

CONVERGENCE OF THE PROXIMAL POINT METHOD FOR METRICALLY REGULAR MAPPINGS

F. J. ARAGÓN ARTACHO¹, A. L. DONTCHEV² AND M. H. GEOFFROY³

Abstract. In this paper we consider the following general version of the proximal point algorithm for solving the inclusion $T(x) \ni 0$, where T is a set-valued mapping acting from a Banach space X to a Banach space Y . First, choose any sequence of functions $g_n : X \rightarrow Y$ with $g_n(0) = 0$ that are Lipschitz continuous in a neighborhood of the origin. Then pick an initial guess x_0 and find a sequence x_n by applying the iteration $g_n(x_{n+1} - x_n) + T(x_{n+1}) \ni 0$ for $n = 0, 1, \dots$. We prove that if the Lipschitz constants of g_n are bounded by half the reciprocal of the modulus of regularity of T , then there exists a neighborhood O of \bar{x} (\bar{x} being a solution to $T(x) \ni 0$) such that for each initial point $x_0 \in O$ one can find a sequence x_n generated by the algorithm which is linearly convergent to \bar{x} . Moreover, if the functions g_n have their Lipschitz constants convergent to zero, then there exists a sequence starting from $x_0 \in O$ which is superlinearly convergent to \bar{x} . Similar convergence results are obtained for the cases when the mapping T is strongly subregular and strongly regular.

Résumé. Nous considérons dans ces travaux une généralisation de l'algorithme du point proximal pour résoudre des inclusions du type $T(x) \ni 0$, où T est une application multivoque agissant entre deux espaces de Banach X et Y . Tout d'abord, on se donne une suite de fonctions $g_n : X \rightarrow Y$ telle que chaque fonction g_n soit lipschitzienne dans un voisinage de l'origine et vérifie $g_n(0) = 0$. Partant d'un point initial x_0 on construit une suite x_n en appliquant la procédure itérative $g_n(x_{n+1} - x_n) + T(x_{n+1}) \ni 0$ for $n = 0, 1, \dots$. On montre que si les constantes de lipschitz des fonctions g_n sont bornées supérieurement par la moitié de l'inverse du module de régularité de T , alors il existe un voisinage O de \bar{x} (\bar{x} étant une solution de $T(x) \ni 0$) tel que pour tout point initial $x_0 \in O$ on peut trouver une suite x_n engendrée par l'algorithme et qui converge linéairement vers \bar{x} . De plus, si la suite des constantes de lipschitz des fonctions g_n converge vers 0, alors il existe une suite partant du point $x_0 \in O$ et qui converge super-linéairement vers \bar{x} . Des résultats de convergence similaires sont obtenus quand l'application multivoque T est fortement sous-régulière et fortement régulière.

1. INTRODUCTION

In this paper we study the convergence of a general version of the proximal point algorithm for solving the inclusion

$$T(x) \ni 0, \tag{1.1}$$

where T is a set-valued mapping acting from a Banach space X to the subsets of a Banach space Y . Choose a sequence of functions $g_n : X \rightarrow Y$ with $g_n(0) = 0$ and consider the following algorithm: given x_0 find a

¹ Department of Statistics and Operations Research, University of Alicante, 03071 Alicante, Spain, francisco.aragon@ua.es. This author is supported by Grant BES-2003-0188 from FPI Program of MEC (Spain).

² Mathematical Reviews, Ann Arbor, MI 48107, USA, ald@ams.org.

³ Laboratoire AOC, Dpt. de Mathématiques, Université Antilles-Guyane, F-97159 Pointe-à-Pitre, Guadeloupe, mgeoffro@univ-ag.fr. This author is supported by Contract EA3591 (France).

sequence x_n by applying the iteration

$$g_n(x_{n+1} - x_n) + T(x_{n+1}) \ni 0 \quad \text{for } n = 0, 1, 2, \dots \quad (1.2)$$

We prove in this work that if \bar{x} is a solution of (1.1) and the mapping T is metrically regular at \bar{x} for 0 and with locally closed graph near $(\bar{x}, 0)$, then, for any sequence of functions g_n that are Lipschitz continuous in a neighborhood U of the origin, the same for all n , and whose Lipschitz constants λ_n have supremum that is bounded by half the reciprocal of the modulus of regularity of T , there exists a neighborhood O of \bar{x} such that for each initial point $x_0 \in O$ one can find a sequence x_n satisfying (1.2) which is linearly convergent to \bar{x} in the norm of X . Moreover, if the functions g_n have their Lipschitz constants λ_n convergent to zero, then, among the sequences obtained by (1.2) with $x_0 \in O$ there exists at least one which is superlinearly convergent to \bar{x} .

