Journal of Computational and Applied Mathematics 279 (2015) 318–335

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

A robust semi-local convergence analysis of Newton's method for cone inclusion problems in Banach spaces under affine invariant majorant condition

O.P. Ferreira

IME/UFG, CP-131, CEP 74001-970- Goiânia, GO, Brazil

ARTICLE INFO

Article history:
Received 14 November 2013
Received in revised form 15 March 2014

Keywords:
Inclusion problems
Newton's method
Majorant condition
Semi-local convergence

ABSTRACT

A semi-local analysis of Newton's method for solving nonlinear inclusion problems in Banach space is presented in this paper. Under an affine majorant condition on the nonlinear function which is associated to the inclusion problem, the robust convergence of the method and results on the convergence rate are established. Using this result we show that the robust analysis of the Newton's method for solving nonlinear inclusion problems under affine Lipschitz-like and affine Smale's conditions can be obtained as a special case of the general theory. Besides for the degenerate cone, which the nonlinear inclusion becomes a nonlinear equation, our analysis retrieves the classical results on semi-local analysis of Newton's method.

Contents

1. Introduction

The idea of solving nonlinear inclusion problems of the form

$$\text{Find } \bar{x} \text{ such that } F(\bar{x}) \in C, \tag{1}$$

E-mail address: orizon@ufg.br.

http://dx.doi.org/10.1016/j.cam.2014.11.019

where C is a nonempty closed convex cone in a Banach space $\mathbb{Y}$ and F is a function from a Banach space $\mathbb{X}$ into $\mathbb{Y}$, plays a huge role in classical analysis and its applications. For instance, the special case in which C is the degenerate cone $\{0\} \subset \mathbb{Y}$, the inclusion problem in (1) corresponds to a nonlinear equation. In the case for which $\mathbb{X} = \mathbb{R}^n$, $\mathbb{Y} = \mathbb{R}^{p+q}$ and $C = \mathbb{R}^p_- \times \{0\}$ are the products of the nonpositive orthant in $\mathbb{R}^p$ with the origin in $\mathbb{R}^q$, the inclusion problem in (1) corresponds to nonlinear systems of p inequalities and q equality, for example see [1–9]. For real-life applications of nonlinear equations solution techniques in different fields see [10–15].

Newton's method and its variant are powerful tools for solving nonlinear equation in Banach space, having a wide range of applications in pure and applied mathematics, see [16,1,4,17,5,7,18–20]. Newton's Method has been extended in order to solve nonlinear systems of equalities and inequalities (see [3,21,9]) and several other different purposes, see [22,23,16,1,24,17,5,25–27,20,28]. In particular, Robinson in [21] (see also [29,8,28]) generalized Newton's method for solving problems of the form (1), which becomes the usual Newton's method to the special case in which C is the degenerate cone $\{0\} \subset \mathbb{Y}$.

One of the usual hypotheses used in convergence analysis of Newton's method is that the Lipschitz continuity of the derivative of the nonlinear operator in question or something like Lipschitz continuity is critical; that is, keeping control of the derivative is an important point in the convergence analysis of Newton's method. These assumptions seem to be actually natural in analysis of Newton's method, [30,17,5,31–36]. On the other hand, in the last few years, a couple of papers have dealt with the issue of convergence analysis of Newton's method and its variants by relaxing the assumption of Lipschitz continuity of the derivative of the function, which define the nonlinear equation in consideration, see [30,37,31,38,39,32,40,33,8,25,41,34,26,27,35,42,36]. In particular, in [32], under a affine majorant condition, a semi-local convergence, as well as results on the convergence rate is established. The advantage of working with a affine majorant condition rests in the fact that it allow to unify converge results on Newton's method without previous relationship, see [32], besides makes the results insensitive with respect to invertible continuous linear mapping, see [4,43]. It is worth mentioning that the affine majorant condition used in [32] is equivalent to Wang's condition (see [42]) whenever the derivative of majorant function is convex.

In [40] a new technique for analyzing the convergence properties of Newton's Method which simplifies the analyses and proof of the results has been introduced. After that, this technique was successfully employed for analyzing Newton's method in difference context, [23,32,25,41,26,27]. In the present work, we will use the technique introduced in [40] to present a semi-local convergence analysis of Newton's method for solving a nonlinear inclusion problems of the form (1). In our analysis, the classical Lipschitz continuity of the derivative of the nonlinear function, which define the nonlinear inclusion in consideration, is relaxed by using a affine majorant function. Although the semi-local convergence analysis of the Newton method for solving nonlinear inclusion under Lipschitz-like (see, [38,8,21]) or Wang's condition (see [29,44,28]) is well known, as far as we know, the *robust* semi-local convergence analysis of Newton's method for solving nonlinear inclusion problems in Banach space under general affine invariant assumptions is a new contribution of this paper.

The basic idea of the analysis is to find a good region for Newton's Method. In this region, the majorant function indeed bounds the nonlinear function associated to the inclusion problem, the behavior of Newton's method is controlled by Newton's iteration of the majorant function and, as a consequence, the region is invariant under Newton's iteration multifunction. Besides, in this region, the analysis provides a clear relationship between the majorant function and the nonlinear function under consideration and allows us to obtain bounds for the Q-quadratic convergence of the method, which depend on the behavior of the majorant function at its smallest zero. It is worth pointing out that, the technique employed also makes it possible to analyze the method in the presence of errors in the initial point, which shows that the assumption on the initial point is robust. Finally, using this convergence result we obtain a robust analysis of Newton's method for solving nonlinear inclusion problems under affine invariant Lipschitz-like and Smale's conditions as a special case of the general theory. We remark that, for the special case in which C is the degenerate cone $\{0\} \subset \mathbb{Y}$, the analysis presented merge in the usual local convergence analysis on Newton's method, see [32].

The organization of the paper is as follows. In Section 1.1, some notations and basic results used in the paper are presented. In Section 2, the main results are stated and in Section 2.1 some properties of the majorant function are established and the main relationships between the majorant function and the nonlinear operator used in the paper are presented. In Section 2.4, the main results are proved and the applications of this results are given in Section 3. Some final remarks are made in Section 4.

1.1. Notation and auxiliary results

The following notations and results are used throughout our presentation. We begin with the following elementary convex analysis result:

Proposition 1. *Let $I \subset \mathbb{R}$ be an interval and $\varphi : I \to \mathbb{R}$ be convex. For any $s_0 \in \operatorname{int}(I)$, the left derivative there exist (in $\mathbb{R}$)*

$$D^-\varphi(s_0) := \lim_{s \to s_0^-} \frac{\varphi(s_0) - \varphi(s)}{s_0 - s} = \sup_{s<s_0} \frac{\varphi(s_0) - \varphi(s)}{s_0 - s}.$$

Moreover, if $s, t, r \in I$, $s < r$, and $s \leqslant t \leqslant r$ then $\varphi(t) - \varphi(s) \leqslant [\varphi(r) - \varphi(s)]\,[(t - s)/(r - s)]$.

Let $\mathbb{X}$ be a Banach space. The *open* and *closed ball* at x with radius $\delta > 0$ are denoted, respectively, by

$$B(x, \delta) := \{y \in \mathbb{X} : \|x - y\| < \delta\} \quad \text{and} \quad B[x, \delta] := \{y \in \mathbb{X} : \|x - y\| \leqslant \delta\}.$$

Proposition 2. *Let $\{z_k\}$ be a sequence in $\mathbb{X}$ and $\Theta > 0$. If $\{z_k\}$ converges to z_* and satisfies*

$$\|z_{k+1} - z_k\| \leq \Theta \|z_k - z_{k-1}\|^2, \quad k = 1, 2, \ldots \tag{2}$$

then $\{z_k\}$ converges Q-quadratically to z_* as follows

$$\limsup_{k\to\infty} \frac{\|z_{k+1} - z_*\|}{\|z_k - z_*\|^2} \leq \Theta.$$

Proof. The proof follows the same pattern as the proof of Proposition 2 of [38], since finite dimensionality plays no role. □

Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces. A set valued mapping $T : \mathbb{X} \rightrightarrows \mathbb{Y}$ is called *sublinear* or *convex process* when its graph is a convex cone, i.e.,

$$0 \in T(0), \qquad T(\lambda x) = \lambda T(x) \quad \forall \lambda > 0, \qquad T(x + x') \supseteq T(x) + T(x'), \quad \forall x, x' \in \mathbb{X}. \tag{3}$$

(sublinear mapping has been extensively studied in [5,45–47]) the following definitions and results about sublinear mappings will be needed: The *domain* and *range* of a sublinear mapping T are defined, respectively, by

$$\operatorname{dom} T := \{d \in \mathbb{X} : Td \neq \varnothing\}, \qquad \operatorname{rge} T := \{y \in \mathbb{Y} : y \in T(x) \text{ for some } x \in \mathbb{X}\},$$

and the *inverse* $T^{-1} : \mathbb{Y} \rightrightarrows \mathbb{X}$ of a sublinear mapping T is another sublinear mapping defined by

$$T^{-1}y := \{d \in \mathbb{X} : y \in Td\}, \quad y \in \mathbb{Y}. \tag{4}$$

The *norm* (or inner norm as is called in see [5]) of a sublinear mapping T is defined by

$$\|T\| := \sup\{\|Td\| : d \in \operatorname{dom} T, \ \|d\| \leqslant 1\}, \tag{5}$$

where $\|Td\| := \inf\{\|v\| : v \in Td\}$ for $Td \neq \varnothing$. We use the convention $\|Td\| = +\infty$ for $Td = \varnothing$, it will be also convenient to use the convention $Td + \varnothing = \varnothing$ for all $d \in \mathbb{X}$.

Lemma 3. *Let $T : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a sublinear mapping with closed graph. Then* $\operatorname{dom} T = \mathbb{X}$ *if and only if* $\|T\| < +\infty$ *and* $\operatorname{rge} T = \mathbb{Y}$ *if and only if* $\|T^{-1}\| < +\infty$.

Proof. See Corollary 5C.2 of [5]. □

Let $S, T : \mathbb{X} \rightrightarrows \mathbb{Y}$ and $U : \mathbb{Y} \rightrightarrows \mathbb{Z}$ be sublinear mappings. The scalar *multiplication*, *addition* and *composition* of sublinear mappings are sublinear mappings defined, respectively, by

$$(\alpha S)(x) := \alpha S(x), \qquad (S + T)(x) := S(x) + T(x), \qquad UT(x) := \bigcup \{U(y) : y \in T(x)\},$$

for all $x \in \mathbb{X}$ and $\alpha > 0$ and the following norm properties there hold:

$$\|\alpha S\| = |\alpha| \, \|S\|, \qquad \|S + T\| \leqslant \|S\| + \|T\|, \qquad \|UT\| \leqslant \|U\| \, \|T\|. \tag{6}$$

Remark 1. Note that definition of the norm in (5) implies that if $\operatorname{dom} T = \mathbb{X}$ and A is a linear mapping from $\mathbb{Z}$ to $\mathbb{X}$ then $\|T(-A)\| = \|TA\|$.

