

IOANNIS K. ARGYROS (Lawton, OK)

A CONVERGENCE ANALYSIS OF NEWTON-LIKE METHODS FOR SINGULAR EQUATIONS USING OUTER OR GENERALIZED INVERSES

Abstract. The Newton–Kantorovich approach and the majorant principle are used to provide new local and semilocal convergence results for Newton-like methods using outer or generalized inverses in a Banach space setting. Using the same conditions as before, we provide more precise information on the location of the solution and on the error bounds on the distances involved. Moreover since our Newton–Kantorovich-type hypothesis is weaker than before, we can cover cases where the original Newton–Kantorovich hypothesis is violated.

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

$$(1) \quad LF(x) = 0,$$

where F is a Fréchet-differentiable operator defined on an open convex subset D of a Banach space X with values in Banach space Y and $L \in L(X, Y)$, the space of bounded linear operators from X into Y .

A large number of problems in applied mathematics and also in engineering are solved by finding 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 equa-

2000 *Mathematics Subject Classification*: 65H10, 65J15, 65G99, 65B05, 65N30, 47H17, 49M15.

Key words and phrases: Newton-like methods, Banach space, local, semilocal convergence analysis, outer inverse, generalized inverse, Newton–Kantorovich hypothesis, Newton–Kantorovich theorem.

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

We consider Newton-like methods

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

to generate a sequence approximating x^* . Here, $F'(x_n)$ denotes the Fréchet-derivative operator F' evaluated at $x = x_n$, $A(x_n) \in L(X, Y)$ is an approximation of $F'(x_n)$ and $A(x_n)^\#$ denotes an outer inverse of $A(x_n)$, i.e., $A(x_n)^\# A(x_n) A(x_n)^\# = A(x_n)^\#$ ($n \geq 0$).

Under some assumptions, Rheinboldt [16] established a convergence theorem for (2) when $A(x_n)^\# = A(x_n)^{-1}$ ($n \geq 0$) which includes the Newton–Kantorovich theorem for the Newton method ($A(x_n) = F'(x_n)$) as a special case. A further generalization was given by Dennis [8]. Yamamoto [19] and others [2]–[5], [7] improved on the error bounds obtained. In the context of outer and generalized inverses, Ben-Israel [6], Deuffhard and Heindl [9], Häubler [12], Yamamoto [20], Nashed and Chen [14] and Argyros [2], [4] have provided Newton–Kantorovich type theorems under various conditions.

Here motivated by the elegant work of Nashed and Chen [14], in particular, we introduce a new semilocal analysis for method (2). It turns out that under the same conditions our Newton–Kantorovich type hypothesis is weaker, and the error bounds on the distances involved as well as the location of the solution are more precise.

Finally, we provide a local analysis for method (2) which compares favorably with earlier ones. In particular in the case of Newton’s method we obtain a larger convergence radius. This observation is important in numerical mathematics, since the method finds applications in projection methods and in the construction of efficient mesh independence refinement strategies [1], [5], [21].

2. Semilocal analysis of Newton-like methods. We need the following result on the convergence of majorizing sequences.

THEOREM 1. *Assume there exist nonnegative parameters $K, M, L, \ell, \mu, \eta, \delta \in [0, 1]$ such that:*

$$(3) \quad L \leq K,$$

$$(4) \quad \ell + 2\mu < 1,$$

and

$$(5) \quad h_\delta \equiv \left(K + L\delta + \frac{4M}{2-\delta} \right) \eta + \delta\ell + 2\mu \leq \delta.$$

Then the iteration $\{t_n\}$ ($n \geq 0$) given by

$$(6) \quad \begin{aligned} t_0 &= 0, \quad t_1 = \eta, \\ t_{n+2} &= t_{n+1} + \frac{K(t_{n+1} - t_n) + 2(Mt_n + \mu)}{2(1 - \ell - Lt_{n+1})} (t_{n+1} - t_n) \quad (n \geq 0) \end{aligned}$$

is nondecreasing, bounded above by t^{**} and converges to some t^* such that

$$(7) \quad 0 \leq t^* \leq \frac{2\eta}{2-\delta} \equiv t^{**}.$$

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

$$(8) \quad t_{n+2} - t_{n+1} \leq \frac{\delta}{2} (t_{n+1} - t_n) \leq \left(\frac{\delta}{2} \right)^{n+1} \eta.$$

