TAIWANESE JOURNAL OF MATHEMATICS
Vol. 18, No. 2, pp. 419-433, April 2014
DOI: 10.11650/tjm.18.2014.3066
This paper is available online at http://journal.taiwanmathsoc.org.tw

ON THE CONVERGENCE OF INEXACT PROXIMAL POINT ALGORITHM ON HADAMARD MANIFOLDS

P. Ahmadi and H. Khatibzadeh*

Abstract. In this paper we consider the proximal point algorithm to approximate a singularity of a multivalued monotone vector field on a Hadamard manifold. We study the convergence of the sequence generated by an inexact form of the algorithm. Our results extend the results of [3, 25] to Hadamard manifolds as well as the main result of [11] with more general assumptions on the control sequence. We also give some application to optimization.

1. INTRODUCTION

Let H be a real Hilbert space with inner product $\langle\ ,\ \rangle$ and norm $\|\ .\ \|$. A possibly multivalued mapping $A : H \longrightarrow 2^H$ is said to be monotone (resp. strongly monotone) operator provided that

$$\langle y_2-y_1, x_2-x_1\rangle \geqslant 0\,(resp.\ \geqslant \alpha \parallel x_1-x_2 \parallel^2)\,,\ \forall x_i \in \mathcal{D}(A),\ \forall y_i \in A(x_i)\,,\ i=1,2\,,$$

where α is a fixed positive real number and $\mathcal{D}(A)$ denotes the domain of A defined by $\mathcal{D}(A) := \{x \in H\ :\ A(x) \neq \varnothing\}$. A is maximal monotone if and only if A is monotone and $R(I + A) = H$, where I is the identity mapping of H. Given any function $\varphi : H \longrightarrow]-\infty, +\infty]$ (not necessarily convex) with the domain $\mathcal{D}(\varphi)$, its subdifferential is defined by

$$\partial\varphi(x) := \{w \in H \mid \varphi(x) - \varphi(y) \leqslant \langle w, x-y\rangle\ ,\ \forall y \in H\}.$$

The function φ is called proper if and only if there exists an $x \in H$ such that $\varphi(x) < +\infty$. It is a well-known result that if φ is a proper, convex and lower semicontinuous function, then $\partial\varphi$ is a maximal monotone operator. We refer the reader to the book

Received March 14, 2013, accepted August 22, 2013.
Communicated by Juan Enrique Martinez-Legaz.
2010 *Mathematics Subject Classification*: 47H05, 49J40.
Key words and phrases: Proximal point algorithm, Maximal monotone operator, Resolvent, Subdifferential, Convergence, Hadamard manifold.
*Corresponding author.

by Morosanu [21] in order to understand monotone operators and subdifferential of convex functions in Hilbert spaces.

One of the most important problems in maximal monotone operator theory is approximation of a zero of the maximal monotone operator and one of the most popular algorithms to find a zero of a maximal monotone operator is the proximal point algorithm. This algorithm was first proposed by Martinet [19] for convex functions. The proximal point algorithm for a maximal monotone operator $A : H \longrightarrow 2^H$, which has been introduced by Rockafellar [25], is the sequence generated by the following process

(1) $$y_n = (1 + \lambda_n A)^{-1}(y_{n-1} + e_n) \quad , \quad n = 1, 2, ...,$$

where $\{\lambda_n\}$ is a positive real sequence and $\{e_n\}$ is a sequence in Hilbert space H. The algorithm (1) is also a discretization of nonhomogeneous first order evolution equation of maximal monotone type. Rockafellar in his seminal paper [25] showed the weak convergence of the sequence $\{y_n\}$ generated by (1) to a zero of A, provided that $\lambda_n \geqslant \lambda > 0$, $\forall n \geqslant 1$, and $\sum_{n=1}^{\infty} \|e_n\| < +\infty$. Brézis and Lions [3] proved the weak convergence of the sequence $\{y_n\}$ with condition $\sum_{n=1}^{\infty} \lambda_n^2 = +\infty$ on the parameter sequence $\{\lambda_n\}$ and the same condition on the error sequence $\{e_n\}$. They also proved some other weak and strong convergence theorems with additional conditions on the maximal monotone operator A. Djafari Rouhani and Khatibzadeh [13] showed that the weak and strong convergence results of Brézis and Lions may be obtained without maximality assumption of the monotone operator A, and when the monotone operator is maximal, the weak and strong convergence point belongs to $A^{-1}(0)$. In fact they proved that the weak and strong convergence theorems for the sequence $\{y_n\}$ are valid without assuming $A^{-1}(0) \neq \varnothing$. The second author [17] and Zaslavski [29] studied the convergence of the sequence $\{y_n\}$ without summability assumption on the error $\{e_n\}$. Convergence analysis of a modified version of the proximal point algorithm under more general error sequence studied in [5, 18, 24, 28]. Recently, the monotone operators have been defined by Németh [23] in single valued case, and by Li, López, Márquez and Wang [11, 12] as well as by Iwamiya and Okochi [15] in set-valued case on Hadamard manifolds.

Since we aim to study multivalued monotone vector fields on Hadamard manifolds, we remind some indispensable backgrounds about Riemannian manifolds from [16] and [26].

