APPROXIMATING SOLUTIONS OF EQUATIONS USING NEWTON'S METHOD WITH A MODIFIED NEWTON'S METHOD ITERATE AS A STARTING POINT

IOANNIS K. ARGYROS*

Abstract. In this study we are concerned with the problem of approximating a locally unique solution of an equation in a Banach space setting using Newton's and modified Newton's methods. We provide weaker convergence conditions for both methods than before [6]–[8]. Then, we combine Newton's with the modified Newton's method to approximate locally unique solutions of operator equations in a Banach space setting. Finer error estimates, a larger convergence domain, and a more precise information on the location of the solution are obtained under the same or weaker hypotheses than before [6]–[8]. Numerical examples are also provided.

MSC 2000. 65H10, 65G99, 47H10, 49M15.

Keywords. Banach space, Newton–Kantorovich method, radius of convergence, Fréchet-derivative, Banach lemma on invertible operators.

1. INTRODUCTION

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

$$(1) \qquad\qquad\qquad F(x) = 0,$$

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

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is driven by the equation $\dot{x} = Q(x)$ for some suitable operator Q, where x is the state. Then the equilibrium states are determined by solving equation (1.1). 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

*Cameron University, Department of Mathematical Sciences, Lawton, OK 73505, USA, e-mail: iargyros@cameron.edu.

(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.

The most popular methods for approximation x^* are undoubtedly Newton's method

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

and the modified Newton's method

$$(3) \qquad y_{n+1} = y_n - F'(y_0)^{-1} F(y_n) \quad (n \geq 0), \quad (y_0 = x_0).$$

There is an extensive literature on the semilocal as well as the local convergence results for both methods under various hypotheses. Such results can be found in [4], [6], [7], and the references there.

The most popular hypotheses are of Newton–Kantorovich type [4], [6], [7]. Indeed, let $x_0 \in D$. Assume there exist constants $\eta > 0$, $\ell > 0$ such that

$$(4) \qquad F'(x_0)^{-1} \in L(Y, X),$$

$$(5) \qquad \|F'(x_0)^{-1} F(x_0)\| \leq \eta,$$

$$(6) \qquad \|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leq \ell \|x - y\| \ \text{ for all } x, y \in D,$$

$$(7) \qquad h = \ell \eta \leq \frac{1}{2},$$

and

$$(8) \qquad \overline{U}(x_0, s^*) = \{x \in X \mid \|x - x_0\| \leq s^*\} \subseteq D,$$

where,

$$(9) \qquad s^* = \frac{1 - \sqrt{1 - 2h}}{\ell} \, .$$

Estimate (7) is the crucial non-optimum sufficient condition for the semilocal convergence of both methods [4], [6], [7] (see also Theorem 1).

Under condition (7) method (2) converges quadratically to x^* (if (7) holds as a strict inequality) whereas method (3) converges linearly to x^*. There are examples in the literature where both methods converge to x^* but condition (7) is violated. Therefore one would expect that there may be conditions weaker than (7). This is the motivation for our study. Note that in view of the Lipschitz condition (6) it follows that there exists $\ell_0 > 0$ such that the center-Lipschitz condition

$$(10) \qquad \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \ell_0 \|x - x_0\| \ \text{ for all } x \in D$$

holds.

In general

$$(11) \qquad\qquad\qquad \ell_0 \leq \ell,$$

holds true, and $\frac{\ell}{\ell_0}$ can be arbitrarily large [2]–[4].

Recently, in [3, pp. 387, Case 3, for $\delta_0 = \delta$], we showed that condition (7) can always be replaced by the weaker

$$(12) \qquad h_1 = \ell_1 \eta \leq \frac{1}{2}, \quad \ell_1 = \frac{1}{8} \left(\ell + 4\,\ell_0 + \sqrt{\ell^2 + 8\,\ell_0\,\ell} \right),$$

in the case of Newton's method (2) (see also, Examples 9–11).

Here, we show that in the case of the modified method (3), for convergence, condition (7) can be replaced by

$$(13) \qquad\qquad\qquad h_0 = \ell_0 \eta \leq \frac{1}{2},$$

and (6) by weaker condition (10). Finer error estimates on the distances involved, a larger convergence domain, and a more precise information on the location of the solution than in earlier results [6] are also obtained this way (see Theorem 3 for method (2), and Theorem 5 for method (3)).

Using (13) (whenever (7) (or (12) do not hold)), we can employ method (3) for a finite number of steps, say N until condition (7) (or (12)) is satisfied for $x_0 = y_N$. Then faster method (2) takes over from method (3).

The above advantages extend to the local convergence of methods (2) and (3) (see Theorems 10 and 11). Numerical examples are also provided. The technique introduced here can extend to other Newton-type iterative methods [1], [3], [4], [5].