Proof. The result clearly holds if $\eta = 0$ or $K = 0$ or $\delta = 0$. So assume $K \neq 0, \eta \neq 0$ and $\delta \neq 0$. We must show that

$$(9) \quad K(t_{i+1} - t_i) + 2(Mt_i + \mu) + \delta\ell + \delta Lt_{i+1} \leq \delta, \quad 1 - \ell - Lt_{i+1} > 0 \quad (i \geq 0).$$

Estimate (8) then follows immediately from (6) and (9). For $i = 0$, (8) and (9) hold by (3)–(6). We also get

$$0 \leq t_2 - t_1 \leq \frac{\delta}{2} (t_1 - t_0).$$

Assume (8) and (9) hold for all $i \leq n + 1$. We have

$$(10) \quad \begin{aligned} & K(t_{i+2} - t_{i+1}) + 2(Mt_{i+1} + \mu) + \delta\ell + \delta Lt_{i+2} \\ & \leq K\eta \left(\frac{\delta}{2} \right)^{i+1} + \delta\ell + 2 \left\{ M \left[t_1 + \frac{\delta}{2} (t_1 - t_0) + \left(\frac{\delta}{2} \right)^2 (t_1 - t_0) \right. \right. \\ & \quad \left. \left. + \cdots + \left(\frac{\delta}{2} \right)^i \eta \right] + \mu \right\} + \delta L \left[t_1 + \frac{\delta}{2} (t_1 - t_0) + \cdots + \left(\frac{\delta}{2} \right)^{i+1} (t_1 - t_0) \right] \\ & \leq \left\{ K \left(\frac{\delta}{2} \right)^{i+1} + \frac{2\delta L}{2-\delta} \left[1 - \left(\frac{\delta}{2} \right)^{i+2} \right] \right\} \eta + \frac{4M}{2-\delta} \left[1 - \left(\frac{\delta}{2} \right)^{i+1} \right] \eta + \delta\ell + 2\mu. \end{aligned}$$

Quantity (10) will be bounded above by δ if we show

$$(11) \quad \delta L \left[\frac{2}{2-\delta} \left(1 - \left(\frac{\delta}{2} \right)^{i+1} \right) - 1 \right] \leq K \left[1 - \left(\frac{\delta}{2} \right)^{i+1} \right]$$

and

$$(12) \quad \frac{4M}{2-\delta} \left[1 - \left(\frac{\delta}{2} \right)^{i+1} \right] \leq \frac{4M}{2-\delta}.$$

But (12) certainly holds by the choice of δ . Instead of (11) we can show

$$\left[\frac{L\delta^2}{2-\delta} - K \right] \left[1 - \left(\frac{\delta}{2} \right)^{i+1} \right] \leq 0$$

or

$$(13) \quad \frac{L\delta^2}{2-\delta} \leq K,$$

which is true by the choice of δ . Hence, estimate (8) holds for all $n \geq 0$.

Moreover we show that

$$t_i \leq t^{**} \quad (i \geq 0).$$

The above estimate holds for $i = 0, 1, 2$ by the initial conditions. Assume it holds for all $i \leq n$. It then follows from (8) that

$$\begin{aligned} t_{i+2} &\leq t_{i+1} + \frac{\delta}{2}(t_{i+1} - t_i) \leq t_i + \frac{\delta}{2}(t_i - t_{i-1}) + \frac{\delta}{2}(t_{i+1} - t_i) \\ &\leq \cdots \leq n + \frac{\delta}{2}\eta + \cdots + \left(\frac{\delta}{2} \right)^{i+1} \eta \leq \frac{[1 - (\delta/2)^{i+2}]\eta}{1 - \delta/2} < \frac{2n}{2 - \delta} = t^{**}. \end{aligned}$$

Moreover,

$$(14) \quad Lt_{i+2} < \frac{2L\eta}{2-\delta} \leq L(1+\delta)\eta \leq 1 - \ell \quad (\text{by (5)}),$$

which shows (9) for all $i \geq 0$. Hence the sequence $\{t_n\}$ ($n \geq 0$) is bounded above by t^{**} .

It also follows from (6) that $\{t_n\}$ ($n \geq 0$) is nondecreasing and as such it converges to some t^* satisfying (7).