Let M be a complete and connected m-dimensional Riemannian manifold, with a Riemannian metric $\langle , \rangle$ and the corresponding norm denoted by $\|.\|$. For $p \in M$ the tangent space at p is denoted by T_pM and the tangent bundle of M by TM. Throughout the paper we assume that M is a complete, simply connected Riemannian manifold of non-positive sectional curvature of dimension m, which is called a Hadamard manifold of dimension m.

Proposition 1.1. ([26, p. 221]). *Let $p \in M$. Then $\exp_p : T_pM \to M$ is a diffeomorphism, and for any two points $p, q \in M$ there exists a unique normalized geodesic joining p to q, which is, in fact, a minimal geodesic.*

Let $[a, b]$ be a closed interval in $\mathbb{R}$, $\gamma : [a, b] \to M$ a smooth curve. The length of γ is defined as

$$L(\gamma) := \int_a^b \|\dot{\gamma}(t)\| dt$$

and the Riemannian distance $d(p, q)$ is defined by

$$d(p, q) := inf\{L(\gamma) |\ \gamma : [0, 1] \to M \text{ is a piecewise smooth curve with } \gamma(0) = p\ , \gamma(1) = q\}\ ,$$

which induces the original topology on M. Furthermore, $d(p, q) = \| \exp_p^{-1} q\|$, for any two points $p, q \in M$ (see [26]).

A geodesic joining p to q in M is said to be minimal if its length equals $d(p, q)$. By definition, a geodesic triangle $\Delta(p_1p_2p_3)$ of a Riemannian manifold is a set consisting of three points p_1, p_2 and p_3, and three minimal geodesics joining these points.

Proposition 1.1 shows that any m-dimensional Hadamard manifold has the same topology and differential structure as the Euclidean space $\mathbb{R}^m$. In fact, Hadamard manifolds and Euclidean spaces have some similar geometrical properties. One of them is described in the following proposition.

Proposition 1.2. ([26, p. 223]) (Comparison theorem for triangles). *Let $\Delta(p_1p_2p_3)$ be a geodesic triangle. Denote by $\gamma_i : [0, l_i] \to M$ the geodesic joining p_i to p_{i+1}, and set $l_i := L(\gamma_i)$, $\alpha_i := \angle(\dot{\gamma}_i(0), -\dot{\gamma}_{i-1}(l_{i-1}))$, where $i = 1, 2, 3\ (mod\ 3)$. Then*

$$\begin{aligned} &\alpha_1 + \alpha_2 + \alpha_3 \leqslant \pi\ , \\ &l_i^2 + l_{i+1}^2 - 2l_il_{i+1} \cos\alpha_{i+1} \leqslant l_{i-1}^2\ . \end{aligned} \tag{2}$$

Since

$$\langle \exp_{p_{i+1}}^{-1} p_i, \exp_{p_{i+1}}^{-1} p_{i+2}\rangle = d(p_i, p_{i+1})d(p_{i+1}, p_{i+2}) \cos\alpha_{i+1}\ ,$$

so the inequality (2) may be rewritten as follows

$$d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2\langle \exp_{p_{i+1}}^{-1} p_i, \exp_{p_{i+1}}^{-1} p_{i+2}\rangle \leqslant d^2(p_{i+2}, p_i). \tag{3}$$

The monotone vector fields were first defined by Németh, [22], and the monotone point-to-set vector fields were first considered by Cruz Neto, Ferreira and Lucambio Pérez, [6]. For some important properties of monotone vector fields we refer to [7, 10]. The following definition extends some notions of the monotonicity, from the corresponding notions in Hilbert spaces (see [4, 20, 21, 30]), to multivalued vector fields on Hadamard manifolds. Let $\mathcal{X}(M)$ denote the set of all multivalued vector fields $A : M \longrightarrow 2^{TM}$ such that $A(x) \subseteq T_xM$ for each $x \in M$ and the domain $\mathcal{D}(A)$ of A is closed and convex, where $\mathcal{D}(A)$ is defined by

$$\mathcal{D}(A) = \{x \in M : \; A(x) \neq \varnothing\} \ .$$

Definition 1.3. ([11]). Let $A \in \mathcal{X}(M)$. Then A is said to be

(i) *monotone* if the following condition holds for any $x, y \in \mathcal{D}(A)$:

$$\langle u, \exp_x^{-1} y\rangle \leqslant \langle v, -\exp_y^{-1} x\rangle \ , \quad \forall u \in A(x) \quad and \quad \forall v \in A(y) \ ; \tag{4}$$

(ii) *strongly monotone* if there exists $\rho > 0$ such that, for any $x, y \in \mathcal{D}(A)$, we have

$$\langle u, \exp_x^{-1} y\rangle - \langle v, -\exp_y^{-1} x\rangle \leqslant -\rho d^2(x, y) \ , \quad \forall u \in A(x) \quad and \quad \forall v \in A(y) \ ; \tag{5}$$

(iii) *maximal monotone* if it is monotone and the following implication holds for any $x \in \mathcal{D}(A)$ and $u \in T_xM$:

$$\langle u, \exp_x^{-1} y\rangle \leqslant \langle v, -\exp_y^{-1} x\rangle \ , \quad \forall y \in \mathcal{D}(A) \quad and \quad v \in A(y) \Longrightarrow u \in A(x). \tag{6}$$

Iwamiya and Okochi have introduced an alternative definition of monotonicity in terms of the distance function between geodesics (see [15]). It has been proved that both definitions are equivalent (see [11]).

