Canad. Math. Bull. Vol. **67** (3), 2024, pp. 781–795
http://dx.doi.org/10.4153/S000843952400016X

On the complexity of extending the convergence domain of Newton's method under the weak majorant condition

Ioannis K. Argyros and Santhosh George

Abstract. The local analysis of convergence for Newton's method has been extensively studied by numerous researchers under a plethora of sufficient conditions. However, the complexity of extending the convergence domain requires very general conditions such as the ones depending on the majorant principle in order to include as large classes of operators as possible. In the present article, such an analysis is developed under the weak majorant condition. The new results extend earlier ones using similar information. Finally, the numerical examples complement the theory.

1 Introduction

Let E_1 and E_2 denote complete normed spaces and $D \subset E_1$ be a nonempty, open, and convex set, and $G : D \subseteq E_1 \longrightarrow E_2$ be a Fréchet differentiable operator. The determination of a locally unique solution $x^* \in D$ for the nonlinear equation of the form

$$G(x) = 0 \tag{1.1}$$

is very challenging and of extreme importance. This is indeed the case, since many applications from diverse disciplines such as Mathematical: Biology; Chemistry; Ecology; Economics; Physics; Scientific Computing; and Engineering can be written in a form like (1.1) using Mathematical Modeling [4–6, 8, 16]. However, the analytical version of the solution x^* is hard or expensive and it can be found only in special cases. That explains the reason why most practitioners and researchers develop iterative method generating a sequence approximating x^* under certain conditions imposed on the initial data.

Newton's method is defined for each $n = 0, 1, 2, \ldots$ by

$$x_0 \in D,\ x_{n+1} = x_n - G'(x_n)^{-1}G(x_n). \tag{1.2}$$

There is a plethora of studies about Newton's method [3, 5, 18, 20]. Practical applications in convex programming can be found in [16]. Other applications can be found in [4, 6, 15, 17, 19, 24]. A usual hypothesis in such studies is some type of Lipchitz, Hölder, or majorant condition on G'. Such hypotheses are important, since they allow some

Received by the editors September 18, 2023; revised January 27, 2024; accepted February 26, 2024.
Published online on Cambridge Core March 1, 2024.
AMS subject classification: **65K15**, **90C30**, **49M15**.
Keywords: Newton's methods, weak majorant condition, Banach space, convergence domain.

control on the derivative, which is important in both the local as well as the semi-local analyses of convergence of iterative methods. The local convergence analysis results are important, since they provide the degree of difficulty in choosing the starting points x_0 so that the sequence $\{x_n\}$ is convergent to x^*.

In particular, we are motivated by the elegant work on the local analysis of convergence for Newton's method in [9] (see also [10–13, 25, 27]) using the majorant principle and optimization considerations.

The new local analysis of convergence uses weaker majorant conditions resulting to the following advantages:

Novelty

(a_1) A larger radius of convergence. That allows a wider selection of starting points x_0 ensuring the convergence of the sequence $\{x_n\}$ to x^*.

(a_2) Tighter upper error bounds on the distances $\|x_n - x^*\|$. This way we use fewer iteration to achieve a predetermined error tolerance denoted by $\varepsilon > 0$. That is, there exists N such that for each $n \geq N$, $\|x_n - x^*\| \leq \varepsilon$.

(a_3) A larger than before domain is found containing x^* as the only solution of the equation $G(x) = 0$.

It is worth noting that such advantages extend the number of classes of equations that can be solved using Newton's method, the advantages $(a_1) - (a_3)$ are obtained under the same computational effort, since in practice the new and tighter majorant functions are special cases of the one used in [9]. The same advantages can be obtained in other studies, e.g., for solving generalized equations using Newton's method [1, 2, 8, 11, 13, 22] as long as they use the majorant conditions.

The rest of the article contains the following: The preliminaries in Section 2 followed by the properties of the majorant functions in Section 3. The local analysis of convergence for Newton's method is developed in Section 4. Finally, the special cases and examples can be found in the concluding Section 5.

2 Preliminaries

Let E_1 and E_2 stand for complete normed spaces and $L(E_1, E_2)$ denote the space of continuous linear operators sending E_1 into E_2. Moreover, define the open ball $U(z, \mu)$ and its closure $U[z, \mu]$, respectively, for $\mu > 0$ as

$$U(z, \mu) = \{v \in E_1 : \|z - v\| < \mu\}$$

and

$$U[z, \mu] = \{v \in E_1 : \|z - v\| \leq \mu\}.$$

A standard auxiliary result on the inverses of operator that are linear is useful.

Lemma 2.1 *(Banach's perturbation lemma) Assume that $T \in L(E_1, E_2)$ and satisfies $\|T - I\| < 1$, where $I : E_1 \longrightarrow E_1$ denotes the identity operator. Then, the following items are valid: the linear operator T is invertible and*

$$\|T^{-1}\| \leq \frac{1}{1 - \|T - I\|}.$$

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

Proof Set $A = I$ and $c = \|T - I\|$ in Lemma 1 of [23, p. 189]. ∎

Some basic properties of convex functions are needed. More information about such functions can be found in [1–3, 7, 14].

