

CONVERGENCE OF THE GAUSS–NEWTON METHOD FOR CONVEX COMPOSITE OPTIMIZATION UNDER A MAJORANT CONDITION*

O. P. FERREIRA[†], M. L. N. GONÇALVES[†], AND P. R. OLIVEIRA[‡]

Abstract. Under the hypothesis that an initial point is a quasi-regular point, we use a majorant condition to present a new semilocal convergence analysis of an extension of the Gauss–Newton method for solving convex composite optimization problems. In this analysis the conditions and proof of convergence are simplified by using a simple majorant condition to define regions where a Gauss–Newton sequence is well behaved.

Key words. convex composite optimization problem, Gauss–Newton methods, majorant condition, semilocal convergence

AMS subject classifications. 49M15, 65H10, 90C53, 90C30

DOI. 10.1137/110841606

1. Introduction. Consider the convex composite optimization problem

$$(1.1) \quad \min h(F(x)),$$

where $h : \mathbb{R}^m \rightarrow \mathbb{R}$ is a real-valued convex and $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuously differentiable. As is well known (see [1, 8, 9, 10] and references therein), a wide variety of applications with this formulation can be found in the mathematical programming literature, e.g., nonlinear inclusions, penalization methods, minimax, and goal programming.

The basic algorithm considered in [1, 8, 10], which is an extension of the Gauss–Newton method for solving the nonlinear least squares problem, will be considered in this paper. The study of (1.1) is related to the convex inclusion problem

$$(1.2) \quad F(x) \in C = \operatorname{argmin} h,$$

because if $x_* \in \mathbb{R}^n$ satisfies the convex inclusion (1.2), then x_* is a solution of (1.1), but if $x_* \in \mathbb{R}^n$ is a solution of (1.1) it does not necessarily satisfy the inclusion convex (1.2). Although a priori our goal is to give criteria that ensure the convergence of the sequence generated by the Gauss–Newton algorithm for a solution of (1.1), we will give criteria that ensure the convergence of that sequence for some $x_* \in \mathbb{R}^n$ satisfying $F(x_*) \in C$ which, in particular, solves (1.1).

In this paper, we are interested in the semilocal convergence analysis, i.e., based on the information at an initial point, criteria are given that ensure the convergence of the sequence generated by the Gauss–Newton algorithm for some $x_* \in \mathbb{R}^n$ with

*Received by the editors July 20, 2011; accepted for publication (in revised form) June 18, 2013; published electronically September 4, 2013.

<http://www.siam.org/journals/siopt/23-3/84160.html>

[†]IME/UFG, Campus II, Caixa Postal 131, CEP 74001-970, Goiânia, GO, Brazil (orizon@mat.ufg.br, maxlmg@mat.ufg.br). The first author was supported in part by CAPES, CNPq grant 471815/2012-8, CNPq grant 303732/2011-3, PRONEX–Optimization (FAPERJ/CNPq), and FAPEG/GO. The second author was supported in part by CAPES, FAPEG/GO, and CNPq grant 471815/2012-8.

[‡]COPPE-Sistemas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21945-970, Brazil (poliveir@cos.ufrj.br). This author was supported in part by CNPq.

$F(x_*) \in C$. Under the hypothesis that the initial point is a quasi-regular point of the inclusion (1.2), we use a majorant condition similar to the one used in [3, 4, 6] to present a new semilocal convergence analysis for the sequence generated by the Gauss–Newton algorithm. The convergence analysis presented here communicates the conditions and proof in quite a simple manner. This is possible thanks to our majorant condition and a demonstration technique in which instead of only looking to the generated sequence, we identify regions where the Gauss–Newton sequence (for the convex composite optimization problem) is well behaved, as compared with the Newton method applied to an auxiliary function associated with the majorant function. This technique was introduced in [5].

The convergence of the sequence generated by the Gauss–Newton algorithm was also studied in [1, 8, 10]. Among these, the criterion introduced by Li and Ng in [8] is the best. Besides the technique used in the demonstration, the main difference from our analysis regarding [8] is that they used Wang’s condition, introduced in [15], in place of our majorant condition. But the formulation using the majorant condition provides a clear relationship between the majorant function and the nonlinear function F under consideration. Besides this, the majorant condition simplifies the proof of convergence.

The organization of our paper is as follows. In section 1.1, we list some notation and two basic results. The Gauss–Newton algorithm is discussed in section 2, in section 2.1 we present an analysis of the majorant and auxiliary functions, and some regularity properties are established in section 2.2. In section 3 the main result is stated, and in section 3.1 it is proved. Some applications of this result are given in section 4.

1.1. Notation and auxiliary result. The following notation and result are used throughout our presentation. Let \mathbb{R}^n be with a norm $\|\cdot\|$. The open and closed ball in \mathbb{R}^n with center x and radius r are denoted, respectively, by $B(x, r)$ and $B[x, r]$. The polar of a closed convex $W \subset \mathbb{R}^n$ is the set $W^\circ := \{z \in \mathbb{R}^n : \langle z, w \rangle \leq 0, w \in W\}$. The distance from a point x to a set $W \subset \mathbb{R}^n$ is given by $d(x, W) := \inf\{\|x - w\| : w \in W\}$. The set of all subsets of \mathbb{R}^n is denoted by $P(\mathbb{R}^n)$ and $\text{Ker}(A)$ represents the kernel of the linear map A . For $S \subset \mathbb{R}^n$ a vector subspace, the number $\dim(S)$ denotes its dimension. If $v \in \mathbb{R}^n$, then $v^\perp = \{u \in \mathbb{R}^n : \langle u, v \rangle = 0\}$. The sum of a point $x \in \mathbb{R}^n$ with a set $X \in P(\mathbb{R}^n)$ is the set given by $y + X = \{y + x : x \in X\}$. We denote by \mathbb{M}_n the vectorial space of the $n \times n$ matrix with the Frobenius norm. Finally, $\mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R}^m)$ is the set of ℓ -times continuous differentiable functions from \mathbb{R}^n to \mathbb{R}^m and, in the case $m = n$ we use the short notation $\mathcal{C}^\ell(\mathbb{R}^n)$.

The following auxiliary results will be needed.

PROPOSITION 1.1. *Let $I \subset \mathbb{R}$ be an interval and $\varphi : I \rightarrow \mathbb{R}$ be convex and differentiable.*

1. *For any $u_0 \in \text{int}(I)$, the application*

$$u \mapsto \frac{\varphi(u_0) - \varphi(u)}{u_0 - u}, \quad u \in I, u \neq u_0,$$

is increasing and there exist (in \mathbb{R})

$$\varphi'(u_0) = \lim_{u \rightarrow u_0^-} \frac{\varphi(u_0) - \varphi(u)}{u_0 - u} = \sup_{u < u_0} \frac{\varphi(u_0) - \varphi(u)}{u_0 - u}.$$

2. *If $u, v, w \in I$, $u < w$, and $u \leq v \leq w$, then*

$$\varphi(v) - \varphi(u) \leq [\varphi(w) - \varphi(u)] \frac{v - u}{w - u}.$$

Proof. See Theorem 4.1.1 on p. 21 of [7]. \square

The proof of the next result is based on [1].

PROPOSITION 1.2. *Let $\{z_k\}$ be a sequence in \mathbb{R}^n and $\Theta > 0$. If $\{z_k\}$ converges to z_* and satisfies*

$$(1.3) \quad \|z_{k+1} - z_k\| \leq \Theta \|z_k - z_{k-1}\|^2, \quad k = 1, 2, \dots,$$

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

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

Proof. For simplicity, let $\epsilon_k = \Theta \|z_{k+1} - z_k\|$ for $k = 0, 1, \dots$. As $\{z_k\}$ converges, for large k , there holds $\epsilon_k < 1/2$. Hence, using the assumption in (1.3), for large k we obtain

$$\|z_{k+j+1} - z_{k+j}\| \leq \Theta^{2^j-1} \|z_{k+1} - z_k\|^{2^j} \leq \epsilon_k^j \|z_{k+1} - z_k\|, \quad j = 1, 2, \dots$$

Thus, after simple algebraic manipulation together with the last inequality it easy to see that

$$(1.4) \quad \|z_{k+i} - z_{k+1}\| \leq \sum_{j=1}^{i-1} \|z_{k+j+1} - z_{k+j}\| \leq \sum_{j=1}^{\infty} \|z_{k+j+1} - z_{k+j}\| = \frac{\epsilon_k}{1 - \epsilon_k} \|z_{k+1} - z_k\|$$

for large k and $i = 2, 3, \dots$. On the other hand, for all $i, k = 1, 2, \dots$ the triangular inequality implies

$$\|z_{k+i} - z_k\| \geq \|z_{k+1} - z_k\| - \|z_{k+i} - z_{k+1}\|.$$

Since, $\epsilon_k < 1/2$ for large k , combining the last two inequalities we have

$$\|z_{k+i} - z_k\|^2 \geq \frac{(1 - 2\epsilon_k)^2}{(1 - \epsilon_k)^2} \|z_{k+i} - z_k\|^2, \quad i = 2, 3, \dots,$$

which together with inequality (1.4) yields

$$\frac{\|z_{k+1} - z_{k+i}\|}{\|z_k - z_{k+i}\|^2} \leq \Theta \frac{1 - \epsilon_k}{(1 - 2\epsilon_k)^2}, \quad i = 2, 3, \dots$$

Therefore, as $\{z_k\}$ converges to z_* and $\lim_{k \rightarrow \infty} \epsilon_k = 0$, the desired result follows from the last inequality by letting i go to ∞ and then taking the limsup as k goes to ∞ . \square

2. Preliminary. In this section we present the algorithm to solve problem (1.1), an analysis of our majorant function, and a brief study of regularity. The results of this section are the main tools used in the proof of convergence of the sequence generated by the Gauss-Newton algorithm.

In order to state the Gauss-Newton algorithm for solving problem (1.1), we need the following definition: For $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$, $\Delta \in (0, +\infty]$, and $x \in \mathbb{R}^n$ define

$$(2.1) \quad D_\Delta(x) := \operatorname{argmin} \{h(F(x) + F'(x)d) : d \in \mathbb{R}^n, \|d\| \leq \Delta\},$$

that is, $D_\Delta(x)$ is the solution set for the following problem:

$$(2.2) \quad \min \{h(F(x) + F'(x)d) : d \in \mathbb{R}^n, \|d\| \leq \Delta\}.$$

Given that $\Delta \in (0, +\infty]$, $\eta \in [1, +\infty)$ and a point $x_0 \in \mathbb{R}^n$, the Gauss–Newton algorithm associated with (Δ, η, x_0) as defined in [1] (see also, [8, 9, 10]) is as follows.

ALGORITHM 2.1.

INITIALIZATION. Take $\Delta \in (0, +\infty]$, $\eta \in [1, +\infty)$ and $x_0 \in \mathbb{R}^n$. Set $k = 0$.

STOP CRITERION. Compute $D_\Delta(x_k)$. If $0 \in D_\Delta(x_k)$, STOP. Otherwise.

ITERATIVE STEP. Compute d_k satisfying

$$d_k \in D_\Delta(x_k), \quad \|d_k\| \leq \eta d(0, D_\Delta(x_k)),$$

and set

$$x_{k+1} = x_k + d_k,$$

$k = k + 1$ and GO TO STOP CRITERION.

Note that since (2.2) is a convex optimization problem in a compact set, it follows that the set $D_\Delta(x)$ is nonempty for all $x \in \mathbb{R}^n$. Therefore, the sequence $\{x_k\}$ generated by Algorithm 2.1 is well defined.

2.1. The majorant condition. In this section, we define the majorant condition for the nonlinear function F , which relaxes the assumption of Lipschitz continuity to F' , used in our analysis. We present an analysis of the behavior of the majorant function and of a certain associated auxiliary function; more details about the majorant condition can be found in [3, 4, 6].

