Filomat 31:14 (2017), 4571–4578
https://doi.org/10.2298/FIL1714571A

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Long Time Behavior of Quasi-convex and Pseudo-convex Gradient Systems on Riemannian Manifolds

P. Ahmadi[a], H. Khatibzadeh[a]

[a]*Department of Mathematics, University of Zanjan, P.O.Box 45371-38791, Zanjan-Iran*

Abstract. In this paper, we study the following gradient system on a complete Riemannian manifold M,

$$\begin{cases} -x'(t) = \mathrm{grad}\varphi(x(t)) \\ x(0) = x_0, \end{cases}$$

where $\varphi : M \to \mathbb{R}$ is a C^1 function with $\mathrm{Argmin}\varphi \neq \emptyset$. We prove that the gradient flow $x(t)$ converges to a critical point of φ if φ is pseudo-convex, or if φ is quasi-convex and M is Hadamard. As an application to minimization, we consider a discrete version of the system to approximate a minimum point of a given pseudo-convex function φ.

1. Introduction

A gradient system is a first order dynamical system of the form

$$\begin{cases} -x'(t) = \mathrm{grad}\varphi(x(t)), \\ x(0) = x_0, \end{cases} \tag{1}$$

where φ is a differentiable real-valued function on a Hilbert space. A trajectory of solution to (1) is called a gradient flow. A well-known result says that if φ is convex with $\mathrm{Argmin}\varphi \neq \emptyset$, then the gradient flow converges weakly to a minimum point of φ. This fact, which is valuable in optimization, was extended by Bruck [2] even for nonsmooth convex functions. In [4] Goudou and Munier studied the asymptotic behavior of (1), when φ is a continuously differentiable quasi-convex function on a Hilbert space H with $\mathrm{Argmin}\varphi \neq \emptyset$. They proved the weak convergence of the gradient flow to a critical point of φ, as well as the strong convergence with some additional conditions on φ. When φ is a pseudo-convex function, any critical point becomes a minimum point and so the gradient flow converges weakly to a minimum point of φ and therefore it solves the unconstrained minimization problem:

$$\min_{x \in H} \varphi(x). \tag{2}$$

2010 *Mathematics Subject Classification*. 34C40; 37C10.
Keywords. Gradiant system, Quasi-convex function, Convergence, Riemannian manifold.
Received: 24 April 2016; Accepted: 08 July 2016
Communicated by Mića Stanković
Corresponding author: H.Khatibzadeh
Email addresses: p.ahmadi@znu.ac.ir (P. Ahmadi), hkhatibzadeh@znu.ac.ir (H. Khatibzadeh)

Let M be a submanifold of a Hilbert space H. Consider the constrained minimization problem:

$$\operatorname*{Min}_{x\in M} \varphi(x). \tag{3}$$

In some cases φ is not quasi-convex on the whole space H, but it becomes quasi-convex (or even convex) on the constrained set M along geodesics. Therefore, as a dynamical approach for studying these kinds of constrained minimization problems, we may consider (1) when φ is defined on a Riemannian manifold M. Munier [7] proved the convergence of the gradient flow of (1) to a minimum point of a convex function φ which is defined on a Riemannian manifold M. The authors [1] considered the nonhomogeneous case of (1) on a Hadamard manifold to study the convergence of the solutions. In this paper, we consider (1) when φ is a quasi-convex function on a Hadamard manifold with Argmin$\varphi \neq \emptyset$. We also prove convergence of the gradient flow of a pseudo-convex function to a minimum point of φ on Riemannian manifolds. Our results extend the related results of [4] to Riemannian or Hadamard manifolds and the results of [1, 7] to quasi-convex or pseudo-convex functions.

2. Preliminaries of Riemannian Geometry

In this section, we recall some important background about Riemannian manifolds from [5] and [9] which is needed in the sequel.

