The Australian Journal of Mathematical Analysis and Applications

http://ajmaa.org

Volume 7, Issue 1, Article 2, pp. 1-11, 2010

AN IMPROVED MESH INDEPENDENCE PRINCIPLE FOR SOLVING EQUATIONS AND THEIR DISCRETIZATIONS USING NEWTON'S METHOD

IOANNIS K. ARGYROS

Received 1 August, 2007; accepted 9 September, 2008; published 3 March, 2010.

CAMERON UNIVERSITY, DEPARTMENT OF MATHEMATICS SCIENCES, LAWTON, OK 73505, USA
iargyros@cameron.edu

ABSTRACT. We improve the mesh independence principle [1] which states that when Newton's method is applied to an equation on a Banach space as well as to their finite–dimensional discretization there is a difference of at most one between the number of steps required by the two processes to converge to within a given error tolerance. Here using a combination of Lipschitz and center Lipschitz continuity assumptions instead of just Lipschitz conditions we show that the minimum number of steps required can be at least as small as in earlier works. Some numerical examples are provided whereas our results compare favorably with earlier ones.

Key words and phrases: Newton's method, Banach space, Lipschitz/center–Lipschitz conditions, Mesh independence principle, Discretized equation–iteration, Local–semilocal convergence.

2000 *Mathematics Subject Classification.* Primary 65B05, 65G99. Secondary 65J15, 65N30, 65N35, 47H17.

ISSN (electronic): 1449-5910

1. INTRODUCTION

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

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

where F is a continuously Fréchet–differentiable operator defined on a 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 (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 method for generating a sequence approximating x is undoubtedly Newton's method:

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

where $F'(x) \in (X, Y)$ $(x \in D)$, the space of bounded linear operators from X into Y, denotes the Fréchet–derivative of operator F [4], [7].
Sufficient conditions for the local and the semilocal convergence of Newton's method (1.2) have been given by many authors under various Lipschitz–type conditions [2]–[8].
A survey of such results can be found in [4], and the references therein.
The iterates x_n $(n \geq 1)$ can rarely be found in infinite–dimensional spaces. That is why in practice equation (1.1) is replaced by a family of discretized equations

$$\mathcal{P}_h(y) = 0 \tag{1.3}$$

indexed by some real number $h > 0$, where $\mathcal{P}_h$ is a nonlinear operator between finite dimensional space $\mathcal{X}_h$ and $\widehat{\mathcal{Y}}_h$. Assume the discretization on X be given by the bounded linear operators $L_h : X \longrightarrow \mathcal{X}_h$. Under certain assumptions, equation (1.3) have solutions $y_h^\star = L_h(x) + \mathcal{O}(h^p)$, found as the limit of Newton's method applied to (1.3) as follows:

$$y_0^h = L_h(x_0), \qquad y_{n+1}^h = y_n^h - \mathcal{P}_h'(y_n^h)^{-1}\, \mathcal{P}_h(y_n^h) \quad (n \geq 0). \tag{1.4}$$

The mesh independence principle shown in [1] state that, when Newton's method (1.2) is applied to nonlinear equation (1.1) as well as to (1.3), then the behavior of the discretized process (1.4) is asymptotically the same as that of (1.2), and consequently, the number of steps required by the two processes to converge to within a given tolerance is essentially the same. The importance of an efficient mesh size strategy base upon the mesh independence principle has been explained in [1], [2], [6] (see also [4]).
Here, motivated by optimization considerations, we show how to improve on the size of h (i.e.

enlarge h) under the same hypotheses and computational cost as in [1].
Our main idea is to introduce and employ the center–Lipschitz condition for the computation of the upper bounds of the inverses of the linear operators used in [1]. This idea has already been used by us in [3]–[5], to provide weaker sufficient conditions than the usual Newton–Kantorovich hypotheses [1], [2], [6]–[8] for the local as well as semilocal convergence of Newton's method (1.2) (see Theorems 2.1 and 2.2).
Numerical examples are also provided.

