*Annales Univ. Sci. Budapest., Sect. Comp. 58 (2025) 31–45*

# EXTENDED NEWTON'S METHOD FOR SOLVING GENERALIZED EQUATIONS UNDER KANTOROVICH'S-TYPE WEAK CONDITIONS

**Ioannis K. Argyros** (Lawton, US)
**Santhosh George** (Mangaluru, India)

Communicated by László Szili
(Received 21 November 2024; accepted 30 December 2024)

**Abstract.** The Kantorovich's methodology has been applied extensively to solve generalized equations using Newton's method. However, the mostly sufficient conditions limit the applicability of this method. But the method may converge even if these conditions are not satisfied. Therefore, it is important to show convergence to a solution under weaker conditions, and if possible without additional conditions. Motivated by optimization considerations and using more precise majorizing sequences we obtain the following advantages over either studies:
**Semi-Local Case**: Extended convergence domain under the same or weaker convergence conditions, and tighter error bounds on the distances involved.
**Local Case**: Enlarged radius of convergence and a finer error analysis. Our approach provides the same advantages in the case of Smale's theory for generalized equations.

## 1. Introduction

In 1948, L.V. Kantorovich published a seminal article for approximating solutions of nonlinear equations using Newton's method defined on abstract

*Key words and phrases*: Generalized equation, Fréchet derivative, Newton's method, Banach space, metric regularity, local-semi-local convergence.
*2010 Mathematics Subject Classification*: 49M15, 47H99, 65J15, 65G99, 65D99.
https://doi.org/10.71352/ac.58.301224
First published online: 8 March 2025

spaces. These results on the one hand provide the existence of a solution as well as sufficient convergence conditions for finding it using Newton's method. The assumptions usually involve the derivative of the operator which is controled by Lipschitz–Hölder or ever more general conditions.

In 1980, S.M. Robinson [20] extended previous results on nonlinear equations for solving generalized equations involving parameters. Later Josephy [14] and Bonnans [7], Dontchev et al. [9, 10, 11, 12] worked on more general generalized equations defined on the finite-dimensional Euclidean space or Hilbert or Banach spaces.

Let the letters $B_1$ and $B_2$ stand for Banach spaces. We are concerned with the problem of finding a solution $x_* \in B_1$ of the generalized equation

$$0 \in g(x) + G(x), \tag{1.1}$$

using Newton's Method (NM) defined by

$$0 \in g(x_m) + g'(x_m)(x_{m+1} - x_m) + G(x_{m+1}), \quad \text{for each } \ m = 0, 1, 2, \dots, \tag{1.2}$$

where $g : B_1 \to B_2$ is a single-valued operator, and $G : B_1 \rightrightarrows B_2$ is a set-valued operator with a closed graph [1, 12, 16].

In this article, we are motivated by optimization considerations and the elegant study by Adly et al. [2] on the local as well as the semi-local convergence analysis of NM which is based on conditions involving $g''$.

But there are even simple scalar cases when $g''$ does not exist limiting the applicability of NM. Consider, the case when $G = 0$. Define $g : D_0 \to \mathbb{R}$ by $g(x) = \mu_1 x^2 \ln x + \mu_2 x^3 + \mu_3 x^2$ if $x \neq 0$ and $g(x)$, if $x = 0$ with $\mu_1 \neq 0$, $\mu_2 + \mu_3 = 0$, and $D_0$ be an interval containing 0 and 1. Notice that $x_* = 1 \in D_0$ solves the equation $g(x) = 0$. But $g''$ is unbounded on $D_0$, since $g''(x)$ does not exist at $x = 0 \in D_0$. Notice also that the operator $g^{''}$ does not appear on NM given by (1.2) .

But there are other limitations in general.

<u>Local analysis of convergence:</u>

$(e_1)$ : The radius of convergence is small and

$(e_2)$ : The error bounds on the distances $\|x_m - x_*\|$ are pessimistic.

<u>Semi-local analysis of convergence:</u>

$(e_3)$ : The convergence domain is small and the limitation $(e_2)$ may also be valid. The following are accomplished by using only the operators on NM:

$(e_1)'$ : The radius of convergence is extended resulting to a wider choice of initial points $x_0$ which can be used to ensure convergence of the method.

$(e_2)'$ : The new error bounds are tighter. Hence the number of iterations to be executed to arrive at a predetermined error tolerance is reduced.

$(e_3)'$ : The convergence domain is extended, since the conditions are weakened.

It is worth noting that these improvements do not require additional conditions and computational effort, since the Lipschitz parameters used are special cases of the ones used previously. The new methodology requires more precise majorizing sequences and uses the concept of metric regularity. Moreover, it is so general that it can be applied to extend the applicability of other methods [6, 17, 18, 23] and under other conditions such as Smale's [22].

The rest of the article contains the preliminaries in Section 2, and the local analysis of convergence in Section 3. The semi-local analysis in Section 4, and the concluding remarks in Section 5.

