

where r_1 is the smallest nonnegative zero of the equation

$$p(t) = 0,$$

guaranteed to exist by (45), *then Newton's method* (2)–(3) *starting from x_0 generates a sequence which converges quadratically to an isolated solution x^* of the equation $F(x) = 0$.*

REMARK 3. Using this theorem we obtain two further sufficiency conditions for the convergence of Newton's method. It is convenient for us to set $\alpha_{m+1}(F, x_0) = \mu(F, x_0)$, and $\alpha_i(F, x_0) = q_i$ (q_i evaluated at x_0), $2 \le i \le m$. Condition (45) can be written as

$$d(F, x_0) \le s_0, \tag{47}$$

where

$$s_0 = s - \left[\frac{q_2}{2!} s^2 + \ldots + \frac{\mu(F, x_0)}{(m+1)!} s^{m+1}\right] > 0 \tag{48}$$

by the definition of s. Define functions h_1, h_2 by

$$h_1(t) = \frac{\mu(F, x_0)m}{(m+1)!} t^{m+1} + \frac{q_m(m-1)}{m!} t^m + \ldots + \frac{q_2}{2!} t^2 + t - s_0, \tag{49}$$

and

$$\begin{aligned} h_2(t) = {} & \frac{\overline{\mu}(x_0)m}{(m+1)!} t^{m+1} \\ & + \frac{\overline{q}_m(x_0)(m-1)}{m!} t^m + \ldots + \frac{\overline{q}_2(x_0)}{2!} t^2 + t - s_0. \end{aligned} \tag{50}$$

Since $h_1(0) = h_2(0) = -s_0 < 0$, we deduce that there exist minimum $t_1 > 0$, $t_2 > 0$ such that

$$h_1(t) \le 0 \quad \text{for all } t \in [0, t_1] \tag{51}$$

and

$$h_2(t) \le 0 \quad \text{for all } t \in [0, t_2]. \tag{52}$$

THEOREM 4. *Let $F \in T$, and $x_0 \in U(x^*, \alpha)$. Then condition* (45) *holds if either*

(a) *$F'(x_0)^{-1}$ exists and $\|x_0 - x^*\| \le \min\{\alpha, t_1\}$, or*
(b) *$F'(x_0)^{-1}$ exists and $\|x_0 - x^*\| \le \min\{\alpha, t_2\}$,*

where t_1 and t_2 are defined in (51) *and* (52) *respectively.*

Proof. Choose $\delta^* > 0$ such that $U(x_0, \delta^*) \subseteq U(x^*, \alpha)$. By (3) and (21), we get (for $\alpha_{m+1}(G, x_0) = \mu(F, x_0)$, and $\alpha_i(F, x_0) = q_i$ (q_i evaluated at x_0), $g \le i \le m$):

$$\begin{aligned} \|F'(x_0)^{-1}F(x_0)\| = \|G(x_0) - x_0\| \le \|F(x_0) - x^*\| + \|x^* - x_0\| \\ \le \frac{\mu(F, x_0)m}{(m+1)!} \|x_0 - x^*\|^{m+1} + \frac{q_m(m-1)}{m!} \|x_0 - x^*\|^m \\ + \ldots + \frac{q_2}{2!} \|x_0 - x^*\|^2 + \|x_0 - x^*\|. \end{aligned} \tag{53}$$

Using (53) to replace $d(F, x_0)$ in (44) and setting $\|x_0 - x^*\| \le t$, we deduce that (45) holds if $h_1(t) \le 0$, which is true by the choice of t_1 and (a). Moreover, by replacing $\mu(G, x_0)$ and q_i, $2 \le i \le m$, using (19) and (20) respectively, condition (45) holds if $h_2(t) \le 0$, which is true by the choice of t_2 and (b). ■

3. Applications. The results obtained here have theoretical and practical value. As an example we consider the operator F in (1); it satisfies the autonomous differential equation of the form [3], [7]

$$F'(x) = Q(F(x)) \quad (x \in D), \tag{54}$$

where $Q : E_2 \to E_1$ is a known $m-1$ times Fréchet-differentiable operator. Using (54) we get $F'(x^*) = Q(F(x^*)) = Q(0)$ and $F''(x^*) = F'(x^*)Q'(F(x^*)) = Q(0)Q'(0)$, etc. That is, without knowing x^* we can use Theorem 1 (for example) to solve the equation $F(x) = 0$, using Newton's method (2)–(3).

