



Proximal subgradient and a characterization of Lipschitz function on Riemannian manifolds

Orizon P. Ferreira¹

*Universidade Federal de Goiás, Instituto de Matemática e Estatística, Campus Samambaia,
Caixa Postal 131, Goiânia, GO 74001-970, Brazil*

Received 6 February 2004

Available online 12 September 2005

Submitted by Steven G. Krantz

Abstract

A characterization of Lipschitz behavior of functions defined on Riemannian manifolds is given in this paper. First, it is extended the concept of proximal subgradient and some results of proximal analysis from Hilbert space to Riemannian manifold setting. A technique introduced by Clarke, Stern and Wolenski [F.H. Clarke, R.J. Stern, P.R. Wolenski, Subgradient criteria for monotonicity, the Lipschitz condition, and convexity, *Canad. J. Math.* 45 (1993) 1167–1183], for generating proximal subgradients of functions defined on a Hilbert spaces, is also extended to Riemannian manifolds in order to provide that characterization. A number of examples of Lipschitz functions are presented so as to show that the Lipschitz behavior of functions defined on Riemannian manifolds depends on the Riemannian metric.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Lipschitz functions; Proximal subgradient; Riemannian manifolds

1. Introduction

Extensions of concepts and techniques from Euclidean space to Riemannian manifold are natural. It has been done frequently in the last few years, with theoretical objectives and also in order to obtain effective algorithms of optimization on Riemannian manifold setting (see [7,8,10]). Re-

E-mail addresses: orizonpf@brturbo.com.br, orizon@mat.ufg.br.

¹ The author was supported in part by FUNAPE/UFG, CAPES, PADCT-CNPq and PRONEX-Otimização (FAPERJ/CNPq).

cently, Azagra et al. [2] have extended several important nonsmooth analysis concepts from Hilbert space to Riemannian manifold.

Our aim is to provide a characterization of Lipschitz behavior of functions defined on Riemannian manifolds in terms of the proximal subgradient. First, we extend the concept of proximal subgradient and some basics results of proximal analysis from Hilbert space to Riemannian manifold setting (see [3–5]). In order to obtain this characterization, we also extend to Riemannian manifolds setting the technique introduced by Clarke et al. [5] for generating proximal subgradient. Let us mention that it has been pointed out by Azagra et al. [2] that the extension to Riemannian manifolds of the concept of proximal subgradient is quite natural.

We will present a number of examples of Lipschitz functions so as to show that the Lipschitz behavior of functions defined on Riemannian manifolds depends on the Riemannian metric. Let us recall the definition of Lipschitz function defined on metric space: Let U be a subset of M , where (M, d) is a metric space. The function $f: U \rightarrow \mathbb{R}$ is said to be Lipschitz of rank $L \geq 0$ if it satisfies

$$|f(p) - f(q)| \leq Ld(p, q), \quad (1)$$

for all $p, q \in U$. Since the Lipschitz properties depend on the distance d defined on M , the Riemannian geometry gives many tools for investigating the Lipschitz behavior of function from a new viewpoint. Let us give an example: define $f: \mathbb{R}_{++}^2 \rightarrow \mathbb{R}$ by

$$f(p) = |\ln(p_1)| + |\ln(p_2)|, \quad \text{where } p = (p_1, p_2).$$

Note that f is not a Lipschitz function on \mathbb{R}_{++}^2 endowed with the Euclidean metric $\langle \cdot, \cdot \rangle$. On the other hand, endowing \mathbb{R}_{++}^2 with a new Riemannian metric $\langle \cdot, \cdot \rangle_G$ defined by $\langle u, v \rangle_G = \langle G(p)v, u \rangle$, where $u, v \in \mathbb{R}^2 = T_p M$ and

$$G(p) = \text{diag}(p_1^{-2}, p_2^{-2}),$$

we obtain a Riemannian manifold $M_G = (M, G)$. Let $\Phi: \mathbb{R}^2 \rightarrow M_G$ be an isometry defined by

$$\Phi(x) = (e^{x_1}, e^{x_2}), \quad \text{where } x = (x_1, x_2).$$

Define $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ by $g(x) = |x_1| + |x_2|$. Since $g(x) = f(\Phi(x))$ is a Lipschitz function on \mathbb{R}^2 , the Proposition 4.1 below asserts that f is also a Lipschitz function on M_G . So, this approach permits us to show that the Lipschitz behavior of functions defined on Riemannian manifolds depends on the Riemannian metric.

