PROXIMAL POINT ALGORITHM ON RIEMANNIAN MANIFOLDS

O. P. FERREIRA\textsuperscript{a,*} and P. R. OLIVEIRA\textsuperscript{b,†}

\textsuperscript{a}IME/Universidade Federal de Goiás, Campus Samambaia- Caixa Postal 131, CEP 74001-970 - Goiânia, GO, Brazil; \textsuperscript{b}COPPE-Sistemas, Universidade Federal, do Rio de Janeiro - Caixa Postal 68511, CEP 21945-970 - Rio de Janeiro, RJ, Brazil

(Received 10 August 1999; In final form 22 August 2000)

In this paper we consider the minimization problem with constraints. We will show that if the set of constraints is a Riemannian manifold of nonpositive sectional curvature, and the objective function is convex in this manifold, then the proximal point method in Euclidean space is naturally extended to solve that class of problems. We will prove that the sequence generated by our method is well defined and converge to a minimizer point. In particular we show how tools of Riemannian geometry, more specifically the convex analysis in Riemannian manifolds, can be used to solve nonconvex constrained problem in Euclidean space.

Keywords: Proximal point; Nonconvex problem; Riemannian manifold

Mathematics Subject Classifications 1991: 49M30, 90C26

1. INTRODUCTION

Extension of the concepts and techniques of Mathematical Programming of the Euclidean space $\mathbb{R}^n$ to Riemannian manifolds is natural. It has been done frequently in recent years, with theoretical objective and also to obtain effective algorithms; see [3–5, 11, 12, 16 and 17]. In particular, we observe that, these extensions make possible to solve

\textsuperscript{*}Corresponding author. e-mail: orizon@mat.ufg.br
\textsuperscript{†}Research of this author was partially supported by CNPq, Brazil. e-mail: poliveir@cos.ufrj.br

ISSN: 0233-1934. Online ISSN: 1029-4945. © 2002 Taylor & Francis Ltd
DOI: 10.1080/02331930290019413
some nonconvex constrained problems in Euclidean space. The proximal point algorithm, introduced by Martinet [10] and Rockafellar [13], has been extended to different contexts, see [6] and [8]. To extend this algorithm to Riemannian manifold is the subject of this paper.

Let $M$ be a complete Riemannian manifold and let $f: M \rightarrow \mathbb{R}$ a convex function. We will consider the optimization problem

$$\min_{x \in M} f(x).$$

(1)

References about convex functions in Riemannian manifold are [12, 15 and 17]. The proximal point algorithm in Riemannian manifold generates, for a starting point $p_0 \in M$, a sequence $\{x_k\} \subset M$ by the iteration

$$x_{k+1} = \operatorname{argmin}_{y \in M} \{f(y) + \lambda_k \rho_x(y)\}$$

(2)

with $\rho_x: M \rightarrow \mathbb{R}$ defined by $\rho_x(y) = (1/2)d^2(x, y)$, where $d$ is the Riemannian distance (to be defined later on), and $\{\lambda_k\}$ is a sequence of positive numbers. Our point of departure will be to show that this extension is natural. The main results are the proof of well definedness of the sequence generated by (2) and the proof of convergence these sequence to a minimizer of $f$, when $M$ is a Riemannian manifold with non positive sectional curvature. We decide for inclusion of some known results, that are necessary for understanding of our work, for the read’s comfort, due to these results be diluted in several references. We include also some proofs. In such case, it is due because we believe that we have simplified these proofs.

2. BASICS CONCEPTS

In this section, we introduce some fundamental properties and notations of Riemannian manifold. These basics facts can be found in any introductory book of Riemannian geometry, for example in [1, 2 and 14].

Let $M$ be a connected manifold. We denote by $T_xM$ the tangent space of $M$ at $x$, by $TM = \bigcup_{x \in M} T_xM$ the tangent bundle of $M$. Let $M$ be endowed with a Riemannian metric $\langle \cdot, \cdot \rangle$, with corresponding norm denoted by $\| \|$, so that $M$ is now a Riemannian manifold. Recall that
PROXIMAL POINT ALGORITHM