DEFINITION 2.2. Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$, $R > 0$, and $x_0 \in \mathbb{R}^n$. A twice-differentiable function $f : [0, R) \rightarrow \mathbb{R}$ is a majorant function for the function F on $B(x_0, R)$ if it satisfies

$$(2.3) \quad \|F'(y) - F'(x)\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|)$$

for any $x, y \in B(x_0, R)$, $\|x - x_0\| + \|y - x\| < R$, and moreover,

- (h1) $f(0) = 0$, $f'(0) = -1$;
- (h2) f' is convex and strictly increasing.

Before presenting some examples of majorant functions, we will give a result which bounds the linearization error of the function F by the error in the linearization on the majorant function.

LEMMA 2.3. Take $x, y \in B(x_0, R)$ and $0 \leq t < v < R$. If $\|x - x_0\| \leq t$ and $\|y - x\| \leq v - t$, then

$$\|F(y) - [F(x) + F'(x)(y - x)]\| \leq f(v) - [f(t) + f'(t)(v - t)] \left(\frac{\|y - x\|}{v - t} \right)^2.$$

Proof. The proof follows the same pattern as the proof of Lemma 7 in [6]. \square

2.1.1. Examples. In this section we present some class of functions with the associated majorant function.

Example 2.4. Let $x_0 \in \mathbb{R}^n$, $R > 0$, and $K > 0$. Consider the following class of Lipschitz functions:

$$\mathcal{L} := \{F \in \mathcal{C}^1(\mathbb{R}^n) : \|F'(y) - F'(x)\| \leq K\|y - x\|, \quad x, y \in B(x_0, R)\}.$$

It is immediate to see that the function $f : [0, R) \rightarrow \mathbb{R}$ given by $f(t) = (K/2)t^2 - t$ is a majorant function for all $F \in \mathcal{L}$ on $B(x_0, R)$.

The next result gives a condition that is easier to check than condition (2.3) whenever the functions under consideration are twice-continuously differentiable.

LEMMA 2.5. *Let $F \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^m)$, $x_0 \in \mathbb{R}^n$, and $R > 0$. If there exists a $f : [0, R] \rightarrow \mathbb{R}$ twice-continuously differentiable and satisfying*

$$\|F''(x)\| \leq f''(\|x - x_0\|), \quad x \in B(x_0, R),$$

then F and f satisfy (2.3).

Proof. The proof follows the same pattern as the proof of Lemma 22 in [3]. \square

Now, we use the last lemma to give two classes of functions satisfying Definition 2.2.

Example 2.6. Let $R > 0$. Consider the following class of functions:

$$\mathcal{F}_1 = \{F \in \mathcal{C}^2(\mathbb{R}^n) : \|F''(x)\| \leq \|x\|^{2/3}, x \in B(0, R)\}.$$

Let $f : [0, R] \rightarrow \mathbb{R}$ be given by

$$f(t) = \frac{9}{40} t^{8/3} - t.$$

It is easy to see that $f(0) = 0$, $f'(t) = (3/5)t^{5/3} - 1$, $f'(0) = -1$, and $f''(t) = t^{2/3} > 0$, hence f satisfies (h1) and (h2) in Definition 2.2. Now, as $f''(t) = t^{2/3}$, using Lemma 2.5 it follows that f is a majorant function for all $F \in \mathcal{F}_1$ on $B(0, R)$. Moreover, the set \mathcal{F}_1 is nonempty, that is, the function $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by

$$F(x) = \frac{9}{50} \left(\|x\|^{5/3}x - \bar{x} \right),$$

where $\bar{x} \in \mathbb{R}^n$, is in \mathcal{F}_1 . Indeed, note that the second derivative of F is given by

$$F''(x)(v, v) = \frac{9}{50} \left[-\frac{5}{9} \|x\|^{-7/3} \langle x, v \rangle^2 x + \frac{5}{3} \|x\|^{-1/3} \|v\|^2 x + \frac{10}{3} \|x\|^{-1/3} \langle x, v \rangle v \right]$$

for all $x, v \in \mathbb{R}^n, x \neq 0$, and $F''(0) = 0$. Hence, from the last inequality we obtain

$$\|F''(x)\| \leq \|x\|^{2/3}, \quad x \in \mathbb{R}^n.$$

Therefore, the statement is proved.

Example 2.7. Consider the class of functions

$$\mathcal{F}_2 = \left\{ F \in \mathcal{C}^2(\mathbb{M}_n) : \|F''(X)\| \leq \|X\|^{3/2}, X \in B(0, R) \right\},$$

where $B(0, R) = \{X \in \mathbb{M}_n : \|X\| < R\}$ and $R > 0$. Let $f : [0, R] \rightarrow \mathbb{R}$ be given by

$$f(t) = \frac{4}{35} t^{7/2} - t.$$

It is easy to see that $f(0) = 0$, $f'(t) = (2/5)t^{5/2} - 1$, $f'(0) = -1$, and $f''(t) = t^{3/2}$, hence f satisfies (h1) and (h2) in Definition 2.2. Now, as $f''(t) = t^{3/2}$, using Lemma 2.5 it follows that f is a majorant function for all $F \in \mathcal{F}_2$ on $B(0, R)$. Moreover, the set \mathcal{F}_2 is nonempty, that is, the function $F : \mathbb{M}_n \rightarrow \mathbb{M}_n$ given by

$$F(X) = \frac{4}{35} \left(\|X\|^{5/2}X - \bar{X} \right),$$

where $\bar{X} \in \mathbb{M}_n$, is in \mathcal{F}_2 . Indeed, note that the second derivative of F is given by

$$F''(X)(V, V) = \frac{4}{35} \left[\frac{5}{4} \|X\|^{-3/2} \langle X, V \rangle^2 X + \frac{5}{2} \|X\|^{-1/2} \|V\|^2 X + 5 \|X\|^{1/2} \langle X, V \rangle V \right]$$

for all $X, V \in \mathbb{M}_n$, $X \neq 0$, and $F''(0) = 0$. Hence, from the last inequality, we obtain

$$\|F''(X)\| \leq \|X\|^{3/2}, \quad X \in \mathbb{M}_n.$$

Therefore, the statement is proved.

We need the following definition and result to present one more example. Consider \mathcal{S} the class of analytic functions from \mathbb{R}^n to \mathbb{R}^m satisfying the Smale condition at $x_0 \in \mathbb{R}^n$, that is,

$$(2.4) \quad \mathcal{S} := \left\{ F : \mathbb{R}^n \rightarrow \mathbb{R}^m : F \text{ is analytic and } \gamma := \sup_{j>1} \left\| \frac{F^{(j)}(x_0)}{j!} \right\|^{1/(j-1)} < +\infty \right\};$$

see [14]. Note that the class \mathcal{S} is nonempty, because all polynomial functions are in \mathcal{S} .

LEMMA 2.8. *If $F \in \mathcal{S}$, then for all $x \in B(x_0, 1/\gamma)$, it holds that*

$$\|F''(x)\| \leq (2\gamma)/(1 - \gamma\|x - x_0\|)^3.$$

Proof. The proof follows the same pattern as the proof of Lemma 21 in [3]. \square

Example 2.9. Consider the class of functions \mathcal{S} defined in (2.4). Let $f : [0, 1/\gamma) \rightarrow \mathbb{R}$ be defined by

$$f(t) = \frac{t}{1 - \gamma t} - 2t.$$

It is straightforward to show that $f(0) = 0$, $f'(t) = 1/(1 - \gamma t)^2 - 2$, $f'(0) = -1$, and $f''(t) = (2\gamma)/(1 - \gamma t)^3$. It follows from the last equalities that f satisfies (h1) and (h2) in Definition 2.2. To prove that the function f satisfies (2.3) with $R = 1/\gamma$ for all $F \in \mathcal{S}$, combine $f''(t) = (2\gamma)/(1 - \gamma t)^3$, with Lemmas 2.5 and 2.8. Therefore, the function f is a majorant function for all $F \in \mathcal{S}$ on $B(x_0, 1/\gamma)$.

Remark 2.10. As pointed out by one of the referees, all the above examples of majorant function are really the same with all following from a straightforward derivation using the assumption that the underlying majorant function has a Puiseux series. Indeed, assuming a Puiseux series with base $N \in \mathbb{N}$, Definition 2.2 implies that f necessarily has the form

$$(2.5) \quad f(t) = -t + \sum_{k=1}^{\infty} \alpha_k t^{(k+N)/N},$$

where the α_k 's must be chosen to yield convexity of f' for $t \geq 0$. For example, Example 2.4 is $N = 1$, $\alpha_1 = K/2$, and $\alpha_k = 0$ for $k = 2, 3, \dots$, Example 2.6 is $N = 3$, $\alpha_5 = 9/40$, and $\alpha_k = 0$ for $k = 2, 3, 4, 6, \dots$, Example 2.7 is $N = 2$, $\alpha_5 = 4/35$, and $\alpha_k = 0$ for $k = 2, 3, 4, 6, \dots$, and Example 2.9 is $N = 1$ and $\alpha_k = \gamma^k$ for $k = 1, 2, 3, \dots$. In particular, using condition in Lemma 2.5, examples of the type

$$(2.6) \quad \mathcal{F} = \left\{ F \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n) : \|F''(x)\| \leq K\|x\|^{(k-N)/N}, x \in B(0, R) \right\}$$

are obtained by taking $k \geq N$ and setting

$$(2.7) \quad f(t) = -t + K \frac{N^2}{k(k+N)} t^{(k+N)/N}.$$

For example, if $F \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ in Example 2.4, then $(K, N, k) = (K, 1, 1)$, Example 2.6 is $(K, N, k) = (1, 3, 5)$, and Example 2.7 is $(K, N, k) = (1, 2, 5)$.

2.1.2. The auxiliary function. To state our main theorem we need a certain auxiliary function associated with the majorant function as in Definition 2.2. We shall see later that the sequence generated by Algorithm 2.1 will be “majorized” by the Newton sequence associated with this auxiliary function.

Let $f : [0, R) \rightarrow \mathbb{R}$ be a function satisfying assumptions (h1) and (h2). Take $\xi > 0, \alpha > 0$ and define the *auxiliary function*

$$(2.8) \quad \begin{aligned} f_{\xi,\alpha} : [0, R) &\rightarrow \mathbb{R}, \\ t &\mapsto \xi + (\alpha - 1)t + \alpha f(t). \end{aligned}$$

Now, consider the following conditions on the auxiliary function $f_{\xi,\alpha}$:

- (h3) there exists $t_* \in (0, R)$ such that $f_{\xi,\alpha}(t) > 0$ for all $t \in (0, t_*)$ and $f_{\xi,\alpha}(t_*) = 0$;
- (h4) $f'_{\xi,\alpha}(t_*) < 0$.

From now on, we assume that (h3) holds. The assumption (h4) will be considered to hold only when explicitly stated.

PROPOSITION 2.11. *The following statements hold:*

- (i) $f_{\xi,\alpha}(0) = \xi > 0, f'_{\xi,\alpha}(0) = -1$;
- (ii) $f'_{\xi,\alpha}$ is convex and strictly increasing.

Proof. The proof is an immediate consequence of the definition in (2.8) and assumptions (h1) and (h2). \square

PROPOSITION 2.12. *The function $f_{\xi,\alpha}$ is strictly convex, and*

$$(2.9) \quad f_{\xi,\alpha}(t) > 0, \quad f'_{\xi,\alpha}(t) < 0, \quad t < t - f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t) < t_*, \quad t \in [0, t_*).$$

Moreover, $f'_{\xi,\alpha}(t_*) \leq 0$.

