J. Math. Anal. Appl. 350 (2009) 147–156

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

# An inertial proximal scheme for nonmonotone mappings

Michel H. Geoffroy [1]

*Laboratoire AOC, Dpt. de Mathématiques, Université Antilles-Guyane, Campus de Fouillole, F-97159 Pointe-à-Pitre, Guadeloupe (FWI), France*

ARTICLE INFO

*Article history:*
Received 22 April 2008
Available online 19 September 2008
Submitted by A. Dontchev

*Keywords:*
Set-valued mapping
Metric regularity
Subregularity
Strong regularity
Generalized equations
Successive approximation
Variational convergences

ABSTRACT

We present an inertial proximal method for solving an inclusion involving a nonmonotone set-valued mapping enjoying some regularity properties. More precisely, we investigate the local convergence of an implicit scheme for solving inclusions of the type $T(x) \ni 0$ where $T$ is a set-valued mapping acting from a Banach space into itself. We consider subsequently the case when $T$ is strongly metrically subregular, metrically regular and strongly regular around a solution to the inclusion. Finally, we study the convergence of our algorithm under variational perturbations.



## 1. Introduction

In [1], Alvarez studied the asymptotic behavior of the solutions of a second-order evolution equation with linear damping and convex potential. More precisely, he considered the following differential system in a Hilbert space $H$:

$$u''(t) + \gamma u'(t) + \nabla\Phi\big(u(t)\big) = 0, \tag{1.1}$$

where $\gamma > 0$ and $\Phi : H \to \mathbb{R}$ is a convex differentiable function. Since most differential equations cannot be solved explicitly, he considered the following implicit discretization of (1.1)

$$\frac{u_{k+1} - 2u_k + u_{k-1}}{h^2} + \gamma \frac{u_{k+1} - u_k}{h} + \nabla\Phi(u_{k+1}) = 0, \tag{1.2}$$

where the positive number $h$ stands for the so-called *step size* parameter. Note that relation (1.2) can be rewritten in the following way

$$u_{k+1} - u_k - \alpha(u_k - u_{k-1}) + \lambda\nabla\Phi(u_{k+1}) = 0, \tag{1.3}$$

where $\alpha$ and $\lambda$ are positive numbers depending on $h$ and $\gamma$. Alvarez called relation (1.3) the *inertial proximal method* and investigated also the case when the function $\Phi$ is not necessarily differentiable by replacing the derivative of $\Phi$ at $u_{k+1}$ in (1.3) with $\partial\Phi(u_{k+1})$ ($\partial\Phi$ denoting the standard convex subdifferential of the function $\Phi$). Next, Alvarez and Attouch [2] studied the case when $\nabla\Phi$ is replaced with a maximal monotone operator $A$. More precisely, they considered the algorithm

$$x_{k+1} - x_k - \alpha_k(x_k - x_{k-1}) + \lambda_k A(x_{k+1}) \ni 0, \tag{1.4}$$

*E-mail address:* michel.geoffroy@univ-ag.fr.

[1] Supported by Contract EA3591 (France).

0022-247X/$ – see front matter 
doi:10.1016/j.jmaa.2008.09.030

where $\alpha_k \in [0,1)$ and $\lambda_k > 0$ for all $k$. Note that when $\alpha_k = 0$ for $k = 0, 1, 2, \ldots$ then (1.4) corresponds to the standard proximal iteration (see, *e.g.*, [20]):

$$\frac{x_{k+1} - x_k}{\lambda_k} + A(x_{k+1}) \ni 0. \tag{1.5}$$

Alvarez and Attouch proved that, under several suitable assumptions involving the sequences $\alpha_k$ and $\lambda_k$, there exists a point $\hat{x} \in A^{-1}(0)$ such that the sequence $x_k$ generated by (1.4) is weakly convergent to $\hat{x}$ whenever the operator $A : H \rightrightarrows H$ is maximal monotone and such that $A^{-1}(0) \neq \emptyset$. To our knowledge, the inertial proximal method has not been investigated outside the (maximal) monotone setting. Carrying out such a study is the purpose of this work. Indeed, there is an interest in considering and studying such a method without monotonicity. First, because monotonicity forces us to work with mappings acting between a space and its dual, which usually yields to restrict the algorithm for mappings on a Hilbert space. Second, because in some cases monotonicity turns out to be rather a strong assumption, excluding several mappings that are metrically regular. As a simple example, consider a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}^n$. From [21, Proposition 12.3], $f$ is monotone if and only if the Jacobian $\nabla f(x)$ is positive-semidefinite at each $x$. On the other hand, a consequence of the Lyusternik–Graves theorem (see Theorem 2.2, Section 2) is that $f$ is metrically regular at some point $\bar{x}$ if and only if $\nabla f(\bar{x})$ is surjective. Therefore, metric regularity does not imply monotonicity.

Furthermore, iterative methods as (1.2) and (1.5) are closely tied to some continuous differential systems. Indeed, as we mentioned it at the very beginning of this introduction, the iteration (1.2) takes its inspiration from (1.1) while relation (1.5) can be interpreted as an implicit one-step discretization method for the following evolution differential inclusion

$$x'(t) + A\big(x(t)\big) \ni 0, \quad \text{a.e. } t \geqslant 0. \tag{1.6}$$

When $A$ is the subdifferential of a closed proper convex function $f : H \to \mathbb{R}$ (*i.e.*, $A = \partial f$) we know that under suitable conditions both the trajectory $\{x(t) : t \to \infty\}$ of (1.6) and the proximal sequence $x_k$ generated by (1.5) converge to a minimizer of $f$. Moreover, in [1], Alvarez proved that whenever the function $\Phi$ is convex and bounded from below, the trajectory $\{u(t) : t \to \infty\}$ defined by (1.1) is minimizing for $\Phi$. If the infimum of $\Phi$ on $H$ is attained, then $u(t)$ converges weakly to a minimizer of $\Phi$. Here also, the implicit discretization (1.2) of (1.1) generates a sequence $x_k$ which converges weakly to a minimizer of the mapping $\Phi$. Thus, one may observe, that in many situations the solution trajectories converge to a critical point of the mapping involved in the differential system, in addition, the asymptotic behavior of these solutions is preserved by the solutions of the corresponding discrete scheme. If the function under consideration is not convex, the discrete method shall converge to a critical point, *i.e.*, a solution to the inclusion $T(x) \ni 0$ where $T$ is some subdifferential that is not a maximal monotone operator. For this reason, it is worth studying discrete algorithms for solving $T(x) \ni 0$ when $T$ is a nonmonotone operator satisfying, for instance, some metric regularity properties (for developments on the metric regularity of subdifferentials one could refer to [3,24]).