The organization of the paper is as follows: in Section 2, we list some basic notations and terminology used in this presentation. In Section 3 we introduce the concept of proximal subgradient of functions defined on a Riemannian manifolds and also prove some results of proximal analysis in this setting. In Section 4 we will give a characterization and some examples of Lipschitz functions defined on Riemannian manifolds. We conclude this paper by making some general comments about extensions of our results from finite to infinite dimensional Riemannian manifolds as applications of the recent results due to Azagra and Ferrera [1].

2. Preliminaries

In this section we recall some notations, definitions and basic properties of Riemannian manifolds which will be used throughout the paper. They can be found in many introductory books on Riemannian geometry, for example, [6,9].

Throughout the paper, M is a smooth manifold and $C^k(U)$ is the class of all k -times continuously differentiable functions on the open subset $U \subseteq M$. We denote by $\mathfrak{X}(M)$ the space of vector fields on M and by T_pM the tangent space of M at p . Let M be endowed with a Riemannian metric $\langle \cdot, \cdot \rangle$, with corresponding norm denoted by $\| \cdot \|$, so that M is now a Riemannian manifold. Let us recall that the metric can be used to define the length of piecewise C^1 curve $c : [a, b] \rightarrow M$ joining p to q , i.e., such that $c(a) = p$ and $c(b) = q$, by $l(c) = \int_a^b \|c'(t)\| dt$. Minimizing this length functional over the set of all such curves, we obtain a distance $d(p, q)$, which induces the original topology on M . Also, the metric induces a map $f \in C^1(M) \mapsto \text{grad } f \in \mathfrak{X}(M)$, which associates to each f its gradient via the rule $\langle \text{grad } f, X \rangle = df(X)$, for all $X \in \mathfrak{X}(M)$. The chain rule generalizes to this setting in the usual way: $(f \circ c)'(t) = \langle \text{grad } f(c(t)), c'(t) \rangle$, for all curve $c \in C^1$. Denoting the Levi-Civita connection associated to $(M, \langle \cdot, \cdot \rangle)$ by ∇ , the hessian of a function $f \in C^2(M)$ is given by $\text{hess } f(X, Y) = \langle \nabla_X \text{grad } f, Y \rangle$, for all $X, Y \in \mathfrak{X}(M)$. So, the chain rule implies $(f \circ c)''(t) = \text{hess } f_{c(t)}(c'(t), c'(t))$, for all curve $c \in C^2$. Let c be a curve joining points p and q in M . For each $t \in [a, b]$, ∇ induces an isometry, relative to $\langle \cdot, \cdot \rangle$, $P(c)_t^a : T_{c(a)}M \rightarrow T_{c(t)}M$, the so-called parallel transport along c from $c(a)$ to $c(t)$. A vector field V along c is said to be parallel if $\nabla_{c'}V = 0$. If c' itself is parallel, then we say that c is geodesic. The geodesic equation $\nabla_{\gamma'}\gamma' = 0$ is a second order nonlinear ordinary differential equation, so the geodesic γ is determined by its position and velocity at one point. It is easy to check that $\|\gamma'\|$ is constant. We say that γ is normalized if $\|\gamma'\| = 1$. The restriction of a geodesic to a closed bounded interval is called a geodesic segment. A geodesic segment joining p to q in M is said to be minimal if its length is equal $d(p, q)$.

A finite dimensional Riemannian manifold is complete if its geodesics are defined for any values of t . The Hopf–Rinow’s theorem asserts that if the Riemannian manifold M is complete, then any pair of points in M can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, (M, d) is a complete metric space and its closed and bounded subsets are compact. In this paper, we assume that all manifolds are complete and finite dimensional. The exponential map $\exp_p : T_pM \rightarrow M$ is defined by $\exp_p v = \gamma_v(1)$, where γ_v is the geodesic defined by its position p and velocity v at p . We can prove that $\exp_p tv = \gamma_v(t)$ for any values of t . Now, for $p \in M$, let

$$i_p = \sup\{r > 0 : \exp_p : B_r(o_p) \rightarrow M \text{ is diffeomorphism}\},$$

where o_p denotes the origin of T_pM and $B_r(o_p) = \{v \in T_pM : \|v - o_p\| < r\}$. Note that if $0 < \delta < i_p$ then $\exp_p B_\delta(o_p) = B_\delta(p)$, where $B_\delta(p) = \{q \in M : d(p, q) < \delta\}$. We say that i_p is the injectivity radius of M at p . It is well known that, for any $p \in M$, the map $d^2(p, \cdot) \in C^\infty(B_{i_p}(p))$, where $B_{i_p}(p) = \{q \in M : d(p, q) < i_p\}$, and $\text{grad } d^2(p, q) = -2 \exp_q^{-1} p$, for all $q \in B_{i_p}(p)$.

