A semilocal convergence of the Secant–type method for solving a generalized equations

S. Hilout [1], **A. Piétrus** [2],

January 10, 2005

Abstract. In this paper we present a study of the existence and the convergence of a secant–type method for solving abstract generalized equations in Banach spaces. With different assumptions for divided differences, we obtain a procedure that have superlinear convergence. This study follows the recent results of semilocal convergence related to the resolution of nonlinear equations (see [11]).

Keywords. Set–valued mapping, generalized equation, super–linear convergence, Aubin continuity, divided difference.

AMS subject classifications. 47H04, 65K10

1 Introduction

This paper is concerned with the problem of approximating a solution of the "abstract" generalized equation

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

where f is a continuous function from X into Y and G is a set–valued map from X to the subsets of Y with closed graph and X, Y are two real or complex Banach spaces. Let us recall that equation (1) is an abstract model for various problems, the reader could refer to [5, 6]. For solving (1), we consider the sequence

$$\begin{cases} x_0 \text{ and } x_1 \text{ are given starting points} \\ y_k = \alpha x_k + (1-\alpha)x_{k-1};\ \alpha \text{ is fixed in } [0,1[\\ 0 \in f(x_k) + [y_k, x_k; f](x_{k+1} - x_k) + G(x_{k+1}) \end{cases} \tag{2}$$

[1]Département de Mathématiques Appliquées et Informatique, Faculté des Sciences et Techniques, BP 523, 23000 Béni–Mellal, Maroc, E–mail said_hilout@yahoo.fr

[2]Laboratoire Analyse, Optimisation, Contrôle, Université des Antilles et de la Guyane, Département de Mathématiques et Informatique, Campus de Fouillole, F–97159 Pointe–à–Pitre, France, E–mail (Alain Piétrus) apietrus@univ–ag.fr

where $[y_k, x_k; f]$ is a first order divided difference of f on the points y_k and x_k. This operator will be defined in section 2.
In [11], the authors consider a similar iterative method like (2) with $\alpha = 0$ to solve nonlinear equations ($G \equiv 0$), they prove the semilocal convergence result of the method under a conditioned divided differences. Analogous results can also be found in [12]. Let us note that a study of the convergence of Steffenson's method to a locally unique solution of a nonlinear operator equation is developped in [1] using a special choice of divided differences. When the single–valued function involved in (1) is differentiable and when the Fréchet derivative is Lipschitz, Dontchev [5, 6] showed that the Newton–type method is locally (quadratically) convergent to a solution of (1) and he prove that this convergence is uniform in the sense that the solution of (1) is stable, i.e., we find similar result when we replace $y = 0$ in (1) by a small perturbation y. Analogous results (superlinear convergence) can be found in [17] when the derivative of f is Hölder.
In [9], we consider a third order iterative method under some assumptions on the first and the second Fréchet derivative of f at the solution of (1), we prove that this method is locally (cubically) convergent.
A combination of Newton's method with the first order divided differences method is developped in [4] to solve a nonlinear equations ($G \equiv 0$) with $f = f_1 + f_2$ where f_1 is a differentiable function and f_2 is continuous but admitting a divided difference. An extension of this method to generalized equations is studied in [10] under an assumption on the second order divided difference.
Recently, Michel Geoffroy in [8] obtained theQ-superlinear convergence of a secant type method for solving (1) assuming the existence of first and second order divided differences and that the solution satisfies a calmness-type property.
Note that in this present work, we don't use the concept of second order divided difference, but only first order divided difference. This means that our method is valid if f possesses a second order divided difference or not.

Here, we show the existence of a sequence defined by (2) which is locally convergent a the solution x^* of (1).
The paper is organized as follows: In section 2, we give some definitions and recall a fixed–point theorem (lemma 2.1) which has been proved in [7]. This fixed point theorem is the main tool to prove the existence and the convergence of the sequence (2). In section 3, we show the existence and the convergence of the sequence defined

by (2). At the end of the paper, we specify the cases $\alpha = 1$ and $\alpha = 0$.

