Int. J. Appl. Math. Sci. Vol. **1**(2004), pp. 13-26

A Convergence Analysis and Applications of Newton-Like Methods Under Generalized Chen–Yamamoto-Type Assumptions

Ioannis K. Argyros
Cameron University
Department of Mathematical Sciences
Lawton, OK 73505, USA
E-mail address: ioannisa@cameron.edu

Abstract In this study we introduce more general Chen–Yamamoto-type conditions to generate a Newton-like method which converges to a locally unique solution of a nonlinear equation in a Banach space containing a non-differentiable term. Using new and precise majorizing sequences we provide local and semilocal results, under different sufficient convergence conditions than before. In some cases our semilocal convergence conditions are weaker than before and the error bounds on the distances more precise, whereas in the local case we can obtain a larger convergence radius. Finally in the general case our results can be reduced to earlier ones.

AMS (MOS) Subject Classification Codes: 65H10, 65B05, 65G99, 65J15, 65N30, 47H17, 49M15, CR:1.5.

Key Words: Newton-like method, Banach space, Fréchet-derivative, majorizing sequence, Chen–Yamamoto conditions, radius of convergence, Newton–Kantorovich hypothesis.

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

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

where F, G are operators defined on a closed ball $\overline{U}(z,R)$ $(R > 0)$, which is a subset of a Banach space X with values in a Banach space Y. F is Fréchet-differentiable on $\overline{U}(z,R)$, while the differentiability of operator G is not assumed.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. 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 use the Newton-like method

$$y_0 \in U(z,R),\ y_{n+1} = y_n - A(y_n)^{-1}[F(y_n) + G(y_n)] \quad (n \geq 0) \tag{2}$$

to generate a sequence converging to x^*. Here, $A(y) \in L(X,Y)$, the space of bounded linear operators from X into Y. Sufficient convergence conditions as well as convergence domains have been given in [5] under assumptions that generalize earlier ones [6]–[16]. We provide local and semilocal convergence results for method (2) under more general conditions. It turns out that this way we have a greater flexibility in the choice of "majorizing" functions/sequences, which in turn can lead to more precise error bounds and weaker sufficient convergence conditions. Finally in the local case we provide a larger convergence radius than before. This observation is important in computational mathematics [1]–[4], [7], [15].

2. Semilocal Convergence of Newton-Like Methods

We assume from now on $A(z)^{-1}$ for some fixed $z \in X$, $R > 0$ exists, and for any $x, y \in \overline{U}(z,r) = \{v \in X \mid \|v - z\| \leq r\} \subseteq \overline{U}(z,R)$:

$$\|A(z)^{-1}[A(x) - A(z)]\| \leq w_0(\|x - z\|) + b, \tag{3}$$

$$\|A(z)^{-1}[F'(x + t(y - x)) - A(x)]\| \leq w(\|x - z\| + t\|y - x\|) - w_1(\|x - z\|) + c, \qquad t \in [0,1], \tag{4}$$

$$\|A(z)^{-1}[G(x) - G(y)]\| \leq w_2(r)\|x - y\|, \tag{5}$$

where, $w(r + t) - w_1(r)$, $t \geq 0$, $w_0(r)$ and $w_2(r)$ are monotonically increasing functions on $[0,R]$ with $w(0) = w_0(0) = w_2(0) = w_1(0) = 0$, and constants b, c are non-negative parameters.

With the above choices, we show a result concerning the convergence of majorizing sequences:

Theorem 1. *Assume*:
there exist $\eta \geq 0$, $b \geq 0$, $c \geq 0$, $\delta \in [0,2)$, $r_0 \in [0,r]$, $r \in [0,R]$ *such that*:

$$h_\delta = 2\int_0^1 w(r_0 + \theta\eta)d\theta - 2w_1(r_0) + 2w_2(r_0 + \eta) + 2c + \delta b + \delta w_0(r_0 + \eta) \leq \delta, \tag{6}$$

$$w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] + b < 1, \tag{7}$$

$$\frac{2\eta}{2-\delta} + r_0 \leq r, \tag{8}$$

and for all $n \geq 0$:

$$\begin{aligned} 2\int_0^1 w &\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right) + r_0 + \theta\left(\frac{\delta}{2}\right)^{n+1}\eta\right] d\theta \\ &- 2w_1\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] \\ &+ 2w_2\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] \\ &+ \delta w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] + 2c + \delta b \leq \delta. \end{aligned} \tag{9}$$

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

$$t_0 = r_0, \quad t_1 = r_0 + \eta, \quad t_{n+2} = t_{n+1} + \frac{\int_0^1 \{w[t_n + \theta(t_{n+1} - t_n)]d\theta - w_1(t_n) + c\}(t_{n+1} - t_n) + \int_{t_n}^{t_{n+1}} w_2(\theta)d\theta}{1 - b - w_0(t_{n+1})} \tag{10}$$

is non-decreasing, bounded above by

$$t^{**} = \frac{2\eta}{2 - \delta} + r_0, \tag{11}$$

and converges to some t^* *such that*

$$0 \leq t^* \leq t^{**}. \tag{12}$$

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

$$0 \leq t_{n+2} - t_{n+1} \leq \frac{\delta}{2}(t_{n+1} - t_n) \leq \left(\frac{\delta}{2}\right)^{n+1} \eta. \tag{13}$$

Proof. We must show:

$$2\int_0^1 w[t_k + \theta(t_{k+1} - t_k)]d\theta - 2w_1(t_k) + 2w_2(t_{k+1}) + 2c + \delta b + \delta w_0(t_{k+1}) \leq \delta, \tag{14}$$

$$0 \leq t_{k+1} - t_k \tag{15}$$

and

$$w_0(t_{k+1}) + b < 1 \tag{16}$$

for all $k \geq 0$.