## 2. Mathematical background

Some standard notation, concepts and results are restated in order to make the article as self-contained as possible. More information about set-valued operators and metric regularity can be found in [12]. We have adopted the notations $F(x,d)$ and $F[x,d]$ to stand for the open and closed balls, respectively with center $x \in B_1$ and of radius $d > 0$. If $x \in B_1$ and $E \subset B_1$. The distance function from $x$ to $E$ is defined as $d(x,E) := \inf_{\bar{x}\in E} \|x - \bar{x}\|$. If $E_1 \subset B_1$ and $E_2 \subset B_2$, the excess of $E_1$ beyond $E_2$ is defined as $e(E_1, E_2) := \inf\{s > 0 : E_1 \subset E_2 + sB_1\}$. The Hausdorff distance between $E_1$ and $E_2$ is given as $d(E_1, E_2) := \max\{e(E_1, E_2), e(E_2, E_1)\}$. A set-valued operator $H : B_1 \rightrightarrows B_2$ relates $x \in B_1$ to a subset $H(x) \subset B_2$. Moreover, the domain and graph of operator $H$ are defined by $dom(H) := \{x \in B_1 : H(x) \neq \emptyset\}$ and $gph(H) := \{(x,y) \in B_1 \times B_2 : y \in H(x)\}$. Furthermore, the inverse of the operator $H$ is the set-valued operator $H^{-1} : B_2 \rightrightarrows B_1$ so that $x \in H^{-1}(y)$ iff $y \in H(x)$.

An important role in the local as well as the semi-local analysis of convergence is played by the metric regularity property of an operator [12].

**Definition 2.1.** An operator $H : B_1 \rightrightarrows B_2$ is metrically regular at $(\tilde{x}, \tilde{y}) \in gph(H)$ with modulus $\xi > 0$ if there exist neighbourhoods $N_1$ of $\tilde{x} \in B_1$ and $N_2$ of $\tilde{y}$ in $B_2$ so that $d(x, H^{-1}(y)) \leq \xi d(y, H(x))$, for each $((x,y) \in N_1 \times N_2)$. The infimum of all parameters $\xi$ is the regularity modulus of $H$ at $\tilde{x}$ for $\tilde{y}$ which we denote as $Reg(H; (\tilde{x}, \tilde{y}))$.

The metric regularity of an operator $H$ is equivalent to the Aubin property which is a Lipschitz-type property of the inverse $H^{-1}$ at $(\tilde{y}, \tilde{x})$ [2]. Note that an operator $H_0 : B_2 \rightrightarrows B_1$ is Lipschitz-type at $(\tilde{y}, \tilde{x}) \in gph(H_0)$ with parameter $\xi$ if there exists a neighbourhood $N_2 \times N_1$ of $(\tilde{y}, \tilde{x})$ of $B_2 \times B_1$ so that $e(H_0(y) \cap N_1, H_0(\bar{y})) \leq \xi \|y - \bar{y}\|$, for each $y, \bar{y} \in N_2$.

Recall that a set-valued operator $H_1 : B_1 \rightrightarrows B_2$ is said to be Lipschitz continuous on a subset $\Omega$ of $B_1$ if there exists a Lipschitz parameter $\ell > 0$ so that

$$d(H_1(x), H_1(\bar{x})) \leq \ell \|x - \bar{x}\|,$$

for each $x, \bar{x} \in \Omega$.

The rest of the results in this section are taken from [2] (see also the more general versions in [1, Theorem 3.2 and Theorem 6.2]). From now we denote by $\Psi : B_1 \rightrightarrows B_2$ a set-valued operator with a closed graph. Let $(\tilde{x}, \tilde{y}) \in gph(\Psi)$.

**Theorem 2.1.** *Suppose that $\Psi$ is metrically regular at $(\tilde{x}, \tilde{y})$ with modulus $\xi > 0$ on a neighbourhood $F(\tilde{x}, a) \times U(\tilde{y}, b)$ of $(\tilde{x}, \tilde{y})$ for some $a > 0$ and $b > 0$. Pick $\delta > 0$, $\ell \in (0, \xi^{-1})$ and take $\theta = \dfrac{\xi}{1 - \xi\ell}$. Let $\alpha > 0, \beta > 0$ be parameters validating*

$$2\alpha + \beta\theta < \min\left\{a, \frac{\delta}{2}\right\}, \quad \beta(\theta + \xi) < \delta, \quad 2\gamma\alpha + \beta(1 + \gamma\theta) < b$$

*for $\gamma := \max\{1, \xi^{-1}\}$. If $C : B_1 \to B_2$ is Lipschitz conditions on $F(\tilde{x}, \delta)$ with parameter $\ell$, and $\Psi + C$ has closed graph, then $\Psi + C$ is metrically regular on $F(\tilde{x}, \alpha) \times F(\tilde{y} + C(\tilde{x}), \beta)$ with modulus $\theta$.*

**Definition 2.2.** Let $\rho > 0, \bar{\rho} > 0$. Define the set

$$S = S(\Psi, x_0, \rho, \bar{\rho}) := \{(x, y) \in B_1 \times B_2 : x \in F[x_0, \rho], d(y, \Psi(x)) < \bar{\rho}\}. \tag{2.1}$$

We say that $\Psi$ is metrically regular on the set $S$ with modulus $\theta > 0$ if

$$d(x, \Psi^{-1}(y)) \leq \theta d(y, \Psi(x)) \text{ for each } (x, y) \in S \tag{2.2}$$

In this case the infimum for all *theta* $> 0$ for which (2.2) is valid is denoted by $Reg(\Psi, x_0, \rho, \bar{\rho})$. Otherwise, we take $Reg(\Psi, x_0, \rho, \bar{\rho}) = \infty$.