Proof. Using Proposition 2.11, the proof follows the same pattern as the proof of Proposition 3 in [6]. \square

In view of the second inequality in (2.9), the Newton iteration map is well defined in $[0, t_*)$. Let us call it

$$(2.10) \quad \begin{aligned} n_{f_{\xi,\alpha}} : [0, t_*) &\rightarrow \mathbb{R}, \\ t &\mapsto t - f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t). \end{aligned}$$

PROPOSITION 2.13. *For each $t \in [0, t_*)$ it holds that $\xi \leq n_{f_{\xi,\alpha}}(t) < t_*$.*

Proof. Proposition 2.12 implies that $f_{\xi,\alpha}$ is convex. Hence, using the first item of Proposition 2.11 it is easy to see, by using convexity properties, that $t - \xi \geq -f_{\xi,\alpha}(t)$. Hence, the above definition implies that

$$n_{f_{\xi,\alpha}}(t) - \xi = t - \frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)} - \xi \geq -f_{\xi,\alpha}(t) - \frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)} = \frac{f_{\xi,\alpha}(t)}{-f'_{\xi,\alpha}(t)} [f'_{\xi,\alpha}(t) + 1], \quad t \in [0, t_*).$$

As Proposition 2.11 implies that $f'_{\xi,\alpha}(0) = -1$ and $f'_{\xi,\alpha}$ is strictly increasing, we have $f'_{\xi,\alpha}(t) + 1 \geq 0$ for all $t \in [0, t_*)$. Therefore, combining the above inequality with the first two inequalities in Proposition 2.12, the first inequality follows. For proving the second inequality combine the last inequality in (2.9) with definition in (2.10). \square

PROPOSITION 2.14. *Newton iteration map $n_{f_{\xi,\alpha}}$ maps $[0, t^*)$ into $[0, t^*)$, and it holds that*

$$t < n_{f_{\xi,\alpha}}(t), \quad t_* - n_{f_{\xi,\alpha}}(t) \leq \frac{1}{2}(t_* - t), \quad t \in [0, t_*).$$

If $f_{\xi,\alpha}$ also satisfies (h4), i.e., $f'_{\xi,\alpha}(t_*) < 0$, then

$$t_* - n_{f_{\xi,\alpha}}(t) \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}(t_* - t)^2, \quad t \in [0, t_*).$$

Proof. The proof follows the same pattern as the proof of Proposition 4 in [6]. \square

The Newton sequence $\{t_k\}$ for solving the equation $f_{\xi,\alpha}(t) = 0$ with starting point $t_0 = 0$ is defined as

$$(2.11) \quad t_0 = 0, \quad t_{k+1} = n_{f_{\xi,\alpha}}(t_k), \quad k = 0, 1, \dots$$

COROLLARY 2.15. *The sequence $\{t_k\}$ is well defined, is strictly increasing, is contained in $[0, t_*)$, and converges Q -linearly to t_* as follows:*

$$t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k), \quad k = 0, 1, \dots$$

If $f_{\xi,\alpha}$ also satisfies assumption (h4), then $\{t_k\}$ converges Q -quadratically to t_* as follows:

$$t_* - t_{k+1} \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \dots;$$

moreover, the following inequality holds:

$$(2.12) \quad t_{k+1} - t_k \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}(t_k - t_{k-1})^2, \quad k = 1, 2, \dots$$

Proof. With the exception of (2.12), all statements of the corollary follow from Proposition 2.14 and (2.12). Now, we are going to prove the inequality (2.12). First note that (2.10) and (2.11) imply $f_{\xi,\alpha}(t_{k-1}) + f'_{\xi,\alpha}(t_{k-1})(t_k - t_{k-1}) = 0$ for all $k = 1, 2, \dots$. Thus, using also the continuity of $f'_{\xi,\alpha}$ we obtain

$$\begin{aligned} t_{k+1} - t_k &= \frac{1}{-f'_{\xi,\alpha}(t_k)} [f_{\xi,\alpha}(t_k) - f_{\xi,\alpha}(t_{k-1}) - f'_{\xi,\alpha}(t_{k-1})(t_k - t_{k-1})] \\ &= \frac{1}{-f'_{\xi,\alpha}(t_k)} \int_{t_{k-1}}^{t_k} f'_{\xi,\alpha}(u) - f'_{\xi,\alpha}(t_{k-1}) \, du. \end{aligned}$$

Since $\{t_k\}$ is strictly increasing and $f'_{\xi,\alpha}$ is convex, it follows from item 2 of Proposition 1.1 that

$$f'_{\xi,\alpha}(u) - f'_{\xi,\alpha}(t_{k-1}) \leq [f'_{\xi,\alpha}(t_k) - f'_{\xi,\alpha}(t_{k-1})] \frac{u - t_{k-1}}{t_k - t_{k-1}}, \quad u \in [t_{k-1}, t_k].$$

Taking into account the positivity of $-1/f'_{\xi,\alpha}(t)$ (second inequality in (2.9)) and combining the two above relations we have

$$t_{k+1} - t_k \leq (-1/f'_{\xi,\alpha}(t_k)) \int_{t_{k-1}}^{t_k} [f'_{\xi,\alpha}(t_k) - f'_{\xi,\alpha}(t_{k-1})] \frac{u - t_{k-1}}{t_k - t_{k-1}} \, du.$$

Direct integration of the last term of the above inequality yields

$$(2.13) \quad t_{k+1} - t_k \leq \frac{1}{2} \left(\frac{f'_{\xi,\alpha}(t_k) - f'_{\xi,\alpha}(t_{k-1})}{-f'_{\xi,\alpha}(t_k)} \right) (t_k - t_{k-1}).$$

As $f_{\xi,\alpha}$ satisfies assumption (h4), $f'_{\xi,\alpha}$ is increasing, and $t_{k-1} < t_k < t_*$, we obtain

$$\begin{aligned} \frac{f'_{\xi,\alpha}(t_k) - f'_{\xi,\alpha}(t_{k-1})}{-f'_{\xi,\alpha}(t_k)} &\leq \frac{1}{-f'_{\xi,\alpha}(t_*)} \frac{f'_{\xi,\alpha}(t_k) - f'_{\xi,\alpha}(t_{k-1})}{t_k - t_{k-1}} (t_k - t_{k-1}) \\ &\leq \frac{f''_{\xi,\alpha}(t_*)}{-f'_{\xi,\alpha}(t_*)} (t_k - t_{k-1}), \end{aligned}$$

where the last inequality follows from item 1 of Proposition 1.1. Combining the above inequality with (2.13) we conclude that (2.12) holds. \square

PROPOSITION 2.16. *The map $[0, t_*) \ni t \rightarrow -f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t)$ is decreasing.*

Proof. Proposition 2.12 implies that $f'_{\xi,\alpha}(t) \neq 0$ for all $t \in [0, t_*)$. Hence, the function in the proposition is well defined. As $f_{\xi,\alpha}$ is twice-differentiable we have

$$\left(\frac{-f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)} \right)' = \frac{f_{\xi,\alpha}(t)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2}{(f'_{\xi,\alpha}(t))^2}, \quad t \in [0, t_*).$$

Thus, it suffices to show that

$$(2.14) \quad f_{\xi,\alpha}(t)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2 \leq 0, \quad t \in [0, t_*).$$

Since $f_{\xi,\alpha}$ is strictly convex (Proposition 2.12) and $f'_{\xi,\alpha}$ is convex (Proposition 2.11), we have

$$0 > f_{\xi,\alpha}(t) + f'_{\xi,\alpha}(t)(t_* - t), \quad f''_{\xi,\alpha}(t) \geq 0, \quad f'_{\xi,\alpha}(t_*) \geq f'_{\xi,\alpha}(t) + f''_{\xi,\alpha}(t)(t_* - t)$$

for all $t \in [0, t_*)$. Using these inequalities and the second inequality in (2.9), we obtain

$$f_{\xi,\alpha}(t)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2 \leq f'_{\xi,\alpha}(t)(t - t_*)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2 \leq -f'_{\xi,\alpha}(t)f'_{\xi,\alpha}(t_*),$$

which combined with Proposition 2.12 yields the inequality in (2.14). Therefore, the proposition is fulfilled. \square

PROPOSITION 2.17. *Let $0 < \bar{\alpha} < \alpha$. For the auxiliary functions $f_{\xi,\bar{\alpha}}$ and $f_{\xi,\alpha}$, consider \bar{t}_* and t_* , its smallest zeros, respectively. Then the following assertions hold:*

- (i) $f_{\xi,\bar{\alpha}} < f_{\xi,\alpha}$ on $(0, R)$;
- (ii) $f'_{\xi,\bar{\alpha}} < f'_{\xi,\alpha}$ on $(0, R)$;
- (iii) $\bar{t}_* < t_*$.

Proof. From (h2) it follows that f' is strictly increasing, which implies that f is strictly convex. Thus, using (h1) we conclude that $f(t) + t > 0$ for all $t \in (0, R)$ and hence the assumption $\alpha > \bar{\alpha} > 0$ implies

$$\bar{\alpha}(t + f(t)) < \alpha(t + f(t)), \quad t \in [0, R).$$

To conclude the proof of item (i), add $\xi - t$ on both sides of the last inequality and use the definition in (2.8).

To prove item (ii), we first use that f' is strictly increasing (h2), as well as the assumption $\alpha > \bar{\alpha}$ to obtain that $(\alpha - \bar{\alpha})(f'(t) - f'(0)) > 0$ for all $t \in (0, R)$. Hence, from (h1) and some algebraic manipulation, we obtain

$$(\bar{\alpha} - 1) + \bar{\alpha}f'(t) < (\alpha - 1) + \alpha f'(t), \quad t \in [0, R).$$

Thus, by using the definition in (2.8), the statement holds.

To establish item (iii), use item (i) and the definition of \bar{t}_* and t_* in (h3). \square

2.2. Regularity. In this section we state the hypothesis on the starting point of the sequence generated by Algorithm 2.1, which we need in our analysis, as well as some related concepts.

Let C be as defined in (1.2), that is, C is the set of all minimum points of h . For $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and $x \in \mathbb{R}^n$, we define the set $D_C(x)$ associated to C as

$$D_C(x) := \{d \in \mathbb{R}^n : F(x) + F'(x)d \in C\}.$$

In the next proposition we state a relation between the sets $D_\Delta(x)$ and $D_C(x)$.

PROPOSITION 2.18. *Let $x \in \mathbb{R}^n$. If $D_C(x) \neq \emptyset$ and $d(0, D_C(x)) \leq \Delta$, then*

$$D_\Delta(x) = \{d \in \mathbb{R}^n : \|d\| \leq \Delta, F(x) + F'(x)d \in C\} \subset D_C(x).$$

As a consequence, $d(0, D_\Delta(x)) = d(0, D_C(x))$.

Proof. By definition of C in (1.2) and $D_\Delta(x)$ in (2.1) it can be seen that

$$\{d \in \mathbb{R}^n : \|d\| \leq \Delta, F(x) + F'(x)d \in C\} \subset D_\Delta(x).$$

Let $d \in D_\Delta(x)$. Since $D_C(x) \neq \emptyset$ and $d(0, D_C(x)) \leq \Delta$, there exists $\bar{d} \in D_C(x)$ such that $\|\bar{d}\| \leq \Delta$ and $F(x) + F'(x)\bar{d} \in C$. Hence, from the definition of C in (1.2) and $D_\Delta(x)$ in (2.1) we obtain $\bar{d} \in D_\Delta(x)$. Therefore, as $\bar{d}, d \in D_\Delta(x)$, and using again the definition of $D_\Delta(x)$ in (2.1), we have

$$h(F(x) + F'(x)d) = h(F(x) + F'(x)\bar{d}).$$

Now, using $F(x) + F'(x)\bar{d} \in C$, the last equality and definition of C , we obtain $F(x) + F'(x)d \in C$, which proves the first statement. The second statement, i.e., $D_\Delta(x) \subset D_C(x)$, can be seen by definition of $D_C(x)$. To conclude the proof, first note that the inclusion $D_\Delta(x) \subset D_C(x)$ implies that

$$(2.15) \quad d(0, D_\Delta(x)) \geq d(0, D_C(x)).$$

Since $D_C(x) \neq \emptyset$ and $d(0, D_C(x)) \leq \Delta$, there exists $\bar{d} \in D_C(x)$ such that

$$\|\bar{d}\| = d(0, D_C(x)) \leq \Delta.$$

Hence, from the definition of C in (1.2) and $D_\Delta(x)$ in (2.1) we conclude that $\bar{d} \in D_\Delta(x)$. Therefore,

$$d(0, D_\Delta(x)) \leq \|\bar{d}\| = d(0, D_C(x))$$

and taking into account (2.15), the proof is concluded. \square

The next definition has been introduced in [8] for studying the Gauss–Newton method (see also [9]).

