Ioannis K. Argyros; Saïd Hilout

Extending the applicability of Newton's method using nondiscrete induction

Czechoslovak Mathematical Journal, Vol. 63 (2013), No. 1, 115–141

Persistent URL: `http://dml.cz/dmlcz/143174`

Czechoslovak Mathematical Journal, 63 (138) (2013), 115–141

EXTENDING THE APPLICABILITY OF NEWTON'S METHOD USING NONDISCRETE INDUCTION

Ioannis K. Argyros, Lawton, Saïd Hilout, Poitiers

(Received November 3, 2011)

Abstract. We extend the applicability of Newton's method for approximating a solution of a nonlinear operator equation in a Banach space setting using nondiscrete mathematical induction concept introduced by Potra and Pták. We obtain new sufficient convergence conditions for Newton's method using Lipschitz and center-Lipschitz conditions instead of only the Lipschitz condition used in F. A. Potra, V. Pták, Sharp error bounds for Newton's process, Numer. Math., 34 (1980), 63–72, and F. A. Potra, V. Pták, Nondiscrete Induction and Iterative Processes, Research Notes in Mathematics, 103. Pitman Advanced Publishing Program, Boston, 1984. Under the same computational cost as before, we provide: weaker sufficient convergence conditions; tighter error estimates on the distances involved and more precise information on the location of the solution. Numerical examples are also provided in this study.

Keywords: Newton's method, Banach space, rate of convergence, semilocal convergence, nondiscrete mathematical induction, estimate function

MSC 2010: 65H10, 65G99, 49M15

1. Introduction

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

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

where F is a Fréchet-differentiable operator defined on a closed and convex subset $\mathscr{D}$ of a Banach space $\mathscr{X}$ with values in a Banach space $\mathscr{Y}$.

Computational sciences have received substantial and significant interest of researchers in recent years in several areas such as engineering sciences, economic equilibrium theory and mathematics. These sciences can solve various problems by passing first through mathematical modelling and then later looking for the solution iteratively [9], [12], [15]. For example, finding a local minimum of a function is connected to solving a set of nonlinear equations. So, numerical methods are crucial and necessary for solving these nonlinear equations. Dynamic systems are also mathematically modeled by nonlinear differential or difference 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 $\dot{x} = \Lambda(x)$, for some suitable operator Λ, where x is the state. Then the equilibrium states are determined by solving the equation (1.1). Note that 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).

In computer graphics, the intersection of two surfaces is also modeled by nonlinear equation and can be complicated in general, because of some closed loops and singularities. This requires finding efficient algorithms for solving this intersection. We often need to compute and display the intersection $\mathscr{C} = \mathscr{A} \cap \mathscr{B}$ of two surfaces $\mathscr{A}$ and $\mathscr{B}$ in $\mathbb{R}^3$ [28]. If the two surfaces are explicitly given by

$$\mathscr{A} = \{(u,v,w)^T : \ w = F_1(u,v)\} \quad \text{and} \quad \mathscr{B} = \{(u,v,w)^T : \ w = F_2(u,v)\},$$

then the solution $x^\star = (u^\star, v^\star, w^\star)^T \in \mathscr{C}$ must satisfy the nonlinear equation

$$F_1(u^\star, v^\star) = F_2(u^\star, v^\star) \quad \text{and} \quad w^\star = F_1(u^\star, v^\star).$$

Hence, we must solve equation of the form (1.1) with $F := F_1 - F_2$. There is a signifiant literature addressing the surface intersection problem [11], [27].

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 control and 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. Finally, note that in computational sciences, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method [4], [12], [15], [19], [26], [29], [46], [66].

Newton's method (NM)

$$x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \quad (n \geqslant 0), \quad (x_0 \in \mathscr{D}), \tag{1.2}$$

is undoubtedly the most popular iterative process for generating a sequence $\{x_n\}$ approximating $x^\star$ [1]–[26], [29]–[68]. Here, $F'(x)$ $(x \in \mathscr{D})$ is the Fréchet-derivative of F at x. There is an extensive literature on local as well as semilocal convergence results of (NM) under various Lipschitz-type conditions. Recent results can be found in [9], [12], [15] and the references there (see also [11], [14], [47], [48]).

Let $x_0 \in \mathscr{D}$ be such that $F'(x_0)^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$, the space of bounded linear operators from $\mathscr{Y}$ into $\mathscr{X}$. We say that $F'(x_0)^{-1}F'(.)$ satisfies the Lipschitz-condition on $\mathscr{D}$ with constant L $(L > 0)$, if

$$\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leqslant L\|x - y\| \quad \text{for all } x, y \in \mathscr{D}. \tag{1.3}$$

Set

$$\|F'(x_0)^{-1}F(x_0)\| \leqslant r_0. \tag{1.4}$$

Then, a sufficient convergence condition for the semilocal convergence of (NM) is the Kantorovich hypothesis (KH), famous for its simplicity and clarity, given by (see [9], [12], [26])

$$H_K = 2Lr_0 \leqslant 1. \tag{1.5}$$

In the scalar case (1.5) coincides with the condition given earlier by Ostrowski [30]–[32]. If strict inequality holds in (1.5), the convergence is quadratic. Otherwise it is only linear. Later Ostrowski [32] obtained sharp a priori estimates. Simpler sharp a priori estimates were provided (using different method and proofs) by Gragg and Tapia [24] and some papers of Pták in [53], [54], [56], [58]. The celebrated method of nondiscrete induction is first used by Pták [55], [57]. Subsequently, Potra and Pták developed in a series of papers and an excellent book [35], [42], [43], [46] the nondiscrete induction and provided a posteriori estimates which are in general better than those given by Gragg and Tapia [24]. Other works on iterative methods and nondiscrete induction can be found in [39], [41], [42], [44], [59].

The hypothesis (1.5) is not a sufficient condition for the convergence of (NM). In Section 5 we provide an example where the hypothesis (1.5) is violated but (NM) (1.2) converges to the solution $x^\star$. Therefore, any hypothesis using the same information (F, x_0, L) weaker than (1.5) will expand applicability of (NM).

Let us report on what has been done in this direction. First of all note that in view of (1.3), $F'(x_0)^{-1}F'(.)$ satisfies a center-Lipschitz condition with constant L_0 $(L_0 > 0)$. That is

$$\|F'(x_0)^{-1}(F'(x) - F'(x_0))\| \leqslant L_0\|x - x_0\| \quad \text{for all} \quad x \in \mathscr{D}. \tag{1.6}$$

Note that in general

$$L_0 \leqslant L \tag{1.7}$$

holds, and L/L_0 can be arbitrarily large [5]–[15] (see Section 5 for Examples). Condition (1.6) is not an additional (to (1.3)) hypothesis, since in practice the computation of the Lipschitz constant L requires that of the center-Lipschitz constant L_0. We can then use (1.6) instead of (1.3) to compute upper bounds on the norms $\|F'(x)^{-1}F'(x_0)\|$. This observation has lead to the following set of advantages ($\mathscr{A}$) in the discrete case when $L_0 < L$ (see [5]–[15]):

▷ a weaker hypothesis than (KH) (1.5);

▷ tighter error bounds on the distances involved;

and

▷ at least as precise information on the location of the solution $x^\star$.

These advantages ($\mathscr{A}$) are obtained under the same information (x_0, F, L).