2. CONVERGENCE ANALYSIS

We will need the following semilocal and local convergence theorems whose proofs can be found in [4, p. 387, Case 3, for $\delta = \delta_0$], and [3], respectively.

Theorem 2.1. *Let $F\colon D \subseteq X \to Y$ be a Fréchet–differentiable operator. Assume:*
there exist a point $x_0 \in D$, and parameters $\eta \geq 0$, $K_0 > 0$, $K > 0$, $\sigma > 0$, such that

$$F'(x_0)^{-1} \in (Y,\ X), \quad \|F'(x_0)^{-1}\| \leq \sigma \tag{2.1}$$

$$\|F'(x_0)^{-1} F(x_0)\| \leq \eta, \tag{2.2}$$

$$\|F'(x) - F'(x_0)\| \leq K_0\ \|x - x_0\|, \tag{2.3}$$

$$\|F'(x) - F'(y)\| \leq K\ \|x - y\| \text{ for all } x, y \in D, \tag{2.4}$$

$$M\ \sigma\ \eta \leq 1, \tag{2.5}$$

and

$$\overline{U}(x_0,) = \{x \in X \mid \|x - x_0\| \leq \} \subseteq D, \tag{2.6}$$

where,

$$M = \frac{1}{4}\ (K + 4\ K_0 + \sqrt{K^2 + 8\ K_0\ K}),$$

$$a = \frac{2}{2-b}, \qquad b = \frac{1}{2}\left[-\frac{K}{K_0} + \sqrt{\left(\frac{K}{K_0}\right)^2 + 8\ \frac{K}{K_0}}\right]\sigma,$$

$$= \lim_{n\to\infty} t_n \leq a\ \eta, \tag{2.7}$$

$$t_0 = 0,\ t_1 = \eta,\ t_{n+2} = t_{n+1} + \frac{K\ \sigma\ (t_{n+1} - t_n)^2}{2\ (1 - K_0\ \sigma\ t_{n+1})}\ (n \geq 0). \tag{2.8}$$

Then, Newton's sequence $\{x_n\}$ $(n \geq 0)$ generated by (1.2) is well defined, remains in $\overline{U}(x_0,)$ for all $n \geq 0$, and converges to a unique solution $x \in \overline{U}(x_0,)$ of equation $F(x) = 0$. Moreover the following estimates hold for all $n \geq 0$

$$\|x_{n+2} - x_{n+1}\| \leq \frac{K\ \sigma\ \|x_{n+1} - x_n\|^2}{2\ [1 - K_0\ \sigma\ \|x_{n+1} - x_0\|]} \leq t_{n+2} - t_{n+1} \tag{2.9}$$

and

$$\|x_n - x\| \leq -t_n. \tag{2.10}$$

Furthermore, if there exists $R >$ such that

$$U(x_0, R) \subseteq D \tag{2.11}$$

and

$$\sigma\ K_0\ (+R) \leq 2, \tag{2.12}$$

then the solution is unique in $U(x_0, R)$.

Theorem 2.2. *Let $F\colon D \subseteq X \to Y$ be a Fréchet–differentiable operator. Assume:*

there exist a simple zero $x \in D$ of equation $F(x) = 0$, and parameters $\gamma > 0$, $\gamma_0 > 0$, $\sigma > 0$, such that:

$$\|F'(x)^{-1}\| \le \sigma$$

$$\|F'(x) - F'(y)\| \le \gamma \, \|x - y\| \quad \text{for all } x, y \in D \tag{2.13}$$

$$\|F'(x) - F'(x)\| \le \gamma_0 \, \|x - x\| \quad \text{for all } x \in D, \tag{2.14}$$

and

$$U^* = \overline{U}(x, \gamma_1) \subseteq D, \tag{2.15}$$

where

$$\gamma_1 = \frac{2}{(2\gamma_0 + \gamma)\, \sigma}\,. \tag{2.16}$$

Then, Newton's sequence $\{x_n\}$ $(n \ge 0)$ generated by (1.2) is well defined, remains in $\overline{U}(x, \gamma_1)$ for all $n \ge 0$ and converges to x. Moreover the following estimates hold for all $n \ge 0$:

$$\|x_{n+1} - x\| \le \frac{\gamma\ \sigma\ \|x_n - \|^2}{2\ (1 - \gamma_0\ \sigma\ \|x_n - x\|)}\,. \tag{2.17}$$