Lemma 4. *Let $S, T : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a sublinear mappings with closed graph such that* $\operatorname{dom} S = \operatorname{dom} T = X$ *and* $\|T^{-1}\| < +\infty$. *Suppose that $\|T^{-1}\| \, \|S\| < 1$ and $(S + T)(x)$ is closed for each $x \in \mathbb{X}$ then* $\operatorname{rge}(S + T)^{-1} = \mathbb{X}$ *and*

$$\|(S + T)^{-1}\| \leqslant \frac{\|T^{-1}\|}{1 - \|T^{-1}\| \, \|S\|}.$$

Proof. These results follow from Theorem 5 of [45] by taking in account Lemma 3. □

Lemma 5. *Let $G : [0, 1] \to \mathbb{Y}$ and $g : [0, 1] \to \mathbb{R}$ be continuous function and $\mathbb{Z}$ a reflexive Banach space. Suppose that $U : \mathbb{Y} \rightrightarrows \mathbb{Z}$ is a sublinear mapping with closed graphic such that* $\operatorname{dom} U \supseteq \operatorname{rge} G$. *If*

$$\|UG(\tau)\| \leqslant g(\tau), \quad \tau \in [0, 1],$$

then there hold

$$U \int_0^1 G(\tau) d\tau \neq \varnothing, \quad \left\| U \int_0^1 G(\tau) d\tau \right\| \leqslant \int_0^1 g(\tau) d\tau.$$

Proof. See Lemma 2.1 of [8]. □

Let $\Omega \subseteq \mathbb{X}$ be an open set and $F : \Omega \to \mathbb{Y}$ a continuously Fréchet differentiable function. The linear map $F'(x) : \mathbb{X} \to \mathbb{Y}$ denotes the Fréchet derivative of $F : \Omega \to \mathbb{Y}$ at $x \in \Omega$. Let $C \subset \mathbb{Y}$ be a nonempty closed convex cone, $z \in \Omega$ and $T_z : \mathbb{X} \rightrightarrows \mathbb{Y}$ a mapping defined as

$$T_z d = F'(z)d - C. \tag{7}$$

It is well known that the mappings T_z and T_z^{-1} are sublinear with closed graphic, $\operatorname{dom} T_z = X$, $\|T_z\| < +\infty$ and $\operatorname{rge} T_z = Y$ if and only if $\|T_z^{-1}\| < +\infty$ (see Lemma 3 and Corollary 4A.7, Corollary 5C.2 and Example 5C.4 of [5]). Note that

$$T_z^{-1}y := \{d \in \mathbb{X} : F'(z)d - y \in C\}, \quad \forall z \in \Omega,\ \forall y \in \mathbb{Y}. \tag{8}$$

Lemma 6. *Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces. $\Omega \subseteq \mathbb{X}$ an open set and $F : \Omega \to \mathbb{Y}$ a continuously Fréchet differentiable function. Then the following inclusion holds*

$$T_z^{-1}F'(v)T_v^{-1}w \subseteq T_z^{-1}w, \quad \forall\, v, z \in \Omega,\ \forall\, w \in \mathbb{Y}.$$

As a consequence,

$$\left\|T_z^{-1}\left[F'(y) - F'(x)\right]\right\| \leq \left\|T_z^{-1}F'(v)T_v^{-1}\left[F'(y) - F'(x)\right]\right\|, \quad \forall\, v, x, y, z \in \Omega.$$

Proof. If $T_v^{-1}w = \varnothing$ for some $v \in \Omega$ and $w \in \mathbb{Y}$ then the inclusion holds trivially. Assume that $T_v^{-1}w \neq \varnothing$, i.e., there exists $d \in T_v^{-1}w$ for each $v \in \Omega$ and $w \in \mathbb{Y}$. Note that definition in (8) and simple algebraic manipulation give

$$T_z^{-1}F'(v)d = \left\{u \in \mathbb{X} : F'(z)u - w \in C + [F'(v)d - w]\right\}, \quad \forall\, d \in \mathbb{X},\ \forall\, w \in \mathbb{Y}.$$

Definition in (8) implies that, for each $v \in \Omega$, $w \in \mathbb{Y}$ and $d \in T_v^{-1}w$ there holds $F'(v)d - w \in C$. Thus, the last equality and (8) imply that

$$T_z^{-1}F'(v)d = T_z^{-1}w, \quad \forall\, w \in \mathbb{Y},\ \forall\, d \in T_v^{-1}w,$$

which implies the desired inclusion. To end the proof, note that the first part of the lemma implies

$$\left\|T_z^{-1}\left[F'(y) - F'(x)\right]u\right\| \leq \left\|T_z^{-1}F'(v)T_v^{-1}\left[F'(y) - F'(x)\right]u\right\|, \quad \forall\, v, x, y, z \in \Omega,\ \forall\, u \in \mathbb{Y}.$$

Hence, the inequality of the lemma follows from the definition of the norm in (5). □

2. Semi-local analysis for Newton's method

Our goal is to state and prove a robust semi-local affine invariant theorem for Newton's method to solve nonlinear inclusion of the form

$$F(x) \in C, \tag{9}$$

where $F : \Omega \to \mathbb{Y}$ is a nonlinear continuously differentiable function, $\mathbb{X}$ and $\mathbb{Y}$ are Banach spaces and $\mathbb{X}$ is reflexive, $\Omega \subseteq \mathbb{X}$ an open set and $C \subset \mathbb{Y}$ a nonempty closed convex cone. To state the theorem we need some definitions.

A nonlinear continuously Fréchet differentiable function $F : \Omega \to \mathbb{Y}$ satisfies *Robson's Condition* at $\tilde{x} \in \Omega$ if

$$\operatorname{rge} T_{\tilde{x}} = \mathbb{Y}, \tag{10}$$

where $T_{\tilde{x}} : \mathbb{X} \rightrightarrows \mathbb{Y}$ is a sublinear mapping as defined in (7).

Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces, $\Omega \subseteq \mathbb{X}$ an open set and $R > 0$ a scalar constant. A continuously differentiable scalar function $f : [0, R) \to \mathbb{R}$ is a *majorant function* at a point $\tilde{x} \in \Omega$ for a continuously differentiable function $F : \Omega \to \mathbb{Y}$ if

$$B(\tilde{x}, R) \subseteq \Omega, \qquad \left\|T_{\tilde{x}}^{-1}\left[F'(y) - F'(x)\right]\right\| \leqslant f'(\|y - x\| + \|x - \tilde{x}\|) - f'(\|x - \tilde{x}\|), \tag{11}$$

for all $x, y \in B(\tilde{x}, R)$ such that $\|x - \tilde{x}\| + \|y - x\| < R$ and satisfies the following conditions:

(h1) $f(0) > 0, f'(0) = -1$;
(h2) f' is convex and strictly increasing;
(h3) $f(t) = 0$ for some $t \in (0, R)$.

We also need of the following condition on the majorant condition f which will be considered to hold only when explicitly stated

(h4) $f(t) < 0$ for some $t \in (0, R)$.

Remark 2. Since $f(0) > 0$ and f is continuous then condition **h4** implies condition **h3**.

Theorem 7. *Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and $\mathbb{X}$ reflexive, $\Omega \subseteq \mathbb{X}$ an open set, $F : \Omega \to \mathbb{Y}$ a continuously Fréchet differentiable function, $C \subset \mathbb{Y}$ a nonempty closed convex cone, $R > 0$ and $f : [0, R) \to \mathbb{R}$ a continuously differentiable function. Suppose that $\tilde{x} \in \Omega$, F satisfies Robson's Condition at $\tilde{x}$, f is a majorant function for F at $\tilde{x}$ and*

$$\left\| T_{\tilde{x}}^{-1}(-F(\tilde{x})) \right\| \leqslant f(0). \tag{12}$$

Then f has a smallest zero $t_* \in (0, R)$, the sequences generated by Newton's Method for solving the inclusion $F(x) \in C$ and the equation $f(t) = 0$, with starting point $x_0 = \tilde{x}$ and $t_0 = 0$, respectively,

$$x_{k+1} \in x_k + \operatorname{argmin}\left\{ \|d\| : F(x_k) + F'(x_k)d \in C \right\}, \qquad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, \quad k = 0, 1, \ldots \tag{13}$$

are well defined, $\{x_k\}$ is contained in $B(\tilde{x}, t_*)$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$ and converges to t_* and satisfies the inequalities

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \qquad \|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} \|x_k - x_{k-1}\|^2, \tag{14}$$

for all $k = 0, 1, \ldots$, and $k = 1, 2, \ldots$, respectively. Moreover, $\{x_k\}$ converges to $x_* \in B[\tilde{x}, t_*]$ such that $F(x_*) \in C$,

$$\|x_* - x_k\| \le t_* - t_k, \qquad t_* - t_{k+1} \le \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \ldots \tag{15}$$

and, therefore, $\{t_k\}$ converges Q-linearly to t_* and $\{x_k\}$ converges R-linearly to x_*. If, additionally, f satisfies **h4** *then the following inequalities hold:*

$$\|x_{k+1} - x_k\| \le \frac{D^- f'(t_*)}{-2f'(t_*)} \|x_k - x_{k-1}\|^2, \qquad t_{k+1} - t_k \le \frac{D^- f'(t_*)}{-2f'(t_*)} (t_k - t_{k-1})^2, \quad k = 1, 2, \ldots, \tag{16}$$

and, as a consequence, $\{x_k\}$ and $\{t_k\}$ converge Q-quadratically to x_* and t_*, respectively, as follows

$$\limsup_{k \to \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \le \frac{D^- f'(t_*)}{-2f'(t_*)}, \qquad t_* - t_{k+1} \le \frac{D^- f'(t_*)}{-2f'(t_*)} (t_* - t_k)^2, \quad k = 0, 1, \ldots. \tag{17}$$

We will use the above result to prove a robust semi-local affine invariant theorem for Newton's method for solving nonlinear inclusion of the form (9). The statement of the theorem is:

Theorem 8. *Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and $\mathbb{X}$ reflexive, $\Omega \subseteq \mathbb{X}$ an open set, $F : \Omega \to \mathbb{Y}$ a continuously Fréchet differentiable function, $C \subset \mathbb{Y}$ a nonempty closed convex cone, $R > 0$ and $f : [0, R) \to \mathbb{R}$ a continuously differentiable function. Suppose that $\tilde{x} \in \Omega$, F satisfies Robson's Condition at $\tilde{x}$, f is a majorant function for F at $\tilde{x}$ satisfying* **h4** *and*

$$\left\| T_{\tilde{x}}^{-1}(-F(\tilde{x})) \right\| \leqslant f(0). \tag{18}$$

Define $\beta := \sup\{-f(t) : t \in [0, R)\}$. Let $0 \le \rho < \beta/2$ and $g : [0, R - \rho) \to \mathbb{R}$,

$$g(t) := \frac{-1}{f'(\rho)} [f(t + \rho) + 2\rho]. \tag{19}$$

Then g has a smallest zero $t_{*,\rho} \in (0, R - \rho)$, the sequences generated by Newton's Method for solving the inclusion $F(x) \in C$ and the equation $g(t) = 0$, with starting point $x_0 = \hat{x}$ for any $\hat{x} \in B(\tilde{x}, \rho)$ and $t_0 = 0$, respectively,

$$x_{k+1} \in x_k + \operatorname{argmin}\left\{ \|d\| : F(x_k) + F'(x_k)d \in C \right\}, \qquad t_{k+1} = t_k - \frac{g(t_k)}{g'(t_k)}, \quad k = 0, 1, \ldots \tag{20}$$

are well defined, $\{x_k\}$ is contained in $B(\tilde{x}, t_{*,\rho})$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_{*,\rho})$ and converges to $t_{*,\rho}$ and satisfies the inequalities

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \quad k = 0, 1, \ldots, \tag{21}$$

$$\|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} \|x_k - x_{k-1}\|^2 \le \frac{D^- g'(t_{*,\rho})}{-2g'(t_{*,\rho})} \|x_k - x_{k-1}\|^2, \quad k = 1, 2, \ldots. \tag{22}$$

Moreover, $\{x_k\}$ converges to $x_* \in B[\tilde{x}, t_{*,\rho}]$ such that $F(x_*) \in C$, satisfies the inequalities

$$\|x_* - x_k\| \le t_{*,\rho} - t_k, \qquad t_{*,\rho} - t_{k+1} \le \frac{1}{2}(t_{*,\rho} - t_k), \quad k = 0, 1, \ldots \tag{23}$$

and the convergence of $\{x_k\}$ *and* $\{t_k\}$ *to* x_* *and* $t_{*,\rho}$*, respectively, is Q-quadratic as follows*

$$\limsup_{k\to\infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \le \frac{D^- g'(t_{*,\rho})}{-2g'(t_{*,\rho})}, \qquad t_{*,\rho} - t_{k+1} \le \frac{D^- g'(t_{*,\rho})}{-2g'(t_{*,\rho})}(t_{*,\rho} - t_k)^2, \quad k = 0, 1, \ldots. \tag{24}$$

Remark 3. Using Lemma 3 it is easy to see that the inequalities (11), (12) and (18) are well defined.