We have provided the following hypothesis instead of (1.5) (see, e.g. [5], [6], [8], [9], [12], [14], [15], [20])

$$H_1 = (5 + 2\sqrt{6})L_0 r_0 \leqslant 1, \tag{1.8}$$

$$H_2 = (L + L_0) r_0 \leqslant 1, \tag{1.9}$$

$$H_3 = 2\overline{L} r_0 \leqslant 1, \tag{1.10}$$

where,

$$\overline{L} = \frac{1}{8}(L + 4L_0 + (L^2 + 8L_0 L)^{1/2}). \tag{1.11}$$

Note that in particular

$$H_K \leqslant 1 \Longrightarrow H_2 \leqslant 1 \Longrightarrow H_3 \leqslant 1, \tag{1.12}$$

but not necessarily vice versa unless if $L_0 = L$. We also have

$$\frac{H_3}{H_K} \to \frac{1}{4} \quad \text{as} \quad \frac{L_0}{L} \to 0, \tag{1.13}$$

$$\frac{H_2}{H_K} \to \frac{1}{2} \quad \text{as} \quad \frac{L_0}{L} \to 0 \tag{1.14}$$

and

$$\frac{H_3}{H_2} \to \frac{1}{2} \quad \text{as} \quad \frac{L_0}{L} \to 0, \tag{1.15}$$

which provide a maximum measure on the expandability of (NM) under the hypotheses (1.8) or (1.9) or (1.10). By comparing (1.5) to (1.8) we get

$$\frac{L}{L_0} \geqslant \frac{5 + 2\sqrt{6}}{2} \quad \text{and} \quad H_K \leqslant 1 \Longrightarrow H_1 \leqslant 1 \tag{1.16}$$

or

$$\frac{L}{L_0} \leqslant \frac{5+2\sqrt{6}}{2} \quad \text{and} \quad H_1 \leqslant 1 \Longrightarrow H_K \leqslant 1. \tag{1.17}$$

Clearly, the first case (1.16) expands the applicability of (NM) when

$$\frac{L}{L_0} > \frac{5+2\sqrt{6}}{2}, \quad H_1 \leqslant 1 \quad \text{and} \quad H_K > 1. \tag{1.18}$$

The hypothesis (1.8) requires the computation of the constant L_0 only, whereas (1.9) and (1.10) require both constants L_0 and L. In [8], Argyros further weakened (1.8) in some sense using

$$H_M = 2L_0 r_0 \leqslant 1, \tag{1.19}$$

which is a sufficient convergence condition for the modified Newton's method (MNM)

$$y_{n+1} = y_n - F'(y_0)^{-1}F(y_n) \quad (n \geqslant 0), \quad (y_0 = x_0 \in \mathscr{D}). \tag{1.20}$$

But this time a certain number of iterates y_n in (1.20) must be computed until $y_N = x_0$ (N is a finite naturel number), for more details, see [8]. We also note that if (1.8) or (1.19) hold, then the convergence of (NM) is shown only to be linear. Note also that (1.19) is the weakest of the H hypotheses given by (1.5) and (1.8)–(1.10).

In this study we are motivated by optimization considerations and the method of nondiscrete mathematical induction as developed by Potra and Pták [42]. We show that the advantages ($\mathscr{A}$) carry over from the discrete to the nondiscrete case using (1.8) or (1.9) or (1.10) instead of (1.5) and smaller rate of convergence ω and corresponding estimate functions s (to be precised in Section 2). Note that ω and s are used to measure the error distances involved.

Potra and Pták [42] defined functions ω (see Figure 1) and s (see Figure 2) by

$$\omega(r) = \frac{1}{2}r^2(r^2+a^2)^{-1/2} \tag{1.21}$$

and

$$s(r) = r + (r^2+a^2)^{1/2} - a, \tag{1.22}$$

where $a \geqslant 0$. Under hypothesis (1.5), Potra and Pták [42] showed that the optimum value for a is given by

$$a = a_P = \Big(\frac{1}{L}\Big(\frac{1}{L} - 2r_0\Big)\Big)^{1/2}. \tag{1.23}$$

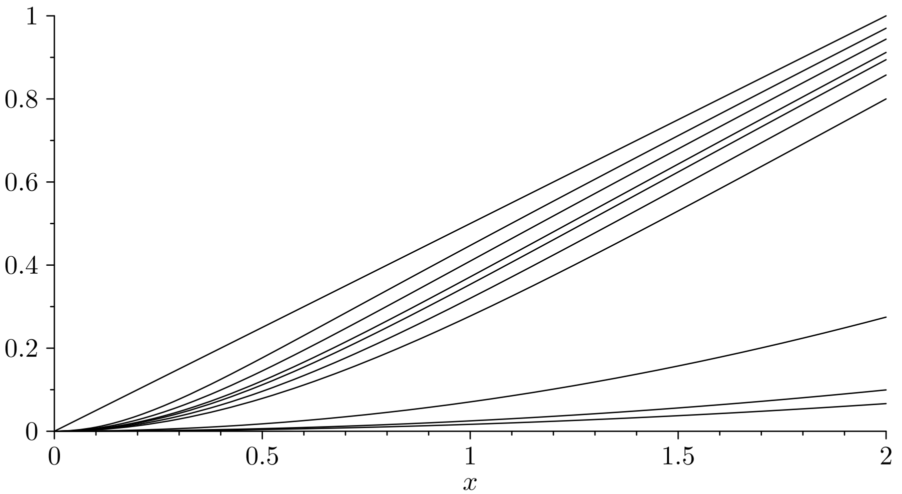

Figure 1. Functions $\omega(r)$ (from top to bottom) on $[0, 2]$ $a = 0, .5, .7, .9, 1, 1.2, 1.5, 7, 20, 30$, respectively.

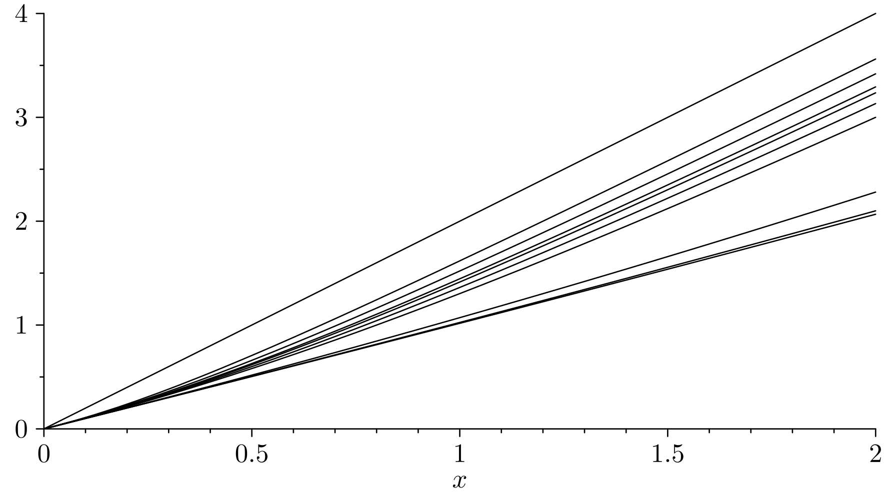

Figure 2. Functions $s(r)$ (from top to bottom) on $[0, 2]$ for $a = 0, .5, .7, .9, 1, 1.2, 1.5, 7, 20, 30$, respectively.

The error bounds are related with functions w and s by

$$d(x_n, x_{n-1}) \leqslant \omega^{(n)}(r_0) \tag{1.24}$$

and

$$d(x_n, x^\star) \leqslant s(\omega^{(n)}(r_0)), \tag{1.25}$$

where $\omega^{(n)}$ is the n-iterate of the function ω so that

$$\omega^{(0)}(r) = r,\ \omega^{(1)}(r) = \omega(r),\ \omega^{(2)}(r) = \omega(\omega(r)), \dots, \omega^{(n)}(r) = \omega(\omega^{(n-1)}(r)).$$

It follows from (1.21)–(1.25) that the larger the parameter "a" is the tighter the estimates (1.24) and (1.25) will be. If (1.9) holds, set:

$$a_1 = \frac{1}{L}((1 - L_0 r_0)^2 - L^2 r_0^2)^{1/2} \geqslant 0. \tag{1.26}$$

Moreover, if (1.19) is satisfied, let

$$a_M = \Big(\frac{1}{L_0}\Big(\frac{1}{L_0} - 2r_0\Big)\Big)^{1/2} \geqslant 0. \tag{1.27}$$

Note that if $L_0 = L$, then $a_M = a_1 = a_P$. Otherwise, we have

$$a_P < a_1 < a_M. \tag{1.28}$$

Other values for the parameter "a" have been given in Sections 3–5.

Our introduction of the center-Lipschitz condition in the discrete case has produced the advantages ($\mathscr{A}$) for other iterative processes such as the Secant method, the directional Newton method, Stirling's method, Steffensen's method and Newton-like methods [5], [9], [11], [12]–[15].

In this study we show that the advantages ($\mathscr{A}$) can carry from discrete to nondiscrete case. In particular we provide using the same information (F, x_0, L) a finer convergence analysis than in [34]–[37] for (NM).

