Revista Colombiana de Matemáticas
Volumen 39 (2005), páginas 97–112

A variant of Newton's method for generalized equations

Jean-Alexis Célia
Pietrus Alain
Université des Antilles et de la Guyane, France

Abstract. In this article, we study a variant of Newton's method of the following form

$$0 \in f(x_k) + h\nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1}),$$

where f is a function whose Frechet derivative is K-lipschitz, F is a set-valued map between two Banach spaces X and Y and h is a constant. We prove that this method is locally convergent to x^* a solution of

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

if the set-valued map $[f(x^*) + h\nabla f(x^*)(. - x^*) + F(.)]^{-1}$ is Aubin continuous at $(0, x^*)$ and we also prove the stability of this method.

Keywords and phrases. Set–valued mapping, generalized equation, linear convergence, Aubin continuity.

2000 Mathematics Subject Classification. Primary: 49J53, 47H04. Secondary: 65K10.

Resumen. En este artículo estudiamos una variante del método de Newton de la forma

$$0 \in f(x_k) + h\nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1})$$

donde, f es una función cuya derivada de Frechet es K-lipschitz, F es una función entre dos espacios de Banach X y Y cuyos valores son conjuntos y h es una constante. Probamos que este método converge localmente a x^*, una solución de

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

si la aplicación $[f(x^*) + h\nabla f(x^*)(. - x^*) + F(.)]^{-1}$ es Aubin continua en $(0, x^*)$. También probamos la estabilidad del método.

1. Introduction

Throughout this article X and Y are two Banach spaces. We consider a generalized equation of the form

$$0 \in f(x) + F(x) \tag{1}$$

where $f : X \to Y$ is Frechet-differentiable and $F : X \to 2^Y$ is a set-valued map with closed graph. Let us note that the equation (1) is an abstract model for various problems.

- When $F = 0$, (1) is an equation,
- when F is the positive orthant in $\mathbb{R}^m$, (1) is a system of inequalities,
- when F is the normal cone to a convex and closed set in X, (1) may represent variational inequalities.

For others examples, the reader could refer to [4].

To solve (1), in [3] and [4], A.L. Dontchev introduced a Newton type sequence of the form

$$0 \in f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad k = 0, 1 \ldots . \tag{2}$$

where $\nabla f(x_k)$ is the Frechet derivative of f at the point x_k, and he also proved the stability of the method (2). The main tool used for obtaining the convergence which is quadratic is the Aubin continuity of $(f+F)^{-1}$ and the Lipschitz property of the Frechet derivative ∇f.

Following up, in [10], A. Pietrus extended this study to the functions f whose Frechet derivative ∇f satisfies the Hölder condition, he showed that the convergence is superlinear and also proved, in [9], the stability of the method (2) in this mild differentiability context.

Let us remark that when $F = \{0\}$ and x^* is a solution of (1) of order $h > 1$, the method (2) is no longer valid. To avoid this drawback, in [7, 8], the authors proposed a variant of the Newton method of the form

$$x_{k+1} = x_k - h\ \nabla f(x_k)^{-1} f(x_k). \tag{3}$$

Following this work, we introduce to solve (2), the following sequence of the form

$$0 \in f(x_k) + h\nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad k = 0, 1 \ldots \tag{4}$$

Let us remark that when $h = 1$, the method (4) is exactly the Newton type method (2).

This paper is organized as follows : in Section 2, we recall a few preliminary results, in Section 3, we show that the method (4) is locally convergent and in Section 4, we prove the stability of this method. In the sequel, all the norms will be denoted by $||.||$, the distance by *dist* and the ball of center x and of radius r by $B_r(x)$.

2. **Preliminaries**

In this section, we collect some definitions and results that we will need to prove our results.

Definition 2.1. *A set-valued map* $\Gamma : Y \to 2^X$ *is said to be* M*-pseudo-lipschitz around* $(y_0, x_0) \in Graph(\Gamma) := \{(y, x) \in Y \times X, x \in \Gamma(Y)\}$ *if there exist neighborhoods* V *of* y_0 *and* U *of* x_0 *such that*

$$\sup_{x \in \Gamma(y_1) \cap U} dist(x, \Gamma(y_2)) \leq M||y_1 - y_2||, \quad \forall\, y_1,\ y_2 \in V. \tag{5}$$

Let A and C be two subsets of X, we recall that the excess e from the set A to the set C is given by $e(C, A) = \sup_{x \in C} \operatorname{dist}(x, A)$. Then, we have an equivalent definition of M–pseudo-Lipschitz property in terms of excess replacing (5) by

$$e(\Gamma(y_1) \cap U, \Gamma(y_2)) \leq M||y_1 - y_2|| \tag{6}$$

in the previous definition.

In [2], the above property is called the Aubin continuity and the maps satisfying this property are called Aubin continuous. In [5], the above property has been used in order to study the problem of inverse for set-valued maps. For more information about the Aubin continuity, the reader could refer to [1, 2, 11, 12].

Lemma 2.1. *Let* (X, ρ) *be a complete metric space, let* Ψ *be a map from* X *into the closed subsets of* X*, let* $\eta_0 \in X$*,* r *and* λ *be such that* $0 \leq \lambda < 1$ *and*

(a) $dist(\eta_0, \Psi(\eta_0)) < r(1 - \lambda)$.
(b) $e(\Psi(x_1) \cap B_r(\eta_0), \Psi(x_2)) \leq \lambda\rho(x_1, x_2), \quad \forall x_1, x_2 \in B_r(\eta_0)$.

Then Ψ *has a fixed point in* $B_r(\eta_0)$*. That is, there exists* $x \in B_r(\eta_0)$ *such that* $x \in \Psi(x)$*. If* Ψ *is single-valued, then* x *is the unique fixed point of* Ψ *in* $B_r(\eta_0)$*.*

The previous lemma which has been proved in [5] is a generalization of a fixed point theorem in Ioffe–Tikhomirov [6] where in (b) the excess e is replaced by the Hausdorff distance. It's clear that when Ψ is single-valued, the theorem is closed to the Picard fixed point theorem.