DEFINITION 2.19. *Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and let $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function. A point $x_0 \in \mathbb{R}^n$ is called a quasi-regular point of the inclusion (1.2), that is, of the inclusion*

$$F(x) \in C = \operatorname{argmin} h := \{z \in \mathbb{R}^m : h(z) \leq h(x), x \in \mathbb{R}^m\},$$

if $r \in (0, +\infty)$ exists as well as an increasing positive-valued function $\beta : [0, r) \rightarrow (0, +\infty)$ such that

$$(2.16) \quad D_C(x) \neq \emptyset, \quad d(0, D_C(x)) \leq \beta(\|x - x_0\|)d(F(x), C), \quad x \in B(x_0, r).$$

Let $x_0 \in \mathbb{R}^n$ be a quasi-regular point of the inclusion (1.2). We denote r_{x_0} the supremum of r such that (2.16) holds for some increasing positive-valued function β on $[0, r)$, that is,

$$(2.17) \quad r_{x_0} := \sup \{r : \exists \beta : [0, r) \rightarrow (0, +\infty) \text{ satisfying (2.16)}\}.$$

Let $r \in [0, r_{x_0})$. The set $\mathcal{B}_r(x_0)$ denotes the set of all increasing positive-valued functions β on $[0, r)$ such that (2.16) holds, that is,

$$\mathcal{B}_r(x_0) := \{\beta : [0, r) \rightarrow (0, +\infty) : \beta \text{ satisfying (2.16)}\}.$$

Define

$$(2.18) \quad \beta_{x_0}(t) := \inf \{\beta(t) : \beta \in \mathcal{B}_{r_{x_0}}(x_0)\}, \quad t \in [0, r_{x_0}).$$

The number r_{x_0} and the function β_{x_0} are called, respectively, the *quasi-regular radius* and the *quasi-regular bound function* of the quasi-regular point x_0 .

2.2.1. Conditions yielding quasi-regularity. In this section we present some examples of the quasi-regular point of the inclusion (1.2). We begin by defining regularity.

DEFINITION 2.20. Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and let $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function with minimizer set C nonempty. A point $x_0 \in \mathbb{R}^n$ is a regular point of the inclusion $F(x) \in C$ if

$$\text{Ker}(F'(x_0)^T) \cap (C - F(x_0))^o = \{0\}.$$

As we know the definition of a quasi-regular point extends the definition of a regular point (see [8, 9]). The following proposition relates these two concepts (see [1, 8]).

PROPOSITION 2.21. Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and $x_0 \in \mathbb{R}^n$ be a regular point of the inclusion $F(x) \in C$. Then there exist constants $r > 0$ and $\beta > 0$ such that

$$D_C(x) \neq \emptyset, \quad d(0, D_C(x)) \leq \beta d(F(x), C), \quad x \in B(x_0, r).$$

Consequently, x_0 is a quasi-regular point with the quasi-regular radius $r_{x_0} \geq r$ and the quasi-regular bound function $\beta_{x_0} \leq \beta$ on $[0, r)$, as defined in (2.17) and (2.18), respectively.

Remark 2.22. Proposition 2.21 implies that each regular point of the inclusion (1.2) is a quasi-regular point of (1.2).

In the next example, we will prove that each point satisfying the Robinson condition (see [8, 10, 11]) is a quasi-regular point for the inclusion (1.2). For this, we need some definitions and results.

Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$, $C \subset \mathbb{R}^m$, be a nonempty closed convex cone and $x \in \mathbb{R}^n$. Define the multifunction $T_x : \mathbb{R}^n \rightarrow P(\mathbb{R}^m)$ as

$$(2.19) \quad T_x d = F'(x)d - C.$$

The multifunction T_x is a convex process from \mathbb{R}^n to \mathbb{R}^m . The convex process has been extensively studied in [12, 13]. As usual, the domain, norm, and inverse of T_x are defined, respectively, by

$$\mathcal{D}(T_x) := \{d \in \mathbb{R}^n : T_x d \neq \emptyset\}, \quad \|T_x\| := \sup \{\|T_x d\| : x \in \mathcal{D}(T_x), \|d\| \leq 1\},$$

$$T_x^{-1}y := \{d \in \mathbb{R}^n : F'(x)d \in y + C\}, \quad y \in \mathbb{R}^m,$$

where $\|T_x d\| := \inf\{\|v\| : v \in T_x d\}$.

The point $x_0 \in \mathbb{R}^n$ satisfies the *Robinson condition*, with respect to C and F , if the multifunction T_{x_0} carries \mathbb{R}^n onto \mathbb{R}^m , that is,

$$(2.20) \quad \forall y \in \mathbb{R}^m \quad \exists d \in \mathbb{R}^n, \exists c \in C; \quad y = F'(x_0)d - c.$$

LEMMA 2.23. *Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and C a nonempty closed convex cone. Suppose that $x_0 \in \mathbb{R}^n$ satisfies the Robinson condition. Then*

$$\|T_{x_0}^{-1}\| < +\infty.$$

Moreover, if S is a linear transformation from \mathbb{R}^n to \mathbb{R}^m such that $\|T_{x_0}^{-1}\| \|S\| < 1$, then the convex process \bar{T} , defined by $\bar{T} := T_{x_0} + S$, carries \mathbb{R}^n onto \mathbb{R}^m , $\|\bar{T}^{-1}\| < +\infty$ and

$$\|\bar{T}^{-1}\| \leq \frac{\|T_{x_0}^{-1}\|}{1 - \|T_{x_0}^{-1}\| \|S\|}.$$

Proof. See Theorem 1 on p. 342 of [11]. \square

LEMMA 2.24. *Let $F \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ and let $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function with minimizer set C nonempty. Suppose that $x_0 \in \mathbb{R}^n$ satisfies the Robinson condition with respect to C and F . Then x_0 is a regular point of the inclusion $F(x) \in C$. Moreover, assume C is a cone, $R > 0$, and $f : [0, R) \rightarrow \mathbb{R}$ is a majorant function for F on $B(x_0, R)$. Let $\xi > 0$, $\beta_0 = \|T_{x_0}^{-1}\|$, the auxiliary function $f_{\xi, \beta_0} : [0, R) \rightarrow \mathbb{R}$,*

$$f_{\xi, \beta_0}(t) := \xi + (\beta_0 - 1)t + \beta_0 f(t),$$

and $r_{\beta_0} := \sup\{t \in [0, R) : f'_{\xi, \beta_0}(t) < 0\}$. Then the quasi-regular radius r_{x_0} and the quasi-regular bound function β_{x_0} satisfy

$$r_{x_0} \geq r_{\beta_0}, \quad \beta_{x_0}(t) \leq \frac{\beta_0}{1 - \beta_0[f'(t) + 1]}, \quad t \in [0, r_{\beta_0}).$$

Proof. Take $y \in \text{Ker}(F'(x_0)^T) \cap (C - F(x_0))^o$. Hence,

$$0 = \langle F'(x_0)^T y, d \rangle = \langle y, F'(x_0)d \rangle, \quad d \in \mathbb{R}^n, \quad \langle y, c - F(x_0) \rangle \leq 0, \quad c \in C.$$

Since x_0 satisfies the Robinson condition, $d \in \mathbb{R}^n$ and $c \in C$ exist, such that $-y - F(x_0) = F'(x_0)d - c$, which combined with the above inequalities gives

$$\langle y, y \rangle = \langle y, c - F(x_0) - F'(x_0)d \rangle = \langle y, c - F(x_0) \rangle \leq 0.$$

So $y = 0$, and we obtain from Definition 2.20 that x_0 is a regular point of the inclusion $F(x) \in C$.

To establish the second part, first take $x \in \mathbb{R}^n$ such that $\|x - x_0\| \leq r_{\beta_0}$. Using that f is a majorant function of F on $B(x_0, R)$, as well as the definitions of β_0 , f_{ξ, β_0} , and r_{β_0} , we obtain

$$(2.21) \quad \begin{aligned} \|T_{x_0}^{-1}\| \|F'(x) - F'(x_0)\| &\leq \beta_0 [f'(\|x - x_0\|) - f'(0)] \\ &= f'_{\xi, \beta_0}(\|x - x_0\|) + 1 < 1. \end{aligned}$$

Using that x_0 satisfies the Robinson condition and the last inequality, it follows from Lemma 2.23 that the convex process

$$T_x d = F'(x)d - C = T_{x_0} d + [F'(x) - F'(x_0)]d, \quad d \in \mathbb{R}^n,$$

carries \mathbb{R}^n onto \mathbb{R}^m and

$$(2.22) \quad \|T_x^{-1}\| \leq \frac{\|T_{x_0}^{-1}\|}{1 - \|T_{x_0}^{-1}\| \|F'(x) - F'(x_0)\|} \leq \frac{\beta_0}{1 - \beta_0 [f'(\|x - x_0\|) - f'(0)]},$$

where the last inequality follows the definition of β_0 and (2.21). Moreover, as T_x carries \mathbb{R}^n onto \mathbb{R}^m , we also have

$$(2.23) \quad D_C(x) = \{d \in \mathbb{R}^n : F(x) + F'(x)d \in C\} \neq \emptyset, \quad x \in B(x_0, r_{\beta_0}).$$

Now, let $d \in T_x^{-1}(c - F(x))$. Using the definition of T_x^{-1} it follows that

$$F'(x)d \in c - F(x) + C = C - F(x),$$

hence we conclude that $F(x) + F'(x)d \in C$, which combined with the definition of $D_C(x)$ yields

$$T_x^{-1}(c - F(x)) \subset D_C(x).$$

Therefore,

$$d(0, D_C(x)) \leq \|T_x^{-1}(c - F(x))\| \leq \|T_x^{-1}\| \|c - F(x)\|, \quad c \in C.$$

The last inequality together with (2.22) implies

$$d(0, D_C(x)) \leq \|T_x^{-1}\| d(F(x), C) \leq \frac{\beta_0}{1 - \beta_0 [f'(\|x - x_0\|) - f'(0)]} d(F(x), C),$$

which combined with (2.23), as well as the definitions of r_{x_0} and β_{x_0} in (2.17) and (2.18), respectively, yields the desired inequalities. \square

Remark 2.25. Lemma 2.24 implies that each point satisfying the Robinson condition, with respect to C and F , is a quasi-regular point of the inclusion (1.2).

Next, we will give an example of a quasi-regular point for the inclusion (1.2) which is not a regular point.

Example 2.26. Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a real-valued convex function such that $\operatorname{argmin} h = \{0\}$. Consider the linear function $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by

$$F(x) = Q(\bar{x} + x),$$

where $n > 2$, $\bar{x} \in \mathbb{R}^n$, and $Q \in \mathbb{M}_n$ such that $\dim(\operatorname{Ker}(Q)) > 1$ and $Q \neq 0$. The point $x_0 \in \mathbb{R}^n \setminus \{-\bar{x}\}$ such that $Q(\bar{x} + x_0) \neq 0$ is a quasi-regular point of the inclusion $F(x) \in \{0\}$, with quasi-regular radius r_{x_0} and quasi-regular bound function β_{x_0} satisfying, respectively,

$$r_{x_0} \geq \frac{\|Q(\bar{x} + x_0)\|}{\|Q\|}, \quad \beta_{\bar{x}}(t) \leq \frac{\|\bar{x} - x_0\| + t}{\|Q(\bar{x} + x_0)\| - \|Q\|t}, \quad t \in [0, \|Q(\bar{x} + x_0)\|/\|Q\|].$$

Indeed, first note that $F'(x)d = Qd$ for all $d \in \mathbb{R}^n$. It is easy to see that

$$-\bar{x} - x \in D_C(x) = \{d \in \mathbb{R}^n : Q(\bar{x} + x) + Qd \in \{0\}\}, \quad x \in \mathbb{R}^n,$$

which implies $d(0, D_C(x)) \leq \|\bar{x} + x\| \leq \|\bar{x} + x_0\| + \|x - x_0\|$. Now, simple calculations yield

$$d(F(x), \{0\}) = \|Q(\bar{x} + x)\| \geq \|Q(\bar{x} + x_0)\| - \|Q\|\|x - x_0\|.$$

Therefore, using the two last inequalities we obtain

$$d(0, D_C(x)) \leq \frac{\|\bar{x} + x_0\| + \|x - x_0\|}{\|Q(\bar{x} + x_0)\| - \|Q\|\|x - x_0\|} d(F(x), \{0\})$$

and

$$x \in B(x_0, \|Q(\bar{x} + x_0)\|/\|Q\|),$$

and the statement is proved.