2 Preliminaries and assumptions

Let us recall that the distance from a point x to a set A in the metric space (Z, ρ) is defined by $dist\,(x, A) = \inf\{\rho(x, y),\ y \in A\}$ and the excess e from the set A from the set C is given by $e(A, C) = \sup\{dist\,(x, A),\ x \in C\}$. Let $\Lambda : X \rightrightarrows Y$ be a set–valued map, we denote by $gph\,\Lambda = \{(x, y) \in X \times Y,\ y \in \Lambda(x)\}$ and $\Lambda^{-1}(y) = \{x \in X,\ y \in \Lambda(x)\}$. We denote by $B_r(x)$ the closed ball centered at x with radius r. The norm in the Banach spaces X and Y are both denoted by $\parallel . \parallel$ and $\mathcal{L}(X, Y)$ is the space of bounded and linear operators from X to Y. ∇f denotes the Fréchet derivative of f.

Definition 2.1 *(Aubin [2]) A set–valued Λ is Pseudo–Lipschitz around $(x_0, y_0) \in gph\,\Lambda$ with modulus M if there exist constants a and b such that*

$$\sup_{z \in \Lambda(y') \cap B_a(y_0)} dist\,(z, \Lambda(y'')) \leq M \parallel y' - y'' \parallel, \quad \text{for all } y' \text{ and } y'' \text{ in } B_b(x_0). \tag{3}$$

Using the excess, we have an equivalent definition replacing the inequality (3) by

$$e(\Lambda(y') \cap B_a(y_0), \Lambda(y'')) \leq M \parallel y' - y'' \parallel, \quad \text{for all } y' \text{ and } y'' \text{ in } B_b(x_0). \tag{4}$$

Characterizations of the Pseudo–Lipschitz property are obtained by Rockafellar using the Lipschitz continuity of the distance function $dist\,(y, \Lambda(x))$ around (x_0, y_0) in [18] and by Mordukhovich in [14] via the concept of coderivative of multiapplications. For more details and applications of this property, the reader could refer to [3, 7, 15, 19].

Definition 2.2 *An operator $[x, y; f] \in \mathcal{L}(X, Y)$ is called a divided difference of first order of the function f on the points x and y in X $(x \neq y)$ if this operator satisfies the followings :*

1. $[x, y; f](y - x) = f(y) - f(x)$.
2. *if f is Fréchet differentiable at x then* $[x, x; f] = \nabla f(x)$.

Lemma 2.1 *Let (Z, ρ) be a complete metric space, let ϕ a set–valued map from Z into the closed subsets of Z, let $\eta_0 \in Z$ and let r and λ be such that $0 \leq \lambda < 1$ and*
(a) $dist\,(\eta_0, \phi(\eta_0)) \leq r(1 - \lambda)$,

(b) $e(\phi(x_1) \cap B_r(\eta_0), \phi(x_2)) \leq \lambda\ \rho(x_1, x_2)\ \forall x_1, x_2 \in B_r(\eta_0)$,
then ϕ *has a fixed–point in* $B_r(\eta_0)$*. That is, there exists* $x \in B_r(\eta_0)$ *such that* $x \in \phi(x)$*. If* ϕ *is single–valued, then* x *is the unique fixed point of* ϕ *in* $B_r(\eta_0)$*.*

The proof of lemma 2.1 is given in [7] employing the standard iterative concept for contracting mapping. This lemma is a generalization of a fixed–point theorem given in [13] where in assertion (b) of the lemma 2.1 the excess e is replaced by the Pompeiu–Hausdorff distance. In the continuation of this work, the distance ρ in lemma 2.1 is replaced by the norm.