In this paper, we study the local convergence of the inertial proximal algorithm for solving the inclusion

$$T(x) \ni 0, \tag{1.7}$$

where $T$ is a set-valued mapping acting from a general Banach space $X$ into itself. The method under consideration is the following: given $x_0$ and $x_1$ in $X$ find a sequence $x_n$ by applying the iteration

$$\lambda_n(x_{n+1} - x_n - e_n) - \lambda_n \mu_n (x_n - x_{n-1}) + T(x_{n+1} - e_n) \ni 0 \quad \text{for } n = 1, 2, \ldots.$$

The sequence $e_n$ is a so-called *error sequence* having to go to zero. In Section 3, we will discuss the very meaning of this sequence and we will also make precise the properties of the sequences $\lambda_n$ and $\mu_n$.

In the sequel, we subsequently, investigate the case when the mapping $T$ is metrically regular, strongly metrically regular and strongly metrically subregular around an element $(\bar{x}, 0)$ in the graph of $T$. Contrary to [2, Theorem 2.1] we obtain the strong convergence of the sequence $x_n$ generated by the above iteration to a solution to the inclusion (1.7). Moreover, most of our results are local since the initial guesses $x_0$ and $x_1$ must be sufficiently close to a solution to the problem.

Throughout, $X$ and $Y$ are Banach spaces. A set-valued mapping $F$ from $X$ into $Y$ is denoted by $F : X \rightrightarrows Y$. Its inverse $F^{-1}$ is defined as $x \in F^{-1}(y) \Leftrightarrow y \in F(x)$. The set $\operatorname{gph} F = \{(x, y) \in X \times Y \mid y \in F(x)\}$ stands for the graph of $F$ and we denote by $d(x, C)$ the distance from a point $x$ to a set $C$, that is, $d(x, C) = \inf_{y \in C} \|x - y\|$. The closed ball of radius $r$ centered at $a$ is denoted by $\mathbb{B}_r(a)$.

The rest of this paper consists of three sections. Section 2 is devoted to some basic definitions of metric regularity. The third section, the main part of this work, deals with the convergence analysis of the inertial proximal method for mappings enjoying metric regularity properties. Finally, we discuss the stability of the algorithm subjected to variational perturbations in Section 4.

## 2. Metric regularity

The concept of metric regularity of set-valued mappings goes back to the end of the 1970s but it sources come from some older classical theorems of differential calculus and linear analysis. The Banach open mapping theorem [6], the tangent space theorem of Lyusternik [18] and the surjection theorem of Graves [14] are among them. The definition of the metric regularity of a set-valued mapping reads as follows.

**Definition 2.1.** A mapping $F : X \rightrightarrows Y$ is said to be metrically regular at $\bar{x}$ for $\bar{y}$ if $F(\bar{x}) \ni \bar{y}$ and there exist some positive constants $\kappa, a$ and $b$ such that

$$d\big(x, F^{-1}(y)\big) \leqslant \kappa d\big(y, F(x)\big) \quad \text{for all } x \in \mathbb{B}_a(\bar{x}),\ y \in \mathbb{B}_b(\bar{y}). \tag{2.1}$$

The infimum of $\kappa$ for which (2.1) holds is the *regularity modulus* denoted $\operatorname{reg} F(\bar{x}|\bar{y})$; the case when $F$ is not metrically regular at $\bar{x}$ for $\bar{y}$ corresponds to $\operatorname{reg} F(\bar{x}|\bar{y}) = \infty$. Smaller values of $\kappa$ correspond to more favorable behavior. The metric regularity of a mapping $F$ at $\bar{x}$ for $\bar{y}$ is known to be equivalent to the Aubin continuity of the inverse $F^{-1}$ at $\bar{y}$ for $\bar{x}$ (see, *e.g.*, [21]). Recall that a set-valued map $\Gamma : Y \rightrightarrows X$ is Aubin continuous at $(\bar{y}, \bar{x}) \in \operatorname{gph} \Gamma$ (see [4]) if there exist positive constants $\kappa, a$ and $b$ such that

$$e\big(\Gamma(y') \cap \mathbb{B}_a(\bar{x}), \Gamma(y)\big) \leqslant \kappa \|y' - y\| \quad \text{for all } y, y' \in \mathbb{B}_b(\bar{y}), \tag{2.2}$$

where $e(A, B)$ denotes the excess from a set $A$ to a set $B$ and is defined as $e(A, B) = \sup_{x \in A} d(x, B)$.

A central result in the theory of metric regularity is the Lyusternik–Graves of which the Banach open mapping theorem is an immediate consequence. It goes back to the works of Dmitruk et al. [8] who proved it in the single-valued framework. In the general form of this theorem we present next, and which is from [12] (see also [15,16]), we use the following convention: we say that a set is locally closed at one of its points if some neighborhood of that point has closed intersection with the set.

**Theorem 2.2** *(Extended Lyusternik–Graves). Consider a mapping $F : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ at which $\operatorname{gph} F$ is locally closed. Consider also a function $g : X \to Y$ which is Lipschitz continuous near $\bar{x}$ with a Lipschitz constant $\delta$. If $\operatorname{reg} F(\bar{x}|\bar{y}) < \kappa < \infty$ and $\delta < \kappa^{-1}$, then*

$$\operatorname{reg}(g + F)\big(\bar{x}|g(\bar{x}) + \bar{y}\big) \leqslant \big(\kappa^{-1} - \delta\big)^{-1}.$$

Now we introduce the next regularity property for set-valued mapping that we will need in the sequel, to this end, let us recall the notion of graphical localization. A *graphical localization* of a mapping $F : X \rightrightarrows Y$ at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is a mapping $\tilde{F} : X \rightrightarrows Y$ such that $\operatorname{gph} \tilde{F} = (U \times V) \cap \operatorname{gph} F$ for some neighborhood $U \times V$ of $(\bar{x}, \bar{y})$.

**Definition 2.3.** A mapping $F : X \rightrightarrows Y$ is strongly metrically regular at $\bar{x}$ for $\bar{y}$ if the metric regularity condition in Definition 2.1 is satisfied by some $\kappa$ and neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$ and, in addition, the graphical localization of $F^{-1}$ with respect to $U$ and $V$ is single-valued. Equivalently, the graphical localization $V \ni y \mapsto F^{-1}(y) \cap U$ is a Lipschitz continuous function whose Lipschitz constant is equal to $\kappa$.