Lemma 2.2 *Let $R > 0$ be a given constant. Assume that $\psi : [0, R) \longrightarrow (-\infty, +\infty)$ is convex and differentiable. Then, the following items are valid:*

(i)

$$\frac{\psi(s) - \psi(\theta s)}{s} \leq \psi'(s)(1 - \theta)$$

for each $s \in (0, R)$ and $\theta \in [0, 1]$;

(ii)

$$\frac{\psi(u_1) - \psi(\theta u_1)}{u_1} \leq \frac{\psi(u_2) - \psi(\theta u_2)}{u_2}$$

for each $u_1, u_2 \in [0, R), u_1 < u_2$ and $\theta \in [0, 1]$.

Proof See Theorem 4.1.1 and Remark 4.1.2 in [14, p. 21]. ∎

We need a relationship between different types of majorant conditions. Let $\bar{\rho} = \sup\{s \in [0, R) : U(x^*, s) \subset E_1\}$.

Definition 2.1 A function $h_0 : [0, R) \longrightarrow (-\infty, +\infty)$ which is twice continuously differentiable is said to be a center-majorant function for G on $U(x^*, \bar{\rho})$, if for each $y \in U(x^*, \bar{\rho})$

$$\|M^{-1}(G'(y) - M)\| \leq h_0'(\|y - x^*\|) - h_0'(0)$$

for some operator $M \in L(E_1, E_2)$ which is invertible, independent of y, and may or may not depend on x^* (see also Remark 2.3(iv)).

(A1) The function h_0' is convex and strictly increasing.

(A2) $h_0(0) = 0$ and $h_0'(0) = -1$.

M is chosen to be any linear invertible operator satisfying this condition (see also Remark 2.3 and Section 5).

Let $\rho \in [0, \bar{\rho}]$ be such that $\rho = \sup\{s \in [0, \bar{\rho}) : h_0'(s) < 0\}$.

Definition 2.2 A function $h : [0, \rho) \longrightarrow (-\infty, +\infty)$ which is twice continuously differentiable is said to be a restricted-majorant function for G on $U(x^*, \rho)$, if for each $\theta \in [0, 1]$, $y \in U(x^*, \rho)$

$$\|M^{-1}(G'(y) - G'(x^* + \theta(y - x^*)))\| \leq h'(\|y - x^*\|) - h'(\theta\|y - x^*\|).$$

Notice that h depends on the function h_0.

(A3) The function h' is convex and strictly increasing.

(A4) $h(0) = 0$ and $h'(0) = -1$.

Notice that the function h_0 depends on x^* and $\bar{\rho}$, whereas the function h on x^*, ρ, and h_0.

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

Definition 2.3 A function $h_1 : [0, R) \longrightarrow (-\infty, +\infty)$ which is twice continuously differentiable is said to be a majorant function for G on $U(x^*, \rho)$, if for each $\theta \in [0,1]$, $y \in U(x^*, \bar{\rho})$

$$\|M^{-1}(G'(y) - G'(x^* + \theta(y - x^*)))\| \leq h_1'(\|y - x^*\|) - h_1'(\theta\|y - x^*\|).$$

(A3)' The function h_1' is convex and strictly increasing.
(A4)' $h_1(0) = 0$ and $h_1'(0) = -1$.

Remark 2.3 (i) It follows by these definitions that

$$h_0'(s) \leq h_1'(s) \text{ and } h'(s) \leq h_1'(s) \text{ for each } s \in [0, \rho). \tag{2.1}$$

(ii) Thus, the results in the literature using only h_1 (for $M = G'(x^*)$) (see [9]) can be replaced by the pair (h_0, h) resulting to finer error distances, a larger convergence radius, and a more precise and larger uniqueness radius for the solution x^*. These advantages are obtained under the same computational cost, since in practice the computation of the function h_1 requires that of h_0 and h as special cases.

(iii) A popular (but not the most flexible) choice for $M = G'(x^*)$. In this case, the solution is simple. However, assumptions allow the determination of a solution that is not necessarily simple.

(iv) The choice of the initial point can be chosen without the actual knowledge of x^*. Suppose that the operator G satisfies the autonomous differential equation [3, 18]:

$$G'(x) = P(G(x)),$$

where P is a continuous operator. By this definition, we obtain $G'(x^*) = P(G(x^*)) = P(0)$, for any solution x^* of the equation $G(x) = 0$. As an example, define $G(x) = e^x - 1$ and choose $P(x) = x + 1$. Then, for the operator P satisfies $G'(x) = P(G(x))$. Therefore, we can choose $M = G'(x^*)$ and $M = P(G(x^*))$ or $M = P(0)$, which is known although x^* is not known.

Definition 2.4 We define the parameters

$$c_1 = \sup\{s \in [0, \rho) : h_0'(s) < 0\},$$

$$c_2 = \sup\left\{s \in [0, c_1) : \frac{h(s) - sh'(s)}{sh_0'(s)} < 1\right\},$$

and

$$c_3 = \sup\{s \in (0, c_2) : h_0(s) < 0\}.$$

Moreover, define the Newton iteration for solving the equation $h(s) = 0$ given by

(2.2)

$$s_0 = |x_0 - x^*|,\ s_1 = \left|\frac{s_0 h_0'(s_0) - h_0(s_0)}{h_0'(s_0)}\right|,\ s_{n+1} = \left|\frac{s_n h'(s_n) - h(s_n)}{h_0'(s_n)}\right| \text{ for each } n = 1, 2, \ldots.$$

Define the radius

$$\rho^* = \min\{\rho, c_2\} \tag{2.3}$$

and the parameter

$$c_4 = \min\{\rho, c_3\}.$$

Next, the main local analysis convergence result for Newton's method is stated.