The paper in organized as follows: In order to make the study as self contained as possible we have summarized some necessary concepts related to the method of nondiscrete mathematical induction in Section 2. The results on the enlargement of the parameter "a" are given in Section 3. The semilocal convergence of (NM) is given in Section 4. In the concluding Section 5 we provide numerical examples to support the claims made in the advantages ($\mathscr{A}$).

2. Nondiscrete mathematical induction and (NM)

Pták inaugurated in his Gatlinburg lecture [55] the method of Nondiscrete Mathematical Induction (NMI). We refer the reader to the excellent monograph by Potra and Pták [46] for more details about the motivation and general principles for (NMI). For $z \in \mathscr{X}$ and $r > 0$, we denote by $\overline{U}(z, r)$ the closed ball centered at z and of radius r. Let $\mathscr{T}$ be either the positive real axis or an interval of the form

$$\mathscr{T} = \{r \in \mathbb{R}\colon\ 0 < r < \alpha\} = (0, \alpha).$$

We need the definition of the rate of convergence.

Definition 2.1. A function $\omega\colon \mathscr{T} \to \mathscr{T}$ is called a rate of convergence on $\mathscr{T}$ if the series

$$\sum_{n=0}^{\infty} \omega^{(n)}(r) \tag{2.1}$$

is convergent for each $r \in \mathscr{T}$. The sum (2.1) is denoted by $s(r)$ and is called the corresponding estimate function. Then we write

$$s(r) = \sum_{n=0}^{\infty} \omega^{(n)}(r) \quad \text{for all } r \in \mathscr{T}. \tag{2.2}$$

Functions ω and s satisfy the functional equation

$$s(r) = r + s(\omega(r)). \tag{2.3}$$

It then follows from (2.3) that (with the exception of pathological cases) we have:

$$\omega(r) = s^{-1}(s(r) - r). \tag{2.4}$$

That is, given s, the function ω can be recovered using the functional equation (2.4). The computation of the function s is very difficult or impossible in general.

We have the following result characterizing rates of convergence.

Proposition 2.2 [46]**.** *Let $\omega\colon \mathscr{T} \to \mathscr{T}$ and $\nu\colon \mathscr{T} \to \mathscr{T}$ be such that*

$$\nu(r) = r + \nu(\omega(r)) \quad \textit{for all } r \in \mathscr{T}. \tag{2.5}$$

Then the following items hold:

(a) *ω is a rate of convergence on $\mathscr{T}$*

and

(b) *if the limit $\nu(0) = \lim\limits_{r \searrow 0} \nu(r)$ exists, then*

$$s(r) = \sum_{n=0}^{\infty} \omega^{(n)}(r) = \nu(r) - \nu(0) \quad \textit{for all } r \in \mathscr{T}. \tag{2.6}$$

It can easily be seen by verifying (2.5), that the function ω given by (1.16) is a rate of convergence on $\mathscr{T}$ with the corresponding estimate function s given by (1.17).

Another example is given for $\delta \in [0, 1)$ by

$$\omega(r) = \delta r \tag{2.7}$$

and

$$s(r) = \frac{r}{1-\delta}. \tag{2.8}$$

Define $G\colon \mathscr{D} \to \mathscr{Y}$ by

$$G(x) = x - F'(x)^{-1}F(x). \tag{2.9}$$

We need the following result relating the (MNI), (2.9) and (NM).

Lemma 2.3 [42], [46]**.**
(1) *Assume that for a given pair* (G, x_0) *there exists a rate of convergence* ω *on an interval* $\mathscr{T}$ *and a family of sets* $\mathscr{Z}(r) \subseteq \mathscr{X}$ *such that the inclusion conditions* $x_0 \in \mathscr{Z}(r_0)$ *for a certain* $r_0 \in \mathscr{T}$ *and*

$$r \in \mathscr{T} \text{ and } x \in \mathscr{Z}(r) \implies G(x) \in U(x,r) \cap \mathscr{Z}(\omega(r)) \tag{2.10}$$

are satisfied.

Then, sequence $\{x_n\}$ *generated by* (NM) *is well defined and converges to a point* $x^\star$*. Moreover,* (1.24), (1.25) *and the following estimate*

$$x_n \in \mathscr{Z}(\omega^{(n)}(r_0)) \tag{2.11}$$

hold.

(2) *If, in addition, for a certain* $n \geqslant 1$*, we have*

$$x_{n-1} \in \mathscr{Z}(d(x_n, x_{n-1})), \tag{2.12}$$

then for this n*, the following estimate holds*

$$d(x_n, x^\star) \leqslant f(d(x_n, x_{n-1})) \tag{2.13}$$

for some function $f\colon [0,\infty) \to [0,\infty)$ *such that*

$$f(r) = s(r) - r. \tag{2.14}$$

Lemma 2.3 is essentially a corollary of the induction theorem (see Proposition 1.7 in [46, p. 5]). This theorem is related to the graph theorem of functional analysis. The closed graph theorem can be seen as a limit case of the induction theorem for an infinitely fast rate of convergence (see, e.g. [46, Theorem 1.15]).

We use the following measure of invertibility

$$d(\mathscr{B}) = \inf_{\|x\|=1} \|\mathscr{B}(x)\| \quad \text{for } \mathscr{B} \in \mathscr{L}(\mathscr{X}, \mathscr{Y}). \tag{2.15}$$

If $\mathscr{B}$ is invertible and $\mathscr{B}^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$, then

$$d(\mathscr{B}) = \|\mathscr{B}^{-1}\|^{-1}. \tag{2.16}$$

We also need the following Banach-type result on invertible operators [4], [9].

Lemma 2.4. *If $\mathscr{B}$ and $\mathscr{C}$ belong in $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ such that $\mathscr{B}$ is boundedly invertible and*

$$d(\mathscr{B}) > \|\mathscr{B} - \mathscr{C}\| \tag{2.17}$$

then $\mathscr{C}$ is also boundedly invertible and

$$d(\mathscr{C}) \geqslant d(\mathscr{B}) - \|\mathscr{B} - \mathscr{C}\|. \tag{2.18}$$

3. Enlarging the parameter "a"

Nondiscrete induction for iterative processes requires verification of inclusion hypotheses in (1) of Lemma 2.3. We shall illustrate how this method works in the case of (NM).

The differences between our approach and the one given by Potra and Pták [42], [46] will also be given in our description that follows.

First, we need to define a suitable nonempty approximate set $\mathscr{Z}$ for some rate of convergence ω. If x is an initial guess, we hope

$$x_+ = x - F'(x)^{-1}F(x) \tag{3.1}$$

to be closer to the solution $x^\star$. Let r be the distance between x and x_+. We must have for $x \in \mathscr{Z}(r)$ that $x_+ \in \mathscr{Z}(\omega(r))$.

Potra and Pták [46, p. 23] used the following approximate set $\mathscr{Z}(r)$ $(r > 0)$ for a rate of convergence ω (first in non-affine invariant form):

$$\begin{aligned}\mathscr{Z}(r) = \{x \in \mathscr{X} \colon\ &\|x - x_0\| \leqslant g(r),\ F'(x) \text{ is invertible}, \\ &\|F'(x)^{-1}F(x)\| \leqslant r \text{ and } d(F'(x_0)^{-1}F(x)) \geqslant h(r)\},\end{aligned} \tag{3.2}$$

where g and h are functions to be determined later. This way they produced a plethora of results on (NM) that have improved the error bounds on the distances $d(x_n, x_{n-1})$ and $d(x_n, x^\star)$ of the discrete case but not the sufficient convergence condition (1.5).