Estimate (13) can then follow immediately from (14), (15) and (16). Using induction on the integer k, we get for $k = 0$

$$2\int_0^1 w[t_0 + \theta(t_1 - t_0)]d\theta - 2w_1(t_1) + 2w_2(t_1) + 2c + \delta b + \delta w_0(t_1) \leq \delta,$$

$$w_0(t_1) + b = w_0(r_0 + \eta) + b < 1$$

by the initial conditions. But (10) gives

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

Assume (14)–(16) hold for all $k \leq n + 1$. Using (6)–(10) we obtain in turn

$$\begin{aligned}
&2\int_0^1 w[t_{k+1} + \theta(t_{k+2} - t_{k+1})]d\theta - 2w_1(t_{k+1}) + 2w_2(t_{k+1}) + 2c + \delta b + \delta w_0(t_{k+1}) \\
&\leq 2\int_0^1 w\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right) + r_0 + \theta\left(\frac{\delta}{2}\right)^{k+1}\eta\right]d\theta \\
&\quad - 2w_1\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right) + r_0\right] + 2w_2\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right) + r_0\right] \\
&\quad + 2c + \delta b + \delta w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right) + r_0\right] + 2c + \delta b \\
&\leq \delta \quad \text{(by (9))}.
\end{aligned}$$

Moreover, we must show
$$t_k \leq t^{**} \tag{17}$$
But we have:
$$t_0 = r_0 \leq t^{**}, \quad t_1 = r_0 + \eta \leq t^{**} \quad \text{and} \quad t_2 \leq r_0 + \eta + \frac{\delta}{2}\eta = r_0 + \frac{2+\delta}{2}\eta \leq t^{**}.$$
Assume (17) holds for all $k \leq n+1$. It follows from (10), (14)–(16):
$$\begin{aligned} t_{k+2} &\leq t_{k+1} + \frac{\delta}{2}(t_{k+1} - t_k) \leq t_k + \frac{\delta}{2}(t_k - t_{k-1}) + \frac{\delta}{2}(t_{k+1} - t_k) \\ &\leq \cdots \leq r_0 + \eta + \frac{\delta}{2}\eta + \left(\frac{\delta}{2}\right)^2 \eta + \cdots + \left(\frac{\delta}{2}\right)^{k+1} \eta \\ &\leq r_0 + \frac{1 - \left(\frac{\delta}{2}\right)^{k+2}}{1 - \frac{\delta}{2}}\eta \leq r_0 + \frac{2\eta}{2-\delta} = t^{**}. \end{aligned}$$
Hence, sequence $\{t_n\}$ $(n \geq 0)$ is bounded above by t^{**}. Furthermore, sequence $\{t_n\}$ $(n \geq 0)$ is monotonically increasing by (24) and as such it converges to some t^* satisfying (12).

That completes the proof of Theorem 1. □

We provide the main result on the semilocal convergence of Newton-like methods using majorizing sequence (10).

Theorem 2. *Assume*:
hypotheses of Theorem 1 hold, and there exists
$$y_0 \in \overline{U}(z, r), \quad r \in [0, R] \tag{18}$$
such that:
$$\|A(y_0)^{-1}[F(y_0) + G(y_0)]\| \leq \eta. \tag{19}$$
Then, sequence $\{y_n\}$ $(n \geq 0)$ generated by Newton-like method (2) is well defined, remains in $\overline{U}(z, t^*)$ for all $n \geq 0$, and converges to a solution x^* of equation $F(x) + G(x) = 0$.

Moreover the following error bounds hold for all $n \geq 0$:
$$\|y_{n+1} - y_n\| \leq t_{n+1} - t_n \tag{20}$$
and
$$\|y_n - x^*\| \leq t^* - t_n. \tag{21}$$
Furthermore the solution x^* is unique in $\overline{U}(z, t^*)$ if
$$\int_0^1 [w(t^* + 2tt^*) - w_1(t^*)]dt + w_2(3t^*) + w_0(t^*) + b + c < 1, \tag{22}$$
and in $\overline{U}(z, R_0)$ for $R_0 > t^*$ if
$$\int_0^1 [w(t^* + t(t^* + R_0)) - w_1(t^*)]dt + w_2(2t^* + R_0) + w_0(t^*) + b + c < 1, \tag{23}$$
and
$$R_0 \leq r. \tag{24}$$