3. **Convergence analysis**

From now on, we make the following assumptions (we recall that x^* denotes a solution of (1)):

(H1) $f : X \to Y$ is a function which is Frechet–differentiable in a open neighborhood Ω of x^*.
(H2) The Frechet derivative ∇f of f is K–lipschitz in Ω with K a constant which is strictly positive.
(H3) $F : X \to 2^Y$ is a set-valued map with closed graph.

(H4) The set-valued map $[f(x^*) + h\nabla f(x^*)(. - x^*) + F(.)]^{-1}$ is M–pseudo-lipschitz at $(0, x^*)$ where $h \in$ IR.

Let us remark that the hypothesis (H2) implies that there exists a constant $L > 0$ such that $||\nabla f(x)|| \leq L$, for every $x \in \Omega$, we will use this assumption, in the sequel of our study.

(H5) The constants M,K, L and h are such that $M\left(\frac{K}{2} + |1 - h|L\right) < 1$.

The first theorem of this paper reads as follows :

Theorem 3.1. *Let x^* be a solution of (1) and suppose that the assumptions (H1)-(H5) are satisfied. Then for every c such that $M\left(\frac{K}{2} + |1 - h|L\right) < c < 1$, one can find $\delta > 0$ such that for every starting point $x_0 \in B_\delta(x^*)$, there exists a sequence $(x_k)_k$ for (1) defined by (4) which satisfies :*

$$||x_{k+1} - x^*|| \leq c||x_k - x^*|| \tag{7}$$

that is, the sequence, $(x_k)_k$ is linearly convergent to x^*.

Proof. Before proving theorem 3.1, we need to introduce some notations. First, define the set-valued map P from X into the subsets of Y by

$$P(x) = f(x^*) + h\nabla f(x^*)(x - x^*) + F(x)\,,$$

with $h \in$ IR and the map Ψ_0 for x_0 fixed in X by

$$x \to \Psi_0(x) = P^{-1}[f(x^*) + h\nabla f(x^*)(x - x^*) - f(x_0) - h\nabla f(x_0)(x - x_0)].$$

Then a fixed point x_1 of Ψ_0 checks $x_1 \in \Psi_0(x_1)$, which may be written as follow :

$$f(x^*) + h\nabla f(x^*)(x_1 - x^*) - f(x_0) - h\nabla f(x_0)(x_1 - x_0) \subset P(x_1)\,,$$

and finally

$$0 \in f(x_0) + h\nabla f(x_0)(x_1 - x_0) + F(x_1)\,, \tag{8}$$

i.e., x_1 is a solution of the equation (4).

The induction will consist in starting point x_k to show that the map

$$x \to \Psi_k(x) = P^{-1}[f(x^*) + h\nabla f(x^*)(x - x^*) - f(x_k) - h\nabla f(x_k)(x - x_k)]\,,$$

has a fixed point x_{k+1}.

This fixed point will satisfy a relation which is similar to (8), replacing x_0 by x_k and x_1 by x_{k+1}. So, repeating this algorithm, we will build a sequence $(x_k)_k$ which will converge to x^*.

Now, we state a result which is the starting point of our algorithm. It is an efficient tool to prove theorem 3.1 and reads as follows.

Proposition 3.1. *Under the assumptions of theorem 3.1, there exists $\delta > 0$ such that for all $x_0 \neq x^*$ and $x_0 \in B_\delta(x^*)$, the map $\Psi_0(x) = P^{-1}[f(x^*) + h\nabla f(x^*)(x - x^*) - f(x_0) - h\nabla f(x_0)(x - x_0)]$ has a fixed point $x_1 \in B_\delta(x^*)$.*

Proof. Without loss of generality, we can suppose that the diameter of Ω noted $\mathrm{diam}(\Omega)$ is less than 1. By hypothesis (H4), there exist positive numbers a and b such that

$$e(P^{-1}(y') \cap B_a(x^*), P^{-1}(y'')) \leq M||y' - y''||, \quad \forall\, y',\ y'' \in B_b(0). \tag{9}$$

fix $\delta > 0$ such that

$$\delta < \min\left\{ a, \frac{-|1-h|L + \sqrt{L^2|1-h|^2 + 2bK\,(1+2|h|)}}{K\,(1+2|h|)}, \frac{1}{M|h|K} \right\}. \tag{10}$$

To prove proposition 3.1, we will show that both assertions (a) and (b) of Lemma 2.1 hold, where $\eta_0 = x^*$, Ψ is the function Ψ_0 defined at the very beginning of this section and where r and *lambda* are numbers to be set.

According to the definition of the excess e, we have

$$\begin{aligned}&\mathrm{dist}(x^*, \Psi_0(x^*)) \leq \\ &\qquad e\left(P^{-1}(0) \cap B_\delta(x^*),\ P^{-1}\left[f(x^*) - f(x_0) - h\triangledown f(x_0)(x^* - x_0)\right]\right).\end{aligned} \tag{11}$$

Moreover, for all $x_0 \in B_\delta(x^*)$, we have

$$\begin{aligned}&||f(x^*) - f(x_0) - h\triangledown f(x_0)(x^* - x_0)|| \\ &\qquad \leq ||f(x^*) - f(x_0) - \triangledown f(x_0)(x^* - x_0)|| + ||(1-h)\triangledown f(x_0)(x^* - x_0)|| \\ &\qquad \leq \frac{K}{2}||x^* - x_0||^2 + |1-h|L||x^* - x_0|| \\ &\qquad \leq \frac{K}{2}\delta^2 + |1-h|L\delta.\end{aligned}$$

Thanks to (10), we have $\frac{K}{2}\delta^2 + |1-h|L\delta < b$, which implies that $(f(x^*) - f(x_0) - h\triangledown f(x_0)(x^* - x_0)) \in B_b(0)$. Combining the last remark, the inequality (9) and the definition of Ψ_0, we get

$$\begin{aligned}\mathrm{dist}\,(x^*,\ ,\Psi_0(x^*)) &\leq M\,||f(x^*) - f(x_0) - h\triangledown f(x_0)(x^* - x_0)|\,| \\ &\leq M\left(\frac{K}{2}||x^* - x_0|| + |1-h|L\right)||x^* - x_0||.\end{aligned}$$

fix c such that $c \geq M\left(\frac{K}{2} + |1-h|L\right)$. Setting $\lambda = M|h|K\delta$, then, $\lambda \in]0,1[$ and one can find δ such that $M\left(\frac{K}{2} + |1-h|L\right) \leq c(1-\lambda)$ since $\mathrm{diam}(\Omega) < 1$.