Let M and N be Riemannian manifolds and let $\Phi : M \rightarrow N$ be an isometry, that is, Φ is $C^\infty(M)$ and, for all $p \in M, u, v \in T_pM, \langle D\Phi_p u, D\Phi_p v \rangle = \langle u, v \rangle$, where $D\Phi_p : T_pM \rightarrow T_{\Phi(p)}N$ is the differential of Φ at p . We can verify that, Φ preserves the Levi-Civita connection. In particular, Φ preserves geodesics, that is, γ is geodesic in M if and only if $\beta = \Phi \circ \gamma$ is geodesic in N , as a consequence $D\Phi_{\gamma(t)}\gamma'(t) = \beta'(t)$. Furthermore, Φ preserves the distance, that is, for all $p, q \in M, d(\Phi(p), \Phi(q)) = d(p, q)$.

Let U be an open subset of M . From now on, we denote by $\mathcal{F}(U)$ the class of all functions $f : M \rightarrow (-\infty, +\infty]$ which are lower semicontinuous on U and $\text{dom}(f) \cap U \neq \emptyset$, where $\text{dom}(f) = \{x \in M : f(x) < +\infty\}$. If $U = M$ we denote \mathcal{F} for $\mathcal{F}(U)$.

3. The proximal subgradient

In this section we introduce the concept of proximal subgradient of functions defined on a Riemannian manifolds and also prove some results of proximal analysis in this setting.

Let us to start with the definition of proximal subgradient of functions defined on a Riemannian manifolds. Let $f \in \mathcal{F}$. A vector $v \in T_p M$ is a proximal subgradient of f at $p \in \text{dom}(f)$ if there exist $\sigma > 0$ and $0 < \delta < i_p$ such that

$$f(q) \geq f(p) + \langle v, \exp_p^{-1} q \rangle - \sigma d^2(p, q), \quad (2)$$

for all $q \in B_\delta(p)$. We denote the set of all proximal subgradients of f at p by $\partial_P f(p)$.

Remark 1. Now, we remark two important properties of proximal subgradient.

- (i) Let $\Phi : M \rightarrow N$ be an isometry and let $f : N \rightarrow \mathbb{R}$, where M and N are Riemannian manifolds. Then, it easy to prove that $v \in \partial_P f(p)$ if and only if, for all $p \in M$,

$$d\Phi_p^{-1} v \in \partial_P (f \circ \Phi)(\Phi^{-1} p).$$

- (ii) Let $p \in U$ be a local minimum of $f \in \mathcal{F}(U)$. Then $0 \in \partial_P f(p)$.

Proposition 3.1. Let $f \in \mathcal{F}(U)$ and $p \in \text{dom}(f)$. A vector $v \in T_p M$ is a proximal subgradient of f at p if and only if there exist $0 < \delta < i_p$ and $h \in C^2(B_\delta(p))$ with $h \leq f$, $h(p) = f(p)$ and $\text{grad} h(p) = v$.

Proof. Let $p \in \text{dom}(f)$ and $v \in \partial_P f(p)$. Since $v \in \partial_P f(p)$ there exist $0 < \delta < i_p$ such that $h \leq f$, where $h(q) = f(p) + \langle v, \exp_p^{-1} q \rangle - \sigma d^2(p, q)$ for all $q \in B_\delta(p)$. It is straightforward to show that $h(p) = f(p)$,

$$\text{grad} d^2(p, q)|_{q=p} = 0 \quad \text{and} \quad \text{grad} \langle v, \exp_p^{-1} q \rangle|_{q=p} = v.$$

Thus, $\text{grad} h(p) = v$. Now, as functions $d^2(p, \cdot)$ and $\langle v, \exp_p^{-1} \cdot \rangle$ are in $C^2(B_\delta(p))$ the statement follows.

For the converse, take $p \in \text{dom}(f)$ and $q \in B_{i_p}(p) \cap U$. Consider the geodesic $\gamma(t) = \exp_p t(\exp_p^{-1} q)$. Since h is in $C^2(B_\delta(p))$, there exists $t_* \in (0, t)$ such that

$$h(\gamma(t)) = h(p) + \langle \text{grad} h(p), \gamma'(0) \rangle t + \frac{1}{2} \text{hess} h_{\gamma(t_*)}(\gamma'(t_*), \gamma'(t_*)) t^2. \quad (3)$$

Owing to definition of γ , we have that $\gamma'(0) = \exp_p^{-1} q$ and $\|\gamma'(t)\| = d(p, q)$ for all t . Now, as $\text{hess} h$ is in $C(B_\delta(p))$, it is easy to see from (3) that there exists $\sigma > 0$ such that

$$h(q) \geq h(p) + \langle \text{grad} h(p), \exp_p^{-1} q \rangle - \sigma d^2(p, q). \quad (4)$$

Finally, because $f(q) \geq h(q)$ for all $q \in B_\delta(p)$ and $h(p) = f(p)$, we conclude from (4) that $v = \text{grad} h(p)$ is a proximal subgradient of f at p . \square

In particular, Proposition 3.1 implies that $\partial_P f(p) \subset D^- f(p)$, where $D^- f(p)$ is the viscosity subdifferential of f at p as defined in [2].