Theorem 2.4 *Assume that the conditions (A1)–(A5) are valid with the radius ρ^* as defined in (2.4). Then, if the starter $x_0 \in U(x^*, \rho^*) - \{x^*\}$, the following assertions hold:*

(i) *The scalar majorant sequence $\{s_n\}$ given by the formula (2.2) converges to zero, belongs in the interval $(0, \rho^*)$, and the sequence $\{\frac{s_{n+1}}{s_n^2}\}$ is strictly decreasing.*

(ii) *The sequence $\{x_n\} \subset U(x^*, \rho^*)$ converges Q-quadratically, to x^* so that*

$$\|x^* - x_{n+1}\| \le \frac{s_{n+1}}{s_n^2}\|x_n - x^*\|^2, \text{ and } \frac{s_{n+1}}{s_n^2} \le \frac{h''(s_0)}{2|h_0'(s_0)|} \tag{2.4}$$

for each $n = 0, 1, 2, \dots$.

(iii) *Additionally, if for $\alpha \in (0, \rho)$, $\frac{h(\alpha)}{(\alpha h_0'(\alpha))-1} = 1$, then $\rho^* = \alpha$ is the largest convergence radius. Moreover, x^* is the unique solution of equation (1.1) in the ball $U(x^*, c_4)$.*

3 The properties of the majorant functions

The parameters $\bar{\rho}, c_1, c_2, c_3$ related to the majorant functions h_0, h and the domain D are shown to exist and be positive in the auxiliary results that follow in this section. Moreover, the properties of the sequence $\{s_n\}$ which appears in Theorem 2.4 are investigated.

Lemma 3.1 *The following items are valid:*

$$\bar{\rho} > 0, c_1 > 0, c_3 > 0$$

and

$$\frac{sh'(s) - h(s)}{h_0'(s)} < 0 \tag{3.1}$$

for each $s \in (0, c_1)$.

Proof By hypothesis $x^* \in D$ which is an open set. Thus, it follows that $\bar{\rho} > 0$. By the second condition in (A_2), $h_0'(0) = -1$, there exists a parameter $\gamma > 0$ so that $h_0'(s) < 0$ for each $s \in (0, \gamma)$ leading us to deduce that $c_1 > 0$. Moreover, by the condition (A_2), $h_0(0) = 0$ and $h_0'(0) = -1$. Then, there exists a parameter $\gamma > 0$ so that $h_0(s) < 0$ for each $s \in (0, \gamma)$, so $c_3 > 0$. By the conditions (A_3), (A_4) and Lemma 2.2,

$$h(s) - sh'(s) < h(0) = 0 \tag{3.2}$$

for each $s \in (0, \bar{\rho})$. But if $s \in (0, c_1)$, then $h_0'(s) < 0$, showing (3.1) by (3.2.) ■

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

It follows by the condition (A_1) and the definition of the parameter c_1 that the real function

$$\psi_{h_0,h} : [0, c_1) \to (-\infty, 0], s \to \frac{sh'(s) - h(s)}{h_0'(s)} \tag{3.3}$$

exists in the interval $[0, c_1)$.

Lemma 3.2 *The following items are valid.*
The function $\frac{|\psi_{h_0,h}(s)|}{s^2}$ is strictly increasing and

$$\frac{|\psi_{h_0,h}(s)|}{s^2} \le \frac{h''(s)}{2|h_0'(s)|} \tag{3.4}$$

for each $s \in (0, c_1)$.

Proof The definition of the function $\psi_{h_0,h}$, $h_0(0) = 0$ and $h_0'(s) < 0$ for each $s \in [0, c_1)$ imply that

$$\begin{aligned}\frac{|\psi_{h_0,h}(s)|}{s^2} &= \left|\frac{sh'(s) - h(s)}{s^2 h_0'(s)}\right| \\ &= \frac{1}{|h_0'(s)|}\left|\frac{sh'(s) - h(s)}{s^2}\right|, \\ &= \frac{1}{h_0'(s)}\int_0^1 \frac{h'(s) - h'(\theta s)}{s} d\theta \\ &\le \frac{1}{|h_0'(s)|}\int_0^1 h''(t)(1-\theta)d\theta. \end{aligned} \tag{3.5}$$

Thus, (3.4) holds, where we also used Lemma 2.2 to deduce that the function

$$s \to \frac{h'(s) - h'(\theta s)}{s}$$

for each $s \in (0, c_1)$, $\theta \in (0, 1)$ is strictly increasing, and positive as well as the function $s \to \frac{1}{|h_0'(s)|}$ is strictly increasing. This makes the right-hand side of (3.5) positive. ∎

Lemma 3.3 *The following items are valid:*

$$c_2 > 0 \tag{3.6}$$

and

$$|\psi_{h_0,h}(s)| < s \tag{3.7}$$

for each $s \in (0, c_2)$.