2. SEMILOCAL CONVERGENCE ANALYSIS

The following semilocal convergence result for methods (2) and (3) can be found in [7]:

THEOREM 1. *Let $F\colon D \subseteq X \to Y$ be a differentiable operator.*
Assume there exist $x_0 \in D$, and constants $\ell > 0$, $\eta > 0$ such that

$$F'(x_0)^{-1} \in L(Y, X),$$
$$\|F'(x_0)^{-1} F(x_0)\| \leq \eta,$$
$$\|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leq \ell \|x - y\| \quad \text{for all } x, y \in D,$$
$$h = \ell \eta \leq \frac{1}{2},$$

and

$$\overline{U}(x_0, s^*) \subseteq D,$$

where

$$s^* = \frac{1 - \sqrt{1 - 2\ell\eta}}{\ell}.$$

Then sequences $\{y_n\}$, $\{x_n\}$ are well defined, remain in $\overline{U}(x_0, s^*)$ for all $n \geq 0$ and converge to a unique solution x^* of equation $F(x) = 0$ in $\overline{U}(x_0, s^*)$. Moreover the following estimates hold:

$$\begin{aligned}
\|y_{n+1} - y_n\| &\leq q^n \|y_1 - y_0\| \leq q^n \eta, \\
\|y_n - x^*\| &\leq \frac{q^n}{1-q} \eta, \\
\|x_{n+2} - x_{n+1}\| &\leq \frac{\ell(s_{n+1} - s_n)^2}{2(1 - \ell s_{n+1})},
\end{aligned}$$

and

$$\|x_n - x^*\| \leq s^* - s_n, \quad s^* = \lim_{n \to \infty} s_n,$$

where,

$$s_0 = 0, \quad s_1 = \eta, \quad s_{n+2} = s_{n+1} + \frac{\ell(s_{n+1} - s_n)^2}{2(1 - \ell s_{n+1})} \quad (n \geq 0),$$

and

(14) $$q = 1 - \sqrt{1 - 2\ell h}.$$

REMARK 2. There is a plethora of estimates on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^*\|$, $\|y_{n+1} - y_n\|$, $\|y_n - x^*\|$ $(n \geq 0)$ [4], [6], [7]. However we decided to list only the estimates related to what we need in this study. In the case of Newton's method (2) we showed in [2], [3] the following improvement of Theorem 1. □

THEOREM 3. [2], [3]. *Let $F: D \subseteq X \to Y$ be a differentiable operator. Assume there exist $x_0 \in D$, and constants $\ell_0 > 0$, $\ell > 0$, $\eta \geq 0$ such that*

$$\begin{aligned}
&F'(x_0)^{-1} \in L(Y, X), \\
&\|F'(x_0)^{-1} F(x_0)\| \leq \eta, \\
&\|F'(x_0)^{-1} [F'(x) - F'(x_0)]\| \leq \ell_0 \|x - x_0\| \quad \text{for all } x \in D, \\
&\|F'(x_0)^{-1} [F'(x) - F'(y)]\| \leq \ell \|x - y\| \quad \text{for all } x, y \in D, \\
&h_1 = \ell_1 \eta \leq \frac{1}{2}, \\
&\overline{U}(x_0, t^*) \subseteq D,
\end{aligned}$$

where,

$$t_0 = 0, \quad t_1 = \eta, \quad t_{n+2} = t_{n+1} + \frac{\ell(t_{n+1} - t_n)^2}{2(1 - \ell_0 t_{n+1})} \quad (n \geq 0),$$

and

$$t^* = \lim_{n \to \infty} t_n \leq \frac{2\eta}{2 - \ell_2} = t_0^*, \quad \ell_2 = \frac{-\frac{\ell}{\ell_0} + \sqrt{\left(\frac{\ell}{\ell_0}\right)^2 + \frac{8\ell}{\ell_0}}}{2}.$$

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

Moreover the following estimates hold for all $n \geq 0$:

$$
\begin{aligned}
\|x_{n+1} - x_n\| &\leq t_{n+1} - t_n, \\
\|x_n - x^*\| &\leq t^* - t_n,
\end{aligned}
$$

$$
(15) \qquad\qquad t_n \leq s_n,
$$

$$
(16) \qquad\qquad t_{n+1} - t_n \leq s_{n+1} - s_n,
$$

and

$$
(17) \qquad\qquad t^* - t_n \leq s^* - s_n.
$$

REMARK 4. Note also that (15) and (16) hold as strict inequalities if $\ell_0 < \ell$ [2]–[4]. Moreover we have:

$$
(18) \qquad\qquad h \leq \frac{1}{2} \Rightarrow h_1 \leq \frac{1}{2},
$$

but not vice versa unless if $\ell_0 = \ell$. That is under the same computational cost we managed to weaken (7) since in practice the computation of ℓ also requires the computation of ℓ_0. Furthermore, in Example 9, we show that (12) holds but condition (7) is violated. $\qquad\square$

Concerning the semilocal convergence of the modified Newton's method we show that (13) replaces condition (7).