**Theorem 2.2.** *Assume $\Psi$ is metrically regular on $S$ with modulus $\xi$. Choose $\ell \in (0, \xi^{-1})$ and take $\theta = \dfrac{\xi}{1 - \xi\ell}$. If $C : B_1 \to B_2$ is Lipschitz continuous on $F[x_0, \rho]$ with parameter $\ell$, and $\Psi + C$ has a closed graph, then the set-valued operator $\Psi + C$ is metrically regular on $S(\Psi + C, x_0, \dfrac{\rho}{4}, \tilde{\rho})$ with modulus $\theta$ for $\tilde{\rho} = \min\left\{\bar{\rho}, \dfrac{\rho}{5\theta}\right\}$.*

## 3. Local analysis of convergence

Let $R > 0$ be a given parameter. Define $\rho = \sup\{s \in [0, R) : F(x_*, s) \subset \subset D_0\}$. The following Lipschitz-type conditions play a role in the local analysis of convergence for NM

**Definition 3.1.** Let $g : F(x_*, \rho) \to B_2$ be a continuously differentiable operator. We say that the operator $g'$ is center-Lipschitz continuous at $x_*$ with modulus $\theta > 0$ if

$$\|g'(x) - g'(x_*)\| \leq \frac{L_0}{\theta}\|x - x_*\| \tag{$A_1$}$$

for each $x \in F(x_*, \rho)$ and some $L_0 > 0$.

**Definition 3.2.** Define $\rho_0 = \dfrac{1}{2L_0}$, and $\rho_* = \min\{\rho_0, \rho\}$ and $F_0 = F(x_*, \rho_*)$. Let $g : F(x_*, \rho_*) \to B_2$ be a continuously differentiable operator. We say that the operator $g'$ is restricted Lipschitz continuous at $x_*$ with modulus $\theta > 0$ if

$$(A_2) \qquad \|g'(x) - g'(y)\| \leq \frac{L}{\theta}\|x - y\|$$

for each $x, y \in F_0$, and some $L > 0$.

**Definition 3.3.** Let $g : F(x_*, \rho) \to B_2$ be a continuously differentiable operator. We say that the operator $g'$ is Lipschitz continuous with modulus $\theta > 0$ if

$$(A_3) \qquad \|g'(x) - g'(y)\| \leq \frac{L_1}{\theta}\|x - y\|$$

for each $x, y \in F(x_*, \rho)$, and some $L_1 > 0$.

**Remark 3.1.** (i) It follows by these definitions that

$$L_0 \leq L_1 \tag{3.1}$$

and

$$L \leq L_1. \tag{3.2}$$

(ii) Notice that the parameters $L_0, L$ and $L_1$ can be chosen as

$$L_0 = \theta \sup_{x \in F[x_*, \rho]} \frac{\|g'(x) - g'(x_*)\|}{\|x - x_*\|},$$

$$L = \theta \sup_{x \in F_0} \frac{\|g'(x) - g'(y)\|}{\|x - y\|},$$

and

$$L_1 = \theta \sup_{x \in F[x_*, \rho]} \frac{\|g'(x) - g'(y)\|}{\|x - y\|}.$$

Moreover, if the operator $g$ is twice continuously differentiable then

$$\overline{L_1} = \theta \sup_{x \in F[x_*, \rho]} \|g''(x)\|.$$

The preceding choice for $\bar{L}$ is used in the local analysis of convergence for NM (see [2, Theorem 3.1 in page 399]). However, in view of (3.1) and (3.2), the parameter $L_1$ can be replaced by the tighter and actually needed parameters $(L_0, L)$. This way we obtain the advantages as already stated in the introduction. Indeed, we can present the local analysis of convergence for NM.

**Theorem 3.1.** *Suppose:*

(a) *The conditions* $(A_1)$ *and* $(A_2)$ *hold; take* $\rho_* = g^*(x_*)(x_*) - g(x_*) \in B_2$.

(b) *The set-valued operator* $\Psi(.) = g'(x_*)(.) + G(.)$ *is metrically regular on the neighbourhood* $V = F[x_*, \rho] \times F[y_*, h]$ *of* $(y_*, h)$ *with modulus* $\theta$. *Take* $\epsilon = \min\{\rho, h, \theta h\}$.