That completes the proof of Theorem 1. ■

REMARK 1. It can easily be seen from the proof of Theorem 1 that (5) can be replaced by the weaker conditions:

$$(5)' \quad h_\delta \leq \delta, \quad \frac{2L\eta}{2-\delta} \leq 1 \quad \text{and} \quad \frac{L\delta^2}{2-\delta} \leq K$$

provided that $\delta \in [0, 2)$.

For comparison purposes we state the elegant semilocal convergence theorem for Newton-like methods using outer inverses (see Theorem 3.1 in [14, p. 241]):

THEOREM 2. *Let $F: D \subseteq X \rightarrow Y$ 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 , $x_0 \in D_0$, a bounded outer inverse $A^\#$ of $A(x_0)$, and parameters $\eta > 0$, $K > 0$, $M \geq 0$, $L \geq 0$, $\mu \geq 0$, $\ell \geq 0$ such that*

$$(15) \quad \|A^\#F(x_0)\| \leq \eta,$$

$$(16) \quad \|A^\#[F'(x) - F'(y)]\| \leq K\|x - y\|,$$

$$(17) \quad \|A^\#[F'(x) - A(x)]\| \leq M\|x - x_0\| + \mu,$$

$$(18) \quad \|A^\#[A(x) - A(x_0)]\| \leq L\|x - x_0\| + \ell$$

for all $x, y \in D_0$;

(b)

$$(19) \quad b \equiv \mu + \ell < 1,$$

$$(20) \quad h \equiv 2\sigma\eta \leq (1 - b)^2,$$

where

$$(21) \quad \sigma := \max(K, M + L),$$

$$(22) \quad \bar{U}(x_0, s^*) \subseteq D_0,$$

$$(23) \quad s^* = \frac{1 - b - \sqrt{(1 - b)^2 - h}}{\sigma}.$$

Then

- (i) *the sequence $\{x_n\}$ ($n \geq 0$) generated by the Newton-like method (2) with*

$$(24) \quad A(x_n)^\# = [I + A^\#(A(x_n) - A(x_0))]^{-1}A^\#$$

is well defined, remains in $U(x_0, s^)$ for all $n \geq 0$ and converges to a solution x^* of $A^\#F(x) = 0$;*

- (ii) *the solution x^* is unique in $\tilde{U}(x_0, s^*) \cap \{R(A^\# + x_0)\}$, where $R(A^\#)$ denotes the range of $A^\#$,*

$$(25) \quad \tilde{U}(x_0, s^*) = \begin{cases} \bar{U}(x_0, s^*) \cap D_0 & \text{if } h = (1 - b)^2, \\ U(x_0, s^{**}) \cap D_0 & \text{if } h < (1 - b)^2, \end{cases}$$

*s^{**} is the larger zero of the function f given by*

$$(26) \quad f(s) = \frac{\sigma}{2}s^2 - (1 - b)s + \eta$$

and

$$(27) \quad R(A^\#) + x_0 = \{x + x_0 : x \in R(A^\#)\}.$$

Moreover, define a function g by

$$(28) \quad g(s) = 1 - Ls - \ell$$

and a sequence $\{s_n\}$ ($n \geq 0$) by

$$(29) \quad s_0 = 0, \quad s_1 = \eta, \quad s_{n+2} = s_{n+1} + \frac{f(s_{n+1})}{g(s_{n+1})} \quad (n \geq 0).$$

Then the following error bounds hold for all $n \geq 0$:

$$(30) \quad \|x_{n+1} - x_n\| \leq s_{n+1} - s_n,$$

$$(31) \quad \|x_n - x^*\| \leq s^* - s_n.$$

We present the main semilocal result on the convergence of Newton-like methods.

THEOREM 3. *Let $D_0, D, X, Y, F, A^\#, A(x_0), A(x)$ be as in Theorem 2. Assume:*

- (a) *there exist nonnegative parameters η, K, M, L, μ, ℓ and δ satisfying (3)–(5) or (3)–(5)';*
- (b) *conditions (15)–(18) hold for all $x, y \in D_0$ and*

$$(32) \quad \bar{U}(x_0, t^*) \subseteq D_0,$$

where t^* is defined in Theorem 1.