Let M be a smooth manifold of dimension n. For $p \in M$, the tangent space at p is denoted by T_pM and the tangent bundle of M by $TM = \bigcup_{p\in M} T_pM$, which is naturally a manifold. We restrict ourselves to real manifolds. Since T_pM is a linear space and has the same dimension of M, the tangent space T_pM is isomorphic to $\mathbb{R}^n$. The manifold M is called a Riemannian manifold if it is endowed with a Riemannian metric g, and in this case, it is denoted by (M, g). In the tangent space T_pM, the inner product of two vectors v and w, is defined by $\langle v, w\rangle_p := g_p(v, w)$, where g_p is the metric at the point p, and the corresponding norm is defined by $\| v \|_p := \sqrt{\langle v, v\rangle_p}$. Whenever there is no confusion, we use the notation $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_p$ and $\|\cdot\| = \|\cdot\|_p$.

Let $[a, b]$ be a closed interval in $\mathbb{R}$ and $\gamma : [a, b] \to M$ 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

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

which induces the original topology on M.

Let ∇ be the Levi-Civita connection on M associated with the Riemannian metric, and γ be a smooth curve in M. A vector field X is said to be parallel along γ if $\nabla_{\dot{\gamma}} X = 0$. A smooth curve γ is a geodesic if $\dot{\gamma}$ itself is parallel along γ. A geodesic joining p to q in M is said to be minimal if its length equals $d(p, q)$.

A Riemannian manifold is complete if for each $p \in M$ all geodesics emanating from p are defined on whole $\mathbb{R}$. If M is complete then by Hopf-Rinow Theorem any pair of points of M can be joined by a minimal geodesic.

Let M be a complete Riemannian manifold. The exponential map $\exp_p : T_pM \to M$ at p is defined by $\exp_p(v) = \gamma_v(1)$ for each $v \in T_pM$, where $\gamma_v(0)$ is the geodesic with $\gamma_v(0) = p$ and $\dot{\gamma}_v(0) = v$. Then $\exp_p(tv) = \gamma_v(t)$, for each real number t.

There is a special type of Riemannian manifolds on which the study of gradient systems yields interesting results. A Riemannian manifold M is said to be a Hadamard manifold if it is complete, simply connected and of non-positive sectional curvature. The following result which is a part of Hadamard-Cartan Theorem from [9, p. 221], shows that any n-dimensional Hadamard manifold has the same topology and differential structure as the Euclidean space $\mathbb{R}^n$.

Theorem 2.1. *Let M be an Hadamard manifold and $x \in M$. Then $\exp_x : T_xM \to M$ is a diffeomorphism, and for any two points $x, y \in M$ there exists a unique normalized geodesic joining x to y, which is in fact, a minimal geodesic (i.e., distance realizing).*

Hadamard manifolds and Euclidean spaces have some similar geometrical properties. One of them is described in the following proposition. By definition, a geodesic triangle $\Delta(p_1p_2p_3)$ in 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 2.2. ([9, 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*

$$\alpha_1 + \alpha_2 + \alpha_3 \leqslant \pi ,$$

$$l_i^2 + l_{i+1}^2 - 2l_il_{i+1}\cos\alpha_{i+1} \leqslant l_{i-1}^2 . \tag{4}$$

Since

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

so the inequality (4) may be rewritten as follows

$$d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2\langle \exp^{-1}_{p_{i+1}} p_i, \exp^{-1}_{p_{i+1}} p_{i+2}\rangle \leqslant d^2(p_{i+2}, p_i). \tag{5}$$

Now we recall three kinds of convexity concepts which we use in the paper; quasi, pseudo and θ-weak convexity. A differentiable function $\varphi : M \to \mathbb{R}$ is said to be a quasi-convex function if it is quasi-convex when restricted to any geodesic $\gamma : [a, b] \subset \mathbb{R} \to M$, which means that

$$\varphi \circ \gamma(ta + (1 - t)b) \leqslant \text{Max}\{\varphi(\gamma(a)), \varphi(\gamma(b))\} \tag{6}$$

holds for any $a, b \in \mathbb{R}$ and $0 \leqslant t \leqslant 1$. Let φ be a quasi-convex function, x and y be two distinct points in M, and without loss of generality suppose that $\text{Max}\{\varphi(x), \varphi(y)\} = \varphi(x)$. Let $\gamma : [0, 1] \to M$ be a minimal geodesic connecting x to y. Then

$$\varphi(\gamma(t)) \leqslant \varphi(\gamma(0)), \quad \forall t \in [0, 1],$$

which shows that

$$\frac{\varphi(\gamma(t)) - \varphi(\gamma(0))}{t} \leqslant 0, \quad \forall t \in (0, 1].$$