Proof By the last two lemmas,

$$\frac{|\psi_{h_0,h}(s)|}{s} = \frac{h(s) - sh'(s)}{h_0'(s)} > 0, \tag{3.8}$$

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

and the function $\frac{|\psi_{h_0,h}(s)|}{s^2}$ is bounded close enough to zero. Thus, we can write

$$\lim_{s\to 0}\frac{|\psi_{h_0,h}(s)|}{s} = \lim_{s\to 0}\left(\frac{|\psi_{h_0,h}(s)|}{s^2}\right)s = 0. \tag{3.9}$$

Hence, by the estimates (3.8) and (3.9), there exists a parameter $\gamma > 0$ so that

$$0 < \frac{h(s) - sh'(s)}{h_0'(s)} < 1.$$

Consequently, by the definition of the parameter c_2, and the preceding double inequality, we get $c_2 > 0$. Moreover, Lemma 3.2 gives that the function $s \to \frac{|\psi_{h_0,h}(s)|}{s^2}$ is strictly increasing for each $s \in (0, c_1)$. Furthermore, the definition of the parameter c_2 implies (3.7) for each $s \in (0, c_2)$. ■

It follows by (3.4) that the scalar sequence $\{s_n\}$ can also be given as

$$0 < s_0 = \|x^* - x_0\|, s_{n+1} = |\psi_{h_0,h}(s_n)| \tag{3.10}$$

for each $n = 0, 1, 2, \ldots$.

Lemma 3.4 *The following items are valid: The scalar sequence $\{s_n\}$ given in (3.10) exists in the interval $(0, c_2)$ for each $n = 0, 1, 2, \ldots$;*

$$s_{n+1} < s_n, \tag{3.11}$$

i.e., the sequence $\{s_n\}$ is strictly decreasing; the sequence $\{\frac{s_{n+1}}{s_n^2}\}$ is strictly decreasing,

$$\lim_{s\to+\infty} s_n = 0 \tag{3.12}$$

and

$$\frac{s_{n+1}}{s_n^2} \le \frac{h''(s_0)}{2|h_0'(s_0)|}. \tag{3.13}$$

Proof Lemma 3.3 and the definition of the sequence $\{s_n\}$ given in (3.10) imply by a simple inductive argument that $s_{m+1} = |\psi_{h_0,h}(s_m)| < s_m$, where we also used the restriction the first formula in (3.10) to deduce $s_0 < \rho^* \le c_2$. Thus, the sequence $\{s_m\}$ exists in the interval $(0, c_2)$, remains in $(0, c_2)$, and satisfies (3.11). Moreover, by the definition (3.10),

$$\frac{s_{m+1}}{s_m^2} = \frac{|\psi_{h_0,h}(s_m)|}{s_m^2} \le \frac{|\psi_{h_0,h}(s_0)|}{s_0^2}, \tag{3.14}$$

since by Lemma 3.2, the sequence $\{\frac{s_{m+1}}{s_m^2}\}$ is strictly decreasing. Then, the estimate (3.14) and $s_m < s_0$ imply

$$s_{m+1} \le cs_m, m = 0, 1, 2, \ldots, \tag{3.15}$$

where $c = \frac{|\psi_{h_0,h}(s_0)|}{s_0} \in [0, 1)$ by Lemma 2.3.

Consequently, by (3.15), we deduce the validity of the item (3.12) the second item in Lemma 2.2.

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

Furthermore, we obtain

$$\frac{|\psi_{h_0,h}(s_0)|}{s_0^2} \leq \frac{|h''(s_0)|}{2|h_0'(s_0)|}. \tag{3.16}$$

Finally, the item (3.13) follows by the estimates (3.14) and (3.16). ∎

Remark 3.5 Let us show the implications of the new conditions $(A_1) - (A_2)$, when compared to the old ones $(A_3)'$ and $(A_4)'$, given in [9] for the specializations

$$h_0(s) = \frac{l_0}{2}s^2 - s, h(s) = \frac{l}{2}s^2 - s$$

and

$$h_1(s) = \frac{l_1}{2}s^2 - s$$

for some parameters $l_0 > 0, l > 0$ and $l_1 > 0$. Notice that

$$l_0 \leq l_1 \tag{3.17}$$

and

$$l \leq l_1. \tag{3.18}$$

Case 1: $l_0 = l = l_1$ (Lipschitz). Then, the definitions of the parameters c_2 and ρ^* imply that

$$\rho_1^* = \min\left\{\bar{\rho}, \frac{2}{3l_1}\right\}, \tag{3.19}$$

where in order to obtain the second element in the preceding set, we solved for the variable s the inequality

$$\frac{\frac{l_1}{2}s^2 - s - s(l_1 s - 1)}{s(l_0 s - 1)} < 1.$$

Case 2: By the condition (A_5),

$$l_0 \leq l. \tag{3.20}$$

Then, the corresponding radius is

$$\rho^* = \min\left\{\bar{\rho}, \frac{2}{2l_0 + l}\right\}, \tag{3.21}$$

where we again solved for s the inequality

$$\frac{\frac{l}{2}s^2 - s - s(ls - 1)}{s(l_0 s - 1)} < 1.$$

Clearly, it follows by (3.17)–(3.20) that

$$\rho_1^* \leq \rho^*. \tag{3.22}$$

Notice that by (3.19) and (3.21) as $\frac{\ell}{\ell_1} \longrightarrow 0$, the new radius is at least as three times larger. The value ρ_1^* is due to Rheinboldt [21] and Traub [24].