(c) *For* $\bar{L} = \frac{L_0+L}{2}$

$$2\bar{L}\rho_* < 1. \tag{3.3}$$

*Define the sequence* $\{s_m\}$ *for some* $x_0 \in F(x_*, \rho_*) - \{x_*\}$,

$$s_0 = \|x_* - x_0\|, \quad s_1 = \frac{L_0 s_0^2}{2(1 - \theta L_0 \rho_*)},$$

*and each* $m = 1, 2, \dots$ *as*

$$s_{m+1} = \frac{L s_m^2}{2(1 - \theta L_0 \rho_*)}. \tag{3.4}$$

*Then, the following assertions hold:*

(i) *The scalar sequence* $\{s_m\}$ *is well defined in the interval* $(0, \|x_* - x_0\|]$, *remains in* $(0, \rho_*)$ *for each* $n = 0, 1, 2, \dots$, *is decreasing and is convergent quadratically to zero.*

(ii) *For some* $x_0 \in F(x_*, \epsilon)$ *and* $\epsilon = \min\{\rho_*, 1, \theta\lambda\}$ *there exists a sequence* $\{x_m\}$ *generated by NM which is quadratically convergent to* $x_*$ *so that for each* $n = 0, 1, 2, \dots$

$$\|x_{m+1} - x_*\| \le s_{n+1} \tag{3.5}$$

*and*

$$\|x_{m+1} - x_*\| < \bar{c}\,\|x_m - x_*\|^2, \tag{3.6}$$

*where*

$$c_0 = \frac{L_0}{2(1 - \theta L_0 \rho_*)}, \quad c = \frac{L}{2(1 - \theta L_0 \rho_*)} \quad \text{and} \quad \bar{c} = \begin{cases} c_0, & m = 1 \\ c\,, & m \neq 1. \end{cases}$$

**Proof.** (i) By the condition (3.3), it follows

$$0 \le L_0 \rho_* < 1. \tag{3.7}$$

In view of (3.7) the sequence $\{s_m\}$ simplifies to $s_0 = \|x_* - x_0\|$, $s_1 = c_0\, s_0^2$ and

$$s_{m+1} = c\, s_m^2, \quad m = 0, 1, 2, \dots. \tag{3.8}$$

Then, $s_1 < s_0$, $s_2 < s_1$, $c_0\, s_0 < 1$ and $c\, s_1 < 1$ by (3.3). By induction assume $s_m < s_{m-1}$ and $\bar{c}\, s_{m-1} < 1$, then $s_{m+1} - s_m = (\bar{c}\, s_m - 1)s_m < 0$ for $s_m > 0$. Thus, the sequence $\{s_m\}$ is decreasing, bounded from below by zero and as such it is convergent to some limit greater or equal to zero. By letting $m \to +\infty$ in (3.4) we see that this limit is zero. Moreover, by (3.8) the convergence is quadratic, since $\bar{c} \neq 0$. But this is absurd, since $x_0 \in F(x_*, \rho_*) - \{x_*\}$.

(ii) Take $\ell = \dfrac{L_0 \rho_*}{\theta} \leq \dfrac{1}{2\theta}$, by (3.3). Set $\bar{\theta} = \dfrac{\theta}{1 - \theta\ell}$ and $v = \dfrac{\epsilon}{4\theta}$. Then, it follows $v\bar{\theta} = v\dfrac{\theta}{1-\theta\ell} \leq 2v\theta = \epsilon/2 < \dfrac{\rho_*}{2}$. So, since $\theta\ell \leq \dfrac{1}{2}$,

$$v(\bar{\theta} + \theta) = v\theta\Big(\frac{1}{1-\theta\ell} + 1\Big) \leq 3\theta v = \frac{3\epsilon}{4} \leq \rho_*.$$

Take $\gamma = \max\Big\{1, \dfrac{1}{\theta}\Big\}$. If $\theta \geq 1$, then $\gamma = 1$. Consequently,

$$v(1 + \gamma\bar{\theta}) \leq v(\theta + \theta) \leq 3\theta v < \lambda.$$

If $\theta < 1$, then $\gamma = \dfrac{1}{\theta}$, so

$$v(1 + \gamma\bar{\theta}) = v\Big(1 + \frac{1}{\theta}\bar{\theta}\Big) = v\Big(1 + \frac{1}{1-\theta\ell}\Big) \leq 3v < \lambda.$$

Pick $\mu > 0$ so that

$$2\mu + v\bar{\theta} < \frac{\rho_*}{2}, \quad v(\bar{\theta} + \theta) < \rho_*, \quad 2\gamma\mu + v(1 + \gamma\bar{\theta}) < \lambda.$$

For any operator $T : B_1 \to B_2$ which is linear and continuous with $\|T\| \leq \ell$, Theorem 2.1 with $a = \rho_*, b = \lambda, \delta = \rho_*$ asserts $\Psi_T = T + \Psi$ is metrically regular on $F[x_*, \mu] \times F[y_* + T(x_*), v]$ with a modulus $\bar{\theta}$. By hypothesis $x_0 \in F(x_*, \epsilon) - \{x_*\}$ and the condition $(A_1)$

$$\|g'(x_0) - g'(x_*)\| \leq \frac{L_0}{\theta}\|x_0 - x_*\| \leq \frac{L_0}{\theta}\epsilon \leq \frac{L_0\rho_*}{\theta} = \ell.$$