Remark 4. It follows from (7) and (8) that definition of the sequence $\{x_k\}$ in (13) is equivalent to the conditions

$$x_{k+1} - x_k \in T_{x_k}^{-1}(-F(x_k)) \quad \text{and} \quad \|x_{k+1} - x_k\| = \left\|T_{x_k}^{-1}(-F(x_k))\right\|, \quad k = 0, 1, \ldots.$$

Remark 5. Theorems 7 and 8 are affine-invariant in the following sense: Letting $A : \mathbb{Y} \to \mathbb{Y}$ be an invertible continuous linear mapping, $\tilde{F} := A \circ F$ and the set $\tilde{C} := A(C)$, the corresponding inclusion problem (9) is given by

$$\tilde{F}(x) \in \tilde{C},$$

and the convex process associated is denoted by $\tilde{T}_{\tilde{x}} d := \tilde{F}(\tilde{x})d - \tilde{C}$. Then $\tilde{T}_{\tilde{x}} = A \circ T_{\tilde{x}}$ and $\tilde{T}_{\tilde{x}}^{-1} = T_{\tilde{x}}^{-1} \circ A^{-1}$. Moreover, the conditions rge $\tilde{T}_{\tilde{x}} = Y$,

$$\left\|\tilde{T}_{\tilde{x}}^{-1}(-\tilde{F}(\tilde{x}))\right\| \leqslant f(0),$$

and the affine majorant condition (Lipschitz-like condition)

$$\left\|\tilde{T}_{\tilde{x}}^{-1}\left[|\tilde{F}'(y) - |\tilde{F}'(x)\right]\right\| \leqslant f'(\|y - x\| + \|x - \tilde{x}\|) - f'(\|x - \tilde{x}\|),$$

for $x, y \in B(\tilde{x}, R)$, $\|x - \tilde{x}\| + \|y - x\| < R$, are equivalent to rge $T_{\tilde{x}} = \mathbb{Y}$, (12) and (11) respectively. Therefore, the assumptions in Theorems 7 and 8 are insensitive with respect to invertible continuous linear mappings. Note that the results of [21] do not have such property. For more details about affine invariant theorems see [4,43].

2.1. Preliminary results

In this section, we will prove all the statements in Theorems 7 and 8 regarding to the majorant function and the sequence $\{t_k\}$ associated. The main relationships between the majorant function and the nonlinear operator will be also established.

2.2. The majorant function

In this section we will study the majorant function f and prove all results regarding only the sequence $\{t_k\}$. Define

$$\bar{t} := \sup\left\{t \in [0, R) : f'(t) < 0\right\}. \tag{25}$$

Proposition 9. *The majorant function* f *has a smallest root* $t_* \in (0, R)$*, is strictly convex and*

$$f(t) > 0, \qquad f'(t) < 0, \quad t < t - f(t)/f'(t) < t_*, \ \forall\, t \in [0, t^*). \tag{26}$$

Moreover, $f'(t_*) \leqslant 0$ *and*

$$f'(t_*) < 0 \iff \exists\, t \in (t_*, R), \qquad f(t) < 0. \tag{27}$$

If, additionally, f *satisfies condition* **h4** *then the following statements hold:*

(i) $f'(t) < 0$ *for any* $t \in [0, \bar{t})$*;*
(ii) $0 < t_* < \bar{t} \le R$*;*
(iii) $\beta = -\lim_{t\to\bar{t}_-} f(t),\ 0 < \beta < \bar{t}$*;*
(iv) *If* $0 \le \rho < \beta/2$ *then* $\rho < \bar{t}/2 < \bar{t}$ *and* $f'(\rho) < 0$.

Proof. See Propositions 2.3 and 5.2 of [39] and Proposition 3 of [32]. □

In view of the second inequality in (26), Newton iteration is well defined in $[0, t_*)$. Let us call it

$$\begin{array}{rcl} n_f : [0, t_*) & \to & \mathbb{R} \\ t & \mapsto & t - f(t)/f'(t). \end{array} \tag{28}$$

Proposition 10. *Newton iteration n_f is strictly increasing, maps $[0, t^*)$ in $[0, t^*)$.*

Proof. See Proposition 4 of [32]. □

The definition of $\{t_k\}$ in Theorem 7 is equivalent to the following one

$$t_0 = 0, \qquad t_{k+1} = n_f(t_k), \quad k = 0, 1, \ldots. \tag{29}$$

Corollary 11. *The sequence $\{t_k\}$ is well defined, is strictly increasing, is contained in $[0, t_*)$ and converges Q-linearly to t_* as second inequality in (15). If f also satisfies* **h4***, then the second inequality in (16) holds and, moreover, $\{t_k\}$ converges Q-quadratically to t_* as in (17).*

Proof. The proof follows the same pattern as the proof of Corollary 2.15 of [38]. □

2.3. Relationship between the majorant function and the nonlinear operator

In this section, we will present the main relationships between the majorant function f and the nonlinear operator F that we need for proving Theorem 7.

Proposition 12. *If $\|x - \tilde{x}\| \leqslant t < \bar{t}$ then* $\operatorname{dom}(T_x^{-1}F'(\tilde{x})) = \mathbb{X}$ *and there holds*

$$\left\|T_x^{-1}F'(\tilde{x})\right\| \leqslant -1/f'(t).$$

As a consequence, $\operatorname{rge} T_x = \mathbb{Y}$.

Proof. Take $0 \leqslant t < \bar{t}$ and $x \in B[\tilde{x}, t]$. Since f is a majorant function for F at $\tilde{x}$, using (11), **h2**, **h1** and Proposition 9 item i we obtain

$$\begin{aligned}\left\|T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})]\right\| &\leqslant f'(\|x - \tilde{x}\|) - f'(0)\\ &\leqslant f'(t) - f'(0)\\ &= f'(t) + 1 < 1.\end{aligned}$$

To simplify the notation define $S = T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})]$. Since $[F'(x) - F'(\tilde{x})]$ is a continuous linear mapping and $T_{\tilde{x}}^{-1}$ is a sublinear mapping with closed graph, it is easy to see that S is a sublinear mapping with closed graph. Moreover, as by assumption $\operatorname{rge} T_{\tilde{x}} = \mathbb{Y}$ we have $\operatorname{dom} S = \mathbb{X}$. Because S has closed graph, it is easy to see that $(S + I)(x)$ has also closed graph for all $x \in \mathbb{X}$, where I is the identity mapping on $\mathbb{X}$. Therefore, applying Lemma 4 with $T = I$ and taking in account above inequality we conclude that $\operatorname{rge}(T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})] + I) = \mathbb{X}$ and

$$\left\|\left(T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})] + I\right)^{-1}\right\| \leqslant \frac{1}{1 - (f'(t) + 1)} = \frac{1}{-f'(t)}. \tag{30}$$

Since $0 \in C$ we have $T_{\tilde{x}}^{-1}F'(\tilde{x})d \ni d$ for all $d \in \mathbb{X}$. Thus, as $T_{\tilde{x}}^{-1}$ is a sublinear mapping, using the last inclusion in (3) we obtain that

$$\left(T_{\tilde{x}}^{-1}F'(x)\right)d \supseteq \left(T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})] + I\right)d, \quad \forall\, d \in \mathbb{X}.$$

In particular, as we know that $\operatorname{rge}(T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})] + I) = \mathbb{X}$, last inclusion implies that

$$\operatorname{rge}\left(T_{\tilde{x}}^{-1}F'(x)\right) = \mathbb{X}. \tag{31}$$

Hence, utilizing again above inclusion and definition of the inverse in (8) we easily conclude that

$$\left(T_{\tilde{x}}^{-1}F'(x)\right)^{-1} v \supseteq \left(T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})] + I\right)^{-1} v, \quad \forall\, v \in \mathbb{Y}.$$

Taking into account the definition of the norm in (5), the last inclusion implies that

$$\left\|\left(T_{\tilde{x}}^{-1}F'(x)\right)^{-1}\right\| \leqslant \left\|\left(T_{\tilde{x}}^{-1}[F'(x) - F'(\tilde{x})] + I\right)^{-1}\right\|. \tag{32}$$

On the other hand, as $F'(x)$ is a linear mapping, using the definitions in (4) and (8) we obtain that

$$\begin{aligned}-\left(T_{\tilde{x}}^{-1}F'(x)\right)^{-1} d &= \{-v \in \mathbb{X} : F'(\tilde{x})d - F'(x)v \in C\}\\ &= \{u \in \mathbb{X} : F'(x)u + F'(\tilde{x})d \in C\}\\ &= T_x^{-1}(-F'(\tilde{x}))d,\end{aligned}$$

for all $d \in \mathbb{X}$. Thus using the last equality, definition of the norm in (5), (31) and Lemma 3 we have

$$\left\|T_x^{-1}(-F'(\tilde{x}))\right\| = \left\|-\left(T_{\tilde{x}}^{-1}F'(x)\right)^{-1}\right\| = \left\|\left(T_{\tilde{x}}^{-1}F'(x)\right)^{-1}\right\| < +\infty. \tag{33}$$

Thus, since $F'(\tilde{x})$ is a linear mapping, Remark 1 and last inequality allow us to conclude that

$$\left\|T_x^{-1}F'(\tilde{x})\right\| = \left\|T_x^{-1}(-F'(\tilde{x}))\right\| < +\infty,$$

which combined with Lemma 3 implies the first statement of the proposition. Moreover, the desired inequality follows by combination of (30), (32), (33) with the last equality.