Proof. We must show estimates (20) and (21). For $n = 0$, (20) is obvious, since
$$\|y_1 - y_0\| = \|A(y_0)^{-1}[F(y_0) + G(y_0)]\| \leq \eta = t_1 - t_0.$$

Suppose (20) holds for all $n = 0, 1, \dots, k+1$; this implies in particular

$$\begin{aligned} \|y_{k+1} - z\| &\le \|y_{k+1} - y_k\| + \|y_k - y_{k-1}\| + \cdots + \|y_1 - y_0\| + \|y_0 - z\| \\ &\le (t_{k+1} - t_k) + (t_k - t_{k-1}) + \cdots + (t_1 - t_0) + r_0 = t_{k+1} \le t^*. \end{aligned}$$

That is $y_{k+1} \in \overline{U}(z, t^*)$. We show (20) holds for $n = k+2$. Using (3) and (7) we get

$$\begin{aligned} \|A(z)^{-1}[A(y_{k+1}) - A(z)]\| &\le w_0(\|y_{k+1} - z\|) + b \\ &\le w_0(t_{k+1}) + b \\ &\le w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right) + r_0\right] + b < 1. \end{aligned} \tag{25}$$

It follows from (25) and the Banach Lemma on invertible operators [8] that $A(y_{k+1})^{-1}$ exists and

$$\begin{aligned} \|A(y_{k+1})^{-1}A(z)\| &\le [1 - b - w_0(\|y_{k+1} - z\|)]^{-1} \le [1 - b - w_0(t_{k+1})]^{-1} \\ &\le [1 - b - w_0(R)]^{-1} = b_0. \end{aligned} \tag{26}$$

Using (2), (4), (5), (10) we obtain in turn

$$\begin{aligned} &\|y_{k+2} - y_{k+1}\| = \|A(y_{k+1})^{-1}[F(y_{k+1}) + G(y_{k+1})]\| \\ &\le \|A(y_{k+1})^{-1}A(z)\|\,\|A(z)^{-1}\{F(y_{k+1}) + G(y_{k+1}) \\ &- A(y_k)(y_{k+1} - y_k) - F(y_k) - G(y_k)\}\| \\ &\le \frac{\{\int_0^1 \|A(z)^{-1}[F'(y_k + t(y_{k+1} - y_k)) - A(y_k)]\|\|y_{k+1} - y_k\|dt + \|A(z)^{-1}[G(y_{k+1} - G(y_k)]\|}{1 - b - w_0(\|y_{k+1} - z\|)} \\ &\le \frac{\int_0^1 [w(\|y_k - z\| + t\|y_{k+1} - y_k\|) - w_1(\|y_k - z\|)]\|y_{k+1} - y_k\|dt + c\|y_{k+1} - y_k\| + \int_{t_k}^{t_{k+1}} w_2(t)dt}{1 - b - w_0(t_{k+1} - t_0)} \\ &\le \frac{\int_0^1 \{w[(t_k) + t(t_{k+1} - t_k)] - w_1(t_k)\}(t_{k+1} - t_k)dt + c(t_{k+1} - t_k) + \int_{t_k}^{t_{k+1}} w_2(t)dt}{1 - b - w_0(t_{k+1})} \\ &= t_{k+2} - t_{k+1}, \end{aligned} \tag{27}$$

which shows (20) for all $n \ge 0$.

Note also that

$$\|y_{k+2} - z\| \le \|y_{k+2} - y_{k+1}\| + \|y_{k+1} - z\| \le t_{k+2} - t_{k+1} + t_{k+1} = t_{k+2} \le t^*.$$

That is, $y_{k+2} \in \overline{U}(z, t^*)$.

It follows from (27) that $\{y_n\}$ $(n \ge 0)$ is a Cauchy sequence in a Banach space X, and as such it converges to some $x^* \in \overline{U}(z, t^*)$ (since $\overline{U}(z, t^*)$ is a closed set). As in (27) we have

$$\|y_{k+2} - y_{k+1}\| \le b_0\|A(z)^{-1}[F(y_{k+1}) + G(y_{k+1})]\| \le b_2\|y_{k+1} - y_k\| \tag{28}$$

where,

$$b_2 = b_0 b_1 \tag{29}$$

and

$$b_1 = \int_w^1 (R + t\eta)dt - w_1(R) + w_2(R) + c \tag{30}$$

By letting $k \to \infty$ in (28), using (26), and the continuity of operators F, G we get

$$b_0\|A(z)^{-1}(F(x^*) + G(x^*))\| = 0,$$

from which we obtain $F(x^*)+G(x^*)=0$ (since $b_0>0$). Estimate (21) follows from (20) by using standard majorization techniques [4], [8].