In the sequel we suppose that, x^* is a solution of (1) and for every distinct points x and y in a neighbourhood V of x^*, there exists a first order divided difference of f at these points. We also make the following assumptions on a neigbourhood V of x^*

($\mathcal{H}1$) There exists $\nu > 0$ such that for all x, y, u and v in V ($x \neq y$ and $u \neq v$)

$$\| [x, y; f] - [u, v; f] \| \leq \nu(\| x - u \|^p + \| y - v \|^p),\ p \in [0, 1]$$

($\mathcal{H}2$) The set–valued map $(f + G)^{-1}$ is M-Pseudo–Lipschitz around $(0, x^*)$.

($\mathcal{H}3$) For all $x, y \in V$, we have $||[x, y; f]|| \leq \kappa$ and $M\kappa < 1$.

Remark 2.1 *The hypothesis (*$\mathcal{H}3$*) implies that the function* f *is* κ*-Lipschitz on* V*.*

When a single–valued function f satisfies the assumption ($\mathcal{H}1$), we say that f has a (ν, p)–Hölder continuous divided differences on V. In [11], the authors showed a semilocal result of convergence of the secant method to solve a nonlinear equation ($G \equiv 0$) under a new condition relaxing the condition ($\mathcal{H}1$) by replacing (in ($\mathcal{H}1$)) the right term of the inequality by $\omega(\| x - u \|, \| y - v \|)$ where ω is a continuous nondecreasing function in its two arguments from $\mathbb{R}_+ \times \mathbb{R}_+$ to $\mathbb{R}_+$.

3 Convergence Analysis

In this section we show the existence of the sequence defined by (2) and we present some results of convergence to the solution x^* of (1) under the previous assumptions.

We need to introduce some notations. First, define the set–valued map $Q : X \rightrightarrows Y$ by

$$Q(x) = f(x^*) + G(x). \tag{5}$$

For $k \in I\!N^*$ and x_k, y_k defined in (2), we consider the application

$$Z_k(x) := f(x^*) - f(x_k) - [y_k, x_k; f](x - x_k). \tag{6}$$

Finally, define the set–valued map $\psi_k : X \rightrightarrows X$ by

$$\psi_k(x) := Q^{-1}(Z_k(x)). \tag{7}$$

Lemma 3.1 *The following are equivalent*

1. *The map* $(f + G)^{-1}$ *is pseudo–Lipschitz around* $(0, x^*)$*;*
2. *The map* $(f(x^*) + G(.))^{-1}$ *isPseudo–Lipschitz around* $(0, x^*)$*.*

Proof of lemma 3.1. The proof is a consequence of corollary 2 ([7]), identifying F and f in corollary with $(f + G)$ and h respectively where $h(.) = -f(.) + f(x^*)$. □

The main result of this study is follow

Theorem 3.1 *Let* x^* *be solution of (1). We suppose that the assumptions* ($\mathcal{H}1$)–($\mathcal{H}3$) *are satisfied. For every* $C > \dfrac{M\nu[(1-\alpha)^p + \alpha^p]}{1 - M\kappa}$*, one can find* $\delta > 0$ *such that for every distinct starting points* x_0 *and* x_1 *in* $B_\delta(x^*)$*, there exists a sequence* (x_k) *defined by (2) which satisfies*

$$\| x_{k+1} - x^* \| \leq C \| x_k - x^* \| \max \{\| x_k - x^* \|^p, \| x_{k-1} - x^* \|^p\}. \tag{8}$$

To prove theorem 3.1, we first prove the following proposition:

Proposition 3.1 *Under the assumptions of theorem 3.1, one can find* $\delta > 0$ *such that for every distinct starting points* x_0 *and* x_1 *in* $B_\delta(x^*)$ *(*x_0*,* x_1 *and* x^* *distincts), the set–valued map* ψ_1 *has a fixed point* x_2 *in* $B_\delta(x^*)$ *satisfying*

$$\| x_2 - x^* \| \leq C \| x_1 - x^* \| \max \{\| x_1 - x^* \|^p, \| x_0 - x^* \|^p\}. \tag{9}$$