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

In the numerical section, examples are developed where (3.17), (3.18), (3.20), and (3.22) are valid as strict inequalities. Moreover, under Case 1, the corresponding sequence $\{u_n\}$ to $\{s_n\}$ is

$$u_{n+1} = \left| \frac{l_1 u_n^2}{2(1 - l_1 u_n)} \right|,$$

for $n = 0, 1, 2, \ldots$,

whereas

$$s_{n+1} = \left| \frac{(2l_0 - l)s_n^2}{2(1 - l_0 s_n)} \right|,$$

for $n = 0, 1, 2, \ldots$. Therefore, by (3.17), (3.18), and these definitions,

$$0 \leq s_{n+1} \leq u_{n+1}, \tag{3.23}$$

for each $n = 0, 1, 2, \ldots$.

Hence, we conclude that the new error bounds $\{s_n\}$ are tighter than the old ones $\{u_n\}$ given in [9]. Notice also that these advantages hold even if h_0, h, and h_1 are chosen in ways other than the Lipschitz, since h_0 and h are always tighter than h_1. Moreover, the aforementioned advantages require no additional computational effort, since in practice the calculation to determine the function h_1 also requires that of the functions h_0 and h as special cases. Moreover, there are advantages concerning the uniqueness of the solution x^* (see Remark 3.5). Furthermore, notice that the conditions on c_1, c_2, c_3, ρ^* are always weaker than the corresponding ones given in [9] for $h_0 = h = h_1$. That is, if the old convergence conditions hold, so do the new ones but not necessarily vice versa. Finally, notice that all other results involving the sequence $\{u_n\}$ and the function can be replaced by $\{s_n\}$ and the functions h_0 and h (see [9] and the references therein).

4 Local Analysis

In this section, we present the association of the majorant functions h_0 and h to the nonlinear operator F. Next, we first present a Banach perturbation lemma of inverses for linear operators.

Lemma 4.1 *Under the hypotheses of Lemmas 3.1–3.4, further assume that*

$$x \in U(x^*, \min\{\bar{\rho}, c_1\}). \tag{4.1}$$

Then, the following items are valid: $G'(x)^{-1} \in L(E_2, E_1)$ *and*

$$\|G'(x)^{-1}M\| \leq \frac{1}{|h_0'(\|x_n - x^*\|)|} \tag{4.2}$$

for each $x \in U(x^*, \min\{\bar{\rho}, c_1\})$. *Thus, in particular,* G' *is invertible on* $U(x^* . \bar{\rho})$.

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

Proof It follows by the condition (4.1) that $h_0'(\|x-x^*\|)<0$. Thus, the condition (A_2) and Definition 2.1 give

$$\|M^{-1}(G'(x)-M)\| \le h_0'(\|x-x^*\|)-h_0'(0) < -h_0'(0) = 1.$$

Hence, by Lemma 2.1, $M^{-1}G'(x) \in L(E_1,E_2)$, so $G'(x)^{-1} \in L(E_2,E_1)$.

Moreover,

$$\begin{aligned}\|G'(x)^{-1}M\| &\le \frac{1}{1-\|M^{-1}(G'(x)-M)\|}\\ &\le \frac{1}{1-(h_0'(\|x-x^*\|)-h_0'(0))} = \frac{1}{|h_0'(\|x-x^*\|)|}.\end{aligned}$$

Finally, the last item is valid due to the inequality $\rho^* \le \min\{\bar{\rho}, c_1\}$. ∎

It is well known that the Newton method at a single point is the solution of the linearization at the point at hand. Thus, it is important to study the linearization error at that point in D:

$$\Lambda_G(u_1,u_2) = G(u_2) - G(u_1) - G'(u_1)(u_2-u_1) \tag{4.3}$$

for each $u_2, u_1 \in D$.

This error is controlled by the error in the linearization of the majorant functions h_0, h, and h_1 defined as follows for each $v_1, v_2 \in [0,R)$:

$$\lambda_{h_0}(v_1,v_2) = h_0(v_2) - h_0(v_1) - h_0'(v_1)(v_2-v_1),$$
$$\lambda_h(v_1,v_2) = h(v_2) - h(v_1) - h'(v_1)(v_2-v_1),$$

and

$$\lambda_{h_1}(v_1,v_2) = h_1(v_2) - h_1(v_1) - h_1'(v_1)(v_2-v_1).$$

Notice that the pair $(\Lambda_G, \lambda_{h_1})$ are used in the local analysis of convergence in [9] (see also [9–13]).

Lemma 4.2 *Assume that $x \in U(x^*,\bar{\rho})$. Then, the following error estimate is valid:*

$$\|M^{-1}\Lambda_G(x,x^*)\| \le \lambda_h(\|x-x^*\|,0) \tag{4.4}$$

for each $x \in U(x^*,\bar{\rho})$.

Proof Simply replace λ_{h_1} by λ_h and Definition 2.4 by Definition 2.2 in the corresponding proof of Lemma 2.10 in [9], provided that $M = G'(x^*)$. ∎

The invertibility of G' is guaranteed by Lemma 4.1 in the ball $U(x^*,\rho^*)$.