Definition 1.4. ([11]). Let $A \in \mathcal{X}(M)$ and $x_0 \in \mathcal{D}(A)$. Then A is said to be upper Koratowski semicontinuous at x_0 if, for any sequences $\{x_k\} \subseteq \mathcal{D}(A)$ and $\{u_k\} \subset TM$ with each $u_k \in A(x_k)$, the relations $lim_{k\to\infty} x_k = x_0$ and $lim_{k\to\infty} u_k = u_0$ imply that $u_0 \in A(x_0)$. A is said to be upper Koratowski semicontinuous on M if it is upper Koratowski semicontinuous at each point $x_0 \in \mathcal{D}(A)$.

In [11, Proposition 3.5] it has been shown that each maximal monotone vector field is upper Koratowski semicontinuous.

Definition 1.5. Let X be a metric space. A map $T : X \to X$ is called nonexpansive if $d(T(x), T(y)) \leqslant d(x, y)$, for all $x, y \in X$.

Definition 1.6. ([12]). Given $\lambda > 0$ and $A \in \mathcal{X}(M)$, the resolvent of A of order λ is the set-valued mapping $J_\lambda : M \to 2^M$ defined by

$$J_\lambda(x) := \{z \in M \ : \ x \in \exp_z \lambda Az\} \ , \quad \forall x \in M \tag{7}$$

It is a well-known result that if $A \in \mathcal{X}(M)$, $\mathcal{D}(A) = M$, the vector field A is maximal monotone if and only if J_λ is single-valued and firmly nonexpansive and $\mathcal{D}(J_\lambda) = M$ (see [12]).

Ferreira and Oliveira in [9] introduced the exact (without error) proximal point algorithm

$$0 \in \lambda_n A(x_{n+1}) - \exp^{-1}_{x_{n+1}} x_n \quad , \quad n = 0, 1, 2, ..., \tag{8}$$

where $\{\lambda_n\}$ is a positive real sequence and $A : M \longrightarrow 2^{TM}$ is a multivalued monotone vector field. Some properties of the proximal sequence for finding singularities of vector fields were shown in [8], and recently, some important properties of the algorithm for optimization problem in Hadamard manifolds were established (see [1, 2, 27]).

Li, López and Márquez have studied the algorithm (8) in [11] as well. The main result of [11] is the convergence of $\{x_n\}$ to a singularity of the monotone vector field A when $\lambda_n \geqslant \lambda > 0$. It extends Rockafellar's result when $e_n \equiv 0$ on Hadamard manifolds setting. Our aim in this paper is to extend the convergence results of Brézis and Lions in Hadamard manifolds. We study the convergence of the sequence generated by

$$0 \in \lambda_n A(y_{n+1}) - \exp^{-1}_{y_{n+1}} y_n + e_n \quad , \quad n = 0, 1, 2, ... \tag{9}$$

to a singularity of the maximal monotone operator A under summability assumption on $\{e_n\}$ and appropriate assumptions on the sequence $\{\lambda_n\}$ and the maximal monotone operator A. Our results extend previous results of [3, 11, 25]. In [11] authors studied the convergence of the sequence given by (8) to a singularity of A under the assumption $\lambda_n \geqslant \lambda > 0$. In this paper our motivation is to study the convergence of the sequence $\{y_n\}$ given by (9) to a singularity of the monotone vector field A under the more general assumptions on the control parameter $\{\lambda_n\}$ and summability assumption on the error sequence $\{e_n\}$. Obviously, more freedom in choosing the parameters $\{\lambda_n\}$ and existence the error sequence in the algorithm given by (9) can be useful from practical and computational point of views. Note that the existence of the sequence $\{x_n\}$ in (8) is guaranteed by the maximal monotonicity of A and Remark 4.4-(ii) of [11].

2. Convergence Results in the Proximal Point Algorithm

We first recall the notion of Fejér convergence and the following related result from [14].

Definition 2.1. Let X be a complete metric space and $K \subseteq X$ be a nonempty set. A sequence $\{x_n\} \subset X$ is called Fejér convergent to K if

$$d(x_{n+1}, y) \leqslant d(x_n, y), \quad \forall y \in K \quad \textit{and} \quad n = 0, 1, 2, ... \, .$$

Lemma 2.2. *Let X be a complete metric space and $K \subseteq X$ be a nonempty set. Let $\{x_n\} \subset X$ be Fejér convergent to K and suppose that any cluster point of $\{x_n\}$ belongs to K. If the set of cluster points of $\{x_n\}$ is nonempty, then $\{x_n\}$ converges to a point of K.*

Theorem 2.3. *Let $A \in \mathcal{X}(M)$ be a maximal monotone vector field such that $A^{-1}(0) \neq \varnothing$. Suppose that $x_0 = y_0 \in \mathcal{D}(A)$. Assume that the sequences $\{x_n\}$ and $\{y_n\}$ are generated by the algorithms (8) and (9), respectively. If $\{x_n\}$ converges to a singularity of A, then $\{y_n\}$ does.*

Proof. Let $\{y_n\}$ converge to a singularity of A, then by (9) we have

$$\exp_{y_k}^{-1} y_{k-1} - e_k \in \lambda_k A(y_k) \quad , \quad k = 1, 2, \dots . \tag{10}$$