Remark 3.1 *The point* x_2 *is a fixed point of* ψ_1 *if and only if the following holds*

$$0 \in f(x_1) + [y_1, x_1; f](x_2 - x_1) + G(x_2). \tag{10}$$

Once x_k *is computed, we show that the function* ψ_k *has a fixed point* x_{k+1} *in* X*. This process allows us to prove the existence of a sequence* (x_k) *satisfying (2).*

Proof of the proposition 3.1. Since the iterate y_1 in (2) is defined by $y_1 = \alpha x_1 + (1-\alpha)x_0$ then it is clear that $y_1 \in B_\delta(x^*)$.
By hypothesis $(\mathcal{H}2)$ and lemma 3.1 there exist positive numbers M, a and b such that

$$e(Q^{-1}(y') \cap B_a(x^*), Q^{-1}(y'')) \leq M \parallel y' - y'' \parallel, \ \forall y', y'' \in B_b(0). \tag{11}$$

Fix $\delta > 0$ such that

$$\delta < \min \left\{ a \ ; \ \sqrt[p+1]{\frac{b}{\nu((1-\alpha)^p + \alpha^p)}} \ ; \ \frac{1}{\sqrt[p]{C}} \ ; \ \frac{b}{2\kappa} \ ; \ \sqrt[p+1]{\frac{b}{2^{p+2}\nu}} \right\}. \tag{12}$$

To prove proposition 3.1 we intend to show that both assertions (a) and (b) of lemma 2.1 hold; where $\eta_0 := x^*$, ϕ is the function ψ_1 defined by (7) and where r and λ are numbers to be set. According to the definition of the excess e, we have

$$dist\,(x^*, \psi_1(x^*)) \leq e\Big(Q^{-1}(0) \cap B_\delta(x^*), \psi_1(x^*)\Big). \tag{13}$$

Moreover, for all points x_0 and x_1 in $B_\delta(x^*)$ (x_0, x_1 and x^* distincts) we have

$$\parallel Z_1(x^*) \parallel = \parallel f(x^*) - f(x_1) - [y_1, x_1; f](x^* - x_1) \parallel .$$

By definition 2.2 and the assumption $(\mathcal{H}1)$ we deduce

$$\begin{array}{rcl} \parallel Z_1(x^*) \parallel & = & \parallel \Big([x^*, x_1; f] - [y_1, x_1; f]\Big)(x^* - x_1) \parallel \\ & \leq & \parallel [x^*, x_1; f] - [y_1, x_1; f] \parallel \parallel x^* - x_1 \parallel \\ & \leq & \nu \parallel x^* - y_1 \parallel^p \parallel x^* - x_1 \parallel \\ & \leq & \nu\Big((1-\alpha) \parallel x^* - x_0 \parallel + \alpha \parallel x^* - x_1 \parallel \Big)^p \parallel x^* - x_1 \parallel \end{array} \tag{14}$$

Thus

$$\parallel Z_1(x^*) \parallel \leq \nu[(1-\alpha)^p \parallel x^* - x_0 \parallel^p + \alpha^p \parallel x^* - x_1 \parallel^p] \parallel x^* - x_1 \parallel . \tag{15}$$

Then (12) yields, $Z_1(x^*) \in B_b(0)$.
Hence from (11) one has

$$\begin{array}{rcl} e\Big(Q^{-1}(0) \cap B_\delta(x^*), \psi_1(x^*)\Big) & = & e\Big(Q^{-1}(0) \cap B_\delta(x^*), Q^{-1}[Z_1(x^*)]\Big) \\ & \leq & M\nu[(1-\alpha)^p \parallel x^* - x_0 \parallel^p + \alpha^p \parallel x^* - x_1 \parallel^p] \parallel x^* - x_1 \parallel \end{array} \tag{16}$$