Then the sequence $\{x_n\}$ ($n \geq 0$) generated by the Newton-like method (2) is well defined, remains in $U(x_0, t^*)$ for all $n \geq 0$ and converges to a unique solution x^* of the equation $A^\#F(x) = 0$ in $\bar{U}(x_0, t^*) \cap D_0$. Moreover, the following error bounds hold for all $n \geq 0$:

$$(33) \quad \|x_{n+1} - x_n\| \leq t_{n+1} - t_n,$$

$$(34) \quad \|x_n - x^*\| \leq t^* - t_n,$$

where the sequence $\{t_n\}$ ($n \geq 0$) is given by (6).

Proof. Simply replace $s^*, \{s_n\}$ ($n \geq 0$) by $t^*, \{t_n\}$ ($n \geq 0$) respectively in the proof of Theorem 3.1 in [14, p. 241] and use Theorem 1 which establishes the convergence of $\{t_n\}$ to t^* .

That completes the proof of Theorem 3. ■

REMARK 2. The point t^* can be replaced by $2\eta/(2 - \delta)$ in (32).

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

THEOREM 4. *Under the hypotheses of Theorems 2 and 3 the following error bounds hold:*

$$(35) \quad t_{n+1} \leq s_{n+1} \quad (n \geq 1),$$

$$(36) \quad t_{n+1} - t_n \leq s_{n+1} - s_n \quad (n \geq 1),$$

$$(37) \quad t^* - t_n \leq s^* - s_n \quad (n \geq 0),$$

$$(38) \quad t^* \leq s^*.$$

Moreover strict inequality holds in (35) and (36) if $K < M + L$.

Proof. We use induction to first show (35) and (36). For $n = 0$ in (6) and (29) we obtain

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

and

$$t_2 \leq s_2.$$

Assume that

$$(39) \quad t_{i+1} \leq s_{i+1}, \quad t_{i+1} - t_i \leq s_{i+1} - s_i.$$

Using (6), (29) and (39) we obtain

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

which shows (35) and (36) for all $n \geq 1$.

Let $j \geq 0$. Then

$$\begin{aligned}
 (40) \quad t_{i+j} - t_i &\leq (t_{i+j} - t_{i+j-1}) + (t_{i+j-1} - t_{i+j-2}) + \cdots + (t_{i+1} - t_i) \\
 &\leq (s_{i+j} - s_{i+j-1}) + (s_{i+j-1} - s_{i+j-2}) + \cdots + (s_{i+1} - s_i) \\
 &\leq s_{i+1} - s_i.
 \end{aligned}$$

By letting $j \rightarrow \infty$ in (40) we obtain (37).

Finally, (37) implies (38) (since $t_1 = s_1 = 0$). It can easily be seen from (6) and (29) that strict inequality holds in (35) and (36) if $K < M + L$.

That completes the proof of Theorem 4. ■

REMARK 3. Due to (38) our Theorem 3 provides a more precise information on the location of the solution x^* than Theorem 2. Note also that $t^* \in [\eta, 2\eta]$, and under the hypotheses of Theorems 2 and 3, $t^* \in [\eta, s^*]$, where $s^* \leq 2\eta$ and is given by (23).

REMARK 4. Let us compare condition (5) with condition (20). For simplicity let us consider only Newton's method. That is, as in Corollary 3.1 in [14] we set $A(x) = F'(x)$ and choose $M = \mu = \ell = 0$, $K = L$ to obtain, using (21),

$$(41) \quad h = 2K\eta \leq 1,$$

which is the famous Newton–Kantorovich hypothesis for the convergence of Newton's method [13]. However in general $L \leq K$. Hence, our condition becomes

$$(42) \quad h_1 = (K + L)\eta \leq 1.$$

Note that

$$(43) \quad h \leq 1 \Rightarrow h_1 \leq 1$$

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

REMARK 5. It is clear from Theorems 2 and 3 that if we replace $U(x_0, t^*) \subseteq D$ by $\bar{U}(x_1, t^* - \eta) \subseteq D$, using the assumptions of Theorem 3 instead of Theorem 2 and proceeding along the well established lines of [18] and [19], we can first refine the error bounds obtained here which as we already showed in Theorem 4 can further improve the corresponding error bounds in [14]. However, we leave the details to the motivated reader.