Obviously, the strong regularity implies the metric regularity by definition. Nevertheless, in some cases, metric regularity and strong regularity are equivalent. In particular, this equivalence holds for mappings of the form of the sum of a smooth function and the normal cone mapping over a polyhedral convex set (see, *e.g.*, [11]). Moreover, for any set-valued mapping that is *locally monotone* near the reference point metric regularity at that point implies, and hence is equivalent to, strong regularity. This is a consequence of a result established by Kenderov [17, Proposition 2.6] about the single-valuedness of lower semicontinuous monotone mappings.

Furthermore, see [21, Proposition 12.54], any maximal monotone mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ which is strongly monotone is strongly regular at the unique solution of $T(x) \ni 0$.

Now we briefly present the strong subregularity which is the last regularity property we consider here.

**Definition 2.4.** A mapping $F : X \rightrightarrows Y$ is said to be strongly metrically subregular at $\bar{x}$ for $\bar{y}$ if $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and there exists a constant $\kappa > 0$ along with a neighborhood $U$ of $\bar{x}$ such that

$$\|x - \bar{x}\| \leqslant \kappa d\big(\bar{y}, F(x)\big) \quad \text{for all } x \in U.$$

This property is equivalent to the "local Lipschitz property at a point" of the inverse mapping, a property first formally introduced in [10]. Note that Dontchev and Rockafellar [13, Definition 5.1] define a strongly subregular set-valued mapping $F : X \rightrightarrows Y$ (at $\bar{x}$ for $\bar{y}$) as a mapping satisfying

$$\|x - \bar{x}\| \leqslant \kappa d\big(\bar{y}, F(x) \cap V\big) \quad \text{for all } x \in U$$

for some neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$. It turns out that this last definition is equivalent to Definition 2.4 (see [3, Remark 3.4]).

For more details on metric regularity and applications to variational problems one can refer to [5,13,16] and the monographs [19,21].

## 3. Convergence analysis

From now on, we assume that the solution set of the inclusion (1.7) is nonempty, *i.e.*, there exists an element $\bar{x}$ in $T^{-1}(0)$ and we consider the following implicit scheme for solving this inclusion:

$$\lambda_n(x_{n+1} - x_n - e_n) - \lambda_n\mu_n(x_n - x_{n-1}) + T(x_{n+1} - e_n) \ni 0 \quad \text{for } n = 1, 2, \ldots. \tag{3.1}$$

In the above, $\lambda_n$ is a non-increasing sequence of positive numbers bounded above by some $\bar{\lambda}$ in $(0, 1)$ and such that $\lambda_n \searrow 0$ while $\mu_n$ stands for a sequence bounded above by some $\bar{\mu} \in (0, 1)$.

The sequence $e_n$ is a so-called *error sequence* having to go to zero. More precisely, a routine computation shows that relation (3.1) is equivalent to

$$x_{n+1} \in (T + \lambda_n I)^{-1}\big(\lambda_n x_n + \lambda_n\mu_n(x_n - x_{n-1})\big) + e_n \quad \text{for } n = 1, 2, \ldots. \tag{3.2}$$

Hence, $e_n$ measures the error made in the computation of $(T + \lambda_n I)^{-1}(\lambda_n x_n + \lambda_n\mu_n(x_n - x_{n-1}))$, incorporating such an error provides a more realistic model from a practical point of view.

To prove some of our convergence results, we will employ the following set-valued generalization (established in [9]) of a local version of Banach fixed point theorem.

**Theorem 3.1.** *Let $(X, d)$ be a complete metric space and consider a set-valued mapping $\Phi : X \rightrightarrows X$, a point $\bar{x} \in X$, and nonnegative scalars $\alpha$ and $\theta$ such that $0 \leqslant \theta < 1$, the sets $\Phi(x) \cap \mathbb{B}_\alpha(\bar{x})$ are closed for all $x \in \mathbb{B}_\alpha(\bar{x})$ and the following conditions hold*:

(i) $d(\bar{x}, \Phi(\bar{x})) < \alpha(1 - \theta)$;
(ii) $e(\Phi(u) \cap \mathbb{B}_\alpha(\bar{x}), \Phi(v)) \leqslant \theta d(u, v)$ *for all* $u, v \in \mathbb{B}_\alpha(\bar{x})$.

*Then $\Phi$ has a fixed point in $\mathbb{B}_\alpha(\bar{x})$. That is, there exists $x \in \mathbb{B}_\alpha(\bar{x})$ such that $x \in \Phi(x)$.*

First, we investigate the case when the mapping $T$ is strongly metrically subregular at $\bar{x}$ for 0. This implies, in particular, that the point $\bar{x}$ is an isolated solution to the inclusion (1.7).

**Proposition 3.2.** *Assume that the mapping $T$ is strongly metrically subregular at $\bar{x}$ for* 0. *Then there is a neighborhood $\Omega$ of $\bar{x}$ such that any sequence $x_n$ generated by* (3.1) *and whose elements are in $\Omega$ strongly converges to $\bar{x}$.*

**Proof.** Let $a$ and $\kappa$ be positive numbers such that the mapping $T$ is strongly metrically subregular at $\bar{x}$ for 0 with constant $\kappa$ and neighborhood $\mathbb{B}_a(\bar{x})$. Now, suppose that (3.1) generates a sequence $x_n$ such that $x_n \in \mathbb{B}_\alpha(\bar{x})$ for all $n$ where $\alpha \in (0, a)$ is such that $x_{n+1} - e_n \in \mathbb{B}_a(\bar{x})$ for $n$ large enough. From the definition of the strong subregularity we have

$$\|x - \bar{x}\| \leqslant \kappa d\big(0, T(x)\big) \quad \text{for all } x \in \mathbb{B}_a(\bar{x}).$$

Then

$$\|x_{n+1} - e_n - \bar{x}\| \leqslant \kappa\, \big\|\lambda_n\mu_n(x_n - x_{n-1}) - \lambda_n(x_{n+1} - x_n - e_n)\big\| \leqslant \kappa\lambda_n\big(2\mu_n a + 2a + \|e_n\|\big).$$

Since both of the sequences $\lambda_n$ and $\|e_n\|$ go to zero we get the desired conclusion. □

The interest of Proposition 3.2 arises from the fact that any sequence (whose elements are close enough to the solution $\bar{x}$) strongly converges to $\bar{x}$; nevertheless it is somewhat lessened by the fact that its statement does not guarantee that the whole sequence $x_n$ stays close to $\bar{x}$ whenever the initial guesses $x_0$ and $x_1$ are near $\bar{x}$. It turns out that we are able to avoid this drawback by strengthening the regularity property of the mapping $T$. Actually the strong metric subregularity of $T$ corresponds to a *calmness-type* property of its inverse $T^{-1}$ (see, *e.g.*, [13]); but if we assume that the set-valued mapping $T^{-1}$ is Lipschitz continuous then we can improve the conclusions of Proposition 3.2. Before stating explicitly this result we have to collect some background material.