Let $x \in \mathscr{X}(r)$, then for $x_+ \in \mathscr{X}(\omega(r))$, the following must hold:

$$\|x_+ - x_0\| \leqslant g(\omega(r)), \tag{3.3}$$

$$d(F'(x_0)^{-1}F'(x_+)) \geqslant h(\omega(r)) \tag{3.4}$$

and

$$\|F'(x_+)^{-1}F(x_+)\| \leqslant \omega(r). \tag{3.5}$$

But we can write

$$\|x_0 - x_+\| \leqslant \|x_+ - x\| + \|x - x_0\| \leqslant r + g(r). \tag{3.6}$$

We also have

$$\begin{aligned} d(F'(x_0)^{-1}F'(x_+)) &\geqslant d(F'(x_0)^{-1}F'(x)) - \|F'(x_0)^{-1}(F'(x_+) - F'(x))\| \\ &\geqslant d(F'(x_0)^{-1}F'(x)) - (\|F'(x_0)^{-1}(F'(x_+) - F'(x_0))\| \\ &\quad + \|F'(x_0)^{-1}(F'(x_0) - F'(x))\|) \\ &\geqslant d(F'(x_0)^{-1}F'(x)) - L_0(\|x_+ - x_0\| + \|x - x_0\|) \\ &\geqslant h(r) - L_0(r + 2g(r)). \end{aligned} \tag{3.7}$$

As long as $h(r) - L_0(r + 2g(r))$ is positive, the Banach lemma on invertible operators [4], [9], [26] and Lemma 2.4 guarantee the existence of $F'(x)^{-1}$ and the estimate

$$\|F'(x)^{-1}F'(x_0)\| \leqslant (h(r) - L_0(r + 2g(r)))^{-1}. \tag{3.8}$$

Using the approximation

$$F(x_+) = F(x_+) - F(x) - F'(x)(x_+ - x) = \int_0^1 (F'(x + t(x_+ - x)) - F'(x))(x_+ - x)\,\mathrm{d}t$$

and (1.3), we get

$$\|F'(x_0)^{-1}F(x_+)\| \leqslant \frac{1}{2}L\|x_+ - x\|^2. \tag{3.9}$$

Then, we have by (3.8) and (3.9)

$$\begin{aligned} \|F'(x_+)^{-1}F(x_+)\| &\leqslant \|F'(x_+)^{-1}\ F'(x_0)\|\|F'(x_0)^{-1}F(x_+)\| \\ &\leqslant \frac{1}{2}L(h(r)-L_0(r+2g(r)))^{-1}r^2. \end{aligned} \tag{3.10}$$

In view of (3.6), (3.8) and (3.10), the conditions (3.3)–(3.5) hold if there exist functions h, g and parameter b satisfying the system of inequations S_{AH}:

$$g(r)+r \leqslant g(\omega(r)), \tag{3.11}$$

$$h(r)-L_0(r+2g(r)) \geqslant h(\omega(r)), \tag{3.12}$$

$$\frac{L}{2}r^2\big(h(r)-L_0(r+2g(r))\big)^{-1} \leqslant \omega(r), \tag{3.13}$$

$$g(r) < b \tag{3.14}$$

and

$$0 < h(r) \leqslant 1. \tag{3.15}$$

The system S_{PP} in [46, p. 25] uses inequation

$$h(r)-Lr \geqslant h(\omega(r)) \tag{3.16}$$

instead of (3.12). The rest of the inequations are the same.

We shall see later that replacing (3.16) by (3.12) is a major modification leading to the advantages ($\mathscr{A}$) already stated in the Introduction of this study.

Next, we shall show that the system S_{AH} is satisfied in two cases when the rate of convergence ω is given by (1.16) or (2.7) and

$$h(r) = L\Big(a+\frac{L-L_0}{L}\big(s(r)-r\big)+\frac{L_0}{L}\big(2s(r_0)-s(r)\big)\Big), \tag{3.17}$$

$$g(r) = s(r_0)-s(r), \tag{3.18}$$

$$b_0 = s(r_0) < b, \tag{3.19}$$

where s is the estimate function corresponding to rate of convergence ω and $a \geqslant 0$ is to be determined later.

In the first case the functions ω, s are given by (1.21) and (1.22), respectively.

Proposition 3.1. *Let $r_0 \geqslant 0$ and $L \geqslant L_0 > 0$. Let also the functions ω, s be given by* (1.21) *and* (1.22), *respectively.*
Assume that (1.9) *and*

$$b_1 = \frac{1}{L}(1 + (L - L_0)r_0 - ((1 - L_0 r_0)^2 - L^2 r_0^2)^{1/2}) < b \tag{3.20}$$

hold.
Then, the system S_{AH} has a solution (h, g, b_1), where

$$\begin{aligned} h(r) = L\Big(a_1 + \frac{L - L_0}{L}((r^2 + a_1^2)^{1/2} - a_1) \\ + \frac{L_0}{L}(a_1 + 2b_1 - r - (r^2 + a_1^2)^{1/2})\Big), \end{aligned} \tag{3.21}$$

$$g(r) = \frac{1 + (L - L_0)r_0}{L} - r - (r^2 + a_1^2)^{1/2} \tag{3.22}$$

and a_1 is given by (1.26).
Moreover, we have

$$x_0 \in \mathscr{X}(r_0). \tag{2.23}$$

P r o o f. By the hypothesis (1.9), $a_1 \geqslant 0$. Indeed, if $L_0 \neq L$, we have:

$$\begin{aligned} (L_0^2 - L^2)\Big(r_0 - \frac{1}{L_0 - L}\Big)\Big(r_0 - \frac{1}{L_0 + L}\Big) \geqslant 0 \Longrightarrow (L_0^2 - L^2)r_0^2 - 2L_0 r_0 + 1 \geqslant 0 \\ \Longrightarrow (1 - L_0 r_0)r_0^2 - L^2 r_0^2 \geqslant 0 \Longrightarrow a_1 \geqslant 0. \end{aligned}$$

If $L_0 = L$, then again we deduce $a_1 \geqslant 0$, since $2Lr_0 \leqslant 1$.

Moreover it can easily be seen by simple substitution that the triplet (h, g, b_1) satisfies the system S_{AH}. Note in particular that (3.21) implies (3.15). Finally, the inclusion (3.23) follows from (3.1) and (3.15). That completes the proof of Proposition 3.1. □

Remark 3.2. If $L_0 = L$, the hypothesis (1.9) reduces to (1.5). In this case we have $a = a_1 = a_P$.

If $L_0 < L$, Proposition 3.1 improves the results in [46] and $a_1 > a_P$ (see also Example 5.1).

In the second case the functions ω, s are given by (2.7) and (2.8), respectively.

Proposition 3.3. *Let $r_0 \geqslant 0$ and $L \geqslant L_0 > 0$. Let also the functions ω, s be given by* (2.7) *and* (2.8)*, respectively, for $\delta = \frac{1}{2}$.*
Assume that (1.9) *for $L \geqslant 3L_0$ or $4L_0 r_0 \leqslant 1$ for $L \leqslant 3L_0$ and*

$$b_2 = 2r_0 < b \tag{3.24}$$

hold.
Then, the system S_{AH} has a solution (h, g, b_2), where

$$h(r) = (L - 3L_0)r + 4L_0 r_0 \tag{3.25}$$

and

$$g(r) = 2(r_0 - r). \tag{3.26}$$

Proof. It is easy to see by substitution that the system S_{AH} is satisfied with the above choices of g, h, b_2 and b. That completes the proof of Proposition 3.3. □

Remark 3.4. It turns out that if the approximate set $\mathscr{X}$ is defined in a way other than (3.2), then the system S_{AH} can be simplified and weaker hypotheses than before are needed (in some cases).

This time, we define

$$\begin{aligned} \mathscr{X}_0(r) = \{x \in \mathscr{X}\colon\ &\|x - x_0\| \leqslant g(r),\ F'(x) \text{ is invertible}, \\ &\|F'(x)^{-1}F(x)\| \leqslant r \quad \text{and} \quad d(F'(x_0)^{-1}F(x)) \geqslant 1 - L_0(r + g(r))\}. \end{aligned} \tag{3.27}$$

The motivation for the introduction of the new approximate set $\mathscr{X}_0$ is due to the estimate

$$\begin{aligned} d(F'(x_0)^{-1}F'(x_+)) &\geqslant d(F'(x_0)^{-1}F'(x_0)) - \|F'(x_0)^{-1}(F'(x_+) - F'(x_0))\| \\ &\geqslant 1 - L_0\|x_+ - x_0\| \geqslant 1 - L_0(r + g(r)). \end{aligned} \tag{3.28}$$

Then, in view of the implications

$$\begin{aligned} \omega(r) \geqslant 0 &\Longrightarrow \omega(r) + s(r_0) - s(\omega(r)) \geqslant r + s(r_0) - s(r) \\ &\Longrightarrow 1 - L_0(r + g(r)) \geqslant 1 - L_0(\omega(r) + g(\omega(r))), \end{aligned} \tag{3.29}$$

the inequation (3.12) can be dropped from the system S_{AH}. Denote the resulting system by $S^{\star}_{AH}$ defined by

$$S^{\star}_{AH} \begin{cases} g(r) + r \leqslant g(\omega(r)), \\ \dfrac{L}{2} r^2 (1 - L_0(r + g(r)))^{-1} \leqslant \omega(r), \\ g(r) < b, \\ 0 < L_0(r + g(r)) \leqslant 1. \end{cases}$$

Then, we can have results corresponding to Propositions 3.1 and 3.3, respectively.