For every fixed k, consider the sequence $\{\xi_n(k)\}$ defined by

$$\xi_0(k) = y_k \quad , \quad \xi_1(k) = J_{\lambda_{k+1}}(y_k) \quad , \dots , \quad \xi_n(k) = J_{\lambda_{k+n}}(\xi_{n-1}(k)) \ .$$

By Theorem 4 of [12], J_λ is nonexpansive, so

$$\begin{aligned} d(\xi_n(k), \xi_{n+1}(k-1)) &= d(J_{\lambda_{k+n}}(\xi_{n-1}(k)), J_{\lambda_{k+n}}(\xi_n(k-1))) \\ &\leqslant d(\xi_{n-1}(k), \xi_n(k-1)) \\ &\leqslant \dots \leqslant d(\xi_0(k), \xi_1(k-1)) \\ &= d(y_k, J_{\lambda_k}(y_{k-1})) \end{aligned}$$

By definition of $J_{\lambda_k}(y_{k-1})$ we have

$$\exp_{J_{\lambda_k}(y_{k-1})}^{-1} y_{k-1} \in \lambda_k A J_{\lambda_k}(y_{k-1}) \ . \tag{11}$$

This together (4) and (10) imply that

$$\langle \exp_{J_{\lambda_k}(y_{k-1})}^{-1} y_{k-1}, \exp_{J_{\lambda_k}(y_{k-1})}^{-1} y_k \rangle \leqslant \langle \exp_{y_k}^{-1} y_{k-1} - e_k, -\exp_{y_k}^{-1} J_{\lambda_k}(y_{k-1}) \rangle$$

Hence by (3), we get

$$d(y_k, J_{\lambda_k}(y_{k-1})) \leqslant \|e_k\|.$$

Therefore

$$d(\xi_n(k), \xi_{n+1}(k-1)) \leqslant \|e_k\|. \tag{12}$$

By the assumption $\{\xi_n(k)\}$ converges to some $\xi(k)$, so (12) implies that

$$d(\xi(k), \xi(k-1)) \leqslant \|e_k\|.$$

Hence $\{\xi(k)\}$ is a Cauchy sequence and so $\{\xi(k)\}$ converges to some a. On the other hand

$$d(y_k, \xi_{n+1}(k-n-1)) \leqslant d(y_k, \xi_1(k-1)) + d(\xi_1(k-1), \xi_2(k-2)) + \cdots + d(\xi_n(k-n), \xi_{n+1}(k-n-1)) \leqslant \sum_{i=k-n}^{k} \|e_i\| \quad , \quad \forall k > n \ ,$$

hence

$$d(y_{n+k}, \xi_{n+1}(k-1)) \leqslant \sum_{i=k}^{k+n} \|e_i\|.$$

By the triangular inequality

$$d(y_{k+n}, a) \leqslant d(y_{k+n}, \xi_{n+1}(k-1)) + d(\xi_{n+1}(k-1), \xi(k-1)) + d(\xi(k-1), a).$$

Taking limsup when $n \to +\infty$ from both sides of this inequality, we get that

$$limsup_{n\to+\infty} d(y_n, a) \leqslant \sum_{i=k}^{+\infty} \|e_i\| + d(\xi(k-1), a).$$

Now the theorem is proved by letting $k \to +\infty$. ∎

Theorem 2.4. *Let $A \in \mathcal{X}(M)$ be maximal monotone such that $A^{-1}(0) \neq \varnothing$. Let $\{\lambda_n\}$ be a sequence of positive real numbers with*

(13) $$\sum_{n=1}^{+\infty} \lambda_n^2 = +\infty \ .$$

If $y_0 \in \mathcal{D}(A)$, then the sequence $\{y_n\}$ generated by (9) converges to a singularity of A.

Proof. By Theorem 2.3 it is enough to show the convergence of the sequence $\{x_n\}$, defined by (8), to a singularity of A. For this purpose, we show that the sequence $\{x_n\}$ is Fejér convergent to $A^{-1}(0)$ and any cluster point of $\{x_n\}$ belongs to $A^{-1}(0)$, then one gets the result by Lemma 2.2. Let $x \in A^{-1}(0)$. By (8), we get

(14) $$\lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n \in A(x_{n+1}) \quad , \quad n = 0, 1, 2, \ldots \ .$$

Hence the monotonicity of A, for $u = \lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n \in A(x_{n+1})$ and $v = 0 \in A(x)$, implies that

(15) $$\langle \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x \rangle \leqslant 0 \ .$$

Consider the geodesic triangle $\Delta(x_n x_{n+1} x)$. By inequality (3) of the comparison theorem for triangles, one obtains

$$d^2(x_{n+1}, x) + d^2(x_{n+1}, x_n) - 2\langle \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x\rangle \leqslant d^2(x_n, x) \ .$$

It follows from (15) that

$$d^2(x_{n+1}, x) \leqslant d^2(x_n, x)$$

which shows that $\{x_n\}$ is Fejér convergent to $A^{-1}(0)$.