Proposition 3.2. *Let $f \in \mathcal{F}(U)$. The following statements hold:*

- (i) *if $f \in C^1(U)$ then we have that $\partial_P f(p) \subset \{\text{grad } f(p)\}$, for all $p \in U$;*
- (ii) *if $f \in C^2(U)$ then we have that $\partial_P f(p) = \{\text{grad } f(p)\}$, for all $p \in U$.*

Proof. For the first statement, take $p \in U$, $u \in T_p M$ and define the geodesic $\gamma(t) = \exp_p tu$. So, for $v \in \partial_P f(p)$, there exist $0 < \sigma$ and $0 < \delta < i_p$ such that

$$f(\gamma(t)) - f(p) \geq \langle v, \exp_p^{-1} \gamma(t) \rangle - \sigma d^2(p, \gamma(t)), \tag{5}$$

for all $0 < t < \delta/\|u\|$. Thus, since $\exp_p^{-1} \gamma(t) = tu$ and $d^2(p, \gamma(t)) = t^2\|u\|^2$ we have from (5) that

$$\frac{f(\gamma(t)) - f(p)}{t} \geq \langle v, u \rangle - t\sigma\|u\|^2, \tag{6}$$

for all $0 < t < \delta/\|u\|$. Because $f \in C^1(U)$, letting t goes to 0 in (6), we obtain

$$\langle \text{grad } f(p), u \rangle \geq \langle v, u \rangle.$$

Now, as the latter inequality holds for all $u \in T_p M$, we have $v = \text{grad } f(p)$.

For the second statement, first note that $f \in C^2(U)$ implies from Proposition 3.1 that $\text{grad } f(p) \in \partial_P f(p)$. Hence, the statement follows from item (i). \square

Lemma 3.1. *Let $f \in \mathcal{F}(U)$ and let $g \in C^2(U)$. Suppose that $p \in \text{dom } f$ and $v \in \partial_P(f + g)(p)$. Then*

$$v - \text{grad } g(p) \in \partial_P f(p).$$

Proof. Since $g \in C^2(U)$, we have from Proposition 3.2(ii) that there exist $0 < \sigma_1$ and $0 < \delta_1 < i_p$ such that

$$-g(q) \geq -g(p) + \langle -\text{grad } g(p), \exp_p^{-1} q \rangle - \sigma_1 d^2(p, q), \tag{7}$$

for all $q \in B_{\delta_1}(p) \cap U$. On the other hand, because $v \in \partial_P(f + g)(p)$ there exists $0 < \sigma_2$ and $0 < \delta_2 < i_p$ such that

$$f(q) + g(q) \geq f(p) + g(p) + \langle v, \exp_p^{-1} q \rangle - \sigma_2 d^2(p, q), \tag{8}$$

for all $q \in B_{\delta_2}(p) \cap U$. Setting $\delta = \min\{\delta_1, \delta_2\}$ it is easy to show from (7) and (8) that

$$f(q) - f(p) \geq \langle v - \text{grad } g(p), \exp_p^{-1} q \rangle - (\sigma_1 + \sigma_2)d^2(p, q),$$

for all $q \in B_{\delta}(p) \cap U$. Therefore, $v - \text{grad } g(p) \in \partial_P f(p)$. \square

Assume that the hypothesis of Lemma 3.1 holds. In particular, if the point p is a minimizer of $f + g$, then from Remark 1(ii) we have that $-\text{grad } g(p)$ is a proximal subgradient of f at p . We will use this simple device for generating proximal subgradient to obtain our main result. Next, we give an application for it, the idea can be translated from the corresponding one for viscosity subdifferential, namely [2, Proposition 4.17] (see also [5, Lemma 2.2]).

Theorem 3.1. Let $f \in \mathcal{F}(U)$. Then $\text{dom}(\partial_P f)$ is dense in $\text{dom}(f)$.