Proposition 3.5. *Under the hypotheses of Proposition* 3.1, $S^{\star}_{AH}$ *has a solution* (g, b_1), *where* g *and* b_1 *are given in Proposition* 3.1.

Moreover, we have

$$x_0 \in \mathscr{Z}_0(r_0). \tag{3.30}$$

P r o o f. It can easily be seen that the pair (g, b_1) satisfies the system $S^{\star}_{AH}$. In particular for the verification of (3.13), we must show

$$1 - L_0(r + g(r)) \geqslant 0 \tag{3.31}$$

and

$$Lr^2 \leqslant 2\omega(r)(1 - L_0(r + g(r))). \tag{3.32}$$

We have

$$1 - L_0(r + g(r)) = 1 - L_0(r + (r^2 + a_1^2)^{1/2}) \geqslant 0 \tag{3.33}$$

by the choice of a_1 and $r \in [0, r_0]$. Hence the estimate (3.31) holds. We also have

$$\begin{aligned} a_1^2 &\leqslant \frac{(1 - L_0 r)^2 - L^2 r^2}{L^2} \Longrightarrow L^2 a_1^2 \leqslant (1 - L_0 r)^2 - L^2 r^2 \Longrightarrow L^2(r^2 + a_1^2) \\ &\leqslant (1 - L_0 r)^2 \Longrightarrow L(r^2 + a_1^2)^{1/2} \leqslant 1 - L_0 r \Longrightarrow L(r^2 + a_1^2)^{1/2} \\ &\leqslant 1 - L_0(r + (r^2 + a_1^2)^{1/2}) + L_0(r^2 + a_1^2)^{1/2} \Longrightarrow Lr^2 \\ &\leqslant r^2(r^2 + a_1^2)^{-1/2}(1 - L_0(r + (r^2 + a_1^2)^{1/2}) + L_0(r^2 + a_1^2)^{1/2}) \Longrightarrow (3.31). \end{aligned}$$

That completes the proof of Proposition 3.5. □

Proposition 3.6. *Let* $r_0 \geqslant 0$ *and* $L \geqslant L_0 > 0$. *Let also* ω, s *be given by* (2.6) *and* (2.8), *respectively, with*

$$\delta = \frac{2L}{L + (L^2 + 8L_0 L)^{1/2}}. \tag{3.34}$$

Suppose that (1.10) *and*

$$b_3 = \frac{r_0}{1 - \delta} < b \tag{3.35}$$

hold.

Then the system $S^{\star}_{AH}$ has a solution (g, b_3), where,

$$g(r) = \frac{1}{1-\delta}(r_0 - r). \tag{3.36}$$

Moreover, we have $x_0 \in \mathscr{Z}_0(r_0)$. Furthemore, $\delta \in [1/2, 1)$.

Proof. We shall show how do we arrive at the hypothesis (1.10) and the choice of δ. The rest shall follow by substituting (g, b_3) in $S^{\star}_{AH}$.

Indeed, we have

$$Lr^2 \leqslant 2\omega(r)(1 - L_0(r + g(r)))$$

or

$$r\Big(L - \frac{2L_0\delta^2}{1-\delta}\Big) \leqslant 0$$

or

$$2L_0\delta^2 + L\delta - L \geqslant 0, \tag{3.37}$$

which is true as equality by (3.34). We must also show

$$g(r) \geqslant 0 \tag{3.38}$$

or

$$s(r_0) \geqslant s(r)$$

or

$$\frac{r}{1-\delta} \leqslant \frac{1}{L_0}$$

or

$$r_0 L_0 \leqslant 1 - \delta = 1 - \frac{2L}{L + (L^2 + 8L_0L)^{1/2}}$$

or

$$r_0 L_0 \leqslant \frac{-L + (L^2 + 8L_0L)^{1/2}}{L + (L^2 + 8L_0L)^{1/2}}$$

or

$$r_0 L_0 (L + (L^2 + 8L_0L)^{1/2})^2 \leqslant 8L_0L$$

or

$$r_0(L + 4L_0 + (L^2 + 8L_0L)^{1/2}) \leqslant 4,$$

which is exactly the hypothesis (1.10). That completes the proof of Proposition 3.6. □

Remark 3.7. If $L = L_0$, then (1.10) reduces to (1.5) and $\delta = 1/2$. If $L_0 < L$, then (1.10) is weaker than (1.9) and (1.5).

4. Semilocal convergence of (NM)

The only difference in the proofs of [46, Sections 1 and 5], [42] and ours is that we use different value of “a” and (1.9) or (1.10) instead of (1.5). Therefore the proofs of semilocal convergence results (corresponding to Propositions 3.1, 3.3 and 3.6) for (NM) are omitted.

For brevity, we only provide estimates of the form (1.24), (1.25) and (2.11). Estimates of the form (2.13) can also follow immediately as in [46], [42] but using different “a” as in Propositions 3.1, 3.3 and 3.6.

Theorem 4.1. *Let $F\colon \mathscr{D} \subseteq X \to \mathscr{Y}$ be a Fréchet-differentiable operator and let $x_0 \in \mathscr{D}$. Assume that*

(i)
$$F'(x_0)^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X});$$

(ii) *$F'(x_0)^{-1}F'$ satisfies the Lipschitz condition with constant L and the center-Lipschitz condition with constant L_0 on $\mathscr{D}$;*

(iii)
$$\|F'(x_0)^{-1}F(x_0)\| \leqslant r_0;$$

(iv) *the hypotheses of Proposition* 3.1 *hold; and*

(v)
$$\overline{U}(x_0, b_1) \subseteq \mathscr{D}.$$

Then the sequence $\{x_n\}$ $(n \geqslant 0)$ generated by (NM) *is well defined, remains in $\overline{U}(x_0, b_1)$ for all $n \geqslant 0$ and converges to a unique solution $x^\star$ of the equation* (1.1) *in $\overline{U}(x_0, b_1)$.*

Moreover, the following error estimates hold for all $n \geqslant 1$:

$$d(x_n, x_{n-1}) \leqslant \omega^{(n)}(r_0) = \frac{2a_1\theta_1(r_0)^{2^n}}{1-\theta_1(r_0)^{2^{n+1}}}, \tag{4.1}$$

$$d(x_n, x^\star) \leqslant s(\omega^{(n)}(r_0)) = \frac{2a_1\theta_1(r_0)^{2^n}}{1-\theta_1(r_0)^{2^n}} \tag{4.2}$$

and

$$d(x_n, x^\star) \leqslant (a_1^2 + \|x_n - x_{n-1}\|^2)^{1/2} - a_1, \tag{4.3}$$

where

$$\theta_1(r) = \frac{(r^2 + a_1^2)^{1/2} - a_1}{r}, \tag{4.4}$$

where a_1 is given by (1.26).