Now, we show that any cluster point of $\{x_n\}$ belongs to $A^{-1}(0)$. Let x' be a cluster point of $\{x_n\}$. Then there exists a subsequence $\{n_k\}$ of $\{n\}$ such that $x_{n_k} \to x'$. Let

$$u_n := \lambda_n^{-1} \exp_{x_n}^{-1} x_{n-1} \quad , \quad n = 1, 2, \ldots \ .$$

Hence $u_n \in A(x_n)$ for each $n \geqslant 1$ by (8). We claim that $u_n \to 0$. The monotonicity of A implies that

$$\langle u_{n+1}, \exp_{x_{n+1}}^{-1} x_n\rangle \leqslant \langle u_n, -\exp_{x_n}^{-1} x_{n+1}\rangle \quad , \quad n = 1, 2, \ldots \ .$$

Hence

$$\begin{aligned} \lambda_{n+1}\|u_{n+1}\|^2 &\leqslant \|u_n\|\|\exp_{x_n}^{-1} x_{n+1}\| \\ &= \|u_n\|\|\exp_{x_{n+1}}^{-1} x_n\| \\ &= \|u_n\|\lambda_{n+1}\|u_{n+1}\|, \end{aligned}$$

which shows that the sequence $\{\|u_n\|\}$ is nonincreasing. The inequality (3), in the geodesic triangle $\Delta(x_n x_{n+1} x)$, and the inequality (15) show that

$$d^2(x_{n+1}, x_n) \leqslant d^2(x_n, x) - d^2(x, x_{n+1}) \ .$$

Hence

$$\lambda_{n+1}^2\|u_{n+1}\|^2 \leqslant d^2(x_n, x) - d^2(x, x_{n+1}).$$

Summing up from $n = 0$ to $n = k$, and by the fact that $\{\|u_n\|\}$ is a nonincreasing sequence, we get that

$$\|u_{k+1}\|^2 \sum_{n=0}^{k} \lambda_{n+1}^2 \leqslant d^2(x_0, x) < \infty \ . \tag{16}$$

By taking limit from the both sides of the inequality (16), when $k \to +\infty$, and using (13) one gets that $u_n \to 0$.

Thus the subsequence $\{u_{n_k}\}$ of $\{u_n\}$ converges to 0 as well. Since $x_{n_k} \to x'$ and A is upper Kuratowski semicontinuous at x' by Proposition 3.5 of [11], $0 \in A(x')$, that is, $x' \in A^{-1}(0)$. Now the theorem is proved by Lemma 2.2. ∎

Theorem 2.5. *Let $A \in \mathcal{X}(M)$ be such that $A^{-1}(0) \neq \varnothing$. Suppose that A is maximal monotone and strongly monotone. Let $\{\lambda_n\}$ be a sequence of positive real numbers such that*

$$\sum_{n=0}^{\infty} \lambda_n = +\infty \ . \tag{17}$$

Let $y_0 \in \mathcal{D}(A)$, then the sequence $\{y_n\}$ generated by (9) converges to a singularity of A.

Proof. By Theorem 2.3 we only need to prove the convergence of the sequence $\{x_n\}$, defined by (8), to a singularity of A. Let $x \in A^{-1}(0)$, so $0 \in A(x)$ and $\lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n \in A(x_{n+1})$ by (8). The strong monotonicity of A and (8) imply that

$$\langle \lambda_n^{-1} \exp_{x_n}^{-1} x_{n-1}, \exp_{x_n}^{-1} x \rangle \leqslant -\rho d^2(x_n, x) \qquad n = 1, 2, \dots \ . \tag{18}$$

Consider the geodesic triangle $\Delta(x_{n-1} x_n x)$. By inequality (3) of the comparision theorem for triangles, we have that

$$d^2(x_{n-1}, x_n) + d^2(x_n, x) - 2\langle \exp_{x_n}^{-1} x_{n-1}, \exp_{x_n}^{-1} x \rangle \leqslant d^2(x, x_{n-1}) \ .$$

It follows from (18) that

$$2\rho \lambda_n d^2(x_n, x) \leqslant d^2(x, x_{n-1}) - d^2(x_n, x) \qquad n = 1, 2, \dots$$

and

$$d^2(x_n, x) \leqslant d^2(x_{n-1}, x) \quad n = 1, 2, \dots \ . \tag{19}$$

Hence

$$2\rho \sum_{n=1}^{k} \lambda_n d^2(x_n, x) \leqslant \sum_{n=1}^{k} (d^2(x, x_{n-1}) - d^2(x_n, x)) \ .$$

Combining this with (19), we obtain

$$2\rho d^2(x_k, x) \sum_{n=1}^{k} \lambda_n \leqslant d^2(x, x_\circ) \qquad k = 1, 2, \dots \ .$$

This together (17) implies that $lim_{k\to\infty} d^2(x_k, x) = 0$, that is $x_n \to x$, as $n \to +\infty$. ■

Definition 2.6. A map $A : M \to 2^{TM}$ is called demipositive if there exists $x_\circ \in A^{-1}(0)$ such that $\Omega(x_\circ) \subset A^{-1}(0)$, where $\Omega(x_\circ)$ is the set of all $p \in M$ for which there are sequences $\{p_n\} \subset M$ and $\{\omega_n\} \subset TM$ such that $\omega_n \in A(p_n)$, $p_n \to p$, $\langle \omega_n, -\exp_{p_n}^{-1} x_\circ \rangle \to 0$ and $\{\|\omega_n\|\}$ is a bounded sequence.