By taking limit in the both sides when $t \to 0^+$, we get

$$\langle \text{grad}\varphi(x), \dot\gamma(0)\rangle \leqslant 0, \tag{7}$$

where gradφ is the vector field metrically equivalent to the differential $d\varphi$, i.e.,

$$\langle \text{grad}\varphi, X\rangle = d\varphi(X) = X\varphi,$$

where X is also a vector field. If M is a Hadamard manifold then the inequality (7) becomes

$$\langle \text{grad}\varphi(x), \exp^{-1}_x y\rangle \leqslant 0. \tag{8}$$

The function φ is called pseudo-convex if the inequality (7) holds strictly. Clearly, any pseudo-convex function is quasi-convex. For pseudo-convex functions any critical point is a minimum point.

The function φ is called θ-weakly convex for $\theta > 0$ iff for each $x, y \in M$ and any geodesic segment $\gamma : [0, d(x, y)] \to M$ with $\gamma(0) = x$ and $\gamma(d(x, y)) = y$ and each $t \in]0, d(x, y)[$

$$\varphi \circ \gamma(t) \leq \frac{t}{d(x, y)}\varphi \circ \gamma(0) + (1 - \frac{t}{d(x, y)})\varphi \circ \gamma(d(x, y)) + \theta t(d(x, y) - t).$$

If φ is also differentiable, by a similar computation as above, we derive

$$\langle \text{grad}\varphi(x), \exp^{-1}_x y\rangle \leq \varphi(y) - \varphi(x) + \theta d^2(x, y).$$

3. Convergence Analysis

Throughout this section, it is assumed that $\varphi : M \to]-\infty, +\infty]$ is a C^1 quasi-convex function, $\varphi \not\equiv +\infty$ and M is a Hadamard manifold. First we recall the notion of Fejér convergence and the following related result from [3].

Definition 3.1. *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*

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

Lemma 3.2. *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.*

Let Argminφ denote the following set

$$\text{Argmin}\varphi := \{x \in M \mid \varphi(x) \leqslant \varphi(y) \quad, \forall y \in M\}.$$

Lemma 3.3. *Let M be a Hadamard manifold and $x : \mathbb{R} \to M$ be a solution to* (1). *If* Argmin$\varphi \neq \emptyset$, *then $d(x(t), p)$ is a nonincreasing function, for each $p \in$* Argminφ.

Proof. By (1) and (8), we have $\langle -x'(t), \exp^{-1}_{x(t)} p\rangle \leqslant 0$, and so

$$\begin{aligned}
\lim_{h\to 0^+} \frac{1}{h}\langle \exp^{-1}_{x(t)} x(t-h), \exp^{-1}_{x(t)} p\rangle &= \langle -\frac{d}{ds}\exp^{-1}_{x(t)} x(s)|_{s=t}, \exp^{-1}_{x(t)} p\rangle \\
&= \langle -d\exp^{-1}_{x(t)}(x(t))\frac{d}{ds}x(s)|_{s=t}, \exp^{-1}_{x(t)} p\rangle \\
&= \langle -x'(t), \exp^{-1}_{x(t)} p\rangle \\
&\leqslant 0.
\end{aligned}$$

Then, by using the inequality (5) of the comparison theorem for the geodesic triangle $\triangle(x(t)x(t-h)p)$, one gets that

$$\lim_{h\to 0^+} \frac{1}{h}(d^2(x(t), x(t-h)) + d^2(x(t), p) - d^2(x(t-h), p)) \leqslant 0,$$

which implies that

$$\frac{d}{dt}d(x(t), p) = \lim_{h\to 0^+}\frac{1}{h}(d(x(t), p) - d(x(t-h), p)) \leqslant 0.$$

□

Lemma 3.4. *Let M be a complete Reimannian manifold and $\varphi : M \to]-\infty, +\infty]$ be a C^1 quasi-convex function. Let $x : \mathbb{R} \to M$ satisfy* (1). *Then $\varphi(x(\cdot))$ is a nonincreasing function.*

Proof. By the definition of gradφ at $x(t)$, we have

$$\frac{d}{dt}\varphi(x(t)) = \langle \text{grad}\varphi(x(t)), x'(t)\rangle = -\|x'(t)\|^2 \leqslant 0.$$