For proof of the last statement of the proposition, first note that definition of the norm in (5) and Lemma 6 give us

$$\|T_x^{-1}\| \leqslant \|T_x^{-1}F'(\tilde{x})T_{\tilde{x}}^{-1}\|.$$

From Lemma 3 the assumption rge $T_{\tilde{x}} = Y$ implies $\|T_{\tilde{x}}^{-1}\| < +\infty$ and first part of the proposition implies $\|T_x^{-1}F'(\tilde{x})\| < \infty$, hence (6) and above inequality yield $\|T_x^{-1}\| < +\infty$ which, by using again Lemma 3, gives the desired result. □

Newton's iteration at a point $x \in \Omega$ happens to be a solution of the linearization of the inclusion $F(y) \in C$ at such a point, namely, a solution of the linear inclusion $F(x) + F'(x)(x - y) \in C$. So, we study the linearization error of F at a point in Ω

$$E(x, y) := F(y) - \left[F(x) + F'(x)(y - x)\right], \quad y,\ x \in \Omega. \tag{34}$$

We will bound this error by the error in the linearization on the majorant function f

$$e(t, s) := f(s) - [f(t) + f'(t)(s - t)], \quad t, s \in [0, R). \tag{35}$$

Lemma 13. *Take $x, y \in B(\tilde{x}, R)$ and $0 \leqslant t < s < R$. If $\|x - \tilde{x}\| \leqslant t$ and $\|y - x\| \leqslant s - t$, then*

$$\left\|T_{\tilde{x}}^{-1}(-E(x, y))\right\| \leqslant e(t, s)\left[\frac{\|y - x\|}{s - t}\right]^2.$$

Proof. As $x, y \in B(\tilde{x}, R)$ and the ball is convex $x + \tau(y - x) \in B(\tilde{x}, R)$, for all $\tau \in [0, 1]$. Since, by assumption, rge $T_{\tilde{x}} = \mathbb{Y}$ we obtain that dom $T_{\tilde{x}}^{-1} = \mathbb{Y}$. On the other hand, taking into account Remark 1 and that $F'(.)$ is a linear mapping on dom F, we conclude

$$\left\|T_{\tilde{x}}^{-1}\left(-[F'(x + \tau(y - x)) - F'(x)](y - x)\right)\right\| \leq \left\|T_{\tilde{x}}^{-1}[F'(x + \tau(y - x)) - F'(x)]\right\| \|(y - x)\|,$$

for all $\tau \in [0, 1]$. Hence, as f is a majorant function for F at $\tilde{x}$, using (11) and last inequality we have

$$\left\|T_{\tilde{x}}^{-1}\left(-[F'(x + \tau(y - x)) - F'(x)](y - x)\right)\right\| \leqslant \left[f'\left(\|x - \tilde{x}\| + \tau\|y - x\|\right) - f'\left(\|x - \tilde{x}\|\right)\right]\|y - x\|,$$

for all $\tau \in [0, 1]$. Thus, as dom $T_{\tilde{x}}^{-1} = \mathbb{Y}$ we apply Lemma 5 with $U = T_{\tilde{x}}^{-1}$, $G(\tau)$ equal to the expression in parentheses on the left hand side of last inequality and $g(\tau)$ equal to the expression on the right hand side of that inequality, obtaining

$$\left\|T_{\tilde{x}}^{-1}\int_0^1 -[F'(x + \tau(y - x)) - F'(x)](y - x)\, d\tau\right\|$$
$$\leqslant \int_0^1 \left[f'\left(\|x - \tilde{x}\| + \tau\|y - x\|\right) - f'\left(\|x - \tilde{x}\|\right)\right]\|y - x\|\, d\tau. \tag{36}$$

Using the convexity of f', that $\|x - \tilde{x}\| < t$, $\|y - x\| < s - t$, $s < R$ and Proposition 1 we have

$$f'\left(\|x - \tilde{x}\| + \tau\|y - x\|\right) - f'\left(\|x - \tilde{x}\|\right) \leqslant f'(t + \tau\|y - x\|) - f'(t)$$
$$\leqslant [f'(t + \tau(s - t)) - f'(t)]\frac{\|y - x\|}{s - t},$$

for any $\tau \in [0, 1]$. Combining the two above inequalities we obtain

$$\left\|T_{\tilde{x}}^{-1}\int_0^1 -[F'(x + \tau(y - x)) - F'(x)](y - x)\, d\tau\right\| \leqslant \int_0^1 [f'(t + \tau(s - t)) - f'(t)]\frac{\|y - x\|^2}{s - t}\, d\tau,$$

which, after performing the integration of the right hand side, taking in account the definition of $e(t, s)$ in (35) and that (34) is equivalent to

$$E(x, y) = \int_0^1 [F'(x + \tau(y - x)) - F'(x)](y - x)\, d\tau,$$

yields the desired result. □

Since $\mathbb{X}$ is reflexive, second part of Proposition 12 guarantees, in particular, that *Newton's step set* $D_{F,C}(x)$ *at* $x \in B(\tilde{x}, t_*)$ *associated to* F *and* C is nonempty, that is,

$$\begin{aligned} D_{F,C}(x) &:= \operatorname{argmin}\left\{\|d\| : F(x) + F'(x)d \in C\right\} \\ &= \operatorname{argmin}\left\{\|d\| : d \in T_x^{-1}(-F(x))\right\} \\ &\neq \varnothing, \end{aligned} \tag{37}$$

and consequently, Newton's iteration multifunction is well defined in $B(\tilde{x}, t_*)$. Let us call $N_{F,C}$ *Newton's iteration multifunction for* F *and* C defined in $B(\tilde{x}, t_*)$:

$$\begin{aligned} N_{F,C} : B(\tilde{x}, t_*) &\rightrightarrows \mathbb{X} \\ x &\mapsto x + D_{F,C}(x). \end{aligned} \tag{38}$$

One can apply a *single* Newton's iteration multifunction on any $x \in B(\tilde{x}, t_*)$ to obtain the set $N_{F,C}(x)$ which may not be contained to $B(\tilde{x}, t_*)$, or even may not in the domain of F. So, this is enough to guarantee the well-definedness of only one iteration. To ensure that Newtonian iteration multifunction may be repeated indefinitely, we need some additional results.

First, define some subsets of $B(\tilde{x}, t_*)$ in which, as we shall prove, Newton iteration (38) is "well behaved".

$$K(t) := \left\{x \in \mathbb{X} : \|x - \tilde{x}\| \leqslant t,\ \left\|T_x^{-1}(-F(x))\right\| \leqslant -\frac{f(t)}{f'(t)}\right\}, \quad t \in [0, t_*), \tag{39}$$

$$K := \bigcup_{t\in[0,t_*)} K(t). \tag{40}$$

In (39), $0 \leqslant t < t_*$, therefore, $f'(t) \neq 0$ and $\operatorname{rge} T_x = \mathbb{Y}$ in $B[\tilde{x}, t] \subset B[\tilde{x}, t_*)$ (Proposition 12). So, the above definitions are consistent.

Lemma 14. *For each* $t \in [0, t_*)$, $x \in K(t)$ *and* $x_+ \in N_{F,C}(x)$ *there hold:*

(i) $\|x_+ - x\| \leq n_f(t) - t$;
(ii) $\|x_+ - \tilde{x}\| \leq n_f(t) < t_*$;
(iii) $\|z - x_+\| \leq -\frac{f(n_f(t))}{f'(n_f(t))}\left[\frac{\|x_+ - x\|}{n_f(t) - t}\right]^2$, *for all* $z \in N_{F,C}(x_+)$.

Proof. Take $t \in [0, t_*)$, $x \in K(t)$. Using definition in (39) and the statements in Proposition 10 we have

$$\|x - \tilde{x}\| \leqslant t, \qquad \|T_x^{-1}(-F(x))\| \leqslant -f(t)/f'(t), \quad t < n_f(t) < t_*. \tag{41}$$

Take $x_+ \in N_{F,C}(x)$. From the definition in (38) we have $x_+ - x \in D_{F,C}(x)$, which taking into account definition in (37) and second inequality in (41) yields

$$\|x_+ - x\| = \|T_x^{-1}(-F(x))\| \leqslant -f(t)/f'(t),$$

and by using definition of n_f in (28) we obtain the item **i**.

Combining first and last inequalities in (41), item **i** and definition of n_f in (28) we have

$$\|x_+ - \tilde{x}\| \leqslant \|x - \tilde{x}\| + \|x_+ - x\| \leqslant t - f(t)/f'(t) = n_f(t) < t_*,$$

which is the inequality in item **ii**.

Now we are going to prove the last item of the lemma. From item **ii** we conclude that

$$N_{F,C}(x) \subseteq B[\tilde{x}, n_f(t)] \subset B(\tilde{x}, t_*), \tag{42}$$

which, in particular, implies that the set $N_{F,C}(x)$ is contained in the domain of the function F. Therefore, we claim that the following relations hold:

$$\varnothing \neq T_{x_+}^{-1}F'(\tilde{x})T_{\tilde{x}}^{-1}(-E(x, x_+)) \subset T_{x_+}^{-1}(-F'(x_+)), \quad \forall x_+ \in N_{F,C}(x) \tag{43}$$

where E is the linearization error of F as defined in (34). Assumption, $\operatorname{rge} T_{\tilde{x}} = \mathbb{Y}$ implies that $\operatorname{dom} T_{\tilde{x}}^{-1} = \mathbb{Y}$. Thus, as $\|x_+ - \tilde{x}\| < t_* \leq \bar{t}$ using that $\operatorname{dom} T_{\tilde{x}}^{-1} = \mathbb{Y}$ together with Proposition 12 the first claim in (43) follows. For proving the inclusion in (43), we first use Lemma 6 for concluding that

$$T_{x_+}^{-1}F'(\tilde{x})T_{\tilde{x}}^{-1}(-E(x, x_+)) \subset T_{x_+}^{-1}(-E(x, x_+)), \quad \forall x_+ \in N_{F,C}(x).$$

Since $x_+ \in N_{F,C}(x)$ it follows from (37) and (38) that $F(x) + F'(x)(x_+ - x) \in C$. Hence, using definition in (8) we easily conclude that

$$T_{x_+}^{-1}(-E(x, x_+)) = T_{x_+}^{-1}(-F'(x_+)), \quad \forall x_+ \in N_{F,C}(x).$$

Therefore, combining two last inclusions we conclude that the inclusion (43) holds, as claimed. On the other hand, since $n_f(t)$ belongs to the domain of f, using the definitions of Newton iterations on (28) and definition of the linearization error (35), we obtain

$$f(n_f(t)) = f(n_f(t)) - \left[f(t) + f'(t)(n_f(t) - t)\right] = e(t, n_f(t)). \tag{44}$$

From (41) we have $\|x - \tilde{x}\| \leqslant t$ and $t < n_f(t) < t_* < R$. Hence, the first inequality in (41) and item **i** allow us to apply Lemma 13 with $y = x_+$ and $s = n_f(t)$ to have

$$\|T_{\tilde{x}}^{-1}(-E(x, x_+))\| \leqslant e(t, n_f(t)) \left[\frac{\|x_+ - x\|}{f(t)/f'(t)}\right]^2 = f(n_f(t)) \left[\frac{\|x_+ - x\|}{f(t)/f'(t)}\right]^2,$$

where, in above expression, was used (44) to obtain the equality. As $x_+ \in N_{F,C}(x)$ implies that $\|x_+ - \tilde{x}\| \leqslant n_f(t) < t_* \leq \bar{t}$, it follows from Proposition 12 that $\operatorname{dom}(T_{x_+}^{-1}F'(\tilde{x})) = \mathbb{X}$ and

$$\|T_{x_+}^{-1}F'(\tilde{x})\| \leqslant -1/f'(n_f(t)).$$

Combining (43) with the two above inequalities and by using property of the norm we obtain

$$\begin{aligned}
\left\|T_{x_+}^{-1}(-F'(x_+))\right\| &\leq \left\|T_{x_+}^{-1}F'(\tilde{x})T_{\tilde{x}}^{-1}(-E(x, x_+))\right\| \\
&\leqslant \left\|T_{x_+}^{-1}F'(\tilde{x})\right\| \left\|T_{\tilde{x}}^{-1}(-E(x, x_+))\right\| \\
&\leqslant -\frac{f(n_f(t))}{f'(n_f(t))} \left[\frac{\|x_+ - x\|}{f(t)/f'(t)}\right]^2,
\end{aligned}$$

hence, as $\|z - x_+\| = \left\|T_{x_+}^{-1}(-F'(x_+))\right\|$ for all $z \in N_{F,C}(x_+)$, the item **iii** is proved. □

Lemma 15. *For each $t \in [0, t_*)$ the following inclusions hold: $K(t) \subset B(\tilde{x}, t_*)$ and*

$$N_{F,C}(K(t)) \subset K\left(n_f(t)\right).$$

As a consequence, $K \subset B(\tilde{x}, t_*)$ and $N_{F,C}(K) \subset K$.