the metric can be used to define the length of piecewise smooth curve $c: [a, b] \rightarrow M$ joining $x'$ to $x$, i.e., such that $c(a) = x'$ and $c(b) = x$, by $L(c) = \int_a^b \|c'(t)\| dt$. Minimizing this length functional over the set of all such curves we obtain a Riemannian distance $d(x', x)$ which induces the original topology on $M$. Let $\nabla$ be the Levi-Civita connection associated to $(M, \langle \cdot, \cdot \rangle)$. A vector field $V$ along $c$ is said to be parallel if $\nabla_c V = 0$. If $c'$ itself is parallel we say that $c$ is a geodesic. The geodesic equation $\nabla_{c'} \gamma' = 0$ is a second order nonlinear ordinary differential equation, then $\gamma = \gamma_a(\cdot, x)$ is determined by its position $x$ and velocity $v$ at $x$. It is easy to check that $\|\gamma'\|$ is constant. We say that $\gamma$ is normalized if $\|\gamma'\| = 1$. The restriction of a geodesic to a closed bounded interval is called a geodesic segment. A geodesic segment joining $x'$ to $x$ in $M$ is said to be minimal if its length equals $d(x', x)$ and this geodesic is called a minimizing geodesic.

A Riemannian manifold is complete if geodesics are defined for any values of $t$. Hopf-Rinow's theorem asserts that if this is the case then any pair of points, say $x'$ and $x$, in $M$ can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, $(M, d)$ is a complete metric space and bounded and closed subsets are compact. In this paper, all manifolds are assumed to be complete. Take $x \in M$, the exponential map $\exp_x: T_x M \rightarrow M$ is defined by $\exp_x v = \gamma_v(I, x)$.

We denote by $R$ the curvature tensor defined by $R(X, Y) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$, where $X$, $Y$ and $Z$ are vector fields of $M$ and $[X, Y] = YX - XY$. Then the sectional curvature with respect to $X$ and $Y$ is given by $K(X, Y) = \langle R(X, Y) Y, X \rangle / (\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2)$, where $\|X\|^2 = \langle X, X \rangle$. If $K(X, Y) \leq 0$ for all $X$ and $Y$, then $M$ is called a Riemannian manifold of nonpositive curvature and we use the short notation $K \leq 0$. Some interesting results are obtained when the sign of curvature is constant.

**Theorem 2.1** Let $M$ be a complete, simply connected Riemannian manifold with nonpositive sectional curvature. Then $M$ is diffeomorphic to the Euclidean space $\mathbb{R}^n$, $n = \dim M$. More precisely, at any point $x \in M$, the exponential mapping $\exp_x: T_x M \rightarrow M$ is a diffeomorphism.

**Proof** See [2 and 14].

A complete simply connected Riemannian manifold of nonpositive sectional curvature is called a Hadamard manifold. The Theorem 2.1
says that if $M$ is Hadamard manifold, then $M$ have the same topology and differential structure of the Euclidean space $R^n$. Furthermore, are known some similar geometrical properties of the Euclidean space $R^n$.

Now we proceed starting with those which we will go to use.

A geodesic triangle $\Delta(x_1x_2x_3)$ of a Riemannian manifold is the set consisting of three distinct points $x_1, x_2, x_3$ called the vertices and three minimizing geodesic segments $\gamma_{i+1}$ joining $x_{i+1}$ to $x_{i+2}$ called the sides, where $i=1, 2, 3(\text{mod } 3)$.

**Theorem 2.2** Let $M$ be a Hadamard manifold and $\Delta(x_1x_2x_3)$ a geodesic triangle. Denote by $\gamma_{i+1}: [0, l_{i+1}] \to M$ geodesic segments joining $x_{i+1}$ to $x_{i+2}$ and set $l_{i+1} = L(\gamma_{i+1}), \theta_{i+1} = \gamma_{i+1}'(0), -\gamma_{i}'(l_i)$, where $i=1, 2, 3(\text{mod } 3)$. Then

$$\theta_1 + \theta_2 + \theta_3 \leq \pi$$

$$l_{i+1}^2 + l_{i+2}^2 - 2l_{i+1}l_{i+2} \cos \theta_{i+2} \leq l_i^2$$

$$l_{i+1} \cos \theta_{i+2} + l_i \cos \theta_i \geq l_{i+2}. \tag{5}$$

**Proof** Inequality (3) and (4) are proved in [14]. Inequality (5) is an immediate consequence of (4).

Let $\gamma: [a, b] \to M$ be a normalized geodesic segment. A differentiable variation of $\gamma$ is by definition a differentiable mapping $\alpha: [a, b] \times (-\varepsilon, \varepsilon) \to M$ satisfying $\alpha(t, 0) = \gamma(t)$. The vector field along of $\gamma$ defined by $V(t) = (\partial\alpha/\partial s)(t, 0)$ is called the variational vector field of $\alpha$. The first variational formula of arc length on $\alpha$ is then given as follows:

$$L'(\gamma) := \frac{d}{ds} L(c_s)|_{s=0} = \langle V, \gamma' \rangle|_a^b,$$ \hspace{1cm} \tag{6}$$

where $c_s(t) = \alpha(t, s)$ with $s \in (-\varepsilon, \varepsilon)$.