Proof. Let $p_0 \in \text{dom}(f)$. Since $f \in \mathcal{F}(U)$, take $\delta > 0$ such that f is bounded below on $B_\delta(p_0) \subset U$. Let $g_\delta : U \rightarrow [0, +\infty]$ be given by

$$g_\delta(p) = \begin{cases} \frac{1}{\delta^2 - d^2(p_0, p)}, & \text{if } p \in B_\delta(p_0), \\ +\infty, & \text{otherwise.} \end{cases}$$

Now, we have that $g_\delta \in C^2(B_\delta(p_0))$ and $(f + g)(p)$ goes to $+\infty$ as p approaches to the boundary of $B_\delta(p_0)$. Moreover, $f + g_\delta \in \mathcal{F}(U)$ is bounded below on $B_\delta[p_0]$, where $B_\delta[p_0]$ denotes the closure of $B_\delta(p_0)$. Thus, as M is a complete Riemannian manifold of finite dimension, there exists $p_\delta \in B_\delta(p_0)$ a minimizer of $f + g_\delta$. Therefore $0 \in \partial_P(f + g_\delta)(p_\delta)$ and from Lemma 3.1 we obtain that $-\text{grad } g_\delta(p_\delta) \in \partial f(p_\delta)$, which implies in particular that $\partial f(p_\delta) \neq \emptyset$. Since we may take δ arbitrarily small, the proof is complete. \square

It is worth pointing out that almost at the same time when this paper was being submitted to publication, the results of this section appeared in the preprint [1] due to D. Azagra and J. Ferrera. In [1] is presented a study of main results concerning proximal calculus on infinite dimensional Riemannian manifolds.

4. A characterization of Lipschitz function

In this section we present a characterization of Lipschitz functions defined on Riemannian manifolds. Also, we give some examples of Lipschitz functions.

Let us begin with the definition of Lipschitz functions. A function $f \in \mathcal{F}(U)$ is said to be Lipschitz on V , of rank $L \geq 0$, if $V \subset \text{dom}(f)$ and there holds

$$|f(p) - f(q)| \leq Ld(p, q), \quad (9)$$

for all $p, q \in V$, where d is the Riemannian distance on M . We denote the set of all Lipschitz functions on V , of rank L , by $\text{Lip}_L(V)$. A function $f \in \mathcal{F}(U)$ is said to be Lipschitz at p , of rank L_p , if there exists $\delta > 0$ such that $f \in \text{Lip}_{L_p}(B_\delta(p))$. Now, f is said to be locally Lipschitz on V if it is Lipschitz at every $p \in V$.

Remark 2. Note that, the Lipschitz properties depend on the Riemannian metric defined on M . In other words, if the metric on M is changed then the set of Lipschitz functions on M becomes different from the previous one.

Example 4.1. Let M be a noncompact Riemannian manifold. A geodesic $\gamma : [0, +\infty) \rightarrow M$ parameterized by arc-length and emanating from p is called a ray emanating from p if $d(\gamma(t), \gamma(s)) = |t - s|$, for all $t, s > 0$. For a ray γ , the Busemann function $b_\gamma : M \rightarrow \mathbb{R}$ is defined by

$$b_\gamma(q) = \lim_{t \rightarrow +\infty} (t - d(q, \gamma(t))),$$

see [9]. Note that $b_\gamma \in \text{Lip}_1(M)$. In fact, for all $p, q \in M$ we have

$$|b_\gamma(q) - b_\gamma(p)| \leq \lim_{t \rightarrow +\infty} |d(q, \gamma(t)) - d(p, \gamma(t))| \leq d(p, q).$$

Example 4.2. The distance function $d_S : M \rightarrow \mathbb{R}$, associated to $S \subset M$, is defined by

$$d_S(p) = \inf\{d(p, s) : s \in S\}. \tag{10}$$

We claim that $d_S \in \text{Lip}_1(M)$. Indeed, take $\epsilon > 0$ and $q \in M$. From (10), there exists $s \in S$ such that $d(q, s) < d_S(q) + \epsilon$. Therefore, we obtain that

$$d_S(p) \leq d(p, s) \leq d(p, q) + d(q, s) < d(p, q) + d_S(q) + \epsilon,$$

for all $p \in M$. Reversing the roles for p and q in the latter inequality and letting ϵ goes to 0, it is easy to see that $d_S \in \text{Lip}_1(M)$.

Example 4.3. Let $f \in \mathcal{F}$ and $\lambda > 0$. Suppose that f is bounded below by the constant c . Then it is easy to show that the function $f_\lambda : M \rightarrow (-\infty, +\infty]$ defined by

$$f_\lambda(p) = \inf\{f(q) + \lambda d^2(p, q)\},$$

is also bounded below by c . Moreover, following the same pattern used to prove the first part of [3, Theorem 5.1], we can prove that f_λ is locally Lipschitz on its domain.