Thus, the Newton iteration operator

$$\begin{aligned} N_G &: U(x^*,\rho^*) \to E_2,\\ x &\to x - G'(x)^{-1}G(x) \end{aligned} \tag{4.5}$$

is well defined on $U(x^*,\rho^*)$.

The next auxiliary lemma develops conditions to guarantee that the Newtonian iterates can be repeated indefinitely.

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

Lemma 4.3 *Let $s \in (0, \rho^*)$. Assume $x \in U(x^*, s)$. Then, the following error estimate is valid:*

$$\|N_G(x) - x^*\| \le \frac{|\psi_{h_0,h}(s)|}{s^2}\|x - x^*\|^2 \tag{4.6}$$

for each $x \in U(x^*, s)$.

Proof Inequality (4.6) is trivially valid for $x = x^*$, since $G'(x^*) = 0$. Thus, we can assume $\|x - x^*\| \in (0, s)$. Lemma 4.1 assures the invertibility of the linear operator $G'(x)$. Then, we have the identity

$$\begin{aligned} x^* - N_G(x) &= -G'(x)^{-1}[G(x^*) - G(x) - G'(x)(x^* - x)] \\ &= -G'(x)^{-1}\Lambda_G(x, x^*). \end{aligned} \tag{4.7}$$

Consequently, by (4.7) and Lemmas 4.1 and 4.2, we get

$$\begin{aligned} \|x^* - N_G(x)\| &\le \|-G'(x)M\|\|M^{-1}\Lambda_G(x, x^*)\| \\ &\le \frac{\lambda_h(\|x - x^*\|, 0)}{|h_0'(\|x - x^*\|)|} \le |\Lambda_{h_0,h}(\|x - x^*\|)|, \end{aligned} \tag{4.8}$$

since $h_0(0) = h(0) = 0$. Then, the application of Lemma 3.2 for $x \in U(x^*, s)$ implies

$$\frac{|\Lambda_{h_0,h}(\|x - x^*\|)|}{\|x - x^*\|} \le \frac{|\lambda_{h_0,h}(s)|}{s^2}. \tag{4.9}$$

Next, by (4.8) and (4.9), we obtain

$$\frac{\lambda_{h_0,h}(\|x - x^*\|, 0)}{|h_0'(\|x - x^*\|)|} \le \frac{|\psi_{h_0,h}(s)|}{s^2}\|x - x^*\|,$$

leading to the validation of the item (4.8). ■

Corollary 4.4 *Let $s \in (0, \rho^*)$. Then, the following are valid:*

$$N_G(U[x^*, s]) \subset U[x^*, |\psi_{h_0,h}(s)|] \tag{4.10}$$

and

$$N_G(U[x^*, \rho^*]) \subset U(x^*, \rho^*). \tag{4.11}$$

Proof By the application of Lemma 4.3 for $x \in U[x^*, s]$ and since $\|x - x^*\| \le s$, we obtain $\|N_G(x) - x^*\| \le |\psi_{h_0,h}(s)|$.

Hence, the item (4.10) is valid. Moreover, Lemma 3.2 and $\rho^* \le c_2$ imply $|\psi_{h_0,h}(s)| \le s$. Therefore, the item (4.11) is also valid. ■

Next, a domain is determined that contains only one solution which is x^*.

Proposition 4.5 *Let $s \in (0, \bar{\rho})$. Assume that zero is the only solution of the equation $h_0(s) = 0$ in the closed interval $[0, s]$, i.e., $h_0(s) < 0$. Then, the limit point x^* is the only solution of the equation $G(x) = 0$ in the closed ball $U[x^*, s]$. Consequently, the limit point x^* is the only solution of the equation $G(x) = 0$ in the open ball $U(x^*, c_3)$.*

Proof Simply replace the tighter function h_0 than h_1 that is actually needed in the proof of the corresponding Lemma 2.13 in [9]. ■

Remark 4.6 If $h_0 = h = h_1$, Lemma 4.1 and Proposition 4.5 reduce to the corresponding Lemmas 2.9 and 2.13, respectively. Otherwise, since

$$|h_0(s)| \leq |h_1(s)|$$

and

$$|h_0(s)| \leq |h_1(s)|.$$

The new results provide tighter upper error bounds on the norms $\|G'(x)^{-1}M\|$ and a larger radius of uniqueness for the solution x^* (see also Remark 3.5). In particular, under the Lipchitz case, we have

$$\|G'(x)^{-1}M\| \leq \frac{1}{1 - l_0\|x - x^*\|} \leq \frac{1}{1 - l_1\|x - x^*\|}.$$

In order to specify c_3, we must solve the inequality $h_0(s) < 0$, leading to

$$c_3 = \frac{2}{l_0}.$$

The corresponding c_3 given in Lemma 2.13 in [9] is obtained if we solve for s the inequality $h_1(s) < 0$, leading to

$$\bar{c}_3 = \frac{2}{l_1}.$$

Consequently, we get

$$\bar{c}_3 \leq c_3.$$

Next, the proof of Theorem 2.4 is developed based on the abovementioned auxiliary results.

