Extending the Kantorovich’s theorem on Newton’s method for solving strongly regular generalized equation

I. K. Argyros\textsuperscript{1} · G. N. Silva\textsuperscript{2}

Received: 4 July 2017 / Accepted: 20 April 2018 / Published online: 27 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract  The aim of this paper, is to extend the applicability of Newton’s method for solving a generalized equation of the type \( f(x) + F(x) \ni 0 \) in Banach spaces, where \( f \) is a Fréchet differentiable function and \( F \) is a set-valued mapping. The novelty of the paper is the introduction of a restricted convergence domain. Using the idea of a weaker majorant, the convergence of the method, the optimal convergence radius, and results of the convergence rate are established. That is we find a more precise location where the Newton iterates lie than in earlier studies. Consequently, the Lipschitz constants are at least as small as the ones used before. This way and under the same computational cost, we extend the semilocal convergence of the Newton iteration for solving \( f(x) + F(x) \ni 0 \). The strong regularity concept plays an important role in our analysis. We finally present numerical examples, where we can solve equations in cases not possible before without using additional hypotheses.

Keywords  Newton’s method · Generalized equation · Weaker majorant condition · Strong regularity · Kantorovich’s theorem

1 Introduction

In the present paper, we deal with the problem of finding a point \( x^* \in X \) satisfying the following generalized equation
where $f : \Omega \to \mathbb{Y}$ is a Fréchet differentiable function, $\mathbb{X}$ and $\mathbb{Y}$ are Banach spaces, $\Omega \subseteq \mathbb{X}$ is an open set and $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ is a set-valued mapping. It is well-known that the generalized equation (1) is an abstract model for a wide range of problems in mathematical programming. See, for instance, [3,5,11,12] as part of a whole. In the case $F \equiv \{0\}$, (1) becomes the nonlinear equation $f(x) = 0$. A particular case of problem (1) is when $F = -C$, where $C \subset \mathbb{X}$ is a nonempty closed convex cone. Thus, problem (1) becomes

$$f(x) \in C.$$  

If $F$ is the normal cone mapping $NC$ of a convex set $C$ in $\mathbb{Y}$, then (1) is a variational inequality problem, which covers a wide range of problems in nonlinear programming, including linear and nonlinear complementary problems; additional comments about such problems can be found in [3,5,7,11,12,18].

Newton’s method for solving (1) utilizes the iteration

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad k = 0, 1, \ldots$$  

for $x_0$, a given initial point. As it is well-known, the generalized equation (1) has considerable scope in classical analysis and its applications. When $F \equiv 0$, the iteration (3) becomes the standard Newton method for solving the nonlinear equation $f(x) = 0$,

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0, \quad k = 0, 1, \ldots$$  

Josephy, in his Ph.D. thesis [11] studied Newton’s method for solving $f(x) + NC(x) \ni 0$, where $f : \Omega \to \mathbb{R}^m$ is continuously differentiable, $\Omega \subseteq \mathbb{R}^n$ is an open set, and $C \subset \mathbb{R}^m$ is a convex set. To validate the definition of the sequence generated by the method, the strong regularity property on $f + NC$, a concept introduced by Robinson in [13], was used. If $\mathbb{X} = \mathbb{Y}$ and $NC = \{0\}$, then strong regularity at $x$ is equivalent to $f'(x)^{-1}$ being a continuous linear operator. An important case is when (1) represents Karush–Kuhn–Tucker’s systems for the standard nonlinear programming problem with a strict local minimizer. In this case, the strong regularity of this system, along with the condition that the primal variable is an optimal solution, is equivalent to the linear independence of the gradients of the active constraints and a strong second-order sufficient optimality condition; see [4, Theorem 6].

A usual assumption used to obtain quadratic convergence of Newton’s method (3), for solving equation (1), is the Lipschitz continuity of $f'$ in a neighborhood of the solution. Indeed, ensuring control of the derivative is an important consideration in the convergence analysis of Newton’s method. On the other hand, a couple of studies have dealt with the issue of convergence analysis of Newton’s method, for solving the equation $f(x) = 0$, by relaxing the assumption of Lipschitz continuity of $f'$, see for example [10,16,17].

The idea of the majorant function has been shown to be an appropriate and powerful tool for the convergence of Newton-like methods. The convergence domain for such methods is small in general. In the present study, we extend the convergence
domain for the Newton’s method. To achieve this goal, we first introduce the center-
Lipschitz condition which determines a subset of the original domain for the mapping
containing the iterates. The majorant functions are then related to the subset instead
of the original domain. This way, the majorant functions are more precise than if they
were depending on the original domain of the mapping as in earlier studies. The new
technique leads to: weaker sufficient convergence conditions, tighter error bounds on the
distance involved and an at least as precise information on the location of the solution.
These advantages are obtained under the same computational cost as in earlier studies,
since in practice the new majorant functions are special cases of the old majorant func-
tions. In [14,15], (1) was considered for F maximal monotone acting between Hilbert
spaces and a Kantorovich-type and local analysis were obtained under majorant con-
dition. In [8], under strong regularity at the solution of (1) and a majorant condition,
it was shown that (3) is locally convergent at a quadratic rate. Besides, another advan-
tage of working with a majorant condition rests in the fact that it allows unifying
of several convergence results pertaining to Newton’s method; see [10,16,17]. The
analysis presented provides a clear relationship between the majorant function and
the function defining the generalized equation. In addition, it allows us to obtain the
optimal convergence radius for the method with respect to the majorant condition and
uniqueness of the solution. The analysis of this method, under Lipschitz’s condition,
is provided as a special case. In the concluding Sect. 4, we have presented examples,
where earlier results [8–11,14,15] cannot be used to solve equations. However, our
results can apply to solve these equations and under the same computational cost.