Hence,

$$\mathrm{dist}(x^*\,, \Psi_0(x^*)) \leq c(1-\lambda)||x^* - x_0||.$$

By setting $\eta_0 = x^*$ and $r = r_0 = M\left(\frac{K}{2}||x^* - x_0|| + |1-h|L\right)||x^* - x_0||$. We can deduce from the last inequalities that assertion (a) of Lemma 2.1 is satisfied.

Now, we show that condition (b) of Lemma 2.1 holds. By (10), we have $r_0 \leq \delta \leq a$. Moreover, for $x \in B_\delta(x^*)$, setting

$$y = f(x^*) + h\triangledown f(x^*)(x - x^*) - f(x_0) - h\triangledown f(x_0)(x - x_0)$$

we have

$$\begin{aligned}
||y|| &\leq ||f(x^*) - f(x_0) - \nabla f(x_0)(x^* - x_0)|| + ||(1-h)\nabla f(x_0)(x^* - x_0)|| \\
&\quad + ||h(\nabla f(x^*) - \nabla f(x_0))(x - x^*)|| \\
&\leq \frac{K}{2}||x^* - x_0||^2 + |1-h|L||x^* - x_0|| + |h|K||x^* - x_0||||x - x^*|| \\
&\leq \left(\frac{K}{2} + |h|K\right)\delta^2 + |1-h|L\delta.
\end{aligned}$$

Then, by (10), we can deduce that for all $x \in B_\delta(x^*)$, $y \in B_b(0)$. It follows that for all $x', x'' \in B_{r_0}(x^*)$ where $r_0 \leq \delta$, we have

$$e\left(\Psi_0(x') \cap B_{r_0}(x^*), \Psi_0(x'')\right) \leq e\left(\Psi_0(x') \cap B_\delta(x^*), \Psi_0(x'')\right)$$

which yields, by (9),

$$\begin{aligned}
e\left(\Psi_0(x') \cap B_{r_0}(x^*), \Psi_0(x'')\right) &\leq M||h\nabla f(x^*)(x' - x'') - h\nabla f(x_0)(x' - x'')|| \\
&\leq M|h|K||x^* - x_0||\ ||x' - x''|| \\
&\leq M|h|K\delta||x' - x''|| \\
&\leq \lambda||x' - x''||
\end{aligned}$$

Thus condition (b) of Lemma 2.1 is satisfied.

Since both conditions of Lemma 2.1 are fulfilled, we can deduce the existence of a fixed point $x_1 \in B_{r_0}(x^*)$ for the map Ψ_0. The fact that $r_0 < \delta$ completes the proof of Proposition 3.1. ☑

According to a previous remark, x_1 is obtained by the equation (4) starting with x_0 and $x_1 \in B_{r_0}(x^*)$. Thus, we have

$$||x_1 - x^*|| \leq r_0 = M\left(\frac{K}{2}||x^* - x_0|| + |1-h|L\right)||x^* - x_0||.$$

Now that we proved proposition 3.1, the proof of theorem 3.1 is straightforward as it is shown below:

Proof of theorem 3.1. Proceeding by induction, keeping $\lambda = M|h|K\delta$ $\eta_0 = x^*$ and setting $r_k = M\left(\frac{K}{2}||x_k - x^*|| + |1-h|L\right)||x_k - x^*||$, the application of proposition 3.1 to the map Ψ_k gives the existence of a fixed point x_{k+1} for Ψ_k, which is an element of $B_{r_k}(x^*)$.

This last fact implies that

$$||x_{k+1} - x^*|| \leq r_k$$

and proceeding as in the proof of the proposition 3.1, we can choose $c < 1$ such that $||x_{k+1} - x^*|| \leq c||x_k - x^*||$. Hence, the proof of theorem 3.1 is complete. ☑

As an illustration of our general results, let us consider the following nonlinear programming problem:

minimize $f_0(x)$

subject to

$$\begin{cases} f_i(x) = 0, & i = 1, \cdots, m \\ f_i(x) \leq 0, & i = m+1, \cdots, p \end{cases}$$

where the function $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \cdots, p$ are twice continuously differentiable on $\mathbb{R}^n$. The lagragian L associated with the minimization problem is defined by

$$L : (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^p \mapsto f_0(x) + \sum_{i=1}^{p} \lambda_i f_i(x),$$

then, the Karush-Kuhn-Tucker first order optimality conditions read as follows:

$$\nabla_x L(x, \lambda) = 0 \tag{12}$$

$$\nabla_\lambda L(x, \lambda) \in N_\Lambda(\lambda) \tag{13}$$

where $N_\Lambda(\lambda)$ denotes the normal cone to the set $\Lambda = \mathbb{R}^m \times \mathbb{R}_+^{p-m}$ at the point λ. Then, it is easy to see that conditions (12) and (13) amount to

$$0 \in (\nabla_x L(x, \lambda), -\nabla_\lambda L(x, \lambda)) + N_C(x, \lambda) \tag{14}$$

where $C = \mathbb{R}^n \times \Lambda$. Moreover, relation (14) can be reformulated in the following way:

$$0 \in f(x, \lambda) + F(x, \lambda), \tag{15}$$

where $f(x, \lambda) = (\nabla_x L(x, \lambda), -\nabla_\lambda L(x, \lambda))$ and $F(x, \lambda) = N_C(x, \lambda)$. Hence, Karush-Kuhn-Tucker optimality system (12),(13) is equivalent to (14) which is a generalized equation of the form of (1) and then can be studied using the method presented in this paper.