Theorem 4.2. *Let $F\colon \mathscr{D} \subseteq X \to \mathscr{Y}$ be a Fréchet-differentiable operator and let $x_0 \in \mathscr{D}$. Assume that*

(1) *the hypotheses* (i)–(iii) *of Theorem* 4.1 *hold;*

(2) *the hypotheses of Proposition* 3.3 *hold and*

(3)
$$\overline{U}(x_0, b_2) \subseteq \mathscr{D}.$$

Then the sequence $\{x_n\}$ $(n \geqslant 0)$ generated by (NM) *is well defined, remains in $\overline{U}(x_0, b_2)$ for all $n \geqslant 0$ and converges to a unique solution $x^\star$ of the equation* (1.1) *in $\overline{U}(x_0, b_2)$.*

Moreover, the following error estimates hold for all $n \geqslant 1$:

$$d(x_n, x_{n-1}) \leqslant \omega^{(n)}(r_0) = \Big(\frac{1}{2}\Big)^n r_0 \tag{4.5}$$

and

$$d(x_n, x^\star) \leqslant s(\omega^{(n)}(r_0)) = \Big(\frac{1}{2}\Big)^{n-1} r_0. \tag{4.6}$$

Theorem 4.3. *Let $F\colon \mathscr{D} \subseteq X \to \mathscr{Y}$ be a Fréchet-differentiable operator and let $x_0 \in \mathscr{D}$. Assume that*

(1) *the hypotheses* (i)–(iii) *of Theorem* 4.1 *hold;*

(2) *the hypotheses of Proposition* 3.6 *hold*

and

(3)
$$\overline{U}(x_0, b_3) \subseteq \mathscr{D}.$$

Then, the conclusions of Theorem 4.2 *hold with b_3, $\frac{1}{2}\delta$ replacing b_2 and $\frac{1}{2}$, respectively.*

Remark 4.4. If $L_0 = L$, the results reduce to the corresponding ones in [42], [46]. Otherwise they constitute an improvement since (1.9) or (1.10) are weaker than (1.5), error estimates are tighter and the information on the location of the solution $x^\star$ is more precise, since our a_1 is larger than a_P. Indeed, under the hypotheseis (1.5), the error bounds in [42], [46] are:

$$d(x_n, x_{n-1}) \leqslant \frac{2a_P\theta_P(r_0)^{2^n}}{1-\theta_P(r_0)^{2^{n+1}}}, \tag{4.7}$$

$$d(x_n, x^\star) \leqslant \frac{2a_P\theta_P(r_0)^{2^n}}{1-\theta_P(r_0)^{2^n}} \tag{4.8}$$

and

$$(4.9) \qquad d(x_n, x^\star) \leqslant (a_P^2 + \|x_n - x_{n-1}\|^2)^{1/2} - a_P,$$

where

$$(4.10) \qquad \theta_P(r) = \frac{(r^2 + a_P^2)^{1/2} - a_P}{r},$$

where a_P is given by (1.23), and

$$(4.11) \qquad \overline{b}_0 = \frac{1}{L} - \Big(\frac{1}{L}\Big(\frac{1}{L} - 2r_0\Big)\Big)^{1/2} < b.$$

Then, we have by (1.21), (1.23), (1.26), (3.20) and (4.11)

$$\theta_1(r) < \theta_P(r), \qquad r \in [0, r_0]$$

and

$$b_1 < \overline{b}_0.$$

Concerning (MNM) defined by (1.20), we have the following semilocal convergence result.

Theorem 4.5. *Let $F\colon \mathscr{D} \subseteq X \to \mathscr{Y}$ be a Fréchet-differentiable operator and let $x_0 \in \mathscr{D}$. Assume that*

(1) *the hypotheses* (i)–(iii) *of Theorem* 4.1 *and* (1.19) *hold*

and

(2)
$$\overline{U}(x_0, b_M) \subseteq \mathscr{D},$$

where
$$b_M = \frac{1}{L_0} - \Big(\frac{1}{L_0}\Big(\frac{1}{L_0} - 2r_0\Big)\Big)^{1/2}.$$

Then the sequence $\{x_n\}$ $(n \geqslant 0)$ generated by (MNM) *given by* (1.20) *is well defined, remains in $\overline{U}(x_0, b_M)$ for all $n \geqslant 0$ and converges to a unique solution $x^\star$ of the equation* (1.1) *in $\overline{U}(x_0, b_M)$.*

Moreover, the estimates (1.24), (1.25) *and*

$$\|x_n - x^\star\| \leqslant s(\|x_n - x_{n-1}\|) - \|x_n - x_{n-1}\|$$

hold, with

$$\omega(r) = \frac{1}{2}L_0 r^2 + r(1 - (L_0^2 a_M^2 + 2L_0 r)^{1/2})$$

and

$$s(r) = \left(a_M^2 + \frac{2r}{L_0}\right)^{1/2} - a_M,$$

where a_M *is given by* (1.27).

Remark 4.6. If $L_0 = L$, Theorem 4.5 reduces to the corresponding one in [42], [46]. Otherwise they constitute an improvement, since $a_P < a_M$, $b_M < \overline{b}_0$ and our functions ω, s are smaller than the ones in [42], [46].

5. Numerical examples

We provide examples where our results apply but earlier ones do not. When all results apply we show that ours provide tighter error bounds and better information on the location of the solution.

Example 5.1. Let $\mathscr{X} = \mathscr{Y} = \mathbb{R}$, equipped with the max-norm, $x_0 = 1$, $\mathscr{D} = [\varrho, 2-\varrho]$, $\varrho \in [0, \frac{1}{2})$ and define the function F on $\mathscr{D}$ by

$$F(x) = x^3 - \varrho. \tag{5.1}$$

Using (1.3), (1.4) and (1.6) we get:

$$r_0 = \frac{1}{3}(1-\varrho), \quad L_0 = 3 - \varrho \quad \text{and} \quad L = 2\ (2-\varrho).$$

Then, we obtain the conditions (1.5) and (1.8), respectively, as follow

$$H_K = \frac{4}{3}(1-\varrho)(2-\varrho) > 1$$

and

$$H_1 = \frac{1}{3}(5+2\sqrt{6})(3-\varrho)(1-\varrho) > 1 \quad \text{for all } \varrho \in \left[0, \frac{1}{2}\right).$$

Hence, there is no guarantee that (NM) converges to $x^\star = \sqrt[3]{\varrho}$, starting at x_0.

However, if we consider our conditions (1.19), (1.8) and (1.10), respectively, we get

$$H_M = \frac{2}{3}(3-\varrho)(1-\varrho) \leqslant 1 \qquad \text{for all } \varrho \in [.418861170, .5),$$

$$H_2 = \frac{1}{3}(7-3\varrho)(1-\varrho) \leqslant 1 \quad \text{for all } \varrho \in [.464816242, .5)$$

and

$$H_3 = \frac{1}{6}(8 - 3\varrho + (5\varrho^2 - 24\varrho + 28)^{1/2})(1-\varrho) \leqslant 1 \quad \text{for all } \varrho \in [.450339002, .5).$$

Next we pick three values of ϱ such that all hypotheses are satisfied, so we can compare the "a" values and the corresponding error bounds.

Case $\varrho = .49999$

By Maple 13, we have the following results

$$x^\star = .7936952346, \quad H_K = 1.000026667 > 1, \quad H_1 = 4.124673776 > 1,$$
$$H_2 = .9166899999 < 1, \quad H_3 = .8877981560 < 1, \quad H_M = .8333533332 < 1,$$
$$a_1 = .1001396659, \quad a_M = .1632888647$$

and

$$b_1 = .2609701491, \quad b_2 = .3333400000, \quad b_3 = .3551178419.$$

We can not compare (4.1) and (4.7) in the case $\varrho \in (.5, 1)$ since (1.5) does not hold and a_P is a complex number in this interval. Note that we have

$$a_P \geqslant 0 \Longleftrightarrow \varrho \in (.5, 2.5).$$

Case $\varrho = .5$

By Maple 13, we have the following results

$$x^\star = .7937005260, \quad H_K = 1, \quad H_2 = .9166666665 < 1,$$
$$a_P = 0, \quad a_1 = .1001542021, \quad b_1 = .5109569091 \quad \text{and} \quad b_2 = .3333333333.$$

Then the convergence is only linear in [42], [46] (see also the estimates (4.7)–(4.9) in Remark 4.4) since $a_P = 0$, but our Theorems 4.1 and 4.2 apply and we can produce the following tables (Tables 1 and 2) for estimating error bounds (4.1), (4.2) and (4.5), (4.6), respectively.

n	x_n	(4.1)	(4.2)
1	.8333333333	.05612119686	.07077314850
2	.8151148834	.01371698569	.01465195163
3	.8059078274	.0009306422249	.0009349659389
4	.8008359800	.000004323620699	.000004323714024
5	.7979271348	9.332457124e-11	9.332457129e-11
6	.7962228487	4.348033039e-20	4.348033039e-20
7	.7952122874	9.438141843e-39	9.438141843e-39
8	.7946089091	4.447068601e-76	4.447068601e-76
9	.7942471777	9.872985216e-151	9.872985216e-151
10	.7940297902	4.866287937e-300	4.866287937e-300

Table 1.

n	x_n	(4.5)	(4.6)
1	.8333333333	.08333333335	.1666666666
2	.8151148834	.04166666666	.08333333335
3	.8059078274	.02083333334	.04166666666
4	.8008359800	.01041666666	.02083333334
5	.7979271348	.005208333335	.01041666666
6	.7962228487	.002604166666	.005208333335
7	.7952122874	.001302083334	.002604166666
8	.7946089091	.0006510416665	.001302083334
9	.7942471777	.0003255208334	.0006510416665
10	.7940297902	.0001627604166	.0003255208334

Table 2.

