On the convergence of extended Newton-type method for solving variational inclusions

M.H. Rashid

Abstract: In this paper, we introduce and study the extended Newton-type method for solving the variational inclusion

$$0 \in f(x) + g(x) + F(x),$$

where $f: \Omega \subseteq X \to Y$ is Fréchet differentiable in a neighborhood $\Omega$ of a point $x$ in $X$, $g: \Omega \subseteq X \to Y$ is linear and differentiable at point $x$, and $F$ is a set-valued mapping with closed graph acting in Banach spaces $X$ and $Y$. Semilocal and local convergence of the extended Newton-type method are analyzed.

Subjects: Mathematics & Statistics; Advanced Mathematics; Pure Mathematics

Keywords: variational inclusions; semilocal convergence; Lipschitz–like mappings; extended Newton-type method; divided difference

AMS Subject Classifications: 49J53, 47H04, 65K10

1. Introduction

In this study, we are concerned with the problem of approximating a solution of a variational inclusions. Let $X$ and $Y$ be Banach spaces. We consider here a variational inclusions problem to find a point $\bar{x} \in \Omega$ satisfying

$$f(x) + g(x) + F(x) \ni 0$$

Received: 10 July 2014
Accepted: 15 October 2014
Published: 24 November 2014

© 2014 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license.
\[ 0 \in f(\bar{x}) + g(\bar{x}) + F(\bar{x}), \]  
(1.1)

where \( f : \Omega \subseteq X \to Y \) is differentiable in a neighborhood \( \Omega \) of a point \( \bar{x} \) in \( X \), \( g : \Omega \subseteq X \to Y \) is linear and differentiable at \( \bar{x} \) but may not differentiable in a neighborhood \( \Omega \) of \( \bar{x} \), and \( F : X \to 2^Y \) is a set-valued mapping with closed graph.

When \( g = 0 \), the inclusion (1.1) reduce to a generalized equation of the form

\[ 0 \in f(\bar{x}) + F(\bar{x}). \]  
(1.2)

Several iterative methods have been presented for solving (1.2). Dontchev (1996b) established a quadratically convergent Newton-type method under a pseudo-Lipschitz property for set-valued mapping when \( \nabla f \) is Lipschitz on a neighborhood of a solution \( x^* \) of (1.2) and subsequently he (1996c) proved the stability of the method. When \( \nabla f \) is Hölder on a neighborhood of \( x^* \), Piétrus (2000a) obtained superlinear convergence by following Dontchev’s method and later he (2000b) proved the stability of this method in this mild differentiability context. For solving (1.2), Hilout, Alexis, and Piétrus (2006) considered the following sequence

\[
\begin{cases}
  x_0 \text{ and } x_1 \text{ are given starting points} \\
  y_k = \alpha x_k + (1-\alpha)x_{k-1}; \quad \alpha \text{ is fixed in } (0,1) \\
  0 \in f(x_k) + [y_k, x_k] f(x_{k+1} - x_k) + F(x_{k+1}),
\end{cases}
\]

where \([y_k, x_k] f\) is the first-order divided difference of \( f \) on the points \( y_k \) and \( x_k \). This operator will be defined in Section 2. They proved the convergence of this method is superlinear when \( f \) is only continuous and differentiable at \( x^* \). Furthermore, it should be mentioned that Argyros (2004) has studied local as well as semilocal convergence analysis for two-point Newton-like methods in a Banach space setting under very general Lipschitz type conditions for solving (Argyros, 2004) in the case when \( F = 0 \).

Alexis and Piétrus (2008) introduced a method for solving the variational inclusions (1.1), which can be defined as follows:

\[ 0 \in f(x_k) + g(x_k) + (\nabla f(x_k) + [2x_{k+1} - x_k, x_k] g)(x_{k+1} - x_k) + F(x_{k+1}), \]  
(1.3)

where \( \nabla f(x) \) denotes the Fréchet derivative of \( f \) at \( x \) and \([x, y] g\) the first order divided difference of \( g \) on the points \( x \) and \( y \); and proved the convergence is superlinear and quadratic when \( \nabla f \) is Lipschitz continuous. Rashid, Wang, and Li (2012) established local convergence results for the method (1.3) under the weaker conditions than Alexis and Pietrus (2008). In particular, Rashid et al. (2012) extended the results by fixing a gap in the proof of corresponding ones (Alexis & Piétrus, 2008, Theorem 1).

Although the method (1.3) guarantees the existence of a convergent sequence \( \{x_k\} \), the points \( x_1, x_2, \ldots \) of the sequence \( \{x_k\} \) are not converges separately. Therefore, for a starting point near to a solution, the sequences generated by the method (1.3) are not uniquely defined. For instance, the convergence result, obtained by Alexis and Piétrus (2008) or Rashid et al. (2012), guarantees the existence of a convergent sequence. Hence, in view of numerical computation, these kind of methods are not convenient in practical application. This drawback motivates us to introduce a method ‘so-called’ extended Newton-type method. The difference between the method (1.3) and our proposed method is that the extended Newton-type method generates a convergent sequence \( \{x_k\} \) whose each point \( x_1, x_2, \ldots \) converge individually but this does not happen for the method (1.3).

Thus, we propose the following extended Newton-type method:
Algorithm 1.1. Let $\eta \geq 1$ and $x_0 \in X$. For $k = 0, 1, 2, \ldots$, having $x_k$, determine $x_{k+1}$ as follows:

Choose $d_k \in D(x_k)$ such that

$$\|d_k\| \leq \eta \text{dist}(0, D(x_k)),$$

and set $x_{k+1} = x_k + d_k$, where the set $D(x)$ is defined for any $x \in X$ by

$$D(x) = \left\{ d \in X : 0 \in f(x) + g(x) + (\nabla f(x) + [2d + x, x, g])d + F(x + d) \right\}.$$

Remark 1.1. If $g = 0$, then the set $D$ will be replaced by the set

$$D'(x) := \left\{ d \in X : 0 \in f(x) + \nabla f(x)d + F(x + d) \right\}.$$

Then the Algorithm 1.1 reduces to the same algorithm corresponding one given by Rashid, Yu, Li, and Wu (2013).

There have been studied many fruitful works on semilocal convergence analysis for the Gauss–Newton method in the case when $F = 0$ and $g = 0$ [see Dedieu and Kim (2002), Dedieu and Shub (2000), Xu and Li (2008), for more details] or when $F = C$ and $g = 0$ [see Li and Ng (2007), for details]. Rashid et al. (2013) have studied semilocal convergence analysis for the Gauss–Newton-type method to solve the generalized Equation (1.2). However, in our best knowledge, there is no study on semilocal convergence analysis discovered for the general case (1.1), even for the method (1.3).

Our purpose here is to analyze the semilocal convergence of the extended Newton-type method defined by Algorithm 1.1. The main tool is the Lipschitz-like property of set-valued mappings, which was introduced by Aubin (1984) in the context of nonsmooth analysis and studied by many mathematicians [see for example, Alexis and Piétrus (2008), Argyros and Hilout (2008), Dontchev (1996a), Hilout et al. (2006), Piétrus (2000b)] and the references therein. The main results are the convergence criteria, established in Section 3, which, based on the attraction region around the initial point, provide some sufficient conditions ensuring the convergence to a solution of any sequence generated by Algorithm 1.1. As a result, local convergence results for the extended Newton-type method are obtained.

This paper is organized as follows: In Section 2, we recall some necessary notations, notions, some preliminary results and also recall a fixed-point theorem which has been proved by Dontchev and Hager (1994). This fixed-point theorem is the main tool to prove the existence of the sequence generated by Algorithm 1.1. In Section 3, we consider the extended Newton-type method as well as the concept of Lipschitz-like property to show the existence and the convergence of the sequence generated by Algorithm 1.1. In the last section, we give a summary of the major results to close our paper.

2. Notations and preliminaries

In this section, we give some notations and collect some results that will be helpful to prove our main results. Throughout this paper, we suppose that $X$ and $Y$ are two real or complex Banach spaces. Let $x \in X$ and $r > 0$. The closed ball centered at $x$ with radius $r$ is denoted by $B_r(x)$. Let $F: X \rightrightarrows 2^Y$ be a set-valued mapping with $\text{dom } F \neq \emptyset$. The domain $\text{dom } F$, the inverse $F^{-1}$ and the graph $\text{gph } F$ of $F$ are, respectively, defined by

$$\text{dom } F := \{ x \in X : F(x) \neq \emptyset \},$$

$$F^{-1}(y) := \{ x \in X : y \in F(x) \} \text{ for each } y \in Y$$

and
Let $A \subseteq X$. The distance function of $A$ is defined by

$$\text{dist}(x, A) := \inf \{\|x - a\| : a \in A\} \quad \text{for each} \quad x \in X,$$

while the excess from the set $A$ to the set $C \subseteq X$ is defined by

$$e(C, A) := \sup \{\text{dist}(x, A) : x \in C\}.$$

All the norms are denoted by $\| \cdot \|$ and the space of linear operators from $X$ to $Y$ is denoted by $L(X, Y)$.

Now, we recall a few definitions, some results and then state the Banach fixed point theorem. We begin with the definition of the first-order divided difference operators. The following definition, is given by Argyros [see, Argyros (2007), introduces the notion of divided differences of nonlinear operators.