Proof. The first inclusion follows trivially from the definition of $K(t)$. Combining items **i** and **iii** of Lemma 14 and taking in account that

$$\|z - \tilde{x}\| = \left\|T_{\tilde{x}}^{-1}(-F'(\tilde{x}))\right\|,$$

for all $z \in N_{F,C}(\tilde{x})$, it is easy to conclude that the following inequality holds

$$\left\|T_{\tilde{x}}^{-1}(-F'(\tilde{x}))\right\| \leq -\frac{f(n_f(t))}{f'(n_f(t))}.$$

This inequality together with item **ii** of Lemma 14 prove the second inclusion. The next inclusion (first on the second sentence), follows trivially from definitions (39) and (40). To verify the last inclusion, take $x \in K$. Then $x \in K(t)$ for some $t \in [0, t_*)$. Using the first part of the lemma, we conclude that $N_{F,C}(x) \subseteq K(n_f(t))$. To end the proof, note that $n_f(t) \in [0, t_*)$ and use the definition of K. □

2.4. Convergence

Finally, we are ready to prove the main results of the paper, namely, Theorems 7 and 8. First we note that, by using (37) and (38), the sequence $\{x_k\}$ (see (13)) satisfies

$$x_{k+1} \in N_{F,C}(x_k), \quad k = 0, 1, \ldots, \tag{45}$$

which is indeed an equivalent definition of this sequence.

Proof of Theorem 7. All statements involving only $\{t_k\}$ were proved in Corollary 11, namely, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$, converges Q-linearly to t_* as in second inequality in (15) and if f satisfies **h4** then $\{t_k\}$ satisfies the second inequality in (16) and converges Q-quadratically as the second inequality in (17).

Now we are going to prove the statements involving the sequence $\{x_k\}$ with starting point $x_0 = \tilde{x}$. From (12) and **h1**, we have

$$x_0 = \tilde{x} \in K(0) \subset K,$$

where the second inclusion follows trivially from (40). Using the above inclusion, the inclusions $N_{F,C}(K) \subset K$ in Lemma 15 and (45) we conclude that the sequence $\{x_k\}$ is well defined and rests in K. From the first inclusion on second part of Lemma 15 we have trivially that $\{x_k\}$ is contained in $B(\tilde{x}, t_*)$.

We will prove, by induction that

$$x_k \in K(t_k), \quad k = 0, 1, \ldots. \tag{46}$$

The above inclusion, for $k = 0$ follows from (12), assumption **h1** and definition of $K(0)$ in (39). Assume now that $x_k \in K(t_k)$. Thus, using Lemma 15, (45) and (29) we conclude that $x_{k+1} \in K(t_{k+1})$, which completes the induction proof of (46).

Now, combining (29), (45), (46) and item **i** of Lemma 14 we obtain first inequality in (14). Since $\{t_k\}$ converges to t_*, the first inequality (14) implies

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

for any $k_0 \in \mathbb{N}$. Hence, $\{x_k\}$ is a Cauchy sequence in $B(\tilde{x}, t_*)$ and so, converges to some $x_* \in B[\tilde{x}, t_*]$. The above inequality also implies that $\|x_* - x_k\| \leqslant t_* - t_k$, for any k. For proving that $F(x_*) \in C$. First, observe that the first inequality in (15) implies that $\|x_* - \tilde{x}\| \le t_* < R$. Hence, as $B(\tilde{x}, R) \subseteq \Omega$ we conclude that $x_* \in \Omega$. Since definition of $\{x_k\}$ in (13) implies that

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

and F is continuous differentiable in Ω, $\{x_k\} \subset B(\tilde{x}, t_*) \subset B(\tilde{x}, R)$, $\{x_k\}$ converges to $x_* \in \Omega$ and $C \subset \mathbb{Y}$ is closed, the result follows by taking limit as k goes to infinite in above inclusion.

In order to prove the second inequality in (14), first note that $x_k \in K(t_k)$, $x_{k+1} \in N_{F,C}(x_k)$ and $t_{k+1} = n_f(t_k)$, for all $k = 0, 1, \ldots$. Thus, apply item **iii** of Lemma 14 with $x = x_{k-1}$, $z = x_{k+1}$, $x_+ = x_k$ and $t = t_{k-1}$ to obtain

$$\|x_{k+1} - x_k\| \le -\frac{f(t_k)}{f'(t_k)} \left[\frac{\|x_k - x_{k-1}\|}{t_k - t_{k-1}} \right]^2,$$

which, using second inequality in (13) yields the desired inequality.

Now, we assume that **h4** holds. Therefore, combining both the second inequalities in (16) and (14), we obtain the first inequality in (16). To establish the first inequality in (17), use the first inequality in (16) and Proposition 2 with $z_k = x_k$ and $\Theta = D^- f'(t_*)/(-2f'(t_*))$. Thus, the proof is concluded. □

Now we are going to prove Theorem 8, but first we need two more additional results.

Proposition 16. *For any $y \in B(\tilde{x}, R)$, the following inequality holds*

$$\left\|T_{\tilde{x}}^{-1}(-F(y))\right\| \le f(\|y - \tilde{x}\|) + 2\|y - \tilde{x}\|.$$

Proof. Let $y \in B(\tilde{x}, R)$. First note that assumption (12) implies that the result holds for $y = \tilde{x}$. Hence assume that $y \neq \tilde{x}$. From definition of the linearization error of F in (34) we have

$$-F(y) = -E(\tilde{x}, y) - F(\tilde{x}) - F'(\tilde{x})(y - \tilde{x}).$$

Above equality together with definition of sublinear mapping in (3) gives us

$$T_{\tilde{x}}^{-1}(-F(y)) \supseteq T_{\tilde{x}}^{-1}(-E(\tilde{x}, y)) + T_{\tilde{x}}^{-1}(-F(\tilde{x})) + T_{\tilde{x}}^{-1}\left(-F'(\tilde{x})(y - \tilde{x})\right).$$

Taking in account properties of the norm, above inclusion implies

$$\left\|T_{\tilde{x}}^{-1}(-F(y))\right\| \le \left\|T_{\tilde{x}}^{-1}(-E(\tilde{x}, y))\right\| + \left\|T_{\tilde{x}}^{-1}(-F(\tilde{x}))\right\| + \left\|T_{\tilde{x}}^{-1}\left(-F'(\tilde{x})(y - \tilde{x})\right)\right\|. \tag{47}$$

Now we are going to bound the three terms in the left hand side of last inequality. Applying Lemma 13 with $x = \tilde{x}$, $t = 0$ and $s = \|y - \tilde{x}\|$ and taking in account that $f'(0) = -1$ we have

$$\left\|T_{\tilde{x}}^{-1}(-E(\tilde{x}, y))\right\| \le f(\|y - \tilde{x}\|) - f(0) + \|y - \tilde{x}\|.$$

Definition of $T_{\tilde{x}}^{-1}$ implies that $-(y - \tilde{x}) \in T_{\tilde{x}}^{-1}\left(-F'(\tilde{x})(y - \tilde{x})\right)$, hence we conclude that

$$\left\|T_{\tilde{x}}^{-1}\left(-F'(\tilde{x})(y - \tilde{x})\right)\right\| \le \|y - \tilde{x}\|.$$

Since assumption (12) implies that second term in (47) is bounded by $f(0)$, thus substituting the two later inequalities into (47) the desired inequality follows. □

Proposition 17. *Let $R > 0$ and $f : [0, R) \to \mathbb{R}$ a continuously differentiable function. Suppose that $\tilde{x} \in \Omega$, f is a majorant function for F at $\tilde{x}$ and satisfies* **h4***. If $0 \le \rho < \beta/2$, where $\beta := \sup\{-f(t) : t \in [0, R)\}$ then for any $\hat{x} \in B(\tilde{x}, \rho)$ the scalar function $g : [0, R-\rho) \to \mathbb{R}$,*

$$g(t) = \frac{-1}{f'(\rho)}[f(t+\rho) + 2\rho],$$

is a majorant function for F at $\hat{x}$ and also satisfies condition **h4***.*

Proof. Since the domain of f is $[0, R)$ and $f'(\rho) \neq 0$ (see Proposition 9 item iv) we conclude that g is well defined. First we will prove that function g satisfies condition **h1**, **h2**, **h3** and **h4**. We trivially have that $g'(0) = -1$. Since f is convex, combining this with **h1** we have $f(t) + t \ge f(0) > 0$, for all $0 \le t < R$, which by using Proposition 9 item iv implies $g(0) = -[f(\rho) + 2\rho]/f'(\rho) > 0$, hence g satisfies **h1**. Using **h2** and Proposition 9 item iv we easily conclude that g also satisfies **h2**. Now, as $\rho < \beta/2$, using Proposition 9 items iii and iv we have

$$\lim_{t \to \bar{t}-\rho} g(t) = \frac{-1}{f'(\rho)}(2\rho - \beta) < 0,$$

which implies that g satisfies **h4** and, as g is continuous and $g(0) > 0$, it also satisfies **h3**.