Proof of Theorem 2.4 It follows by (2.2) and (4.5) that the Newton iteration $\{x_n\}$ satisfies the identity

$$x_{n+1} = N_G(x_n), \text{ for each } n = 0, 1, 2, \ldots. \tag{4.12}$$ ■

Next, notice that all the items concerning the scalar sequence $\{s_n\}$ are shown in Lemma 3.4. Moreover, the Newton iteration $\{x_n\}$ is well defined and belongs in $U(x^*, \rho^*)$. Indeed, for $x_0 \in U(x^*, \rho^*)$ and, since $\rho^* \leq c_1$, the formula (4.12), the item $N_G(U(x^*, \rho^*)) \subset U(x^*, \rho^*)$ in Corollary 4.4 (see (4.11)) and Lemma 4.1 validate the claim. Next, we are going to determine that $\lim_{m\to+\infty} x_m = x^*$. Let us show that the scalar sequence $\{s_m\}$ majorizes the Newton sequence $\{x_m\}$, i.e.,

$$\|x^* - x_m\| \leq s_m, \text{ for each } m = 0, 1, 2, \ldots. \tag{4.13}$$

By the choice $s_0 = \|x^* - x_0\|$, inequation (4.13) holds if $m = 0$. We employ mathematical induction and assume that $\|x^* - x_m\| \leq s_m$. Then, by Lemma 3.4, $\{s_m\} \subset (0, \rho^*)$.

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

Furthermore, the application of the definition of the sequence $\{s_n\}$ given by the formula (3.10) and the identity (4.12) gives

$$\|x^* - x_{m+1}\| \le \|x^* - N_G(x_m)\| \le |\psi_{h_0,h}(s_m)| = s_{m+1}. \tag{4.14}$$

The proof of the induction for inequation (4.13) is complete. Furthermore, by (4.13) and (3.12), $\lim_{m\to+\infty} x_m = x^*$, since $\lim_{m\to+\infty} s_m = 0$. The claim about the optimality of the convergence radius ρ^* is given in Lemma 2.15 in [9]. The first inequality in (2.4) follows by Lemma 3.4 and (4.12) since

$$\|x^* - x_{m+1}\| = \|x^* - N_G(x_m)\| \le \frac{|\psi_{h_0,h}(s_m)|}{s_m^2}\|x^* - x_m\|^2,$$

for each $m = 0, 1, 2, \ldots$. Thus, the last inequation and the definition of the majorant sequence $\{s_m\}$ show the validity of the first inequation in (2.4). Finally, the uniqueness of the solution x^* is shown in Proposition 4.5.

5 Special cases and examples

It is worth noticing that, if we simply specialize the functions h_0 and h, the results extend the applicability of the classical cases [Lipchitz] (see also Remark 3.5), under the Smale [23] or Wang [25, 26] conditions and under the Nesterov et al. conditions [16]. We leave the details to the motivational reader.

Next, we present two examples to test the convergence conditions and further validate the theoretical results. We have chosen $M = G'(x^*)$.

Example 5.1 Consider, the choice $B_1 = B_2 = \mathbb{R}^3$ and $D = P[0,1]$. Define the operator on D as

$$G(t) = \left(\frac{e-1}{2}t_1^2 + t_1, e^{t_2} - 1, t_3\right)^T \quad \text{for} \quad t = (t_1, t_2, t_3)^T \tag{5.1}$$

and $t_1, t_2, t_3 \in \mathbb{R}$. Clearly, $t^* = (0,0,0)^T$ solves the equations. Then, by the definition (5.1), it follows that the derivative according to Fréchet G' of the operator F is defined as

$$G'(t) = \begin{bmatrix} (e-1)t_1 + 1 & 0 & \\ 0 & e^{t_2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We have $h_0(s) = (e-1)s$, $h(s) = e^{\frac{1}{e-1}}s$, and $h_1(s) = es$. Then, $\rho_1^* = 0.2453$ [9] and $\rho^* = 0.3827$. Thus, the previous radius of convergence ρ_1^* is smaller than the new one ρ^*. This allows for a wider choice of initial guesses x_0, and other benefits as already mentioned under novelty in the introduction of this article.

Example 5.2 Let $K[0,1]$ stand the space of continuous functions mapping the interval $[0,1]$ into the real number system. Let $B_1 = B_2 = K[0,1]$ and $D = P[x^*, 1]$ with $x^*(\mu) = 0$. The operator G is defined on $K[0,1]$ as

https://doi.org/10.4153/S000843952400016X Published online by Cambridge University Press

$$G(z)(\mu) = z(\mu) - 6\int_0^1 \mu z(\tau)^3 d\tau.$$

Then, the definition of the derivative according to Fréchet [4–6, 22] gives for the function G

$$G'(z(w))(\mu) = w(\mu) - 18\int_0^1 \mu\tau z(\tau)^2 w(\tau) d\tau$$

for each $w \in K[0,1]$. Then, the conditions (A1)–(A5) of Theorem 2.4 are validated, since $G'(x^*(\mu)) = I$ provided that $h_0(s) = h(s) = 12s$ and $h_1(s) = 24s$. Then, $\rho_1^* = 0.0278$ and $\rho^* = 0.0556$. Notice that as in Example 5.1, $\rho_1^* < \rho^*$.