THEOREM 5. *Let $F \colon D \subseteq X \to Y$ be a differentiable operator.*
Assume there exist $x_0 \in D$ and constants $\ell_0 > 0$, $\eta > 0$, such that

$$
\begin{aligned}
&F'(x_0)^{-1} \in L(Y, X), \\
&\|F'(x_0)^{-1}F(x_0)\| \leq \eta, \\
&\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \ell_0\|x - x_0\| \quad \text{for all } x \in D,
\end{aligned}
$$

$$
h_0 = \ell_0\eta < \frac{1}{2},
$$

and

$$
(19) \qquad\qquad \overline{U}(x_0, s_0^*) \subseteq D,
$$

where,

$$
s_0^* = \frac{1 - \sqrt{1 - 2\ell_0\eta}}{\ell_0}.
$$

Then sequence $\{y_n\}$ ($n \geq 0$) generated by the modified Newton's method (3) is well defined, remains in $\overline{U}(x_0, s_0^*)$ for all $n \geq 0$ and converges to a unique solution x^* of equation $F(x) = 0$ in $\overline{U}(x_0, s_0^*)$.
Moreover the following estimates hold for all $n \geq 0$:

$$
\|y_{n+1} - y_n\| \leq q_0^n \, \|y_1 - y_0\|
$$

and

$$
\|y_n - x^*\| \leq \frac{q_0^n}{1 - q_0}\eta,
$$

where

$$q_0 = 1 - \sqrt{1 - 2\ell_0\eta}\,.$$

Proof. We shall show that the assumptions of the contraction mapping principle are satisfied for the operator

(20) $$P(x) = x - F'(x_0)^{-1}F(x) \quad \text{on} \quad \overline{U}(x_0, s_0^*).$$

Let $x, y \in \overline{U}(x_0, s_0^*)$. Then we can obtain the identity

$$\begin{aligned}
P(x) - P(y) &= x - y - F'(x_0)^{-1}(F(x) - F(y)) \\
&= F'(x_0)^{-1}\int_0^1 \{F'(x_0) - F'[y + t(x - y)]\}(x - y)\mathrm{d}t.
\end{aligned}$$

This identity together with (10) implies the estimate

$$\begin{aligned}
\|P(x) - P(y)\| &\leq \ell_0 \int_0^1 [(1-t)\|x - x_0\| + t\|y - x_0\|]\,\mathrm{d}t\ \|x - y\| \\
(21) &\leq \ell_0\, s_0^*\, \|x - y\| = q_0\, \|x - y\|.
\end{aligned}$$

Consequently, P is a contraction operator in the ball $\overline{U}(x_0, s_0^*)$. To complete the proof, it remains to show that

$$P\overline{U}(x_0, s_0^*) \subseteq \overline{U}(x_0, s_0^*).$$

Let $x \in \overline{U}(x_0, s_0^*)$. Then by (20) we can obtain in turn

$$\begin{aligned}
\|P(x) - x_0\| &\leq \|P(x) - P(x_0)\| + \|P(x_0) - x_0\| \\
&\leq \left\| F'(x_0)^{-1}\int_0^1 \{F'(x_0) - F'[x_0 + t(x - x_0)]\}(x - x_0)\mathrm{d}t \right\| + \eta \\
&\leq \ell_0 \int_0^1 t\,\mathrm{d}t\|x - x_0\|^2 + \eta \leq \frac{\ell_0}{2}(s_0^*)^2 + \eta = s_0^*,
\end{aligned}$$

by the choice of s_0^*. That completes the proof of Theorem 5. $\qquad\square$

REMARK 6. Note that by (21) the operator P satisfies a Lipschitz condition with constant q_0 in the ball $\overline{U}(x_0, s_0^*)$. The modified Newton's method thus converges at the rate of a geometric progression with quotient q_0. $\qquad\square$

The above analysis of method (3) relates to the simplest case. More subtle arguments (see, e.g. Kantorovich and Akilov [6]) show that Theorem 5 remains valid if the sign $<$ in (19) is replace by $\leq$. Therefore from now on we can replace (19) by (13) in Theorem 5.

REMARK 7. If $\ell_0 = \ell$ Theorems 3 and 5 reduce to Theorem 1. Otherwise these theorems constitute improvements of it. Indeed see (15)–(18), and notice that

$$q_0 < q$$

and

$$s_0^* < s^*.$$

Notice also that (7) (or (12)) implies (13) and if $t^* \leq s_0^*$ the quadratic convergence of method (2) is guaranteed. Moreover s_0^* given in closed form can then in practice replace t^*. Furthermore if $s_0^* < t^*$ then there exists $N > 1$ such that $x_n \in \overline{U}(x_0, s_0^*)$ for $n \geq N$ and then again s_0^* can replace t^*. $\square$

Next we show that we can start with method (3) and after a finite number of steps continue with faster method (2):