As in [1], [2], [4], let $S \subseteq X$ such that

$$x \in S^*,\ x_n \in S^*,\ x_n - x \in S^*,\ x_{n+1} - x_n \in S^* \quad (n \ge 0). \tag{2.18}$$

Consider the family

$$\{\mathcal{P}_h, L_h, \hat{L}_h\}, \quad h > 0, \tag{2.19}$$

where,

$$P_h\colon D_h \subset X_h \to \hat{Y}_h, \tag{2.20}$$

$$L_h\colon X \to X_h, \quad \hat{L}_h\colon Y \to \hat{Y}_h \tag{2.21}$$

such that:

$$L_h(S^* \cap U^*) \subseteq D_h. \tag{2.22}$$

The discretization family (2.19) is called: Lipschitz–center Lipschitz uniform if there exist $\rho > 0$, $K_0 > 0$, $K > 0$ such that

$$\overline{U}(L_h(x), \rho) \subseteq D_h, \tag{2.23}$$

$$\|\mathcal{P}_h'(u) - \mathcal{P}_h'(L_h(x))\| \le K_0\ \|u - L_h(x)\|, \quad u \in \overline{U}(L_h(x), \rho) \tag{2.24}$$

$$\|\mathcal{P}_h'(u) - \mathcal{P}_h'(v)\| \le K\ \|u - v\|, \quad u, v \in \overline{U}(L_h(x), \rho). \tag{2.25}$$

Moreover (2.19) is called: bounded if there exists a constant $q > 0$ such that

$$\|L_h(u)\| \le q\ \|u\|, \qquad u \in S^*, \tag{2.26}$$

stable: if there exists a $\sigma > 0$ such that

$$\|\mathcal{P}_h'(L_h(u))^{-1}\| \le \sigma, \quad u \in S^* \cap U^*, \tag{2.27}$$

and consistent of order p: if there exist $c_0 > 0$, $c_1 > 0$, $c_2 > 0$ such that

$$\|\hat{L}_h(F(x)) - \mathcal{P}_h(L_h(x))\| \le c_0\ h^p, \tag{2.28}$$

$$\|\hat{L}_h(F(x)) - \mathcal{P}_h(L_h(x))\| \le c_1\ h^p,\ \ x \in S^* \cap U^*, \tag{2.29}$$

and

$$\|\hat{L}_h(F'(x))(y) - \mathcal{P}'_h(L_h(x))L_h(y)\| \leq c_2\ h^p, \tag{2.30}$$

$x \in S^* \cap U^*$, $y \in S^*$.

We can show the following result relating (1.1), (1.2) with (1.3), (1.4) respectively.

Theorem 2.3. *Let $F\colon D \subseteq X \to Y$ be an operator satisfying hypotheses of Theorem 2.2 such that a Lipschitz, center–Lipschitz uniform discretization (2.19) exists which is bounded, stable and consistent of order p. Then,*

(a) equation (1.3) has a locally unique solution

$$y_h^* = L_h(x) + O(h^p), \tag{2.31}$$

for all h such that:

$$0 < h \leq h_0 = \min\left\{\left(\frac{\rho}{a\ c_0\ \sigma}\right)^{1/p}, \left(\frac{1}{c_0\ \sigma^2\ M}\right)^{1/p}\right\}; \tag{2.32}$$

(b) there exist $h_1 \in (0, h_0]$, $r_1 \in (0, r^*]$ such that Newton's method (1.4) converges to y_h^*; and for all $k \geq 0$

$$y_k^h \;=\; L_h(x_k) + O(h^p), \tag{2.33}$$

$$P_h(y_k^h) \;=\; \hat{L}_h(F(x_k)) + O(h^p) \tag{2.34}$$

$$y_k^h - y_h^* \;=\; L_h(x_k - x) + O(h^p) \tag{2.35}$$

for all $h \in (0, h_1]$ and $x_0 \in U(x, r_1)$.