To complete the proof, it remains to prove that g satisfies (11). First of all, note that for any $\hat{x} \in B(\tilde{x}, \rho)$, from Proposition 9 item iv we have $\|\hat{x} - \tilde{x}\| < \rho < \bar{t}$ and by using Proposition 12 we obtain that

$$\left\|T_{\hat{x}}^{-1}F'(\tilde{x})\right\| \le \frac{-1}{f'(\rho)}. \tag{48}$$

Because $B(\tilde{x}, R) \subseteq \Omega$, for any $\hat{x} \in B(\tilde{x}, \rho)$ we trivially have $B(\hat{x}, R-\rho) \subset \Omega$. Now, take $x, y \in \mathbb{X}$ such that

$$x, y \in B(\hat{x}, R-\rho), \qquad \|x - \hat{x}\| + \|y - x\| < R - \rho.$$

Hence $x, y \in B(\tilde{x}, R)$ and $\|x - \tilde{x}\| + \|y - x\| < R$. Thus, applying the inequality of Lemma 6 with $z = \hat{x}$ and $v = \tilde{x}$, using the property of the norm in (6), (48) and (11) we have

$$\begin{aligned}
\left\|T_{\hat{x}}^{-1}\left[F'(y) - F'(x)\right]\right\| &\le \left\|T_{\hat{x}}^{-1}F'(\tilde{x})T_{\tilde{x}}^{-1}\left[F'(y) - F'(x)\right]\right\| \\
&\le \left\|T_{\hat{x}}^{-1}F'(\tilde{x})\right\| \left\|T_{\tilde{x}}^{-1}\left[F'(y) - F'(x)\right]\right\| \\
&\le \frac{-1}{f'(\rho)}\left[f'\left(\|y - x\| + \|x - \tilde{x}\|\right) - f'(\|x - \tilde{x}\|)\right].
\end{aligned}$$

On the other hand, since f' is convex, the function $s \mapsto f'(t+s) - f'(s)$ is increasing for $t \ge 0$ and as $\|x - \tilde{x}\| \le \|x - \hat{x}\| + \rho$ we conclude that

$$f'\left(\|y - x\| + \|x - \tilde{x}\|\right) - f'\left(\|x - \tilde{x}\|\right) \le f'\left(\|y - x\| + \|x - \hat{x}\| + \rho\right) - f'\left(\|x - \hat{x}\| + \rho\right).$$

Hence, combining the two above inequalities with the definition of the function g we obtain

$$\|T_{\hat{x}}^{-1}\left[F'(y) - F'(x)\right]\| \le g'(\|y - x\| + \|x - \hat{x}\|) - g'(\|x - \hat{x}\|),$$

implying that the function g satisfies (11), which complete the proof of the proposition. □

We are now ready to prove Theorem 8, its proof is obtained by combining Theorem 7 with Propositions 16 and 17.

Proof of Theorem 8. Since $\hat{x} \in B(\tilde{x}, \rho)$, from Proposition 9 item iv we have $\|\hat{x} - \tilde{x}\| < \rho < \bar{t}$ and by using Proposition 12 we obtain that

$$\left\|T_{\hat{x}}^{-1}F'(\tilde{x})\right\| \le \frac{-1}{f'(\rho)}, \tag{49}$$

moreover, the point $\hat{x}$ satisfies Robinson's Condition, namely,

$$\operatorname{rge} T_{\hat{x}} = \mathbb{Y}. \tag{50}$$

Hence, applying inclusion in Lemma 6 with $z = \hat{x}$, $v = \tilde{x}$ and $w = -F(\hat{x})$, using property of the norm in (6), inequality (49) and Proposition 16 with $y = \hat{x}$ we obtain

$$\begin{aligned}
\left\|T_{\hat{x}}^{-1}(-F(\hat{x}))\right\| &\le \left\|T_{\hat{x}}^{-1}F'(\tilde{x})T_{\tilde{x}}^{-1}(-F(\hat{x}))\right\| \\
&\le \left\|T_{\hat{x}}^{-1}F'(\tilde{x})\right\| \left\|T_{\tilde{x}}^{-1}(-F(\hat{x}))\right\| \\
&\le \frac{-1}{f'(\rho)}\left[f\left(\|\hat{x} - \tilde{x}\|\right) + 2\|\hat{x} - \tilde{x}\|\right].
\end{aligned}$$

As $f' \geq -1$, the function $t \mapsto f(t) + 2t$ is (strictly) increasing. Combining this fact with $\|\hat{x} - \tilde{x}\| < \rho$, the above inequality and definition of the function g we conclude that

$$\left\| T_{\hat{x}}^{-1}(-F(\hat{x})) \right\| \leq g(0).$$

Therefore, since (50) implies that $\hat{x}$ satisfies Robinson's Condition and Proposition 17 implies that g is a majorant function for F at $\hat{x}$ satisfying condition **h4**, the last inequality allows us to apply Theorem 7 for F and the majorant function g at point $\hat{x}$ for obtaining the desired result. □

3. Special case

The affine majorant condition is crucial for our analysis. It is worth pointing out that to construct a majorizing function for a given nonlinear function is a very difficult problem and this is not our aim in this moment. On the other hand, 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 the affine invariant Lipschitz-like and Smale's conditions, respectively. In this sense, the results obtained in Theorems 7 and 8 unify the convergence analysis for the classes of inclusion problems involving these functions.

3.1. Convergence result under affine invariant Lipschitz-like condition

In this section, we will present a robust convergence theorem on Newton's method for solving nonlinear inclusion problem under affine invariant Lipschitz-like condition, in particular, the result include as special case the Theorem 2 of [21]. Under the Lipschitz-Like condition, Theorem 7 becomes:

Theorem 18. *Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and $\mathbb{X}$ reflexive, $\Omega \subseteq \mathbb{X}$ an open set and $F : \Omega \to \mathbb{Y}$ a continuously differentiable function and $C \subset \mathbb{Y}$ a nonempty closed convex cone. Take $\tilde{x} \in \Omega$, $L > 0$ and $b > 0$. Suppose that $\tilde{x} \in \Omega$, F satisfies Robson's Condition at $\tilde{x}$,*

$$B(\tilde{x}, R) \subseteq \Omega, \qquad \left\| T_{\tilde{x}}^{-1}\left[F'(y) - F'(x)\right] \right\| \leq L\|y - x\|, \quad \forall x, y \in B(\tilde{x}, R),$$

and $2bL \leq 1$. Moreover, assume that

$$\left\| T_{\tilde{x}}^{-1}(-F(\tilde{x})) \right\| \leqslant b.$$

Define, the scalar function $f : [0, +\infty) \to \mathbb{R}$ as $f(t) := Lt^2/2 - t + b$. Then f has $t_* := (1 - \sqrt{1 - 2bL})/L$ as a smallest zero, the sequences generated by Newton's Method for solving the inclusion $F(x) \in C$ and the equation $f(t) = 0$, with starting point $x_0 = \tilde{x}$ and $t_0 = 0$, respectively,

$$x_{k+1} \in x_k + \operatorname{argmin}\left\{\|d\| : F(x_k) + F'(x_k)d \in C\right\}, \qquad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, \quad k = 0, 1, \ldots,$$

are well defined, $\{x_k\}$ is contained in $B(\tilde{x}, t_*)$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$ and converge to t_* and satisfy the inequalities

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \qquad \|x_{k+1} - x_k\| \leq \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2}\|x_k - x_{k-1}\|^2,$$

for all $k = 0, 1, \ldots,$ and $k = 1, 2, \ldots,$ respectively. Moreover, $\{x_k\}$ converges to $x_* \in B[\tilde{x}, t_*]$ such that $F(x_*) \in C$,

$$\|x_* - x_k\| \leq t_* - t_k, \qquad t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \ldots$$

and, therefore, $\{t_k\}$ converges Q-linearly to t_* and $\{x_k\}$ converges R-linearly to x_*. Additionally, if $2bL < 1$ then the following inequalities hold

$$\|x_{k+1} - x_k\| \leq \frac{L}{2\sqrt{1 - 2bL}}\|x_k - x_{k-1}\|^2, \qquad t_{k+1} - t_k \leq \frac{L}{2\sqrt{1 - 2bL}}(t_k - t_{k-1})^2, \quad k = 0, 1, \ldots,$$

and, as a consequence, $\{x_k\}$ and $\{t_k\}$ converge Q-quadratically to x_* and t_*, respectively, as follow

$$\limsup_{k\to\infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \leq \frac{L}{2\sqrt{1 - 2bL}}, \qquad t_* - t_{k+1} \leq \frac{L}{2\sqrt{1 - 2bL}}, \quad k = 0, 1, \ldots.$$

Proof. It easy to see that $f : [0, +\infty) \to \mathbb{R}$ defined by $f(t) := Lt^2/2 - t + b$ is a majorant function to F in $\tilde{x}$, with the smallest zero equal to t_*. Therefore, the result follows from Theorem 7. □

Under the affine invariant Lipschitz-Like condition, the robust theorem on Newton's method for solving nonlinear inclusion problems, namely, Theorem 8 becomes:

Theorem 19. *Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and X reflexive, $\Omega \subseteq \mathbb{X}$ an open set and $F : \Omega \to \mathbb{Y}$ a continuously differentiable function and $C \subset \mathbb{Y}$ a nonempty closed convex cone. Take $\tilde{x} \in \Omega$, $L > 0$ and $b > 0$. Suppose that $\tilde{x} \in \Omega$, F satisfies Robson's Condition at $\tilde{x}$,*

$$B(\tilde{x}, R) \subseteq \Omega, \qquad \left\| T_{\tilde{x}}^{-1} \left[F'(y) - F'(x)\right] \right\| \le L\|y - x\|, \quad \forall x, y \in B(\tilde{x}, R),$$

and $2bL < 1$. Moreover, assume that

$$\left\| T_{\tilde{x}}^{-1}(-F(\tilde{x})) \right\| \leqslant b.$$

Let $0 \le \rho < (1 - 2Lb)/(4L)$ and the quadratic polynomial $g : [0, +\infty) \to \mathbb{R}$ defined by

$$g(t) := \frac{-1}{L\rho - 1} \left[\frac{L}{2}(t + \rho)^2 - (t + \rho) + b + 2\rho \right].$$

Then the quadratic polynomial g has the smallest zero given by

$$t_{*,\rho} := \left(1 - \rho L - \sqrt{1 - 2L(b - 2\rho)}\right)/L,$$

the sequences generated by Newton's Method for solving the inclusion $F(x) \in C$ and the equation $g(t) = 0$, with starting point $x_0 = \hat{x}$ for any $\hat{x} \in B(\tilde{x}, \rho)$ and $t_0 = 0$, respectively,

$$x_{k+1} \in x_k + \operatorname{argmin}\left\{\|d\| : F(x_k) + F'(x_k)d \in C\right\}, \qquad t_{k+1} = t_k - \frac{g(t_k)}{g'(t_k)}, \quad k = 0, 1, \ldots$$

are well defined, $\{x_k\}$ is contained in $B(\tilde{x}, t_{*,\rho})$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_{*,\rho})$ and converge to $t_{*,\rho}$ and satisfy the inequalities

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \quad k = 0, 1, \ldots,$$
$$\|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2}\|x_k - x_{k-1}\|^2 \le \frac{L}{2\sqrt{1 - 2L(b - 2\rho)}}\|x_k - x_{k-1}\|^2, \quad k = 1, 2, \ldots.$$

Moreover, $\{x_k\}$ converges to $x_* \in B[\tilde{x}, t_{*,\rho}]$ such that $F(x_*) \in C$, satisfies the inequalities

$$\|x_* - x_k\| \le t_{*,\rho} - t_k, \qquad t_{*,\rho} - t_{k+1} \le \frac{1}{2}(t_{*,\rho} - t_k), \quad k = 0, 1, \ldots$$

and the convergence of $\{x_k\}$ and $\{t_k\}$ to x_* and $t_{*,\rho}$, respectively, is Q-quadratic as follows

$$\limsup_{k\to\infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \le \frac{L}{2\sqrt{1 - 2L(b - 2\rho)}}, \qquad t_{*,\rho} - t_{k+1} \le \frac{L}{2\sqrt{1 - 2L(b - 2\rho)}}(t_{*,\rho} - t_k)^2,$$

for $k = 0, 1, \ldots$.