PROPOSITION 8. *Under hypotheses* (4)–(6), (10), (13), *and* (19), *for* $x_0 = y_0$, *define*

$$\alpha = \frac{1}{1 - \ell_0 \, s_0^*},$$
$$L = \alpha \, \ell,$$

for a fixed integer N

$$\ell_2 = \sup_{x \in \overline{U}(y_N, r_N)} \frac{\| F'(y_0)^{-1} \, [F'(x) - F'(y_N)] \|}{\| x - y_N \|},$$

$$L_0 = \alpha \, \ell_2 \leq L,$$

$$\overline{L} = \begin{cases} L & if \quad L_0 = L \\[2mm] \frac{1}{8} \, (L + 4 \, L_0 + \sqrt{L^2 + 8 \, L_0 \, L}) & if \quad L_0 < L, \end{cases}$$

$$N = \left[- \frac{\ln 2 \, \overline{L} \, \alpha^2 \, \eta}{\ln q_0} \right] + 1,$$

$$r_N = \begin{cases} \frac{1 - \sqrt{1 - 2 \, \overline{L} \, \eta_N}}{\overline{L}} & if \quad L_0 = L \\[2mm] \frac{2 \, \eta_N}{2 - L_{2\,N}} & if \quad L_0 < L, \end{cases}$$

$$\eta_N = \alpha \, q_0^N \, \eta,$$

and

$$L_{2\,N} = \frac{-\frac{L}{L_0} + \sqrt{\left(\frac{L}{L_0}\right)^2 + \frac{8\,L}{L_0}}}{2}, \qquad for \quad L_0 \neq 0,$$

where $[r]$ *denotes the integer part of real number* r.

Set

$$\overline{x_0} = y_N.$$

Moreover, assume:

(22) $$\overline{U}(y_N, r_N) \subseteq D.$$

Then the following hold:

(23) $$\| F'(y_N)^{-1} \, F(y_N) \| \leq \eta_N,$$

(24) $$\| F'(y_N)^{-1} \, [F'(x) - F'(y)] \| \leq L \, \| x - y \|,$$

(25) $$\| F'(y_N)^{-1} \, [F'(x) - F'(y_N)] \| \leq L_0 \, \| x - y_N \|,$$

(26) $$H_N = \overline{L} \, \eta_N \leq \frac{1}{2};$$

Newton's method (2), starting at $x_0 = \overline{x_0}$ converges to a unique solution x^{**} of equation $F(x) = 0$ in $\overline{U}(y_N, r_N)$,

and

$$N_0 \leq N_1,$$

where

$$N_0 = N \qquad \text{for} \quad L_0 = L,$$

and

$$N_1 = N \qquad \text{for} \quad L_0 < L.$$

Moreover, if the inclusion

(27) $$\overline{U}(y_N, r_N) \subseteq \overline{U}(y_0, s_0^*),$$

hold, then

$$x^* = x^{**}.$$

Note that parameter L_0 is independent of N if $L_0 = L$, and the inclusion (27) holds if and only if

$$\| y_N - y_0 \| + r_N \leq s_0^*.$$

If $\| y_N - y_0 \| \leq s_0^*$, then by the definition of r_N, there exists an integer N_0, such that (27) holds. In this case $N^* = \{N, N_0\}$ should replace N in the Proposition. Hypothesis (22) can now be dropped, since it follows from (19) and (27).

   Proof. Using Theorem 5, and the estimates

$$\begin{aligned}
\| F'(y_N)^{-1} \, F(y_N) \| \quad &\leq \quad \| F'(y_N)^{-1} \, F'(y_0) \| \, \| F'(y_0)^{-1} \, F(y_N) \| \\[2mm]
&\leq \quad \tfrac{1}{1-\ell_0 \, \|y_N-y_0\|} \, \| F'(y_0)^{-1} \, F(y_N) \| \\[2mm]
&\leq \quad \tfrac{\|F'(y_0)^{-1} \, F(y_N)\|}{1-\ell_0 \, s_0^*} \leq \alpha \, q_0^N \, \eta = \eta_N,
\end{aligned}$$

$$\begin{aligned}
\| F'(y_N)^{-1} \, [F'(x) - F'(y)] \| \quad &\leq \quad \alpha \, \| F'(y_0)^{-1} \, [F'(x) - F'(y)] \| \\
&\leq \quad \alpha \, \ell \, \| x - y \| = L \, \| x - y \|,
\end{aligned}$$

$$\begin{aligned}
\| F'(y_N)^{-1} \, [F'(x) - F'(y_N)] \| \quad &\leq \quad \alpha \, \| F'(y_0)^{-1} \, [F'(x) - F'(y_N)] \| \\
&\leq \quad \alpha \, \ell_2 \, \| x - y_N \| = L_0 \, \| x - y_N \|,
\end{aligned}$$

we obtain

$$H_N = \overline{L} \, \eta_N = \alpha^2 \, \overline{L} \, q_0^N \, \eta \leq \frac{1}{2}$$

and

$$N_0 \leq N_1,$$

by the choice of N.

It follows by Theorems 1 and 3, with L_0, L, η_N, r_N, r_N replacing ℓ_0, ℓ, η, s^*, t_0^* respectively, that there exists a unique solution x^{**} of equation $F(x) = 0$ in $\overline{U}(y_N, r_N)$.