Proof. We showed in Theorem 2.1 that when

$$\alpha(h) \;=\; M\ \sigma\|\mathcal{P}'_h(L_h(x))^{-1}\mathcal{P}_h(L_h(x))\| \leq 1, \tag{2.36}$$

$$r(h) \;\leq\; a\ \|\mathcal{P}'_h(L_h(x))^{-1}\mathcal{P}_h(L_h(x))\| \leq \rho, \tag{2.37}$$

then equation (1.3) has a unique solution y_h^* in $\overline{U}(L_h(x), r(h))$. Using (2.27), (2.28), (2.36), and (2.37) we get in turn

$$\begin{aligned} \alpha(h) &= M\ \sigma^2\|\mathcal{P}_h(L_h(x))\| \\ &= M\ \sigma^2\|\mathcal{P}_h(L_h(x)) - \hat{L}_h(F(x))\| \\ &\leq M\ \sigma^2\ c_0\ h^p \leq 1, \end{aligned} \tag{2.38}$$

and

$$r(h) \leq a\ \sigma\ c_0\ h^p \leq \rho, \tag{2.39}$$

which hold by the choice of h given by (2.32). Hence (2.31) follows from

$$\|y_h^* - L_h(x)\| \leq r(h) \leq a\ \sigma\ c_0\ h^p. \tag{2.40}$$

By Theorem 2.2, Newton's method (1.4) converges to y_h^*, if

$$\|L_h(x_0) - y_h^*\| < \frac{2}{(2K_0 + K)\ \|\mathcal{P}'_h(y_h^*)^{-1}\|}, \tag{2.41}$$

and

$$\overline{U}(y_h^*, \|L_h(x_0) - y_h^*\|) \subseteq \overline{U}(L_h(x), \rho). \tag{2.42}$$

Estimate (2.42) holds, if

$$\|y_h^* - L_h(x)\| + \|L_h(x_0) - y_h^*\| \leq \rho. \tag{2.43}$$

By (2.26) and (2.40) we can have

$$\begin{aligned} \|L_h(x_0) - y_h^*\| &\leq \|L_h(x_0) - L_h(x)\| + \|L_h(x) - y_h^*\| \\ &\leq q\ \|x_0 - x\| + a\ \sigma\ c_0\ h^p. \end{aligned} \tag{2.44}$$

Therefore (2.43) holds, if

$$q\ \|x_0 - x\| + 2\ a\ \sigma c_0\ h^p \leq \rho. \tag{2.45}$$

Using the identity and the Banach Lemma on invertible operators [4], [7]

$$\mathcal{P}_h'(y_h^*) = \mathcal{P}_h'(L_h(x))\big[I - \mathcal{P}_h'(L_h(x))^{-1}(\mathcal{P}_h'(L_h(x)) - \mathcal{P}_h'(y_h^*))\big], \tag{2.46}$$

we get

$$\|\mathcal{P}_h'(y_h^*)^{-1}\| \leq \frac{\|\mathcal{P}_h'(L_h())^{-1}\|}{1 - K_0\ \|\mathcal{P}_h'(L_h(x))^{-1}\|\ \|L_h(x) - y_h^*\|} \tag{2.47}$$

$$\leq \frac{\sigma}{1 - a\ K_0\ \sigma^2\ c_0\ h^p}. \tag{2.48}$$

Hence (2.41) holds if

$$q\ \|x_0 - x\| + 2\ a\ \sigma\ c_0\ h^p < \frac{2\ (1 - a\ K_0\ c_0\ \sigma^2\ h^p)}{(2\ K_0 + K)\ \sigma}. \tag{2.49}$$

Choose:

$$h_2 = \min\left\{\left(\frac{\rho}{4\ a\ c\ \sigma}\right)^{1/p}, \left[\frac{1}{2\ (3\ K_0 + K)\ \sigma^2\ a\ c}\right]^{1/p}\right\}, \quad c = \max\{c_0,\ c_1\}, \tag{2.50}$$

and

$$r_2 = \min\left\{\frac{\rho}{2\ q}, \frac{1}{(2\ K_0 + K)\ q\ \sigma}\right\}. \tag{2.51}$$