To show uniqueness in $\overline{U}(z,t^*)$, let y^* be a solution of equation (1) in $\overline{U}(z,t^*)$. Then as in (27) we get in turn

$$\begin{aligned}\|y^*-y_{k+1}\| &\le \|A(y_k)^{-1}A(z)\|\Big\{\int_0^1 [\|A(z)^{-1}[F'(y_k+t(y^*-y_k)-A(y_k)]\|dt \\ &\quad + \|A(z)^{-1}(G(y_k)-G(y^*))\|\Big\} \\ &\le \frac{\int_0^1[w(\|y_k-z\|+t\|y^*-y_k\|)-w_1(\|y_k-z\|)]\|y^*-y_k\|dt + c\|y^*-y_k\| + \int_{\|y_k-z\|}^{\|y_k-z\|+\|y^*-y_k\|} w_2(t)dt}{1-b-w_0(t^*)} \\ &\le \frac{\int_0^1[w((1+2t)t^*)-w_1(t^*)]dt + c + w_2(3t^*)}{1-b-w_0(t^*)}\|y^*-y_k\| \\ &< \|y^*-y_k\| \to 0 \ \text{ as } k\to\infty \text{ (by (22))}. \end{aligned} \tag{31}$$

That is $y^*=x^*$. If $y^*\in\overline{U}(z,R_0)$ then as in (31) we obtain

$$\begin{aligned}\|y^*-y_{k+1}\| &\le \frac{\int_0^1[w(t^*+t(t^*+R_0))-w_1(t^*)]dt + c + w_2(2t^*+R_0)}{1-b-w_0(t^*)}\|y^*-y_k\| \\ &< \|y^*-y_k\| \to 0 \ \text{ as } k\to\infty \text{ (by (23))}. \end{aligned} \tag{32}$$

Hence, again we get $y^*=x^*$.

That completes the proof of Theorem 2. □

Remark 1. Condition (9) (similarly for (7)) can be replaced by the stronger but easier to check

$$\begin{aligned} 2\int_0^1 w\left[\frac{2\eta}{2-\delta}+\frac{\theta\delta}{2}\eta+r_0\right]d\theta - 2w_1\left(\frac{2\eta}{2-\delta}+r_0\right)+2w_2\left(\frac{2\eta}{2-\delta}+r_0\right) \\ + \delta w_0\left(\frac{2\eta}{2-\delta}+r_0\right)+2c+\delta b \le \delta. \end{aligned} \tag{33}$$

Note that conditions of the form (6)–(9) are Newton–Kantorovich-type hypotheses (see also (71)) which are always present in the study of sufficient convergence conditions of Newton-like methods [2], [4], [5], [6], [11]–[16].

In order for us to compare with earlier results as in [5, p. 40], we assume that $A(z)^{-1}$ exists and for any $x,y\in\overline{U}(z,r)\subseteq\overline{U}(z,R)$, $0<\|A(z)^{-1}[F(z)+G(z)]\|\le\overline{\eta}$,

$$\|A(z)^{-1}(A(x)-A(z))\| \le \overline{w}_0(\|x-z\|)+\overline{b}, \tag{34}$$

$$\begin{aligned}\|A(z)^{-1}[F'(x+t(y-x))-A(x)]\| &\le \overline{w}(\|x-z\|+t\|y-x\|) \\ &\quad - \overline{w}_0(\|x-z\|)+\overline{c}, \quad t\in[0,1]\end{aligned} \tag{35}$$

and (5) hold, where $\overline{w}(r+t)-\overline{w}_0(r)$, $t\ge 0$ are monotonically increasing functions with $\overline{w}(0)=\overline{w}_0(0)=w_2(0)=0$, $\overline{w}_0(r)$ is differentiable, $w_0'(r)>0$ at every point of $[0,R]$, ($\eta>0$ given by (18)) and parameters $\overline{b}$, $\overline{c}$ satisfy

$$\overline{b}\ge 0, \quad \overline{c}\ge 0 \quad \text{and} \quad \overline{b}+\overline{c}\le 1. \tag{36}$$

As in [5] set

$$\phi(r) = \eta - r + \int_0^r \overline{w}(t)dt, \tag{37}$$

$$\psi(r) = \int_0^r w_2(t)dt, \tag{38}$$

$$\chi(r) = \phi(r) + \psi(r) + (\bar{b} + \bar{c})r. \tag{39}$$

Denote the minimal value of $\chi(r)$ in $[0, R]$ by χ^*, and the minimal point by r^*. If $\chi(R) \leq 0$, denote the unique zero of χ by $r_0^* \in (0, r^*]$. Define scalar sequence $\{r_n\}$ $(n \geq 0)$ by

$$r_0 \in [0, R], \quad r_{n+1} = r_n + \frac{u(r_n)}{g(r_n)} \quad (n \geq 0), \tag{40}$$

where,

$$u(r) = \chi(r) - x^* \tag{41}$$

and

$$g(r) = 1 - \overline{w}_0(r) - \bar{b}. \tag{42}$$

We the above notation they showed:

Theorem 3. *Suppose that $\chi(R) \leq 0$. Then equation (1) has a solution $x^* \in \overline{U}(z, r_0^*)$, which is unique in*

$$\tilde{U} = \begin{cases} \overline{U}(z, R) & \text{if } \chi(R) < 0 \text{ or } \chi(R) = 0 \text{ and } r_0^* = R \\ U(z, R) & \text{if } \chi(R) = 0 \text{ and } r_0^* < R. \end{cases} \tag{43}$$