If $g_n(u) = \lambda_n u$ and then we assume $Y = X$, we obtain that classical proximal point algorithm

$$0 \in \lambda_n(x_{n+1} - x_n) + T(x_{n+1}) \quad \text{for } n = 0, 1, 2, \dots, \quad (1.3)$$

first proposed by Martinet [16] and thoroughly explored in a subsequent path-breaking paper by Rockafellar [18] for the case when X is a Hilbert space and T is a *monotone* mapping. In particular, Rockafellar (see [18], Proposition 3) showed that when x_{n+1} is an approximate solution of (1.3) and T is maximal monotone, then for a sequence of positive scalars λ_n the iteration (1.3) produces a sequence x_n which is weakly convergent to a solution to (1.1) for any starting point $x_0 \in X$.

In the last three decades a number of authors have considered generalizations and modifications of the proximal point algorithm, and have also found applications of this method to specific variational problems. Most of the rapidly growing body of the literature on this subject has been concentrated on various versions of the algorithm for solving inclusions involving monotone mappings, and specifically, on monotone variational inequalities. We mention here the more recent papers by Solodov and Svaiter [20], Auslender and Teboulle [2], a series of papers by Kaplan and Tichatschke [13], Ahn et al. [1], Yang and He [22] and Bauschke et al. [3]; see also the references therein. Weaker form of monotonicity have been considered first in Spingarn [21], for details see Iusem et al. [12]. A convergence result of a different type was recently derived by Pennanen [17] who replaced the global condition of monotonicity by the local condition of strong regularity and observed that under this condition the Yosida regularization of the mapping T can be locally monotone even when T is not, and this local monotonicity is sufficient to show local convergence. In all papers mentioned above X is assumed to be a Hilbert space and T acts from X to itself.

In this paper we consider the proximal point method from a different perspective, by employing recent developments on regularity properties of set-valued mappings most of which can be found in the paper [10]. First, we consider the inclusion (1.1) with a mapping T acting between two arbitrary Banach spaces X and Y . Second, the possibility to choose different functions g_n enhances the efficacy of the algorithm. And third, the only assumption on the mapping T is that it is metrically regular at a reference solution. The convergence result we obtain is also local: to obtain a convergent sequence the initial guess must be sufficiently close to the solution.

In Section 2 we present some background material on metric regularity. Section 3 gives a statement and a proof of the main Theorem 3.1. The following Section 4 is about convergence of the proximal point algorithm (1.2) under different regularity assumptions for the mapping T . Namely, if T is *strongly subregular* we show that when (1.2) generates a sequence x_n whose elements are sufficiently close to the reference solution \bar{x} , then this sequence is linearly convergent to \bar{x} (or even superlinearly, depending on the functions g_n). Then we consider the case when T is *strongly regular*, which is more than metric regularity, and obtain uniqueness of the sequence x_n satisfying (1.2), which hence converges to \bar{x} in the way described in Theorem 3.1. This latter result partially recovers Theorem 9 in [17].

2. METRIC REGULARITY

Let X and Y be Banach spaces, let F be a set-valued mapping from X into the subsets of Y , denoted $F : X \rightrightarrows Y$ and let $(\bar{x}, \bar{y}) \in \text{gph } F$. Here $\text{gph } F = \{(x, y) \in X \times Y \mid y \in F(x)\}$ is the graph of F . We denote

by $d(x, C)$ the distance from a point x to a set C , that is, $d(x, C) = \inf_{y \in C} \|x - y\|$. $\mathcal{B}_r(a)$ denotes the closed ball of radius r centered at a and F^{-1} is the inverse of F defined as $x \in F^{-1}(y) \Leftrightarrow y \in F(x)$.