Case $\varrho = .52$

By Maple 13, we have the following results

$$x^\star = .8041451517, \quad H_K = .9472000000 < 1, \quad H_1 = 3.927915060 > 1,$$
$$H_2 = .8703999998 < 1, \quad H_3 = .8438043214 < 1, \quad H_M = .7935999998 < 1,$$
$$a_1 = .1262055091, \quad a_M = .1831905919, \quad a_P = .07762922486$$

and

$$b_1 = .2375782746, \quad b_2 = .3200000000, \quad b_3 = .3402436781.$$

We can now compare our results of Theorem 4.1 (see also the estimates (4.1)–(4.3) with the ones in [42], [46] (see also the estimates (4.7)–(4.9)).

n	x_n	(4.1)	(4.7)
1	.8400000000	.07006258936	.07909638766
2	.8229000192	.01700313396	.02822534747
3	.8144944601	.001135124615	.004822385995
4	.8099973466	.000005104594030	.0001494969517
5	.8074963585	1.032319444e-10	1.439489908e-7
6	.8060774320	4.222016307e-20	1.334633444e-13
7	.8052635983	7.062061639e-39	1.147278253e-25
8	.8047939593	1.975853311e-76	8.477782648e-50
9	.8045219996	1.546682206e-151	4.629235881e-98
10	.8043641969	9.477501671e-302	1.380267862e-194

Comparison Table 3.

Comparison Tables 3 and 4 show that our error bounds (4.1) and (4.2) are finer than (4.7) and (4.8) given in [42], [46].

n	x_n	(4.2)	(4.8)
1	.8400000000	.08820595264	.1122937620
2	.8229000192	.01814336327	.03319737437
3	.8144944601	.001140229312	.004972026896
4	.8099973466	.000005104697262	.0001496409008
5	.8074963585	1.032319444e-10	1.439491243e-7
6	.8060774320	4.222016307e-20	1.334633444e-13
7	.8052635983	7.062061639e-39	1.147278253e-25
8	.8047939593	1.975853311e-76	8.477782648e-50
9	.8045219996	1.546682206e-151	4.629235881e-98
10	.8043641969	9.477501671e-302	1.380267862e-194

Comparison Table 4.

Finally, we provide examples where the inequality between the Lipschitz and the center-Lipschitz constants is strict (i.e., $L_0 < L$).

Example 5.2. Define the scalar function F by $F(x) = d_0 x + d_1 + d_2 \sin e^{d_3 x}$, $x_0 = 0$, where d_i, $i = 0, 1, 2, 3$ are given parameters. Then it can easily be seen that for d_3 large and d_2 sufficiently small, L/L_0 can be arbitrarily large.

Example 5.3. Let $\mathscr{X} = \mathscr{Y} = \mathscr{C}[0,1]$, equipped with the max-norm. Consider the following nonlinear boundary value problem [9]

$$\begin{cases} u'' = -u^3 - \gamma u^2, \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 \mathscr{Q}(s,t)(u^3(t) + \gamma u^2(t))\,\mathrm{d}t \tag{5.2}$$

where $\mathscr{Q}$ is the Green function:

$$\mathscr{Q}(s,t) = \begin{cases} t(1-s), & t \leqslant s, \\ s(1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0 \leqslant s \leqslant 1} \int_0^1 |\mathscr{Q}(s,t)|\,\mathrm{d}t = \frac{1}{8}.$$

Then the problem (5.2) is in the form (1.1), where, $F\colon \mathscr{D} \to \mathscr{Y}$ is defined as

$$[F(x)](s) = x(s) - s - \int_0^1 \mathscr{Q}(s,t)(x^3(t) + \gamma x^2(t))\,\mathrm{d}t.$$

If we set $u_0(s) = s$ and $\mathscr{D} = U(u_0, R)$, then since $\|u_0\| = 1$, it is easy to verify that $U(u_0, R) \subset U(0, R + 1)$. If $2\gamma < 5$, then the operator F' satisfies the conditions of Theorem 4.1, with

$$r_0 = \frac{1+\gamma}{5-2\gamma}, \quad L = \frac{\gamma + 6R + 3}{4}, \quad L_0 = \frac{2\gamma + 3R + 6}{8}.$$

Note that $L_0 < L$.

Other applications and examples including the solution of nonlinear Chandrasekhar-type integral equations appearing in radiative transfer are also found in [9], [15].

Conclusion

For approximating a solution of a nonlinear operator equation in a Banach space setting, we provided new results for (NM) and (MNM) using the concept of (NMI) introduced by Potra and Pták [42], [46]. We obtained new sufficient convergence conditions for (NM) and (MNM) using Lipschitz and center-Lipschitz conditions on the Fréchet-derivative of the operator involved instead of only the Lipschitz condition used in [42], [46]. Our results extend the applicability of these methods studied in [42], [46]. Numerical examples are also provided in this study.

References

[1] *S. Amat, C. Bermúdez, S. Busquier, J. Gretay*: Convergence by nondiscrete mathematical induction of a two step secant's method. Rocky Mt. J. Math. *37* (2007), 359–369.

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

[3] *S. Amat, S. Busquier, J. M. Gutiérrez, M. A. Hernández*: On the global convergence of Chebyshev's iterative method. J. Comput. Appl. Math. *220* (2008), 17–21.

[4] *I. K. Argyros*: The Theory and Application of Abstract Polynomial Equations. St. Lucie/CRC/Lewis Publ. Mathematics series, Boca Raton, Florida, USA, 1998.

[5] *I. K. Argyros*: A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space. J. Math. Anal. Appl. *298* (2004), 374–397.

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

[7] *I. K. Argyros*: Concerning the "terra incognita" between convergence regions of two Newton methods. Nonlinear Anal., Theory Methods Appl. *62* (2005), 179–194.

[8] *I. K. Argyros*: Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point. Rev. Anal. Numér. Théor. Approx. *36* (2007), 123–137.

[9] *I. K. Argyros*: Computational Theory of Iterative Methods. Studies in Computational Mathematics 15. Elsevier, Amsterdam, 2007.