3. **Convex Analysis**

In this section, we introduce some fundamental properties and notations of convex analysis in Hadamard manifolds which are used later. The option by inclusion of proofs is because we believe that we have simplified them, using definitions and ordinary arguments of the
convex analysis in Euclidean space $\mathbb{R}^n$. Reference of the convex analysis, in Euclidean space $\mathbb{R}^n$ is [7] and in Riemannian manifold are [1, 12, 14 – 17].

### 3.1. Convex Set

Let be $C$ a nonempty subset of Hadamard manifold $M$. The subset $C$ of $M$ is said to be convex if for arbitrary points $x'$ and $x$ in $C$, the geodesic segment $\gamma$ joining $x'$ and $x$ is contained in $C$, that is, if $x', x \in C$ and $\gamma$ is the geodesic such that $x' = \gamma(a)$ and $x = \gamma(b)$, then $\gamma((1-t)a + tb) \in C$, for all $t \in [0, 1]$.

Now, in what follows, suppose that $C$ is a closed convex subset of $M$. Then for fixed $x' \in M$, there exists a closest point in $C$ to $x'$. In fact, for fixed $x' \in M$, fix $x_0 \in C$ and define the set $A_{x_0} = \{x \in M \mid d(x', x) \leq d(x', x_0)\}$. Since the map $x \mapsto d(x', x)$ is continuous and the set $C \cap A_{x_0}$ is compact, there exists a $y_{x'} \in C$ such that $d(x', y_{x'}) \leq d(x', x)$, for all $x \in C$. We proceed now to prove the uniqueness of the closest point in $C$ to $x'$. The next result helps us.

**Proposition 3.1** Let $x'$ in $M$. If $y_{x'}$ in $C$ is such that $d(x', y_{x'}) \leq d(x', x)$, then

$$\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} x \rangle \leq 0,$$

for all $x \in C$.

**Proof** If $x' \in C$ the result is trivial. Suppose that $x' \notin C$. Set $l = d(x', y_{x'})$ and let $\gamma: [0, l] \to M$ be the geodesic such that $\gamma(0) = x'$ and $\gamma(l) = y_{x'}$. Note that $\gamma'(l) = -\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} x' \rangle$. Now suppose, on the contrary, that there exists a point $\bar{x} \in C$ such that $\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} \bar{x} \rangle > 0$. Consider the variation $\alpha: [0, l] \times (-\varepsilon, \varepsilon) \to M$ of the geodesic $\gamma$, defined by $\alpha(t, s) = \exp_{x'} t(\exp_{y_{x'}}^{-1} \beta(s))$, where $\beta: (-\varepsilon, l+\varepsilon) \to M$ is the geodesic such that $\beta(0) = y_{x'}$ and $\beta(l) = \bar{x}$. This way the variational vector field $V(t) = (\partial \alpha/\partial s)(t, 0)$ of $\alpha$ satisfy $V(0) = 0$ and $V(l) = \beta'(0) = \exp_{y_{x'}}^{-1} \bar{x}$. Then from (6) it follows that

$$L'(\gamma) = \frac{d}{ds} L(c_s)|_{s=0} = \langle -\exp_{y_{x'}}^{-1} x', \|\exp_{y_{x'}}^{-1} x'\|, \exp_{y_{x'}}^{-1} \bar{x} \rangle < 0,$$

where $c_s(\cdot) = \alpha(\cdot, s)$ and $s \in (-\varepsilon, l\varepsilon)$. From last inequality there exists $\delta > 0$ such that $L(c_\delta) < L(\gamma)$ for any $0 < s < \delta$, then $d(x', \beta(s)) \leq$
$d(x', y')$ for any $0 < s < \delta$. Since $C$ is convex, it follows that $\beta(s) \in C$ for any $0 < s < \delta$. Then the inequality $d(x', \beta(s)) \leq d(x', y')$, for any $0 < s < \delta$, contradicts the definition of $y'$, and our proof is complete.

**Proposition 3.2** Let $x' \in M$. There exists only one $y_x$ in $C$ such that $d(x', y_x) = d(x, x)$ for all $x \in C$.