Proof. The proof it follows from Theorem 8 by noting that $f : [0, +\infty) \to \mathbb{R}$ defined by $f(t) := Lt^2/2 - t + b$ is a majorant function to F in $\tilde{x}$ and $\beta = (1 - 2Lb)/(4L)$. □

3.2. Convergence result under affine invariant Smale's condition

In this section, we will present a robust convergence theorem on Newton's method for solving nonlinear inclusion problem under affine invariant Smale's condition. For the degenerated cone, i.e., $C = \{0\}$, this is the corollary of Proposition 3 pp. 195 of [48], see also Proposition 1 pp. 157 and Remark 1 pp. 158 of [1].

First of all we give a condition more easy to check then condition (11), when the functions under consideration are two times continuously differentiable. For state the condition, we need some definitions. Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and $\Omega \subseteq \mathbb{X}$ a open set. The norm of a nth multilinear map $B : \mathbb{X} \times \cdots \times \mathbb{X} \to \mathbb{X}$ is defined by

$$\|B\| := \sup\left\{\|B(v_1, \ldots, v_n)\| : v_1, \ldots, v_n \in \mathbb{X}, \ \|v_i\| = 1, \ i = 1, \ldots, n\right\}.$$

In particular, the norm of the nth derivative of $F : \Omega \to \mathbb{Y}$ at a point $x \in \Omega$ is given by

$$\|F^{(n)}(x)\| = \sup\left\{\|F^{(n)}(x)(v_1, \ldots, v_n)\| : v_1, \ldots, v_n \in \mathbb{X}, \ \|v_i\| = 1, \ i = 1, \ldots, n\right\}.$$

Let $T : \mathbb{X} \rightrightarrows \mathbb{Y}$ be sublinear mapping. The composition $TF^{(n)}(x) : \mathbb{X} \times \cdots \times \mathbb{X} \rightrightarrows \mathbb{Y}$ is defined by $TF^{(n)}(x)(v_1, \ldots, v_n) := T(F^{(n)}(x)(v_1, \ldots, v_n))$. Then, for $F^{(n)}(x)(v_1, \ldots, v_n) \in \operatorname{dom} T$ there hold:

$$\|TF^{(n)}(x)\| = \sup\left\{\left\|T(F^{(n)}(x)(v_1, \ldots, v_n))\right\| : v_1, \ldots, v_n \in \mathbb{X}, \ \|v_i\| = 1, \ i = 1, \ldots, n\right\},$$

where $\|T(F^{(n)}(x)(v_1, \ldots, v_n))\| := \inf\{\|u\| : u \in T(F^{(n)}(x)(v_1, \ldots, v_n))\}$.

Lemma 20. *Let $\mathbb{X}$ and $\mathbb{Y}$ be a Banach spaces such that $\mathbb{X}$ is reflexive, $\Omega \subseteq \mathbb{X}$ and $F : \Omega \to \mathbb{Y}$ a continuous function, two times continuously differentiable on $int(\Omega)$. Suppose that $\tilde{x} \in \Omega$ and $\operatorname{rge} T_{\tilde{x}} = \mathbb{Y}$. If there exists $f : [0, R) \to \mathbb{R}$ twice continuously differentiable such that*

$$\|T_{\tilde{x}}^{-1} F''(x)\| \leqslant f''(\|x - \tilde{x}\|), \tag{51}$$

for all $x \in \Omega$ such that $\|x - \tilde{x}\| < R$. Then F and f satisfy (11).

Proof. Take $x, y \in B(\tilde{x}, R)$ such that $\|x - \tilde{x}\| + \|y - x\| < R$. As the ball is convex $x + \tau(y - x) \in B(\tilde{x}, R)$, for all $\tau \in [0, 1]$. Since, by assumption, $\operatorname{rge} T_{\tilde{x}} = \mathbb{Y}$ we obtain that $\operatorname{dom} T_{\tilde{x}}^{-1} = \mathbb{Y}$, hence properties of the norm and assumption (51) imply that

$$\left\| T_{\tilde{x}}^{-1} \left(F''(x + \tau(y - x))(y - x) \right) \right\| \le f''(\|(x - \tilde{x}) + \tau(y - x)\|)\|y - x\|, \quad \forall\, \tau \in [0, 1].$$

Thus, as $\operatorname{dom} T_{\tilde{x}}^{-1} = \mathbb{Y}$ and $\mathbb{X}$ is reflexive, we apply Lemma 5 with $U = T_{\tilde{x}}^{-1}$, $G(\tau)$ equal to the expression in parentheses on the left hand side of last inequality and $g(\tau)$ equal to the expression on the right hand side of that inequality, obtaining

$$\left\| T_{\tilde{x}}^{-1} \int_0^1 F''(x + \tau(y - x))(y - x) d\tau \right\| \leqslant \int_0^1 f''(\|(x - \tilde{x}) + \tau(y - x)\|)\|y - x\| d\tau,$$

which, after performing the above integrations yields the desired result. □

Theorem 21. *Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and $\mathbb{X}$ reflexive, $\Omega \subseteq \mathbb{X}$ an open set and $F : \Omega \to \mathbb{Y}$ a analytic function and $C \subset \mathbb{Y}$ a nonempty closed convex cone. Suppose that $\tilde{x} \in \Omega$, F satisfies Robson's Condition at $\tilde{x}$,*

$$\gamma := \sup_{k>1} \left\| \frac{T_{\tilde{x}}^{-1} F^{(k)}(\tilde{x})}{k!} \right\|^{\frac{1}{k-1}} < +\infty,$$

$B(\tilde{x}, 1/\gamma) \subseteq \Omega$, there exists $b > 0$ such that

$$\left\| T_{\tilde{x}}^{-1}(-F(\tilde{x})) \right\| \leqslant b,$$

and $\alpha := b\gamma \leqslant 3 - 2\sqrt{2}$. Define, the analytic function $f : [0, 1/\gamma) \to \mathbb{R}$ as $f(t) := t/(1 - \gamma t) - 2t + b$. Then f has

$$t_* := \left(\alpha + 1 - \sqrt{(\alpha + 1)^2 - 8\alpha}\right)/(4\gamma),$$

as a smallest zero, the sequences generated by Newton's Method for solving the analytics inclusion $F(x) \in C$ and the equation $f(t) = 0$, with starting point $x_0 = \tilde{x}$ and $t_0 = 0$, respectively,

$$x_{k+1} \in x_k + \operatorname{argmin}\left\{\|d\| : F(x_k) + F'(x_k)d \in C\right\}, \qquad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, \quad k = 0, 1, \ldots,$$

are well defined, $\{x_k\}$ is contained in $B(\tilde{x}, t_*)$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$ and converges to t_* and satisfies the inequalities

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \qquad \|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2}\|x_k - x_{k-1}\|^2,$$

for all $k = 0, 1, \ldots,$ and $k = 1, 2, \ldots,$ respectively. Moreover, $\{x_k\}$ converges to $x_* \in B[\tilde{x}, t_*]$ such that $F(x_*) \in C$,

$$\|x_* - x_k\| \le t_* - t_k, \qquad t_* - t_{k+1} \le \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \ldots$$

and, therefore, $\{t_k\}$ converges Q-linearly to t_* and $\{x_k\}$ converge R-linearly to x_*. Additionally, if $\alpha < 3 - 2\sqrt{2}$ then the following inequalities hold

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

and, as a consequence, $\{x_k\}$ and $\{t_k\}$ converge Q-quadratically to x_* and t_*, respectively, as follows

$$\limsup_{k\to\infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \le \frac{\gamma}{(1 - \gamma t_*)[2(1 - \gamma t_*)^2 - 1]},$$
$$t_* - t_{k+1} \le \frac{\gamma}{(1 - \gamma t_*)[2(1 - \gamma t_*)^2 - 1]}(t_* - t_k), \quad k = 0, 1, \ldots.$$

Proof. Use Lemma 20 to prove that $f : [0, 1/\gamma) \to \mathbb{R}$ defined by $f(t) = t/(1-\gamma t) - 2t + b$, is a majorant function to F in $\tilde{x}$, with root equal to t_*, see [42]. Therefore, the result follows from Theorem 7. □

Under the affine invariant Smale's Condition, the robust theorem on Newton's method for solving nonlinear inclusion, namely, Theorem 8 becomes:

Theorem 22. *Let $\mathbb{X}$, $\mathbb{Y}$ be Banach spaces and $\mathbb{X}$ reflexive, $\Omega \subseteq \mathbb{X}$ an open set and $F : \Omega \to \mathbb{Y}$ a analytic function and $C \subset \mathbb{Y}$ a nonempty closed convex cone. Suppose that $\tilde{x} \in \Omega$, F satisfies Robson's Condition at $\tilde{x}$,*

$$\gamma := \sup_{k>1} \left\| \frac{T_{\tilde{x}}^{-1} F^{(k)}(\tilde{x})}{k!} \right\|^{\frac{1}{k-1}} < +\infty,$$

$B(\tilde{x}, 1/\gamma) \subseteq \Omega$, there exists $b > 0$ such that

$$\left\| T_{\tilde{x}}^{-1}(-F(\tilde{x})) \right\| \leqslant b,$$

and $\alpha := b\gamma < 3 - 2\sqrt{2}$. Let $0 \le \rho < [\sqrt{2}(3-\alpha) - 3]/(2\gamma\sqrt{2})$ and define $g : [0, 1/\gamma - \rho) \to \mathbb{R}$ as

$$g(t) := (t+\rho)/(1-\gamma(t+\rho)) - 2(t+\rho) + b + 2\rho.$$

Then the scalar analytic function g has a smallest zero given by

$$t_{*,\rho} := \left(\alpha + 1 - 2\rho\gamma - \sqrt{(\alpha+1-2\rho\gamma)^2 - 8\alpha - 8\rho\gamma(1-\alpha)}\right)/(4\gamma),$$

the sequences generated by Newton's Method for solving the analytics inclusion $F(x) \in C$ and equation $g(t) = 0$, with starting point $x_0 = \hat{x}$ for any $\hat{x} \in B(\tilde{x}, \rho)$ and $t_0 = 0$, respectively,

$$x_{k+1} \in x_k + \operatorname{argmin}\left\{\|d\| : F(x_k) + F'(x_k)d \in C\right\}, \qquad t_{k+1} = t_k - \frac{g(t_k)}{g'(t_k)}, \quad k = 0, 1, \ldots \tag{52}$$

are well defined, $\{x_k\}$ is contained in $B(\tilde{x}, t_{*,\rho})$, $\{t_k\}$ is strictly increasing, is contained in $[0, t_{*,\rho})$ and converges to $t_{*,\rho}$ and satisfies the inequalities

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \quad k = 0, 1, \ldots,$$
$$\|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2}\|x_k - x_{k-1}\|^2 \le \frac{\gamma}{(1-\gamma(t_{*,\rho}+\rho))[2(1-\gamma(t_{*,\rho}+\rho))^2 - 1]}\|x_k - x_{k-1}\|^2,$$

for $k = 1, 2, \ldots$. Moreover, $\{x_k\}$ converges to $x_* \in B[\tilde{x}, t_{*,\rho}]$ such that $F(x_*) \in C$, satisfies the inequalities

$$\|x_* - x_k\| \le t_{*,\rho} - t_k, \qquad t_{*,\rho} - t_{k+1} \le \frac{1}{2}(t_{*,\rho} - t_k), \quad k = 0, 1, \ldots$$

and the convergence of $\{x_k\}$ and $\{t_k\}$ to x_* and $t_{*,\rho}$, respectively, are Q-quadratic as follows

$$\limsup_{k\to\infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \le \frac{\gamma}{(1-\gamma(t_{*,\rho}+\rho))[2(1-\gamma(t_{*,\rho}+\rho))^2 - 1]},$$
$$t_{*,\rho} - t_{k+1} \le \frac{\gamma}{(1-\gamma(t_{*,\rho}+\rho))[2(1-\gamma(t_{*,\rho}+\rho))^2 - 1]}(t_{*,\rho} - t_k)^2, \quad k = 0, 1, \ldots.$$