Now, we will prove that the point x_0 given above is not a regular point for the inclusion $F(x) \in \{0\}$. First, as $F'(x_0) = Q$ and $\dim(\text{Ker}(Q)) > 1$, we conclude that $\dim(\text{Ker}(F'(x_0)^T)) > 1$. On the other hand, since $F(x_0) \neq 0$, we obtain $\{F(x_0)\}^\perp \subset (-F(x_0))^0$, where $\dim(\{F(x_0)\}^\perp) = n - 1$. Hence, $\text{Ker}(F'(x_0)^T) \cap (-F(x_0))^0 \neq \{0\}$, which proves the statement.

For additional examples of quasi-regular points of the inclusion (1.2) which are not regular points see [8].

3. Semilocal analysis for the Gauss–Newton method. In this section our goal is to state and prove a semilocal theorem for the sequence generated by Algorithm 2.1 in order to solve problem (1.1). Under the hypothesis that the initial point is a quasi-regular point of the inclusion (1.2) and the nonlinear function F satisfies the majorant condition in Definition 2.2, we will prove convergence of the sequence to a point $x_* \in B[x_0, t_*]$ such that $F(x_*) \in C$ and in particular that x_* solves (1.1). The statement of the theorem is as follows.

THEOREM 3.1. *Let $F \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. Assume that $R > 0$, $x_0 \in \mathbb{R}^n$, and $f : [0, R] \rightarrow \mathbb{R}$ is a majorant function for F on $B(x_0, R)$. Take the constants $\alpha > 0$ and $\xi > 0$ and consider the auxiliary function $f_{\xi, \alpha} : [0, R] \rightarrow \mathbb{R}$,*

$$f_{\xi, \alpha}(t) := \xi + (\alpha - 1)t + \alpha f(t).$$

If $f_{\xi, \alpha}$ satisfies (h3), i.e., t_ is the smallest zero of $f_{\xi, \alpha}$, then the sequence generated by the Newton method for solving $f_{\xi, \alpha}(t) = 0$, with starting point $t_0 = 0$,*

$$(3.1) \quad t_{k+1} = t_k - f'_{\xi, \alpha}(t_k)^{-1} f_{\xi, \alpha}(t_k), \quad k = 0, 1, \dots,$$

is well defined, $\{t_k\}$ is strictly increasing, contained in $[0, t_)$, and it converges Q -linearly to t_* . Let $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$, and $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function with minimizer set C nonempty. Suppose that $x_0 \in \mathbb{R}^n$ is a quasi-regular point of the inclusion*

$$F(x) \in C$$

with the quasi-regular radius r_{x_0} and the quasi-regular bound function β_{x_0} as defined in (2.17) and (2.18), respectively. If $d(F(x_0), C) > 0$, $t_ \leq r_{x_0}$,*

$$(3.2) \quad \Delta \geq \xi \geq \eta \beta_{x_0}(0) d(F(x_0), C), \quad \alpha \geq \sup \left\{ \frac{\eta \beta_{x_0}(t)}{\eta \beta_{x_0}(t) [f'(t) + 1] + 1} : \xi \leq t < t_* \right\},$$

then the sequence generated by Algorithm 2.1, denoted by $\{x_k\}$, is contained in $B(x_0, t_*)$,

$$(3.3) \quad F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C, \quad k = 0, 1, \dots,$$

satisfies the inequalities

$$(3.4) \quad \|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \quad \|x_{k+1} - x_k\| \leq \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} \|x_k - x_{k-1}\|^2,$$

for all $k = 0, 1, \dots$, and $k = 1, 2, \dots$, respectively, and converges to a point $x_* \in B[x_0, t_*]$ such that $F(x_*) \in C$,

$$(3.5) \quad \|x_* - x_k\| \leq t_* - t_k, \quad k = 0, 1, \dots,$$

and the convergence is R -linear. If, additionally, $f_{\xi, \alpha}$ satisfies (h4), then the following inequalities hold:

$$(3.6) \quad \|x_{k+1} - x_k\| \leq \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)} \|x_k - x_{k-1}\|^2, \quad t_{k+1} - t_k \leq \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)} (t_k - t_{k-1})^2$$

for all $k = 1, 2, \dots$. Moreover, the sequences $\{x_k\}$ and $\{t_k\}$ converge Q -quadratically to x_* and t_* , respectively, as follows:

$$(3.7) \quad \limsup_{k \rightarrow \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \leq \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)}, \quad t_* - t_{k+1} \leq \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)} (t_* - t_k)^2$$

for all $k = 0, 1, \dots$

Remark 3.2. If

$$\alpha > \bar{\alpha} := \sup \left\{ \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f'(t) + 1] + 1} : \xi \leq t < t_* \right\},$$

then the sequence $\{x_k\}$ converges Q -quadratically to x_* . To prove this assertion, note that through item (iii) of Proposition 2.17, we have $\bar{t}_* < t_*$. Hence, using $f'_{\xi, \bar{\alpha}}$ strictly increasing and item (ii) of Proposition 2.17, we obtain

$$f'_{\xi, \bar{\alpha}}(\bar{t}_*) < f'_{\xi, \bar{\alpha}}(t_*) < f'_{\xi, \alpha}(t_*),$$

which combined with Proposition 2.12 implies that $f'_{\xi, \bar{\alpha}}(t_*) < 0$, that is, $f_{\xi, \bar{\alpha}}$ satisfies (h4). So, the statement is correct if $f_{\xi, \alpha}$ is replaced by $f_{\xi, \bar{\alpha}}$ in Theorem 3.1.

All statements in Theorem 3.1 for the sequence $\{t_k\}$ were proved in Corollary 2.15. Now we are going to prove the statements for the sequence $\{x_k\}$. From now on, we assume that the hypotheses of Theorem 3.1 hold, with the exception of (h4), which will be considered to hold only when explicitly stated.

3.1. Proof of convergence. As we saw in section 2, $D_\Delta(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$; therefore the sequence $\{x_k\}$ is well defined. But this is not enough to prove the convergence of sequence $\{x_k\}$ to some point $x_* \in \mathbb{R}^n$ such that $F(x_*) \in C$, because we have no relationship between the set of search directions $D_\Delta(x)$ to the set of solutions of the linearized inclusion

$$F(x) + F'(x)d \in C, \quad \|d\| \leq \Delta.$$

Now, if we prove that $D_\Delta(x) \subset D_C(x)$ for suitable points, then we can use the results of regularity to relate the sets mentioned above. First, we define some subsets of $B(x_0, t_*)$ in which, as we shall prove, the desired inclusion holds for all points in these subsets.

$$(3.8) \quad K(t) := \left\{ x \in \mathbb{R}^n : \|x - x_0\| \leq t, \eta d(0, D_C(x)) \leq -\frac{f_{\xi, \alpha}(t)}{f'_{\xi, \alpha}(t)} \right\}, \quad t \in [0, t_*),$$

$$(3.9) \quad K := \bigcup_{t \in [0, t_*)} K(t).$$

In (3.8) we assume that $0 \leq t < t_*$; therefore it follows from Proposition 2.12 that $f'_{\xi, \alpha}(t) \neq 0$. So, the above definitions are consistent.

PROPOSITION 3.3. *If $x \in K$, then*

$$D_\Delta(x) = \{d \in \mathbb{R}^n : F(x) + F'(x)d \in C, \|d\| \leq \Delta\} \subset D_C(x)$$

and

$$d(0, D_\Delta(x)) = d(0, D_C(x)).$$

Proof. From Proposition 2.18 it is sufficient to prove that $D_C(x) \neq \emptyset$ and $d(0, D_C(x)) \leq \Delta$ for all $x \in K$. Let $x \in K$; then $x \in K(t)$ for some $t \in [0, t_*)$, which implies that $x \in B(x_0, t_*)$. Since $t_* \leq r_{x_0}$ and x_0 is a quasi-regular point, it follows from Definition 2.19 and the definition of the quasi-regular radius in (2.17) that $D_C(x) \neq \emptyset$. By hypothesis $\eta \geq 1$ and $\xi \leq \Delta$. Thus, as $x \in K(t)$, by using the definition in (3.8), Proposition 2.16, and Proposition 2.11 we obtain

$$d(0, D_C(x)) \leq \eta d(0, D_C(x)) \leq -\frac{f_{\xi, \alpha}(t)}{f'_{\xi, \alpha}(t)} \leq -\frac{f_{\xi, \alpha}(0)}{f'_{\xi, \alpha}(0)} = \xi \leq \Delta,$$

which proves the desired result. \square

For each $x \in \mathbb{R}^n$, we define the set $\bar{D}_\Delta(x)$ as

$$(3.10) \quad \bar{D}_\Delta(x) := \{d \in D_\Delta(x) : \|d\| \leq \eta d(0, D_\Delta(x))\}.$$

As $D_\Delta(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$, we have $\bar{D}_\Delta(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$ and consequently the Gauss–Newton iteration multifunction is well defined. Let us call G_F the Gauss–Newton iteration multifunction for F in $B(x_0, t_*)$:

$$(3.11) \quad \begin{aligned} G_F : B(x_0, t_*) &\rightarrow P(\mathbb{R}^n), \\ x &\mapsto x + \bar{D}_\Delta(x). \end{aligned}$$

We shall prove that the Gauss–Newton iteration multifunction is well behaved on the subsets defined in (3.8), but first we need the following technical result.

LEMMA 3.4. *For each $t \in [0, t_*)$, $x \in K(t)$, and $y \in G_F(x)$ it holds that*

- (i) $\|y - x\| \leq n_{f_{\xi, \alpha}}(t) - t$;
- (ii) $\|y - x_0\| \leq n_{f_{\xi, \alpha}}(t) < t_*$;
- (iii) $\eta d(0, D_C(y)) \leq -\frac{f_{\xi, \alpha}(n_{f_{\xi, \alpha}}(t))}{f'_{\xi, \alpha}(n_{f_{\xi, \alpha}}(t))} \left(\frac{\|y - x\|}{n_{f_{\xi, \alpha}}(t) - t} \right)^2$.

Proof. Since $t \in [0, t_*)$ and $x \in K(t)$, by using the definition in (3.8), Proposition 3.3, and the first two statements in Proposition 2.14, we obtain

$$(3.12) \quad \|x - x_0\| \leq t, \eta d(0, D_\Delta(x)) = \eta d(0, D_C(x)) \leq -\frac{f_{\xi, \alpha}(t)}{f'_{\xi, \alpha}(t)}, \quad t < n_{f_{\xi, \alpha}}(t) < t_*.$$

Now, as $y \in G_F(x)$ there exists $d \in \bar{D}_\Delta(x)$ such that $y = x + d$. Using the definition of the set $\bar{D}_\Delta(x)$ in (3.10) and the second inequality in (3.12) it follows that

$$\|d\| \leq \eta d(0, D_\Delta(x)) = \eta d(0, D_C(x)) \leq -f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t).$$

Since $d = y - x$, the last inequality together with the definition in (2.10) implies item (i).

Triangular inequality combined with the first inequality in (3.12), item (i), and the last inequality in (3.12) yields

$$(3.13) \quad \|y - x_0\| \leq \|y - x\| + \|x - x_0\| \leq n_{f_{\xi,\alpha}}(t) < t_*,$$

which proves item (ii).