Moreover, if inclusion (27) holds by the uniqueness of the solution x^* in $\overline{U}(y_0, s_0^*)$, we deduce $x^* = x^{**}$.

That completes the proof of Proposition 8. □

Let us provide an example.

EXAMPLE 9. Let $X = Y = \mathbf{R}$, $D = [a, 2 - a]$, $a \in \left[0, \frac{1}{2}\right)$ and define scalar function F on D by

$$(28) \qquad\qquad F(x) = x^3 - a.$$

Choose $y_0 = 1$. Using (5), (6), (10) and (28), we obtain

$$(29) \qquad \eta = \frac{1}{3}(1 - a), \quad \ell_0 = 3 - a \ \text{ and } \ \ell = 2(2 - a).$$

The Newton–Kantorovich hypothesis (7) becomes

$$(30) \qquad\qquad h = \frac{2}{3}(1 - a)(2 - a) > \frac{1}{2}$$

for all $a \in \left[0, \frac{1}{2}\right)$. That is according to Theorem 1 there is no guarantee that either methods (2) or (3) starting at $x_0 = y_0 = 1$ converge to x^*.

However according to Theorem 3 condition (12) becomes:

$$(31) \qquad\qquad h_1 \leq \frac{1}{2},$$

provided that

$$(32) \qquad\qquad a \in \left[.450339002, \frac{1}{2}\right).$$

Using condition (13) we can do even better since

$$(33) \qquad\qquad h_0 = \frac{1}{3}(1 - a)(3 - a) \geq \frac{1}{2}$$

provided that

$$(34) \qquad\qquad a \in \left[\frac{4 - \sqrt{10}}{2}, \frac{1}{2}\right)$$

which improves the choice for a given by (32). However only linear and not quadratic convergence is guaranteed.

Let us now use $a = .49$. In particular (7) does not hold since for $\eta = .17$, $\ell = 3.02$

$$h_1 = .5134 > \frac{1}{2}.$$

However (13) holds since for $\ell_0 = 2.51$:

$$h_0 = .4267 \leq \frac{1}{2} \,.$$

We get

$$q_0 = .617116205, \quad \alpha = 2.61175848, \quad s_0^* = .24586303,$$

(35) $$N = [4.0325] + 1 = 5.$$

Moreover we obtain

$$\overline{x}_0 = y_4 = .78911736.$$

Finally note that $x^* = .788373516$.

Our motivation for introducing condition (10) instead of (6) for the convergence modified Newton's method (3) can also be seen in the following examples.

EXAMPLE 10. Let $X = Y = \mathbf{R}$, $D = [0, \infty)$, $x_0 = 1$ and define function F on D by

(36) $$F(x) = \frac{x^{1+\frac{1}{i}}}{1+\frac{1}{i}} + c_1 x + c_2,$$

where c_1, c_2 are real parameters and $i > 2$ an integer. Then $F'(x) = x^{\frac{1}{i}} + c_1$ is not Lipschitz on D. However, center Lipschitz condition (10) holds for $\ell_0 = (1 + c_1)^{-1}$ $(c_1 \neq -1)$.

Indeed, we have

$$
\begin{aligned}
\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| &= (1+c_1)^{-1}\left|x^{\frac{1}{i}} - x_0^{\frac{1}{i}}\right| \\
&= \frac{(1+c_1)^{-1}|x-x_0|}{x_0^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}} \\
&\leq \ell_0 |x - x_0|.
\end{aligned}
$$

EXAMPLE 11. We consider the integral equation

(37) $$u(s) = f(s) + \lambda \int_a^b G(s,t)u(t)^{1+\frac{1}{n}}\mathrm{d}t, \quad n \in \mathbf{N}.$$

Here, f is a given continuous function satisfying $f(s) > 0$, $s \in [a, b]$, λ is a real number, and the kernel G is continuous and positive in $[a, b] \times [a, b]$.

For example, when $G(s, t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$
\begin{aligned}
u'' &= \lambda u^{1+\frac{1}{n}}, \\
u(a) &= f(a), \ \ u(b) = f(b).
\end{aligned}
$$

These type of problems have been considered in [2]–[7].

Equations of the form (37) generalize equations of the type

(38) $$u(s) = \int_a^b G(s,t)u(t)^n\mathrm{d}t$$

studies in [4], [6], [7].

Instead of (37) we can try to solve the equation $F(u) = 0$ where

$$F\colon \Omega \subseteq C[a,b] \to C[a,b], \quad \Omega = \{u \in C[a,b]\colon u(s) \geq 0, s \in [a,b]\},$$

and

$$F(u)(s) = u(s) - f(s) - \lambda \int_a^b G(s,t)u(t)^{1+\frac{1}{n}}\, \mathrm{d}t.$$

The norm we consider is the max-norm.