4. Stability of the method

Now, we prove that the hypothesis (H4) of Aubin continuity is sufficient to obtain the stability of the method. In fact, this section is about the method for solving equations involving set-valued maps and parameters.

We consider the generalized equation of the following form:

$$y \in f(x) + F(x) \tag{16}$$

where y is a parameter, f is a function which is Frechet-differentiable and F is a set-valued map.

The second important result of this paper follows.

Theorem 4.1. *Let x^* be a solution of (1), and suppose that the assumptions (H1)-(H5) are satisfied except (H2) that we replace by a weaker assumption* (H'2). *The Frechet derivative ∇f of f is continuous in Ω. Then, there exist positive constants σ, b and c such that for every $y \in B_b(0)$ and for every*

$x_0 \in B_\sigma(x^*)$, there exists a sequence $(x_k)_k$ starting to x_0 which converges to a solution x of (16).

Moreover, if x_0 is a solution of (16) for $y = y_0$ then the limit x satisfies

$$||x - x_0|| \leq c||y - y_0||.$$

Proof. To prove this theorem, we will show the following Lemma.

Lemma 4.1. *Let $(x^*, y^*) \in Graph(f + F)$ and suppose that the assumptions (H1), (H'2) and (H3) are satisfied. If the map $P_{x^*}(.) = [f(x^*) + h\triangledown f(x^*)(. - x^*) + F(.)]^{-1}$ is Aubin-continuous at (y^*, x^*) then there exist positive constants α, β, M such that for every $x \in B_\alpha(x^*)$,*

$$e\left(P_x(y') \cap B_\alpha(x^*), P_x(y'')\right) \leq M\,||y' - y''|\,|, \quad \forall y', y'' \in B_\beta(y^*).$$

Proof. The map $P_{x^*}(.)$ is Aubin-continuous at (y^*, x^*) let a, b and M' be the associated constants.

Choose $\epsilon > 0$ such that $M'\epsilon|h| < 1$ and $\alpha > 0$ such that $||\triangledown f(x) - \triangledown f(x^*)|| < \epsilon$, for every $x \in B_\alpha(x^*)$. Moreover, by hypothesis (H'2), there exists $L > 0$ such that $||\triangledown f(x)|| \leq L$, for every $x \in B_\alpha(x^*)$.

Take $\alpha > 0$ smaller if necessary so that $2\alpha \leq a$ and $\alpha\left(\epsilon + |1 - h|L + 3\epsilon|h|\right) < b$. Further, choose $\beta > 0$ such that

$$\beta + \alpha\left(\epsilon + 3\epsilon|h| + |1 - h|L\right) \leq b \text{ and } \frac{2M'\beta}{1 - M'\epsilon h} \leq \alpha. \tag{17}$$

Let $x \in B_\alpha(x^*)$, let $y', y'' \in B_\beta(y^*)$ and let $x' \in P_x(y') \cap B_\alpha(x^*)$. Denote $x_1 = x'$. Then

$$x_1 \in P_{x^*}\left(y' - f(x) - h\triangledown f(x)(x_1 - x) + f(x^*) + h\triangledown f(x^*)(x_1 - x^*)\right) \cap B_a(x^*)\,,$$

and

$$||x - x_1|| \leq ||x - x^*|| + ||x^* - x_1|| \leq 2\alpha.$$

Using (17), we obtain

$$\begin{aligned}
&||y' - f(x) - h\triangledown f(x)(x_1 - x) + f(x^*) + h\triangledown f(x^*)(x_1 - x^*) - y^*|| \\
&\quad = ||y^* - y' + f(x) - f(x^*) - \triangledown f(x^*)(x - x^*) \\
&\qquad + (1 - h)\triangledown f(x^*)(x - x^*) - h\triangledown f(x^*)(x_1 - x) + h\triangledown f(x)(x_1 - x)|| \\
&\quad \leq ||y' - y^*|| + ||f(x) - f(x^*) - \triangledown f(x^*)(x - x^*)|| \\
&\qquad + |1 - h|\; ||\triangledown f(x^*)||\; ||x - x^*|| + |h|\; ||\triangledown f(x^*) - \triangledown f(x)|||||(x - x_1)|| \\
&\quad \leq \beta + \epsilon||x - x^*|| + |1 - h|\; L\; ||x - x^*|| + |h|\; \epsilon\; ||x - x_1|| \\
&\quad \leq \beta + \alpha(\epsilon + 2|h|\; \epsilon + |1 - h|\mathrm{L}) \leq b\,;
\end{aligned}$$

the same inequality holds for y''.

For these estimates and from the Aubin-continuity of P_{x^*}, we obtain that there exists an element $x_2 \in P_{x^*}\big(y'' - f(x) - h\triangledown f(x)(x_1 - x) + f(x^*) + h\triangledown f(x^*)$ $(x_1 - x^*)\big)$ that is,

$$y'' \in f(x) + h\triangledown f(x)(x_1 - x) + h\triangledown f(x^*)(x_2 - x_1) + F(x_2)\,,$$

and such that

$$||x_2 - x_1|| \le M'||y' - y''||.$$

Proceeding by induction, suppose that there exist an integer $n > 2$ and points $x_2, x_3, \dots, x_n$ with

$$y' \in f(x) + h\nabla f(x)(x_{i-1} - x) + h\nabla f(x^*)(x_i - x_{i-1}) + F(x_i)\,,$$

and

$$\begin{aligned} ||x_i - x_{i-1}|| &\le (M'\epsilon|h|)^{i-2}\ \ ||x_2 - x_1||, \quad i = 3, 4, \dots, n. \\ &\le (M'\epsilon|h|)^{i-2}\ \ M'||y' - y''||. \end{aligned}$$