Thus, $\Psi_0 = g'(x_0) + G = (g'(x_0) - g'(x_*)) + \Psi$ is metrically regular in the neighbourhood $F[x_*, \mu] \times F[y_0, v]$ of $(x_*, y_0)$ with modulus $\theta_0 = \dfrac{\theta}{1-\theta\ell}$ provided $y_0 = g'(x_0)(x_*) - g(x_*)$. Take $u_0 = g'(x_0)(x_0) - g(x_0)$. Then, we get in turn by $(A_2)$ that

$$\begin{aligned}
\|u_0 - y_0\| &= \|g'(x_0)(x_0) - g(x_0) - (g'(x_0)(x_*) - g(x_*))\| = \\
&= \|g(x_*) - g(x_0) - g'(x_0)(x_* - x_0)\| = \\
&= \left\| \int_0^1 [g'(x_0 + \tau(x_* - x_0)) - g'(x_0)]\, d\tau (x_* - x_0) \right\| \leq \\
&\leq \frac{L}{2\theta}\|x_* - x_0\|^2 \leq \frac{L}{2\theta}\rho_*\epsilon \leq \frac{\epsilon}{4\theta} = v,
\end{aligned}$$

so $u_0 \in F[y_0, v]$. But the operator $\Psi_0$ is metrically regular on $F[x_*, \mu] \times F[y_0, v]$. Thus, we get in turn

$$d(x_*, \Psi_0^{-1}(u_0)) \le \theta_0 d(u_0, \Psi_0(x_*)) \le \theta_0 d(u_0, y_0) \le \theta_0 \frac{L}{2\theta}\|x_* - x_0\|^2 = \\ = \frac{L}{2(1-\theta\ell)}\|x_* - x_0\|^2 \le \frac{L}{2(1-\theta\ell)} s_0^2 = s_1.$$

By the choice of $\ell$ and (3.3), we get

$$d(x_*, \Psi_0^{-1}(u_0)) < \frac{1}{2\rho_*}\|x_* - x_0\|^2.$$

Hence, we can pick $x_1 \in \Psi_0^{-1}(u_0)$ which satisfies

$$\|x_* - x_1\| \le s_1$$

and

$$\|x_* - x_1\| < \frac{1}{2\rho_*}\|x_* - x_0\|^2.$$

Notice also that

$$g'(x_0)(x_0) - g(x_0) = u_0 \in \Psi(x_1) = g'(x_0)(x_1) + G(x_1).$$

So, $x_1$ is a Newton iteration produces by (1.1) if $m = 0$. Moreover, the choice of $x_0$ gives

$$\|x_* - x_1\| < \frac{1}{2\rho_*}\|x_* - x_0\|^2 \le \frac{1}{2}\|x_* - x_0\| < \epsilon.$$

Concluding to $x_1 \in F(x_*, \epsilon)$. The preceding calculations can be repeated with $x_{m-1}$ replacing $x_0$ to obtain $x_m$. Then, we get

$$\|x_* - x_m\| \le s_m$$

and

$$\|x_* - x_m\| < \left(\frac{1}{2\rho_*}\|x_* - x_0\|\right)^{2^m - 1}\|x_* - x_0\| \le \\ \le \left(\frac{1}{2}\right)^{2^m - 1}\|x_* - x_0\|,$$

showing the quadratic convergence of the sequence $\{x_m\}$. ■

**Remark 3.2.** (i) The corresponding convergence condition and sequence $\{\bar{s}_n\}$ given in Theorem 3.1 in [2] are

(3.9) $$2L_1 r < 1$$

and

(3.10) $$\bar{s}_0 = \|x_0 - x_*\|, \quad \bar{s}_{n+1} = \frac{L_1}{2(1-\theta L_1)}\bar{s}_n^2.$$

provided that the operator $g$ is twice differentiable. Then, it follows from (3.1), (3.2), (3.3) and (3.9) that

$$2L_1 r < 1 \implies 2\bar{L}\rho_* < 1 \tag{3.11}$$

but not necessarily vice versa unless if $L_0 = L = L_1$. Moreover, by (3.1), (3.2), (3.4) and (3.5) we deduce

$$0 \leq s_n \leq \bar{s}_n \quad \text{for each} \quad n = 1, 2, \ldots . \tag{3.12}$$

Items (3.11) and (3.12) justify the advantages as already stated in the introduction. It is also worth noticing that

$$L_0 = L_0(\theta, x_*, \rho),$$

$$L = L(\theta, x_*, \rho_0)$$

and

$$L_1 = L_1(\theta, x_*, \rho).$$

Moreover, in practice the computation of the parameter $L_1$ requires that of $L_0$ and $L$ as special cases. Consequently, the advantages of the new approach are obtained under less computational cost, since only $g'$ is computed instead of $g''$ which may not even exist.