Hence $\varphi(x(\cdot))$ is a nonincreasing function. □

Theorem 3.5. *Let M be a Hadamard manifold and $\varphi : M \to]-\infty, +\infty]$ be a C^1 quasi-convex function. Let* Argmin$\varphi \neq \emptyset$ *and $x : \mathbb{R} \to M$ satisfy (1). Then $\lim_{t\to+\infty} x(t) = p$ and $\lim_{t\to+\infty} \varphi(x(t)) = \varphi(p)$, where p is a critical point of φ.*

Proof. First we claim that $x(t)$ converges to some point $p \in M$ as $t \to +\infty$. For any positive fixed real number t and any $s \in [0, t]$, we have

$$\varphi(x(t)) \leqslant \varphi(x(s)),$$

by Lemma 3.4. Quasi-convexity of φ and (1) imply that

$$\langle -x'(s), \exp^{-1}_{x(s)} x(t)\rangle = \langle \mathrm{grad}\varphi(x(s)), \exp^{-1}_{x(s)} x(t)\rangle \leqslant 0.$$

Hence

$$\lim_{h\to 0^+} \frac{1}{h}\langle \exp^{-1}_{x(s)} x(s-h), \exp^{-1}_{x(s)} x(t)\rangle \leqslant 0.$$

This together with the inequality (5) of the comparison theorem for triangles in the geodesic triangle $\triangle(x(s)x(s-h)x(t))$ show that

$$\lim_{h\to 0^+} \frac{1}{h}(d^2(x(s), x(t)) - d^2(x(s-h), x(t))) \leqslant 0.$$

So

$$\frac{d}{ds} d^2(x(s), x(t)) \leqslant 0.$$

Thus the function $d^2(x(\cdot), x(t))$ is nonincreasing on $[0, t]$. Hence for every $s_1, s_2 \in [0, t]$, where $s_1 \leqslant s_2$, we have

$$d^2(x(s_2), x(t)) \leqslant d^2(x(s_1), x(t)).$$

Let K be the set of all cluster points of $\{x(t) | \ t \in \mathbb{R}^+\}$, that is nonempty by Lemma 3.3. Suppose that $q \in K$. Then there exists an increasing sequence $\{t_n\}$ of positive real numbers such that $\lim_{n\to+\infty} x(t_n) = q$. Hence for any t_n and any $s_1, s_2 \in [0, t_n]$, where $s_1 \leqslant s_2$, we have

$$d^2(x(s_2), x(t_n)) \leqslant d^2(x(s_1), x(t_n)).$$

Taking limit from both sides of the above inequality, when $n \to +\infty$, we get

$$d^2(x(s_2), q) \leqslant d^2(x(s_1), q).$$

Thus $\{x(\cdot)\}$ converges to some point $p \in K$ by Lemma 3.2, which proves our claim.

Now we show that p is a critical point of φ.

$$\begin{aligned}
\int_0^{+\infty} \|x'(t)\|^2 dt &= \int_0^{+\infty} \langle -\mathrm{grad}\varphi(x(t)), x'(t)\rangle dt \\
&= -\int_0^{+\infty} \frac{d}{dt}\varphi(x(t)) dt \\
&= \varphi(x(0)) - \lim_{t\to+\infty} \varphi(x(t)) \\
&< +\infty.
\end{aligned}$$

This shows that $\liminf_{t\to+\infty} \|x'(t)\| = 0$. Hence, by (1),

$$\lim_{n\to+\infty} \|x'(t_n)\| = \lim_{n\to+\infty} \|\mathrm{grad}\varphi(x(t_n))\| = 0,$$

for some increasing sequence $\{t_n\}$ of positive real numbers. Since φ is C^1, we get

$$\mathrm{grad}\varphi(p) = \lim_{t\to+\infty} \mathrm{grad}\varphi(x(t_n)) = 0.$$

Thus p is a critical point of φ. □

In the following theorem, we show that the conclusion of Theorem 3.5 remains true when φ is a pseudo-convex function even on a complete Riemannian manifold (not necessarily with nonpositive sectional curvature). Although the proof of the following theorem is similar to that of Proposition 1 of [7], we facilitate the reader with the following proof.