Proof. Use Lemma 20 to prove that $f : [0, 1/\gamma) \to \mathbb{R}$ defined by $f(t) = t/(1-\gamma t) - 2t + b$, is a majorant function to F in $\tilde{x}$ (see [42]). The proof it follows from Theorem 8 by noting that f is a majorant function to F in $\tilde{x}$ and $\beta = [\sqrt{2}(3-\alpha) - 3]/(2\gamma\sqrt{2})$. □

4. Final remarks

Let us finally make a few brief comments on computational aspects of Newton's method for solving the nonlinear inclusion (9). Note that the first equality in (13) implies

$$x_{k+1} = x_k + d_k, \qquad F(x_k) + F'(x_k)d_k \in C, \quad k = 0, 1, \ldots$$

and for the degenerated cone $C = \{0\}$ the above iteration formally becomes Newton's iteration for solving nonlinear equation $F(x) = 0$, that is

$$x_{k+1} = x_k + d_k, \qquad F(x_k) + F'(x_k)d_k = 0, \quad k = 0, 1, \ldots. \tag{53}$$

Since the solution of the linear equation (53) is computationally expensive, namely, at each iteration the derivative at x_k must be computed and stored. To circumvent drawbacks like this, Dembo, Eisenstat and Steihaug introduced in [49] the inexact Newton's method for solving nonlinear equation $F(x) = 0$. The inexact Newton's methods for solving nonlinear equation $F(x) = 0$ is any method which, given an initial point x_0, generates a sequence $\{x_k\}$ as follows:

$$x_{k+1} = x_k + d_k, \qquad \|F(x_k) + F'(x_k)d_k\| \leq \theta_k \|F(x_k)\|, \quad k = 0, 1, \ldots,$$

for a suitable forcing sequence $\{\theta_k\}$, which is used to control the level of accuracy. Hence, solutions of practical problems are obtained by computational implementations of the inexact Newton's methods. Therefore, we extend the inexact Newton's methods for solving nonlinear inclusion as any method which, given an initial point x_0, generates a sequence $\{x_k\}$ as follows:

$$x_{k+1} = x_k + d_k, \qquad d(0, F(x_k) + F'(x_k)d_k - C) \leq \theta_k d(0, F(x_k) - C), \quad k = 0, 1, \ldots,$$

for a suitable forcing sequence $\{\theta_k\}$, where $d(x, D)$ denotes the distance from a point $x \in \mathbb{X}$ to the subset $D \subset \mathbb{Y}$; that is $d(x, D) = \inf\{\|x - x'\| : x' \in D\}$. The analysis of these methods under Lipschitz-like and majorant conditions will be done in the future.

Acknowledgment

This work was done while the author were visiting The Institute of Computational Mathematics and Scientific/ Engineering Computing, Academy of Mathematics and Systems Science at Chinese Academy of Science. He is grateful to host institution for the congenial scientific atmosphere that it has provided during his visit. The author was supported in part by CAPES (Projeto 019/2011- Cooperação Internacional Brasil–China), CNPq Grant 471815/2012-8, CNPq Grant 303732/2011-3, PRONEX-Optimization(FAPERJ/CNPq) and FAPEG/GO.

References

[1] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer-Verlag, New York, 1997.
[2] J.E. Dennis Jr., R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, in: Classics in Applied Mathematics, SIAM, Philadelphia, 1996.
[3] J.W. Daniel, Newton's method for nonlinear inequalities, Numer. Math. 21 (1973) 381–387.
[4] P. Deuflhard, Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms, in: Series Computational Mathematics, vol. 35, Springer, Heidelberg, 2004.
[5] A.L. Dontchev, R.T. Rockafellar, Implicit Functions and Solution Mappings. A View From Variational Analysis, in: Springer Monographs in Mathematics, Springer, Dordrecht, 2009.
[6] Y. He, J. Sun, Error bounds for degenerate cone inclusion problems, Math. Oper. Res. 30 (3) (2005) 701–717.
[7] S.G. Krantz, H.R. Parks, The Implicit Function Theorem: History, Theory, and Applications, Birkhäuser, Boston, 2002.
[8] C. Li, K.F. Ng, Convergence analysis of the Gauss–Newton method for convex inclusion and convex-composite optimization problems, J. Math. Anal. Appl. 389 (1) (2012) 469–485.
[9] B.N. Psenicnyi, A Newton method for the solution of systems of equalities and inequalities, Mat. Zametki 8 (1970) 635–640.
[10] K.W. Chau, Application of a PSO-based neural network in analysis of outcomes of construction claims, Autom. Constr. 16 (5) (2007) 642–646.
[11] C.T. Cheng, K.W. Chau, J. Sun, Long-term prediction of discharges in Manwan Reservoir using artificial neural network models, in: Lecture Notes in Computer Science, vol. 3498, 2005, pp. 1040–1045.
[12] Zhi-Kai Huang, Kwok-Wing Chau, A new image thresholding method based on Gaussian mixture model, Appl. Math. Comput. 205 (2) (2008) 899–907.
[13] R. Taormina, K.W. Chau, R. Sethi, Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon, Eng. Appl. Artif. Intell. 25 (8) (2012) 1670–1676.
[14] C.L. Wu, K.W. Chau, Y.S. Li, Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques, Water Resour. Res. 45 (8) (2009), http://dx.doi.org/10.1029/2007WR006737.
[15] J. Zhang, K.W. Chau, Multilayer ensemble pruning via novel multi-sub-swarm particle swarm optimization, J. UCS 15 (4) (2009) 840–858.
[16] J.F. Bonnans, Local analysis of Newton-type methods for variational inequalities and nonlinear programming, Appl. Math. Optim. 29 (2) (1994) 161–186.
[17] A.L. Dontchev, R.T. Rockafellar, Newton's method for generalized equations: a sequential implicit function theorem, Math. Program. Ser. B 123 (1) (2010) 139–159.
[18] J. Moser, A new techniques for the construction of solutions of nonlinear differential equations, Proc. Natl. Acad. Sci. USA 47 (1961) 1824–1831.
[19] J. Nash, The embedding problem for Riemannian manifolds, Ann. of Math. 63 (1956) 20–63.
[20] D. Sun, A regularization Newton method for solving nonlinear complementarity problems, Appl. Math. Optim. 40 (3) (1999) 315–339.
[21] S.M. Robinson, Extension of Newton's method to nonlinear functions with values in a cone, Numer. Math. 19 (1972) 341–347.
[22] R. Adler, J.-P. Dedieu, J. Margulies, M. Martens, M. Shub, Newton method on Riemannian manifolds and a geometric model for the human spine, IMA J. Numer. Anal. 22 (2002) 1–32.
[23] F. Alvarez, J. Botle, J. Munier, A unifying local convergence result for Newton's method in Riemannian manifolds, Found. Comput. Math. 8 (2) (2008) 197–226.
[24] J.-P. Dedieu, P. Priouret, G. Malajovich, Newton's method on Riemannian manifolds: covariant alpha theory, IMA J. Numer. Anal. 23 (2003) 395–419.
[25] C. Li, J.H. Wang, Kantorovich's theorems of Newton's method for mappings and optimization problems on Lie groups, IMA J. Numer. Anal. 31 (1) (2011) 322–347.
[26] C. Li, J.H. Wang, Newton's method for sections on Riemannian manifolds: generalized covariant α-theory, J. Complexity 24 (3) (2008) 423–451.
[27] C. Li, J.H. Wang, Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition, IMA J. Numer. Anal. 26 (2) (2006) 228–251.
[28] J.-H. Wang, S. Huang, C. Li, Extended Newton's method for mappings on Riemannian manifolds with values in a cone, Taiwanese J. Math. 13 (2B) (2009) 633–656.
[29] C. Li, K.F. Ng, Convergence analysis of the Gauss–Newton method for convex inclusion problems and convex composite optimization, Preprint, 2013.
[30] J. Chen, W. Li, Convergence behaviour of inexact Newton methods under weak Lipschitz condition, J. Comput. Appl. Math. 191 (1) (2006) 143–164.
[31] O.P. Ferreira, Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle, IMA J. Numer. Anal. 29 (3) (2009) 746–759.
[32] O.P. Ferreira, B.F. Svaiter, Kantorovich's majorants principle for Newton's method, Comput. Optim. Appl. 42 (2009) 213–229.

[33] Z. Huang, The convergence ball of Newton's method and the uniqueness ball of equations's under Hölder-type continuos derivatives, Comput. Math. Appl. 47 (2004) 247–251.
[34] C. Li, W.P. Shen, Local convergence of inexact methods under the Holder condition, J. Comput. Appl. Math. 222 (2008) 544–560.
[35] P.D. Proinov, General local convergence theory for a class of iterative processes and its applications to Newton's method, J. Complexity 25 (1) (2009) 8–62.
[36] X. Wang, Convergence of Newton's method and uniqueness of the solution of equation in Banach space, IMA J. Numer. Anal. 20 (2000) 123–134.
[37] O.P. Ferreira, Local convergence of Newton's method under majorant condition, J. Comput. Appl. Math. 235 (2011) 1515–1522.
[38] O.P. Ferreira, M.L.N. Gonçalves, P.R. Oliveira, Convergence of the Gauss–Newton method for convex composite optimization under a majorant condition, SIAM J. Optim. 23 (3) (2013) 1757–1783.
[39] O.P. Ferreira, B.F. Svaiter, A robust Kantorovich's theorem on the inexact Newton method with relative residual error tolerance, J. Complexity 28 (3) (2012) 346–363.
[40] O.P. Ferreira, B.F. Svaiter, Kantorovich's theorem on Newton's method in Riemannian manifolds, J. Complexity 18 (1) (2002) 304–329.
[41] C. Li, J.-H. Wang, J.-P. Dedieu, Smale's point estimate theory for Newton's method on Lie groups, J. Complexity 25 (2) (2009) 128–151.
[42] X. Wang, Convergence of Newton's method and inverse function theorem in Banach space, Math. Comp. 68 (225) (1999) 169–186.
[43] P. Deuflhard, G. Heindl, Affine invariant convergence for Newton's method and extensions to related methods, SIAM J. Numer. Anal. 16 (1) (1979) 1–10.
[44] C. Li, K.F. Ng, Majorizing functions and convergence of the Gauss–Newton method for convex composite optimization, SIAM J. Optim. 18 (2007) 613–642.
[45] S.M. Robinson, Normed convex processes, Trans. Amer. Math. Soc. 174 (1972) 127–140.
[46] R.T. Rockafellar, Monotone Processes of Convex and Concave Type, in: Memoirs of the American Mathematical Society, vol. 77, American Mathematical Society, Providence, RI, 1967.
[47] R.T. Rockafellar, Convex Analysis, in: Princeton Mathematical Series, vol. 28, Princeton University Press, Princeton, NJ, 1970.
[48] S. Smale, Newton method estimates from data at one point, in: R. Ewing, K. Gross, C. Martin (Eds.), The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, Springer-Verlag, New York, 1986, pp. 185–196.
[49] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (1982) 400–408.