Then (2.41) and (2.42) hold for all $h \in (0, h_2]$ and $x_0 \in U(x, r_2)$. That is for these choices of h and x_0, Newton's method (1.4) converges to y_h^*. Define

$$h_1 = \min\left\{h_2, \left[\frac{1}{8\ \sigma^2\ (c_1 + c_2)\ (2\ K_0 + K)}\right]^{1/p}\right\}, \tag{2.52}$$

$$r_1 = \min\left\{r_2, \frac{1}{4\ K\ \sigma\ q}\right\}. \tag{2.53}$$

With the above choices it can easily be seen that the small root of the quadratic equation in λ

$$\frac{\sigma}{1 - K_0\ \sigma\ \lambda}\left[\frac{K}{2}\ \lambda^2 + 2\ K\ q\ \lambda\ \|x_0 - x\| + (c_1 + c_2)\ h^p\right] = \lambda \tag{2.54}$$

denoted by d is positive, and satisfies

$$d \leq 4\ \sigma\ (c_1 + c_2)\ h^p. \tag{2.55}$$

We now show using induction on n that for $h \in (0, h_1)$, $x_0 \in \overline{U}(x, r_1)$, and all $n \geq 0$

$$\|y_n^h - L_h(x_n)\| \leq d \tag{2.56}$$

holds.

For $n = 0$ (2.56) holds. Assume (2.56) holds for $n = 0, 1, \ldots, k$. Using (1.2) and (1.4) we obtain the identity

$$\begin{aligned} y_{k+1}^h - L_h(x_{k+1}) \quad = \quad & \mathcal{P}_h'(y_k^h)^{-1}\big\{[\mathcal{P}_h'(y_k^h)(y_k^h - L_h(x_i)) \\ & - \mathcal{P}_h(y_k^h) + \mathcal{P}_h(L_h(x_k))] \\ & + [\mathcal{P}_h'(y_k^h) - \mathcal{P}_h'(L_h(x_k))L_h(F'(x_k)^{-1}F(x_k))] \\ & + [\mathcal{P}_h'(L_h(x_k))L_h(F'(x_k)^{-1}F(x_k)) - \hat{L}_h(F(x_k))] \\ & + [\hat{L}_h(F(x_k)) - \mathcal{P}_h(L_h(x_k))]\big\}. \end{aligned} \tag{2.57}$$

As in (2.47) we get

$$\|\mathcal{P}_h'(y_k^h)^{-1}\| \le \frac{\sigma}{1 - K_0\ \sigma\ \|y_k^h - L_h(x_k)\|} \le \frac{\sigma}{1 - K_0\ \sigma\ d}\,. \tag{2.58}$$

We can get in turn using Taylor's formula, (2.25) and definition of d given in (2.54):

$$\begin{aligned} & \|\mathcal{P}_h'(y_k^h)(y_k^h - L_h(x_k)) - \mathcal{P}_h(y_k^h) + P_h(L_h(x_k))\| \\ & \le \| \int_0^1 \left[\mathcal{P}_h'(L_h(x_k) + t\,(y_k^h - L_k(x_k)) - \mathcal{P}_h'(y_k^h)\right] (y_k^h - L_h(x_k))\ dt\| \\ & \le \frac{K}{2}\|y_k^h - L_h(x_k)\|^2 \le \frac{K}{2}d^2. \end{aligned} \tag{2.59}$$

Using (2.25), and (2.26) we obtain:

$$\begin{aligned} & \|(\mathcal{P}_h'(y_k^h) - \mathcal{P}_h'(L_h(x_k))(L_h(F'(x_k)^{-1}F(x_k)))\| \\ & \le\ K\ q\ \|y_k^h - L_h(x_k)\|\ \|x_{k+1} - x_k\| \le 2\ K\ q\ d\ \|x_0 - x\|, \end{aligned} \tag{2.60}$$