REMARK 6. As in [14] (see Theorem 2 above), suppose that

$$(44) \quad [I + (A(x) - A(x_0))A^+]^{-1}A(x) \text{ maps } N(A(x_0)) \text{ into } R(A(x_0))$$

whenever $I + (A(x) - A(x_0))A^+$ is invertible for some $x \in D$, where $N(A)$ denotes the null space of A and A^+ the generalized inverse of $A(x_0)$. Then by Lemma 2.4 in [14], $A(x_n) = [I + A^+(x_0)(A(x_n) - A(x_0))]^{-1}A^+(x_0)$ is a generalized inverse. Hence by Lemma 2.4 of [14] and Theorem 3 we establish a semilocal convergence theorem for Newton-like methods using generalized inverses. In the finite-dimensional case (x, y both finite), condition (44) can be replaced by

$$(45) \quad \text{rank}(A(x)) \leq \text{rank}(A(x_0)) \quad (x \in D).$$

We now complete this section with three simple numerical examples. In the first and third example we show that hypothesis (41) fails whereas (42) holds. In the second example used also in [15] we compare estimates (30), (31) and (33), (34), respectively. Let $A(x) = F'(x)$, $A(x)^{-1} = A^\#(x)$ ($x \in D$) in the next three examples.

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

$$(46) \quad F(x) = \frac{1}{6}x^3 - \left(\frac{2^{3/2}}{6} + .23\right).$$

Using (15)–(18), (40) and (41) and setting $\mu = \ell = M = 0$, we obtain

$$(47) \quad \eta = .23, \quad K = 2.4142136, \quad L = 1.914213562, \\ h = 2K\eta = 1.1105383 > 1$$

and from (5) for $\delta = 1$

$$(48) \quad (K + L_0)\eta = .995538247 < 1.$$

That is, the Newton–Kantorovich hypothesis (41) is violated. However since (48) holds, Theorem 3 guarantees the convergence of Newton’s method to x^* .

EXAMPLE 2. Let $X = Y = \mathbb{R}$, $x_0 = 1.3$,

$$D = [x_0 - 2\|F'(x_0)^{-1}F(x_0)\|, x_0 + 2\|F'(x_0)^{-1}F(x_0)\|]$$

and define a function F on D by

$$(49) \quad F(x) = \frac{1}{3}(x^3 - 1).$$

As in Example 1 and setting $\mu = \ell = M = 0$, we obtain

$$\eta = .236094674, \quad K = 2.097265501, \quad L = 1.817863519, \\ h = 2K\eta = .990306428 < 1, \quad (K + L_0)\eta = .92434111 < 1 \quad (\text{for } \delta = 1), \\ t^* = .369677842, \quad s^* = .429866445.$$

That is, we provide a better information on the location of the solution x^* since

$$(50) \quad \bar{U}(x_0, t^*) \subset \bar{U}(x_0, s^*).$$

We can tabulate the following results:

Comparison table

x_n	Estimates (34)	Estimates (33)	Estimates (31)	Estimates (30)
$x_1 = 1.0639053254$.236094674	.133583172	.236094674	.193771771
$x_2 = 1.0037617275$.102400629	.031182539	.115780708	.0779910691
$x_3 = 1.0000140800$.028585756	.002596783	.053649732	.024342893
$x_4 = 1.0000000002$.002575575	.000021208	.020186667	.004156226
$x_5 = 1$.000021207	.000000001	.003987206	.00016902
$x_6 = 1$.000000001	0	.000166761	.000002259

EXAMPLE 3. Let $X = Y = \mathbb{R}$, $x_0 = 1$ and define a function F by

$$F(x) = x^3 - a \quad \text{for all } a \in [0, 1/2), \quad x \in [a, 2 - a].$$

Using (15), (16) and (18) we get

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

The Newton–Kantorovich hypothesis (41) does not hold since

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

However (42) holds for all $a \in [(5 - \sqrt{13})/3, 1/2)$ since

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

and so the Newton’s method (2) converges to the solution $x^* = \sqrt[3]{a}$ of $F(x) = 0$.