Here is such a case:

Example 1. Let $E_1 = E_2 = \mathbb{R}$, $D = U(0,1)$, and define a function F on D by

$$F(x) = e^x - 1. \tag{55}$$

We can set $Q(x) = x + 1$ $(x \in D)$. Then F satisfies (54). Let $m = 2$; then $\alpha = 1$, $x^* = 0$, $\alpha_2 = 1$, $\alpha_3 = e$ and by (28) or (29) we get

$$\delta_0^2 = .411254048. \tag{56}$$

For $m = 3$, $\alpha_2 = \alpha_3 = 1$ and $\alpha_4 = e$. Using (28) we get

$$\delta_0^3 = .480112. \tag{57}$$

To compare our results with earlier ones, note that in Theorem 3.7 of [8, p. 111] the condition is

$$\|x_0 - x^*\| < \min\{\sigma, 2//(3\varrho)\} = \varrho_0, \tag{58}$$

where σ, ϱ are such that $U(x^*, \sigma) \subseteq D$, and

$$\|F'(x^*)^{-1}(F'(x) - F'(y))\| \le \varrho\|x - y\| \quad \text{for all } x, y \in U(x^*, \sigma). \tag{59}$$

Letting $\sigma = \alpha = 1$, by (59) we get $\varrho = e$, and condition (58) becomes

$$\|x_0 - x^*\| < \varrho_0 \equiv .245253. \tag{60}$$

Comparing (56), (57) and (60) we observe that (56) or (57) allow a wider choice of initial guesses x_0 than (60). For example, if we choose $x_0 = .4$, Theorem 3.7 of [8, p. 111] cannot guarantee that Newton's method (2)–(3) starting from $x_0 = .4$ converges to $x^* = 0$, which is the solution of the equation $F(x) = 0$ where F is given by (55). However, due to (56) or (57), our Theorem 1 guarantees convergence in this case.

REMARK 4. Our analysis can be simplified if instead of (22) we consider the following estimate: since $x \in U(x^*, \alpha)$ there exist γ_1, γ_2 such that

$$(61)\qquad 2\left[\frac{m\alpha_{m+1}}{(m+1)!}\|x_0 - x^*\|^{m-1} + \frac{(m-1)\alpha_m}{m!}\|x_0 - x^*\|^{m-2} + \ldots + \frac{\alpha_2}{2!}\right] \le \gamma_1$$

and

$$(62)\qquad \frac{\alpha_{m+1}}{m!}\|x_0 - x^*\|^{m-1} + \ldots + \alpha_2 \le \gamma_2.$$

Hence estimate (22) can be written as

$$(63)\qquad \|G(x) - x^*\| \le \frac{\gamma_1}{2(1 - \gamma_2\|x - x^*\|)}\|x - x^*\|^2,$$

and for $\gamma^* = \max\{\gamma_1, \gamma_2\}$,

$$(64)\qquad \|G(x) - x^*\| \le \frac{\gamma^*}{2(1 - \gamma^*\|x - x^*\|)}\|x - x^*\|^2.$$

The convergence condition of Theorem 3.7 in [8, p. 111] and (63), (64), becomes respectively

$$(65)\qquad \|x_0 - x^*\| \le \min\{\alpha, \gamma\}, \qquad \gamma = \frac{2}{\gamma_1 + 2\gamma_2},$$

and

$$(66)\qquad \|x_0 - x^*\| \le \min\left\{\sigma, \frac{2}{3\gamma^*}\right\}.$$

In particular, estimate (66) is similar to (58), and if $\gamma < \varrho$, then (65) allows a wider range for the initial guess x_0 than (58). Returning back to the numerical example we can have

$$\delta = b = .565444814, \qquad \gamma_1 = 2.024692242, \qquad \gamma_2 = 1.7685192,$$
$$\gamma = .359600299, \qquad \gamma^* = \gamma_1, \qquad \frac{2}{3\gamma^*} = .329268144.$$

That is, both (65) and (66) provide a wider range for the initial guess x_0 than (58). Moreover, based on the stronger (but easier to check) condition (65) or (66) we can generate most of the results in [8].