Let

$$D = \bigcup_{r \in [0, r^*)} \left\{ y \in \overline{U}(z, r) \mid \|A(y)^{-1}[F(y) + G(y)]\| \leq \frac{u(r)}{g(r)} \right\}. \tag{44}$$

Then, for any $y_0 \in D$, sequence $\{y_n\}$ $(n \geq 0)$ generated by Newton-like method (2) is well defined, remains in $\overline{U}(z, r^*)$ and satisfies

$$\|y_{n+1} - y_n\| \leq r_{n+1} - r_n, \tag{45}$$

and

$$\|y_n - x^*\| \leq r^* - r_n. \tag{46}$$

provided that r_0 is chosen in (40) so that $r_0 \in R_{y_0}$, where for $y \in D$

$$R_y = \left\{ r \in [0, r^*) \mid \|A(y)^{-1}(F(y) + G(y))\| \leq \frac{u(r)}{g(r)}, \|y - z\| \leq r \right\}, \tag{47}$$

or if we compute x_1 using (28) and set $y_0 = x_1$.

Remark 2. (a) Conditions (3), (4) are more general than (34), (35), respectively, since we can simply choose:

$$w_0(r) = \overline{w}_0(r),\ w(r) = \overline{w}(r),\ w_0(r) = w_1(r),\ r \in [0, R],\ b = \bar{b} \text{ and } c = \bar{c}. \tag{48}$$

(b) Other choices for y_0 are given by (44) and (47).

(c) Note that the hypotheses on function $\overline{w}_0$ is stronger than the "corresponding" function w_1.

(d) According to Theorem 3, iteration (40) converges to r^* (even if $r_0 = 0$) not r_0^*.

Remark 3. Our error bounds are finer in many interesting cases. Let us choose:

$$\overline{w}_0(r) = w_0(r) = w_1(r),\ w(r) = \overline{w}(r),\ r \in [0, R],\ \bar{b} = b \text{ and } \bar{c} = c. \tag{49}$$

Then we can show:

Theorem 4. *Under the hypotheses of Theorems 2 and 3, further assume*:

$$t_1 < r_1. \tag{50}$$

Then, the following hold:

$$\begin{aligned} t_n &< r_n \quad (n \geq 1), & (51)\\ t_{n+1} - t_n &< r_{n+1} - r_n \quad (n \geq 0), & (52)\\ t^* - t_n &\leq r^* - r_n \quad (n \geq 0) & (53) \end{aligned}$$

and

$$t^* \leq r^*. \tag{54}$$

Proof. Inequalities (53), (54) follow immediately from (51), (52) respectively. Hence we must show (51) and (52). By hypothesis (50), inequality (51) holds for $n = 1$. Using (10) and (40) we obtain in turn

$$\begin{aligned} t_2 - t_1 &= \frac{\int_0^1 \{w[t_0 + \theta(t_1 - t_0)]d\theta - w_1(t_0) + c\}(t_1 - t_0) + \int_{t_0}^{t_1} w_2(t)dt}{1 - b - w_0(t_1)} \\ &< \frac{\int_0^1 \{w[r_0 + \theta(r_1 - r_0)]d\theta - w_1(r_0) + c\}(r_1 - r_0) + \int_{r_0}^{r_1} w_2(t)dt}{1 - b - w_0(r_1)} \\ &= \frac{u(r_1) - u(r_0) + g(r_0)(r_1 - r_0)}{1 - b - w_0(r_1 - r_0)} \\ &= \frac{u(r_1)}{g(r_1)} = r_1 - r_0. \end{aligned} \tag{55}$$

Assume:

$$t_{k+1} < r_{k+1} \tag{56}$$

and

$$t_{k+1} - t_k < r_{k+1} - r_k \tag{57}$$

hold for all $k \leq n$.

Using (10), (40), (56) and (57) we obtain

$$\begin{aligned} t_{k+2} - t_{k+1} &= \frac{\int_0^1 \{w[t_k + \theta(t_{k+1} - t_k)]d\theta - w_1(t_k) + c\}(t_{k+1} - t_k) + \int_{t_k}^{t_{k+1}} w_2(t)dt}{1 - b - w_0(t_{k+1})} \\ &< \frac{\int_0^1 \{w[r_k + \theta(r_{k+1} - r_k)]d\theta - w_1(r_k) + c\}(r_{k+1} - r_k) + \int_{r_k}^{r_{k+1}} w_2(t)dt}{1 - b - w_0(r_{k+1})} \\ &= \frac{u(r_{k+1}) - u(r_k) + g(r_k)(r_{k+1} - r_k)}{g(r_{k+1})} \\ &= \frac{u(r_{k+1})}{g(r_{k+1})} = r_{k+2} - r_{k+1}. \end{aligned}$$