Since $\|y - x_0\| < t_*$ and $t_* \leq r_{x_0}$ we obtain by the quasi-regularity assumption

$$D_C(y) \neq \emptyset, \quad d(0, D_C(y)) \leq \beta_{x_0}(\|y - x_0\|)d(F(y), C).$$

As $x \in K(t) \subset K$ and $y - x = d \in D_\Delta(x)$, it follows from Proposition 3.3 that

$$F(x) + F'(x)(y - x) \in C.$$

Therefore, taking into account that $\eta \geq 1$, by using the above inequality and the last inclusion it is easy to conclude that

$$\eta d(0, D_C(y)) \leq \eta \beta_{x_0}(\|y - x_0\|)\|F(y) - F(x) - F'(x)(y - x)\|.$$

On the other hand, from item (i) we have $\|y - x\| \leq n_{f_{\xi,\alpha}}(t) - t$ and, as $\|x - x_0\| \leq t$, by using Lemma 2.3 we have

$$\|F(y) - F(x) - F'(x)(y - x)\| \leq [f(n_{f_{\xi,\alpha}}(t)) - f(t) - f'(t)(n_{f_{\xi,\alpha}}(t) - t)] \left(\frac{\|y - x\|}{n_{f_{\xi,\alpha}}(t) - t} \right)^2.$$

Hence, combining the above two inequalities we conclude that

$$(3.14) \quad \eta d(0, D_C(y)) \leq \eta \beta_{x_0}(\|y - x_0\|) [f(n_{f_{\xi,\alpha}}(t)) - f(t) - f'(t)(n_{f_{\xi,\alpha}}(t) - t)] \left(\frac{\|y - x\|}{n_{f_{\xi,\alpha}}(t) - t} \right)^2.$$

The definition in (2.10) implies that $f_{\xi,\alpha}(t) + f'_{\xi,\alpha}(t)(n_{f_{\xi,\alpha}}(t) - t) = 0$. So, we have

$$f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) = f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) - f_{\xi,\alpha}(t) - f'_{\xi,\alpha}(t)(n_{f_{\xi,\alpha}}(t) - t).$$

By using the definition in (2.8) and after simple algebraic manipulation, the last equality becomes

$$f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) = \alpha (f(n_{f_{\xi,\alpha}}(t)) - f(t) - f'(t)(n_{f_{\xi,\alpha}}(t) - t)).$$

So, as β_{x_0} is an increasing function, by a simple combination of (3.13), (3.14), and the last equality, we obtain

$$(3.15) \quad \eta d(0, D_C(y)) \leq \frac{\eta \beta_{x_0}(n_{f_{\xi,\alpha}}(t))}{\alpha} f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) \left(\frac{\|y - x\|}{n_{f_{\xi,\alpha}}(t) - t} \right)^2.$$

From Proposition 2.13 we have $\xi \leq n_{f_{\xi,\alpha}}(t) < t_*$. Thus, combining (3.2), (h1), and (h2), we obtain after simple calculations that

$$\alpha\eta\beta_{x_0}(n_{f_{\xi,\alpha}}(t))[f'(n_{f_{\xi,\alpha}}(t)) + 1] + \alpha \geq \eta\beta_{x_0}(n_{f_{\xi,\alpha}}(t)).$$

Hence, using $f'_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) = (\alpha - 1) + \alpha f'(n_{f_{\xi,\alpha}}(t))$ and the second inequality in (2.9), the last inequality becomes

$$\eta\beta_{x_0}(n_{f_{\xi,\alpha}}(t))/\alpha \leq -1/f'_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)),$$

which combined with (3.15) yields item (iii). \square

In the next lemma we prove the desired result, namely, that the Gauss–Newton iteration multifunction is well behaved on the subsets defined in (3.8).

LEMMA 3.5. *For each $t \in [0, t_*)$, the following inclusions hold: $K(t) \subset B(x_0, t_*)$ and*

$$G_F(K(t)) \subset K(n_{f_{\xi,\alpha}}(t)).$$

As a consequence, $K \subset B(x_0, t_*)$ and $G_F(K) \subset K$.

Proof. The first inclusion follows trivially from the definition of $K(t)$. Take $x \in K(t)$ and $y \in G_F(x)$. Combining items (i) and (iii) of Lemma 3.4 we have

$$\eta d(0, D_C(y)) \leq -\frac{f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t))}{f'_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t))}.$$

The last inequality together with item (ii) of Lemma 3.4 and the definition in (3.8) show us that $y \in K(n_{f_{\xi,\alpha}}(t))$, which proves the second inclusion.

The next inclusion, first on the second sentence, follows trivially from definitions (3.8) and (3.9). To verify the last inclusion, take $x \in K$. Therefore, $x \in K(t)$ for some $t \in [0, t_*)$. Using the first part of the lemma, we conclude that $G_F(x) \subset K(n_{f_{\xi,\alpha}}(t))$. To end the proof, note that $n_{f_{\xi,\alpha}}(t) \in [0, t_*)$ and use the definition of K . \square

Finally, we are ready to prove the main result of this section, which is an immediate consequence of the latter results. First, note that definitions (3.10) and (3.11) imply that the sequence $\{x_k\}$ satisfies

$$(3.16) \quad x_{k+1} \in G_F(x_k), \quad k = 0, 1, \dots,$$

which is indeed an equivalent definition of this sequence.

COROLLARY 3.6. *The sequence $\{x_k\}$ which is contained in $B(x_0, t_*)$, converges to a point $x_* \in B[x_0, t_*]$ such that $F(x_*) \in C$. Moreover, $\{x_k\}$ and $\{t_k\}$ satisfy (3.3), (3.4), and (3.5). Furthermore, if $f_{\xi,\alpha}$ also satisfies assumption (h4), then $\{x_k\}$ satisfies the first inequality in (3.6) and converges Q -quadratically to x_* as the first inequality in (3.7).*

Proof. Since $x_0 \in B(x_0, t_*) \subseteq B(x_0, r_{x_0})$, by using the quasi-regularity assumption, $\eta \geq 1$, the first inequality in (3.2), and Proposition 2.11, we obtain

$$D_C(x_0) \neq \emptyset, \quad \eta d(0, D_C(x_0)) \leq \eta\beta_{x_0}(0)d(F(x_0), C) \leq \xi = -\frac{f_{\xi,\alpha}(0)}{f'_{\xi,\alpha}(0)}.$$

Therefore,

$$x_0 \in K(0) \subset K,$$

where the second inclusion follows trivially from (3.9). Using the above inclusion, the inclusions $G_F(K) \subset K$ (Lemma 3.5), and (3.16), we conclude that the sequence $\{x_k\}$ rests in K and, in particular, we have $\{x_k\}$ contained in $B(x_0, t_*)$. Since $\{x_k\} \subset K$, by combining Proposition 3.3 and Algorithm 2.1, the inclusion in (3.3) follows. Now, we prove by induction that

$$(3.17) \quad x_k \in K(t_k), \quad k = 0, 1, \dots$$

The above inclusion, for $k = 0$, is the first result in this proof. Assume that $x_k \in K(t_k)$. From (2.11) we have $t_{k+1} = \eta_{f_{\xi,\alpha}}(t_k)$ and, as $x_k \in K(t_k)$, Lemma 3.5 implies that $G_F(x_k) \subset K(t_{k+1})$, which taking into account (3.16) completes the induction proof.

A simple combination of Algorithm 2.1 with (3.17), Proposition 3.3, and (3.8) yields

$$(3.18) \quad \|x_{k+1} - x_k\| \leq \eta d(0, D_\Delta(x_k)) = \eta d(0, D_C(x_k)) \leq -\frac{f_{\xi,\alpha}(t_k)}{f'_{\xi,\alpha}(t_k)}, \quad k = 0, 1, \dots,$$

which, using (3.1), becomes

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \quad k = 0, 1, \dots$$

So, the first inequality in (3.4) holds. On the other hand, as $\{t_k\}$ converges to t_* , the above inequalities imply that

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

for any $k_0 \in \mathbb{N}$. Hence, $\{x_k\}$ is a Cauchy sequence in $B(x_0, t_*)$ and so converges to some $x_* \in B[x_0, t_*]$. Moreover, the above inequality also implies (3.5), i.e., $\|x_* - x_k\| \leq t_* - t_k$, for any k . As C is closed, $\{x_k\}$ converges to x_* ,

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

and F is a continuously differentiable function; therefore, we have $F(x_*) \in C$.

In order to prove the second inequality in (3.4), first note that $x_k \in K(t_k)$ and $t_{k+1} = \eta_{f_{\xi,\alpha}}(t_k)$, for all $k = 0, 1, \dots$. Thus, take an arbitrary k and apply item (iii) of Lemma 3.4 with $y = x_k$, $x = x_{k-1}$, and $t = t_{k-1}$ to obtain

$$\eta d(0, D_C(x_k)) \leq -\frac{f_{\xi,\alpha}(t_k)}{f'_{\xi,\alpha}(t_k)} \left(\frac{\|x_k - x_{k-1}\|}{t_k - t_{k-1}} \right)^2,$$

which, using (3.1) and the first inequality in (3.18), yields the desired inequality.

Now, we assume that (h4) holds. Therefore, combining the second inequalities in (3.4) and (3.6), we obtain the first inequality in (3.6). To establish the first inequality in (3.7), use the first inequality in (3.6) and Proposition 1.2 with $z_k = x_k$ and $\Theta = f''_{\xi,\alpha}(t_*)/(-2f'_{\xi,\alpha}(t_*))$. Thus, the proof is concluded. \square

Therefore, it follows from Corollaries 2.15 and 3.6 that all statements in Theorem 3.1 are valid.

4. Special cases. In this section, we present special cases for Theorem 3.1. They include the case where the starting point x_0 of the Gauss–Newton sequence is a regular point of the inclusion (1.2) and the case where x_0 satisfies the Robinson condition. Moreover, we present results of convergence under the Lipschitz and Smale conditions.

4.1. Convergence result for regular starting point. In this section we present a correspondent theorem to Theorem 3.1, namely, we assume that the starting point x_0 of the Gauss–Newton sequence is a regular point of the inclusion (1.2); see [1] and references therein. We also present results of convergence under the Lipschitz and Smale conditions.