(since $\|x_{k+1} - x\| \le \|x_k - x\|$)

$$\|\mathcal{P}_h'(L_h(x_k))L_h(F'(x_k)^{-1}F(x_k)) - \hat{L}_h(F(x_k))\| \le c_2\ h^p, \tag{2.61}$$

and

$$\|\hat{L}_h(F(x_i)) - \mathcal{P}_h(L_h(x_k))\| \le c_1\ h^p. \tag{2.62}$$

By (2.55) and (2.57)–(2.61) we get

$$\|y_{k+1}^h - L_h(x_{k+1})\| \le d \le 4\ \sigma\ (c_1 + c_2)\ h^p, \tag{2.63}$$

where d satisfies (2.54).
Moreover by the Lipschitz continuity of $\mathcal{P}_h'$ there exists b such that

$$\|\mathcal{P}_h'(x)\| \le b, \qquad x \in U(L_h(x), \rho). \tag{2.64}$$

Therefore we can have

$$\begin{aligned} & \|\mathcal{P}_h(y_k^h) - \hat{L}_h(F(x_k))\| \le \|\mathcal{P}_h(y_k^h) - \mathcal{P}_h(L_h(x_k))\| \\ & \quad + \|\mathcal{P}_h(L_h(x_k)) - \hat{L}_h(F(x_k))\| \le b\|y_k^h - L_h(x_k)\| + c_1 h^p \\ & \le\ 4\ \sigma\ b\ (c_1 + c_2)\ h^p + c_1\ h^p = c_3\ h^p, \end{aligned} \tag{2.65}$$

where

$$c_3 = 4\ \sigma\ b\ (c_1 + c_2) + c.$$

Furthermore by (2.40), (2.56), and (2.62) we get

$$\begin{aligned} \|y_k^h - y_h^* - L_h(x_k - x)\| \quad \le \quad & \|y_k^h - L_h(x_k)\| + \|y_h^* - L_h(x)\| \\ \le \quad & 4\ \sigma\ (c_1 + c_2)\ h^p + a\ \sigma\ c_0\ h^p = c\ h^p, \end{aligned} \tag{2.66}$$

where

$$c = \sigma\ [4\ (c_1 + c_2) + a\ c].$$

That completes the proof of Theorem 2.3. □

The following result is the second part of the mesh independence principle.

Theorem 2.4. *Assume:*

(a) *hypotheses of Theorem 2.3 hold;*
(b) *there exists $\delta > 0$ such that*

$$\liminf_{h>0} \|L_h(x)\| \geq \delta \; \|x\| \quad \textit{for } x \in S^*. \tag{2.67}$$

Then for some $\overline{r} \in (0, r_1]$ and any fixed $\varepsilon > 0$, $x_0 \in \overline{U}(x, \overline{r})$ there exists $\overline{h} = \overline{h}(\varepsilon, x_0) \in (0, h_1]$ such that

$$|\min\{n \geq 0, \|x_n - x\| < \varepsilon\} - \min\{n \geq 0, \|y_n^h - y_h^*\| < \varepsilon\}| \leq 1 \tag{2.68}$$

for all $h \in (0, \overline{h}]$.

Proof. Let k be the unique integer satisfying

$$\|x_{k+1} - x\| < \varepsilon \leq \|x_k - x\|, \tag{2.69}$$

and $h_3 > 0$ (depending on x_0) such that

$$\|L_h(x_k - x)\| \geq \delta \; \|x_k - x\| \quad \text{for all } h \in (0, h_3). \tag{2.70}$$

Define

$$\overline{r} = \max\left\{r_1, \frac{\beta}{2 \; \sigma \; q \; (K + \beta \; K_0)}\right\}, \; \beta = \min\{\frac{1}{q}, \delta, 2 \; q\}, \tag{2.71}$$

and

$$\overline{h} = \min\left\{h_1, h_3, \left[\frac{\beta}{2 \; \sigma \; c \; (K + K_0 \; \beta)}\right]^{1/p}, \left(\frac{\delta \; \varepsilon}{2 \; c}\right)^{1/p}\right\}. \tag{2.72}$$