That completes the proof of Theorem 4. □

Remark 3. If (50) holds, the proof of Theorem 2 can be used to show (52)–(55) if:

$$\begin{aligned} w(r+t) - w_1(r) &\leq \overline{w}(r+t) - \overline{w}_0(r) \quad (t \geq 0), & (58)\\ b + w_0(r) &\leq \overline{b} + \overline{w}_0(r), \quad r \in [0, R], & (59) \end{aligned}$$

and

$$c \leq \overline{c}. \tag{60}$$

Remark 4. Conditions (7), (9) hold in many interesting cases. Assume:

$$A(x) = F'(x), \quad G(x) = 0, \tag{61}$$
$$\|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leq \ell\|x - y\|, \tag{62}$$
$$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \ell_0\|x - x_0\| \tag{63}$$

for all $x, y \in \overline{U}(z, r)$, $z = y_0$, and $r_0 = 0$. Then, we can set

$$b = c = \overline{b} = \overline{c} = 0, \tag{64}$$

$$\begin{aligned} w_1(r) &= w(r) = \overline{w}(r) = \overline{w}_0(r) = \ell r, \\ w_2(r) &= 0, \quad \text{and} \quad w_0(r) = \ell_0 r, \quad r \in [0, R] \end{aligned} \tag{65}$$

for some $\ell_0 \geq 0$, $\ell \geq 0$ with

$$\ell_0 \leq \ell. \tag{66}$$

(Note that Theorem 3 requires $\ell > 0$ but not Theorem 2.) That is, we consider the famous Newton–Kantorovich method [4], [8].

Condition (6) becomes

$$h_\delta = (\ell + \delta\ell_0)\eta \leq \delta. \tag{67}$$

Case 1. Let us restrict $\delta \in [0, 1]$. Hypotheses (9) now becomes

$$2\int_0^1 \ell\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right) + \theta\left(\frac{\delta}{2}\right)^{n+1}\eta\right]d\theta - 2\ell\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right)\right]$$
$$+ \delta\ell_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right)\right] \leq 2\ell\int_0^1 \theta\eta d\theta + \delta\ell_0\eta$$

or

$$\left[\frac{\ell_0\delta^2}{2-\delta} - \ell\right]\left[1 - \left(\frac{\delta}{2}\right)^{n+1}\right] \leq 1,$$

which is true for all $n \geq 0$ by the choice of δ. Note also

$$w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{n+1}\right)\right] < \frac{2\ell_0\eta}{2-\delta} \leq 2\ell_0\eta \leq 1.$$

It turns out that (67) can be weakened even further. Indeed:

Case 2. It follows from Case 1 that (6), (7), (9) reduce to (67),

$$\frac{2\ell_0\eta}{2-\delta} \leq 1, \tag{68}$$

and

$$\frac{\ell_0\delta^2}{2-\delta} \leq \ell, \tag{69}$$

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

Condition (67) for $\delta = 1$ becomes

$$h_1 = (\ell + \ell_0)\eta \leq 1. \tag{70}$$

Using (41) and (42) we get

$$\chi^* = 0, \quad r_0^* = R = r^* = \frac{1-\sqrt{1-h}}{\ell} \tag{71}$$

provided that the famous Newton–Kantorovich hypothesis holds [8]:

$$h = 2\ell\eta \leq 1. \tag{72}$$

Note that

$$h \leq 1 \Rightarrow h_1 \leq 1 \tag{73}$$

but not vice versa (unless if $\ell_0 = \ell$).

Moreover $\frac{\ell}{\ell_0}$ can be arbitrarily large. Indeed:

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

$$F(x) = c_0 x + c_1 + c_2 \sin e^{c_3 x}, \tag{74}$$

where c_i, $i = 0, 1, 2, 3$ are given parameters. It can easily be seen using (74) for c_3 large and c_2 sufficiently small $\frac{\ell}{\ell_0}$ can be arbitrarily large. That is (67) or (70) may hold but not (71).

Example 2. Let $X = Y = \mathbf{R}$, $\overline{U}(x_0, R) = \overline{U}(\sqrt{2}, 1)$ and define function F on $\overline{U}$ by

$$F(x) = \frac{1}{6}x^3 - \left(\frac{2^{3/2}}{6} + .23\right). \tag{75}$$

It can easily be seen that

$$\begin{aligned} \eta &= .23, \quad \ell = 2.4142136, \quad \ell_0 = 1.914213562, \\ h &= 1.1105383 > 1 \quad \text{and} \quad h_1 = .995538247 < 1. \end{aligned}$$

That is there is no guarantee that Newton's method starting at x_0 converges to $x^* = 1.614507018$ since (72) is violated. However since (70) holds our results guarantee $\lim_{n\to\infty} x_n = x^*$.

A more interesting example is given by the following:

Example 3. Let $X = Y = \mathbf{R}$, $x_0 = 1$, and define

$$F(x) = x^3 - \alpha, \quad G(x) = 0, \quad \text{for all } \alpha \in \left[0, \frac{1}{2}\right), \quad x \in [\alpha, 2-\alpha].$$

The Newton–Kantorovich hypothesis (72) does not hold since

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

That is there is no guarantee that Newton's method (2) converges to the solution $x^* = \sqrt[3]{\alpha}$ of equation $F(x) = 0$. However (70) holds for all $\alpha \in \left[\frac{5-\sqrt{13}}{3}, \frac{1}{2}\right)$ since for $\ell_0 = 3 - \alpha$,