THEOREM 4.1. *Let $F \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. Assume that $R > 0$, $x_0 \in \mathbb{R}^n$, and $f : [0, R) \rightarrow \mathbb{R}$ is a majorant function for F on $B(x_0, R)$. Take the constants $\alpha > 0$ and $\xi > 0$ and consider the auxiliary function $f_{\xi,\alpha} : [0, R) \rightarrow \mathbb{R}$,*

$$f_{\xi,\alpha}(t) = \xi + (\alpha - 1)t + \alpha f(t).$$

If $f_{\xi,\alpha}$ satisfies (h3), i.e., t_ is the smallest zero of $f_{\xi,\alpha}$, then the sequence generated by the Newton method for solving $f_{\xi,\alpha}(t) = 0$, with starting point $t_0 = 0$,*

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \dots,$$

is well defined, $\{t_k\}$ is strictly increasing, contained in $[0, t_)$, and it converges Q -linearly to t_* . Let $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$, and $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function with minimizer set C nonempty. Suppose that $x_0 \in \mathbb{R}^n$ is a regular point of the inclusion $F(x) \in C$ with associated constants $r > 0$ and $\beta > 0$. If $d(F(x_0), C) > 0$, $t_* \leq r$,*

$$\Delta \geq \xi \geq \eta\beta d(F(x_0), C), \quad \alpha \geq \eta\beta/(\eta\beta[f'(\xi) + 1] + 1),$$

then the sequence generated by Algorithm 2.1, denoted by $\{x_k\}$, is contained in $B(x_0, t_)$,*

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

satisfies the inequalities

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

for $k = 0, 1, \dots$, and $k = 1, 2, \dots$, respectively, converges to a point $x_ \in B[x_0, t_*]$ such that $F(x_*) \in C$,*

$$\|x_* - x_k\| \leq t_* - t_k, \quad k = 0, 1, \dots,$$

and the convergence is R -linear. If, additionally, $f_{\xi,\alpha}$ satisfies (h4), then the following inequalities hold:

$$\|x_{k+1} - x_k\| \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} \|x_k - x_{k-1}\|^2, \quad t_{k+1} - t_k \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} (t_{t_k} - t_{k-1})^2$$

for all $k = 1, 2, \dots$. Moreover, the sequences $\{x_k\}$ and $\{t_k\}$ converge Q -quadratically to x_* and t_* , respectively, as follows:

$$\limsup_{k \rightarrow \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}, \quad t_* - t_{k+1} \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \dots$$

Proof. Since x_0 is a regular point for the inclusion, we have from Proposition 2.21 that x_0 is a quasi-regular point for the inclusion $F(x) \in C$ with the quasi-regular radius $r_{x_0} \geq r$. So, taking into account the assumption $t_* \leq r$ we obtain

$$t_* < r_{x_0}.$$

Moreover, Proposition 2.21 also implies the quasi-regular bound function

$$(4.1) \quad \beta_{x_0}(t) \leq \beta, \quad t \in [0, r).$$

Since $\Delta \geq \xi \geq \eta\beta d(F(x_0), C)$ and the last inequality implies that $\beta_{x_0}(0) \leq \beta$, we have

$$\Delta \geq \xi \geq \eta\beta_{x_0}(0)d(F(x_0), C).$$

Now, combining the assumptions $0 < \xi$ and $t_* \leq r$ with Proposition 2.13 we conclude that $0 < \xi < t_* \leq r$. So, using (4.1), $f'(0) = -1$, f' as strictly increasing, and $\eta \geq 1$, after simple algebraic manipulation we obtain

$$\frac{\eta\beta}{\eta\beta[f'(\xi) + 1] + 1} \geq \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f'(t) + 1] + 1}, \quad t \in [\xi, t_*).$$

Hence, the assumption $\alpha \geq \eta\beta/(\eta\beta[f'(\xi) + 1] + 1)$ and the last inequality imply that

$$\alpha \geq \sup \left\{ \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f'(t) + 1] + 1} : \xi \leq t < t_* \right\}.$$

Therefore, F and x_0 satisfy all assumptions in Theorem 3.1 and consequently the statements of the theorem are satisfied. \square

Under the Lipschitz condition, Theorem 4.1 becomes the following.

THEOREM 4.2. *Let $F \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. Assume that $x_0 \in \mathbb{R}^n$, $R > 0$, and $K > 0$ such that*

$$\|F'(y) - F'(x)\| \leq K\|y - x\|, \quad x, y \in B(x_0, R).$$

Take the constants $\alpha > 0$ and $\xi > 0$ and consider the auxiliary function $f_{\xi,\alpha}: [0, R) \rightarrow \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \xi - t + (\alpha K t^2)/2.$$

If $2\alpha K \xi \leq 1$, then $t_* = (1 - \sqrt{1 - 2\alpha K \xi})/(\alpha K)$ is the smallest zero of $f_{\xi,\alpha}$, the sequence generated by the Newton method for solving $f_{\xi,\alpha}(t) = 0$, with starting point $t_0 = 0$,

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \dots,$$

is well defined, $\{t_k\}$ is strictly increasing, contained in $[0, t_*)$, and it converges Q -linearly to t_* . Let $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ and $h: \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex

function with minimizer set C nonempty. Suppose that $x_0 \in \mathbb{R}^n$ is a regular point of the inclusion $F(x) \in C$ with associated constants $r > 0$ and $\beta > 0$. If $d(F(x_0), C) > 0$, $t_* \leq r$,

$$\Delta \geq \xi \geq \eta\beta d(F(x_0), C), \quad \alpha \geq \eta\beta / (K\eta\beta\xi + 1),$$

then the sequence generated by Algorithm 2.1, denoted by $\{x_k\}$, is contained in $B(x_0, t_*)$,

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

satisfies the inequalities

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

for $k = 0, 1, \dots$, and $k = 1, 2, \dots$, respectively, converging to a point $x_* \in B[x_0, t_*]$ such that $F(x_*) \in C$,

$$\|x_* - x_k\| \leq t_* - t_k, \quad k = 0, 1, \dots,$$

and the convergence is R -linear. If, additionally, $2\alpha K\xi < 1$, then the following inequalities hold:

$$\|x_{k+1} - x_k\| \leq \frac{\alpha K}{2\sqrt{1 - 2\alpha K\xi}} \|x_k - x_{k-1}\|^2, \quad t_{k+1} - t_k \leq \frac{\alpha K}{2\sqrt{1 - 2\alpha K\xi}} (t_k - t_{k-1})^2$$

for all $k = 1, 2, \dots$. Moreover, the sequences $\{x_k\}$ and $\{t_k\}$ converge Q -quadratically to x_* and t_* , respectively, as follows:

$$\limsup_{k \rightarrow \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \leq \frac{\alpha K}{2\sqrt{1 - 2\alpha K\xi}}, \quad t_* - t_{k+1} \leq \frac{\alpha K}{2\sqrt{1 - 2\alpha K\xi}} (t_* - t_k)^2$$

for all $k = 0, 1, \dots$.

Proof. It is promptly proved that $f : [0, R) \rightarrow \mathbb{R}$ defined by $f(t) = Kt^2/2 - t$ is a majorant function for the function F on $B(x_0, R)$; see Example 2.4. Hence, for

$$f_{\xi, \alpha}(t) = \xi - t + (\alpha K t^2)/2 = \xi + (\alpha - 1)t + \alpha f(t),$$

and $2\alpha K\xi \leq 1$, we conclude that $f_{\xi, \alpha}$ satisfies (h3), $t_* = (1 - \sqrt{1 - 2\alpha K\xi})/(\alpha K)$ is its smallest root, and

$$\alpha \geq \frac{\eta\beta}{\eta\beta[f'(\xi) + 1] + 1}, \quad \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)} = \frac{\alpha K}{2\sqrt{1 - 2\alpha K\xi}}.$$

Therefore, taking α , $f_{\xi, \alpha}$, and t_* as defined above and by noting that assumption $2\alpha K\xi < 1$ implies that $f_{\xi, \alpha}$ satisfies (h4), all the statements of the theorem follow from Theorem 4.1. \square

Under the Smale condition (see [14]), Theorem 4.1 becomes the following.

THEOREM 4.3. *Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be an analytic function. Assume that $x_0 \in \mathbb{R}^n$ and*

$$\gamma := \sup_{j > 1} \left\| \frac{F^{(j)}(x_0)}{j!} \right\|^{1/(j-1)} < +\infty.$$

Take the constants $\alpha > 0$ and $\xi > 0$ and consider the auxiliary function $f_{\xi,\alpha} : [0, 1/\gamma) \rightarrow \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \frac{\alpha\gamma}{1-\gamma t}t^2 - t + \xi.$$

If $\xi\gamma \leq 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$ then

$$t_* = \frac{1 + \gamma\xi - \sqrt{(1 + \gamma\xi)^2 - 4(1 + \alpha)\gamma\xi}}{2(1 + \alpha)\gamma}$$

is the smallest zero of $f_{\xi,\alpha}$, the sequence generated by the Newton method for solving $f_{\xi,\alpha}(t) = 0$, with starting point $t_0 = 0$,

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1}f_{\xi,\alpha}(t_k), \quad k = 0, 1, \dots,$$

is well defined, $\{t_k\}$ is strictly increasing, contained in $[0, t_*)$, and it converges Q -linearly to t_* . Let $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ and $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function with minimizer set C nonempty. Suppose that $x_0 \in \mathbb{R}^n$ is a regular point of the inclusion $F(x) \in C$ with associated constants $r > 0$ and $\beta > 0$. If $d(F(x_0), C) > 0$, $t_* \leq r$,

$$\Delta \geq \xi \geq \eta\beta d(F(x_0), C), \quad \alpha \geq \frac{\eta\beta(1-\gamma\xi)^2}{\eta\beta + (1-\eta\beta)(1-\gamma\xi)^2},$$

then the sequence generated by Algorithm 2.1, denoted by $\{x_k\}$, is contained in $B(x_0, t_*)$,

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

satisfies the inequalities

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

for $k = 0, 1, \dots$, and $k = 1, 2, \dots$, respectively, converging to a point $x_* \in B[x_0, t_*]$ such that $F(x_*) \in C$,

$$\|x_* - x_k\| \leq t_* - t_k, \quad k = 0, 1, \dots,$$

and the convergence is R -linear. If, additionally, $\xi\gamma < 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$, then the following inequalities hold:

$$\|x_{k+1} - x_k\| \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} \|x_k - x_{k-1}\|^2, \quad t_{k+1} - t_k \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} (t_k - t_{k-1})^2$$

for all $k = 1, 2, \dots$. Moreover, the sequences $\{x_k\}$ and $\{t_k\}$ converge Q -quadratically to x_* and t_* , respectively, as follows:

$$\limsup_{k \rightarrow \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}, \quad t_* - t_{k+1} \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} (t_* - t_k)^2, \quad k = 0, 1, \dots$$

Proof. From Example 2.9 we know that $f : [0, 1/\gamma) \rightarrow \mathbb{R}$ defined by $f(t) = t/(1-\gamma t) - 2t$ is a majorant function for F on $B(x_0, 1/\gamma)$. Hence, for

$$f_{\xi,\alpha}(t) = \frac{\alpha\gamma}{1-\gamma t}t^2 - t + \xi = \xi + (\alpha - 1)t + \alpha f(t)$$

and $\xi\gamma \leq 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$, we conclude that $f_{\xi,\alpha}$ satisfies (h3) and

$$\alpha \geq \frac{\eta\beta}{\eta\beta[f'(\xi) + 1] + 1}, \quad t_* = \frac{1 + \gamma\xi - \sqrt{(1 + \gamma\xi)^2 - 4(1 + \alpha)\gamma\xi}}{2(1 + \alpha)\gamma}$$

is its smallest root. Therefore, taking α , $f_{\xi,\alpha}$ and t_* as defined above and by noting that assumption $\xi\gamma < 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$ implies that $f_{\xi,\alpha}$ satisfies (h4), all the statements of the theorem follow from Theorem 4.1. \square

4.2. Convergence result under the Robinson condition . In this section we present a correspondent theorem to Theorem 3.1, namely, we assume that the starting point x_0 of the Gauss-Newton sequence satisfies the Robinson condition, see [8] and [11]. Under the Robinson condition, we also present results of convergence for the Lipschitz and Smale conditions.