Proposition 4.1. Let M and N be Riemannian manifolds and let $\Phi : M \rightarrow N$ be an isometry. A function $f : N \rightarrow \mathbb{R}$ is Lipschitz on N , of rank $L \geq 0$, if and only if $g : M \rightarrow \mathbb{R}$, defined by $g(p) = f(\Phi(p))$, is Lipschitz on M of rank $L \geq 0$.

Proof. Since isometries preserve the Riemannian distance the result follows. \square

Example 4.4. Let $f : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$ be defined by $f(p) = \sum_{i=1}^n \ln(p_i)$, where $p = (p_1, \dots, p_n)$. It is simple to show that the function f is not Lipschitz in \mathbb{R}_{++}^n endowed with the Euclidean metric $\langle \cdot, \cdot \rangle$. Let $G(p)$ be a $n \times n$ matrix defined by

$$G(p) = \text{diag}(p_1^{-2}, \dots, p_n^{-2}).$$

Now, endowing \mathbb{R}_{++}^n with the Riemannian metric $\langle u, v \rangle = \langle G(p)v, u \rangle$, we obtain a complete Riemannian manifold M_G . Let \mathbb{R}^n be endowed with the Euclidean metric. It is straightforward to show that the function $\Phi : \mathbb{R}^n \rightarrow M_G$ defined by

$$\Phi(x) = (e^{x_1}, \dots, e^{x_n}),$$

is an isometry, where $x = (x_1, \dots, x_n)$. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be defined by $g(x) = \sum_{i=1}^n x_i$. Because Φ is an isometry, the function g is Lipschitz of rank $L = 1$ on \mathbb{R}^n and $g(x) = f(\Phi(x))$, we have from Proposition 4.1 that f is also Lipschitz on M_G of rank $L = 1$.

Proposition 4.2. Let U be an open totally convex subset of M and let $f \in \mathcal{F}(U)$. If f is locally Lipschitz on $\text{dom } f$, then the following statements hold:

- (i) $\text{dom } f = U$;
- (ii) if $L_p = L$ for all $p \in U$ then $f \in \text{Lip}_L(U)$, where L_p is the Lipschitz constant of f at p .

Proof. For (i). First note that $\text{dom}(f) \subset U$. It remains to show that $U \subset \text{dom}(f)$. For that, let $q \in U$. Now, take $p \in \text{dom}(f)$ and a minimal geodesic segment γ such that $\gamma(0) = p$ and

$\gamma(1) = q$. Since U is totally convex it follows that $\gamma([0, 1]) \subset U$. We claim that f is finite at q , i.e., $q \in \text{dom}(f)$. Suppose not. Thus we have that $0 < t_* < 1$, where

$$t_* := \sup\{t \in (0, 1]: f(\gamma(t)) < +\infty\}.$$

Let $t' \in (0, t_*)$, then $\gamma([0, t']) \subset \text{dom } f$. Since $\gamma([0, t'])$ is compact and f is locally Lipschitz on $\text{dom } f$, we can take $0 = t_0 < t_1 < \dots < t_n = t'$ and positive numbers $\delta_0 \dots \delta_n$ such that

$$\gamma([t_i, t_{i+1}]) \subset B_{\delta_i}(\gamma(t_i)) \subset U \quad \text{and} \quad f \in \text{Lip}_{L_i}(B_{\delta_i}(\gamma(t_i))),$$

for all $i = 0, \dots, n - 1$, where L_i is the Lipschitz constant of f at $\gamma(t_i)$. Therefore, by local Lipschitz property and noting that the geodesic segment γ is minimal, we obtain

$$\begin{aligned} f(\gamma(t')) &= f(p) + \sum_{i=0}^n (f(\gamma(t_{i+1})) - f(\gamma(t_i))) \\ &\leq f(p) + \sum_{i=0}^n L_i d(\gamma(t_i), \gamma(t_{i+1})) \\ &\leq f(p) + Ld(p, \gamma(t')), \end{aligned} \tag{11}$$

where $L = \max\{L_i: i = 0, \dots, n - 1\}$. Hence, as f is lower semicontinuous, letting t' go to t_* , we have that

$$f(\gamma(t_*)) < +\infty,$$

so $\gamma(t_*) \in \text{dom}(f)$. According to f is Lipschitz at $\gamma(t_*)$, definition of t_* is violated. Therefore, we obtain that f is finite on the entire segment $\gamma([0, 1])$. In particular, $q \in \text{dom}(f)$, so $U \subset \text{dom}(f)$ and the first statement follows.