The derivative F' is given by

$$F'(u)v(s) = v(s) - \lambda\left(1 + \frac{1}{n}\right)\int_a^b G(s,t)u(t)^{\frac{1}{n}}v(t)\mathrm{d}t, \quad v \in \Omega.$$

First of all, we notice that F' does not satisfy a Lipschitz-type condition in Ω. Let us consider, for instance, $[a,b] = [0,1]$, $G(s,t) = 1$ and $y(t) = 0$. Then $F'(y)v(s) = v(s)$ and

$$\|F'(x) - F'(y)\| = |\lambda|\left(1 + \frac{1}{n}\right)\int_0^1 x(t)^{\frac{1}{n}}\mathrm{d}t.$$

If F' were a Lipschitz function, then

$$\|F'(x) - F'(y)\| \leq L_1\|x - y\|,$$

or, equivalently, the inequality

$$(39) \qquad \int_0^1 x(t)^{\frac{1}{n}}\, \mathrm{d}t \leq L_2 \max_{x \in [0,1]} x(s),$$

would hold for all $x \in \Omega$ and for a constant L_2. But this is not true. Consider, for example, the functions

$$x_j(t) = \tfrac{t}{j}, \quad j \geq 1, \quad t \in [0,1].$$

If these are substituted into (39)

$$\frac{1}{j^{1/n}\left(1+\frac{1}{n}\right)} \leq \frac{L_2}{j} \Leftrightarrow j^{1-\frac{1}{n}} \leq L_2\left(1 + \frac{1}{n}\right), \quad \forall j \geq 1.$$

This inequality is not true when $j \to \infty$.

Therefore, condition (6) fails in this case. However, condition (10) holds. To show this, let $x_0(t) = f(t)$ and $\alpha = \min_{s \in [a,b]} f(s)$, $\alpha > 0$. Then, for $v \in \Omega$,

$$\|[F'(x) - F'(x_0)]v\|$$

$$= |\lambda|\left(1 + \frac{1}{n}\right)\max_{s \in [a,b]}\left|\int_a^b G(s,t)\big(x(t)^{\frac{1}{n}} - f(t)^{\frac{1}{n}}\big)v(t)\mathrm{d}t\right|$$

$$\leq |\lambda|\left(1 + \frac{1}{n}\right)\max_{s \in [a,b]}\frac{G(s,t)|x(t) - f(t)|}{x(t)^{(n-1)/n} + x(t)^{(n-2)/n}f(t)^{1/n} + \cdots + f(t)^{(n-1)/n}}\mathrm{d}t\|v\|.$$

Hence,

$$
\begin{aligned}
\|F'(x) - F'(x_0)\| &\leq \frac{|\lambda|(1 + \frac{1}{n})}{\alpha^{(n-1)/n}} \max_{s \in [a,b]} \int_a^b G(s,t)\mathrm{d}t \|x - x_0\| \\
&\leq K \|x - x_0\|,
\end{aligned}
$$

where $K = \frac{|\lambda|(1 + \frac{1}{n})}{\alpha^{(n-1)/n}} N$, $N = \max\limits_{s \in [a,b]} \int_a^b G(s,t)\mathrm{d}t$, and $\ell_0 = \|F'(x_0)^{-1} K\|$. Finally note that condition (13) is satisfied for sufficiently small λ.

3. LOCAL CONVERGENCE ANALYSIS

In order for us to cover the local convergence of methods (2) and (3) we start the theorem [7]:

THEOREM 12. *Let $F \colon D \subseteq X \to Y$ be a differentiable operator. Assume there exist $x^* \in D$ and a constant $K > 0$ such that:*

(40) $F'(x^*)^{-1} \in L(Y, X), \quad F(x^*) = 0,$

(41) $\|F'(x^*)^{-1}[F'(x) - F'(y)]\| \leq K \|x - y\| \ \ \text{for all } x, y \in D,$

and

(42) $$\overline{U}(x^*, r_{RN}) \subseteq D$$

where,

(43) $$r_{RN} = \frac{2}{3\,K} \, .$$

Then

(a) *sequence $\{x_n\}$ generated by Newton's method (2) is well defined, remains in $\overline{U}(x^*, r_{RN})$ for all $n \geq 0$, converges to x^* provided that $x_0 \in U(x^*, r_{RN})$ and*

(44) $$\|x_{n+1} - x^*\| \leq \frac{K \|x_n - x^*\|^2}{2(1 - K \|x_n - x^*\|)} \quad (n \geq 0).$$

   If (58) is replaced by

(45) $$\overline{U}(x^*, r_{RM}) \subseteq D,$$

   where,

(46) $$r_{RM} = \frac{2}{5\,K},$$

   then

(b) *sequence $\{y_n\}$ generated by modified Newton's method (3) is well defined, remains in $\overline{U}(x^*, r_{RM})$ for all $n \geq 0$, converges to x^* provided that $x_0 \in U(x^*, r_{RM})$, and*

(47) $$\|y_{n+1} - x^*\| \leq \frac{K \left[\|x^* - y_0\| + \frac{1}{2}\|y_n - x^*\|\right]}{1 - K \|x^* - y_0\|} \|y_n - x^*\| \quad (n \geq 0).$$

Proof. The proof of (a) can be found in [8], whereas the proof of (b) is a special case of part (b) in our Theorem 11 that follows.