The remainder of this paper is structured as follows. In Sect. 2, some notations and
important results used throughout the paper are presented. In Sect. 3, the main result
is stated and proved. In Sect. 4, some applications of this result are given.

2 Preliminaries

The following notations and results are used throughout the paper. The open and
closed balls at x with radius δ ≥ 0 are denoted, respectively, by B(x, δ) = {y ∈ X : ∥x − y∥ < δ} and B[x, δ] = {y ∈ X : ∥x − y∥ ≤ δ}. We denote by $\mathcal{L}(X, Y)$ the space consisting of all continuous linear mappings A : X → Y and the norm of
A by ∥A∥ := sup {∥Ax∥ : ∥x∥ ≤ 1}. Let Ω ⊆ X be an open set and f : Ω → Y
be Fréchet differentiable at all x ∈ Ω. The Fréchet derivative of f at x is the linear
mapping $f'(x) : X → Y$, which is continuous. The graph of the set-valued mapping
F : X → Y is the set gph F := {(x, y) ∈ X × Y : y ∈ F(x)}. The domain and
the range of F are, respectively, the sets dom F = {x ∈ X : F(x) ≠ ∅} and
rge F = {y ∈ Y : y ∈ F(x) for some x}. The inverse of F is the set-valued mapping
F⁻¹ : Y → X defined by F⁻¹(y) = {x ∈ X : y ∈ F(x)}. The partial linearization
of f + F at x ∈ Ω is the set-valued mapping $L_f(x, ·) : Ω → Y$ defined by

\[ L_f(x, y) := f(x) + f'(x)(y − x) + F(y). \]  (5)

Definition 1 Let Ω ⊆ X be open and nonempty. The mapping T : Ω → Y is called
strongly regular at x for y, when y ∈ T(x) and there exist $r_x > 0$, $r_y > 0$, and
\( \lambda > 0 \) such that \( B(x, r_x) \subset \Omega \), the mapping \( z \mapsto T^{-1}(z) \cap B(x, r_x) \) is single-valued from \( B(y, r_y) \) to \( B(x, r_x) \) and Lipschitzian on \( B(y, r_y) \) with modulus \( \lambda \), i.e.,

\[
\|T^{-1}(u) \cap B(x, r_x) - T^{-1}(v) \cap B(x, r_x)\| \leq \lambda \|u - v\|, \quad \text{for all } u, v \in B(y, r_y).
\]

Since \( z \mapsto T^{-1}(z) \cap B(x, r_x) \) in Definition 1 is single-valued, for the sake of simplicity, we have used the notation \( w = T^{-1}(z) \cap B(x, r_x) \) instead of \( \{w\} := T^{-1}(z) \cap B(x, r_x) \). Hereafter, we use this simplified notation. For a detailed discussion on Definition 1; see [5, 6, 13]. The next lemma is a type of Banach Perturbation Lemma, its proof is similar to [13, Theorem 2.4] and is omitted here.

**Lemma 1** Let \( X, Y \) be Banach spaces, \( a_0 \) be a point of \( Y \), \( F : X \rightarrow Y \) be a set-valued mapping and \( A_0 : X \rightarrow Y \) be a bounded linear mapping. Suppose that \( x \in X \) and \( 0 \in A_0 x + a_0 + F(x) \). Assume that \( A_0 + a_0 + F \) is strongly regular at \( x \) for 0 with modulus \( \lambda > 0 \). Then, there exist \( r_x > 0, r_{a_0} > 0 \), and \( r_0 > 0 \) such that, for any \( A \in B(A_0, r_{a_0}) \subset L(X, Y) \) and \( a \in B(a_0, r_0) \subset Y \) letting \( T(A, a, \cdot) : B(x, r_x) \rightarrow Y \) be defined as \( T(A, a, x) := Ax + a + F(x) \), the mapping \( y \mapsto T(A, a, y)^{-1} \cap B(\bar{x}, r_\bar{x}) \) is single-valued from \( B(0, r_0) \subset Y \) to \( B(\bar{x}, r_\bar{x}) \). Moreover, for each \( A \in B(A_0, r_{a_0}) \) and \( a \in B(a_0, r_0) \) there holds \( \lambda \|A - A_0\| < 1 \) and the mapping \( y \mapsto T(A, a, y)^{-1} \cap B(\bar{x}, r_\bar{x}) \) is also Lipschitzian on \( B(0, r_0) \) as follows

\[
\left\| T(A, a, y_1)^{-1} \cap B(\bar{x}, r_\bar{x}) - T(A, a, y_2)^{-1} \cap B(\bar{x}, r_\bar{x}) \right\| \leq \frac{\lambda}{1 - \lambda \|A - A_0\|} \|y_1 - y_2\|, \quad \forall \; y_1, y_2 \in B(0, r_0).
\]

Next, we establish a corollary to Lemma 1, which plays an important role in the sequel.