Then

$$\begin{aligned} ||x_n - x^*|| &\le \sum_{j=2}^{n} ||x_j - x_{j-1}|| + ||x_1 - x^*|| \\ &\le \sum_{j=2}^{n} (M'\epsilon|h|)^{j-2}||x_2 - x_1|| + \alpha \\ &\le \sum_{j=2}^{n} (M'\epsilon|h|)^{j-2}M'||y' - y''|| + \alpha \\ &\le \frac{2M'\beta}{1 - M'\epsilon|h|} + \alpha \\ &\le 2\alpha \qquad \text{according to (17).} \end{aligned}$$

We obtain for both $y = y'$ and $y = y''$,

$$\begin{aligned} &||y - f(x) - h\nabla f(x)(x_n - x) + f(x^*) + h\nabla f(x^*)(x_n - x^*) - y^*|| \\ &\quad\le\ \beta + \alpha(\epsilon + 3|h|\ \epsilon + |1 - h|\text{Ł}) \\ &\quad\le b\,. \end{aligned}$$

Then there exists an

$$x_{n+1} \in\ P_{x^*}\left(y'' - f(x) - h\nabla f(x)(x_n - x) + f(x^*) + h\nabla f(x^*)(x_n - x^*)\right)$$

that is,

$$y'' \in\ f(x) + h\nabla f(x)(x_n - x) + h\nabla f(x^*)(x_{n+1} - x_n) + F(x_{n+1})\,, \tag{18}$$

such that

$$\begin{aligned} ||x_{n+1} - x_n|| &\le M'\ ||h\nabla f(x)(x_{n-1} - x_n) + h\nabla f(x^*)(x_n - x_{n-1})|| \\ &\le\ M'|h|\ ||(\nabla f(x) - \nabla f(x^*))(x_n - x_{n-1})|| \\ &\le\ M'|h|\ \epsilon\ ||x_n - x_{n-1}|| \\ &\le\ M'|h|\ \epsilon\ (M'\epsilon|h|)^{n-2}M'||y' - y''|| \\ &\le\ M'||y' - y''||\ |h|\ (M'\epsilon|h|)^{n-1}\,, \end{aligned}$$

this induction step is complete.

Thus $(x_n)_n$ is a Cauchy sequence, hence there exists x'' such that $x_n \to x''$ as $n \to \infty$. Moreover, passing to the limit in (18), we get $x'' \in P_x(y'')$ and

$$\begin{aligned}
||x' - x''|| &\leq \lim_{n\to\infty} \sup \sum_{i=2}^{n} ||x_i - x_{i-1}|| \\
&\leq \lim_{n\to\infty} \sup \sum_{i=2}^{n} (M'\epsilon|h|)^{i-2} ||x_2 - x_1|| \\
&\leq \frac{M'}{1 - M'\epsilon|h|} \, ||y' - y''|| \, .
\end{aligned}$$

Hence, the lemma holds with $M = \dfrac{M'}{1 - M'\epsilon|h|}$. ☑

Proof of theorem 4.1. Let α, β and M be the constants in Lemma 4.1 and let $Q_x(.) = [f(x) + h\nabla f(x)(. - x) + F(.)]^{-1}$, for $x \in B_\alpha(x^*)$ and $h \in \mathbb{R}$. Let $\epsilon > 0$ satisfy $M\epsilon < 1$ and choose $a > 0$ such that $B_a(x^*) \subset \Omega$and $||\nabla f(x') - \nabla f(x'')|| \leq \epsilon$,whenever $x', x'' \in B_a(x^*)$. Let $L > 0$ be such that $||\nabla f(x)|| \leq L$, for $x \in B_a(x^*)$. We can take ϵ in such way that $M(\epsilon + |1 - h|L) < 1$. Choose $\sigma > 0$ such that

$$\sigma \leq \alpha \quad \text{and} \quad 2(\epsilon + |1 - h| \; L)\sigma < \beta \tag{19}$$

and let $b > 0$ satisfy

$$\begin{aligned}
&b\,(1 + M\;(\epsilon + |1 - h|\; L)) + 2\,(\epsilon + |1 - h|\; L)\,\sigma \leq \beta \qquad \text{and} \\
&\frac{Mb + 2\sigma}{1 - M\,(\epsilon + |1 - h|\; L)} \leq \; a.
\end{aligned} \tag{20}$$

Let $x_0 \in B_\sigma(x^*)$. Then

$$x^* \in \; Q_{x_0}\,(-f(x^*) + f(x_0) + h\nabla f(x_0)(x^* - x_0)) \cap B_\alpha(x_0)\,.$$

Further,

$$\begin{aligned}
&||f(x^*) - f(x_0) - h\nabla f(x_0)(x^* - x_0)|| \\
&\qquad \leq ||f(x^*) - f(x_0) - \nabla f(x_0)(x^* - x_0)|| + |1 - h| \; ||\nabla f(x_0)|| \; ||x^* - x_0|| \\
&\qquad \leq \epsilon||x_0 - x^*|| + |1 - h| \; L||x^* - x_0|| \\
&\qquad \leq \sigma(\epsilon + |1 - h| \; L) \leq \beta, \qquad \text{according to (19).}
\end{aligned}$$