(ii) Theorem 3.1 in [2] improved a related result by A. Dontchev [10, see Theorem 1] in the sense that the domain of initial points $x_0$ to assure convergence is explicitly determined. Therefore our Theorem 3.1 improves also Theorem 1 in [10].

## 4. Semi-local analysis of convergence

The role of $x_*$ is exchanged by $x_0$ in this section. Define

$$\lambda = \sup \{ s \in [0, R) : F(x_0, s) \subset D \}.$$

As in the local case certain Lipschitz-type conditions are developed that play a role in the semi-local analysis of convergence for NM.

**Definition 4.1.** Let $g : F(x_0, \lambda) \to B_2$ be a continuously differentiable operator. We say that the operator $g'$ is center-Lipschitz continuous at $x_0$ with modulus $\theta > 0$ if

$$|g'(x) - g'(x_0)\| \leq \frac{M_0}{\theta} \|x - x_0\| \tag{$C_1$}$$

for each $x \in F(x_0, \lambda)$, and some $M_0 > 0$.

Define $\lambda_0 = \dfrac{1}{M_0}$ and $F_1 = F(x_0, \lambda_*)$, where $\lambda_* = \min\{\lambda, \lambda_0\}$.

**Definition 4.2.** Let $g : F(x_0, \lambda_*) \to B_2$ be a continuously differentiable operator. We say that the operator $g'$ is restricted Lipschitz at $x_0$ with modulus $\theta > 0$ if

$$(C_2) \qquad \|g'(x) - g'(y)\| \leq \frac{M}{\theta}\|x - y\|$$

for each $x, y \in F_1$, and some $M > 0$.

**Definition 4.3.** Let $g : F(x_0, \lambda) \to B_2$ be a continuously differentiable operator. We say that the operator $g'$ is Lipschitz continuous at $x_0$ with modulus $\theta > 0$ if

$$(C_3) \qquad \|g'(x) - g'(y)\| \leq \frac{M_1}{\theta}\|x - y\|$$

for each $x, y \in F(x_0, \lambda)$ and some $M_1 > 0$.

**Remark 4.1.** The comparisons between the parameters $M_0, M, M_1$ and $\bar{M}_1$ as identical to $L_0, L, L_1$ and $\bar{L}_1$, respectively are omitted (see Remark 3.1), where

$$\bar{M}_1 = \theta \sup_{x \in F[x_0, \lambda]} \|g''(x)\|,$$

provided that $g''$ exists and is bounded.

Next, we define the scalar sequence $\{s_m\}$ for $s_0 = 0$, $s_1 = ß$, for some $ß \in \in [0, \lambda]$ to be determined later, $s_2 = s_1 + \dfrac{M_0(s_1 - s_0)^2}{2(1 - M_0 s_1)}$ and for each $m = 1, 2, \ldots$

$$(4.1) \qquad s_{m+2} = s_{m+1} + \frac{M(s_{n+1} - s_m)^2}{2(1 - M_0 s_{m+1})}.$$

The sequence $\{s_m\}$ is shown to be majorizing for $\{x_m\}$ in Theorem 4.1. However, we first develop a general auxiliary convergence result for it.

**Lemma 4.1.** *Suppose that* $\dfrac{1}{M_0} \leq \lambda$ *and for each* $m = 0, 1, 2, \ldots$

$$(4.2) \qquad M_0 s_m < 1.$$

*Then, the sequence* $\{s_m\}$ *generated by the formula* (4.1) *is nondecreasing as well convergent to some* $s_* \in \left[0, \dfrac{1}{M_0}\right]$.

**Proof.** The sequence $\{s_m\}$ is convergent to some $ß_* \in \left[ß, \dfrac{1}{M_0}\right]$, as nondecreasing, and bounded from above by $\dfrac{1}{M_0}$. ■

The limit $ß_*$ is the unique least upper bound of the sequence $\{s_m\}$.

For $\theta > 0, \epsilon > 0$ and $y \in N_1$ with $F[y, \epsilon] \subset N_1$, define the parameter

$$ß = ß(\theta, y) := \theta d(0, g(y) + G(y)).$$

The semi-local analysis of NM is developed in the next result.

**Theorem 4.1.** *Let $x \in N_1$, $\lambda > 0$, $\bar{\lambda} > 0$ such that the following conditions hold:*

(i) *The conditions $(C_1)$ and $(C_2)$ are valid.*

(ii) *$\Psi = g'(x) + G$ is metrically regular on $S = S(\Psi, x, 4\lambda, s)$ with modulus $\theta > Reg(\Psi, x, 4\lambda, s)$.*

(iii) *$d(0, Q(x)) < s$, where $Q = g + G$.*
*and the conditions of the Lemma 4.1 hold.*