Definition 2.1. The mapping F is said to be metrically regular at \bar{x} for \bar{y} if there exists a constant $\kappa > 0$ such that

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)) \text{ for all } (x, y) \text{ close to } (\bar{x}, \bar{y}). \quad (2.1)$$

The metric regularity is a basic quantitative property of mappings in variational analysis which goes deep in the roots of both theoretical and numerical optimization. This concept is present implicitly in the classical Banach open mapping principle: a linear and bounded mapping $A : X \rightarrow Y$ is metrically regular (at any point in its graph) if and only if A maps X onto Y . The infimum of κ for which (2.1) holds is the *regularity modulus* denoted $\text{reg } F(\bar{x}|\bar{y})$; the case when F is not metrically regular at \bar{x} for \bar{y} corresponds to $\text{reg } F(\bar{x}|\bar{y}) = \infty$. In the simplest case, the regularity modulus of a linear and bounded mapping A satisfies $\text{reg } A = \sup \{ d(0, A^{-1}(y)) \mid y \in \mathcal{B} \}$ where \mathcal{B} is the closed unit ball. If in addition A^{-1} is single-valued, then $\text{reg } A = \|A^{-1}\|$. If a mapping F has closed and convex graph then F is metrically regular at \bar{x} for \bar{y} if and only if $\bar{y} \in \text{int rge } F$ (Robinson-Ursescu theorem). In particular, the mapping $F(x) = f(x) + \mathbb{R}_+^m$, where f is a function and \mathbb{R}_+^m is the positive orthant in \mathbb{R}^m so that the inclusion $F(x) \ni 0$ describes an inequality system, is metrically regular at a given point exactly when the Mangasarian-Fromovitz condition holds at that point.

The metric regularity of a mapping F is known to be equivalent to two other properties: the openness of linear rate of F and the Aubin property of the inverse F^{-1} . For recent advances on metric regularity and applications to variational problems, see [11], [9] and [10] and the book [19].

A central result in the theory of metric regularity is the Lyusternik-Graves which roughly says that the metric regularity is stable under perturbations of order higher than one. In the general form of this theorem we present next and which is from [9], we use the following convention: we say that a set $C \subset X$ is locally closed at $z \in C$ if there exists $a > 0$ such that the set $C \cap \mathcal{B}_a(z)$ is closed.

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

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

In the proof of our main result (Theorem 3.1) we employ the following set-valued generalization of the Banach fixed point theorem proved in [5]:

Lemma 2.3. (Fixed Point) *Let (X, ρ) be a complete metric space, and consider a set-valued mapping $\Phi : X \rightrightarrows X$, a point $\bar{x} \in X$, and nonnegative scalars α and θ be such that $0 \leq \theta < 1$, the sets $\Phi(x) \cap \mathcal{B}_\alpha(\bar{x})$ are closed for all $x \in \mathcal{B}_\alpha(\bar{x})$ and the following conditions hold:*

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

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

In the above, $e(A, B)$ denotes the excess from a set A to a set B : $e(A, B) = \sup_{x \in A} d(x, B)$.

3. CONVERGENCE OF THE PROXIMAL POINT ALGORITHM

Here we present the statement and a proof of our main result:

Theorem 3.1. *Consider a mapping $T : X \rightrightarrows Y$ and let \bar{x} be a solution of the inclusion $T(x) \ni 0$. Let $\text{gph } T$ be locally closed at $(\bar{x}, 0)$ and let T be metrically regular at \bar{x} for 0. Choose a sequence of functions $g_n : X \rightarrow Y$ with $g_n(0) = 0$ which are Lipschitz continuous in a neighborhood U of 0, the same for all n , with Lipschitz constants λ_n satisfying*

$$\sup_n \lambda_n < \frac{1}{2 \text{reg } T(\bar{x}|0)}. \quad (3.1)$$

Then there exists a neighborhood O of \bar{x} such that for any $x_0 \in O$ there exists a sequence x_n generated by the proximal point algorithms (1.2) which is linearly convergent to \bar{x} . Moreover, if g_n is chosen so that $\lambda_n \rightarrow 0$, then for any $x_0 \in O$ there exists a sequence x_n generated by (1.2) which is superlinearly convergent to \bar{x} .

Proof. Let $\lambda := \sup_n \lambda_n$, then from (3.1) there exists $\kappa > \text{reg } T(\bar{x}|0)$ such that $\kappa\lambda < 0.5$. Then one can choose γ which satisfies $((\kappa\lambda)^{-1} - 1)^{-1} < \gamma < 1$. Let a be such that the mapping T is metrically regular at \bar{x} for 0 with a constant κ and neighborhoods $\mathcal{B}_a(\bar{x})$ and $\mathcal{B}_{2\lambda a}(0)$ and $\mathcal{B}_{2a}(0) \subset U$.