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

Remark 5. The conclusions of Theorem 3 hold if (35) is replaced by the more general condition:

$$\|A(z)^{-1}[F'(x + t(y-x)) - A(x)]\| \leq \overline{w}(\|x-z\| + t\|y-x\|) - \overline{w}_1(\|x-z\|) + \overline{c}_0, \tag{76}$$

where function $\overline{w}_1$ and constant $\overline{c}_0$ have the properties of $\overline{w}_0$ and $\overline{c}$ respectively, provided that

$$\overline{w}_0(r) \le \overline{w}_1(r), \quad r \in [0, R]. \tag{77}$$

If $\overline{w}_1(r) = \overline{w}_0(r)$ $r \in [0, R]$ and $\overline{c}_0 = \overline{c}$, condition (76) reduces to (35). Moreover if strict inequality holds in (77) we obtain more precise error bounds. Indeed, let us denote by $\{s_n\}$ the sequence using (77). That is $\{s_n\}$ is given by

$$s_0 = r_0, \; s_1 = r_1, \; s_{n+1} - s_n = \frac{u(s_n) - u(s_{n-1}) + (1 - \overline{w}_1(s_{n-1}) - \overline{b})(s_n - s_{n-1})}{g(s_n)} \; (n \ge 1).$$

It can easily be seen using induction on n (see also the proof of Theorem 4 that follows) that

$$s_0 = r_0 \;\; = \;\; r_0, \quad s_1 = r_1, \tag{78}$$

$$s_{n+1} - s_n \;\; < \;\; r_{n+1} - r_n \quad (n \ge 1), \; s^* - s_n \le r^* - r_n \; (n \ge 0), \; s^* = \lim_{n\to\infty} s_n, \tag{79}$$

$$s_n < r_n \quad (n \ge 2) \quad \text{and} \quad s^* \le r^*. \tag{80}$$

Furthermore condition (77) allows us more flexibility in choosing the functions and constants. As an example, let us consider the Newton–Kantorovich method and assume (61)–(63). Then we can choose: $\overline{w}_0(r) = \ell_0 r$, $\overline{w}(r) = \overline{w}_1(r)$, $w_2(r) = 0$, $r \in [0, R]$ and $\overline{b} = \overline{c} = \overline{c}_0 = 0$. We have:

$$r_{n+1} = r_n + \frac{\ell(r_n - r_{n-1})^2}{2(1 - \ell r_{n-1})} \quad \text{and} \quad s_{n+1} = s_n + \frac{\ell(s_n - s_{n-1})^2}{2(1 - \ell_0 s_{n-1})} \; (n \ge 1). \tag{81}$$

Condition (77) becomes $\ell_0 \le \ell$, and in case $\ell_0 < \ell$ estimates (79), (80) hold.

3. Local Convergence of Newton-Like Methods

In order to cover the local case, let us assume x^* is a simple zero of equation (1), $A(x^*)^{-1}$ exists and for any $x, y \in \overline{U}(x^*, r) \subseteq \overline{U}(x^*, R)$:

$$\|A(x^*)^{-1}[A(x) - A(x^*)]\| \;\; \le \;\; v_0(\|x - x^*\|) + \beta, \tag{82}$$

$$\|A(x^*)^{-1}[F'(x + t(y - x)) - A(x)]\| \;\; \le \;\; v(\|x - x^*\| + t\|y - x\|) - v_1(\|x - x^*\|) + \gamma, \quad t \in [0, 1] \tag{83}$$

and

$$\|A(x^*)^{-1}[G(x) - G(y)]\| \le v_2(r)\|x - y\|, \tag{84}$$

where, v_0, β, v, v_1, γ, v_2 are as w_0, b, w, w_1, c, w_2 respectively. In order for us to compare our results with earlier ones we only consider the case $r_0 = 0$, $x_0 = z$ in (2), and call the corresponding sequence $\{x_n\}$ instead of $\{y_n\}$. Then along the lines of (31) but using (82)–(84) we can show the following local results for Newton-like methods:

Theorem 5. *Assume*:

there exists a minimal solution $\alpha^* \in [0, r]$ *of equation*

$$f(h) = 0, \tag{85}$$

where,

$$f(h) = \int_0^1 [v((1 + t)h) - v_1(h)]dt + v_2(h) + v_0(h) + \beta + \gamma - 1. \tag{86}$$

Then, sequence $\{x_n\}$ $(n \ge 0)$ *generated by Newton-like method (28) is well defined, remains in* $\overline{U}(x^*, \alpha^*)$ *for all* $n \ge 0$ *and converges to* x^*, *provided that* $x_0 \in U(x^*, \alpha^*)$.

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

$$\|x^* - x_{n+1}\| \leq p_{n+1}, \tag{87}$$

where,

$$p_{n+1} = \frac{\int_0^1 [v((1+t)\|x_n - x^*\|) - v_1(\|x_n - x^*\|)]dt\|x_n - x^*\| + \gamma\|x_n - x^*\| + \int_0^{\|x_n - x^*\|} v_2(t)dt}{1 - \beta - v_0(\|x_n - x^*\|)}, \tag{88}$$

$(n \geq 0)$.