The Lipschitz continuity of a set-valued mapping can be formulated in terms of Pompeiu–Hausdorff distance between sets, but it can also be expressed through the regular distance function as stated below.

**Definition 3.3.** Let $X$ and $Y$ be two Banach spaces. Consider a closed-valued mapping $F : Y \rightrightarrows X$ and a nonempty subset $D$ of the domain of $F$ (recall that the domain of $F$ is the set $\operatorname{dom} F = \{y \in Y \mid F(y) \neq \emptyset\}$). Then $F$ is Lipschitz continuous relative to $D$ with constant $\kappa$ if

$$d\big(x, F(y)\big) \leqslant \kappa d\big(y, F^{-1}(x) \cap D\big) \quad \text{for all } x \in X \text{ and } y \in D. \tag{3.3}$$

An important class of Lipschitz continuous set-valued mappings is given by *polyhedral convex mappings*, *i.e.*, mappings whose graph is a polyhedral convex set. More precisely, Walkup and Wets [23] showed in 1969 that any polyhedral mapping is Lipschitz continuous relative to its domain. Note that polyhedral mappings are a useful tool in linear programming and, in particular, they may describe the solution set of a system of inequalities.

**Lemma 3.4.** *Let $S : X \rightrightarrows X$ be a set-valued mapping, assume that its inverse $S^{-1}$ is Lipschitz continuous relative to its domain* $\operatorname{dom} S^{-1}$ *with constant $\kappa$ and such that $S^{-1}(0) = \{\bar{x}\}$. Let $\lambda_n$ be the sequence defined at the very beginning of the present section. If, in addition, the sequence $\lambda_n$ is chosen such that $\kappa\bar{\lambda} < 1$ then, for all positive integers n and all $x \in X$ we have*

$$\|x - \bar{x}\| \leqslant \frac{\kappa}{1 - \kappa\lambda_n} d\big(\lambda_n\bar{x}, (S + \lambda_n I)(x)\big), \tag{3.4}$$

*where I denotes the identity mapping on X.*

**Proof.** We denote by $D$ the domain of the set-valued mapping $S^{-1}$. Since $S^{-1}(0) = \{\bar{x}\}$, the definition of the Lipschitz continuity of $S^{-1}$ yields

$$\|x - \bar{x}\| \leqslant \kappa d\big(0, S(x) \cap D\big), \quad \forall x \in X. \tag{3.5}$$

Fix a positive integer $n$ and let $x \in X$. If $(S + \lambda_n I)(x) = \emptyset$ then assertion (3.4) holds and there is nothing more to prove. Otherwise pick $z \in (S + \lambda_n I)(x)$; it follows that $z - \lambda_n x \in S(x) \cap D$ then using (3.5) we get $\|x - \bar{x}\| \leqslant k\|z - \lambda_n x\|$. Hence,

$$\|x - \bar{x}\| \leqslant k\big\|z - \lambda_n\bar{x} - \lambda_n(x - \bar{x})\big\|.$$

Thus we have proved that, for any $z \in (S + \lambda_n I)(x)$, one has

$$\|x - \bar{x}\| \leqslant \frac{\kappa}{1 - \kappa\lambda_n}\|z - \lambda_n\bar{x}\|.$$

And it yields

$$\|x - \bar{x}\| \leqslant \frac{\kappa}{1 - \kappa\lambda_n} d\big(\lambda_n\bar{x}, (S + \lambda_n I)(x)\big),$$

that is, assertion (3.4) holds. □

We are now able to state a convergence result of our method in the case when the mapping $T^{-1}$ is Lipschitz continuous.

**Proposition 3.5.** *Assume that the mapping $T^{-1}$ is Lipschitz continuous relative to its domain with constant $\kappa$ and such that $T^{-1}(0) = \{\bar{x}\}$. If, in addition, the sequence $\lambda_n$ is chosen such that $\kappa\bar{\lambda} < 1$ then, any bounded sequence $x_n$ satisfying* (3.1) *strongly converges to $\bar{x}$.*

**Proof.** Let $x_n$ be a bounded sequence satisfying (3.1). Then one has

$$\lambda_n\mu_n(x_n - x_{n-1}) + \lambda_n x_n \in (T + \lambda_n I)(x_{n+1} - e_n) \quad \text{for } n = 1, 2, \ldots. \tag{3.6}$$

Moreover from Lemma 3.4 we get, for $n = 1, 2, \ldots$

$$\|x_{n+1} - e_n - \bar{x}\| \leqslant \frac{\kappa}{1 - \kappa\lambda_n} d\big(\lambda_n\bar{x}, (T + \lambda_n I)(x_{n+1} - e_n)\big).$$

Hence, thanks to relation (3.6), we infer

$$\|x_{n+1} - e_n - \bar{x}\| \leqslant \frac{\kappa\lambda_n}{1 - \kappa\lambda_n}\big\|\bar{x} - x_n - \mu_n(x_n - x_{n-1})\big\|, \quad \text{for } n = 1, 2, \ldots$$

which yields

$$\|x_{n+1} - \bar{x}\| \leqslant \frac{\kappa\lambda_n}{1 - \kappa\lambda_n}\big\|\bar{x} - x_n - \mu_n(x_n - x_{n-1})\big\| + \|e_n\|, \quad \text{for } n = 1, 2, \ldots \tag{3.7}$$

that is, the sequence $x_n$ converges to $\bar{x}$. □

Proposition 3.5 clearly improves the conclusions of Proposition 3.2 since the sequence $x_n$ does not need to be close to $\bar{x}$. Besides, thanks to relation (3.7), we know that whenever the initial guesses $x_0$ and $x_1$ are in a neighborhood $\Omega$ of $\bar{x}$, the whole sequence $x_n$ lies in $\Omega$ (under some suitable technical conditions on the sequences $\lambda_n$ and $\mu_n$).