Theorem 2.7. *Let $A \in \mathcal{X}(M)$ be a maximal monotone and demipositive multi-valued vector field. Suppose that $\{\lambda_n\}$ is a sequence of positive real numbers such that*

$$\sum_{n=1}^{\infty} \lambda_n = +\infty \ . \tag{20}$$

Let $y_0 \in \mathcal{D}(A)$, then the sequence $\{y_n\}$ generated by (9) converges to a singularity of A.

Proof. By Theorem 2.3, we only need to prove that the sequence $\{x_n\}$, defined by (8), is convergent to a singularity of A. Let

$$u_n := \lambda_n^{-1} \exp_{x_n}^{-1} x_{n-1} \quad , \quad n = 1, 2, \ldots \ .$$

Hence $u_n \in A(x_n)$ for each $n \geqslant 1$ by (8). Since A is monotone, the same proof of that of Theorem 2.4 shows that $\{x_n\}$ is Fejér convergent to $A^{-1}(0)$, and $\{\|u_n\|\}$ is a bounded sequence. By Lemma 2.2, we need only to show that any cluster point of $\{x_n\}$ belongs to $A^{-1}(0)$. Since A is demipositive, there exists $x_\circ \in A^{-1}(0)$ such that $\Omega(x_\circ) \subset A^{-1}(0)$. First we verify the following assertion. For any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\forall n \geqslant N \ , \ \exists m \in \mathbb{N} \ , \ N \leqslant m \leqslant n \ ; \ d(x_n, x_m) < \varepsilon \ \ \textit{and} \ \langle u_m, -\exp_{x_m}^{-1} x_\circ \rangle < \varepsilon \ .$$

By (8), we obtain

$$\lambda_k \langle u_k, \exp_{x_k}^{-1} x_\circ \rangle = \langle \exp_{x_k}^{-1} x_{k-1}, \exp_{x_k}^{-1} x_\circ \rangle, \qquad k = 1, 2, \ldots \ ,$$

and so

$$\lambda_k \langle u_k, -\exp_{x_k}^{-1} x_\circ \rangle \leqslant \frac{1}{2} d^2(x_{k-1}, x_\circ) - \frac{1}{2} d^2(x_k, x_\circ) \qquad k = 1, 2, \ldots \ .$$

Hence

$$\sum_{k=1}^{+\infty} \lambda_k \langle u_k, -\exp_{x_k}^{-1} x_\circ \rangle < \infty \ . \tag{21}$$

Let

$$P_\varepsilon = \{k \in \mathbb{N} : \langle u_k, -\exp_{x_k}^{-1} x_\circ \rangle \geqslant \varepsilon\} \ .$$

Since $\sum_{k \in P_\varepsilon} \lambda_k < \infty$ by (21), so

$$\sum_{k \in P_\varepsilon} d(x_k, x_{k-1}) = \sum_{k \in P_\varepsilon} \| \exp_{x_k}^{-1} x_{k-1} \| = \sum_{k \in P_\varepsilon} \lambda_k \|u_k\| < \infty \ .$$

Thus there exists $N_1 \in \mathbb{N}$ such that

$$\sum_{k \geqslant N_1, \ k \in P_\varepsilon} d(x_k, x_{k-1}) < \varepsilon, \tag{22}$$

and so there exists $N \geqslant N_1$, by (21), such that

$$\langle u_N, -\exp_{x_N}^{-1} x_\circ \rangle < \varepsilon \ .$$

Hence for any $n \geqslant N$, if $n \notin P_\varepsilon$ then assume that $m = n$, and if $n \in P_\varepsilon$ then let m be the largest integer number $k < n$ such that $k \notin P_\varepsilon$. Therefore $m \geqslant N$ and $\{m+1, ..., n\} \subseteq P_\varepsilon$, and so by (22) we obtain that

$$d(x_n, x_m) \leqslant \sum_{k=m+1}^{n} d(x_k, x_{k-1}) < \varepsilon$$

and the assertion is proved.

Let x' be a cluster point of $A^{-1}(0)$. So there exists a subsequence $\{n_k\}$ of $\{n\}$ such that $x_{n_k} \to x'$. By the assertion just proved, there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $x_{n_j} \to x'$ and $\langle u_{n_j}, -\exp_{x_{n_j}}^{-1} x_\circ \rangle \to 0$. By demipositivity of A, one gets that $x' \in A^{-1}(0)$ and the proof is complete by Lemma 2.2. ∎

3. Application to Optimization

Recall that M is a Hadamard manifold. Let $f : M \to]-\infty, +\infty]$ be a proper, lower semicontinuous and geodesically convex function. The domain of f, $D(f) = \{x \in M \mid f(x) < \infty\}$, is a closed convex subset of M. Consider the following minimization problem.

$$Min_M f(x) \tag{23}$$

If f is defined on a finite dimensional Hilbert space H, then M can be the constraint set of minimization f on H. Then the problem (23) can be a constraint minimization problem. The subdifferential of f at x is defined by

$$\partial f(x) = \{u \in T_x M \ : \ \langle u, \exp_x^{-1} y \rangle \leqslant f(y) - f(x) \ , \ \forall y \in M\}.$$

If $\mathcal{D}(\partial f) \neq \varnothing$, the subdifferential $\partial f(.)$ is a monotone and upper Kuratowski semicontinuous multivalued vector field, and if $D(f) = M$, then ∂f is a maximal monotone vector field (see Theorem 5.1 of [11]).