It follows from condition (41) that there exists $K_0 \in (0, K)$ such that:

$$(48) \qquad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq K_0 \, \|x - x^*\| \quad \text{for all } x \in D.$$

Then using a combination of conditions (41) and (48) for method (2), and only condition (48) for method (3) we can show: $\qquad\qquad\square$

THEOREM 13. *Let $F \colon D \subseteq X \to Y$ be a differentiable operator.*
Assume there exist $x^* \in D$ and constants $K_0 > 0$, $K > 0$ such that:

$$F'(x^*)^{-1} \in L(Y, X), \quad F(x^*) = 0,$$
$$\|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq K_0 \, \|x - x^*\| \quad \text{for all } x \in D,$$
$$\|F'(x^*)^{-1}[F'(x) - F'(y)]\| \leq K \, \|x - y\| \quad \text{for all } x, y \in D,$$

and

$$(49) \qquad\qquad\qquad \overline{U}(x^*, r_{AN}) \subseteq D$$

where,

$$(50) \qquad\qquad\qquad r_{AN} = \tfrac{2}{2\,K_0 + K} \, .$$

 (a) *Then sequence $\{x_n\}$ generated by Newton's method (2) is well defined, remains in $\overline{U}(x^*, r_{AN})$ for all $n \geq 0$, converges to x^* provided that $x_0 \in U(x^*, r_{AN})$ and*

$$(51) \qquad\qquad \|x_{n+1} - x^*\| \leq \tfrac{K \, \|x_n - x^*\|^2}{2(1 - K_0 \, \|x_n - x^*\|)} \quad (n \geq 0).$$

 Using only the center-Lipschitz condition, and if (49) is replaced by

$$(52) \qquad\qquad\qquad \overline{U}(x^*, r_{AM}) \subseteq D,$$

 and (48), where,

$$(53) \qquad\qquad\qquad r_{AM} = \tfrac{2}{5\,K_0} \, ,$$

 then,

 (b) *sequence $\{y_n\}$ generated by the modified Newton's method (3) is well defined, remains in $\overline{U}(x^*, r_{AM})$ for all $n \geq 0$, converges to x^* provided that $x_0 \in U(x^*, r_{AM})$,*
  and

$$(54) \qquad \|y_{n+1} - x^*\| \leq \tfrac{K_0 \left[\|x^* - y_0\| + \frac{1}{2}\|y_n - x^*\| \right] \|y_n - x^*\|}{1 - K_0 \, \|x^* - y_0\|} \quad (n \geq 0).$$

Proof. (a) The proof can be found in [2].
(b) Let $x \in \overline{U}(x^*, r_{AM})$. Then using (48) we get

$$(55) \qquad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq K_0 \, \|x - x^*\| \leq K_0 \, r_{AM} < 1$$

by the choice of r_{AM}. It follows from (55) and the Banach Lemma on invertible operators [4], [6] that $F'(x)^{-1}$ exists, and

$$(56) \qquad \|F'(x)^{-1}F'(x^*)\| \leq \tfrac{1}{1 - K_0 \, \|x - x^*\|} \leq \tfrac{1}{1 - K_0 \, r_{AM}} \, .$$

In particular by hypothesis $y_0 \in U(x^*, r_{AM}) \subseteq \overline{U}(x^*, r_{AM})$.

Let us assume $y_k \in \overline{U}(x^*, r_{AM})$ for $k = 0, 1, \ldots, n$. Then using (3), (48) and (53) we obtain in turn

$$
\begin{aligned}
y_{n+1} - x^* &= y_n - x^* - F'(y_0)^{-1}F(y_n) \\
&= F'(y_0)^{-1}[F'(y_0)(y_n - x^*) - F(y_n)] \\
&= -F'(y_0)^{-1}[F(y_n) - F(x^*) - F'(y_0)(y_n - x^*)] \\
&= -F'(y_0)^{-1}F'(x^*)F'(x^*)^{-1}\left\{ \int_0^1 [F'(x^* + t(y_n - x^*)) - F'(x^*)] \right.
\end{aligned}
$$

$$
(57) \qquad\qquad \left. + [F'(x^*) - F'(y_0)] \right\}(y_n - x^*)\mathrm{d}t,
$$

and

$$
\begin{aligned}
\|y_{n+1} - x^*\| &\leq \frac{K_0 \int_0^1 [\|x^* - y_0\| + t\|y_n - x^*\|]}{1 - K_0\,\|x^* - y_0\|}\|y_n - x^*\|\mathrm{d}t \\
&= \frac{K_0\,[\|x^* - y_0\| + \frac{1}{2}\|y_n - x^*\|]}{1 - K_0\,\|x^* - y_0\|}\|y_n - x^*\|
\end{aligned}
$$

$$
(58) \qquad\qquad < \|y_n - x^*\| \leq r_{AM}
$$

which shows (54), and $\lim_{n \to \infty} y_n = x^*$.