Next comes a result regarding the convergence of the algorithm (3.1) for metrically regular mappings.

**Theorem 3.6.** *Assume that the mapping $T$ is metrically regular at $\bar{x}$ for 0 with neighborhoods $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(0)$ and is such that its regularity modulus $\operatorname{reg} T(\bar{x}|0)$ is less than $1/(1+\bar{\lambda})$. In addition, we suppose that the graph of $T$ is locally closed at $(\bar{x}, 0)$ and that the following conditions are both fulfilled*:

(i) $\bar{\lambda}(2a(1+\bar{\mu}) + \|e_n\|) \leqslant b$, *for* $n = 1, 2, \ldots,$
(ii) $\bar{\lambda}(a(1+2\bar{\mu}) + \|e_n\|) + \|e_n\|/\kappa \leqslant a$, *for* $n = 1, 2, \ldots.$

*Then there is a neighborhood $\Omega$ of $\bar{x}$ such that for any elements $x_0$ and $x_1$ in $\Omega$ there exists a sequence $x_n$ generated by* (3.1)*, whose elements are in $\Omega$, and which is strongly convergent to $\bar{x}$.*

**Proof.** There is a positive constant $\kappa$ such that the mapping $T$ is metrically regular at $\bar{x}$ for 0 with a constant $\operatorname{reg} T(\bar{x}|0) < \kappa < 1/(1+\bar{\lambda})$ and neighborhoods $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(0)$. Now pick $x_0$ and $x_1$ in $\mathbb{B}_a(\bar{x})$, then for any $x \in \mathbb{B}_a(\bar{x})$ we have

$$\begin{aligned}\big\|-\lambda_1(x - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big\| &\leqslant \lambda_1\|x - x_1\| + \lambda_1\|e_1\| + \lambda_1\mu_1\|x_1 - x_0\| \\ &\leqslant \lambda_1\big(2a(1+\mu_1) + \|e_1\|\big) \\ &\leqslant \bar{\lambda}\big(2a(1+\bar{\mu}) + \|e_1\|\big).\end{aligned}$$

Thanks to condition (i) in Theorem 3.6 we have $\bar{\lambda}(2a(1+\bar{\mu}) + \|e_1\|) \leqslant b$. Hence,

$$-\lambda_1(x - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0) \in \mathbb{B}_b(0). \tag{3.8}$$

Now we consider the mapping $\Phi_1$ defined by

$$\Phi_1 : X \ni x \longmapsto T^{-1}\big(-\lambda_1(x - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big) + e_1.$$

We intend to prove that $\Phi_1$ admits a fixed point $x_2$ in some neighborhood of $\bar{x}$. To this end, we show that the mapping $\Phi_1$ satisfies the assumptions of the Banach fixed point theorem (cf. Theorem 3.1). Using (3.8) and the metric regularity of $T$ at $\bar{x}$ for 0 we get

$$\begin{aligned}d\big(\bar{x}, \Phi_1(\bar{x})\big) &= d\big(\bar{x}, T^{-1}\big(-\lambda_1(\bar{x} - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big) + e_1\big) \\ &\leqslant d\big(\bar{x}, T^{-1}\big(-\lambda_1(\bar{x} - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big)\big) + \|e_1\| \\ &\leqslant \kappa d\big(-\lambda_1(\bar{x} - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0), T(\bar{x})\big) + \|e_1\| \\ &\leqslant \kappa\big\|-\lambda_1(\bar{x} - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big\| + \|e_1\| \\ &\leqslant \kappa\left(\lambda_1\|\bar{x} - x_1 - e_1\| + \lambda_1\mu_1\|x_1 - x_0\| + \frac{\|e_1\|}{\kappa}\right).\end{aligned}$$

Then, setting $\alpha_1 := \lambda_1\|\bar{x} - x_1 - e_1\| + \lambda_1\mu_1\|x_1 - x_0\| + \|e_1\|/\kappa$ and keeping in mind that $\kappa < 1/(1+\bar{\lambda}) \leqslant 1/(1+\lambda_1)$ we get

$$d\big(\bar{x}, \Phi_1(\bar{x})\big) \leqslant \alpha_1(1 - \kappa\lambda_1). \tag{3.9}$$

Moreover, we have $\alpha_1 \leqslant \bar{\lambda}(\|\bar{x} - x_1 - e_1\| + \bar{\mu}\|x_1 - x_0\|) + \|e_1\|/\kappa$. Then,

$$\alpha_1 \leqslant \bar{\lambda}\big(a(1+2\bar{\mu}) + \|e_1\|\big) + \|e_1\|/\kappa.$$

From assertion (ii) in Theorem 3.6 it follows that $\alpha_1 \leqslant a$.

Further, let $u, v \in \mathbb{B}_{\alpha_1}(\bar{x})$ and remind that, from the very beginning of this proof, we have

$$\big\|-\lambda_1(x - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big\| \leqslant b, \quad \text{for all } x \in \mathbb{B}_a(\bar{x}).$$

Since $\alpha_1 \leqslant a$ both $u$ and $v$ are in $\mathbb{B}_a(\bar{x})$ thus

$$\big\|-\lambda_1(u - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big\| \leqslant b, \tag{3.10}$$

and

$$\big\|-\lambda_1(v - x_1 - e_1) + \lambda_1\mu_1(x_1 - x_0)\big\| \leqslant b. \tag{3.11}$$

Hence, from relations (3.10) and (3.11) together with the Aubin continuity of $T^{-1}$ at $(0, \bar{x})$ we have

$$e\big(\Phi_1(u) \cap \mathbb{B}_{\alpha_1}(\bar{x}), \Phi_1(v)\big) \leqslant \kappa\lambda_1\|u - v\|. \tag{3.12}$$

Therefore, from (3.9) and (3.12), there exists a fixed point $x_2 \in \Phi_1(x_2) \cap \mathbb{B}_{\alpha_1}(\bar{x})$, *i.e.*,

$$x_2 \in \mathbb{B}_{\alpha_1}(\bar{x}) \quad \text{and} \quad \lambda_1(x_2 - x_1 - e_1) - \lambda_1\mu_1(x_1 - x_0) + T(x_2 - e_1) \ni 0. \tag{3.13}$$

If $x_2 = \bar{x}$ there is nothing more to prove. Otherwise, for any $x \in \mathbb{B}_a(\bar{x})$, we have

$$\left\| -\lambda_2(x - x_2 - e_2) + \lambda_2\mu_2(x_2 - x_1) \right\| \leqslant \bar{\lambda}\left(2a(1+\bar{\mu}) + \|e_2\|\right) \leqslant b.$$