For (ii). Let $p, q \in U$. With an analogous argument used to obtain (11) we can show that

$$f(q) \leq f(p) + Ld(p, q).$$

Now, by reversing the roles of p and q , it easy to conclude that $f \in \text{Lip}_L(U)$, and the second statement is proved. \square

Theorem 4.1. *Let U be an open totally convex subset of M and let $f \in \mathcal{F}(U)$. Then f is a Lipschitz function on U , of rank $L \geq 0$, if and only if*

$$\|v\| \leq L, \tag{12}$$

for all $v \in \partial_P f(p)$ and all $p \in U$. As a consequence, f is constant on U if and only if $\partial_P f(p) \subset \{0\}$, for all $p \in U$.

Proof. First we assume that f is a Lipschitz function, of rank L , on U . Take $p \in U$ and $v \in \partial_P f(p)$. Let γ be the normalized geodesic defined by

$$\gamma(t) = \exp_p t \frac{v}{\|v\|}.$$

Then, from (2), there exist constants $0 < \sigma$ and $0 < \delta < i_p$ such that

$$f(\gamma(t)) - f(p) \geq \langle v, \exp_p^{-1} \gamma(t) \rangle - \sigma d^2(p, \gamma(t)), \tag{13}$$

for all $0 < t < \delta$. Since f is a Lipschitz function of rank L , $t = d(p, \gamma(t))$ and $\exp_p^{-1} \gamma(t) = t(v/\|v\|)$, we obtain from (13) that

$$L \geq \left\langle v, \frac{v}{\|v\|} \right\rangle - \sigma t = \|v\| - \sigma t,$$

for all $0 < t < \delta$. Thus, taking limit as t goes to 0 in the above inequality, we obtain (12).

For the converse, let $p_0 \in \text{dom}(f)$. Since $f \in \mathcal{F}(U)$, take $\delta > 0$ such that f is bounded below on $B_{3\delta}(p_0) \subset U$. Let $K > L$, $q \in B_\delta(p_0)$ and the function $g : U \rightarrow [0, +\infty]$ be defined by

$$g(p) = \begin{cases} Kd(p, q) & \text{if } p \in B_\delta[q], \\ Kd(p, q) + \frac{(d(p, q) - \delta)^2}{2\delta - d(p, q)} & \text{if } p \in B_{2\delta}(q) \setminus B_\delta(q), \\ +\infty & \text{otherwise,} \end{cases}$$

where $B_\delta[q]$ denotes the closure of $B_\delta(q)$. It is easy to show that g is lower semicontinuous, $g \in C^2(B_{2\delta}(q) \setminus \{q\})$, and for each $p \in B_{2\delta}(q) \setminus \{q\}$ there holds

$$\|\text{grad } g(p)\| \geq K > L. \tag{14}$$

Now, note that $f + g$ is in $\mathcal{F}(U)$, goes to $+\infty$ as p goes to the boundary of $B_{2\delta}(q)$ and is bounded below on $B_{2\delta}(q)$. Therefore, as M is a complete Riemannian manifold of dimension finite, there exists $p_* \in B_{2\delta}(q)$ a minimizer for $f + g$. First assume that $p_* \neq q$. Since $0 \in \partial_P(f + g)(p_*)$ we obtain from Lemma 3.1 that $-\text{grad } g(p_*) \in \partial_P f(p_*)$, so (14) contradicts (12). Consequently $p_* = q$, hence

$$f(q) \leq (f + g)(p) \leq f(p) + Kd(p, q),$$

for all $p \in B_\delta(q)$. Now, by changing the roles of q and p in the above argument, it is easy to see that

$$|f(q) - f(p)| \leq Kd(p, q),$$

for all $p, q \in B_\delta(p_0)$. Letting K go to L , in the latter inequality, we conclude that for any $p_0 \in \text{dom } f$ there exists $\delta > 0$ such that $f \in \text{Lip}_L B_\delta(p_0)$. Thus we have shown that f is locally Lipschitz on $\text{dom}(f)$, with the same rank L for all point. Therefore, the statement follows from Proposition 4.2 and the first part of the theorem is proved. The second part is an immediate consequence of the first part. \square

Example 4.5. Let S_{++}^n be the set of positive definite matrices endowed with the Frobenius metric defined by $\langle U, V \rangle = \text{tr}(VU)$. Let $f : S_{++}^n \rightarrow \mathbb{R}$ be defined by $f(X) = \ln \det X$.

