Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Functional Analysis 220 (2005) 304–361
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds

Daniel Azagra*, Juan Ferrera, Fernando López-Mesas

Departamento de Análisis Matemático, Facultad de Matemáticas, Universidad Complutense, 28040 Madrid, Spain

Received 14 November 2003; accepted 4 October 2004
Communicated by G. Pisier
Available online 8 December 2004

Abstract

We establish some perturbed minimization principles, and we develop a theory of subdifferential calculus, for functions defined on Riemannian manifolds. Then we apply these results to show existence and uniqueness of viscosity solutions to Hamilton–Jacobi equations defined on Riemannian manifolds.

MSC: 49J52; 58E30

Keywords: Subdifferential; Riemannian manifolds

1. Introduction

The aim of this paper is threefold. First, we extend some perturbed minimization results such as the smooth variational principle of Deville, Godefroy and Zizler, and other almost-critical-point-spotting results, such as approximate Rolle's type theorems, to the realm of Riemannian manifolds. Second, we introduce a definition of subdifferential for functions defined on Riemannian manifolds, and we develop a theory of subdifferentiable calculus on such manifolds that allows most of the known

* Corresponding author. Fax: +34 91 3944726.
E-mail addresses: daniel_azagra@mat.ucm.es (D. Azagra), ferrera@mat.ucm.es (J. Ferrera), fLopez_mesas@mat.ucm.es (F. López-Mesas)..

0022-1236/$ - see front matter
doi:10.1016/j.jfa.2004.10.008

applications of subdifferentiability to be extended to Riemannian manifolds. For instance, we show that every convex function on a Riemannian manifold (that is, every function which is convex along geodesics) is everywhere subdifferentiable (on the other hand, every continuous function is superdifferentiable on a dense set, hence convex functions are differentiable on dense subsets of their domains). Third, we also use this theory to prove existence and uniqueness of viscosity solutions to Hamilton–Jacobi equations defined on Riemannian manifolds. Let us introduce some of these results.

It is known that the classic Rolle's theorem fails in infinite-dimensions, that is, in every infinite-dimensional Banach space with a C^1 smooth (Lipschitz) bump function there are C^1 smooth (Lipschitz) functions which vanish outside a bounded open set and yet have a nonzero derivative everywhere inside this set; see [7] and the references therein. In fact, the failure of Rolle's theorem infinite dimensions takes on a much more dramatic form in a recent result of Azagra and Cepedello Boiso [4]: the smooth functions *with no critical points* are dense in the space of continuous functions on every Hilbert manifold (this result may in turn be viewed as a very strong approximate version for infinite-dimensional manifolds of the Morse–Sard theorem). So, when we are given a smooth function on an infinite-dimensional Riemannian manifold we should not expect to be able to find any critical point, whatever the overall shape of this function is, as there might be none. This important difference between finite and infinite dimensions forces us to consider approximate substitutes of Rolle's theorem and the classic minimization principles, looking for the existence of arbitrarily small derivatives (instead of vanishing ones) for every function satisfying (in an approximate manner) the conditions of the classical Rolle's theorem. This is what the papers [5,6] deal with, one in the differentiable case (showing for instance that if a differentiable function oscillates less than 2ε on the boundary of a unit ball then there is a point inside the ball such that the derivative of the function has norm less than or equal to ε), and the other in the subdifferentiable one. More generally, a lot of perturbed minimization (or variational) principles have been studied, perhaps the most remarkable being Ekeland's principle, Borwein–Preiss' principle, and Deville–Godefroy–Zizler's smooth variational principle. See [26,27] and the references therein.

There are many important applications of those variational principles. Therefore, it seems reasonable to look for analog of these perturbed minimizations principles within the theory of Riemannian manifolds. In Section 3, we prove some almost-critical-point spotting results. First, we establish an approximate version of Rolle's theorem which holds for differentiable mappings defined on subsets of arbitrary Riemannian manifolds. Then we give a version of Deville–Godefroy–Zizler smooth variational principle which holds for those complete Riemannian manifold M which are *uniformly bumpable* (meaning that there exist some numbers $R > 1, r > 0$ such that for every point $p \in M$ and every $\delta \in (0, r)$ there exists a function $b : M \to [0, 1]$ such that $b(x) = 0$ if $d(x, p) \geqslant \delta$, $b(p) = 1$, and $\sup_{x\in M} \|db(x)\|_x \leqslant R/\delta$. Of course every Hilbert space is uniformly bumpable, and there are many other examples of uniformly bumpable manifolds: as we will see, every Riemannian manifold which has strictly positive injectivity and convexity radii is uniformly bumpable). For those Riemannian manifolds we show that, for every lower semicontinuous function $f : M \to (-\infty, \infty]$ which is bounded below, there exists a C^1 smooth function $\varphi : M \to \mathbb{R}$, which is arbitrarily small and has an arbitrarily small derivative everywhere, such that $f - \varphi$ attains a strong global minimum at some $p \in M$.