Remark 6. Note that Theorem 5 can be proved using the weaker conditions

$$\begin{aligned} &\|A(x^*)^{-1}[F'(x + t(x^* - x)) - A(x)]\| \\ &\quad \leq \overline{v}(\|x - x^*\|(1+t)) - \overline{v}_1(\|x - x^*\|) + \overline{\gamma}, \quad t \in [0,1] \end{aligned} \tag{89}$$

and

$$\|A(x^*)^{-1}[G(x) - G(x^*)]\| \leq \overline{v}_2(r)\|x - x^*\| \tag{90}$$

for all $x \in \overline{U}(x^*, r) \subseteq \overline{U}(x^*, R)$, instead of (83) and (84) respectively, where, $\overline{v}$, $\overline{v}_1$, $\overline{\gamma}$, $\overline{v}_2$ are as v, v_1, γ, and v_2.

Remark 7. As an application let us again consider Newton's method: $F'(x) = A(x)$, $G(x) = 0$, and assume:

$$\|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq q_0\|x - x^*\| \tag{91}$$

and

$$\|F'(x^*)^{-1}[F'(x) - F'(y)]\| \leq q\|x - y\| \tag{92}$$

for all $x, y \in \overline{U}(x^*, r) \subseteq \overline{U}(x^*, R)$.

Then we can set

$$\beta = \gamma = 0,\ v_2(r) = 0,\ v_0(r) = q_0 r,\ v(r) = v_1(r) = qr,\ r \in [0, R]. \tag{93}$$

Using (86) we get

$$\alpha^* = \frac{2}{2q_0 + q}. \tag{94}$$

Local results were not given in [5]. However Rheinboldt in [13] showed that under only (92) that the convergence radius is given by

$$q_1 = \frac{2}{3q}. \tag{95}$$

Since in general

$$q_0 \leq q, \tag{96}$$

we conclude

$$q_1 \leq \alpha^*. \tag{97}$$

The corresponding error bounds are

$$\|x_{n+1} - x^*\| \leq \delta_n, \tag{98}$$

$$\|x_{n+1} - x^*\| \leq \delta_n^1, \tag{99}$$

where,

$$\delta_n = \frac{q\|x_n - x^*\|^2}{2[1 - q_0\|x_n - x^*\|]} \tag{100}$$

and

$$\delta_n^1 = \frac{q\|x_n - x^*\|^2}{2[1 - q\|x_n - x^*\|]}\,. \tag{101}$$

That is

$$\delta_n \le \delta_n^1 \quad (n \ge 0). \tag{102}$$

If strict inequality holds in (96) then (97) and (19) hold as strict inequalities also (see also Example 4 that follows).

Remark 8. As noted in [1], [4], [7], and [15] 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, and most efficient mesh refinement strategies.

Remark 9. 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 [4], [8]:

$$F'(x) = T(F(x)), \tag{103}$$

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

We complete this section with a numerical example.

Example 4. Let $X = Y = \mathbf{R}$, $D = \overline{U}(x^*, R) = U(0, 1)$, $G = 0$, $A(x) = F'(x)$, and define function F on D by

$$F(x) = e^x - 1. \tag{104}$$

Then it can easily be seen that we can set $T(x) = x + 1$ in (103). Hence we set $q = e$. Moreover since $x^* = 0$ we obtain in turn

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

and for $x \in U(0, 1)$, $\|F'(x) - F'(x^*)\| \le (e - 1)\|x - x^*\|$. Tht is, $q_0 = e - 1$. Using (94) and (96) we obtain respectively:

$$q_1 = .245252961$$

and

$$\alpha^* = .254028662.$$

That is, our convergence radius α^* is larger than the corresponding q_1 due to Rheinboldt [13], and our error bounds (100) finer than (101). This observation is very important in computational mathematics (see also Remark 8).

References

[1] Argyros, I.K., On a new Newton–Mysovskii-type theorem with applications to inexact Newton-like methods and their discretizations, *IMA J. Numer. Anal.* **18** (1997), 37–56.

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

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

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

[5] Chen, X. and Yamamoto, T., Convergence domains of certain iterative methods for solving nonlinear equations, *Numer. Funct. Anal. Optimiz.* **10**, (1 and 2) (1989), 37–48.

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

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

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

[9] Miel, G.J., Majorizing sequences and error bounds for iterative methods, *Math. Comput.* **34**, 149 (1980), 185–202.

[10] Moret, I., A note on Newton-type iterative methods, *Computing* **33** (1984), 65–73.

[11] Potra, F.A., On the convergence of a class of Newton-like methods. In: *Iterative Solution of Nonlinear Systems of Equations* (R. Ansarge and W. Toening, eds.), Lecture Notes in Math. **953**, Springer, Berlin, Heidelberg, New York, 1982, pp. 125–137.

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

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

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

[15] Ypma, T.J., Local convergence of inexact Newton methods, *SIAM J. Numer. Anal.* **21**, 3 (1984), 583–590.

[16] Zabrejko, P.P. and Nguen, D.F., The majorant method in the theory of Newton approximations and the Ptak error estimates, *Numer. Funct. Anal. and Optimiz.* **9** (1987), 671–684.