**Proof** If $x' \in C$ the result is trivial. Suppose that $x' \notin C$. Assume, on the contrary, that there exists $y_x$ and $\tilde{y}_x$ in $C$, with $y_x \neq \tilde{y}_x$ and such that $d(x', y_x) = d(x', \tilde{y}_x) \leq d(x', x)$ for all $x \in M$. Now consider the geodesic triangle $\Delta(x', y_x, \tilde{y}_x)$ and set $\theta$, $\tilde{\theta}$ and $\theta'$ its internal angles. Note that $\theta = \frac{1}{2} (\exp_{x'}^{-1} x, \exp_{y_x}^{-1} \tilde{y}_x)$, $\tilde{\theta} = \frac{1}{2} (\exp_{y_x}^{-1} x, \exp_{x'}^{-1} \tilde{y}_x)$, and $\theta' = \frac{1}{2} (\exp_{\tilde{y}_x}^{-1} x, \exp_{y_x}^{-1} \tilde{y}_x)$. Since $y_x \neq \tilde{y}_x$ from Proposition 3.1 we have that $\theta \geq \frac{\pi}{2}$, $\tilde{\theta} \geq \frac{\pi}{2}$. As $d(x', y_x) = d(x', \tilde{y}_x)$ and $C$ is a convex set it follows that $\theta' > 0$. Therefore, $\theta + \tilde{\theta} + \theta' > \pi$ and this inequality contradicts the Theorem 2.2, which completes the proof of Proposition.

In the proof of the Proposition 3.2 is fundamental that $M$ is a Hadamard manifold. For example, the Proposition 3.2 is false in the Euclidean sphere. The unique point given by Proposition 3.2, is called projection of $x'$ onto the convex $C$ and is denoted by $p_C(x')$.

**Corollary 3.1** Let $x' \in M$. There exists a unique projection $p_C(x')$. Furthermore, the following inequality holds

$$\langle \exp_{p_C(x')}^{-1} x', \exp_{p_C(x')}^{-1} x \rangle \leq 0,$$

for all $x \in C$.

**Proof** (see [1]) It is an immediate consequence of Propositions 3.1 and 3.2.

Next, we introduce the concept of support subspace which will play an important role in the main result of this section.

**Definition 3.1** Suppose that, the boundary $\partial C$ of $C$ is nonempty, $x' \in \partial C$, $s \in T_{x'} M$ and $s \neq 0$. The subspace of $T_{x'} M$ defined by

$$S_{s, x'} := \{ v \in T_{x'} M / \langle s, v \rangle = 0 \}$$

is said to be support to the set $C$ at $x'$ if $\langle s, \exp_{x'}^{-1} x \rangle \leq 0$, for all $x \in C$. 

Lemma 3.1 Let $\bar{x}$ in $M$ such that $\bar{x} \notin C$. Then there exists $s \in T_{\bar{x}}M$ such that

$$\sup \{-s, \exp_{\bar{x}}^{-1} x : x \in C\} < 0.$$  \hspace{1cm} (10)

Proof Take $x \in C$. Consider the geodesic triangle $\Delta(\bar{x}, x, pC(\bar{x}))$. Set, $\beta = \langle \exp_{\bar{x}}^{-1} pC(\bar{x}), \exp_{\bar{x}}^{-1} x \rangle$ and $\theta = \langle \exp_{pC(\bar{x})}^{-1} \bar{x}, \exp_{pC(\bar{x})}^{-1} x \rangle$. By (5) we have

$$d(\bar{x}, x) \cos \beta + d(x, pC(\bar{x})) \cos \theta \geq d(\bar{x}, pC(\bar{x})).$$ \hspace{1cm} (11)

Now consider in $T_{\bar{x}}M$ the geodesic triangle $\Delta(0, \exp_{\bar{x}}^{-1} x, \exp_{\bar{x}}^{-1} pC(\bar{x}))$ and let $s := \exp_{\bar{x}}^{-1} pC(\bar{x})$ and $\alpha = \langle -s, \exp_{\bar{x}}^{-1} x - s \rangle$. Noting that in $T_{\bar{x}}M$ the inequality (5) holds with equality, we get

$$||\exp_{\bar{x}}^{-1} x|| \cos \beta + ||\exp_{\bar{x}}^{-1} x - s|| \cos \alpha = ||s||.$$ \hspace{1cm} (12)

Since $d(\bar{x}, pC(\bar{x})) = ||\exp_{pC(\bar{x})}^{-1} \bar{x}||$ and $||s|| = ||\exp_{\bar{x}}^{-1} pC(\bar{x})||$ it follows from (11) and (12) that $d(x, pC(\bar{x})) \cos \theta \geq ||\exp_{\bar{x}}^{-1} x - s|| \cos \alpha$. Note that $d(x, pC(\bar{x})) = ||\exp_{\bar{x}}^{-1} x||$. Then this implies $\langle -s, \exp_{\bar{x}}^{-1} x - s \rangle \leq \langle \exp_{pC(\bar{x})}^{-1} \bar{x}, \exp_{pC(\bar{x})}^{-1} x \rangle$. From Corollary 3.1, $\langle \exp_{pC(\bar{x})}^{-1} \bar{x}, \exp_{pC(\bar{x})}^{-1} x \rangle \leq 0$, then by this, and the last inequality, we have $\langle -s, \exp_{\bar{x}}^{-1} x - s \rangle \leq 0$, which is the same that $\langle -s, \exp_{\bar{x}}^{-1} x \rangle \leq -||s||^2$, this implies the statement of the Proposition.