Pick $x_0 \in \mathcal{B}_a(\bar{x})$. For any $x \in \mathcal{B}_a(\bar{x})$ we have

$$\| -g_0(x - x_0) \| = \| g_0(x - x_0) - g_0(0) \| \leq \lambda_0 \| x_0 - x \| \leq 2a\lambda_0 \leq 2\lambda a.$$

We will now show that the mapping $\Phi_0(x) = T^{-1}(-g_0(x - x_0))$ satisfies the assumptions of the fixed-point result in Lemma 2.3. First, by using the assumptions that T is metrically regular at \bar{x} for 0, $0 \in T(\bar{x})$ and $g_n(0) = 0$, we have

$$\begin{aligned} d(\bar{x}, \Phi_0(\bar{x})) &= d(\bar{x}, T^{-1}(-g_0(\bar{x} - x_0))) \leq \kappa d(-g_0(\bar{x} - x_0), T(\bar{x})) \\ &\leq \kappa \| g_0(0) - g_0(\bar{x} - x_0) \| \leq \kappa \lambda_0 \| x_0 - \bar{x} \| \\ &\leq \kappa \lambda_0 a < a(1 - \kappa \lambda_0). \end{aligned}$$

Further, for any $u, v \in \mathcal{B}_a(\bar{x})$, by the metric regularity of T ,

$$\begin{aligned} e\left(\Phi_0(u) \cap \mathcal{B}_a(\bar{x}), \Phi_0(v)\right) &= \sup_{x \in T^{-1}(-g_0(u-x_0)) \cap \mathcal{B}_a(\bar{x})} d(x, T^{-1}(-g_0(v-x_0))) \\ &\leq \sup_{x \in T^{-1}(-g_0(u-x_0)) \cap \mathcal{B}_a(\bar{x})} \kappa d(-g_0(v-x_0), T(x)) \\ &\leq \kappa \| -g_0(u-x_0) - (-g_0(v-x_0)) \| \leq \kappa \lambda_0 \| u - v \|. \end{aligned}$$

Hence there exists a fixed point $x_1 \in \Phi_0(x_1) \cap \mathcal{B}_a(\bar{x})$, i.e.,

$$x_1 \in \mathcal{B}_a(\bar{x}) \text{ and } 0 \in g_0(x_1 - x_0) + T(x_1) \quad (3.2)$$

If $x_1 = \bar{x}$ there is nothing more to prove. Assume $x_1 \neq \bar{x}$. For any $x \in \mathcal{B}_a(\bar{x})$, we have

$$\| -g_1(x - x_1) \| \leq 2\lambda_1 a \leq 2\lambda a.$$

Let

$$\alpha_1 = \gamma \| x_1 - \bar{x} \|. \quad (3.3)$$

Since $\gamma < 1$ we have $\alpha_1 < a$. Consider the mapping $\Phi_1(x) = T^{-1}(-g_1(x - x_1))$. By (3.3), the metric regularity of T and the choice of γ

$$\begin{aligned} d(\bar{x}, \Phi_1(\bar{x})) &= d(\bar{x}, T^{-1}(-g_1(\bar{x} - x_1))) \leq \kappa d(-g_1(\bar{x} - x_1), T(\bar{x})) \\ &\leq \kappa \| -g_1(\bar{x} - x_1) + g_1(0) \| \leq \kappa \lambda_1 \| x_1 - \bar{x} \| \\ &< \alpha_1(1 - \kappa \lambda_1). \end{aligned}$$

For $u, v \in \mathcal{B}_{\alpha_1}(\bar{x})$, again by the metric regularity of T , we obtain

$$\begin{aligned} e\left(\Phi_1(u) \cap \mathcal{B}_{\alpha_1}(\bar{x}), \Phi_1(v)\right) &\leq \sup_{x \in T^{-1}(-g_1(u-x_1)) \cap \mathcal{B}_{\alpha_1}(\bar{x})} d(x, T^{-1}(-g_1(v-x_1))) \\ &\leq \sup_{x \in T^{-1}(-g_1(u-x_1)) \cap \mathcal{B}_{\alpha_1}(\bar{x})} \kappa d(-g_1(v-x_1), T(x)) \\ &\leq \kappa \| -g_1(u-x_1) - (-g_1(v-x_1)) \| \leq \kappa \lambda_1 \| u - v \|. \end{aligned}$$