Then consider the mapping $\Phi_2 : x \longmapsto T^{-1}(-\lambda_2(x - x_2 - e_2) + \lambda_2\mu_2(x_2 - x_1)) + e_2$. By the metric regularity of $T$ at $\bar{x}$ for 0 one has

$$\begin{aligned} d\left(\bar{x}, \Phi_2(\bar{x})\right) &= d\left(\bar{x}, T^{-1}\left(-\lambda_2(\bar{x} - x_2 - e_2) + \lambda_2\mu_2(x_2 - x_1)\right) + e_2\right) \\ &\leqslant \kappa\left(\lambda_2\|\bar{x} - x_2 - e_2\| + \lambda_2\mu_2\|x_2 - x_1\| + \frac{\|e_2\|}{\kappa}\right) \\ &\leqslant \alpha_2(1 - \kappa\lambda_2), \end{aligned}$$

where $\alpha_2 := \lambda_2\|\bar{x} - x_2 - e_2\| + \lambda_2\mu_2\|x_2 - x_1\| + \|e_2\|/\kappa$.

Take $u$ and $v$ in $\mathbb{B}_{\alpha_2}(\bar{x})$. The Aubin continuity of $T^{-1}$ at $(0, \bar{x})$ yields

$$e\left(\Phi_2(u) \cap \mathbb{B}_{\alpha_2}(\bar{x}), \Phi_1(v)\right) \leqslant \kappa\lambda_2\|u - v\|. \tag{3.14}$$

Hence, by Theorem 3.1, there exists $x_3 \in \Phi_2(x_3) \cap \mathbb{B}_{\alpha_2}(\bar{x})$ and we have

$$\|x_3 - \bar{x}\| \leqslant \lambda_2\|\bar{x} - x_2 - e_2\| + \lambda_2\mu_2\|x_2 - x_1\| + \|e_2\|/\kappa.$$

The induction step is now clear. Starting with two iterates $x_{n-1}$ and $x_n$ in $\mathbb{B}_a(\bar{x})$ we prove that the mapping

$$\Phi_n : X \ni x \longmapsto T^{-1}\left(-\lambda_n(x - x_n - e_n) + \lambda_n\mu_n(x_n - x_{n-1})\right) + e_n$$

admits a fixed point $x_{n+1} \in \mathbb{B}_{\alpha_n}(\bar{x})$ where

$$\alpha_n := \lambda_n\|\bar{x} - x_n - e_n\| + \lambda_n\mu_n\|x_n - x_{n-1}\| + \|e_n\|/\kappa \leqslant a,$$

that is, $x_{n+1}$ is such that

$$\lambda_n(x_{n+1} - x_n - e_n) - \lambda_n\mu_n(x_n - x_{n-1}) + T(x_{n+1} - e_n) \ni 0 \tag{3.15}$$

and

$$\|x_{n+1} - \bar{x}\| \leqslant \lambda_n\|\bar{x} - x_n - e_n\| + \lambda_n\mu_n\|x_n - x_{n-1}\| + \|e_n\|/\kappa. \tag{3.16}$$

Therefore we are able to construct a sequence $x_n$ satisfying conditions (3.15) and (3.16). Passing to the limit in (3.16) when $n$ goes to $\infty$ we obtain the strong convergence of the sequence $x_n$ to the solution $\bar{x}$ to the inclusion (1.7). □

The following theorem is of interest since it asserts that, whenever the mapping $T$ is strongly metrically regular at $\bar{x}$ for 0, we are able to define explicitly the sequence $x_n$, whose existence is ensured by Theorem 3.6.

**Theorem 3.7.** *Assume that the mapping $T$ is strongly metrically regular at $\bar{x}$ for 0 with neighborhoods $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(0)$ and is such that its regularity modulus $\operatorname{reg} T(\bar{x}|0)$ is less than $1/(1+\bar{\lambda})$. In addition, we suppose that the graph of $T$ is locally closed at $(\bar{x}, 0)$ and that the following conditions are both fulfilled:*

(i) $\bar{\lambda}(2a(1+\bar{\mu}) + \|e_n\|) \leqslant b$, *for* $n = 1, 2, \ldots,$
(ii) $\bar{\lambda}(a(1+2\bar{\mu}) + \|e_n\|) + \|e_n\|/\kappa \leqslant a$, *for* $n = 1, 2, \ldots.$

*Then the conclusion of Theorem 3.6 holds and the sequence $x_n$ is uniquely determined by the following equality*

$$x_{n+1} = (T + \lambda_n I)^{-1}\left(\lambda_n x_n + \lambda_n\mu_n(x_n - x_{n-1})\right) + e_n, \quad \textit{eventually.} \tag{3.17}$$

**Proof.** By repeating the proof of the Theorem 3.6 we get the existence of a sequence $x_n$ whose elements are in some ball centered at $\bar{x}$ and satisfy (3.1) with the same properties as in Theorem 3.6. It remains to show that the sequence $x_n$ is unique. To this end it suffices to prove that relation (3.17) holds. Since the sequence $x_n$ satisfies (3.1) we have

$$x_n \in (T + \lambda_n I)^{-1}\left(\lambda_n x_n + \lambda_n\mu_n(x_n - x_{n-1})\right) + e_n \quad \text{for all } n.$$

Then, to complete the proof, we only need to show that the set-valued mapping $(T + \lambda_n I)^{-1}$ has a single-valued graphical localization in some suitable neighborhoods. For this purpose, we are going to use the same argument that has been used in [13, Theorem 4.3] to study the stability of strong regularity with respect to single-valued perturbations. Since $T$ is strongly metrically regular at $\bar{x}$ for 0, it has the single-valued Lipschitzian inverse property there, *i.e.*, there exist positive constants $\alpha$ and $\beta$ such that the graphical localization

$$y \ni \mathbb{B}_\beta(0) \mapsto T^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x}) \tag{3.18}$$

is Lipschitz continuous with some positive constant $\kappa$ and single-valued on $\mathbb{B}_\beta(0)$. In the sequel, for any $y \in \mathbb{B}_\beta(0)$, we denote by $s(y)$ the unique element of $T^{-1}(y) \cap \mathbb{B}_\alpha(0)$. Then we consider the following graphical localization of the mapping $(T + \lambda_n I)^{-1}$:

$$y \ni \mathbb{B}_\delta(0) \mapsto (T + \lambda_n I)^{-1}(y) \cap \mathbb{B}_\delta(\bar{x}), \tag{3.19}$$

where $\delta$ is a positive number satisfying $\delta < \min\{\beta/3, \alpha\}$. Now, it is our purpose to show that the mapping defined in (3.19) is single-valued. To this end we assume that it is not single-valued, *i.e.*, there is an element $y \in \mathbb{B}_\delta(0)$ such that there exist two distinct points $x$ and $x'$ both of them in $(T + \lambda_n I)^{-1}(y) \cap \mathbb{B}_\delta(\bar{x})$. Thus we have in particular $x \in T^{-1}(y - \lambda_n x) \cap \mathbb{B}_\delta(\bar{x})$. Moreover, since $\lambda_n$ is a non-increasing sequence going to 0, for $n$ large enough, we have

$$\|y - \lambda_n x\| \leqslant \|y\| + \lambda_n\|x - \bar{x}\| + \lambda_n\|\bar{x}\| \leqslant (1 + \lambda_n)\delta + \lambda_n\|\bar{x}\| \leqslant 3\delta < \beta.$$

Hence, $y - \lambda_n x \in \mathbb{B}_\beta(0)$; then observing that $\mathbb{B}_\delta(\bar{x}) \subset \mathbb{B}_\alpha(\bar{x})$ and using the single-valuedness of the graphical localization given in (3.18) we obtain that $x = s(y - \lambda_n x)$. A similar argument yields $x' = s(y - \lambda_n x')$ and it follows

$$\|x - x'\| = \left\|s(y - \lambda_n x) - s(y - \lambda_n x')\right\| \leqslant \kappa\left\|(y - \lambda_n x) - (y - \lambda_n x')\right\| \leqslant \kappa\lambda_n\|x - x'\| < \|x - x'\|,$$

which is absurd then the graphical localization of $(T + \lambda_n I)^{-1}$ in (3.19) is single-valued. To complete the proof it remains to show that $\lambda_n x_n + \lambda_n \mu_n(x_n - x_{n-1}) \in \mathbb{B}_\delta(0)$ for $n$ large enough. This is a straightforward consequence of the convergence of $\lambda_n$ to zero and the boundedness of the sequence $x_n$. □

## 4. The perturbed algorithm

In this section we shall show that the inertial proximal algorithm is stable with respect to certain *variational perturbations* of the mappings $T$. By variational perturbations of $T$ we mean a sequence $T_n : X \rightrightarrows X$ of set-valued mappings converging to $T$ in some sense. We are thus led to associate to the inclusion (1.7) the following inertial proximal method where the mapping $T$ has been replaced with $T_n$:

$$\lambda_n(x_{n+1} - x_n - e_n) - \lambda_n\mu_n(x_n - x_{n-1}) + T_n(x_{n+1} - e_n) \ni 0 \quad \text{for } n = 1, 2, \ldots. \tag{4.1}$$

Before stating our first result we briefly present the convergence notion we shall use in the sequel. The lower and upper limits of a sequence $A_n$ of subsets of a normed space, with unit ball $\mathbb{B}$, are defined by:

$$\liminf_n A_n = \left\{x \in X \;\Big|\; \limsup_{n\to\infty} d(x, A_n) = 0\right\} = \{x \in X \mid \exists x_n \in A_n \text{ with } x_n \to x\};$$

$$\limsup_n A_n = \left\{x \in X \;\Big|\; \liminf_{n\to\infty} d(x, A_n) = 0\right\} = \{x \in X \mid \exists n_1 < n_2 < \cdots \text{ in } \mathbf{N},\ \exists x_{n_k} \in A_{n_k} \text{ with } x_{n_k} \to x\}.$$

**Definition 4.1.** A sequence of subsets $A_n$ is said to *set-converge* to a subset $A$, written $A_n \to A$, provided $\limsup_n A_n \subset A \subset \liminf_n A_n$.

Set convergence in this sense is known more specifically as Painlevé–Kuratowski convergence. A sequence $A_n$ of subsets of $X$ is said to be *lower set-convergent* to $A$ if $A \subset \liminf_n A_n$ and *upper set-convergent* to $A$ if $\limsup_n A_n \subset A$. Obviously, a sequence $A_n$ set-converges to $A$ if and only if it is both lower and upper set-convergent to $A$.

The following statement asserts that, regardless of the regularity properties of the mapping $T$, any cluster point of an arbitrary sequence satisfying (4.1) is a solution to the inclusion (1.7) whenever the sequence $\operatorname{gph} T_n$ upper-set converges to $\operatorname{gph} T$.

**Proposition 4.2.** *Let $T_n$ be a sequence of set-valued mappings such that* $\operatorname{gph} T_n$ *upper set-converges to* $\operatorname{gph} T$ *and let $x_n$ be a sequence satisfying* (4.1)*. Then any cluster point of $x_n$ is a solution to the inclusion* (1.7)*.*

**Proof.** Let $\hat{x}$ be a cluster point of the sequence $x_n$ defined in Proposition 4.2. Then there is a subsequence $x_{n_k}$ of $x_n$ converging to $\hat{x}$ and satisfying

$$\left(x_{n_k+1} - e_{n_k}, \lambda_{n_k}\mu_{n_k}(x_{n_k} - x_{n_k-1}) - \lambda_{n_k}(x_{n_k+1} - x_{n_k} - e_{n_k})\right) \in \operatorname{gph} T_{n_k}, \quad \text{for all } k. \tag{4.2}$$

Moreover, the sequence $(x_{n_k+1} - e_{n_k}, \lambda_{n_k}\mu_{n_k}(x_{n_k} - x_{n_k-1}) - \lambda_{n_k}(x_{n_k+1} - x_{n_k} - e_{n_k}))$ converges to $(\hat{x}, 0)$, thus, $(\hat{x}, 0) \in \limsup \operatorname{gph} T_n$. Keeping in mind that $\operatorname{gph} T_n$ upper set-converges to $\operatorname{gph} T$ we complete the proof. □

To conclude, we show a result similar to Theorem 3.2, that is, a convergence result for the perturbed inertial proximal method when the mapping $T$ is strongly metrically subregular. To this end we assume that the sequence $T_n$ converges to $T$ in the following sense:

$$\sup_{x\in U}\big|d\big(0,T_n(x)\big)-d\big(0,T(x)\big)\big|\to 0,\quad \text{for some neighborhood } U \text{ of } \bar{x}. \tag{4.3}$$

Note that relation (4.3) can be seen as a uniform local Wijsman convergence of the sequence $T_n(x)$ to $T(x)$. We recall that a sequence of subsets $A_n$ is said to *Wijsman-converge* to a subset $A$ if for every $u\in X$, $\lim_{n\to\infty} d(u,A_n)=d(u,A)$ (for details about the Wijsman convergence and variational convergences in general see, *e.g.*, [7,22]).