**Corollary 2** Let \( X, Y \) be Banach spaces, \( \Omega \subset X \) be open and nonempty, \( f : \Omega \rightarrow Y \) be continuous with the Fréchet derivative \( f' \) continuous, and \( F : X \rightarrow Y \) be a set-valued mapping. Suppose that \( x_0 \in \Omega \) and \( L_f(x_0, \cdot) : \Omega \rightarrow Y \) is strongly regular at \( x_1 \in \Omega \) for 0 with modulus \( \lambda > 0 \). Then, there exist \( r_{x_1} > 0, r_0 > 0 \), and \( r_{x_0} > 0 \) such that, for each \( x \in B(x_0, r_{x_0}) \), there holds \( \lambda \|f'(x) - f'(x_0)\| < 1 \), the mapping \( z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1}) \) is single-valued from \( B(0, r_0) \) to \( B(x_1, r_{x_1}) \) and Lipschitzian as follows

\[
\left\| L_f(x, u)^{-1} \cap B(x_1, r_{x_1}) - L_f(x, v)^{-1} \cap B(x_1, r_{x_1}) \right\| \leq \frac{\lambda}{1 - \lambda \|f'(x) - f'(x_0)\|} \|u - v\|, \quad \forall \; u, v \in B(0, r_0).
\]

**Proof** See [9]. \( \square \)

### 3 Kantorovich’s theorem for Newton’s method

In this section, our objective is to state and prove Kantorovich’s theorem for Newton’s method for solving (1). To state the theorem, we need to set some important constants. We refer to the real numbers
as the three constants given by Corollary 2.

Further, we assume that Lipschitz continuity of \( f' \) is relaxed, i.e., we assume that \( f' \) satisfies the conditions stated in the next definitions.

**Definition 2** Let \( X, Y \) be Banach spaces, \( \Omega \subset X \) be open, \( f : \Omega \to Y \) be continuous with Fréchet derivative \( f' \) continuous in \( \Omega \). Let \( x_0 \in \Omega, R > 0 \), and \( \kappa := \sup\{t \in [0, R) : B(x_0, t) \subset \Omega\} \). A twice continuously differentiable function \( \psi : [0, R) \to \mathbb{R} \) is a majorant function for \( f \) on \( B(x_0, \kappa) \) with modulus \( \lambda > 0 \), if it satisfies the following inequality

\[
\lambda \| f'(y) - f'(x) \| \leq \psi'(\|y - x\| + \|x - x_0\|) - \psi'(\|x - x_0\|),
\]

for all \( x, y \in B[x_0, \kappa] \) and \( \|y - x\| + \|x - x_0\| < R \). Moreover, there hold:

- (C1) \( \psi(0) > 0, \psi'(0) = -1 \), for each \( t \in [0, R) \);
- (C2) \( \psi' \) is strictly increasing and convex;
- (C3) \( \psi(t) = 0 \) for some \( t \in (0, R) \).

**Definition 3** Let \( X, Y \) be Banach spaces, \( \Omega \subset X \) be open, \( f : \Omega \to Y \) be continuous with Fréchet derivative \( f' \) continuous in \( \Omega \). Let \( x_0 \in \Omega, R > 0 \), and \( \kappa := \sup\{t \in [0, R) : B(x_0, t) \subset \Omega\} \). A twice continuously differentiable function \( \psi_0 : [0, R) \to \mathbb{R} \) is a center majorant function for \( f \) on \( B(x_0, \kappa) \) with modulus \( \lambda > 0 \), if it satisfies the following inequality

\[
\lambda \| f'(y) - f'(x_0) \| \leq \psi_0'(\|x - x_0\|) - \psi_0'(0),
\]

for all \( x, y \in B[x_0, \kappa] \). Moreover, suppose that

- (C1) \( \psi_0(0) > 0, \psi_0'(0) = -1, \psi_0(t) \leq \psi(t), \) \( \psi_0'(t) \leq \psi'(t); \)
- (C2) \( \psi_0' \) is strictly increasing and convex;
- (C3) \( \psi_0(t) = 0 \) for some \( t \in (0, R) \).

Notice that there is a significant difference between the proof of our Theorem 3 and the corresponding one in [9], since we use a more flexible and accurate function \( \psi_0 \) instead of the less precise function \( \psi \) leading to the already aforementioned advantages and under the same computational cost. The statement of the main result is:

**Theorem 3** Let \( x_0 \in \Omega, R > 0 \) and \( \kappa := \sup\{t \in [0, R) : B(x_0, t) \subset \Omega\} \). Assume that \( L_f(x_0, \cdot) : \Omega \to Y \) is strongly regular at \( x_1 \in \Omega \) for 0 with modulus \( \lambda > 0 \) and there exists a function \( \psi_0 : [0, R) \to \mathbb{R} \) that satisfies the conditions in Definition 3. Then, the sequence \( \{t_n^0\} \) generated by Newton’s method for solving \( \psi_0(t) = 0 \)

\[
t_{n+1}^0 = t_n^0 - \frac{\psi_0(t_n^0)}{\psi_0'(t_n^0)}, \quad k = 0, 1, \ldots,
\]

is strictly increasing, \( \{t_n^0\} \subset (0, t_0^*) \) and converges to \( t_0^* \), where \( t_0^* \) is the smallest solution of \( \psi_0(t) = 0 \) in \([0, R)\). Furthermore, suppose that there exists a function
\( \psi : [0, \ R) \rightarrow \mathbb{R} \) that satisfies the conditions in Definition 2 for all \( x, y \in B_0 := B[x_0, \kappa] \cap B[x_0, t_0^+] \) and
\[
\|x_1 - x_0\| \leq \psi(0). \quad (10)
\]