**Definition 2.1** An operator belonging to the space $L(X, Y)$ is called the first order divided difference of the operator $g:X \rightarrow Y$ on the points $x$ and $y$ in $X (x \neq y)$ if the following properties hold:

(a) $[x, y; g](y - x) = g(y) - g(x)$ for $x \neq y$;

(b) if $g$ is Fréchet differentiable at $x \in X$, then $[x, y; g] = \nabla g(x)$.

Recall from Rashid et al. (2013) the notions of pseudo-Lipschitz and Lipschitz-like set-valued mappings. These notions were introduced by Aubin [see, Aubin (1984), Aubin and Frankowska (1990), for more details] and have been studied extensively.

**Definition 2.2** Let $\Gamma: Y \rightrightarrows 2^X$ be a set-valued mapping and let $(\bar{y}, \bar{x}) \in \text{gph} \Gamma$. Let $r_x > 0$, $r_y > 0$ and $M > 0$. Then $\Gamma$ is said to be

(a) Lipschitz-like on $\mathcal{B}_{r_y}(\bar{y})$ relative to $\mathcal{B}_{r_x}(\bar{x})$ with constant $M$ if the following inequality holds:

$$e(\Gamma(y_1) \cap \mathcal{B}_{r_y}(\bar{y}), \Gamma(y_2)) \leq M\|y_1 - y_2\| \quad \text{for any} \quad y_1, y_2 \in \mathcal{B}_{r_y}(\bar{y})$$

(b) pseudo-Lipschitz around $(\bar{y}, \bar{x})$ if there exist constants $r'_x > 0$, $r'_y > 0$ and $M' > 0$ such that $\Gamma$ is Lipschitz-like on $\mathcal{B}_{r'_y}(\bar{y})$ relative to $\mathcal{B}_{r'_x}(\bar{x})$ with constant $M'$.

**Remark 2.1** $\Gamma$ is Lipschitz-like on $\mathcal{B}_{r_y}(\bar{y})$ relative to $\mathcal{B}_{r_x}(\bar{x})$ with constant $M$ is equivalent to the following statement: if for every $y_1, y_2 \in \mathcal{B}_{r_y}(\bar{y})$ and for every $x_1 \in \Gamma(y_1) \cap \mathcal{B}_{r_x}(\bar{x})$, there exists $x_2 \in \Gamma(y_2)$ such that

$$\|x_1 - x_2\| \leq M\|y_1 - y_2\|.$$

The following lemma is useful and it has been taken from [Rashid et al., 2013, Lemma 2.1].

**Lemma 2.1** Let $\Gamma: Y \rightrightarrows 2^X$ be a set-valued mapping and let $(\bar{y}, \bar{x}) \in \text{gph} \Gamma$. Assume that $\Gamma$ is Lipschitz-like on $\mathcal{B}_{r_y}(\bar{y})$ relative to $\mathcal{B}_{r_x}(\bar{x})$ with constant $M$. Then

$$\text{dist}(x, \Gamma(y)) \leq M\text{dist}(y, \Gamma^{-1}(x))$$

holds for every $x \in \mathcal{B}_{r_x}(\bar{x})$ and $y \in \mathcal{B}_{r_y}(\bar{y})$ satisfying $\text{dist}(y, \Gamma^{-1}(x)) \leq \frac{r_y}{3}$. 
We end this section with the following lemma. This lemma is a fixed-point statement which has been proved by Dontchev and Hager (1994) and employing the standard iterative concept for contracting mapping. This lemma is used to prove the existence of the sequence generated by Algorithm 1.1.

**Lemma 2.2** Let $\Phi \colon X \rightrightarrows 2^X$ be a set-valued mapping. Let $\eta_0 \in X$, $r > 0$ and $0 < \lambda < 1$ be such that
\[
\text{dist}(\eta_0, \Phi(\eta_0)) < r(1 - \lambda) \tag{2.1}
\]
and
\[
\eta(\Phi(x_1) \cap \Pi_r(\eta_0), \Phi(x_2)) \leq \lambda \| x_1 - x_2 \| \quad \text{for any } x_1, x_2 \in \Pi_r(\eta_0). \tag{2.2}
\]
Then $\Phi$ has a fixed point in $\Pi_r(\eta_0)$, that is, there exists $x \in \Pi_r(\eta_0)$ such that $x \in \Phi(x)$. If $\Phi$ is additionally single-valued, then the fixed point of $\Phi$ in $\Pi_r(\eta_0)$ is unique.

The previous lemma is a generalization of a fixed-point theorem which has been given by Ioffe and Tikhomirov (1979), where in assertion (b) the excess $e$ is replaced by Hausdorff distance.

### 3. Convergence analysis of extended Newton-type method

Throughout this section, we suppose that $f : \Omega \subseteq X \rightarrow Y$ is a Fréchet differentiable function on a neighborhood $\Omega$ of $x$ with its derivative denoted by $\nabla f$, $g : \Omega \subseteq X \rightarrow Y$ is linear and differentiable at $x$, and let $F : X \rightrightarrows Y$ be set-valued mapping with closed graph. We prove the existence and convergence of the sequences generated by extended Newton-type method, defined by the Algorithm 1.1, on a neighborhood $\Omega$ of a point $x$.

Let $x \in X$ and define the mapping $Q_x$ by
\[
Q_x(\cdot) = f(x) + g(\cdot) + \nabla f(x)(\cdot - x) + F(\cdot).
\]
Then
\[
D(x) = \left\{ d \in X : 0 \in Q_x(x + d) \right\}, \text{ since } g(x + d) = g(x) + [2d + x, x; g]d.
\]
We remark that
\[
g(x) + [2d + x, x; g]d = g(x) - \frac{1}{2}[2d + x, x; g(x + (2d + x))]
\]
\[
= g(x) - \frac{1}{2}(g(x) - g(2d + x)) = \frac{1}{2}g(x) + \frac{1}{2}g(2d + x)
\]
\[
= \frac{1}{2}(g(2x + 2d)) = g(x + d).
\]
Furthermore, we have the following equivalence
\[
z \in Q_x^{-1}(y) \iff y \in f(x) + g(z) + \nabla f(x)(z - x) + F(z), \text{ for any } z \in X \text{ and } y \in Y. \tag{3.1}
\]
In particular,
\[
\tilde{x} \in Q_x^{-1}(\tilde{y}) \quad \text{for each } (\tilde{x}, \tilde{y}) \in \text{gph}(f + g + F). \tag{3.2}
\]
Let $(\tilde{x}, \tilde{y})$ be in $\text{gph}(f + g + F)$ and let $r_x > 0$, $r_y > 0$. Throughout the whole paper, we assume that $B_{r_x}(x) \subseteq \Omega \cap \text{dom} F$, the function $g$ is Fréchet differentiable at $x$ and admits a first-order divided difference satisfying the following condition:
there exist $\nu > 0$ such that for all $x,y,u$ and $v \in B_{r_x}(\bar{x})$ ($x \neq y, u \neq v$),

$$\|x,y;g\| - |u,v;g| \leq \nu(\|x - u\| + \|y - v\|)$$

and the mapping $Q^{-1}_x(\cdot)$ is Lipschitz-like on $B_{r_x}(\bar{x})$ relative to $B_{r_x}(\bar{x})$ with constant $M$, that is,

$$e(Q^{-1}_x(y_1) \cap B_{r_x}(\bar{x}), Q^{-1}_x(y_2)) \leq M\|y_1 - y_2\| \quad \text{for any } y_1, y_2 \in B_{r_x}(\bar{y}).$$

(3.3)

Let $\varepsilon > 0$ and write

$$r := \min \left\{ r_x - 2\varepsilon r_x, \frac{r_x(1 - M\varepsilon)}{4M} \right\}. \quad (3.4)$$

Then

$$\tilde{r} < 0 \iff \varepsilon < \min \left\{ \frac{r_x}{2r_x}, \frac{1}{M} \right\}. \quad (3.5)$$

The following lemma plays a crucial role for convergence analysis of the extended Newton-type method. The proof is a refinement of the one for (Rashid et al., 2013, Lemma 3.1).