Let $y \in B_b(0)$. From Lemma 4.1, there exists $x_1 \in Q_{x_0}(y)$, i.e,

$$y \in \; f(x_0) + h\nabla f(x_0)(x_1 - x_0) + F(x_1)\,,$$

such that

$$\begin{aligned}
||x_1 - x^*|| &\leq M||y + f(x^*) - f(x_0) - h\nabla f(x_0)(x^* - x_0)|| \\
&\leq M\big[||y|| + ||f(x^*) - f(x_0) - \nabla f(x_0)(x^* - x_0)|| \\
&\quad + |1-h|\ ||\nabla f(x_0)||\ ||x^* - x_0||\big] \\
&\leq M\left[b + \epsilon\ ||x_0 - x^*|| + |1-h|\ L\ ||x_0 - x^*||\right] \\
&\leq Mb + M\sigma(\epsilon + |1-h|\ L) \\
&\leq Mb + \sigma, \qquad \text{according to } M(\epsilon + |1-h|\ L) < 1, \ \text{that is} \\
||x_1 - x^*|| &\leq a\,, \qquad \text{according to (20).}
\end{aligned}$$

Then,

$$||x_1 - x_0|| \leq ||x_1 - x^*|| + ||x^* - x_0|| \leq \ Mb + M\sigma(\epsilon + |1-h|\ L) + \sigma. \tag{21}$$

Note that

$$x_1 \in \ Q_{x_1}\left(y + f(x_1) - f(x_0) - h\nabla f(x_0)(x_1 - x_0)\right) \cap B_\alpha(x_1)\,.$$

Thanks to $M(\epsilon + |1-h|L) < 1$ and (20), we have

$$\begin{aligned}
&||y + f(x_1) - f(x_0) - h\nabla f(x_0)(x_1 - x_0)|| \\
&\quad \leq ||y|| + \epsilon\ ||x_1 - x_0|| + |1-h|\ L\ ||x_1 - x_0|| \\
&\quad \leq b + (\epsilon + |1-h|\ L)\ ||x_1 - x_0|| \\
&\quad \leq b + (\epsilon + |1-h|\ L)(Mb + M(\epsilon + |1-h|\ L)\sigma + \sigma) \\
&\quad \leq b(1 + M(\epsilon + |1-h|\ L)) + (\epsilon + |1-h|\ L)(M(\epsilon + |1-h|\ L)\sigma + \sigma) \\
&\quad \leq b(1 + M(\epsilon + |1-h|\ L)) + 2\sigma(\epsilon + |1-h|\ L) \\
&\quad \leq \ \beta\,.
\end{aligned}$$

Then, from Lemma 4.1, there exists an $x_2 \in \ Q_{x_1}(y)$ such that

$$\begin{aligned}
||x_2 - x_1|| &\leq M||f(x_1) - f(x_0) - h\nabla f(x_0)(x_1 - x_0)|| \\
&\leq M(\epsilon + |1-h|\ L)\ ||x_1 - x_0||\,.
\end{aligned}$$

Further,

$$\begin{aligned}
||x_2 - x^*|| &\leq ||x_2 - x_1|| + ||x_1 - x_0|| + ||x_0 - x^*|| \\
&\leq [1 + M(\epsilon + |1-h|\ L)]||x_1 - x_0|| + \sigma\,,
\end{aligned}$$

now,

$$
\begin{aligned}
&\big(1+M(\epsilon+|1-h|L)\big)\,(Mb+M(\epsilon+|1-h|L)\sigma+\sigma)\\
&\quad-\frac{Mb+M(\epsilon+|1-h|\ L)\sigma+\sigma}{1-M(\epsilon+|1-h|L)}\\
&=\frac{[1+M(\epsilon+|1-h|L)][1-M(\epsilon+|1-h|L)][Mb+M(\epsilon+|1-h|L)\sigma+\sigma]}{1-M(\epsilon+|1-h|\ L)}\\
&\quad-\frac{Mb+M(\epsilon+|1-h|L)\sigma+\sigma}{1-M(\epsilon+|1-h|L)}\\
&=\frac{[1-M^2(\epsilon+|1-h|L)^2-1][Mb+M(\epsilon+|1-h|L)\sigma+\sigma]}{1-M(\epsilon+|1-h|L)}\\
&=-\frac{M^2[\epsilon+|1-h|L]^2][Mb+M(\epsilon+|1-h|L)\sigma+\sigma]}{1-M(\epsilon+|1-h|L)}\\
&\leq\ 0\,.
\end{aligned}
$$

Then using (20), we obtain

$$
\begin{aligned}
||x_2-x^*|| &\leq \frac{Mb+M(\epsilon+|1-h|L)\sigma+\sigma}{1-M(\epsilon+|1-h|L)}+\sigma \\
&\leq \frac{Mb+M(\epsilon+|1-h|L)\sigma+\sigma+\sigma-M(\epsilon+|1-h|\ L)\sigma}{1-M(\epsilon+|1-h|L)}\\
&\leq \frac{Mb+2\sigma}{1-M(\epsilon+|1-h|\ L)}\\
&\leq a\,.
\end{aligned}
\tag{22}
$$

Suppose that for some integer $n>2$, the points $x_2, x_3, \ldots, x_n$ are obtained by the method (4) in which 0 is replaced by y, that is, $x_i \in Q_{x_{i-1}}(y)$ and

$$||x_i-x_{i-1}|| \leq (M(\epsilon+|1-h|\ L))^{i-1}\,||x_1-x_0||, \quad i=3,4,\ldots,n.$$

Then, by repeating the argument in (22), we obtain that $x_i \in B_a(x^*)$, for $i=3,4,\ldots,n$. Further, we have

$$
\begin{aligned}
&||y+f(x_n)-f(x_{n-1})-h\triangledown f(x_{n-1})(x_n-x_{n-1})||\\
&\quad\leq\ b+\epsilon\ ||x_n-x_{n-1}||+|1-h|\ L\ ||x_n-x_{n-1}||\\
&\quad\leq\ b+(\epsilon+|1-h|\ L\ )\ (M(\epsilon+|1-h|\text{L}))^{n-1}\ ||x_1-x_0||\\
&\quad\leq\ b+(\epsilon+|1-h|\ L\ )\ ||x_1-x_0||\,, \qquad \text{according to } M(\epsilon+|1-h|\ L)<1\\
&\quad\leq\ b+(\epsilon+|1-h|\ L\ )\ (Mb+M(\epsilon+|1-h|\ L)\sigma+\sigma)\\
&\quad\leq\ \beta\,,
\end{aligned}
$$

(see front).