THEOREM 4.4. *Let $F \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. Assume that $R > 0$, $x_0 \in \mathbb{R}^n$, and $f : [0, R) \rightarrow \mathbb{R}$ is a majorant function for F on $B(x_0, R)$. Take the constants $\alpha > 0$ and $\xi > 0$ and consider the auxiliary function $f_{\xi,\alpha} : [0, R) \rightarrow \mathbb{R}$,*

$$f_{\xi,\alpha}(t) = \xi + (\alpha - 1)t + \alpha f(t).$$

If $f_{\xi,\alpha}$ satisfies (h3), i.e., t_ is the smallest zero of $f_{\xi,\alpha}$, then the sequence generated by the Newton method for solving $f_{\xi,\alpha}(t) = 0$, with starting point $t_0 = 0$,*

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \dots,$$

is well defined, $\{t_k\}$ is strictly increasing, contained in $[0, t_)$, and it converges Q -linearly to t_* . Let $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ and $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued convex function with minimizer set C nonempty. Suppose that C is a cone and $x_0 \in \mathbb{R}^n$ satisfies the Robinson condition with respect to C and F . Let $\beta_0 = \|T_{x_0}^{-1}\|$. If $d(F(x_0), C) > 0$, $t_* \leq r_{\beta_0} := \sup\{t \in [0, R) : \beta_0 - 1 + \beta_0 f'(t) < 0\}$,*

$$\Delta \geq \xi \geq \eta\beta_0 d(F(x_0), C), \quad \alpha \geq \frac{\eta\beta_0}{1 + (\eta - 1)\beta_0[f'(\xi) + 1]},$$

then the sequence generated by Algorithm 2.1, denoted by $\{x_k\}$, is contained in $B(x_0, t_)$,*

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

satisfies the inequalities

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

for $k = 0, 1, \dots$, and $k = 1, 2, \dots$, respectively, converging to a point $x_ \in B[x_0, t_*]$ such that $F(x_*) \in C$,*

$$\|x_* - x_k\| \leq t_* - t_k, \quad k = 0, 1, \dots,$$

and the convergence is R -linear. If, additionally, $f_{\xi,\alpha}$ satisfies (h4), then the following inequalities hold:

$$\|x_{k+1} - x_k\| \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} \|x_k - x_{k-1}\|^2, \quad t_{k+1} - t_k \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)} (t_k - t_{k-1})^2$$

for all $k = 1, 2, \dots$. Moreover, the sequences $\{x_k\}$ and $\{t_k\}$ converge Q -quadratically to x_* and t_* , respectively, as follows

$$\limsup_{k \rightarrow \infty} \frac{\|x_* - x_{k+1}\|}{\|x_* - x_k\|^2} \leq \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)}, \quad t_* - t_{k+1} \leq \frac{f''_{\xi, \alpha}(t_*)}{-2f'_{\xi, \alpha}(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \dots$$

Proof. Since $x_0 \in \mathbb{R}^n$ satisfies the Robinson condition, with respect to C and F , we have from Lemma 2.24 that x_0 is a quasi-regular point of the inclusion $F(x) \in C$ with the quasi-regular radius $r_{x_0} \geq r_{\beta_0}$. So, taking into account the assumption $t_* \leq r_{\beta_0}$ we obtain

$$t_* < r_{x_0}.$$

Moreover, Lemma 2.24 also implies that the quasi-regular bound function β_{x_0} satisfies

$$(4.2) \quad \beta_{x_0}(t) \leq \frac{\beta_0}{1 - \beta_0[f'(t) + 1]}, \quad t \in [0, r_{\beta_0}).$$

Since $\Delta \geq \xi \geq \eta\beta_0d(F(x_0), C)$ and the last inequality implies that $\beta_{x_0}(0) \leq \beta_0$, we have

$$\Delta \geq \xi \geq \eta\beta_{x_0}(0)d(F(x_0), C).$$

Now, combining the assumptions $0 < \xi$ and $t_* \leq r_{\beta_0}$ with Proposition 2.13 we conclude that $0 < \xi < t_* \leq r_{\beta_0}$. So, using (4.2), $f'(0) = -1$, f' as strictly increasing and $\eta \geq 1$; after simple algebraic manipulation we obtain

$$\eta[f'(t) + 1] + \frac{1}{\beta_{x_0}(t)} \geq \frac{1}{\beta_0} + (\eta - 1)[f'(t) + 1] \geq \frac{1}{\beta_0} + (\eta - 1)[f'(\xi) + 1], \quad t \in [\xi, t_*),$$

or equivalently,

$$\frac{\eta\beta_0}{1 + (\eta - 1)\beta_0[f'(\xi) + 1]} \geq \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f'(t) + 1] + 1}, \quad t \in [\xi, t_*).$$

Hence, the assumption $\alpha \geq \eta\beta_0/[1 + (\eta - 1)\beta_0(f'(\xi) + 1)]$ and the last inequality imply that

$$\alpha \geq \sup \left\{ \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f'(t) + 1] + 1} : \xi \leq t < t_* \right\}.$$

Therefore, F and x_0 satisfy all assumptions in Theorem 3.1 and so statements of the theorem follow. \square

Remark 4.5. If in Theorem 4.4 we assume that there exist $R > 0$ and $K > 0$ such that

$$\|F'(y) - F'(x)\| \leq K\|y - x\|, \quad x, y \in B(x_0, R),$$

then $f : [0, R) \rightarrow \mathbb{R}$ defined by $f(t) = Kt^2/2 - t$ is a majorant function for F on $B(x_0, R)$ (see Example 2.4) and (h3), (h4), and t_* become

$$2\alpha K\xi \leq 1, \quad 2\alpha K\xi < 1, \quad t_* = (1 - \sqrt{1 - 2\alpha K\xi})/(\alpha K),$$

and α satisfies

$$\alpha \geq \frac{\eta\beta_0}{1 + (\eta - 1)K\beta_0\xi}.$$

In particular, if $C = \{0\}$ and $n = m$, the Robinson condition is equivalent to the condition that $F'(x_0)^{-1}$ is nonsingular. Hence, for $\eta = 1$ Theorem 4.4 becomes a semilocal convergence result for the Newton method under the Lipschitz condition; see [2].

Remark 4.6. If in Theorem 4.4 we assume that F is an analytic function and

$$\gamma := \sup_{n>1} \left\| \frac{F^{(n)}(x_0)}{n!} \right\|^{1/(n-1)} < +\infty,$$

then $f : [0, 1/\gamma] \rightarrow \mathbb{R}$ defined by $f(t) = t/(1 - \gamma t) - 2t$ is a majorant function for the function F on $B(x_0, 1/\gamma)$ (see Example 2.9), (h3), (h4), and t_* become

$$\xi\gamma \leq 1 + 2\alpha - 2\sqrt{\alpha(1 + \alpha)}, \quad \xi\gamma < 1 + 2\alpha - 2\sqrt{\alpha(1 + \alpha)},$$

$$t_* = \frac{1 + \gamma\xi - \sqrt{(1 + \gamma\xi)^2 - 4(1 + \alpha)\gamma\xi}}{2(1 + \alpha)\gamma},$$

and α satisfies

$$\alpha \geq \frac{\eta\beta_0(1 - \gamma\xi)^2}{(\eta - 1)\beta_0 + [1 - \beta_0(\eta - 1)](1 - \gamma\xi)^2}.$$

In particular, if $C = \{0\}$ and $n = m$, the Robinson condition is equivalent to the condition that $F'(x_0)^{-1}$ is nonsingular. Hence, for $\eta = 1$ Theorem 4.4 becomes a semilocal convergence result for the Newton method under the Smale condition; see [14].

5. Conclusion. In the following numerical example from Li and Ng [8], the importance of considering the quasi-regularity bound functions rather than a constant bound is illustrated. More specifically, the next example shows that Theorem 4.2 is not applicable; however, Theorem 4.4, which is based in the quasi-regular bound function, is applicable.

Example 5.1. Let $(\sqrt{3} - 1)/4 < \tau \leq 1/4$ and consider the scalar functions h and F defined by

$$h(x) := |x|, \quad F(x) := \tau - x + x^2, \quad x \in \mathbb{R}.$$

Then, $C = \{0\}$. It is easy to check that $F'(x) = -1 + 2x$, $\|F'(x) - F'(y)\| = 2|x - y|$ for all $x, y \in \mathbb{R}$. Thus, $K = 2$. Let $x_0 = 0$. It is immediate that $T_{x_0}d = -d$ for all $d \in \mathbb{R}$ (see the definition in (2.19)). So, T_{x_0} carries \mathbb{R} onto \mathbb{R} , i.e., x_0 satisfies the Robinson condition and $\beta_0 = \|T_{x_0}^{-1}\| = 1$. Hence, taking $\eta = 1$ and $\Delta = +\infty$, we have

$$d(F(x_0), C) = \tau > 0, \quad \alpha = 1, \quad t_* = \frac{1 - \sqrt{1 - 4\tau}}{2} \leq \frac{1}{2} = r_{\beta_0}$$

and

$$\xi = \eta\beta_0d(F(x_0), C) = \tau \leq \frac{1}{4}.$$

Therefore, we conclude that Theorem 4.4 is applicable with initial point $x_0 = 0$; see also Remark 4.5.

Now, we will show that Theorem 4.2 is not applicable. Indeed, as x_0 satisfies the Robinson condition, using Lemma 2.24 we conclude that $x_0 = 0$ is a regular point for the inclusion (1.2). Hence, Proposition 2.21 implies that there exist $r > 0$ and $\beta > 0$ such that

$$(5.1) \quad D_C(x) \neq \emptyset, \quad d(0, D_C(x)) \leq \beta d(F(x), C), \quad |x| < r.$$

As $K = 2$, $d(F(x_0), C) = \tau$, $\eta = 1$, and $\Delta = +\infty$. Hence, taking $\xi = \eta\beta d(F(x_0), C) = \beta\tau$, $\alpha = \beta/(2\beta\xi + 1)$ to apply Theorem 4.2, it is necessary that

$$(5.2) \quad 2\beta\xi \leq 1, \quad t_* = \frac{1 + 2\beta\xi - \sqrt{1 - (2\beta\xi)^2}}{2\beta} \leq r.$$

Now, conditions in (5.1) imply that $1/(1 - 2|x|) \leq \beta$ for all $|x| < r < 1/2$. So, $1/(1 - 2r) \leq \beta$, or equivalently $2\beta r < \beta - 1$, which combined with the second inequality in (5.2) and simple calculus imply that $(8\xi^2 - 4\xi + 1)\beta^2 + 4(2\xi - 1)\beta + 3 \leq 0$. The last inequality implies $\xi \leq (\sqrt{3} - 1)/4$. Since, $\xi = \beta\tau \geq \tau/(1 - 2r) > \tau > (\sqrt{3} - 1)/4$, we have a contradiction.

REFERENCES

- [1] J. V. BURKE AND M. C. FERRIS, *A Gauss-Newton method for convex composite optimization*, Math. Program., 71 (1995), pp. 179–194.
- [2] J. E. DENNIS, JR., AND R. B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Classics in Appl. Math. 16, SIAM, Philadelphia, 1996.
- [3] O. P. FERREIRA, M. L. N. GONÇALVES, AND P. R. OLIVEIRA, *Local convergence analysis of the Gauss-Newton method under a majorant condition*, J. Complexity, 27 (2011), pp. 111–125.
- [4] O. P. FERREIRA, M. L. N. GONÇALVES, AND P. R. OLIVEIRA, *Local convergence analysis of inexact Gauss-Newton like methods under majorant condition*, J. Comput. Appl. Math., 236 (2012), pp. 2487–2498.
- [5] O. P. FERREIRA AND B. F. SVAITER, *Kantorovich's theorem on Newton's method in Riemannian manifolds*, J. Complexity, 18 (2002), pp. 304–329.
- [6] O. P. FERREIRA AND B. F. SVAITER, *Kantorovich's majorants principle for Newton's method*, Comput. Optim. Appl., 42 (2009), pp. 213–229.
- [7] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Convex Analysis and Minimization Algorithms*. I, Grundlehren Math. Wiss. 305, Springer-Verlag, Berlin, 1993.
- [8] C. LI AND K. F. NG, *Majorizing functions and convergence of the Gauss-Newton method for convex composite optimization*, SIAM J. Optim., 18 (2007) pp. 613–642.
- [9] C. LI AND K. F. NG, *Convergence analysis of the Gauss-Newton method for convex inclusion and convex-composite optimization problems*, J. Math. Anal. Appl., 389 (2012), pp. 469–485.
- [10] C. LI AND X. WANG, *On convergence of the Gauss-Newton method for convex composite optimization*, Math. Program., 91 (2002), pp. 349–356.
- [11] S. M. ROBINSON, *Extension of Newton's method to nonlinear functions with values in a cone*, Numer. Math., 19 (1972), pp. 341–347.
- [12] R. T. ROCKAFELLAR, *Monotone Processes of Convex and Concave Type*, Mem. Amer. Math. Soc. 77, AMS, Providence, RI, 1967.
- [13] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton Math. Ser. 28, Princeton University Press, Princeton, NJ, 1970.
- [14] S. SMALE, *Newton's method estimates from data at one point*, in The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics, Springer, New York, 1986, pp. 185–196.
- [15] X. WANG, *Convergence of Newton's method and uniqueness of the solution of equations in Banach space*, IMA J. Numer. Anal., 20 (2000), pp. 123–134.