We recall the following Lemma from [17] which is necessary to prove the following theorem.

Lemma 3.1. *Suppose that $\{a_n\}$ and $\{b_n\}$ be two positive real sequences such that $\{a_n\}$ is nonincreasing and converges to zero, and $\sum_{n=1}^{+\infty} a_n b_n < +\infty$. Then $(\sum_{k=1}^{n} b_k) a_n \to 0$ as $n \to +\infty$.*

Theorem 3.2. *Let f be a proper, lower semicontinuous, and convex function on M and $A = \partial f$. Let $\{\lambda_n\}$ be a sequence of positive real numbers such that*

$$\sum_{n=0}^{\infty} \lambda_n = +\infty \ . \tag{24}$$

Let $y_0 \in \mathcal{D}(A)$, then the sequence $\{y_n\}$ generated by (9) converges to a singularity of A, which is a minimum point of f (by definition of f). In addition, if $e_n \equiv 0$ then $f(y_n) - f(x) = o((\sum_{i=1}^{n} \lambda_i)^{-1})$, where x is a minimum point of f.

Proof. By Theorem 2.3, we need only to verify that the sequence $\{x_n\}$, defined by (8), is convergent to a singularity of A. For this purpose, we first prove that $\{x_n\}$ is Fejér convergent to $A^{-1}(0)$. Let $x \in A^{-1}(0)$ and $n \geqslant 0$. Then $0 \in A(x)$ and $\lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n \in A(x_{n+1})$ by (8). So the monotonicity of A implies that

$$\langle \lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x \rangle \leqslant \langle 0, -\exp_x^{-1} x_{n+1} \rangle = 0. \tag{25}$$

Consider the geodesic triangle $\Delta(x_n x_{n+1} x)$. By inequality (3) of the comparision theorem for triangles, we have that

$$d^2(x_{n+1}, x) + d^2(x_{n+1}, x_n) - 2\langle \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x \rangle \leqslant d^2(x_n, x) \ .$$

It follows from (25) that

$$d^2(x_{n+1}, x) \leqslant d^2(x_n, x) \qquad n = 0, 1, 2, \dots \ .$$

Thus $\{x_n\}$ is Fejér convergent to $A^{-1}(0)$. To complete the proof, we need only to prove that $A^{-1}(0)$ contains each cluster point of $\{x_n\}$ by Lemma 2.2. For this purpose, we first show that $\{f(x_n)\}$ is a nonincreasing sequence, and $f(x_n) \to f(x)$. Since $\lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n \in A(x_{n+1})$, so by definition of the subdifferential, we have

$$\langle \lambda_{n+1}^{-1} \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x_n \rangle \leqslant f(x_n) - f(x_{n+1}) \quad n = 0, 1, 2, \dots \ .$$

Hence $f(x_n) - f(x_{n+1}) \geqslant 0$ for each $n \geqslant 0$, and so $\{f(x_n)\}$ is a nonincreasing sequence.

By definition of the subdifferential and the inequality (3) in the geodesic triangle $\Delta(x_{n-1} x_n x)$, we obtain that

$$\begin{aligned} f(x_n) - f(x) &\leqslant -\langle \lambda_n^{-1} \exp_{x_n}^{-1} x_{n-1}, \exp_{x_n}^{-1} x \rangle \\ &= \frac{1}{2} \lambda_n^{-1} (d^2(x, x_{n-1}) - d^2(x_{n-1}, x_n) - d^2(x_n, x)). \end{aligned}$$

Hence

$$2\lambda_n (f(x_n) - f(x)) \leqslant d^2(x, x_{n-1}) - d^2(x_n, x), \tag{26}$$

and so

$$2(f(x_k) - f(x)) \sum_{n=0}^{k} \lambda_n \leqslant d^2(x_\circ, x) \qquad k = 1, 2, \dots .$$

Taking limit in the previous inequality when $k \to +\infty$ and using (24), we obtain that $f(x_k) \to f(x)$.

Now, let $y_\circ$ be a cluster point of $\{x_n\}$, so there exists a subsequence $\{n_k\}$ of $\{n\}$, such that $x_{n_k} \to y_\circ$, hence by the lower semicontinuity of f, we have

(27) $$f(y_\circ) \leqslant liminf_{k\to\infty} f(x_{n_k}) = f(x) \ .$$

On the other hand $0 \in A(x)$ implies that $x \in S_f$, where

$$S_f = \{x \in M : \ f(x) \leqslant f(y) \ , \ \forall y \in M\} \ .$$

Thus by (27), we obtain that

$$f(y_\circ) \leqslant f(y) \quad , \quad \forall y \in M \ .$$

This means that $0 \in A(y_\circ)$, that is, $y_\circ \in A^{-1}(0)$.

For the rate of convergence, summing up in the both sides of inequality (26) from $n = 1$ to $+\infty$; we get $\sum_{n=1}^{+\infty} \lambda_n (f(x_n) - f(x)) < \infty$. Now the result is obtained by Lemma 3.1 and the assumptions. ∎

Other applications in variational inequalities and saddle point problems can be found in [11].

Acknowledgments

The authors are grateful to the referee for valuable suggestions leading to the improvement of the paper.