Then from,

$$x_n \in Q_{x_n}\,(y+f(x_n)-f(x_{n-1})-h\triangledown f(x_{n-1})(x_n-x_{n-1}))\cap B_\alpha(x_n) \tag{23}$$

and from Lemma 4.1, we conclude that there exists a sequence

$$x_{n+1} \in Q_{x_n}(y) \tag{24}$$

satisfying

$$\begin{aligned} ||x_{n+1} - x_n|| &\leq M||f(x_n) - f(x_{n-1}) - h\nabla f x_{n-1})(x_n - x_{n-1})|| \\ &\leq M\epsilon\, ||x_n - x_{n-1}|| + M\, |1-h|\, L\, ||x_n - x_{n-1}|| \\ &\leq M(\epsilon + |1-h|\, L)\, ||x_n - x_{n-1}|| \\ &\leq M(\epsilon + |1-h|\, L)\, \left(M(\epsilon + |1-h|\, L\right)^{n-1}\, ||x_1 - x_0|| \\ &\leq \left(M(\epsilon + |1-h|\, L\right)^{n}\, ||x_1 - x_0||\,. \end{aligned}$$

Then, there exists a sequence $(x_n)_n$ satisfying (4) when 0 is replaced by y, (x_n) is a Cauchy sequence, and, passing to the limit in (24), we obtain that $(x_n)_n$ is linearly convergent to a solution x to (16).

Let $y_0 \in B_b(0)$ and $x_0 \in (f+F)^{-1}(y_0) \cap B_\sigma(x^*)$. Then $x_0 \in Q_{x_0}(y) \cap B_\alpha(x^*)$. From Lemma 4.1, we obtain that there exists $x_1 \in Q_{x_0}(y)$ such that

$$||x_1 - x_0|| \;\leq\; M||y - y_0||.$$

By repeating the argument between (21) and (24), we obtain a sequence $(x_n)_n$ satisfying (23) and (24) and which converges to a solution $x \in \left(f+F\right)^{-1}(y)$.

Moreover

$$\begin{aligned} ||x_n - x_0|| &\leq \sum_{i=1}^{n} ||x_i - x_{i-1}|| \\ &\leq \sum_{i=1}^{n} \left(M(\epsilon + |1-h|\, L)\right)^{i-1}\, ||x_1 - x_0|| \\ &\leq \frac{M}{1 - M(\epsilon + |1-h|\, L)}\, ||y - y_0||\,. \end{aligned}$$

Passing to the limit with n and taking $c = \dfrac{M}{1 - M(\epsilon + |1-h|L)}$, we complete the proof. ☑

The last result of this paper is the following theorem.

Theorem 4.2. *Let x^* be a solution of (1) and suppose that the assumptions (H1), (H'2), (H3) and (H4) are satisfied. Then there exist positive constants σ, b, and $\alpha_1 < 1$ such that for every $y \in B_b(0)$, for every $x_0 \in B_\sigma(x^*)$, there exists a Newton sequence $(x_k)_k$ starting from x_0 which linearly converges to a solution $x \in (f+F)^{-1}(y)$, i.e*

$$||x_{k+1} - x|| \;\leq\; \alpha_1 ||x_k - x||\,.$$

Proof. Let $P_{x^*}(.) = [f(x^*) + h\nabla f(x^*)(. - x^*) + F(.)]^{-1}$ be a set-valued map which is Aubin-continuous at $(0, x^*) \in \mathit{Graph}(P_{x^*})$ with modulus c. Then, from the definition, there exist constants $\delta > 0$ such that for every $y \in B_\delta(0)$, there exists $x \in P_x(y) \cap B_{c||y||}(x^*)$, that is, $x \in (f+F)^{-1}(y) \cap B_{c||y||}(x^*)$.

Let α, β, M be the constants in the statement of Lemma 4.1 and K the Lipschitz constant of ∇f on Ω. Without loss of generality, we can suppose that $\alpha < 1$.

Choose positive constants σ and b such that $B_{cb}(x^*) \subset \Omega$, and

(i) $\sigma \leq \frac{\alpha}{2}$,
(ii) $b \leq \min\{\frac{\beta}{2},\ \delta\}$,
(iii) $cb \leq \frac{\alpha}{2}$,
(iv) $cb + \sigma \leq \min\left\{\frac{2(-M|1-h|L+1)}{MK}, \frac{-M|1-h|L+\sqrt{M^2L^2|1-h|^2+MK\alpha}}{MK}, \frac{-|1-h|L+\sqrt{L^2|1-h|^2+K\beta}}{K}\right\}$.

Let $x_0 \in B_\sigma(x^*)$, $y \in B_b(0)$ and let $x \in (f+F)^{-1}(y) \cap B_{c||y||}(x^*)$. Then $||x - x^*|| \leq cb \leq \alpha$. Note that

$$x \in P_{x_0}\left(y - f(x) + f(x_0) + h\nabla f(x_0)(x - x_0)\right) \cap B_\alpha(x^*)$$

and

$$\begin{aligned}
&||y - f(x) + f(x_0) + h\nabla f(x_0)(x - x_0)|| \\
&\quad \leq ||y|| + ||f(x) - f(x_0) - \nabla f(x_0)(x - x_0)|| + |1-h|\ ||\nabla f(x_0)||\ ||x - x_0|| \\
&\quad \leq\ b + \frac{K}{2}\ ||x - x_0||^2 + |1-h|\ L\ ||x - x_0|| \\
&\quad \leq\ b + \frac{K}{2}\ (cb+\sigma)^2 + |1-h|\ L\ (cb+\sigma).
\end{aligned}$$