**Theorem 4.3.** *Let $T_n$ be a sequence of set-valued mappings satisfying* (4.3)*. If the mapping $T$ is strongly metrically subregular at $\bar{x}$ for* 0 *then there exists a neighborhood $\Omega$ of $\bar{x}$ such that any sequence $x_n$ satisfying* (4.1) *and whose elements are in $\Omega$ eventually, strongly converges to $\bar{x}$.*

**Proof.** Let $T$ be strongly subregular at $\bar{x}$ for 0 with constants $\kappa$ and $a$, that is,

$$\|x-\bar{x}\|\leqslant \kappa d\big(0,T(x)\big),\quad \text{for all } x\in\mathbb{B}_a(\bar{x}). \tag{4.4}$$

Adjust $a$ if necessary such that $\mathbb{B}_a(\bar{x})\subset U$ where $U$ is the neighborhood of $\bar{x}$ introduced in (4.3). Take a sequence $x_n$ satisfying (4.1) and whose elements are in a ball $\mathbb{B}_\alpha(\bar{x})$ where $0<\alpha<a$. Since the sequence $e_n$ goes to zero, for $n$ large enough, $x_{n+1}-e_n\in\mathbb{B}_a(\bar{x})$ and we have

$$\|x_{n+1}-e_n-\bar{x}\|\leqslant \kappa d\big(0,T(x_{n+1}-e_n)\big). \tag{4.5}$$

From property (4.3), we have that for all $\varepsilon>0$, there exists an integer $N$ such that for all $n\geqslant N$ one has

$$\big|d\big(0,T_n(x_{n+1}-e_n)\big)-d\big(0,T(x_{n+1}-e_n)\big)\big|<\varepsilon.$$

Hence for all $\varepsilon>0$,

$$\|x_{n+1}-e_n-\bar{x}\|\leqslant \kappa d\big(0,T_n(x_{n+1}-e_n)\big)+\varepsilon,\quad \text{eventually.}$$

Since $-\lambda_n(x_{n+1}-x_n-e_n)+\lambda_n\mu_n(x_n-x_{n-1})\in T_n(x_{n+1}-e_n)$ we get

$$\|x_{n+1}-e_n-\bar{x}\|\leqslant \kappa\lambda_n\big\|-(x_{n+1}-x_n-e_n)+\mu_n(x_n-x_{n-1})\big\|+\varepsilon,\quad \text{eventually.}$$

It follows that $x_n$ converges to $\bar{x}$. □

## Acknowledgment

The author would like to express his gratitude to the anonymous referee for his/her valuable comments and suggestions which have improved the original manuscript.

## References

[1] F. Alvarez, On the minimizing property of a second dissipative system in Hilbert spaces, SIAM J. Control Optim. 38 (4) (2000) 1102–1119.
[2] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (2001) 3–11.
[3] F.J. Aragón Artacho, M.H. Geoffroy, Characterization of metric regularity of subdifferentials, J. Convex Anal. 15 (2) (2008) 365–380.
[4] J.P. Aubin, H. Frankowska, Set-Valued Analysis, Birkhäuser, Boston, 1990.
[5] D. Azé, A unified theory for metric regularity of multifunctions, J. Convex Anal. 13 (2) (2006) 225–252.
[6] S. Banach, Théorie des opérations linéaires, Monografje Matematyczne, Warsaw, 1932; reprint, Chelsea, New York, 1963.
[7] G. Beer, Topologies on Closed and Closed Convex Sets, Kluwer Academic Publishers, 1993.
[8] A.V. Dmitruk, A.A. Milyutin, N.P. Osmolovskii, Lyusternik's theorem and the theory of extrema, Russian Math. Surveys 35 (1980) 11–51.
[9] A.L. Dontchev, W.W. Hager, An inverse mapping theorem for set-valued maps, Proc. Amer. Math. Soc. 121 (1994) 481–489.
[10] A.L. Dontchev, Characterizations of Lipschitz stability in optimization, in: Recent Developments in Well-posed Variational Problems, in: Math. Appl., vol. 331, Kluwer Acad. Publ., Dordrecht, 1995, pp. 95–115.
[11] A.L. Dontchev, R.T. Rockafellar, Characterization of strong regularity of variational inequalities over polyhedral convex sets, SIAM J. Optim. 6 (1996) 1087–1105.
[12] A.L. Dontchev, A.S. Lewis, R.T. Rockafellar, The radius of metric regularity, Trans. Amer. Math. Soc. 355 (2002) 493–517.
[13] A.L. Dontchev, R.T. Rockafellar, Regularity and conditioning of solution mappings in variational analysis, Set-Valued Anal. 12 (2004) 79–109.
[14] L.M. Graves, Some mappings theorems, Duke Math. J. 17 (1950) 111–114.
[15] A.D. Ioffe, On the local surjection property, Nonlinear Anal. 11 (1987) 565–592.
[16] A.D. Ioffe, Metric regularity and subdifferential calculus, Uspekhi Mat. Nauk 55 (3 (333)) (2000) 103–162; English translation: Russian Math. Surveys 55 (2000) 501–558.
[17] P. Kenderov, Semi-continuity of set-valued monotone mappings, Fund. Math. 88 (1975) 61–69.
[18] L.A. Lyusternik, On conditional extrema of functionals, Mat. Sb. 41 (1934) 390–401.

[19] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I: Basic Theory, Grundlehren Math. Wiss., vol. 330, Springer-Verlag, Berlin, 2006.
[20] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.
[21] R.T. Rockafellar, R.J.-B. Wets, Variational Analysis, Springer-Verlag, Berlin, 1997.
[22] Y. Sonntag, C. Zalinescu, Set convergences. An attempt of classification, Trans. Amer. Math. Soc. 340 (1) (1993) 199–226.
[23] D.W. Walkup, R.J.-B. Wets, A Lipschitzian characterization of convex polyhedra, Proc. Amer. Math. Soc. 23 (1969) 167–173.
[24] R. Zhang, J. Treiman, Upper-Lipschitz multifunctions and inverse subdifferentials, Nonlinear Anal. 24 (2) (1995) 273–286.