It is easy to see that the function f is not Lipschitz on S_{++}^n . Now, endowing S_{++}^n with the Riemannian metric

$$\langle\langle U, V \rangle\rangle = \langle X^{-1}UX^{-1}, V \rangle,$$

we obtain a complete Riemannian manifold. We denote by M this Riemannian manifold. Because the gradient of f on S_{++}^n is $\nabla f(X) = -X^{-1}$ we have that the gradient of f on M is given by

$$\text{grad } f(X) = X\nabla f(X)X = -X.$$

So $\|\text{grad } f(X)\|^2 = \langle\langle \text{grad } f(X), \text{grad } f(X) \rangle\rangle = 1$. Now, from Theorem 4.1 and Proposition 3.2 it follows that f is Lipschitz on M of rank $L = 1$.

Example 4.6. Let $\Omega = \{p = (p_1, p_2) \in \mathbb{R}^2: p_2 > 0\}$ and let $f: \Omega \rightarrow \mathbb{R}$ be given by

$$f(p) = \ln(p_2).$$

It is easy to see that f is not a Lipschitz function on Ω with respect to Euclidean metric $\langle \cdot, \cdot \rangle$. Let G be a 2×2 matrix defined by $G(p) = (g_{ij}(p))$, where

$$g_{11}(p) = g_{22}(p) = \frac{1}{p_2^2}, \quad g_{12}(p) = g_{21}(p) = 0.$$

Endowing Ω with the Riemannian metric $\langle \cdot, \cdot \rangle$ defined by

$$\langle u, v \rangle = \langle G(p)v, u \rangle = v^T G(p)u,$$

we obtain a complete Riemannian manifold, namely, the upper half-plane model of the hyperbolic space \mathbb{H}^2 . The gradient of f in \mathbb{H}^2 is given by

$$\text{grad } f(p) = G(p)^{-1} \nabla f(p) = (0, p_2),$$

where ∇f is the gradient of f in Ω . It is simple to show that $\|\text{grad } f(p)\|^2 = 1$. Therefore, by Theorem 4.1 and Proposition 3.2 it follows that f is Lipschitz on \mathbb{H}^2 of rank $L = 1$.

5. Final remarks

It has been pointed out to us by the referee that several important concepts of nonsmooth analysis have been extended from Hilbert space to Riemannian manifold setting due to Azagra and Ferrera [1] and Azagra et al. [2]. In particular, in [1] was presented a study of the main results concerning proximal calculus on infinite dimensional Riemannian manifolds.

In this paper we have obtained results only for finite dimensional Riemannian manifolds. However, our main result, Theorem 4.1, can be extended to infinite dimensional Riemannian manifolds in view of [1, Theorem 15], after a more accurate analysis of our proof. As a consequence, we can obtain [2, Theorem 4.18] (Deville's mean value inequality), since Proposition 3.1 implies that $\partial_p f(p) \subset D^- f(p)$, where $D^- f(p)$ is the viscosity subdifferential of f at p .

Acknowledgment

The author thanks the anonymous referee for bringing the papers [1,2] to his attention and for his(her) constructive comments which helped to substantially improve the presentation of the paper.

References

- [1] D. Azagra, J. Ferrera, Proximal calculus on Riemannian manifolds, with applications to fixed point theory, preprint, math.DG/0403465 v 1, 2004.
- [2] D. Azagra, J. Ferrera, F. López-Mesas, Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds, *J. Funct. Anal.* 220 (2005) 304–361.
- [3] F.H. Clarke, Yu.S. Ledyev, R.J. Stern, P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer-Verlag, 1998.
- [4] F.H. Clarke, Yu.S. Ledyev, P.R. Wolenski, Proximal analysis and minimization principles, *J. Math. Anal. Appl.* 196 (1995) 722–735.
- [5] F.H. Clarke, R.J. Stern, P.R. Wolenski, Subgradient criteria for monotonicity, the Lipschitz condition, and convexity, *Canad. J. Math.* 45 (1993) 1167–1183.
- [6] M.P. do Carmo, *Riemannian Geometry*, Birkhäuser, 1992.
- [7] Y.E. Nesterov, M.J. Todd, On the Riemannian geometry defined by self-concordant barriers and interior-point methods, *Found. Comput. Math.* 2 (2002) 333–361.

- [8] T. Rapcsák, *Smooth Nonlinear Optimization in R^n* , Kluwer Academic, 1997.
- [9] T. Sakai, *Riemannian Geometry*, Transl. Math. Monogr., vol. 149, Amer. Math. Soc., Providence, RI, 1996.
- [10] C. Udriste, *Convex Functions and Optimization Methods on Riemannian Manifolds*, Math. Appl., vol. 297, Kluwer Academic, 1994.