Hence, by Lemma 2.3, there exists $x_2 \in \Phi_1(x_2) \cap B_{\alpha_1}(\bar{x})$ which, by (3.3), satisfies

$$\|x_2 - \bar{x}\| \leq \gamma \|x_1 - \bar{x}\|.$$

The induction step is now clear. If $x_n \in B_a(\bar{x})$, we obtain $\| -g_n(x - x_n) \| \leq 2\lambda a$. Then for $\alpha_n = \gamma \|x_n - \bar{x}\|$, by applying Lemma 2.3 to $\Phi_n(x) = T^{-1}(-g_n(x - x_n))$ we obtain the existence of $x_{n+1} \in B_{\alpha_n}(\bar{x})$ such that $0 \in g_n(x_{n+1} - x_n) + T(x_{n+1})$. Thus, we establish that

$$\|x_{n+1} - \bar{x}\| \leq \gamma \|x_n - \bar{x}\| \quad \text{for all } n. \tag{3.4}$$

Since $\gamma < 1$, the sequence x_n converges linearly to \bar{x} .

Let $\lambda_n \rightarrow 0$. Choosing the sequence

$$\gamma_n = \frac{n}{n+1} ((\kappa\lambda_n)^{-1} - 1)^{-1} + \frac{1}{n+1}$$

and repeating the above argument with γ_n instead γ at each step, we obtain (3.4) with γ_n , and since $\gamma_n \rightarrow 0$, the proof is complete. \square

In view of the Extended Lyusternik-Graves theorem, the proximal point algorithm (1.2) represents an iteratively applied perturbation of the mapping T which is small enough to preserve the metric regularity of T . Note that, trivially, the functions g_n in (1.2) can be chosen identically zero, and this does not contradict Theorem 3.1. On the other hand, the possibility for choosing the sequence g_n gives more freedom to enhance the convergence and the method.

4. STRONG SUBREGULARITY AND STRONG REGULARITY

Another concept which has been recognized as a basic regularity property much later than the metric regularity was studied in [10] under the name of *strong metric subregularity*.

Definition 4.1. A mapping $F : X \rightrightarrows Y$ is strongly subregular at \bar{x} for \bar{y} if $F(\bar{x}) \ni \bar{y}$ and there exists $\kappa \in [0, \infty)$ along with neighborhoods U of \bar{x} and V of \bar{y} such that

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

This property is equivalent to the ‘‘local Lipschitz property at a point’’ of the inverse mapping, a property first formally introduced in [4] where a stability result parallel to the Lyusternik-Graves theorem was proved. By a theorem of Robinson, a mapping F acting in finite dimensions whose graph is the union of polyhedral sets is strongly subregular at \bar{x} for \bar{y} if and only if \bar{x} is an isolated point of $F^{-1}(\bar{y})$. At the end of this section we will give another illustration of the strong subregularity.

Next comes a result regarding the convergence of the proximal point algorithm (1.2) for strongly subregular mappings:

Theorem 4.2. Consider \bar{x} be a solution of the inclusion $T(x) \ni 0$ and let T be strongly subregular at \bar{x} for 0 with a constant κ . Choose a sequence of functions $g_n : X \rightarrow Y$ with $g_n(0) = 0$ which are Lipschitz continuous in a neighborhood U of 0, the same for all n , with a Lipschitz constant λ_n and such that $\kappa\lambda_n < 1$ for $n = 0, 1, 2, \dots$. Then there exists a neighborhood O of \bar{x} such that for every sequence x_n generated by the iteration (1.2) and whose elements are in O for all n we have

$$\|x_{n+1} - \bar{x}\| \leq ((\kappa\lambda_n)^{-1} - 1)^{-1} \|x_n - \bar{x}\| \quad \text{for } n = 0, 1, 2, \dots \tag{4.1}$$

If, in addition, $\sup_n \kappa\lambda_n < 0.5$, the sequence x_n is linearly convergent to \bar{x} . Moreover, if g_n is chosen so that $\lambda_n \rightarrow 0$, then x_n is superlinearly convergent to \bar{x} .