**Lemma 3.1** Suppose that $Q^{-1}_x(\cdot)$ is Lipschitz-like on $B_{r_x}(\bar{y})$ relative to $B_{r_x}(\bar{x})$ with constant $M$ and that

$$\sup_{y \in B_{r_x}(\bar{x})} \|\nabla f(x) - \nabla f(\bar{x})\| \leq \varepsilon < \min \left\{ \frac{r_x}{2r_x}, \frac{1}{M} \right\}. \quad (3.6)$$

Let $x \in B_{r_x}(\bar{x})$. Then $Q^{-1}_x(\cdot)$ is Lipschitz-like on $B_{r_x}(\bar{x})$ relative to $B_{r_x}(\bar{x})$ with constant $\frac{M}{1 - M\varepsilon}$, that is,

$$e(Q^{-1}_x(y_1) \cap B_{r_x}(\bar{x}), Q^{-1}_x(y_2)) \leq \frac{M}{1 - M\varepsilon}\|y_1 - y_2\| \quad \text{for any } y_1, y_2 \in B_{r_x}(\bar{y}).$$

**Proof** Noted that (3.5) and (3.6) imply $\tilde{r} > 0$. Now let

$$y_1, y_2 \in B_{r_x}(\bar{y}) \quad \text{and} \quad x' \in Q^{-1}_x(y_2) \cap B_{r_x}(\bar{x}). \quad (3.7)$$

It suffices to show that there exist $x'' \in Q^{-1}_x(y_2)$ such that

$$\|x' - x''\| \leq \frac{M}{1 - M\varepsilon}\|y_1 - y_2\|.$$

To this end, we shall verify that there exists a sequence $\{x_k\} \subset B_{r_x}(\bar{x})$ such that

$$y_k \in f(x) + g(x_k) + \nabla f(x)(x_{k-1} - x) + \nabla f(\bar{x})(x_k - x_{k-1}) + F(x_k),$$

and

$$\|x_k - x_{k-1}\| \leq M\|y_1 - y_2\|(M\varepsilon)^{k-2} \quad (3.8)$$

hold for each $k = 2, 3, 4, \ldots$. We proceed by mathematical induction on $k$. Write

$$z_i := y_i - f(x) - \nabla f(x)(x_{i-1} - x) + f(\bar{x}) + \nabla f(\bar{x})(x_1 - x_0) \quad \text{for each } i = 1, 2.$$ 

Note by (3.7) that

$$\|x - x'\| \leq \|x - \bar{x}\| + \|\bar{x} - x'\| \leq r_x.$$

It follows, from (3.7) and the relation $r \leq r_x - 2\varepsilon r_x$ by (3.4), that
\[
\|z_i - y\| \leq \|y_i - y\| + \|f(x) - f(x) - \nabla f(x)(x - x)\| + \|\nabla f(x) - \nabla f(x)(x - x')\|
\]
\[
\leq \bar{r} + \epsilon (\|x - x\| + \|x - x'\|)
\]
\[
\leq \bar{r} + \epsilon \left( \frac{r_2}{2} + r_2 \right) \leq r_2.
\]

That is \( z_i \in B_{\frac{r_2}{2}}(y) \) for each \( i = 1, 2 \). Denote \( x_i := x' \). Then \( x_i \in Q_\delta^{-1}(y_i) \) by (3.7) and it follows from (3.1) that
\[
y_i \in f(x) + g(x_i) + \nabla f(x)(x_i - x) + F(x_i)
\]
which can be rewritten as
\[
y_i + f(x) + \nabla f(x)(x_i - x)) \in f(x) + g(x_i) + \nabla f(x)(x_i - x) + F(x_i) + f(x) + \nabla f(x)(x_i - x).
\]
This, by the definition of \( z_i \), means that
\[
z_i \in f(x) + g(x_i) + \nabla f(x)(x_i - x) + F(x_i).
\]
Hence, \( x_i \in Q_\delta^{-1}(z_i) \) by (3.1). This, together with (3.7), implies that
\[
x_i \in Q_\delta^{-1}(z_i) \cap B_{r_2} (x).
\]
By the assumed Lipschitz-like property of \( Q_\delta^{-1} (\cdot) \) and noting that \( z_1, z_2 \in B_{r_2} (y) \), it follows from (3.3) that there exists \( x_2 \in Q_\delta^{-1} (z_2) \) such that
\[
\|x_2 - x_1\| \leq M\|z_1 - z_2\| = M\|y_1 - y_2\|.
\]
Moreover, by the definition of \( z_2 \) and \( x_1 = x' \), we have
\[
x_2 \in Q_\delta^{-1} (z_2) = Q_\delta^{-1} (y_2 - f(x) - \nabla f(x)(x_1 - x) + f(x) + \nabla f(x)(x_1 - x))
\]
which together with (3.1) implies that
\[
y_2 \in f(x) + g(x_2) + \nabla f(x)(x_1 - x) + \nabla f(x)(x_2 - x_1) + F(x_2).
\]
This shows that (3.8) and (3.9) are true with constructed \( x_1, x_2 \).

We assume that \( x_1, x_2, \ldots, x_n \) are constructed such that (3.8) and (3.9) are true for \( k = 2, 3, \ldots, n \).
We need to construct \( x_{n+1} \) such that (3.8) and (3.9) are also true for \( k = n + 1 \). For this purpose, write
\[
z_i^n := y_i - f(x) - \nabla f(x)(x_{n-1} - x) + f(x) + \nabla f(x)(x_{n-1} - x) \quad \text{for each } i = 0, 1.
\]
Then, by the inductional assumption,
\[
\|z_0^n - z_1^n\| = \|\nabla f(x) - \nabla f(x)(x_n - x_{n-1})\|
\]
\[
\leq \epsilon \|x_n - x_{n-1}\| \leq \|y_1 - y_2\||(M\epsilon)^{3^{-1}}.
\]
(3.10) Since \( \|x_1 - x\| \leq \frac{r_3}{2} \) by (3.7) and \( \|y_1 - y_2\| \leq 2\bar{r} \) by (3.7), it follows from (3.9) that
\[ \|x_n - \bar{x}\| \leq \sum_{k=2}^{n} \|x_k - x_{k-1}\| + \|x_1 - \bar{x}\| \]
\[ \leq 2M\bar{r} \sum_{k=2}^{n} (Me)^{k-2} + \frac{r_x}{2} \]
\[ \leq \frac{2M\bar{r}}{1-Me} + \frac{r_x}{2}. \]
By (3.4), we have \( \bar{r} \leq \frac{r_x(1-Me)}{4M} \) and so
\[ \|x_n - \bar{x}\| \leq r_x. \] (3.11)
Consequently,
\[ \|x_n - x\| \leq \|x_n - \bar{x}\| + \|\bar{x} - x\| \leq \frac{3}{2} r_x. \] (3.12)
Furthermore, using (3.7) and (3.12), one has that, for each \( i = 0,1, \)
\[ \|z^0_i - \bar{y}\| \]
\[ \leq \|y_2 - \bar{y}\| + \|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})\| + \|\nabla f(x) - \nabla f(\bar{x})\| \] \[ \leq \bar{r} + \varepsilon (\|x - \bar{x}\| + \|x - x_{n+i-1}\|) \leq \bar{r} + \varepsilon \left( \frac{r_x}{2} + \frac{3r_x}{2} \right) \]
\[ = \bar{r} + 2\varepsilon r_x. \]
It follows from the definition of \( \bar{r} \) in (3.4) that \( z^0_i \in B_{r_x}(\bar{y}) \) for each \( i = 0,1. \) Since assumption (3.8) holds for \( k = n, \) we have
\[ y_2 \in f(x) + g(x_n) + \nabla f(x)(x_{n-1} - x) + \nabla f(\bar{x})(x_n - x_{n-1}) + F(x_n) \]
which can be written as
\[ y_2 + f(\bar{x}) + \nabla f(\bar{x})(x_n - x_{n-1} - \bar{x}) \in f(x) + \nabla f(x)(x_{n-1} - x) + f(\bar{x}) + g(x_n) \]
\[ + \nabla f(\bar{x})(x_n - x_{n-1}) + F(x_n) + \nabla f(\bar{x})(x_n - x_{n-1} - \bar{x}); \]
i.e. \( z^n_0 \in f(\bar{x}) + g(x_n) + \nabla f(\bar{x})(x_n - \bar{x}) + F(x_n) \) by the definition of \( z^n_0. \) This together with (3.1) and (3.11) yields that
\[ x_n \in Q^{-1}_\delta(z^n_0) \cap B_{r_x}(\bar{x}). \]
Using (3.3) again, there exists an element \( x_{n+1} \in Q^{-1}_\delta(z^n_1) \) such that
\[ \|x_{n+1} - x_n\| \leq M\|z^n_0 - z^n_1\| \leq M\|y_2 - y_3\|(Me)^{n-1}, \] (3.13)
where the last inequality holds by (3.10). By the definition of \( z^n_1, \) we have
\[ x_{n+1} \in Q^{-1}_\delta(z^n_1) = Q^{-1}_\delta(y_2 - f(x) - \nabla f(x)(x_n - x) + f(\bar{x}) + \nabla f(\bar{x})(x_n - \bar{x})); \]
which together with (3.1) implies
\[ y_2 \in f(x) + g(x_{n+1}) + \nabla f(x)(x_n - x) + \nabla f(\bar{x})(x_{n+1} - x_n) + F(x_{n+1}). \]
This together with (3.13) completes the induction step and ensure the existence of a sequence \( \{x_n\} \) satisfying (3.8) and (3.9).
Since $M_\varepsilon < 1$, we see from (3.9) that $\{x_k\}$ is a Cauchy sequence and hence it is convergent, to say $x''$, that is, $x'':=\lim_{k\to\infty}x_k$. Note that $F$ has closed graph. Then, taking limit in (3.8), we get $y_2 \in f(x)+g(x'')+\nabla f(x)(x''-x)+F(x'')$, that is, $x'' \in Q_{x''}^{-1}(y_2)=[f(x)+g(\cdot)+\nabla f(x)(\cdot-x)+F(\cdot)]^{-1}(y_2)$.