We proceed now to prove the main result of this section.

Theorem 3.1 Let $x' \in \partial C$. There exists a support subspace to $C$ at $x'$.

Proof Consider the unit tangent bundle $UM := \bigcup_{x \in M} U_x M$, where $U_x M = \{u \in T_x M; ||u|| = 1\}$. Take the sequence $\{x_k\}$ in $M$ such that $x_k \notin C$, for $k = 1, 2, \ldots$ and $\lim_{k \to \infty} x_k = x'$. For each $x_k$, by Lemma 3.1 there exists $s_k \in T_{x_k} M$, such that

$$\langle -s_k, \exp_{x_k}^{-1} x \rangle < 0,$$ \hspace{1cm} (13)

for all $x \in C$. Without loss of generality, we can assume that $||s_k|| = 1$, i.e., that the sequence $\{s_k\}$ is in $UM$. Since $\lim_{k \to \infty} x_k = x'$, we have that the subset $B := \bigcup_{k=1}^\infty U_{x_k} M$ of $UM$ is compact. Note that $\{s_k\} \in B$, there exists a subsequence $\{s_{k_j}\}$ of $\{s_k\}$ and $\bar{s} \in T_{x'\bar{x}} M$ such that
lim_{k \to \infty} s_k = \bar{s}. From (13) we get
\[ \langle -s_k, \exp_{x_k}^{-1} x \rangle < 0, \]  
for all \( x \in C \). Now taking limit in the inequality (14) we obtain
\[ \langle -\bar{s}, \exp_{x'}^{-1} x \rangle \leq 0 \] for all \( x \in C \). Therefore \( S_{\bar{s}, x'} \) defined by (9), where \( s = -\bar{s} \), is the support subspace to \( C \) at \( x' \).

### 3.2. Convex Function

Let \( M \) be a Riemannian manifold. A function \( f: M \to \mathbb{R} \) is said to be convex (respectively, strictly convex) if the composition \( f \circ \gamma: \mathbb{R} \to \mathbb{R} \) is convex (respectively, strictly convex) for any geodesic \( \gamma \) of \( M \). This definition implies that all convex functions are continuous.

**Theorem 3.2** A function \( f: M \to \mathbb{R} \) is convex if and only if its epigraph
\[ \text{epi}(f) = \{ (x, r) \in M \times \mathbb{R} / f(x) \geq r \} \]
is a convex subset in \( M \times \mathbb{R} \).

**Proof** The proof is similar to that for \( M = \mathbb{R}^n \). It is enough to observe that \( \alpha = (\gamma, \beta) \) is a geodesic of \( M \) if and only if \( \gamma \) is a geodesic of \( M \) and \( \beta \) is a geodesic of \( R \).

Let \( M \) be a Hadamard manifold, let \( f: M \to \mathbb{R} \) be a convex function. Take \( x' \in M \), the vector \( s \in T_{x'} M \) is said to be subgradient of \( f \) at \( x' \in M \) if
\[ f(x) \geq f(x') + \langle s, \exp_{x'}^{-1} x \rangle, \]  
for all \( x \in M \). The set of all subgradients of \( f \) at \( x' \) is called the subdifferential of \( f \) at \( x' \) and is denoted by \( \partial f(x') \). An analytical proof of the next result appear in [17], we show a geometric proof.