Proof. Let a and b be such that the mapping T is strongly metrically subregular at \bar{x} for 0 with constant κ and neighborhoods $\mathcal{B}_a(\bar{x})$ and $\mathcal{B}_b(0)$ and $\mathcal{B}_{2a}(0) \subset U$. Adjust a if necessary so that $2\lambda_n a \leq b$ for all n . Then, by the definition of the strong subregularity, we have

$$\|x - \bar{x}\| \leq \kappa d(0, T(x) \cap \mathcal{B}_b(0)) \quad \text{for all } x \in \mathcal{B}_a(\bar{x}).$$

Now, suppose that (1.2) generates a sequence x_n such that $x_n \in \mathcal{B}_a(\bar{x})$ for all n . Then

$$\| -g_n(x_{n+1} - x_n) \| = \|g_n(0) - g_n(x_{n+1} - x_n)\| \leq \lambda_n \|x_{n+1} - x_n\| \leq 2\lambda_n a \leq b.$$

Since $-g_n(x_{n+1} - x_n) \in T(x_{n+1}) \cap \mathcal{B}_b(0)$, we obtain

$$\begin{aligned} \|x_{n+1} - \bar{x}\| &\leq \kappa \| -g_n(x_{n+1} - x_n) \| \leq \kappa \lambda_n \|x_{n+1} - x_n\| \\ &\leq \kappa \lambda_n (\|x_{n+1} - \bar{x}\| + \|x_n - \bar{x}\|). \end{aligned}$$

This gives us (4.1). The remaining two statements follow directly from (4.1). \square

In order to introduce the next regularity property, we need the notion of graphical localization. A *graphical localization* of a mapping $F : X \rightrightarrows Y$ at $(\bar{x}, \bar{y}) \in \text{gph } F$ is a mapping $\tilde{F} : X \rightrightarrows Y$ such that $\text{gph } \tilde{F} = (U \times V) \cap \text{gph } F$ for some neighborhood $U \times V$ of (\bar{x}, \bar{y}) .

Definition 4.3. 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 κ such that, in addition, the graphical localization of F^{-1} with respect to U and V is single-valued.

Strong regularity implies metric regularity by definition. The simplest case is the linear and bounded mapping from \mathbb{R}^n to itself represented by an $n \times n$ matrix A . This mapping is strongly regular precisely when the matrix A is nonsingular, and metrically regular when A is merely surjective. But if a square matrix is surjective, that is, of full rank, it must be nonsingular. We see here an instance of a mapping when *metric regularity automatically implies strong regularity*, that is, metric regularity and strong regularity are equivalent. It turns out, see [7], that this equivalence holds for more general mappings of the form of the sum of a smooth function and the normal cone mapping over a polyhedral convex set. On a related matter, strong regularity is stable under small Lipschitz perturbations in the sense of the extended Lyusternik-Graves theorem; the corresponding result is due to Robinson, for more see [10].

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 deeper result by Kenderov [14] regarding single-valuedness of lower semicontinuous monotone mappings. A version of this result for mapping having the Aubin property was proved in [6] and later generalized for ‘‘premonotone’’ mappings in [15]. Thus, for locally premonotone mappings, the metric regularity is equivalent to the strong regularity. One should also note that the strong monotonicity of a mapping implies strong regularity, and this is another link of the present work to papers dealing with convergence of the proximal point method or some of its modifications for strongly monotone mappings.

When the mapping T is strongly regular, the convergent sequence x_n whose existence is proved in Theorem 3.1 becomes unique. The precise result is as follows:

Theorem 4.4. *Consider a mapping $T : X \rightrightarrows Y$ and let \bar{x} be a solution of the inclusion $T(x) \ni 0$. Let $\text{gph } T$ be locally closed at $(\bar{x}, 0)$ and let T be strongly regular at \bar{x} for 0 . Choose a sequence of functions $g_n : X \rightarrow Y$ with $g_n(0) = 0$ which are Lipschitz continuous in a neighborhood U of 0 , the same for all n , with a Lipschitz constant λ_n such that the inequality (3.1) holds. Then there exists a neighborhood O of \bar{x} such that for any $x_0 \in O$ there exists exactly one sequence x_n in O generated by the proximal point algorithm (1.2); moreover, this sequence is linearly convergent to \bar{x} . If in addition g_n is chosen so that $\lambda_n \rightarrow 0$, then for any $x_0 \in O$ there exists a unique in O sequence x_n generated by (1.2) which is superlinearly convergent to \bar{x} .*