By (13), we get

$$\begin{aligned} dist\,(x^*,\psi_1(x^*)) &\leq M\nu[(1-\alpha)^p \parallel x^*-x_0 \parallel^p +\alpha^p \parallel x^*-x_1 \parallel^p] \parallel x^*-x_1 \parallel \\ &\leq M\nu[(1-\alpha)^p+\alpha^p] \parallel x^*-x_1 \parallel \max\,\{\parallel x_1-x^* \parallel^p, \parallel x_0-x^* \parallel^p\} \end{aligned} \tag{17}$$

Since $C(1-M\kappa) > M\nu[(1-\alpha)^p+\alpha^p]$ there exists $\lambda \in [M\kappa, 1[$ such that $C(1-\lambda) \geq M\nu[(1-\alpha)^p+\alpha^p]$ and

$$dist\,(x^*,\psi_1(x^*)) \leq C(1-\lambda) \parallel x^*-x_1 \parallel \max\,\{\parallel x_1-x^* \parallel^p, \parallel x_0-x^* \parallel^p\} \tag{18}$$

By setting $r := r_1 = C \parallel x^*-x_1 \parallel \max\,\{\parallel x_1-x^* \parallel^p, \parallel x_0-x^* \parallel^p\}$ we can deduce from the inequality (18) that the assertion (a) in lemma 2.1 is satisfied.
Now, we show that condition (b) of lemma 2.1 is satisfied.
By (12) we have $r_1 \leq \delta \leq a$ and moreover for $x \in B_\delta(x^*)$ we have

$$\begin{aligned} \parallel Z_1(x) \parallel &= \parallel f(x^*)-f(x_1)-[y_1,x_1;f](x-x_1) \parallel \\ &\leq \parallel f(x^*)-f(x) \parallel + \parallel [x,x_1;f]-[y_1,x_1;f] \parallel \parallel x-x_1 \parallel \end{aligned} \tag{19}$$

Using the assumptions $(\mathcal{H}1)$ and $(\mathcal{H}3)$ we obtain

$$\begin{aligned} \parallel Z_1(x) \parallel &\leq \kappa \parallel x^*-x \parallel +\nu \parallel x-y_1 \parallel^p \parallel x-x_1 \parallel \\ &\leq \kappa \parallel x^*-x \parallel +\nu(\parallel x-x^* \parallel + \parallel x^*-y_1 \parallel)^p \parallel x-x_1 \parallel \\ &\leq \kappa\delta+\nu(2\delta)^p 2\delta = \kappa\delta+\nu 2^{p+1}\delta^{p+1}. \end{aligned} \tag{20}$$

Then by (12) we deduce that for all $x \in B_\delta(x^*)$ we have $Z_1(x) \in B_b(0)$. Then it follows that for all $x', x'' \in B_{r_0}(x^*)$ we have

$$e(\psi_1(x') \cap B_{r_1}(x^*), \psi_1(x'')) \leq e(\psi_1(x') \cap B_\delta(x^*), \psi_1(x'')),$$

which yields by (11)

$$\begin{aligned} e(\psi_1(x') \cap B_{r_1}(x^*), \psi_1(x'')) &\leq M \parallel Z_1(x')-Z_1(x'') \parallel \\ &\leq M \parallel [y_1,x_1;f](x''-x') \parallel \\ &\leq M \parallel [y_1,x_1;f] \parallel \parallel x''-x' \parallel \end{aligned} \tag{21}$$

Using $(\mathcal{H}3)$ and the fact that $\lambda \geq M\kappa$, we obtain

$$e(\phi_0(x') \cap B_{r_1}(x^*), \psi_1(x'')) \leq M\kappa \parallel x''-x' \parallel \leq \lambda \parallel x''-x' \parallel \tag{22}$$

and 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 $x_2 \in B_{r_1}(x^*)$ for the map ψ_1. Then the proof of proposition 3.1 is complete. □