Let
\[
\beta := \sup_{t \in [0, R)} -f(t), \quad t_* := \min \psi^{-1}([0]).
\]

Additionally, for the constants \( r_0 \) and \( r_{x_0} \), suppose that the following inequalities hold:
\[
t_* \leq r_{x_0}, \quad \frac{\psi''(t_*)}{2\lambda} \psi(0)^2 < r_0. \quad (11)
\]

Then, the sequences generated by Newton’s method for solving \( 0 \in f(x) + F(x) \) and \( \psi(t) = 0 \), with starting point \( x_0 \) and \( t_0 = 0 \), defined respectively by,
\[
x_{k+1} := L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_1}), \quad t_{k+1} = t_k - \frac{\psi(t_k)}{\psi'(t_k)}, \quad k = 0, 1, \ldots , \quad (12)
\]
are well defined, \( \{t_k\} \) is strictly increasing, \( \{t_k\} \subset (0, t_*) \) and converges to \( t_* \), and \( \{x_k\} \subset B(x_0, t_*) \) and converges to \( x_* \in B[x_0, t_*] \), which is the unique solution of \( 0 \in f(x) + F(x) \) in \( B[x_0, t_*] \cap B[x_1, r_{x_1}] \). Moreover, \( \{x_k\} \) and \( \{t_k\} \) satisfies
\[
\|x_* - x_k\| \leq t_* - t_k, \quad \|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2} \|x_* - x_k\|^2, \quad (13)
\]
for all \( k = 0, 1, \ldots \), and the sequences \( \{x_k\} \) and \( \{t_k\} \) converge \( Q \)-linearly as follows
\[
\|x_* - x_{k+1}\| \leq \frac{1}{2} \|x_* - x_k\|, \quad t_* - t_{k+1} \leq \frac{1}{2} (t_* - t_k), \quad k = 0, 1, \ldots . \quad (14)
\]

Additionally, if the following condition holds

\( \textbf{(C4)} \quad \psi'(t_*) < 0, \)

then the sequences, \( \{x_k\} \) and \( \{t_k\} \) converge \( Q \)-quadratically as follows
\[
\|x_* - x_{k+1}\| \leq \frac{\psi''(t_*)}{-2\psi'(t_*)} \|x_* - x_k\|^2, \quad \|x_* - x_{k+1}\| \leq \frac{\psi''(t_*)}{-2\psi'(t_*)} \|x_* - x_k\|^2 \quad (15)
\]
\[
t_* - t_{k+1} \leq \frac{\psi''(t_*)}{-2\psi'(t_*)} (t_* - t_k)^2, \quad t_* - t_{k+1} \leq \frac{\psi''(t_*)}{-2\psi'(t_*)} (t_* - t_k)^2 \quad (16)
\]

Remark 1 (a) In the earlier study [9] the following condition was studied instead of (8):
\[
\lambda \| f'(y) - f'(x) \| \leq \psi'(\|y - x\| + \|x - x_0\|) - \psi'(\|x - x_0\|), \quad (17)
\]
Extending the Kantorovich’s theorem on Newton’s method...

for all $\tau \in [0, 1]$, $x \in B(x_0, \kappa)$ where $\psi_1 : [0, R) \to \mathbb{R}$ is twice continuously differentiable with $\psi_1(0) > 0$, $\psi_1'(0) = -1$ and $\psi_1'$ convex and strictly increasing. The corresponding error estimates for

$$ u_{k+1} = \left| \frac{u_k \psi_1'(u_k) - \psi_1(u_k)}{\psi_1'(u_k)} \right|, \quad u_0 = \|x^* - x_0\| $$  \hspace{1cm} (18)

are:

$$ \|x^* - x_{k+1}\| \leq \frac{u_{k+1}}{u_k} \|x^* - x_k\|^2, \quad \frac{u_{k+1}}{u_k} \leq \frac{\psi_1''(u_0)}{2|\psi_1'(0)|} $$  \hspace{1cm} (19)

$$ \|x^* - x_k\| \leq u_0 \left( \frac{u_1}{u_0} \right)^{2^{k-1}}, \quad k = 0, 1, \ldots, \text{ and} $$  \hspace{1cm} (20)

the optimal convergence radius is $r = \rho_1$ if $\rho_1 < \kappa$ and solves

$$ \frac{\psi_1(t) - t \psi_1'(t)}{t \psi_1'(t)} = 1. $$  \hspace{1cm} (21)