Therefore, we obtain

\[
\|x'-x''\| \leq \limsup_{n \to \infty} \sum_{k=2}^{n} \|x_k-x_{k-1}\| \\
\leq \limsup_{n \to \infty} \sum_{k=2}^{n} (M_\varepsilon)^{k-2}M\|y_1-y_2\| \\
\leq \frac{M}{1-M_\varepsilon}\|y_1-y_2\|.
\]

That is,

\[
e(Q_{x''}^{-1}(y_1) \cap B_\varepsilon(\bar{x}), Q_{x''}^{-1}(y_2)) \leq \frac{M}{1-M_\varepsilon}\|y_1-y_2\|.
\]

This completes the proof of the Lemma 3.1. \hfill \Box

For our convenience, we define the mapping $Z_x:X \to Y$, for each $x \in X$, by

$Z_x(\cdot):=f(\bar{x})+g(\cdot)+\nabla f(\bar{x})(\cdot-\bar{x})-f(x)-g(x)-(\nabla f(x)+[2(\cdot)-x, x; g])(\cdot-x)$

and the set-valued mapping $\Phi_x:X \rightrightarrows 2^X$ by

$\Phi_x(\cdot):=Q_{x''}^{-1}[Z_x(\cdot)].$ (3.14)

Then for any $x', x'' \in X$, we have

\[
\|Z_x(x')-Z_x(x'')\| = \|g(x')-g(x'')-[2x'-x, x; g](x'-x)+[2x''-x, x; g](x''-x) \\
+ (\nabla f(x)-\nabla f(x'))(x'-x'')\| \\
\leq \|g(x')-g(x'')-2x'-x, x; g(x'-x)\| + \|2x''-x, x; g\| \\
- (2x'-x, x; g)((x'-x)) + \|\nabla f(x)-\nabla f(x')\||x'-x''|| \\
\leq (\|x', x; g\| - \|2x'-x, x; g\| + \|\nabla f(x)-\nabla f(x')\|)|x'-x''| \\
+ \|2x''-x, x; g\| - \|2x'-x, x; g\||x'-x||.
\]

(3.15)

Our first main theorem is as follows, which provides some sufficient conditions ensuring the convergence of the extended Newton-type method with initial point $x_0$.

**THEOREM 3.1** Suppose that $\eta > 1$ and that $Q_{x''}^{-1}(\cdot)$ is Lipschitz-like on $\overline{B}_\varepsilon(y)$ relative to $\overline{B}_\varepsilon(x)$ with constant $M$. Let

\[
\varepsilon \geq \sup_{x, x' \in \overline{B}_\varepsilon(x)} \|\nabla f(x)-\nabla f(x')\| \tag{3.16}
\]

and let $r$ be defined in (3.4). Let $\nu > 0$ and $\delta > 0$ be such that

(a) $\delta \leq \min \left\{ \frac{r}{4}, \frac{r}{3(\varepsilon + 3\nu)}, \frac{r}{80M_\varepsilon}, \frac{r_\nu}{17(\varepsilon + 3\nu)} \right\},$

(b) $6\eta M_\varepsilon(\varepsilon + 3\nu) \leq 1 - M_\varepsilon,$

(c) $\|\hat{y}\| < (\varepsilon + 3\nu)\delta.$
Suppose that
\[
\lim_{x \to x^*} \text{dist}(\bar{y}, f(x) + g(x) + F(x)) = 0. \tag{3.17}
\]

Then there exists some \( \tilde{\delta} > 0 \) such that any sequence \( \{x_n\} \) generated by Algorithm 1.1 with initial point \( x_0 \) converges to a solution \( x^* \) of (1.1), that is, \( x^* \) satisfies \( 0 \in f(x^*) + g(x^*) + F(x^*) \).

**Proof** Letting that \( q = \frac{\eta M(\varepsilon + 3\nu)}{1 - M\varepsilon} \). Then by the relation \( 6\eta M(\varepsilon + 3\nu) \leq 1 - M\varepsilon \) from assumption (b), we obtain
\[
q = \frac{\eta M(\varepsilon + 3\nu)}{1 - M\varepsilon} \leq \frac{1}{6}.
\]

Take \( 0 < \delta \leq \tilde{\delta} \) such that
\[
\text{dist}(0, f(x_0) + g(x_0) + F(x_0)) \leq (\varepsilon + 3\nu)\delta \quad \text{for each } x_0 \in \mathbb{B}_2(\tilde{x}).
\tag{3.18}
\]

(Noting that such \( \tilde{\delta} \) exists by (3.17) and assumption (c)). Let \( x_0 \in \mathbb{B}_2(\tilde{x}) \). We will proceed by mathematical induction to show that Algorithm 1.1 generates at least one sequence and any sequence \( \{x_n\} \) generated by Algorithm 1.1 satisfies the following assertions:

\[
\|x_n - \bar{x}\| \leq 2\delta
\tag{3.19}
\]

and
\[
\|x_{n+1} - x_n\| \leq q^{n+1}\delta
\tag{3.20}
\]

hold for each \( n = 0, 1, 2, \ldots \). For this purpose, we define
\[
r_x = \frac{5}{2} \left( M(\varepsilon + 3\nu)\|x - \bar{x}\| + \|x - \bar{x}\| + M\|y\| \right) \quad \text{for each } x \in X.
\tag{3.21}
\]

Then, thanks to the fact that \( 6\eta M(\varepsilon + 3\nu) \leq 1 - M\varepsilon \) by assumption (b) and \( \|y\| < (\varepsilon + 3\nu)\delta \) by assumption (c). Since \( \eta > 1 \), (3.21) yields that
\[
r_x < 5M(\varepsilon + 6\nu\delta)\delta + M(\varepsilon + 3\nu)\delta
\leq 6M\varepsilon \delta + 33\nu\delta^2 < 6M\varepsilon \delta + 33\nu\delta
< 11M\varepsilon \delta + 33\nu\delta = 11M(\varepsilon + 3\nu)\delta
< \frac{11}{6\eta} \leq 2\delta \quad \text{for each } x \in \mathbb{B}_2(\tilde{x}).
\tag{3.22}
\]

Note that (3.19) is trivial for \( n = 0 \). To show (3.20) holds for \( n = 0 \), first we need to show that \( x_1 \) exists. To complete this, we have to prove that \( D(x_0) \neq \emptyset \) by applying Lemma 2.2 to the map \( \Phi_{x_0} \) with \( \eta_0 = \bar{x} \). Let us check that both assertions (2.1) and (2.2) of Lemma 2.2 hold with \( r := r_{x_0} \) and \( \lambda := \frac{3}{5} \). Noting that \( \bar{x} \in Q_x(\tilde{y}) \cap \mathbb{B}_{2\delta}(\tilde{x}) \) by (3.2) and according to the definition of the excess \( e \) and the mapping \( \Phi_{x_0} \) in (3.14), we obtain
\[
\text{dist}(\bar{x}, \Phi_{x_1}(\tilde{x})) \leq e(Q_{x_1}(\tilde{y}) \cap \mathbb{B}_{2\delta}(\tilde{x}), \Phi_{x_1}(\tilde{x})) \leq e(Q_{x_1}(\tilde{y}) \cap \mathbb{B}_{2\delta}(\tilde{x}), Q_{x_1}(\tilde{x}))
\leq e(Q_{x_1}(\tilde{y}) \cap \mathbb{B}_{r_x}(\tilde{x}), Q_{x_1}(\tilde{x}))
\tag{3.23}
\]

(noting that \( \mathbb{B}_{2\delta}(\tilde{x}) \subseteq \mathbb{B}_{r_x} \)). By the choice of \( \varepsilon \), we have
\[
\|Z_{x_0}(x) - \bar{y}\| = \|f(\bar{x}) + g(x) + \nabla f(\bar{x})(x - \bar{x}) - f(x_0) - g(x_0) \\
- \nabla f(x_0) + [2x - x_0, x_0; g](x - x_0) - \bar{y}\| \\
\leq \|f(\bar{x}) - f(x_0) - \nabla f(x_0)(x - x_0)\| + \|\nabla f(x_0) - \nabla f(\bar{x})(\bar{x} - x)\| \\
+ \|g(x) - g(x_0) - [2x - x_0, x_0; g](x - x_0)\| + ||\bar{y}\| \\
\leq \varepsilon (\|\bar{x} - x_0\| + ||\bar{x} - x\|) + \|x_0, x_0; g\| - [2x - x_0, x_0; g]\|x - x_0\| \\
+ ||\bar{y}\| \\
\leq \varepsilon (\|\bar{x} - x_0\| + ||\bar{x} - x\|) + \nu(\|x_0 - x\| + ||x - x_0||) + \|x - x_0\| \\
+ ||\bar{y}\|. \\
(3.24)
\]

Note that \(\|x_0 - \bar{x}\| \leq \delta \leq 17(\varepsilon + 3\nu)\delta \leq r_y\) by assumption (a) and \(\|\bar{y}\| \leq (\varepsilon + 3\nu)\delta\) by assumption (c), it follows from (3.24) that, for each \(x \in \mathbb{B}_{r_y}(\bar{x}),\)

\[
\|Z_{x_0}(x) - \bar{y}\| \leq 3\varepsilon\delta + 27\nu\delta^2 + (\varepsilon + 3\nu)\delta < 3\varepsilon\delta + 27\nu\delta + (\varepsilon + 3\nu)\delta \\
\leq 9(\varepsilon + 3\nu)\delta + (\varepsilon + 3\nu)\delta = 10(\varepsilon + 3\nu)\delta \\
\leq r_y. \\
(3.25)
\]

In particular, letting \(x = \bar{x}\) in (3.24). Then we have that

\[
\|Z_{x_0}(\bar{x}) - \bar{y}\| \leq \varepsilon (\|\bar{x} - x_0\| + \nu(2\|x_0 - \bar{x}\| + ||\bar{x} - x_0||) ||\bar{x} - x_0|| + ||\bar{y}\| \\
= (\varepsilon + 3\nu||\bar{x} - x_0||) ||\bar{x} - x_0|| + ||\bar{y}\| \\
\leq (\varepsilon + 3\nu)\delta + ||\bar{y}\| < (\varepsilon + 3\nu)\delta + ||\bar{y}\| \\
\leq 2(\varepsilon + 3\nu)\delta \leq r_y. \\
(3.26)
\]

and hence \(Z_{x_0}(x) \in \mathbb{B}_{r_y}(\bar{y}).\)

Hence, by (3.21), (3.23), (3.26) and the assumed Lipschitz-like property, we have

\[
dist(\bar{x}, \Phi_{z_0}(\bar{x})) \leq M||y - Z_{x_0}(\bar{x})|| \\
\leq M(\varepsilon + 3\nu||\bar{x} - x_0||) ||\bar{x} - x_0|| + M||\bar{y}\| \\
= (1 - \frac{3}{5})r_y = (1 - \lambda) r;
\]

that is, the assertion (2.1) of Lemma 2.2 is satisfied.