**Theorem 3.3** Let \( M \) be a Hadamard manifold and let \( f: M \to \mathbb{R} \) be convex. Then, for any \( x' \in M \), there is \( s \in T_{x'} M \) such that
\[ f(x) \geq f(x') + \langle s, \exp_{x'}^{-1} x \rangle, \]  
for all \( x \in M \). In other words, the subdifferential \( \partial f(x') \) of \( f \) at \( x' \in M \) is nonempty.
Proof Since $f$ is continuous, from Theorem 3.2 the $\text{epi}(f)$ is a closed convex subset of $M \times \mathbb{R}$. Moreover, the boundary of $\text{epi}(f)$ is $\partial(\text{epi}(f)) = \{(x, f(x)) : x \in M\}$. The exponential map on $M \times \mathbb{R}$ at $(x', f(x'))$ is defined by $\exp_{(x', f(x'))}^{-1}(x, r) := (\exp_{x'}^{-1}x, r - f(x'))$. Then, from Theorem 3.1, we can take a nontrivial supporting subspace $S_{((s, \alpha), (x', f(x')))}$ to $\text{epi}(f)$ at $(x', f(x'))$. Therefore

$$\langle s, \exp_{x'}^{-1}x \rangle + \alpha(r - f(x')) = \langle (s, \alpha), (\exp_{x'}^{-1}x, r - f(x')) \rangle \leq 0,$$  \hspace{1cm} (18)

for all $(x, r) \in \text{epi}(f)$. Let $\bar{x} = \exp_{x'} s$ and let $\bar{r} > f(\bar{x})$. Taking $x = \bar{x}$ and $r = \bar{r}$ into (18) we obtain $\|s\|^2 \leq \alpha(f(x') - \bar{r})$. It follows that $\alpha \neq 0$, because the supporting subspace is nontrivial. Without loss of generality we can take $\alpha = -1$. Then taking $r = f(x)$ and $\alpha = -1$ at (18) we have $f(x) \geq f(x') + \langle s, \exp_{x'}^{-1}x \rangle$ for all $x \in M$, which is the statement of the Theorem. 

The Riemannian distance plays a fundamental role in the next section. We proceed now stating a result which we will go to use. Let $M$ be a Hadamard manifold. For any $x' \in M$ we can define the exponential inverse map $\exp_{x'}^{-1} : M \rightarrow T_{x'} M$ which is $C^\infty$; as $d(x', x) = \|\exp_{x'}^{-1}x\|$, then the map $\rho_{x'} : M \rightarrow \mathbb{R}$ defined by

$$\rho_{x'}(x) = \frac{1}{2} d^2(x, x')$$ \hspace{1cm} (19)

is $C^\infty$ too.

**Proposition 3.3** Let $M$ be a Hadamard manifold. Let $x' \in M$, then the map $\rho_{x'}$, defined in (19), is strictly convex and its gradient at $x$ is $\text{grad} \rho_{x'}(x) = -\exp_{x'}^{-1}x'$.

**Proof** see in [14 and 15].

All proofs in this section are analogous to the proofs of its similar facts in Euclidean space $\mathbb{R}^n$. That is happen because Hadamard manifolds and Euclidean space $\mathbb{R}^n$ have similar geometrical properties.

### 4. REGULARIZATION

Let $M$ be a Hadamard manifold and let $f : M \rightarrow \mathbb{R}$ be a convex function. For $\lambda > 0$, the *Moreau-Yosida regularization* $f_\lambda : M \rightarrow \mathbb{R}$ of $f$ is
defined by

\[ f_\lambda(x) = \inf_{y \in M} \{f(y) + \lambda \rho_x(y)\}, \tag{20} \]

where \( \rho_x \) is given by (19).

In order to prove that the function (20) is well defined, some definitions and results are necessary.

**Definition 4.1** A function \( h: M \to R \) is said to be 1-coercive at \( x \in M \) if

\[ \lim_{d(x,y) \to +\infty} \frac{h(y)}{d(x,y)} = +\infty. \tag{21} \]

If \( h: M \to R \) is a 1-coercive function at \( x \in M \), then it is easy to show that the minimizer set of \( h \) is nonempty.

**Lemma 4.1** If \( f: M \to R \) is convex and \( \lambda > 0 \), then the function \( (f + \lambda \rho_x): M \to R \) is 1-coercive at \( x \in M \).

**Proof** Since \( f \) is convex, from Theorem 3.3, it follows that there exists \( s \in T_xM \) such that

\[
\frac{(f + \lambda \rho_x)(y)}{d(x,y)} \geq \frac{f(x)}{d(x,y)} + \left( s, \frac{\exp^{-1}_x y}{d(x,y)} \right) + \lambda \frac{\rho_x(y)}{d(x,y)}
\]

\[
\geq \frac{f(x)}{d(x,y)} + \left( s, \frac{\exp^{-1}_x y}{d(x,y)} \right) + \frac{\lambda}{2} d(x,y).
\]

This inequality gives

\[
\lim_{d(x,y) \to +\infty} \frac{(f + \lambda \rho_x)(y)}{d(x,y)} = +\infty,
\]

since \( \|\exp^{-1}_x y\| = d(x,y) \) and \( \lambda > 0 \).