By comparing (8) and (7) to (17), we see that

$$ \psi'(t) \leq \psi_1'(t) $$  \hspace{1cm} (22)

and

$$ \psi_0'(t) \leq \psi_1'(t) $$  \hspace{1cm} (23)

since, $B_0 \subseteq B(x^*, \kappa)$. Define functions $\varphi$ and $\varphi_1$ on $B[0, R]$ by

$$ \varphi(t) = \frac{\psi(t) - t \psi'(t)}{t \psi_0'(t)} - 1 $$

and

$$ \varphi_1(t) = \frac{\psi_1(t) - t \psi_1'(t)}{t \psi_1'(t)} - 1. $$

Then, we have in turn that

$$ \varphi(t) = \frac{\int_0^1 [\psi'(\tau t) - \psi'(t)]d\tau}{\psi_0'(t)} - 1 \leq \frac{\int_0^1 [\psi_1'(\tau t) - \psi_1(t)]d\tau}{\psi_1'(t)} - 1 = \varphi_1(t). $$  \hspace{1cm} (24)

In particular, we have that

$$ \varphi(\rho_1) \leq \varphi_1(\rho_1) = 0 $$  \hspace{1cm} (25)
so \( \rho_1 \leq \rho \). Moreover, in view of (22)–(24), the new error estimates are tighter than the corresponding ones given by (18)–(20) leading to more precise error bounds on the distances \( \|x^* - x_{k+1}\| \) and \( \|x^* - x_k\| \), i.e., at least as few iterates to obtain a desired error tolerance. Moreover, the information on the uniqueness of the solution is more precise, since

\[ \bar{\sigma} \leq \tilde{\sigma}, \]

where \( \sigma_1 := \sup\{t \in (0, r) : \psi_1(t) < 0\}, \bar{\sigma} = \min\{r_1, \rho_1\}, r_1 = \min\{\kappa, \rho_1\} \), if

\[ \psi_0(t) \leq \psi_1(t), \quad t \in [0, R). \]  

(26)

Condition (26) is assumed without loss of generality. It is also worth noticing that the preceding advantages are obtained under the same computational cost as in [9], since in practice the computation of function \( \psi_1 \) involves the computations of functions \( \psi_0 \) and \( \psi \) as special cases (see also the numerical example). Finally, notice that \( \psi = \psi(\psi_0) \). That is, \( \psi \) depends on function \( \psi_0 \). The construction of \( \psi \) was not possible before without \( \psi_0 \), since only \( \psi_1 \) was used [9].

(b) (C′1) If \( \psi'(t) \leq \psi_0'(t), \ t \in [0, t_0^*) \) holds instead of (C1), then clearly, the conclusions of Theorem 3 hold with (C′1) replacing (C1) and function \( \psi_0 \) replacing \( \psi \) in the conclusions of this theorem. Moreover, the advantages stated in (a) over the results in [9] hold with \( \psi_0 \) replacing \( \psi \).

Remark 2 All the results about the sequences \( \{t_n^0\} \) and \( \{t_n\} \) are easily obtained as in [9].

In Sect. 4, we present particular instances of Theorem 3 for some classes of functions in the above examples. Hereafter, we assume that all the assumptions in Theorem 3 hold.

3.1 Some results

Now, we are going to establish some relationships between \( \psi_0 \) and \( f + F \). The next result is a consequence of Corollary 2.

Proposition 4 For any \( x \in B(x_0, t_*), \) the mapping \( z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1}) \) is single-valued from \( B(0, r_0) \) to \( B(x_1, r_{x_1}) \) and there holds

\[ \left\| L_f(x, u)^{-1} \cap B(x_1, r_{x_1}) - L_f(x, v)^{-1} \cap B(x_1, r_{x_1}) \right\| \leq -\frac{\lambda}{\psi_0'(\|x - x_0\|)} \|u - v\|, \quad \forall \ u, v \in B(0, r_0). \]

Proof Definitions of \( r_{x_1}, r_0, \) and \( r_{x_0} \) in (6) together with Corollary 2 imply that, for any \( x \in B(x_0, r_{x_0}) \), the mapping \( z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1}) \) is single-valued from \( B(0, r_0) \) to \( B(x_1, r_{x_1}) \) and there holds

\( \square \) Springer
Extending the Kantorovich’s theorem on Newton’s method...

\[
\left\| L_f(x, u)^{-1} \cap B(x_1, r_{x_1}) - L_f(x, v)^{-1} \cap B(x_1, r_{x_1}) \right\| \\
\leq \frac{\lambda}{1 - \lambda \| f'(x) - f'(0) \|} \| u - v \|, \tag{27}
\]

for all \( u, v \in B(0, r_0) \). Since \( \| x - x_0 \| < t_\ast \) thus \( \psi'_0(\| x - x_0 \|) < 0 \). Hence, (7) together with \( C_1^0 \) imply that

\[
\lambda \| f'(x) - f'(0) \| \leq \psi'_0(\| x - x_0 \|) - \psi'_0(0) < 1, \quad \forall x \in B(x_0, t_\ast),
\]

and then, using (11), i.e., \( t_\ast \leq r_{x_0}, (27) \) and \( C_1^0 \), the inequality of the proposition follows. \( \square \)

The above proposition guarantees that, for \( x \in B(x_0, t_\ast) \), the mapping \( z \mapsto L_f(x, z)^{-1} \cap B(x_1, r_{x_1}) \) is single-valued from \( B(0, r_0) \) to \( B(x_1, r_{x_1}) \). Thus, we define the Newton iteration mapping \( N_{f + F} : B(x_0, t_\ast) \to X \) by

\[
N_{f + F}(x) := L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}). \tag{28}
\]