That completes the proof of Theorem 13. $\qquad\qquad\square$

REMARK 14. In general

$$
(59) \qquad\qquad\qquad\qquad K_0 \leq K
$$

holds and $\frac{K}{K_0}$ can be arbitrarily large. If $K_0 = K$ Theorem 13 reduces to Theorem 10. Otherwise Theorem 13 improves Theorem 12 under the same hypotheses for method (2), and the same or less computational cost for method (3); finer estimates on the distances $\|x_n - x^*\|$ $(n \geq 0)$ are obtained and the radius of convergence is enlarged. In particular, we have

$$
(60) \qquad\qquad\qquad r_{RN} \quad < \quad r_{AN}
$$

$$
(61) \qquad\qquad\qquad r_{RM} \quad < \quad r_{AM}.
$$

Moreover, since

$$
(62) \qquad\qquad\qquad r_{RM} < r_{RN},
$$

iterates from method (3) cannot be used to find the initial guess x_0 for faster method (2).

Examples where estimate (59) holds as in strict inequality can be found in [2]–[4]. However using Theorem 13 we can achieve this as follows: Let $p \in (0, 1)$ be the desired rate of convergence for method (3). Then by (54) we have for $\|y_0 - x^*\| \leq \frac{2p}{3\,K_0 + 2\,p\,K_0}$ that:

$$
\|y_{n+1} - x^*\| \leq p\|y_n - x^*\|.
$$

Choose:

(63)
$$M = \left\lceil \frac{\ln \frac{5}{2} \frac{K_0}{K_0 + K}}{\ln p} \right\rceil + 1,$$

then it can easily be seen that

$$p^M r_{AM} \leq r_{AN}$$

and consequently according to (a) of Theorem 11 we can set

$$x_0 = y_M.$$

In case of $K_0 = K$ according to (a) of Theorem 10 we can set

$$x_0 = y_{M_1},$$

where,

$$M_1 = \left\lceil \frac{\ln \frac{5}{3}}{\ln p} \right\rceil + 1$$

(simply set $K_0 = K$ in (63)). Note that

$$M \leq M_1.$$

Finally we observe

$$r_{RN} < r_{AM}$$

if

$$\frac{K}{K_0} > \frac{5}{3},$$

which can hold, since $\frac{K}{K_0}$ can be arbitrarily large [2]. □

The ideas presented here can be extended to other Newton-type iterative methods [1], [3], [4], [5] along the same lines.

4. CONCLUSION

The famous for its simplicity and clarity Newton–Kantorovich hypothesis (7) is the crucial sufficient semilocal convergence condition for both the quadratically convergent Newton's method (2), and the linearly convergent modified Newton's method (3) [6]. There exist simple numerical examples to show that both methods converge even when condition (7) is violated [4], [6]. In fact, it is common practice even condition (7) is violated for a certain initial guess, to still use Newton's method (2) for a few iterates until condition (7) is satisfied. However, this approach is considered a shot in the dark, and it is not working in general [4], [7]. Here we have introduced a certain approach that works in case condition (7)is violated. First, we showed that condition (13) is a weaker sufficient convergence hypothesis for the modified Newton's method (3) than (7). That is we extended the convergence region of method (3). Then, if (7) is violated but (13) holds, we start with slower method (3) until we reach, (after a finite number of steps) a certain iterate for which condition (7) also holds true. We then continue with faster method (2) using this iterate.

REFERENCES

[1] AMAT, S., BUSQUIER, S. and CANDELA, V., *A class of quasi-Newton generalized Steffensen methods on Banach spaces*, J. Comput. Appl. Math., **149**, pp. 397–408, 2002.

[2] ARGYROS, I.K., *On the Newton–Kantorovich hypothesis for solving equations*, J. Comput. Appl. Math., **169**, pp. 315–332, 2004.

[3] ARGYROS, I.K., *A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space*, J. Math. Anal. Applic., **298**, 2, pp. 374–397, 2004.

[4] ARGYROS, I.K., *Computational theory of iterative methods*, Studies in Computational Mathematics, **15** (Editors: C.K. Chui and L. Wuytack), Elsevier Publ. Co., New York, 2007.

[5] GUTIERREZ, J.M., HERNANDEZ, M.A. and SALANOVA, M.A., *Accessibility of solutions by Newton's method*, Intern. J. Comput. Math., **57**, pp. 239–247, 1995.

[6] KANTOROVICH, L.V. and AKILOV, G.P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[7] KRASNOSEL′SKII, M.A., VAINIKKO, G.M., ZABREIKO, P.P., RUTITSKII, YA.B. and STETSENKO, V.YA., *Approximate Solution of Operator Equations*, Wolters-Noordhoff Publ., Groningen, 1972.

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