[10] *I. K. Argyros*: On a class of Newton-like methods for solving nonlinear equations. J. Comput. Appl. Math. *228* (2009), 115–122.
[11] *I. K. Argyros*: A semilocal convergence analysis for directional Newton methods. Math. Comput. *80* (2011), 327–343.
[12] *I. K. Argyros, S. Hilout*: Efficient Methods for Solving Equations and Variational Inequalities. Polimetrica Publisher, Milano, Italy, 2009.
[13] *I. K. Argyros, S. Hilout*: Enclosing roots of polynomial equations and their applications to iterative processes. Surv. Math. Appl. *4* (2009), 119–132.
[14] *I. K. Argyros, S. Hilout*: Extending the Newton-Kantorovich hypothesis for solving equations. J. Comput. Appl. Math. *234* (2010), 2993–3006.
[15] *I. K. Argyros, S. Hilout, M. A. Tabatabai*: Mathematical Modelling with Applications in Biosciences and Engineering. Nova Publishers, New York, 2011.
[16] *W. Bi, Q. Wu, H. Ren*: Convergence ball and error analysis of the Ostrowski-Traub method. Appl. Math., Ser. B (Engl. Ed.) *25* (2010), 374–378.
[17] *E. Cătinaş*: The inexact, inexact perturbed, and quasi-Newton methods are equivalent models. Math. Comput. *74* (2005), 291–301.
[18] *X. Chen, T. Yamamoto*: Convergence domains of certain iterative methods for solving nonlinear equations. Numer. Funct. Anal. Optimization *10* (1989), 37–48.
[19] *P. Deuflhard*: Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms. Springer Series in Computational Mathematics 35. Springer, Berlin, 2004.
[20] *J. A. Ezquerro, J. M. Gutiérrez, M. A. Hernández, N. Romero, M. J. Rubio*: The Newton method: from Newton to Kantorovich. Gac. R. Soc. Mat. Esp. *13* (2010), 53–76. (In Spanish.)
[21] *J. A. Ezquerro, M. A. Hernández*: On the R-order of convergence of Newton's method under mild differentiability conditions. J. Comput. Appl. Math. *197* (2006), 53–61.
[22] *J. A. Ezquerro, M. A. Hernández*: An improvement of the region of accessibility of Chebyshev's method from Newton's method. Math. Comput. *78* (2009), 1613–1627.
[23] *J. A. Ezquerro, M. A. Hernández, N. Romero*: Newton-type methods of high order and domains of semilocal and global convergence. Appl. Math. Comput. *214* (2009), 142–154.
[24] *W. B. Gragg, R. A. Tapia*: Optimal error bounds for the Newton-Kantorovich theorem. SIAM J. Numer. Anal. *11* (1974), 10–13.
[25] *M. A. Hernández*: A modification of the classical Kantorovich conditions for Newton's method. J. Comput. Appl. Math. *137* (2001), 201–205.
[26] *L. V. Kantorovich, G. P. Akilov*: Functional Analysis. Transl. from the Russian. Pergamon Press, Oxford, 1982.
[27] *S. Krishnan, D. Manocha*: An efficient surface intersection algorithm based on lower-dimensional formulation. ACM Trans. on Graphics. *16* (1997), 74–106.
[28] *G. Lukács*: The generalized inverse matrix and the surface-surface intersection problem. Theory and Practice of Geometric Modeling, Lect. Conf., Blaubeuren/FRG 1988. 1989, pp. 167–185.
[29] *J. M. Ortega, W. C. Rheinboldt*: Iterative Solution of Nonlinear Equations in Several Variables. Computer Science and Applied Mathematics. Academic Press, New York, 1970.
[30] *A. M. Ostrowski*: Sur la convergence et l'estimation des erreurs dans quelques procédés de résolution des équations numériques. Gedenkwerk D. A. Grave, Moskau, 1940, pp. 213–234. (In French.)
[31] *A. M. Ostrowski*: La méthode de Newton dans les espaces de Banach. (The Newton method in Banach spaces). C. R. Acad. Sci., Paris, Sér. A *272* (1971), 1251–1253. (In French.)

[32] *A. M. Ostrowski*: Solution of Equations in Euclidean and Banach Spaces. 3rd ed. of solution of equations and systems of equations. Pure and Applied Mathematics, 9. Academic Press, New York, 1973.
[33] *I. Păvăloiu*: Introduction in the Theory of Approximation of Equations Solutions. Dacia Ed. Cluj-Napoca, 1976.
[34] *F. A. Potra*: A characterization of the divided differences of an operator which can be represented by Riemann integrals. Math., Rev. Anal. Numér. Théor. Approximation, Anal. Numér. Théor. Approximation *9* (1980), 251–253.
[35] *F. A. Potra*: An application of the induction method of V. Pták to the study of regula falsi. Apl. Mat. *26* (1981), 111–120.
[36] *F. A. Potra*: The rate of convergence of a modified Newton's process. Apl. Mat. *26* (1981), 13–17.
[37] *F. A. Potra*: An error analysis for the secant method. Numer. Math. *38* (1982), 427–445.
[38] *F. A. Potra*: On the convergence of a class of Newton-like methods. Iterative solution of nonlinear systems of equations, Proc. Meeting, Oberwolfach 1982, Lect. Notes Math. 953, pp. 125–137.
[39] *F. A. Potra*: On the a posteriori error estimates for Newton's method. Beitr. Numer. Math. *12* (1984), 125–138.
[40] *F. A. Potra*: On a class of iterative procedures for solving nonlinear equations in Banach spaces. Computational Mathematics, Banach Cent. Publ. *13* (1984), 607–621.
[41] *F. A. Potra*: Sharp error bounds for a class of Newton-like methods. Libertas Math. *5* (1985), 71–84.
[42] *F. A. Potra, V. Pták*: Sharp error bounds for Newton's process. Numer. Math. *34* (1980), 63–72.
[43] *F. A. Potra, V. Pták*: Nondiscrete induction and a double step secant method. Math. Scand. *46* (1980), 236–250.
[44] *F. A. Potra, V. Pták*: On a class of modified Newton processes. Numer. Funct. Anal. Optimization *2* (1980), 107–120.
[45] *F. A. Potra, V. Pták*: A generalization of regula falsi. Numer. Math. *36* (1981), 333–346.
[46] *F. A. Potra, V. Pták*: Nondiscrete Induction and Iterative Processes. Research Notes in Mathematics, 103. Pitman Advanced Publishing Program, Boston, 1984.
[47] *P. D. Proinov*: General local convergence theory for a class of iterative processes and its applications to Newton's method. J. Complexity *25* (2009), 38–62.
[48] *P. D. Proinov*: New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems. J. Complexity *26* (2010), 3–42.
[49] *V. Pták*: Some metric aspects of the open mapping and closed graph theorems. Math. Ann. *163* (1966), 95–104.
[50] *V. Pták*: A quantitative refinement of the closed graph theorem. Czech. Math. J. *24* (1974), 503–506.
[51] *V. Pták*: A theorem of the closed graph type. Manuscr. Math. *13* (1974), 109–130.
[52] *V. Pták*: Deux théoremes de factorisation. C. R. Acad. Sci., Paris, Sér. A *278* (1974), 1091–1094.
[53] *V. Pták*: Concerning the rate of convergence of Newton's process. Commentat. Math. Univ. Carol. *16* (1975), 699–705.
[54] *V. Pták*: A modification of Newton's method. Čas. Pěst. Mat. *101* (1976), 188–194.
[55] *V. Pták*: Nondiscrete mathematical induction and iterative existence proofs. Linear Algebra Appl. *13* (1976), 223–238.
[56] *V. Pták*: The rate of convergence of Newton's process. Numer. Math. *25* (1976), 279–285.
[57] *V. Pták*: Nondiscrete mathematical induction. Gen. Topol. Relat. mod. Anal. Algebra IV, Proc. 4th Prague topol. Symp. 1976, Part A, Lect. Notes Math. 609. 1977, pp. 166–178.

[58] *V. Pták*: What should be a rate of convergence? RAIRO, Anal. Numér. *11* (1977), 279–286.
[59] *V. Pták*: Stability of exactness. Commentat. math., spec. Vol. II, dedic. L. Orlicz (1979), 283–288.
[60] *V. Pták*: A rate of convergence. Numer. Funct. Anal. Optimization *1* (1979), 255–271.
[61] *V. Pták*: Factorization in Banach algebras. Stud. Math. *65* (1979), 279–285.
[62] *H. Ren, Q. Wu*: Convergence ball of a modified secant method with convergence order 1.839... Appl. Math. Comput. *188* (2007), 281–285.
[63] *W. C. Rheinboldt*: A unified convergence theory for a class of iterative processes. SIAM J. Numer. Anal. *5* (1968), 42–63.
[64] *R. A. Tapia*: The Kantorovich theorem for Newton's method. Am. Math. Mon. *78* (1971), 389–392.
[65] *Q. Wu, H. Ren*: A note on some new iterative methods with third-order convergence. Appl. Math. Comput. *188* (2007), 1790–1793.
[66] *T. Yamamoto*: A convergence theorem for Newton-like methods in Banach spaces. Numer. Math. *51* (1987), 545–557.
[67] *P. P. Zabrejko, D. F. Nguen*: The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates. 9 (1987), 671–684.
[68] *A. I. Zinčenko*: Some approximate methods of solving equations with non-differentiable operators. Dopovidi Akad. Nauk Ukraïn. RSR (1963), 156–161. (In Ukrainian.)

Authors' addresses: Ioannis K. Argyros, Cameron University, Department of Mathematics Sciences, Lawton, OK 73505, USA, e-mail: iargyros@cameron.edu; Saïd Hilout, Poitiers University, Laboratoire de Mathématiques et Applications, Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France, e-mail: said.hilout@math.univ-poitiers.fr.