6 Conclusion

A very general theory for studying the convergence of Newton-type methods is developed for generating sequences approximating a solution of a generalized equation involving set-valued operators. Both the local as well as the semi-local analyses of convergence rely on the weaker conditions and the concept of generalized continuity. Moreover, we provide upper error bounds on the norms $\|x_{n+1} - x_n\|$ and $\|x^* - x_n\|$. In particular, the semi-local analysis of convergence is based on majorizing sequences for $\{x_n\}$ generated by the method (1.2). It was shown that even specializations of the operators involved lead to better results when compared to existing ones. The future direction of our research involves the application of the developed theory on other methods [1, 2, 6–8, 17].

Acknowledgements We would like to express our gratitude to the anonymous reviewers for the constructive criticism of this paper.

References

[1] S. Adly, H. Van Ngai, and V. V. Nguyen, *Newton's method for solving generalized equations: Kantorovich's and Smale's approaches*. J. Math. Anal. Appl. **439**(2016), no. 1, 396–418.

[2] F. J. Aragón Artacho, A. Belyakov, A. L. Dontchev, and M. López, *Local convergence of quasi-Newton methods under metric regularity*. Comput. Optim. Appl. **58**(2014), no. 1, 225–247.

[3] I. K. Argyros, *Convergence and applications of Newton-type iterations*, Springer, New York, 2008.

[4] I. K. Argyros, *The theory and application of iteration methods*, 2nd ed., Engineering Series, CRC Press, Boca Raton, FL, 2022.

[5] I. K. Argyros and S. George, *On the complexity of extending the convergence region for Traub's method*. J. Complex. **56**(2020), 101423.

[6] J. F. Bonnans, *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*. Appl. Math. Optim. **29**(1994), no. 20, 161–186.

[7] R. Cibulka, A. Dontchev, and M. H. Geoffroy, *Inexact Newton methods and Dennis–Moré theorems for nonsmooth generalized equations*. SIAM J. Control. Optim. **53**(2015), no. 2, 1003–1019.

[8] A. L. Dontchev and R. T. Rockafellar, *Implicit functions and solution mappings*, 2nd ed., Springer Series in Operations Research and Financial Engineering, Springer, New York, 2014.

[9] O. P. Ferreira, *Local convergence of Newton's method from the view point of the majorant principle*. IMA J. Numer. Anal. **29**(2009), no. 3, 746–759.

[10] O. P. Ferreira and B. F. Svaiter, *Kantorovich's theorem on Newton's method in Riemannian manifolds*. J. Complex. **18**(2002), no. 1, 304–329.

[11] O. P. Ferreira, C. Jean-Alexis, A. Pietrus, and G. N. Silva, *On Newton's method for solving generalized equations*. J. Complex. **74**(2023), 101697.

[12] O. P. Ferreira and G. N. Silva, *Local generalized analysis of Newton's method for solving strong regular generalized equations*. J. Math. Anal. Appl. **458**(2018), no. 1, 481–496.
[13] O. P. Ferreira and R. B. Svaiter, *Kantorovich's majorant principle for Newton's method*. Comput. Optim. Appl. **42**(2009), 213–229.
[14] J. B. Hiriart-Verryty and C. Lemarechal, *Convex analysis and minimization algorithms, part I*, Springer, Berlin, 1993.
[15] A. A. Magreńan and I. K. Argyros, *A contemporary study of iterative methods: convergence dynamics and applications*, Academic Press, Elsevier, 2018.
[16] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.
[17] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer, New York, 2006.
[18] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970.
[19] B. T. Polyak, *Newton's method and its use in optimization*. Eur. J. Oper. Res. **181**(2007), 1086–1096.
[20] P. D. Proinov, *Semi-local convergence of two iterative methods for simultaneous computation of polynomial zeros*. C. R. Acad. Bulgare Sci. **59**(2006), no. 7, 705–712.
[21] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*. In: A. N. Tikhonov, F. Kuhnert, N. N. Kuznecov, A. Moszynski, and A. Wakulicz (eds.), Mathematical models and numerical methods, Vol. 3, Banach Center, Warsaw, 1977, pp. 129–142.
[22] S. M. Robinson, *Generalized equations*. In: Mathematical programming: the state of the art (Bonn, 1982), Springer, Berlin, 1983, pp. 346–367.
[23] S. Smale, *Newton method estimates from data at one point*. In: R. Ewing, K. Gross, and C. Martin (eds.), The merging of disciplines: new directions in pure, applied and computational mathematics, Springer, New York, 1986, pp. 185–196.
[24] J. Traub and H. Wozniakowski, *Convergence and complexity of Newton iteration for operator equations*. J. Assoc. Comput. Math. **26**(1979), no. 2, 250–258.
[25] J. Wang and W. Ouyang, *Newton's method for solving generalized equations without Lipschitz condition*. J. Optim. Theory Appl. **192**(2022), no. 2, 510–532.
[26] X. Wang, *Convergence of Newton's method and uniqueness of the solution of equation in Banach space*. IMA J. Numer. Anal. **20**(2000), 123–134.
[27] P. P. Zabrejko and D. F. Nguen, *The majorant method in the theory of Newton–Kantorovich approximations and the Pták error estimates*. Numer. Funct. Anal. Optim. **9**(1987), nos. 5–6, 671–684.

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

Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, Mangalore 575 025, India
e-mail: sgeorge@nitk.edu.in