Proof. One can prove this result in parallel to the proof of Theorem 3.1 by using the standard contraction mapping theorem instead of the one for set-valued mapping. Another way to prove it is to use Theorem 3.1

and to show that the sequence x_n , which is shown to exist in the ball $B_a(\bar{x})$, is unique in this ball, where a is chosen to match the size of the neighborhood of \bar{x} in the definition of strong regularity. Indeed, suppose that, from a given x_n there exist two points x_{n+1} and x'_{n+1} in $B_a(\bar{x})$ that are obtained by (1.2). Then, the assumed strong regularity of T implies that x_{n+1} satisfies

$$x_{n+1} = T^{-1}(-g_n(x_{n+1} - x_n)) \cap B_a(\bar{x})$$

and the same for x'_{n+1} . Let κ and λ_n as in the proof of Theorem 3.1. Then

$$\|x_{n+1} - x'_{n+1}\| \leq \kappa \lambda_n \|x_{n+1} - x'_{n+1}\| < \|x_{n+1} - x'_{n+1}\|,$$

which is absurd. □

For an illustration, consider the optimization problem

$$\min f(x) \quad \text{subject to } x \in C, \tag{4.2}$$

where C is a nonempty polyhedral (and hence convex and closed) subset of \mathbb{R}^n and the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is twice continuously differentiable everywhere. Let \bar{x} be a local minimum of (4.2); then \bar{x} satisfies the *basic first-order necessary optimality condition*:

$$\nabla f(x) + N_C(x) \ni 0. \tag{4.3}$$

Any solution x of (4.3) is a *stationary point* for problem (4.2). If the function f is convex, the converse statement is also true: every stationary point is a local minimizer. When \bar{x} satisfies (4.3) and the *second-order sufficient condition*

$$\langle u, \nabla^2 f(\bar{x})u \rangle > 0 \quad \text{for all nonzero } u \in \bar{K}, \tag{4.4}$$

then \bar{x} is a local minimizer of (4.2). Here \bar{K} is the critical cone to C associated with \bar{x} defined as

$$\bar{K} = T_C(\bar{x}) \cap (\nabla f(\bar{x}))^\perp,$$

where $T_C(x)$ is the tangent cone to C at x . Having x to satisfy (4.3) and (4.4) is equivalent to the existence of $\alpha > 0$ and $\varepsilon > 0$ such that

$$f(x) \geq f(\bar{x}) + \alpha|x - \bar{x}|^2 \quad \text{for all } x \in C \text{ with } |x - \bar{x}| \leq \varepsilon. \tag{4.5}$$

The condition (4.5) is known as the *quadratic growth condition*.

Finally, we say that the *strong second-order sufficient condition* for local optimality holds at \bar{x} when

$$\langle u, \nabla^2 f(x)u \rangle > 0 \quad \text{for all nonzero } u \in \bar{K} - \bar{K}. \tag{4.6}$$

Recall that $\bar{K} - \bar{K}$ is the smallest subspace of \mathbb{R}^n that contains the critical cone \bar{K} .

The next two results that follow from [8] show the close connection between the properties of subregularity and regularity and the optimality conditions for the problem (4.2).

Proposition 4.5. *Let \bar{x} solve (4.3), that is, \bar{x} is a stationary point of problem (4.2). Then the following are equivalent:*

- (i) *The second-order sufficient condition (4.4) holds at \bar{x} ;*
- (ii) *The point \bar{x} is a local minimizer of (4.2) and the mapping $\nabla f(\bar{x}) + N_C$ is strongly subregular at \bar{x} for 0.*

We can add to the two equivalent conditions in Proposition 4.6 the quadratic growth condition (4.5), which sheds some more light on the result.

The second proposition exhibits once more the striking parallel between the second-order sufficient optimality and the metric regularity of minimizers.