Furthermore, if (4), (5) and (58) hold, our analysis can be based on the following variations of (22):

$$(67)\qquad \|G(x) - x^*\| \le \frac{\frac{\mu(F,x)m}{(m+1)!}\|x - x^*\|^{m+1} + \ldots + \frac{q_2}{2!}\|x - x^*\|^2}{1 - \varrho\|x - x^*\|},$$

and

$$\|G(x)-x^*\| \le \frac{\varrho}{2[1-\alpha_2\|x-x^*\|-\ldots-\frac{\alpha_{m+1}}{m!}\|x-x^*\|^m]}\|x-x^*\|^2. \tag{68}$$

REMARK 5. The results obtained here with slight modifications can be extended to hold in the following cases:

CASE 1. Replace condition (5) by

$$\|F'(z)^{-1}(F^{(m)}(x)-F^{(m)}(z))\| \le \alpha_{m+1}\|x-z\|^\lambda, \qquad \lambda \ge 0. \tag{69}$$

Then the polynomial f in (6) is given by

$$\begin{aligned} f(t) = {} & \left(\frac{1}{\lambda+1}\right)\left(\frac{1}{\lambda+2}\right)\cdots\left(\frac{1}{\lambda+m}\right)\alpha_{m+1}t^{\lambda+m}+\frac{\alpha_m}{m!}t^m \\ & + \ldots + \frac{\alpha_2}{2!}t^2 - t + d. \end{aligned} \tag{70}$$

The function associated with the numerator in (22) is given by

$$\begin{aligned} \widetilde{f}(t) = {} & \frac{(\lambda+m-1)}{(\lambda+1)(\lambda+2)\ldots(\lambda+m)}t^{\lambda+m}+\frac{(m-1)\alpha_m}{m!}t^m \\ & + \ldots + \frac{\alpha_2}{2!}t^2 - t + d. \end{aligned} \tag{71}$$

CASE 2. Replace condition (5) by

$$\|F'(z)^{-1}(F^{(m)}(x)-F^{(m)}(z))\| \le w(\|x-z\|), \tag{72}$$

where w is an increasing positive function on $[0,\alpha]$ with $\lim_{t\to 0} w(t) = 0$ (see [3]). Then the polynomial f in (6) is given by

$$\begin{aligned} f(t) = {} & \int_0^1\ldots\int_0^1 w(\theta_1\theta_2\theta_3\ldots\theta_{m-1}t)\theta_{m-2}^1\theta_{m-3}^2 \\ & \ldots\theta_2^{m-3}\theta_1^{m-2}(1-\theta_1)\,d\theta_1\,d\theta_2\ldots d\theta_{m-1}t^m \\ & + \frac{\alpha_m}{m!}t^m + \ldots + \frac{\alpha_2}{2!}t^2 - t + d \\ = {} & \int_0^1\ldots\int_0^1 w(\theta_2\ldots\theta_{m-1}v_1)(t-v_1)\theta_{m-2}^1\theta_{m-3}^2 \\ & \ldots\theta_2^{m-3}v_1^{m-2}\,dv_1\,d\theta_2\,d\theta_3\ldots d\theta_{m-1} \\ & + \frac{\alpha_m}{m!}t^m + \ldots + \frac{\alpha_2}{2!}t^2 - t + d \\ = {} & \int_0^{v_{m-2}}\int_0^{v_{m-3}}\ldots\int_0^{v_1}\int_0^t w(v_{m-1})(t-v_1)\,dv_1\,dv_2\ldots dv_{m-1} \\ & + \frac{\alpha_m}{m!}t^m + \ldots + \frac{\alpha_2}{2!}t^2 - t + d. \end{aligned} \tag{73}$$