*Then, the sequence $\{x_m\}$ is well defined, in $N_1$, remains in $N_1$ for each $m = 0, 1, 2, \dots$ and is convergent to a solution $x_* \in F\left[x_0, \frac{1}{M_0}\right]$ of the generalized equation (1.1) so that*

$$\|x_{m+1} - x_m\| \leq s_{m+1} - s_m \tag{4.3}$$

*and*

$$\|x_* - x_m\| \leq s_* - s_m. \tag{4.4}$$

**Proof.** Simply exchange the sequence $\{s_m\}$ by the sequence $\{t_m\}$ used in [2, Theorem 3.4, page 401] which is defined for $t_0 = 0, t_1 = ß$ and each $m = 1, 2, \cdots$ by

$$t_{m+1} = t_m + \frac{M_1(t_m - t_{m-1})^2}{2(1 - M_1 t_m)}. \tag{4.5}$$ ■

**Remark 4.2.** As in Remark 3.2 a comparison is given between the present results and the corresponding ones in [2, see Theorem 3.4]. We have the following improvements:

$(I_1)$ The conditions $(C_1)$ and $(C_2)$ are weaker than $(C_3)$ (including the case when $M_1$ is replaced by $\bar{M}_1$).

$(I_2)$ The convergence condition in [2, Theorem 3.4 ] can be written as

$$p = ßM_1 \leq \frac{1}{2}. \tag{4.6}$$

This is a Kantorovich-type condition for the convergence of NM. Clearly, the conditions of the Lemma 4.1 are weaker than (4.6). However, there is a more direct comparison. In a series of papers [3, 4, 5], we weakened the Kantorovich sufficient semi-local convergence condition [15] for solving the equation $g(x) = 0$.

But the same majorant sequence $\{s_m\}$ is used (see e.g. [4]) and the corresponding convergence condition is

$$p_0 = ßK \leq \frac{1}{2}, \tag{4.7}$$

where $K = \frac{1}{8}\Big(4M_0 + \sqrt{K_0M_0 + 8M_0^2} + \sqrt{M_0K_0}\Big)$, where $K_0 = M$ or $M_1$. Notice that

$$K \leq M_1, \tag{4.8}$$

so

$$p \leq \frac{1}{2} \implies p_0 \leq \frac{1}{2} \tag{4.9}$$

but not necessarily vice versa unless if $K = M_1$. Implication (4.9) shows that the convergence conditions of the present paper are weaker than those in [2, 13, 15, 17]. Moreover, notice that as $\frac{M_0}{M_1} \to 0$, then $\frac{p_0}{p} \to \infty$. Therefore, the new conditions can be infinitely many times weaker than (4.6). Examples where $\frac{M_0}{M_1} \to 0$, $M < M_1$, $M_0 < M_1$ can be found in [3, 4, 5, 6].

$(I_3)$ Suppose that conditions (4.6) and (4.7) hold. Then, by (4.1), (4.5), and the definition of $M_0, M$ and $M_1$ a simple inductive argument implies

$$0 \leq s_m \leq t_m,$$
$$0 \leq s_{m+1} - s_m \leq t_{m+1} - t_m$$

and

$$s_* = \lim_{m \to +\infty} s_m \leq t_* = \lim_{m \to +\infty} t_m,$$

where $t_* = \frac{1}{M_1}(1 - \sqrt{1-2p})$. Thus, the new error estimates (4.3) and (4.4) are tighter than the ones in [2, 13, 15, 17] using the sequence $\{t_m\}$. Notice also that since $s_* \leq t_*$, the new information about the location of the solution is more precise. Moreover, the Tapia-type error bounds given [13, pages 402 and 403] can certainly be written in the new and more precise by simply exchanging (4.9), $\{t_n\}, \{t_*\}, p$ by (4.7), $\{s_n\}, \{s_*\}$ and $p_0$, respectively. Furthermore, the analysis of NM under Smale conditions [8, 22] can also be immediately extended under the new methodology in a similar way. We leave the details to the motivated reader.

## 5. Conclusions

The local as well as the semi-local analysis of convergence for NM using the Kantorovich theory for solving the generalized equation (1.1) exhibits the limitations $(e_1) - (e_3)$ as listed in the Introduction. In the present article motivated by optimization considerations; using more precise majorizing sequences and only conditions on the operators on NM we extended the applicability of NM as stated in $(e_1)' - (e_3)'$ (see also Remark 3.2 and Remark 4.2). This way, we can solve a larger class of problems than before [2, 13, 15, 17]. In our future research the Lipschitz-type conditions on $g'$ will be replaced by generalized continuity conditions [6]. This way the applicability of NM shall be extended even further. It is worth noting that the methodology of this article can be used to extend the applicability of methods other than Newton's [6, 15, 16, 17, 21, 23] along the same lines.