Now, we show that the assertion (2.2) of Lemma 2.2 holds. To end this, let \(x', x'' \in \mathbb{B}_{r_y}(\bar{x}).\) Then we have that \(x', x'' \in \mathbb{B}_{r_y}(\bar{x}) \subseteq \mathbb{B}_{r_y}(\bar{x}) \subseteq \mathbb{B}_{r_y}(\bar{x})\) by (3.22) and assumption (a), and \(Z_{x_0}(x'), Z_{x_0}(x'') \in \mathbb{B}_{r_y}(\bar{y})\) by (3.25). This together with the assumed Lipschitz-like property implies that

\[
e(\Phi_{x_0}(x') \cap \mathbb{B}_{r_y}(\bar{x}), \Phi_{x_0}(x'')) \leq e(\Phi_{x_0}(x') \cap \mathbb{B}_{r_y}(\bar{x}), \Phi_{x_0}(x'')) \\
= e(Q_{x_0}^{-1}[Z_{x_0}(x') \cap \mathbb{B}_{r_y}(\bar{x}), Q_{x_0}^{-1}[Z_{x_0}(x'')]]) \\
\leq M\|Z_{x_0}(x') - Z_{x_0}(x'')\|. \\
(3.27)
\]

Using (3.15) and the choice of \(x_0,\) we have

\[
\|Z_{x_0}(x') - Z_{x_0}(x'')\| \leq \left( \|x'\|, x_0; g\| - \|2x'' - x_0, x_0; g\| + ||f(x) - f(x_0)|| \right) ||x' - x''|| \\
+ \|\|2x'' - x_0, x_0; g\| - \|2x' - x_0, x_0; g\|\| ||x' - x_0|| \\\n\leq \left( \nu\left( \|x_0 - x''\| + ||x' - x_0|| \right) + \varepsilon \right) ||x' - x''|| \\
+ 2\nu ||x'' - x_0|| ||x' - x_0|| \\\n\leq (\varepsilon + 12\nu\delta) ||x' - x''|| \leq (\varepsilon + 16\nu\delta) ||x' - x''||.
\]
It follows, from \( \delta \leq \frac{3 - 5M \varepsilon}{80M \varepsilon} \) as in assumption (a) together with (3.27) that
\[
\bar{d} \leq \frac{3}{2} \left\| x' - x'' \right\| \leq \frac{3}{2} \left\| x' - x'' \right\| = \lambda \left\| x' - x'' \right\|.
\]
This yields that the assertion (2.2) of Lemma 2.2 is satisfied. Since both assertions of Lemma 2.2 are fulfilled, we can say that the Lemma 2.2 is applicable and hence we can conclude that there exists \( \hat{x}_1 \in \mathcal{B}_{r}(\hat{x}) \) satisfying \( \hat{x}_1 \in \Phi_{\xi}(\hat{x}_1) \). This means that \( 0 \in f(x_0) +\bar{g}(x_0) + [\nabla f(x_0) + [2\hat{x}_1 - x_0, x_0; g)](\hat{x}_1 - x_0) + F(\hat{x}_1) \), i.e. \( D(x_0) \neq \emptyset \). Since \( \eta > 1 \) and \( D(x_0) \neq \emptyset \), we can choose \( d_0 \in D(x_0) \) such that
\[
\|d_0\| \leq \eta \, \text{dist}(0,D(x_0)).
\]
By Algorithm 1.1, \( x_i := x_0 + d_0 \) is defined. Furthermore, by the definition of \( D(x_0) \), we can write
\[
D(x_0) = \left\{ d_0 \in X : 0 \in f(x_0) +\bar{g}(x_0) + [\nabla f(x_0) + [2d_0 + x_0, x_0; g]]d_0 + F(x_0 + d_0) \right\}
\]
and so
\[
\text{dist}(0,D(x_0)) = \text{dist}(x_0, Q^{-1}_{\xi}(0)). \tag{3.28}
\]
Now, we show that (3.20) holds also for \( n = 0 \). Note by (3.16) that
\[
\varepsilon \geq \sup_{x \in \mathcal{B}_{\frac{1}{2}}(x_0)} \| \nabla f(x) - \nabla f(\hat{x}) \|
\]
and note also that \( \hat{f} > 0 \) by assumption (a). Therefore, (3.5) satisfies (3.6). Hence, by the assumed Lipschitz-like property of \( Q^{-1}_{\xi}(\cdot) \), it follows from Lemma 3.1 that the mapping \( Q^{-1}_{\xi}(\cdot) \) is Lipschitz-like on \( \mathcal{B}_{\frac{1}{2}}(y) \) relative to \( \mathcal{B}_{\frac{1}{2}}(x) \) with constant \( M \frac{1}{1 - M \varepsilon} \) for each \( x \in \mathcal{B}_{\frac{1}{2}}(x) \). In particular, \( Q^{-1}_{\xi}(\cdot) \) is Lipschitz-like on \( \mathcal{B}_{\frac{1}{2}}(y) \) relative to \( \mathcal{B}_{\frac{1}{2}}(x) \) with constant \( M \frac{1}{1 - M \varepsilon} \) as \( x_0 \in \mathcal{B}_{\frac{1}{2}}(x) \). This means that \( \bar{f} > 0 \) by assumption (a) and by the choice of \( \delta \). Furthermore, assumptions (a) and (c) imply that
\[
\|\hat{y}\| < (\varepsilon + 3\nu)\delta \leq \frac{\hat{f}}{3}, \tag{3.29}
\]
and hence (3.18) implies that
\[
\text{dist}(0,Q^{-1}_{\xi}(x_0)) = \text{dist}(0,f(x_0) +\bar{g}(x_0) + F(x_0)) \leq (\varepsilon + 3\nu)\delta \leq \frac{\hat{f}}{3} \tag{3.30}
\]
It is noted earlier that \( x_0 \in \mathcal{B}_{\frac{1}{2}}(x) \) and \( 0 \in \mathcal{B}_{\frac{1}{2}}(\hat{y}) \) by (3.29). Thus, we can apply Lemma 2.1 and utilizing it, we get
\[
\text{dist}(x_0, Q^{-1}_{\xi}(0)) \leq M \frac{1}{1 - M \varepsilon} \, \text{dist}(0,Q^{-1}_{\xi}(x_0)).
\]
This together with (3.28) gives that
\[
\text{dist}(0,D(x_0)) = \text{dist}(x_0, Q^{-1}_{\xi}(0)) \leq M \frac{1}{1 - M \varepsilon} \, \text{dist}(0,Q^{-1}_{\xi}(x_0)). \tag{3.31}
\]
According to Algorithm 1.1 and using (3.10) and (3.31), we have

\[ \|d_0\| \leq \eta \text{dist}(0,D(x_0)) \leq \frac{\eta M}{1-M \varepsilon} \text{dist}(0,\mathcal{Q}_{x}(x_0)) \]

\[ \leq \frac{\eta M(3+3\nu)}{1-M \varepsilon} = q \delta. \]

This implies that

\[ \|x_1-x_0\| = \|d_0\| \leq q \delta \]

and therefore, (3.20) is hold for \( n = 0 \).

We assume that \( x_1,x_2,\ldots,x_k \) are constructed so that (3.19) and (3.20) are hold for \( n = 0,1,2,\ldots,k-1 \). We will show that there exists \( x_{k+1} \) such that (3.19) and (3.20) are also hold for \( n = k \). Since (3.19) and (3.20) are true for each \( n \leq k-1 \), we have the following inequality

\[ \|x_k-x\| \leq \sum_{i=0}^{k-1} \|d_i\| + \|x_0-x\| \leq \delta \sum_{i=0}^{k-1} q^{i+1} + \delta \leq \frac{\delta q}{1-q} + \delta \leq 2 \delta. \]

This shows that (3.19) holds for \( n = k \). Now with almost the same argument as we did for the case when \( n = 0 \), we can show that (3.20) hold for \( n = k \). The proof is complete.