References

1. G. C. Bento, O. P. Ferreira and P. R. Oliveira, Local convergence of the proximal point method for a special class of nonconvex functions on Hadamard manifolds, *Nonlinear Anal.*, **73** (2010), 564-572.
2. G. C. Bento, O. P. Ferreira and P. R. Oliveira, Proximal point method for a special class of nonconvex functions on Hadamard manifolds, *Optimization*, (2012), 1-31.
3. H. Brézis and P. L. Lions, Produits infinis de resolvants, *Israel J. Math.*, **29** (1970), 329-345.
4. F. E. Browder, Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces, *Trans. Amer. Math. Soc.*, **118** (1965), 338-351.

5. L. Ceng, Y. Liou and E. Naraghirad, *Iterative approaches to find zeros of maximal monotone operators by hybrid approximate proximal point methods*, Fixed Point Theory and Applications, 2011, Article ID 282171, 18 pages.

6. J. X. Da Cruz Neto, O. P. Ferreira and L. R. Lucambo Pérez, Monotone point-to-set vector fields, *Bulk. J. Geom. Appl.*, **5** (2000), 69-79.

7. J. X. Da Cruz Neto, O. P. Ferreira and L. R. Lucambo Pérez, Contribution to the study of monotone vector fields, *Acta Math. Hung.*, **94(4)** (2002), 307-320.

8. J. X. Da Cruz Neto, O. P. Ferreira, L. R. Lucambo Pérez and S. Z. Németh, Convex and monotone transformable mathematical programming problems and a proximal-like point method, *J. Global Optim.*, **35(1)** (2006), 53-69.

9. O. P. Ferreira and P. R. Oliveira, Proximal point algorithm on Riemannian manifolds, *Optimization*, **51(2)** (2012), 257-270.

10. O. P. Ferreira, L. R. Lucambo Pérez and S. Z. Németh, Singularities of monotone vector fields and an extragradient-type algorithm, *J. Glob. Optim.*, **31** (2005), 133-151.

11. C. Li, G. López and V. Martín-Márquez, Monotone vector fields and proximal point algorithm on Hadamard manifolds, *J. London Math. Soc.*, **79(2)** (2009), 663-683.

12. C. Li, G. López, V. Martín-Márquez and J. Wang, Resolvents of set-valued monotone vector fileds in Hadamard manifolds, *Set-Valued Anal.*, **19** (2011), 361-383.

13. B. Djafari Rouhani and H. Khatibzadeh, On the proximal point algorithm, *J. Optim. Theory Appl.*, **137** (2008), 411-417.

14. O. P. Ferreira and P. R. Oliveira, Proximal point algorithm on Riemannian manifolds, *Optimization*, **51** (2002), 257-270.

15. T. Iwamiya and H. Okochi, Monotonicity, resolvents and Yoshida approximation on Hilbert manifolds, *Nonlinear Anal.*, **54** (2003), 205-214.

16. J. Jost, Riemannian Geometry and Geometric Analysis, *Springer*, 2002.

17. H. Khatibzadeh, Some remarks on the proximal point algorithm, *J. Optim. Theory Appl.*, **153** (2012), 769-778.

18. H. Khatibzadeh and S. Ranjbar, On the strong convergence of Halpern type proximal point algorithm, *J. Optim. Theory Appl.*, **158** (2013), 385-396.

19. B. Martinet, Régularisation d'inéquations variationelles par approximations successives, *Revue Fr. Inform. Rech. Oper.*, **4** (1970), 154-159.

20. G. J. Minty, On the monotonicity of the gradient of a convex function, *Pacific J. Math.*, **14** (1964), 243-247.

21. G. Morosanu, *Nonlinear Evolution Equations and Applications*, Mathematics and its Applications, Vol. 26, D. Reidel, Dordrecht, The Netherlands, 1988.

22. S. Z. Németh, Monotone vector fields, *Publ. Math. Debrecen*, **54** (1999), 437-449.

23. S. Z. Németh, Variational inequalities on Hadamard manifolds, *Nonlinear Anal.*, **52** (2003), 1491-1498.

24. X. Qin, S. Kang and Y. Cho, Approximate zeros of monotone operators by proximal point algorithms, *J. Glob. Optim.*, **46** (2010), 75-87.

25. R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.*, **14** (1976), 877-898.

26. T. Sakai, *Riemannian Geometry, Translations of Mathematical Monographs*, Vol. 149, Amenrican Mthematical Society, Providence, RI, 1996.

27. G. Tang, L. Zhou and N. Huang, The proximal point algorithm for pseudomonotone variational inequalities on Hadamard manifolds, *Optim. Lett.*, (2012).

28. F. Wang and H. Cui, On the contraction proximal point algorithms with multi-parameters, *J. Glob. Optim.*, **54** (2012), 485-491.

29. A. J. Zaslavski, Maximal monotone operators and proximal point algorithm in the presence of computational errors, *J. Optim. Theory Appl.*, **150** (2011), 20-32.

30. E. Zeidler, *Nonlinear Functional Analysis and Applications*, Part IIB: Nonlinear monotone operators, Springer, New York, 1990.

P. Ahmadi and H. Khatibzadeh
Department of Mathematics
University of Zanjan
P. O. Box 45195-313
Zanjan, Iran
E-mail: P.ahmadi@znu.ac.ir
hkhatibzadeh@znu.ac.ir