By (2.65) and (2.71) we can get

$$\|y_{k+1}^h - y_h^*\| \leq \|L_h(x_{k+1} - x)\| + c \; h^p \leq q\varepsilon + \frac{\beta \; \varepsilon}{2} < 2 \; q \; \varepsilon. \tag{2.73}$$

Moreover from Theorem 2.2 we get

$$\begin{aligned} \|y_{k+2}^h - y_h^*\| &\leq \frac{K\sigma \|y_{k+1}^h - y_h^*\|^2}{2 \; [1 - K_0 \; \sigma \; \|y_{k+1}^h - y_h^*\|]} \\ &\leq \frac{K \; \sigma \; \|y_0^h - y_h^*\|}{2 \; (1 - K_0 \; \sigma \; \|y_0^h - y_h^*\|)} \|y_{k+1} - y_h^*\| \\ &< \frac{K \; \sigma \; (q \; \overline{r} + c \; \overline{h}^p)}{1 - K_0 \; \sigma \; (q \; \overline{r} + c \; \overline{h}^p)} \; q \; \varepsilon \leq b \; q \; \varepsilon < \varepsilon. \end{aligned} \tag{2.74}$$

By (2.65) and (2.69)

$$\varepsilon \leq \|x_k - x\| \leq \frac{1}{\delta} \; \|L_h(x_k - x)\| \leq \frac{1}{\delta}(\|y_k^h - y_h^*\| + c\overline{h}^p), \tag{2.75}$$

or

$$\|y_k^h - y_h^*\| \geq \delta \; \varepsilon - c \; \overline{h}^p \geq \delta \; \varepsilon - \frac{\delta \; \varepsilon}{2} = \frac{\delta \; \varepsilon}{2}. \tag{2.76}$$

Furthermore if $\|y_{k-1}^h - y_h^*\| < \varepsilon$ as in (2.73) we get

$$\|y_k^h - y_h^*\| < \frac{1}{2} \; \beta\varepsilon \leq \frac{\delta \; \varepsilon}{2} \tag{2.77}$$

contradicting (2.75). Hence we get

$$\|y^h_{k-1} - y^*_h\| \geq \varepsilon. \tag{2.78}$$

The result now follows from (2.68), (2.73) and (2.77).
That completes the proof of Theorem 2.4. ∎

Remark 2.1. (a) Theorem 2.1 reduces to the famous semilocal Newton–Kantorovich theorem [7], when $K = K_0$. In general

$$K_0 \leq K \tag{2.79}$$

holds. If $K_0 < K$, then our hypothesis (2.5) may be satisfied, but the Newton–Kantorovich hypothesis

$$2\ K\ \eta \leq 1 \tag{2.80}$$

may not. In this case the error bounds on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x\|$ $(n \geq 0)$ are finer and the information on the location of the solution more precise (see also [3]).

(b) Theorem 2.2 reduces to Theorem 1 in [1] when $\gamma = \gamma_0$. In general

$$\gamma_0 \leq \gamma, \tag{2.81}$$

and in the case when strict inequality holds in (2.80), then our convergence radius γ_1 given by (2.16) is larger than the corresponding one given in [1] by

$$\gamma_2 = \frac{2}{3\gamma}\,. \tag{2.82}$$

Hence we have a wider choice of initial guesses x_0, and our error bounds on the distances $\|x_n - x\|$ $(n \geq 0)$ are finer.

(c) Theorems 2.3, 2.4 reduce to Theorem 2 and Corollary 1 in [1], when

$$K = K_0 \quad \text{and} \quad c_0 = c_1. \tag{2.83}$$

Note though that

$$K_0 \leq K \tag{2.84}$$

and

$$c_0 \leq c_1. \tag{2.85}$$

In case (2.83) or (2.84) holds as strict inequalities then it is clear that our smallest integer n_1 satisfying $\|x_n - x\| < \varepsilon$ is smaller than the corresponding integer n_2 given in references mentioned above. Hence we require less computational steps to achieve the same error tolerance ε than before. The ratios in relationships (2.33)–(2.35) are also finer.