In particular, in the case when \( x \) is a solution of (1.2), i.e. \( \bar{y} = 0 \), Theorem 3.1 is reduced to the following corollary, which gives the local convergent result for the extended Newton-type method.

**Corollary 3.1** Suppose that \( \eta > 1 \) and \( x \) satisfies \( 0 \in f(x) + g(x) + F(x) \). Let \( Q_{x}^{-1}(\cdot) \) be pseudo-Lipschitz around \( (0,\bar{x}) \). Let \( \bar{r} > 0 \), \( \nu > 0 \) and suppose that \( \nabla f \) is continuous on \( \mathbb{B}_r(\bar{x}) \) and that

\[ \lim_{x \to \bar{x}} \text{dist}(0,f(x)+g(x)+F(x)) = 0. \]

Then there exists some \( \delta \) such that any sequence \( \{x_n\} \) generated by Algorithm 1.1 with initial point in \( \mathbb{B}_r(\bar{x}) \) converges to a solution \( x' \) of (1.1), that is, \( x' \) satisfies that \( 0 \in f(x') + g(x') + F(x') \).

**Proof** Let \( Q_{x}^{-1}(\cdot) \) is pseudo-Lipschitz around \( (0,\bar{x}) \). Then there exist constants \( r_{\bar{x}}, \bar{r}_{\bar{x}} \) and \( M \) satisfy the following condition:

\[ e(Q_{x}^{-1}(y_1) \cap \mathbb{B}_{\bar{x}}(\bar{x}),Q_{x}^{-1}(y_2)) \leq M\|y_1-y_2\|, \text{ for every } y_1,y_2 \in \mathbb{B}_{\bar{x}}(0). \]

(3.32)

Thus, by the definition of Lipschitz-like property, we can say that \( Q_{x}^{-1}(\cdot) \) is Lipschitz-like on \( \mathbb{B}_r(0) \) relative to \( \mathbb{B}_{\bar{r}}(\bar{x}) \) with constant \( M \) which satisfy (3.32). Then, for each \( 0 < \bar{r} \leq \bar{r}_{\bar{x}} \), one has that

\[ e(Q_{x}^{-1}(y_1) \cap \mathbb{B}_{\bar{x}}(\bar{x}),Q_{x}^{-1}(y_2)) \leq M\|y_1-y_2\|, \text{ for every } y_1,y_2 \in \mathbb{B}_{\bar{x}}(0), \]

that is, \( Q_{x}^{-1}(\cdot) \) is Lipschitz-like on \( \mathbb{B}_{\bar{r}}(0) \) relative to \( \mathbb{B}_{\bar{r}}(\bar{x}) \) with constant \( M \). Let \( \varepsilon \in (0,1) \) be such that \( M((6\eta+1)s+3\nu) \leq 1 \). By the continuity of \( \nabla f \), we can choose \( r_{\bar{x}} \in (0,\bar{r}_{\bar{x}}) \) such that \( r_{\bar{x}} \leq \bar{r} \), \( r_{\bar{x}} - 2\varepsilon r_{\bar{x}} > 0 \) and

\[ \varepsilon \geq \sup_{x,\bar{x} \in \mathbb{B}_{r}(\bar{x})} \|\nabla f(x) - \nabla f(x')\|. \]

Then

\[ \bar{r} = \min\left\{ r_{0} - 2\varepsilon r_{\bar{x}}, \frac{r_{\bar{x}}(1-M \varepsilon)}{4M} \right\} < 0 \]

and
Rashid, Cogent Mathematics (2014), 1: 980600
http://dx.doi.org/10.1080/23311835.2014.980600

Thus it is routine to check that inequalities (a)–(c) of Theorem 3.1 are satisfied. Therefore, Theorem 3.1 is applicable to complete the proof. □

In the following theorem, we show that if \( \nabla f \) is Lipschitz continuous around \( \bar{x} \), then the sequence generated by Algorithm 1.1 converges quadratically.

**Theorem 3.2** Let \( \eta > 1 \) and suppose that \( Q^{-1}_\bar{x} (\cdot) \) is Lipschitz-like on \( \mathbb{B}_r (\bar{y}) \) relative to \( \mathbb{B}_r (\bar{x}) \) with constant \( M \) and that \( \nabla f \) is Lipschitz continuous on \( \mathbb{B}_r (\bar{x}) \) with Lipschitz constant \( L \). Let

\[
\tilde{r} := \min \left\{ \frac{r_\alpha}{4}, \frac{\bar{r}}{3(\varepsilon + 3\nu)}, \frac{r_0}{17(\varepsilon + 3\nu)}, \frac{3 - 5M\varepsilon}{80M\nu} \right\}.
\]  

Let \( \nu > 0, \delta > 0 \) be such that

(a) \( \delta \leq \min \left\{ \frac{r_\alpha}{4}, 6\bar{r}, 1, \frac{4\bar{r}}{33(L + 6\nu)} \right\}, \)

(b) \( (M + 1)(L + 6\nu)(\eta\delta + 2\bar{r}) \leq 2, \)

(c) \( ||\bar{y}|| < \frac{(L + 6\nu)\delta^2}{4}. \)

Suppose that

\[
\lim_{x \to \bar{x}} \text{dist}(\bar{y}, f(x) + g(x) + F(x)) = 0. \tag{3.34}
\]

Then there exist some \( \tilde{\delta} > 0 \) such that any sequence \( \{x_n\} \) generated by Algorithm 1.1 with initial point in \( \mathbb{B}_r (\bar{x}) \) converges quadratically to a solution \( x^* \) of (1.1).

**Proof** Setting

\[
q_\delta := \frac{\eta M(L + 6\nu)\delta}{2(1 - M\bar{r})}, \tag{3.35}
\]

Then, thanks to the assumption (b) for allowing us to write the fact that

\[
M(L + 6\nu)\eta\delta + 2M\bar{r} < (M + 1)(L + 6\nu)(\eta\delta + 2\bar{r}) \leq 2.
\]

It follows from (3.35) that

\[
q_\delta := \frac{\eta M(L + 6\nu)\delta}{2(1 - M\bar{r})} \leq 1. \tag{3.36}
\]

Taking \( 0 < \delta \leq \tilde{\delta} \) such that

\[
\text{dist}(0, f(x_0) + g(x_0) + F(x_0)) \leq \frac{(L + 6\nu)\delta^2}{4} \quad \text{for each } x_0 \in \mathbb{B}_\delta (\bar{x}). \tag{3.37}
\]

It is noting that such \( \tilde{\delta} \) exists by (3.34) and assumption (c). Let \( x_0 \in \mathbb{B}_\delta (\bar{x}) \). To complete the proof of this theorem, we use almost similar argument that we used for completing the proof of Theorem 3.1.
We show that Algorithm 1.1 generates at least one sequence and such sequence \( \{ x_n \} \) generated by Algorithm 1.1 satisfies the following assertions:

\[
\| x_n - \bar{x} \| \leq 2\delta; \quad (3.38)
\]

and

\[
\| d_n \| \leq q \left( \frac{1}{2} \right)^n \delta \quad (3.39)
\]

hold for each \( n = 0, 1, 2, \ldots \). Let

\[
r_x := \frac{3}{2} \left( M(L + 6\nu)\| x - \bar{x} \|^2 + 2M\| y \| \right) \quad \text{for each } x \in X. \quad (3.40)
\]

Owing to the fact \( \delta \leq \frac{r_x}{4} \) in assumption (a) and \( \eta > 1 \), by assumption (b) we can write as follows

\[
9(M + 1)(L + 6\nu)\delta = (M + 1)(L + 6\nu)(\delta + 8\delta) \leq (M + 1)(L + 6\nu)(\eta\delta + 2r_x) \leq 2.
\]

This gives

\[
M(L + 6\nu)\delta \leq \frac{2}{9} \quad \text{and} \quad (L + 6\nu)\delta \leq \frac{2}{9} \quad (3.41)
\]

and hence by \( \delta \leq \frac{1}{9} \delta \) in assumption (a) together with second inequality of (3.41), we get

\[
\| y \| \leq \frac{(L + 6\nu)\delta^2}{4} \leq \frac{2}{9} \cdot \frac{1}{4} \cdot \frac{1}{3} = \frac{r_x}{3}. \quad (3.42)
\]

thanks to assumption (c). Utilizing the first inequality from (3.41) and assumption (c) together, we obtain from (3.40) that

\[
r_x < \frac{3}{2} \left( \frac{4M(L + 6\nu)\delta^2 + M(L + 6\nu)\delta^2}{2} \right)
\]

\[
= \frac{27}{4} M(L + 6\nu)\delta^2 \leq 2\delta \quad \text{for each } x \in \mathbb{B}_{2\delta}(\bar{x}). \quad (3.43)
\]

Note that (3.38) is trivial for \( n = 0 \). In order to show that (3.39) hold for \( n = 0 \), we need to prove