Theorem 3.6. *Let M be a complete Riemannian manifold and $\varphi : M \to]-\infty, +\infty]$ be a C^1 pseudo-convex function. Let* $\mathrm{Argmin}\varphi \neq \emptyset$ *and $x : \mathbb{R} \to M$ satisfy* (1). *Then $\lim_{t\to+\infty} x(t) = p$ and $\lim_{t\to+\infty} \varphi(x(t)) = \varphi(p)$, where p is a critical point of φ.*

Proof. Let p be an arbitrary fixed point in $\mathrm{Argmin}\varphi$. First we show that the function $t \mapsto d(x(t), p)$ decreases. For every $t > 0$ there is some vector $u(t) \in T_{x(t)}M$ such that

$$\exp_{x(t)}(u(t)) = p$$
$$d(x(t), p) = \|u(t)\|_{x(t)}.$$

Consider the geodesic $\gamma(s) = \exp_{x(t)}(su(t))$. We have $\gamma(0) = x(t)$, $\dot{\gamma}(0) = u(t)$ and $\gamma(1) = p$. Since $\varphi \circ \gamma$ is pseudo-convex, we get:

$$\langle \mathrm{grad}\varphi(x(t)), u(t)\rangle_{x(t)} < 0.$$

As both paths $h \mapsto x(t + h)$ and $h \mapsto \exp_{x(t)}(-h\ \mathrm{grad}\varphi(x(t)))$ are C^1, and have the same initial condition of orders 0 and 1, we have

$$d(x(t + h), \exp_{x(t)}(-h\ \mathrm{grad}\varphi(x(t)))) = o(h). \tag{9}$$

An argument of the same type gives

$$d(\exp_{x(t)}(-h\ \mathrm{grad}\varphi(x(t))), \exp_{x(t)}(h\lambda u(t))) = \| - h\ \mathrm{grad}\varphi(x(t)) - h\lambda u(t)\|_{x(t)} + o(h), \tag{10}$$

where λ is an arbitrary positive real number. Let $\lambda = \frac{\|\mathrm{grad}\varphi(x(t))\|^2_{x(t)}}{-\langle u(t), \mathrm{grad}\varphi(x(t))\rangle_{x(t)}}$. Then $\lambda > 0$, and $\langle -\mathrm{grad}\varphi(x(t)) - \lambda u(t), -\mathrm{grad}\varphi(x(t))\rangle_{x(t)} = 0$. Hence

$$\| - h\ \mathrm{grad}\varphi(x(t)) - h\lambda u(t)\|_{x(t)} = h(\lambda\|u(t)\|^2_{x(t)} - \|\mathrm{grad}\varphi(x(t))\|^2)^{\frac{1}{2}}.$$

Finally,

$$d(\exp_{x(t)}(h\lambda u(t)), p) = (1 - h\lambda)\|u(t)\|_{x(t)}. \tag{11}$$

Now construct a broken minimizing geodesic α joining $x(t+h)$ to $\exp_{x(t)}(-h\ \mathrm{grad}\varphi(x(t)))$, then to $\exp_{x(t)}(h\lambda u(t))$ and then to p. Therefore by combining (9), (10) and (11), we have

$$L(\alpha) = \|u(t)\|_{x(t)} - h[\lambda\|u(t)\|_{x(t)} - ((\lambda\|u(t)\|_{x(t)})^2 - \|\mathrm{grad}\varphi(x(t))\|^2)^{\frac{1}{2}}] + o(h).$$

Since $(\lambda\|u(t)\|_{x(t)})^2 > (\lambda\|u(t)\|_{x(t)})^2 - \|\mathrm{grad}\varphi(x(t))\|^2 \geqslant 0$, so the bracket just above is positive. Thus, for small enough h, we have

$$d(x(t + h), p) \leqslant L(\alpha) \leqslant \|u(t)\|_{x(t)} = d(x(t), p),$$

which shows that the function $t \mapsto d(x(t), p)$ decreases. This implies that the set $\{x(t) |\ t \in \mathbb{R}^+\}$ is bounded in M. Hence by Hopf-Rinow Theorem [5, p.26] there exists some real sequence $\{t_k\}$ such that $x(t_k) \to p$, when $k \to +\infty$. Then $d(x(t), p) \to 0$, when $t \to +\infty$. □

Remark 3.7. *We don't know, whether Theorem 3.6 is true for quasi-convex functions or not. It may be the subject of future researches.*

A well-known result says that any quasi-convex function on a compact Riemannian manifold should be constant, and so Theorem 3.6 is satisfied for quasi-convex functions on compact Riemannian manifolds. Here we give a simple example on a non-compact Riemannian manifold for Theorem 3.5.