**Lemma 4.2** If \( f: M \to R \) is convex then, for any \( x \in M \) and \( \lambda > 0 \) there exists a unique point, denoted by \( p_\lambda(x) \), such that

\[ f(p_\lambda(x)) + \lambda \rho_x(p_\lambda(x)) = f_\lambda(x). \tag{22} \]

characterized by

\[ \lambda(\exp^{-1}_{p_\lambda(x)}x) \in \partial f(p_\lambda(x)). \tag{23} \]
Proof From Proposition 3.3 we have that \( p \) is strictly convex and, therefore \( f + \lambda p \) is strictly convex. Thus, it has at most one minimum point. On the other hand, from Lemma 4.1, \( f + \lambda p \) is 1-coercive then it has at least one minimum. Therefore \( f + \lambda p \) has a unique minimum, and the equality (22) follows from (20). Denoting by \( p_{\lambda}(x) \) the minimum point of \( f + \lambda p \) implies that \( 0 \in \partial ( f + \lambda p ) ( p_{\lambda}(x) ) \). However \( \partial ( f + \lambda p ) (y) = \partial f(y) + \{ -\lambda \exp_{x}^{-1} x \} \), this way (23) is a characterization for \( p_{\lambda}(x) \).

\[ \text{5. PROXIMAL POINT ALGORITHM} \]

Let \( M \) be a Hadamard manifold. Let \( f : M \to \mathbb{R} \) be a convex function. The proximal point algorithm generates, for a starting point \( p_{0} \in M \), a sequence \( \{ x_{k} \} \subset M \) by the iteration

\[
x_{k+1} = p_{\lambda_{k}}(x_{k})
= \arg \min_{y \in M} \{ f(y) + \lambda_{k} p_{x_{k}}(y) \}, \tag{24}
\]

where \( p_{x_{k}} \) is defined by (19) and \( \{ \lambda_{k} \} \) is a sequence of positive numbers. In the particular case in which \( M = \mathbb{R}^{n} \) we have \( p_{x_{k}}(y) = \| x_{k} - y \| \). In this case, the iteration (24) reduces to

\[
x_{k+1} = \arg \min_{y \in \mathbb{R}^{n}} \{ f(y) + \lambda_{k} \| x_{k} - y \|^{2} \}.
\]

Therefore the proximal point algorithm on Riemannian manifold is a natural generalization of the proximal point algorithm on \( \mathbb{R}^{n} \) introduced by Martinet [10] and [13].

**Theorem 5.1** Let \( f : M \to \mathbb{R} \) be a convex function, where \( M \) is a Hadamard manifold. The sequence \( \{ x_{k} \} \) generated by (24) is well defined, and characterized by

\[ \lambda_{k} ( \exp_{x_{k+1}}^{-1} x_{k} ) \in \partial f(x_{k+1}). \]

Proof The result follows immediately from Lemma 4.2 and the definition of the sequence \( \{ x_{k} \} \).
6. CONVERGENCE OF THE PROXIMAL POINT ALGORITHM

In this section is proved that the sequence given by (24) converges, if \( \sum_{k=0}^{\infty} (1/\lambda_k) = +\infty \) and the minimizer set \( O^* \neq \emptyset \). For this, again, some definitions and results are necessary. Let \((M, d)\) be a complete metric space. The sequence \( \{x_k\} \subset M \) is said to be Fejér convergent to the nonempty set \( U \subset M \) when

\[
d(x_{k+1}, y) \leq d(x_k, y),
\]

for all \( y \in U \) and \( k \geq 0 \).

**Lemma 6.1** Let \((M, d)\) be a complete metric space. If \( \{x_k\} \subset M \) is Fejér convergent to a nonempty set \( U \subset M \), then \( \{x_k\} \) is bounded. Furthermore, if a cluster point \( x \) of \( \{x_k\} \) belongs to \( U \), then \( \lim_{k \to +\infty} x_k = x \).

**Proof** (see [9]) Take \( u \in U \). The inequality (25) implies \( d(x_k, u) \leq d(x_0, u) \) for all \( k \), therefore \( \{x_k\} \) is bounded. Now let \( \{x_{k_j}\} \) be a subsequence of \( \{x_k\} \) such that \( \lim_{k \to +\infty} x_{k_j} = x \). As \( x \in U \), by (25), the sequence of positive numbers \( \{d(x_k, x)\} \) is decreasing and it has a subsequence, namely \( \{d(x_{k_j}, x)\} \), which converges to 0. Then the whole sequence converges to 0, i.e., \( 0 = \lim_{k \to +\infty} d(x_k, x) \) implying \( x = \lim_{k \to +\infty} x_k \).