From (iv), it follows that $\frac{K}{2}(cb+\sigma)^2 + |1-h|L(cb+\sigma) - \frac{\beta}{2} \leq 0$ and using (ii), we obtain $||y - f(x) + f(x_0) + h\nabla f(x_0)(x - x_0)|| \leq \beta$. Now, from Lemma 4.1, there exists $x_1 \in P_{x_0}(y)$, i.e

$$y \in f(x_0) + h\nabla f(x_0)(x_1 - x_0) + F(x_1)$$

such that

$$\begin{aligned}
||x - x_1|| \leq\ & M||f(x) - f(x_0) - h\nabla f(x_0)(x - x_0)|| \\
\leq\ & M||f(x) - f(x_0) - \nabla f(x_0)(x - x_0)|| \\
& + M|1-h|\ ||\nabla f(x_0)||\ ||x - x_0|| \\
\leq\ & M\left[\frac{K}{2}\ ||x - x_0||^2 + |1-h|\ L\ ||x - x_0||\right].
\end{aligned}$$

Hence,

$$\begin{aligned}||x_1 - x^*|| \le\ & ||x_1 - x|| + ||x - x^*|| \\ \le\ & M[\frac{K}{2}\ ||x - x_0||^2 + |1-h|\ L\ ||x - x_0||\ ] + cb \\ \le\ & M[\frac{K}{2}\ (cb+\sigma)^2 + |1-h|\ L\ (cb+\sigma)\ ] + cb\,.\end{aligned}$$

From (iv), it follows that $M\left[\frac{K}{2}\ (cb+\sigma)^2 + |1-h|\ L\ (cb+\sigma)\right] - \frac{\alpha}{2}\ \le\ 0$ and using (iii), we obtain $||x_1 - x^*||\ \le\ \alpha$.

Further, $x \in\ P_{x_1}\,(y - f(x) + f(x_1) + h\nabla f(x_1)(x - x_1)) \cap B_\alpha(x^*)$ and from

$$\begin{aligned}||x - x_1|| \le\ & \frac{M}{2}K(cb+\sigma)^2 + M\ |1-h|\ L(cb+\sigma) \\ \le\ & \frac{\alpha}{2} \\ \le\ & \alpha\,,\end{aligned}$$

according to (iv) we have

$$\begin{aligned}&||y - f(x) + f(x_1) + h\nabla f(x_1)(x - x_1)|| \\ &\quad \le ||y|| + ||f(x) - f(x_1) - \nabla f(x_1)(x - x_1)|| + |1-h|\ ||\nabla f(x_1)||\ ||x - x_1|| \\ &\quad \le b + \frac{K}{2}\left(\frac{MK}{2}(cb+\sigma)^2 + M\ |1-h|\ L(cb+\sigma)\right)^2 \\ &\qquad + |1-h|\ L\left(\frac{MK}{2}(cb+\sigma)^2 + M\ |1-h|\ L(cb+\sigma)\right) \\ &\quad \le b + \frac{K}{2}(cb+\sigma)^2 + |1-h|\ L(cb+\sigma), \qquad \text{according to (iv)} \\ &\quad \le \frac{\beta}{2} + \frac{\beta}{2} =\ \beta\,, \qquad \text{according to (ii) and (iv).}\end{aligned}$$

Then, there exists $x_2 \in\ P_{x_1}(y)$ with

$$\begin{aligned}||x_2 - x|| \le\ & M||f(x) - f(x_1) - h\nabla f(x_1)(x - x_1)|| \\ \le\ & \frac{MK}{2}\ ||x - x_1||^2 + M\ |1-h|\ L\ ||x - x_1||\,.\end{aligned}$$

We have $\alpha_1 = \frac{MK}{2}(cb+\sigma) + M|1-h|L \le 1$ according to (iv). This implies that $||x_2 - x||\ \le\ \alpha_1||x_1 - x||$.

Proceeding by induction, we complete the proof of theorem 4.2. ☑

References

[1] J-P. Aubin, Lipschitz behavior of solutions to convex minimization problems, *Math. Oper. Res.* **9** (1984), 87–111.

[2] J-P. Aubin & H. Frankowska, *Set-valued Analysis*, Birkhäuser, Boston, 1990.

[3] A. L. Dontchev, Local convergence of the newton method for generalized equations, *C. R. Acad. Sc.* **1** (1996), 327–329.
[4] A. L. Dontchev, Uniform convergence of the newton method for Aubin continuous maps, *Serdica. Math. J.* **22** (1996), 385–398.
[5] A. L. Dontchev & W. W. Hager, An inverse mapping theorem for set-valued maps, *Proc. Amer. Math. Soc.* **121** (1994), 481–489.
[6] A. D. Ioffe & V. M. Tikhomirov, *Theory of Extremal Problems*, North Holland, Amsterdam, 1979.
[7] J. M. Ortega & W. C Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New-York and London, 1970.
[8] A. M Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, Academic Press, New-York and London, 1970.
[9] A. Pietrus, Does Newton's method for set-valued maps converges uniformly in mild differentiability context?, *Rev. Colombiana Mat.* **32** (2000), 49–56.
[10] A. Pietrus, Generalized equations under mild differentiability conditions, *Rev. S. Acad. Cienc. Exact. Fis. Nat.* **94** no. 1 (2000), 15–18.
[11] R. T. Rockafellar, Lipschitzian properties of multifunctions, *Nonlinear Anal.* **9** (1984) 867–885.
[12] R. T. Rockafellar & R. Wets, *Variational analysis*, Ser. Com. Stu. Math., Springer, 1998.

(Recibido en julio de 2005. Aceptado en agosto de 2005)

Laboratoire Analyse, Optimisation, Contrôle,
Département de Mathématiques et Informatique
Université des Antilles et de la Guyane
Campus de Fouillole, F–97159 Pointe–à–Pitre
France
e-mail: celia.jean-alexis@univ-ag.fr
e-mail: apietrus@univ-ag.fr