It is easy to see that the definition of \( N_{f + F} \) can be rewritten as the following inclusions

\[
0 \in f(x) + f'(x)(N_{f + F}(x) - x) + F(N_{f + F}(x)), \quad N_{f + F}(x) \in B(x_1, r_{x_1}),
\]

\[
x \in B(x_0, t_\ast). \tag{29}
\]

Therefore, one can apply a single Newton iteration on any \( x \in B(x_0, t_\ast) \) to obtain \( N_{f + F}(x) \), which may not belong to \( B(x_0, t_\ast) \). Thus, this is adequate to ensure the well-definedness of only one Newton iteration. To ensure that Newtonian iterations may be repeated indefinitely or, in particular, invariant on subsets of \( B(x_0, t_\ast) \), we need some additional results. First, define some subsets of \( B(x_0, t_\ast) \), in which, as we shall prove, Newton iteration mapping (28) are “well behaved”. Define

\[
K(t) := \left\{ x \in \Omega : \| x - x_0 \| \leq t, \quad \| L_f(x, 0)^{-1} \cap B(x_1, r_{x_1}) - x \| \leq \frac{\psi(t)}{\psi'_0(t)} \right\},
\]

\[
t \in [0, t_\ast),
\]

\[
K := \bigcup_{t \in [0, t_\ast)} K(t). \tag{30}
\]

**Proposition 5** For each \( 0 \leq t < t_\ast \) we have \( K(t) \subset B(x_0, t_\ast) \) and \( N_{f + F}(K(t)) \subset K(\psi(t)) \). As a consequence, \( K \subset B(x_0, t_\ast) \) and \( N_{f + F}(K) \subset K \).

**Proof** The proof is similar as in [9]. \( \square \)

### 3.2 Convergence analysis

To prove the convergence results, which are consequences of the above results, first, we note that the definition (28) implies that the sequence \( \{ x_k \} \) defined in (12), can be formally stated as

\[
\square \text{ Springer}
\]
\[ x_{k+1} = N_{f+F}(x_k), \quad k = 0, 1, \ldots, \tag{32} \]

or equivalently as,

\[ 0 \in f(x_k) + f'(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad x_{k+1} \in B(x_1, r_{x_1}), \quad k = 0, 1, \ldots. \tag{33} \]

First, we show that the sequence \( \{x_k\} \) generated by Newton’s method converges to \( x_* \in B[x_0, t_*] \), a solution of the generalized equation (1), and is well behaved with respect to the set defined in (30).

**Proof of Theorem 3** Since the mapping \( x \mapsto L_f(x_0, x) \) is strongly regular at \( x_1 \) for 0, thus Corollary 2 implies that

\[ x_1 = L_f(x_0, 0)^{-1} \cap B(x_1, r_{x_1}) \]

and the first Newton iterate is well defined. Using, \( C_0^0 \), (10), (30) and (31) it is easy to see that

\[ \{x_0\} = K(0) \subset K. \tag{34} \]

From Proposition 5 we have that \( N_{f+F}(K) \subset K \). Thus, (34) and (32) imply that the sequence \( \{x_k\} \) is well defined and \( x_k \in K \). From the first inclusion in the second part of Proposition 5, we have that \( \{x_k\} \subset B(x_0, t_*) \). By induction, we can prove that

\[ x_k \in K(t_k), \quad k = 0, 1 \ldots \tag{35} \]

Now, using (35) and (30), combined with (32), (28), and (12), we have

\[
\|x_{k+1} - x_k\| = \|L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_1}) - x_k\| \\
\leq -\frac{\psi(t_k)}{\psi_0(t_k)} \leq -\frac{\psi(t_k)}{\psi'(t_k)} = t_{k+1} - t_k, \quad k = 0, 1 \ldots, \tag{36}
\]

Taking into account that \( \{t_k\} \) converges to \( t_* \), we can easily conclude from the above inequality that

\[
\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\| \leq \sum_{k=k_0}^{\infty} t_{k+1} - t_k = t_* - t_{k_0} < +\infty,
\]

for any \( k_0 \in \mathbb{N} \). Hence, we conclude that \( \{x_k\} \) is a Cauchy sequence in \( B(x_0, t_*) \) and thus it converges to some \( x_* \in B[x_0, t_*] \). Therefore, using (36) again, we conclude that \( \|x_* - x_k\| \leq t_* - t_k \), for all \( k = 0, 1 \ldots \). Since \( f \) and \( f' \) are continuous in \( \Omega \), \( B[x_0, t_*] \subset \Omega \) and \( F \) has a closed graph, it is easy to conclude that \( f(x_*) + F(x_*) \ni 0 \). The others statements of the theorem are similar to [9]. \( \square \)

### 4 Special cases

In this section, we will present some special cases of Theorem 3. It is worth pointing out that there exist some classes of well known functions which a majorant function
is available, below we will present two examples, namely, the classes of functions satisfying a Lipschitz-like and Smale’s conditions, respectively. In this sense, the results obtained in Theorem 3 unify the convergence analysis of Newton’s method for the classes of generalized equations involving these functions, for instance, Theorem 2 of [11] due to Josephy and, a particular instance of Theorem 2 of [3] due to Dontchev and a version of Smale’s theorem on Newton’s method for analytical functions, see [2].