Example 3.8. *Let $H = \{(x, y) \in \mathbb{R}^2 | \ y > 0\}$ be the Poincare half plan, which is a Hadamard manifold with constant sectional curvature -1. The function $\varphi : H \to \mathbb{R}$, $\varphi(x, y) = x^2$ is a quasi-convex function, since its sublevel sets are geodesically convex (it is not convex on H). Clearly, $\mathrm{Argmin}\varphi = \{(0, y)| \ y > 0\}$. Consider the natural coordinate system on H. Then $\mathrm{grad}\varphi(x, y) = 2x\frac{\partial}{\partial x}$, and $x(t) = (e^{-2t}, c)$ is a solution to the system (1), where c is a positive constant. Hence $\lim_{t\to+\infty} x(t) = (0, c) \in \mathrm{argmin}\varphi$, as is predicted by Theorem 3.5.*

4. Application to Pseudo-convex Minimization

Consider the following constrained minimization problem

$$\mathrm{Min}_{x\in M}\varphi(x), \tag{12}$$

where the constraint set M is a Riemannian submanifold of $\mathbb{R}^n$. Even when φ is not pseudo-convex on $\mathbb{R}^n$ it may be pseudo-convex on M along geodesics. Therefore the non-pseudoconvex and constrained problem (12) on $\mathbb{R}^n$ can be considered as a pseudo-convex and non-constrained one on M. By the results of the previous section specially Theorem 3.5, and using the fact that any critical point of a pseudo-convex function is a minimum point, the trajectory of (1) converges to a minimum point of φ. This gives us a dynamical approach to pseudo-convex minimization problem (12). Since continuous trajectories are not defined for computer softwares, it is appropriate to consider discretization of (1) in order to approximate a minimum point of φ. There are two ways for discretization of (1), backward and forward Euler discretizations. Backward Euler discretization has been considered by Quiroz, Quispe and Oliveira [8]. Here we consider forward Euler discretization of (1) and prove the existence of the generated sequence as well as its convergence to a minimum point of φ, with some suitable assumptions on φ such as quasi-convexity (more general than pseudo-convexity) and weak convexity on φ. Forward discretization of (1) is in the form

$$\lambda_k \exp^{-1}_{x_k} x_{k-1} = \mathrm{grad}\varphi(x_k), \tag{13}$$

where λ_k is the step-size. First we show for a θ-weakly convex function φ and suitable parameters λ_k the sequence x_k in (13) exists.

Proposition 4.1. *Suppose M is a Hadamard manifold and $\varphi : M \to \mathbb{R}$ is a θ-weakly convex differentiable function. Then for each $k \geq 1$ and a given $x_{k-1} \in M$ and $\lambda_k \geq \lambda > 2\theta$, there exists x_k satisfying* (13).

Proof. For a given $x_{k-1} \in M$ and $\lambda_k > \lambda > \theta$, define

$$A_k(x) = \mathrm{grad}\varphi(x) - \lambda_k exp^{-1}_x x_{k-1}.$$

First we prove A_k is strongly monotone (see Definition 3.1 of [6]).

$$\begin{aligned}
&\langle A_k x, \exp^{-1}_x y\rangle + \langle A_k y, \exp^{-1}_y x\rangle \\
&= \langle \mathrm{grad}\varphi(x), \exp^{-1}_x y\rangle - \lambda_k\langle \exp^{-1}_x x_k, \exp^{-1}_x y\rangle \\
&+\langle \mathrm{grad}\varphi(y), \exp^{-1}_y x\rangle - \lambda_k\langle \exp^{-1}_y x_k, \exp^{-1}_y x\rangle \\
&\leq \varphi(y) - \varphi(x) + \theta d^2(x, y) + \varphi(x) - \varphi(y) + \theta d^2(x, y) \\
&+\frac{\lambda_k}{2}\{d^2(x_k, y) - d^2(x_k, x) - d^2(x, y) + d^2(x_k, x) - d^2(x_k, y) - d^2(x, y)\} \\
&= 2\theta d^2(x, y) - \lambda_k d^2(x, y) \leq -(\lambda - 2\theta)d^2(x, y).
\end{aligned}$$

Therefore A_k is strongly monotone vector field. By Theorem 4.3 of [6], there exists x_k such that $A_k(x_k) = 0$. Equivalently

$$\lambda_k \exp^{-1}_{x_k} x_{k-1} = \mathrm{grad}\varphi(x_k).$$

□

Now we verify convergence of the discrete trajectory generated by (13).