Define $\overline{h_0}$, $\overline{h_1}$, $\overline{h_2}$, $\overline{h_3}$, $\overline{\overline{h}}$ used in [1] as h_0, h_1, h_2, h_3, $\overline{h}$ respectively by setting $K = K_0$. Then, we have $\overline{h_0} \leq h_0$, $\overline{h_1} \leq h_1$, $\overline{h_2} \leq h_2$, $\overline{h_3} \leq h_3$, $\overline{\overline{h_0}} \leq \overline{h}$. If $K_0 < K$, then strict inequality can hold. Hence, the mesh size h has been enlarged.

Note that the improvements made through our Theorems 2.1–2.4 are achieved under the same hypotheses as before, since in practice, the computation of K requires that of K_0. Note also that $\dfrac{K}{K_0}$ and $\dfrac{\gamma}{\gamma_0}$ can be arbitrarily large [3]–[5].

Remark 2.2. If (2.66) is replaced by the stronger but standard in most discretization studies condition

$$\lim_{h \to 0} \|L_h(x)\| = \|x\| \text{ uniformly for } x \in S^*, \tag{2.86}$$

then Theorem 2.4 still holds but $\overline{h}_1$ does not depend on x_0. Note also that (2.66) follows from (2.85).

Remark 2.3. As noted in [1]–[8] 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 problems where the trapezoidal rule, the box scheme and allocation methods for boundary value problems are involved.

Remark 2.4. 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 (see Argyros [5], Kantorovich et al. [7]):

$$F'(x) = T(F(x)), \tag{2.87}$$

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 (1.1).

We complete our study with two numerical examples. In Example 2.5 we show that under the same hypotheses as in Theorem 1 of [1] we can obtain a larger radius of convergence. Whereas in Example 2.6 we show that $\frac{K}{K_0}$ can be arbitrarily large.

Example 2.5. Let $X = Y =$, $D = U(0,1)$, and define function F on D by

$$F(x) = e^x - 1. \tag{2.88}$$

Then it can easily be seen that we can set $T(x) = x + 1$ in (2.86). Using (2.87) we get $\gamma = e$ and $\gamma_0 = e - 1$. Hence (2.16), (2.81) give

$$\gamma_2 = .245252961 < .324947231 = \gamma_1. \tag{2.89}$$

(See also Remark 2.1 (b), (c).) Hence Theorem 1 in [1] cannot guarantee the convergence of (1.2) to $x = 0$ when $x_0 \in [\gamma_2, \gamma_1)$ but our Theorem 2.2 can.

Example 2.6. Let $X = Y =$ and define function F on by

$$F(x) = \theta_0 x + \theta_1 + \theta_2 \sin e^{\theta_3 x}, \tag{2.90}$$

where θ_i, $i = 0, 1, 2, 3$ are given parameters. Using (2.89) it can easily be seen that for θ_3 large and θ_2 sufficiently small $\frac{K}{K_0}$ may be arbitrarily large. That is (2.5) may be satisfied but not (2.79), in which case Theorem 2.2 cannot apply.

REFERENCES

[1] E.L. ALLGOWER, K. BÖHMER, F.A. POTRA and W.C. RHEINBOLDT, A mesh independence principle for operator equations and their discretizations, *SIAM J. Numer. Anal.*, **23** (1986), pp. 160–169.

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

[3] I.K. ARGYROS, On the Newton–Kantorovich hypothesis for solving equations, *J. Comp. Appl. Math.*, **11** (2004), pp. 103–110.

[4] 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), pp. 374–397.

[5] I.K. ARGYROS, *Computational Theory of Iterative Methods*, **Studies in Computational Mathematics**, 15, Elsevier, 2007, New York, U.S.A.

[6] P.N. BROWN, A local convergence theory for combined inexact–Newton/finite–difference projection methods, *SIAM J. Numer. Anal.*, **24** (1987), pp. 407–434.

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

[8] G.M. VAINIKKO, Galerkin's perturbation method and the general theory for approximate methods for non-linear equations, *USSR Comp. Math. and Math. Phys.*, **4** (1967), pp. 723–751.