**Theorem 6** Let \( \Omega \subset \mathbb{R}^n \), \( x_0 \in \Omega \), \( \lambda > 0 \), and \( f : \Omega \to \mathbb{R}^n \) be an analytic function. Let \( F : \mathbb{R}^n \to \mathbb{R}^n \) be a set-valued mapping with a closed graph. Suppose that \( L_f(x_0, \cdot) : \Omega \to \mathbb{R}^n \) at \( x_0 \), is strongly regular at \( x_1 \in \Omega \) for 0 with modulus \( \lambda > 0 \), \( B(x_0, 1/\gamma) \subseteq \Omega \), where \( \gamma = \sup_{n \geq 1} ||f^{(n)}(x_0)||/n!||^{1/(n-1)} < +\infty \) and there exists \( b > 0 \) such that \( ||x_1 - x_0|| \leq b \) and \( by \leq 3 - 2\sqrt{2} \). Suppose that

\[
\frac{4^3 \gamma b^2}{\lambda \left(3 - by + \sqrt{(by + 1)^2 - 8by}\right)^3} < r_0,
\]

hold, where \( t_* = (by + 1 - \sqrt{(by + 1)^2 - 8by})/4\gamma \). Then, the sequence generated by Newton’s method for solving \( f(x) + F(x) \equiv 0 \) with starting point \( x_0 \), \( x_{k+1} := L_f(x_k, 0)^{-1} \cap B(x_1, r_{x_k}) \), for all \( k = 0, 1, \ldots \), is well defined, \( \{x_k\} \) is contained in \( B(x_0, t_*) \), and converges to the point \( x_* \), which is the unique solution of \( f(x) + F(x) \equiv 0 \) in \( B[x_0, t_0] \cap B[x_1, r_{x_1}] \), where \( r_{x_1} \) is fixed in (6). Moreover, letting \( \psi_0 : [0, R) \to \mathbb{R} \), \( \psi : [0, 1/\gamma_0) \to \mathbb{R} \) be defined by \( \psi_0(t) = t/(1-\gamma_0t) - 2t + b \), \( \psi(t) = t/(1-\gamma t) - 2t + b \), with \( \gamma_0 \leq \gamma \), the sequence \( \{x_k\} \) converges \( Q \)-linearly as follows \( ||x_* - x_{k+1}|| \leq \|x_* - x_k\|/2 \), for all \( k = 0, 1, \ldots \). Additionally, if \( by < 3 - 2\sqrt{2} \), then \( \{x_k\} \) converges \( Q \)-quadratically as follows

\[
||x_* - x_{k+1}|| \leq \frac{\gamma_0}{(1-\gamma_0t_*)[2(1-\gamma t_*)^2 - 1]} ||x_* - x_k||^2, \quad k = 0, 1, \ldots
\]

**Proof** Note that \( \psi_0 \) and \( \psi \) satisfy \( C_1^0, C_2^0, C_3^0, C_4^0 \), and (11). Therefore, the result follows from the Theorem 3. \( \square \)

**Remark 3** Similarly, we obtain the improvements of Wang’s theory, if we choose for some \( \gamma_0 \leq \gamma_1 \) and \( \gamma_0 \leq \gamma \) the functions

\[
\psi_0(t) = \frac{t}{1-\gamma_0t} - 2t + b, \quad (37)
\]

\[
\psi(t) = \frac{t}{1-\gamma t} - 2t + b \quad (38)
\]

and

\[
\psi_1(t) = \frac{t}{1-\gamma_1t} - 2t + b, \quad (39)
\]

which can easily be seen to satisfy the conditions of Theorem 3 and those of Remark 1 (a) (or (b) if \( \gamma \leq \gamma_0 \).
The rest of the results in [9] can be improved as long as $B_0$ is a strict subset of $B[x_0, \kappa]$ along the same lines, since again the $\psi$ function defined on $B_0$ is at least as tight as the $\psi$ function defined on $B[x_0, \kappa]$ used in [9].

We present an academic and motivational numerical example, where the previously stated advantages in Remark 1 are obtained, when $F \equiv \{0\}$.

**Example 1** Let $X = Y = \mathbb{R}$, $x_0 = 1$, $p \in [0, \frac{1}{2})$ and $\Omega = \bar{U}(x_0, 1 - p)$. Define function $f$ on $\Omega$ by

$$f(x) = x^3 - p.$$  (40)