3. Local analysis of Newton’s method. We can show the following local result for Newton-like methods:

THEOREM 5. *Let $F: D \subseteq X \rightarrow Y$ 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 (1), a bounded outer inverse $A_1^\#$ of $A(x^*)$ and nonnegative parameters \bar{K} , \bar{L} , \bar{M} , $\bar{\mu}$, $\bar{\ell}$ such that*

$$(51) \quad \|A_1^\#[F'(x) - F'(y)]\| \leq \bar{K}\|x - y\|,$$

$$(52) \quad \|A_1^\#[F'(x) - A(x)]\| \leq \bar{M}\|x - x^*\| + \bar{\mu},$$

$$(53) \quad \|A_1^\#[A(x) - A(x^*)]\| \leq \bar{L}\|x - x^*\| + \bar{\ell}$$

for all $x, y \in D$;

- (b) *the equation*

$$(54) \quad \left(\frac{\bar{K}}{2} + \bar{M} + \bar{L}\right)r + \bar{\mu} + \bar{\ell} - 1 = 0$$

has a minimal nonnegative zero r^ satisfying*

$$(55) \quad \bar{L}r + \bar{\ell} < 1,$$

$$(56) \quad U(x^*, r^*) \subseteq D.$$

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

$$\begin{aligned}
 (57) \quad & \|x^* - x_{n+1}\| \\
 & \leq \frac{1}{1 - \bar{L}\|x^* - x_n\| - \bar{\ell}} \left[\frac{\bar{K}}{2} \|x^* - x_n\| + (\bar{M}\|x^* - x_n\| + \bar{\mu}) \right] \|x^* - x_n\| \\
 & < \frac{(\bar{K}/2 + \bar{M})r^* + \bar{\mu}}{1 - \bar{L}r^* - \bar{\ell}} \|x^* - x_n\|.
 \end{aligned}$$

Proof. This follows exactly as in the last part of the proof of Theorem 3.1 (uniqueness). Indeed using induction on $k \geq 0$ and (51)–(56) we obtain

$$\begin{aligned}
 \|x^* - x_{k+1}\| &= \|x^* - x_k + A(x_n)^\# F(x_k) - A(x_k)^\# F(x^*)\| \\
 &= \|A(x_k)^\# (A(x_k)(x^* - x_k) + F(x_k) - F(x^*))\| \\
 &\leq \|A(x_n)^\# A(x^*)\| \left\{ \int_0^1 \|A_1^\# [F'(x_k + t(x^* - x_k)) - F'(x_k)]\| dt \right. \\
 &\quad \left. + \|A_1^\# (F'(x_k) - A(x_k))\| \right\} \|x^* - x_k\| \\
 &\leq \frac{1}{1 - \bar{L}\|x^* - x_k\| - \bar{\ell}} \left[\frac{\bar{K}}{2} \|x^* - x_k\| + (\bar{M}\|x^* - x_k\| + \bar{\mu}) \right] \|x^* - x_k\| \\
 &< \frac{(\bar{K}/2 + \bar{M})r^* + \bar{\mu}}{1 - \bar{L}r^* - \bar{\ell}} \|x^* - x_k\|,
 \end{aligned}$$

which shows (57) and $x_n \in U(x^*, r^*)$ ($n \geq 0$) (since $x_0 \in U(x^*, r^*)$). Moreover by the choice of r^* there exists $\alpha \in [0, 1)$ such that

$$\|x^* - x_{n+1}\| \leq \alpha \|x^* - x_n\| \leq \alpha^{n+1} \|x^* - x_0\|.$$

Hence, we deduce $\lim_{n \rightarrow \infty} x_n = x^*$.

That completes the proof of Theorem 5. ■

REMARK 7. Note that local results were not given in [14].

REMARK 8. Let $A(x) = F'(x)$, $A(x)^{-1} = A^\#(x)$ (if $A(x)^{-1}$ exists) ($x \in D$) and set $\bar{M} = \bar{\mu} = \bar{\ell} = 0$. Then the radius of convergence is

$$(58) \quad r^* = \frac{2}{2\bar{L} + \bar{K}}.$$

To compare our results with earlier results on local convergence and convergence radii, let us consider Rheinboldt's result in [17], where only (16) is used. In this case

$$(59) \quad r_R^* = \frac{2}{3\bar{K}} \leq r^*.$$

Note that if $\bar{L} < \bar{K}$ then strictly inequality holds in (59). Hence under the hypotheses of Theorem 5 we obtain a larger convergence radius than

before as well as finer error bounds. This observation is important in computational mathematics and finds applications in projection methods and in the construction of efficient/cheap mesh independence refinement strategies [1], [4], [21].

REMARK 9. As noted in [1], [5], [8], [9], [12], [25] 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 $A(x) = F'(x)$, $A(x)^{-1} = A^\#(x)$ ($x \in D$).