Proposition 4.6. *Let \bar{x} solve (4.3). Then the following are equivalent:*

- (i) *The strong second-order sufficient condition (4.6) holds at \bar{x} ;*
- (ii) *The mapping $\nabla f(\bar{x}) + N_C$ is strongly regular at \bar{x} for 0 and \bar{x} is a local minimum.*

One can now easily translate the convergence results for the proximal point algorithm obtained in this paper for solving the optimization problem (4.2).

REFERENCES

- [1] P. N. ANH, L. D. MUU, V. H. NGUYEN, J. J. STRODIOT, Using the Banach contraction principle to implement the proximal point method for multivalued monotone variational inequalities. *J. Optim. Theory Appl.*, **124** (2005) 285–306.
- [2] A. AUSLENDER, M. TEBoulLE, Lagrangian duality and related multiplier methods for variational inequality problems. *SIAM J. Optim.*, **10** (2000) 1097–1115.
- [3] H. H. BAUSCHKE, J. V. BURKE, F. R. DEUTSCH, H. S. HUNDAL, J. D. VANDERWERFF, A new proximal point iteration that converges weakly but not in norm. *Proc. Amer. Math. Soc.*, **133** (2005) 1829–1835.
- [4] A. L. DONTCHEV, Characterizations of Lipschitz stability in optimization. *Recent developments in well-posed variational problems*, 95–115, *Math. Appl.*, 331, Kluwer Acad. Publ., Dordrecht, 1995.
- [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. L. DONTCHEV, W. W. HAGER, Implicit functions, Lipschitz maps, and stability in optimization. *Math. Oper. Res.*, **19** (1994) 753–768.
- [7] A. L. DONTCHEV, R. T. ROCKAFELLAR, Characterization of strong regularity of variational inequalities over polyhedral convex sets, *SIAM J. Optim.*, **6** (1996) 1087–1105.
- [8] A. L. DONTCHEV, R. T. ROCKAFELLAR, Ample parameterization of variational inclusions. *SIAM J. Optim.*, **12** (2001) 170–187.
- [9] A. L. DONTCHEV, A. S. LEWIS, R. T. ROCKAFELLAR, The radius of metric regularity, *Trans. AMS.*, **355** (2002) 493–517.
- [10] A. L. DONTCHEV, R. T. ROCKAFELLAR, Regularity and conditioning of solution mappings in variational analysis, *Set-Valued Anal.*, **12** (2004) 79–109.
- [11] A. D. IOFFE, Metric regularity and subdifferential calculus, *Uspekhi Mat. Nauk*, **55** (2000), no. 3 (333) 103–162; English translation *Math. Surveys*, **55** (2000) 501–558.
- [12] A. N. IUSEM, T. PENNANEN, B. F. SVAITER, Inexact variants of the proximal point algorithm without monotonicity. *SIAM J. Optim.*, **13** (2003) 1080–1097.
- [13] A. KAPLAN, R. TICHATSCHKE, Proximal-based regularization methods and successive approximation of variational inequalities in Hilbert spaces. *Well-posedness in optimization and related topics (Warsaw, 2001)*. *Control Cybernet.*, **31** (2002) 521–544.
- [14] P. KENDEROV, Semi-continuity of set-valued monotone mappings. *Fund. Math.*, **88** (1975) 61–69.
- [15] A. B. LEVY, R. A. POLIQUIN, Characterizing the single-valuedness of multifunctions. *Set-Valued Anal.*, **5** (1997) 351–364.
- [16] B. MARTINET, Régularisation d'inéquations variationnelles par approximations successives. *Rev. Française Informat. Recherche Opérationnelle* **4** (1970) Ser. R-3, 154–158.
- [17] T. PENNANEN, Local convergence of the proximal point algorithm and multiplier methods without monotonicity. *Math. Oper. Res.*, **27** (2002) 170–191.
- [18] R. T. ROCKAFELLAR, Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.*, **14** (1976) 877–898.
- [19] R. T. ROCKAFELLAR, R. J.-B. WETS, *Variational Analysis*, Springer-Verlag, Berlin, 1997.
- [20] M. V. SOLODOV, B. F. SVAITER, A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Anal.*, **7** (1999) 323–345.
- [21] J. E. SPINGARN, Submonotone mappings and the proximal point algorithm. *Numer. Funct. Anal. Optim.*, **4** (1981/82) 123–150.
- [22] Z. YANG, B. HE, A relaxed approximate proximal point algorithm. *Ann. Oper. Res.*, **133** (2005) 119–125.