\( D(x_0) \neq \emptyset \). The nonemptiness of \( D(x_0) \) will ensure us to deduce the existence of the point \( x_1 \). To complete this, we will apply Lemma 2.2 to the map \( \Phi_{x_0} \) with \( n_0 = \bar{x} \). Let us check that both assertions (2.1) and (2.2) of Lemma 2.2 hold with \( r_x := r_{x_0} \) and \( \lambda := \frac{2}{3} \). Noting that \( \bar{x} \in Q_{\bar{x}}^{-1}(\bar{y}) \cap \mathbb{B}_{2\delta}(\bar{x}) \) by (3.2) and according to the definition of the excess \( e \) and the mapping \( \Phi_{x_0} \) by (3.14), we obtain

\[
\text{dist}(\bar{x}, \Phi_{x_0}(\bar{x})) \leq e(Q_{\bar{x}}^{-1}(\bar{y}) \cap \mathbb{B}_{n_0}(\bar{x}, \Phi_{x_0}(\bar{x})) \leq e(Q_{\bar{x}}^{-1}(\bar{y}) \cap \mathbb{B}_{2\delta}(\bar{x}, \Phi_{x_0}(\bar{x}))
\]

\[
\leq e(Q_{\bar{x}}^{-1}(\bar{y}) \cap \mathbb{B}_{n_0}(\bar{x}, \Phi_{x_0}(\bar{x})) \cup Z_{n_0}(\bar{x})]. \quad (3.44)
\]

By the assumed Lipschitz continuity of \( \nabla f \) and for each \( x \in \mathbb{B}_{2\delta}(X) \subseteq \mathbb{B}_{\frac{\delta}{2}}(\bar{x}) \), we obtain that
\[\|Z_{x_n}(x) - \tilde{y}\| = \|f(x) + g(x) + \nabla f(\bar{x})(x - \bar{x}) - f(x_0) - g(x_0)\]
\[\quad - (\nabla f(x_0) + [2x - x_0, x_0]g)(x - x_0) - \tilde{y}\|
\leq \|f(x) - f(x_0) - \nabla f(x_0)(\bar{x} - x_0)\| + \|\nabla f(x_0)(\bar{x} - x)\|
\quad + \|g(x) - g(x_0) - [2x - x_0, x_0]g\| |x - x_0| + \|\tilde{y}\|
\leq \frac{L}{2} \|\bar{x} - x_0\|^2 + L \|x_0 - \bar{x}\| \|\bar{x} - x\| + \|x_0 - x_0\| + \|\tilde{y}\|
\leq \frac{L}{2} \|\bar{x} - x_0\|^2 + L \|x_0 - \bar{x}\| \|\bar{x} - x\| + \nu \left(2\|x_0 - x\| + ||x_0 - x_0\|\right)\|x - x_0\| + \|\tilde{y}\|
\leq \frac{L}{2} (\delta^2 + 4\delta^2) + 3\nu (3\delta^2) + \|\tilde{y}\| = \frac{5L\delta^2}{2} + 27\nu\delta^2 + \|\tilde{y}\|
\leq \frac{9}{2} (L + 6\nu)\delta^2 + \|\tilde{y}\|. \tag{3.45}\]

It follows, from the facts $33(L + 6\nu)\delta \leq 4r_0$, $\delta \leq 1$ and $\|\tilde{y}\| < \frac{(L + 6\nu)\delta^2}{4}$, respectively, in assumptions (a) and (c), that

\[\|Z_{x_n}(x) - \tilde{y}\| \leq \frac{9}{2} (L + 6\nu)\delta^2 + \frac{(L + 6\nu)\delta^2}{4} = \frac{19}{4} (L + 6\nu)\delta^2
\leq \frac{19}{4} (L + 6\nu)\delta \leq r_0. \tag{3.46}\]

This shows that $Z_{x_n}(x) \in B_{r_0}(\tilde{y})$. In particular, let $x = x$ in (3.45). Then it is easily shown that

\[Z_{x_n}(x) \in B_{r_0}(\tilde{y}) \quad \text{and} \quad \|Z_{x_n}(x) - \tilde{y}\| \leq \frac{(L + 6\nu)}{2} \|x - x_0\|^2 + \|\tilde{y}\|. \tag{3.47}\]

Using assumed Lipschitz-like property and (3.47) in (3.44), we have

\[\text{dist}(\bar{x}, \Phi_{x_n}(\bar{x})) \leq M\|\tilde{y} - Z_{x_n}(x)\| \leq \frac{M(L + 6\nu)}{2} \|x - x_0\|^2 + M\|\tilde{y}\|
\]
\[= \left(1 - \frac{2}{3}\right) r_{x_0} = (1 - \lambda)x_0\]

that is, the assertion (2.1) of Lemma 2.2 is satisfied.

Now, we show that assertion (2.2) of Lemma 2.2 holds. To end this, let $x', x'' \in B_{r_0}(\bar{x})$. Then we have that $x', x'' \in B_{r_0}(x) \subseteq B_{r_0}(\bar{x}) \subseteq B_{r_0}(x)$ by (3.43) and $Z_{x_n}(x'), Z_{x_n}(x'') \in B_{r_0}(\tilde{y})$ by (3.46). This together with the assumed Lipschitz-like property implies that

\[e(\Phi_{x_n}(x') \cap B_{r_0}(\bar{x}), \Phi_{x_n}(x'')) \leq e(\Phi_{x_n}(x') \cap B_{r_0}(\bar{x}), \Phi_{x_n}(x'))
\]
\[= e(Q_{x_n}^{-1}[Z_{x_n}(x')] \cap B_{r_0}(\bar{x}), Q_{x_n}^{-1}[Z_{x_n}(x'')])
\]
\[\leq M\|Z_{x_n}(x') - Z_{x_n}(x'')\|
\]

By the choice of $x_0$, (3.15) yields that

\[\|Z_{x_n}(x') - Z_{x_n}(x'')\| \leq \left(\|x' - x''\|^2 + 2\|x' - x''\|\|x' - x\| + L \|x - x_0\|\|x' - x''\|
\]
\[+ 2\nu \|x' - x\| \|x' - x_0\| + (L + 12\nu)\|x' - x''\| \leq 2(L + 6\nu)\|x' - x''\|.\]
Combining above inequality and first inequality from (3.41), we obtain that
e(Φ, (x') \cap Br_r(x), Φ, (x'')) ≤ 2M(1 + 6v)\delta ∥x' - x''∥ ≤ 3M(1 + 6v)\delta ∥x' - x''∥
\leq \frac{2}{3} ∥x' - x''∥ = \lambda ∥x' - x''∥.

This means that the assertion (2.2) of Lemma 2.2 is also satisfied. Since both assertions of Lemma 2.2 are fulfilled, we can conclude that Lemma 2.2 is applicable to deduce the
existence of a point \( \hat{x}_1 \in B_{r_0}(\hat{x}) \) such that \( \hat{x}_1 \in Φ_{x_0}(\hat{x}_1) \). This implies that
0 \in f(x_0) + g(x_0) + (\nabla f(x_0) + [2\hat{x}_1 - x_0, x_0; g])(\hat{x}_1 - x_0) + F(\hat{x}_1), \) that is
\( D(x_0) \neq \emptyset \). Since \( \eta > 1 \) and
\( D(x_0) \neq \emptyset \), we can choose \( d_0 \in D(x_0) \) such that
\[ \|d_0\| \leq \eta \text{ dist}(0, D(x_0)). \]

By Algorithm 1.1, \( x_1 = x_0 + d_0 \) is defined. Furthermore, by the definition of \( D(x_0) \), we can write
\[ D(x_0) : = \{ d_0 \in X : 0 \in f(x_0) + g(x_0) + (\nabla f(x_0) + [2d_0 + x_0; x_0; g])d_0 + F(x_0 + d_0) \} \]
\[ = \{ d_0 \in X : 0 \in f(x_0) + g(x_0 + d_0) + \nabla f(x_0)d_0 + F(x_0 + d_0) \} \]
and so
\[ \text{dist}(0, D(x_0)) = \text{dist}(x_0, Q_{x_0}^{-1}(0)). \] (3.48)

Now we are ready to show that (3.39) is hold for \( n = 0 \). Since \( \nabla f \) is Lipschitz continuous on \( B_{\frac{r}{2}}(\hat{x}) \) with
Lipschitz constant \( L \), we have
\[ L_r \geq \sup_{x, x' \in B_{\frac{r}{2}}(\hat{x})} \|\nabla f(x') - \nabla f(x)\| \geq \sup_{x \in B_{\frac{r}{2}}(\hat{x})} \|\nabla f(x) - \nabla f(\hat{x})\|. \] (3.49)