Proof of theorem 3.1. Proceeding by induction, keeping $\eta_0 = x^*$ and setting

$$r := r_k = C \parallel x^* - x_k \parallel \max \{\parallel x_k - x^* \parallel^p, \parallel x_{k-1} - x^* \parallel^p\},$$

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 gives the inequality (8) and the proof of theorem 3.1 is complete. □

4 Concluding remarks

When $\alpha = 1$, our method is no longer valid, but if we suppose that f is Fréchet differentiable (2) is equivalent to a Newton-type method to solve (1). In this case conditions on ∇f give quadratic convergence (see [5]) and superlinear convergence (see [16]) and in the two cases the convergence is uniform (see [6] and [17]).
When $\alpha = 0$ the sequence (2) seems to the method introduced by M. Geoffroy and A. Piétrus in [10]. Let us note that the problem studied in [10] can be seen as a perturbation of (1) by a Fréchet differentiable function. In both cases, we obtain a superlinear convergence using different assumptions, but in this paper the existence of second order divided differences is not required.

References

[1] I.K. Argyros, A new convergence theorem for steffen's method on Banach spaces and applications, Southwest J. of Pure and Appl. Math., 01, (1997), p. 23–29.

[2] J.P. Aubin, Lipschitz behavior of solutions to convex minimization problems, Mathematics of Operations Research, 9, (1984), p. 87–111.

[3] J.P. Aubin, H. Frankowska, Set–valued analysis, Birkhäuser, Boston, 1990.

[4] E. Cătinas, On some iterative methods for solving nonlinear equations, Rev. Anal. Numé. Théor. Approx., 23, (1994), p. 17–53.

[5] A.L. Dontchev, Local convergence of the Newton method for generalized equation, C.R.A.S, Paris, 322, Serie I, (1996), p. 327–331.

[6] A.L. Dontchev, Uniform convergence of the Newton method for Aubin continuous maps, Serdica Math. J., 22, (1996), p. 385–398.

[7] A.L. Dontchev, W.W. Hager, An inverse function theorem for set–valued maps, Proc. Amer. Math. Soc, 121, 1994, p. 481–489.

[8] M. Geoffroy, A secant type method for variational inclusions, Preprint.

[9] M. Geoffroy, S. Hilout, A. Piétrus, Acceleration of convergence in Dontchev's iterative method for solving variational inclusions, Serdica Math. J., 29, (2003), p. 45–54.

[10] M. Geoffroy, A. Piétrus, Local convergence of some iterative methods for generalized equations, J. Math. Anal. Appl., 290, (2004), p. 497–505.

[11] M.A. Hernández, M.J. Rubio, Semilocal convergence of the secant method under mild convergence conditions of differentiability, Comput. and Math. with Appl., 44, (2002), p. 277–285.

[12] M.A. Hernández, M.J. Rubio, ω–conditioned divided differences to solve nonlinear equations, Monografías del Semin. Matem. García de Galdeano, 27, (2003), p. 323–330.

[13] A.D. Ioffe, V.M. Tikhomirov, Theory of extremal problems, North Holland, Amsterdam, 1979.

[14] B.S. Mordukhovich, Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis, Trans. Amer. Math. Soc., 343, (1994), p. 609–657.

[15] B.S. Mordukhovich, Complete characterization of openess metric regularity and Lipschitzian properties of multifunctions, Trans. Amer. Math. Soc., 340, (1993), p. 1–36.

[16] A. Piétrus, Generalized equations under mild differentiability conditions, Rev. Real. Acad. Ciencias de Madrid, 94, (1), (2000), p.15-18

[17] A. Piétrus, Does Newton's method for set–valued maps converges uniformly in mild differentiability context?, Rev. Colombiana Mat., 32, (2000), p. 49–56.

[18] R.T. Rockafellar, Lipschitzian properties of multifunctions, Nonlinear Analysis, 9, (1984), p. 867–885.

[19] R.T. Rockafellar, R.J–B. Wets, Variational analysis, A Series of Comprehentives Studies in Mathematics, Springer, 317, (1998).