REMARK 10. 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])

$$(60) \quad F'(x) = T(F(x)),$$

where $T: Y \rightarrow X$ 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 $X = Y = \mathbb{R}$, $D = U(0, 1)$ and define a function F on D by

$$(61) \quad F(x) = e^x - 1.$$

Then it can easily be seen that we can set $T(x) = x + 1$ in (60). Since $F'(x^*) = 1$, we get $\|F'(x) - F'(y)\| \leq e\|x - y\|$. Hence we set $\bar{K} = e$. Moreover since $x^* = 0$ we obtain

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

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

$$\|F'(x) - F'(x^*)\| \leq (e - 1)\|x - x^*\|.$$

That is, $\bar{L} = e - 1$. Using (59) and (58) we obtain respectively:

$$r_R^* = .245252961, \quad r^* = .254028662.$$

That is, our convergence radius r^* is larger than the corresponding one r_R^* due to Rheinboldt [17], which allows a wider choice of initial guesses x_0 . Note also that $\bar{L} < \bar{K}$. Hence our error bounds on $\|x_n - x^*\|$ ($n \geq 0$) are also finer.

References

- [1] F. L. Allgower, K. Böhmer, F. A. Potra and W. C. Rheinboldt, *A mesh independence principle for operator equations and their discretizations*, SIAM J. Numer. Anal. 23 (1986), 160–169.
- [2] I. K. Argyros, *Convergence domains for some iterative processes in Banach spaces using outer and generalized inverses*, J. Comput. Anal. Appl. 1 (1999), 87–104.
- [3] —, *Local convergence of inexact Newton-like iterative methods and applications*, Comput. Math. Appl. 39 (2000), 69–75.
- [4] —, *Advances in the Efficiency of Computational Methods and Applications*, World Sci., River Edge, NJ, 2000.
- [5] I. K. Argyros and F. Szidarovszky, *The Theory and Applications of Iteration Methods*, CRC Press, Boca Raton, FL, 1993.
- [6] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.
- [7] X. Chen and T. Yamamoto, *Convergence domains of certain iterative methods for solving nonlinear equations*, Numer. Funct. Anal. Optim. 10 (1989), 37–48.
- [8] J. E. Dennis, Jr., *On Newton-like methods*, Numer. Math. 11 (1968), 324–330.
- [9] P. Deuffhard and G. Heindl, *Convergence theorems for Newton's method and extensions to related methods*, SIAM J. Numer. Anal. 16 (1979), 1–10.
- [10] J. M. Gutiérrez, *A new semilocal convergence theorem for Newton's method*, J. Comput. Appl. Math. 79 (1997), 131–145.
- [11] J. M. Gutiérrez, M. A. Hernandez and M. A. Salanova, *Accessibility of solutions by Newton's method*, Internat. J. Comput. Math. 57 (1995), 239–247.
- [12] W. M. Häubler, *A Kantorovich-type convergence analysis for the Gauss–Newton methods*, Numer. Math. 48 (1986), 119–125.
- [13] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1964.
- [14] M. Z. Nashed and X. Chen, *Convergence of Newton-like methods for singular operator equations using outer inverses*, Numer. Math. 66 (1993), 235–257.
- [15] F. A. Potra and V. Ptak, *Sharp error bounds for Newton's process*, *ibid.* 34 (1980), 67–72.
- [16] W. C. Rheinboldt, *A unified convergence theory for a class of iterative processes*, SIAM J. Numer. Anal. 5 (1968), 42–63.
- [17] —, *An adaptive continuation process for solving systems of nonlinear equations*, in: Banach Center Publ. 3, PWN–Polish Sci. Publ., 1977, 129–142.
- [18] T. Yamamoto, *A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions*, Numer. Math. 49 (1987), 203–230.
- [19] —, *A convergence theorem for Newton-like methods in Banach spaces*, *ibid.* 51 (1987), 545–557.
- [20] —, *Uniqueness of the solution in a Kantorovich-type theorem of Häubler for the Gauss–Newton method*, Japan J. Appl. Math. 6 (1989), 77–81.
- [21] T. J. Ypma, *Local convergence of inexact Newton's method*, SIAM J. Numer. Anal. 21 (1984), 583–590.

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

Received on 6.5.2003;
 revised version on 23.9.2004

(1688)