**Lemma 6.2** Let \( M \) be a Hadamard manifold and let \( f: M \to R \) a convex function. If the sequence \( \{x_k\} \) is generated by (24), then the following inequality holds

\[
d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x) - f(x_{k+1})),
\]

for all \( x \in M \).

**Proof** Take \( x \in M \). Consider the geodesic triangle \( \Delta(x_k, x_{k+1}, x) \). Set \( \theta = \angle \langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle \), by Theorem 2.2 we have

\[
d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2d(x_k, x_{k+1})d(x_{k+1}, x) \cos \theta \leq d^2(x_k, x).
\]

Since \( \langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle = d(x_k, x_{k+1})d(x_{k+1}, x) \cos \theta \), so that

\[
d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle \leq d^2(x_k, x).
\]
This inequality, Theorem 5.1, and the definition of $\partial f(x_{k+1})$ imply

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + 2\langle \exp^{-1}_{x_{k+1}} x_k, \exp^{-1}_{x_{k+1}} x \rangle$$

$$\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x) - f(x_{k+1})), $$

which is what we want.

**Theorem 6.1** Let $\{x_k\}$ be the sequence generated by (24). If the sequence $\{\lambda_k\}$ is such that $\sum_{k=0}^{\infty} (1/\lambda_k) = +\infty$, then $\lim_{k \to +\infty} f(x_k) = f^*$, where $f^* = \inf_{x \in M} f(x)$. If, in addition, the minimizer set $U^*$ is nonempty, then $\lim_{k \to +\infty} x_k = x_*$ and $x_* \in U^*$.

**Proof** If $x_k \notin U^*$, substituting $x$ by $x_k$ in Lemma 6.2 we have that $f(x_{k+1}) < f(x_k)$, then the sequence $\{f(x_k)\}$ is strictly decreasing. Now we must prove that $\lim_{k \to +\infty} f(x_k) = f^*$. Assume, on the contrary, that $\lim_{k \to +\infty} f(x_k) > f^*$. Then there exist $x \in M$ and $\delta > 0$ such that $f(x) - f(x_k) > \delta$, for all $k$. This inequality and Lemma 6.2 imply $d^2(x_{k+1}, x) < d^2(x_k, x) - (2\delta/\lambda_k)$ for all $k$, and least inequality implies

$$\sum_{k=0}^{j} \frac{1}{\lambda_k} \leq \frac{1}{2\delta} (d^2(x_0, x) - d^2(x_0, x_{j+1})) < \frac{1}{2\delta} d^2(x_0, x)$$

for all $j$, which contradicts the equality $\sum_{k=0}^{\infty} (1/\lambda_k) = +\infty$. If the set $U^*$ is nonempty, given $\bar{x} \in U^*$, then $f(\bar{x}) \leq f(x_k)$ for all $k$. From this inequality and by substituting $\bar{x}$ into (26) we obtain $d^2(x_{k+1}, \bar{x}) < d^2(x_k, \bar{x})$, therefore the sequence $\{x_k\}$ is Fejér convergent to $U^*$. From Lemma 6.1 we have that the sequence $\{x_k\}$ is bounded, and because $M$ is a complete manifold, it has a convergent subsequence. Let $\{x_{h_j}\}$ be a subsequence of $\{x_k\}$ such that $\lim_{k \to +\infty} x_{k_j} = x_*$. Since $f$ is continuous and $\lim_{k \to +\infty} f(x_k) = f^*$ it follows that $\lim_{k \to +\infty} f(x_{h_j}) = x_*$, this implies $x \in U^*$. Therefore the cluster point $x_*$ of $\{x_k\}$ belongs to $U^*$, again by Lemma 6.1 $\lim_{k \to +\infty} x_k = x_*$. 

**7. FINAL REMARKS**

The proximal point algorithm for minimization proposed in [6] and [13] solves unconstrained convex problems in $R^n$. Our algorithm solves constrained non convex problems in $R^n$ when the constraint set is a
Riemannian manifold. Then, we can say that our algorithm generalizes, in a certain sense, the algorithm presented in [6] and [13], namely, if the Riemannian manifold is the Euclidean space $\mathbb{R}^n$, then both algorithms are the same. We remark that the techniques used in the definition and convergence proof of the algorithm impose restrictions about the manifold, that is, we needed that the Riemannian manifold has non-positive curvature. Still it remains to remove this hypothesis, namely, in which manifolds, other than Hadamard ones, is it possible to define proximal point method.

References