Then, we have by (17) that $\lambda = \frac{1}{3}$, $\kappa = 1 - p$ and $\psi_1(t) = (2 - p)t^2 - t + \frac{1}{3}(1 - p)$. The sufficient convergence Kantorovich condition is given by

$$h_1 = \frac{4}{3}(2 - p)(1 - p) \leq 1$$  (41)

which however is not satisfied for all $p \in [0, \frac{1}{2})$. Hence, there is no guarantee that Newton’s method starting at $x_0 = 1$ converges to $x^* = \sqrt[3]{p}$ under the earlier study [9]. Let us see that we can get under new approach. Condition (8) is satisfied, if we choose

$$\psi_0(t) = \frac{3 - p}{2}t^2 - t + \frac{1}{3}(1 - p).$$

The corresponding sufficient convergence condition is given by

$$h_0 = 4 \left(\frac{3 - p}{2}\right) \frac{1}{3}(1 - p) \leq 1$$  (42)

which is satisfied provided that $p \in S := \left[\frac{4 - \sqrt{10}}{2}, \frac{1}{2}\right)$. Notice that (42) is the sufficient condition for the modified Newton’s method $f(x_k) + f'(x_0)(x_{k+1} - x_k) = 0$. In this case

$$t_0^* = \frac{1 - \sqrt{1 - h_0}}{3 - p} < \kappa,$$  (43)

so $B[x_0, \kappa] \cap B[x_0, t_0^*] = B[x_0, t_0^*]$. Therefore, condition (7) is satisfied, if we choose

$$\psi(t) = (1 + t_0^*)t^2 - t + \frac{1}{3}(1 - p).$$

The sufficient convergence condition is given by

$$h = 4(1 + t_0^*) \frac{1}{3}(1 - p) \leq 1,$$  (44)

which is satisfied provided that $p \in \left[1 - \frac{3}{4(1 + t_0^*)}, \frac{1}{2}\right]$ or if $p \in S$. Notice also that we have

$$\psi_0(t) < \psi(t) < \psi_1(t) \quad \text{for each} \quad t \in [0, \frac{1}{2}).$$  (45)
If we allow $p \in [0, 1)$, say e.g. $p = 0.6$, then conditions (41), (42) and (44) are satisfied. However, as already noted in Remark 1, in view of (45) the other advantages of our approach over the ones in [9] hold.

Hence, we have extended the applicability of the method for solving equation $f(x) = 0$.

5 Final remarks

In this paper, a semi local convergence result to Newton’s method for solving (1) with a restricted convergence domain has been obtained, extending the applicability of the method under the same computational cost as in [9]. This technique can be used on other iterative methods [1]. In the future, we aim to study this method using the approach of this paper under a weak assumption rather than strong regularity, namely, the regularity metric or strong metric subregularity; see [5,6].

References

1. Argyros, I.K., Magreñán, Á.A.: Iterative Methods and Their Dynamics with Applications: A Contemporary Study. CRC Press, New York (2017)
2. Blum, L., Cucker, F., Shub, M., Smale, S.: Complexity and real computation. Springer, New York (1998). With a foreword by Richard M. Karp
3. Dontchev, A.L.: Local analysis of a Newton-type method based on partial linearization. In: The Mathematics of Numerical Analysis (Park City, UT, 1995). Lectures in Applied Mathematics, vol. 32, pp. 295–306. American Mathematical Society, Providence (1996)
4. Dontchev, A.L., Rockafellar, R.T.: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. 6, 1087–1105 (1996)
5. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings. Springer Monographs in Mathematics. A View from Variational Analysis. Springer, Dordrecht (2009)
6. Dontchev, A.L., Rockafellar, R.T.: Newton’s method for generalized equations: a sequential implicit function theorem. Math. Program. 123(1, Ser. B), 139–159 (2010)
7. Ferreira, O.P., Silva, G.N.: Inexact Newton’s method for nonlinear functions with values in a cone. Appl. Anal. (2018). https://doi.org/10.1080/00036811.2018.1430779
8. Ferreira, O.P., Silva, G.N.: Local convergence analysis of Newton’s method for solving strongly regular generalized equations. J. Math. Anal. Appl. 458(1), 481–496 (2018)
9. Ferreira, O.P., Silva, G.N.: Kantorovich’s theorem on Newton’s method for solving strongly regular generalized equation. SIAM J. Optim. 27(2), 910–926 (2017)
10. Ferreira, O.P., Svaiter, B.F.: Kantorovich’s majorants principle for Newton’s method. Comput. Optim. Appl. 42(2), 213–229 (2009)
11. Josephy, N.: Newton’s Method for Generalized Equations and the PIES Energy Model. University of Wisconsin-Madison (1979)
12. Robinson, S.M.: Extension of Newton’s method to nonlinear functions with values in a cone. Numer. Math. 19, 341–347 (1972)
13. Robinson, S.M.: Strongly regular generalized equations. Math. Oper. Res. 5(1), 43–62 (1980)
14. Silva, G.N.: Kantorovich’s theorem on Newton’s method for solving generalized equations under the majorant condition. Appl. Math. Comput. 286, 178–188 (2016)
15. Silva, G.N.: Local convergence of Newton’s method for solving generalized equations with monotone operator. Appl. Anal. (2017). https://doi.org/10.1080/00036811.2017.1299860
16. Wang, X.: Convergence of Newton’s method and inverse function theorem in Banach space. Math. Comput. 68(225), 169–186 (1999)
17. Zabrejko, P.P., Nguyen, D.F.: The majorant method in the theory of Newton–Kantorovich approximations and the Pták error estimates. Numer. Funct. Anal. Optim. 9(5–6), 671–684 (1987)
18. Zhang, Y., Wang, J., Guu, S.: Convergence criteria of the generalized Newton method and uniqueness of solution for generalized equations. J. Nonlinear Convex Anal. 16(7), 1485–1499 (2015)