Note that \( \bar{r} > 0 \) by assumption (a). Therefore, (3.5) and (3.49) imply that assumption (3.6) is satisfied
with \( \epsilon := L_r \). Since \( Q_{x_0}^{-1}(-) \) is Lipschitz-like on \( B_{\frac{r}{2}}(\hat{y}) \) relative to \( B_{\frac{r}{2}}(\hat{x}) \), it follows from Lemma 3.1 that
for each \( x \in B_{\frac{r}{2}}(\hat{x}) \) the mapping \( Q_{x_0}^{-1}(-) \) is Lipschitz-like on \( B_{\frac{r}{2}}(\hat{y}) \) relative to \( B_{\frac{r}{2}}(\hat{x}) \) with constant
\( \frac{M}{1 - ML_{x}} \). In particular, \( Q_{x_0}^{-1}(-) \) is Lipschitz-like on \( B_{\frac{r}{2}}(\hat{y}) \) relative to \( B_{\frac{r}{2}}(\hat{x}) \) with constant
\( \frac{M}{1 - ML_{x}} \) as \( x_0 \in B_{\frac{r}{2}}(\hat{x}) \subseteq B_{2\frac{r}{2}}(\hat{x}) \subseteq B_{2\frac{r}{2}}(\hat{x}) \) by assumption (a). Furthermore, (3.37) implies that
\[ \text{dist}(0, Q_{x_0}^{-1}(x_0)) = \text{dist}(0f(x_0) + g(x_0) + F(x_0)) \leq \frac{\bar{r}}{3}. \]
It is noted earlier that \( x_0 \in B_{\frac{r}{2}}(\hat{x}) \) and \( 0 \in B_{\frac{r}{2}}(\hat{y}) \) by (3.42). Thus, Lemma 2.1 is applied to get the following inequality:
\[ \text{dist}(x_0, Q_{x_0}^{-1}(0)) \leq \frac{M \text{dist}(0, Q_{x_0}^{-1}(x_0))}{1 - ML_{x}} = \frac{M \text{dist}(0f(x_0) + g(x_0) + F(x_0))}{1 - ML_{x}}. \]
Furthermore, by (3.48), we have that
\[ \text{dist}(0, D(x_0)) = \text{dist}(x_0, Q_{x_0}^{-1}(0)) \leq \frac{M \text{dist}(0f(x_0) + g(x_0) + F(x_0))}{1 - ML_{x}}. \] (3.50)
According to Algorithm 1.1 and using (3.36), (3.37), and (3.50), we have
\[ \|d_0\| \leq \eta \text{ dist}(0,D(x_0)) \]
\[ \leq \frac{\eta M \text{ dist}(0,g(x_0)+f(x_0))}{(1-MLr_1)} \]
\[ \leq \frac{\eta M(L+6\nu)\delta^2}{4(1-MLr_1)} = q \left( \frac{1}{2} \right) \delta. \]

This yields that
\[ \|x_1-x_0\| = \|d_0\| \leq q \left( \frac{1}{2} \right) \delta \]
and therefore, (3.39) is true for \( n = 0 \). We assume that \( x_1,x_2,\ldots,x_k \) are constructed and (3.38), and (3.39) are true for \( n = 0,1,2,\ldots,k-1 \). We show that there exists \( x_{k+1} \) such that (3.38) and (3.39) are also hold for \( n = k \). Since (3.38) and (3.39) are true for each \( n \leq k-1 \), we have the following inequality:
\[ \|x_k-x_\ast\| \leq \sum_{i=0}^{k-1} \|d_i\| + \|x_0-x_\ast\| \leq \eta_0 \delta \sum_{i=0}^{k-1} \left( \frac{1}{2} \right)^i + \delta \leq 2\delta. \]

This shows that (3.38) holds for \( n = k \). Now with almost the same argument as we did for the case when \( n = 0 \), we can show that (3.39) holds for \( n = k \). The proof is complete. \( \square \)

Consider the special case when \( \bar{x} \) is a solution of (1.1) (that is, \( \bar{y} = 0 \)) in Theorem 3.2. We have the following corollary, which gives the local quadratic convergence result for the extended Newton-type method. The proof of this corollary is similar to that we did for Corollary 3.1.

**Corollary 3.2** Suppose that \( \bar{x} \) satisfies \( 0 \in f(\bar{x})+g(\bar{x})+F(\bar{x}) \) and that \( Q^{-1}_f(\cdot) \) is pseudo-Lipschitz around \((0,\bar{x})\). Let \( \eta > 1, \nu > 0, \check{\rho} > 0 \) and suppose that \( \nabla f \) is Lipschitz continuous on \( \mathcal{B}_r(\bar{x}) \) with Lipschitz constant \( L \). Suppose that
\[ \lim_{x \to \bar{x}} \text{ dist}(0,f(x)+g(x)+F(x)) = 0. \]

Then there exist some \( \delta > 0 \) such that any sequence \( \{x_n\} \) generated by Algorithm 1.1 with initial point in \( \mathcal{B}_{\check{\rho}}(\bar{x}) \) converges quadratically to a solution \( \bar{x}^\ast \) of (1.1).

### 4. Concluding remarks
When \( \eta > 1 \), we have established semi-local and local convergence results for the extended Newton-type method under the assumptions that \( Q^{-1}_f(\cdot) \) is Lipschitz-like and \( \nabla f \) is continuous. In particular, if \( \nabla f \) is additionally Lipschitz continuous, we further show that the extended Newton-type method is quadratically convergent. For the case where \( \eta = 1 \), the question, whether the results are true for the extended Newton-type method, is a little bit complicated. However, from the proof of the main theorems, one sees that all the results obtained in the present paper remain true provided that, for any \( x \in \Omega \subseteq X \), the following implication holds:
\[ D(x) \neq \emptyset \implies \exists \ d \in D(x) \text{ such that } \|d\| = \min_{d \in D(x)} \|d\|. \]

To see the detail proof of the above implication, one can refer to (Rashid et al., 2013).

**Funding**
The authors received no direct funding for this research.

**Author details**
M.H. Rashid\(^1\)
E-mail: harun_math@ru.ac.bd

\(^1\) Faculty of Science, Department of Mathematics, University of Rajshahi, Rajshahi-6205, Bangladesh.

**Citation information**
Cite this article as: On the convergence of extended Newton-type method for solving variational inclusions, M.H. Rashid, Cogent Mathematics (2014), 1: 980600.
References
Alexis, C. J., & Pietrus, A. (2008). On the convergence of some methods for variational inclusions. Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A, Matemáticas, 102, 355–361.
Argyros, I. K. (2004). A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space. Journal of Mathematical Analysis and Applications, 298, 374–397.
Argyros, I. K. (2007). Computational theory of iterative methods. Studies of computer mathematics. Amsterdam: Elsevier.
Argyros, I. K., & Hilout, S. (2008). Local convergence of Newton-like methods for generalized equations. Applied Mathematics and Computer, 197, 507–514.
Aubin, J. P. (1984). Lipschitz behavior of solutions to convex minimization problems. Mathematics of Operations Research, 9, 87–111.
Aubin, J. P., & Frankowska, H. (1990). Set-valued analysis. Boston: Birkhäuser.
Dedieu, J. P., & Kim, M. H. (2002). Newton's method for analytic systems of equations with constant rank derivatives. Journal of Complexity, 18, 187–209.
Dedieu, J. P., & Shub, M. (2000). Newton's method for overdetermined systems of equations. Mathematics of Computation, 69, 1099–1115.
Dontchev, A. L. (1996a). The Graves theorem revisited. Journal of Convex Analysis, 3, 45–53.
Dontchev, A. L. (1996b). Local convergence of the Newton method for generalized equation. C.R.A.S Paris Series-I, 322, 327–331.
Dontchev, A. L. (1996c). Uniform convergence of the Newton method for Aubin continuous maps. Serdica Mathematical Journal, 22, 385–398.
Dontchev, A. L., & Hager, W. W. (1994). An inverse mapping theorem for set-valued maps. Proceedings of the American Mathematical Society, 121, 481–498.
Hilout, S., Alexis, C. J., & Pietrus, A. (2008). A semilocal convergence of the secant-type method for solving a generalized equations. Positivity, 10, 673–700.
Ioffe, A. D., & Tikhomirov, V. M. (1979). Theory of extremal problems. Studies in mathematics and its applications. Amsterdam: North-Holland.
Li, C., & Ng, K. F. (2007). Majorizing functions and convergence of the Gauss-Newton method for convex composite optimization. SIAM Journal on Optimization, 18, 613–642.
Pietrus, A. (2000a). Generalized equations under mild differentiability conditions. Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales – Serie A: Matemáticas, 94(1), 15–18.
Pietrus, A. (2000b). Does Newton's method for set-valued maps converges uniformly in mild differentiability context? Revista Colombiana de Matemáticas, 32, 49–56.
Rashid, M. H., Wong, J. H., & Li, C. (2012). Convergence analysis of a method for variational inclusions. Applicable Analysis, 91, 1943–1956. Retrieved from: http://dx.doi.org/10.1080/00036811.2011.618127
Rashid, M. H., Yu, S. H., Li, C., & Wu, S. Y. (2013). Convergence analysis of the Gauss–Newton-type method for Lipschitz-like mappings. Journal of Optimization Theory and Applications, 158, 216–233. doi:10.1007/s10957-012-0206-3
Xu, X. B., & Li, C. (2008). Convergence criterion of Newton's method for singular systems with constant rank derivatives. Journal of Mathematical Analysis and Applications, 345, 689–701.

Cogent OA (ISSN: 2331–1835) is published by Cogent OA, part of Taylor & Francis Group.
Publishing with Cogent OA ensures:
• Immediate, universal access to your article on publication
• High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
• Download and citation statistics for your article
• Rapid online publication
• Input from, and dialog with, expert editors and editorial boards
• Retention of full copyright of your article
• Guaranteed legacy preservation of your article
• Discounts and waivers for authors in developing regions
Submit your manuscript to a Cogent OA journal at www.CogentOA.com

© 2014 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license.
You are free to:
Share — copy and redistribute the material in any medium or format
Adapt — remix, transform, and build upon the material for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.
Under the following terms:
Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.
You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
No additional restrictions
You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Rashid, Cogent Mathematics (2014), 1: 980600
http://dx.doi.org/10.1080/23311835.2014.980600