**Data availibility Statement:** No data is used in this study.

**Competing interests:** The authors declare that they have no conflicts of interest.

## References


[1] **Adly, S., R. Cibulka and H.V. Ngai,** Newton's method for solving inclusions using set-valued approximations, *SIAM J. Optim.*, **25(1)** (2015), 159–184. https://doi.org/10.1137/130926730

[2] **Adly, S., H. Van Ngai, and V.V. Nguyen,** Newton's method for solving generalized equations: Kantorovich's and Smale's approaches, *J. Math. Ana. Appl.*, **439(1)** (2016), 396–418. https://doi.org/10.1016/j.jmaa.2016.02.047

[3] **Argyros, I.K.,** On the Newton–Kantorovich hypothesis for solving equations, *J. Comput. Math.*, **169** (2004), 315–332. https://doi.org/10.1016/j.cam.2004.01.029

[4] **Argyros, I.K. and S. Hilout,** Weaker conditions for the convergence of Newton's method, *J. Complexity*, **28** (2012), 364–387. https://doi.org/10.1016/j.jco.2011.12.003

[5] **Argyros, I.K. and S. Hilout,** On an improved convergence analysis of Newton's method, *Appl. Math. Comp.*, **225** (2013), 372–386. https://doi.org/10.1016/j.amc.2013.09.049

[6] **Argyros, I.K.,** *The Theory and Applications of Iteration Methods*, second edition, Engineering series, Boca Raton, Florida, USA, 2022. https://doi.org/10.1201/9781003128915

[7] **Bonnans, J.F.,** Local analysis of Newton-type methods for variational inequalities and nonlinear programming, *Appl. Math. Optim.*, **29** (1994), 161–186. https://doi.org/10.1007/BF01204181

[8] **Dedieu, J.P.,** *Points Fixes, Zéros et la Méthode de Newton, Mathématiques & Applications*, vol. 54, Springer, Berlin, Heidelberg, 2006.

[9] **Dontchev, A.L.,** Local convergence of the Newton method for generalized equations, *C. R. Acad. Sci. Paris Sér. I Math.*, **322(4)** (1996), 327–331.

[10] **Dontchev, A.L.,** Local analysis of Newton-type method based on partial linearization, in: *Letters in Applied Mathematics*, vol. 32, Amer. Math. Soc., 1996, pp. 295–306.

[11] **Dontchev, A.L.,** Uniform convergence of the Newton method for Aubin continuous maps, *Serdica Math. J.*, **22(3)** (1996), 283–296.

[12] **Dontchev, A.L. and R.T. Rockafellar,** *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer Monographs in Mathematics, Springer, New York, 2009.
https://doi.org/10.1007/978-0-387-87821-8

[13] **Gragg, W.B. and R.A. Tapia,** Optimal error bounds for the Newton–Kantorovich theorem, *SIAM J. Numer. Anal.*, **11(1)** (1974), 10–13.
https://doi.org/10.1137/0711002

[14] **Josephy, N.H.,** *Newton's Method for Generalized Equations, Technical Summary Report*, Mathematics Research Center, University of Wisconsin, Madison, 1979.

[15] **Kantorovich, L.V. and G.P. Akilov,** *Analyse Fonctionnelle,* Mir, Moscou, 1981, translated from Russian.

[16] **Mordukhovich, B.S.,** *Variational Analysis and Generalized Differentiation I: Basic Theory, Grundlehren der Mathematischen Wissenschaften (A Series of Comprehensive Studies in Mathematics),* vol. 330, Springer, Berlin, Heidelberg, 2006.

[17] **Ortega, J.M. and W.C. Rheinboldt,** *Iterative Solution of Nonlinear Equations in Several Variables*, Classics in Applied Mathematics, vol. 30, Society for Industrial and Applied Mathematics, 2000.
https://doi.org/10.1137/1.9780898719468

[18] **Rashid, M.H., S.H. Yu, C. Li and S.Y. Wu,** Convergence analysis of the Gauss–Newton-type method for Lipschitz-like mappings, *J. Optim. Theory Appl.*, **158** (2013), 216–233.
https://doi.org/10.1007/s10957-012-0206-3

[19] **Rheinboldt, W.C.,** An adaptive continuation process for solving systems of nonlinear equations, in: *Mathematical Models and Numerical Methods* (A.N.Tikhonov et al. eds.) (1977), 129–142, Banach Center Publ. 3., Warsaw Poland. https://doi.org/10.4064/-3-1-129-142

[20] **Robinson, S.M.,** Strongly regular generalized equations, *Math. Oper. Res.*, **5(1)** (1980), 43–62. https://doi.org/10.1287/moor.5.1.43

[21] **Shakhno, S.M. and O.P. Gnatyshyn,** On an iterative algorithm of order 1.839... for solving the nonlinear least squares problems, *Applied Mathematics and Computation*, **161(1)** (2005), 253–264.
https://doi.org/10.1016/j.amc.2003.12.025

[22] **Smale, S.,** Newton's method estimates from data at one point, in: *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics*, Springer, New York, 1986, pp. 185–196. https://doi.org/10.1007/978-1-4612-4984-9_13

[23] **Traub, J.F.,** *Iterative Methods for the Solution of Equations*, Prentice Hall, New Jersey, U.S.A, 1964.

**Ioannis K. Argyros**
https://orcid.org/0000-0002-9189-9298
Department of Computing and Mathematical Sciences
Cameron University
Lawton, OK 73505
USA
iargyros@cameron.edu

**Santhosh George**
https://orcid.org/0000-0002-3530-5539
Department of Mathematical and Computational Sciences
National Institute of Technology Karnataka
Surathkal
India
sgeorge@nitk.edu.in