Theorem 4.2. *Suppose that φ is quasi-convex and θ-weakly convex. If $\lambda_k \geq \lambda > 2\theta$ is bounded from above, then x_k given by* (13) *converges to a critical point of φ.*

Proof. Suppose $\varphi(x_k) > \varphi(x_{k-1})$, then by quasi-convexity of φ

$$\langle \text{grad}\varphi(x_k), \exp^{-1}_{x_k} x_{k-1}\rangle \leq 0,$$

which is a contradiction by (13). Therefore $\varphi(x_k)$ is nonincreasing. Let $n \leq m$. We have $\varphi(x_m) \leq \varphi(x_n)$ again by quasi-convexity of φ

$$\langle \text{grad}\varphi(x_n), \exp^{-1}_{x_n} x_m\rangle \leq 0.$$

By (13), we have

$$\langle \exp^{-1}_{x_n} x_{n-1}, \exp^{-1}_{x_n} x_m\rangle \leq 0.$$

By (5), we get

$$d^2(x_{n-1}, x_n) + d^2(x_m, x_n) - d^2(x_{n-1}, x_m) \leq 0.$$

It implies that

$$d^2(x_m, x_n) \leq d^2(x_{n-1}, x_m).$$

Therefore $d^2(x_n, x_m)$ is nonincreasing for $0 \leq n \leq m$. For $0 \leq n \leq k \leq m$ we have

$$d^2(x_k, x_m) \leq d^2(x_n, x_m).$$

Let K be the set of all cluster points of $\{x_n\}$. Suppose that $x_{n_l} \to q \in K$, then for each $k, n \leq n_l$

$$d^2(x_k, x_{n_k}) \leq d^2(x_n, x_{n_k}).$$

Letting $l \to +\infty$ we get

$$d^2(x_k, q) \leq d^2(x_n, q).$$

Therefore $x_n \to p$ as $n \to +\infty$. Since λ_n is bounded, $\lambda_n d^2(x_n, x_{n-1}) \to 0$. Hence $\text{grad}\varphi(x_k) \to 0$. Since $\text{grad}\varphi$ is continuous, $\text{grad}\varphi(p) = 0$. □

Corollary 4.3. *In Theorem 4.2, if φ is pseudo-convex and $Argmin\varphi \neq \emptyset$, then x_k converges to a minimum point of φ.*

References

[1] P. Ahmadi and H. Khatibzadeh, Convergence and rate of convergence of a non-autonomous gradient system on Hadamard manifolds, *Lob. J. Math.* Vol. 35, **3** (2014), 165-171.

[2] R. E. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, *J. Func. Anal.* **18** (1975), 15-26.

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

[4] X. Goudou and J. Munier, The gradient and heavey ball with friction dynamical systems: the quasiconvex case, *Math. Program., Ser. B*, **116** (2009), 173-191.

[5] J. Jost, Riemannian Geometry and Geometric Analysis, *Springer* 2002.

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

[7] J. Munier, Steepest descent method on a Riemannian manifold: the convex case, *Balkan J. Geom. Appl.* **12**, no. 2 (2007), 98106.

[8] E. A. Papa Quiroz, E. M. Quispe and P. R. Oliveira, Steepest descent method with a generalized Armijo search for quasiconvex functions on Riemannian manifolds, J. Math. Anal. Appl. 341 (2008), 467-477.

[9] T. Sakai, Riemannian Geometry, Translations of Mathematical Monographs 149 (American Mathematical Society, Providence, RI, 1996).

[10] C. Udriste, Convex Functions and Optimization Methods on Riemannian Manifolds, *Mathematics and its Applications*, **297** (*Kluwer, Dordrecht*, 1994).