Finally, the function associated with the numerator in (22) is given by

$$\widetilde{f}(t) = \int_0^1 \dots \int_0^1 w(\theta_1\theta_2\dots\theta_{m-1}t)\theta_{m-2}^1\theta_{m-3}^2 \dots \theta_2^{m-3}\theta_1^{m-1}\, d\theta_1\, d\theta_2 \dots d\theta_{m-1} t^m + \frac{(m-1)\alpha_m}{m!}t^m + \dots + \frac{\alpha_2}{2!}t^2 - t + d$$
$$= \int_0^1 \dots \int_0^t w(\theta_2\dots\theta_{m-1}v_1)v_1^{m-1}\, dv_1\, \theta_{m-2}^1\theta_{m-3}^2 \dots \theta_2^{m-3} d\theta_2 d\theta_3 \dots d\theta_{m-1} + \frac{(m-1)\alpha_m}{m!}t^m + \dots + \frac{\alpha_2}{2!}t^2 - t + d$$
$$= \int_0^{v_{m-2}}\int_0^{v_{m-3}} \dots \int_0^{v_1}\int_0^t w(v_{m-1})v_1\, dv_1\, dv_2 \dots dv_{m-1} + \frac{(m-1)\alpha_m}{m!}t^m + \dots + \frac{\alpha_2}{2!}t^2 - t + d. \tag{74}$$

We complete this study with another interesting example where we compute favorably (29) to (58).

Example 2. Consider the system of equations $F(x,y) = 0$, where $F : \mathbb{R}^2 \to \mathbb{R}^2$ and $F(x,y) = (xy-1, xy+x-2y)$. Then

$$F'(x,y) = \begin{bmatrix} y & x \\ y+1 & x-2 \end{bmatrix},$$

and

$$F'(x,y)^{-1} = \frac{1}{x+2y}\begin{bmatrix} 2-x & x \\ y+1 & -y \end{bmatrix},$$

provided that (x,y) does not belong on the straight line $x+2y=0$. The second derivative is a bilinear operator on $\mathbb{R}^2$ given by the matrix

$$F''(x,y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ \hline 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We consider the max-norm in $\mathbb{R}^2$. Moreover, in $L(\mathbb{R}^2,\mathbb{R}^2)$ we use for

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

the norm

$$\|A\| = \max\{|a_{11}| + |a_{12}|, |a_{21}| + |a_{22}|\}.$$

As in [3] we define the norm of a bilinear operator B on $\mathbb{R}^2$ by

$$\|B\| = \sup_{\|z\|=1} \max_i \sum_{j=1}^{2} \Big| \sum_{k=1}^{2} b_i^{jk} z_k \Big|,$$

where

$$z = (z_1, z_2) \quad \text{and} \quad B = \begin{bmatrix} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ \hline b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22} \end{bmatrix}.$$

Using (4), (5), (29), (58), (59), for $m = 2$ and $(x^*, y^*) = (1, 1)$, we get $\varrho = 4//3$, $\varrho_0 = .5$, $\alpha_2 = 1$, $\alpha_3 = 0$, and $\delta_0^2 = 2//3$. Since $\varrho_0 < \delta_0^2$, a remark similar to the one at the end of Example 1 can now follow.

References

[1] I. K. Argyros, *Error bounds for the modified secant method*, BIT 30 (1990), 92–100.

[2] —, *On the solution of equations with nondifferentiable and Ptak error estimates*, ibid., 752–754.

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

[4] P. Deuflhard, *A stepsize control for continuation methods and its special applications to multiple shooting techniques*, Numer. Math. 33 (1979), 115–146.

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

[6] J. Hackl, H. Wacker and W. Zulehner, *An efficient stepsize control for continuation methods*, BIT 20 (1980), 475–485.

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

[8] T. J. Ypma, *Affine invariant convergence results for Newton's method*, BIT 22 (1982), 108–118.

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

Received on 24.2.2000;
revised version on 8.5.2000 (1530)