This result leads up to one of the main topics of this paper: subdifferentiability of functions on Riemannian manifolds, since, according to the definition we are going to give of subdifferential, this implies that such f is subdifferentiable at the point p. We will say that a function $f : M \to (-\infty, \infty]$ is subdifferentiable at p provided there exists a C^1 smooth function $\varphi : M \to \mathbb{R}$ such that $f - \varphi$ attains a local minimum at p. The set of the derivatives $d\varphi(p)$ of all such functions φ will be called subdifferential of f at p, a subset of T^*M_p which will be denoted by $D^- f(p)$. Of course, when M is $\mathbb{R}^n$ or a Hilbert space, this definition agrees with the usual one. Apart from being a useful generalization of the theory of subdifferentiability of convex functions, this notion of subdifferentiability plays a fundamental role in the study of Hamilton–Jacobi equations in $\mathbb{R}^n$ and infinite-dimensional Banach spaces. Not only is this concept necessary to understand the notion of *viscosity solution* (introduced by Crandall and Lions, see [11–20]); from many results concerning subdifferentials one can also deduce relatively easy proofs of the existence, uniqueness and regularity of viscosity solutions to Hamilton–Jacobi equations; see, for instance, [21,22,25,38]. We refer to [23,26] for an introduction to subdifferential calculus in Banach spaces and its applications (especially Hamilton–Jacobi equations).

Section 4 is devoted to the study of subdifferentials of functions defined on manifolds. We start by giving other equivalent definitions of subdifferentiability and superdifferentiability, including a local one through charts, which sometimes makes it easy to translate some results already established in the $\mathbb{R}^n$ or the Banach space cases to the setting of Riemannian manifolds. We also show that a function f is differentiable at a point p if and only if f is both subdifferentiable and superdifferentiable at p. Next, we study the elementary properties of this subdifferential with respect to sums, products and composition, including direct and inverse fuzzy rules. We finish this section by establishing two mean value theorems, and showing that lower semicontinuous functions are subdifferentiable on dense subsets of their domains.

In Section 5, we study the links between convexity and (sub)differentiability of functions defined on Riemannian manifolds. Recall that a function $f : M \to \mathbb{R}$ defined on a Riemannian manifold M is said to be convex provided $f \circ \sigma$ is convex, for every geodesic σ. The papers [34–37] provide a very good introduction to convexity on Riemannian manifolds and the geometrical implications of the existence of global convex functions on a Riemannian manifold; for instance it is shown in [34] that every two-dimensional manifold which admits a global convex function which is locally nonconstant must be diffeomorphic to the plane, the cylinder, or the open Möbius strip. Among other things, we show in this section that every convex function defined on a Riemannian manifold is everywhere subdifferentiable, and is differentiable on a dense set (when the manifold is finite-dimensional, the set of points of nondifferentiability has measure zero).

Finally, in Section 6 we study some Hamilton–Jacobi equations defined on Riemannian manifolds (either finite or infinite-dimensional). Examples of Hamilton–Jacobi equations arise naturally in the setting of Riemannian manifolds, see [1] in relation to Lyapounov theory and optimal control. However, we do not know of any work that has studied nonsmooth solutions, in general, or viscosity solutions, in particular, to Hamilton–Jacobi equations defined on Riemannian manifolds. This may be due to the lack of a theory of nonsmooth calculus for functions defined on Riemannian manifolds. Here we will show how the subdifferential calculus and perturbed minimization principles that we develop in the previous sections can be applied to get results on existence and uniqueness of viscosity solutions to equations of the form

$$\begin{cases} u + F(du) = f \\ u \text{ bounded}, \end{cases}$$

where $f : M \to \mathbb{R}$ is a bounded uniformly continuous function, and $F : T^*M \to \mathbb{R}$ is a function defined on the cotangent bundle of M which satisfies a uniform continuity condition. The manifold M must also satisfy that it has positive convexity and injectivity radii (this condition is automatically met by every compact manifold, for instance). We also prove some results about "regularity" (meaning differentiability almost everywhere) of the viscosity solutions to some of these equations. Finally, we study the equation $\|du(x)\|_x = 1$ for all $x \in \Omega$, $u(x) = 0$ for all $x \in \partial\Omega$, where Ω is a bounded open subset of M, and we show that $x \mapsto d(x, \partial\Omega)$ is the unique viscosity solution to this equation (which has no classical solution).

2. Preliminaries and tools

In this section, we recall some definitions and known results about Riemannian manifolds which will be used later on.

We will be dealing with functions defined on Riemannian manifolds (either finite or infinite-dimensional). A Riemannian manifold (M, g) is a C^∞ smooth manifold M modelled on some Hilbert space H (possibly infinite-dimensional), such that for every $p \in M$ we are given a scalar product $g(p) = g_p := \langle \cdot, \cdot \rangle_p$ on the tangent space $TM_p \simeq H$ so that $\|x\|_p = (\langle x, x \rangle_p)^{1/2}$ defines an equivalent norm on TM_p for each $p \in M$, and in such a way that the mapping $p \in M \mapsto g_p \in \mathcal{S}^2(M)$ is a C^∞ section of the bundle $\sigma_2 : \mathcal{S}_2 \to M$ of symmetric bilinear forms.

If a function $f : M \longrightarrow \mathbb{R}$ is differentiable at $p \in M$, the norm of the differential $df(p) \in T^*M_p$ at the point p is defined by

$$\|df(p)\|_p = \sup\{df(p)(v) : v \in TM_p, \|v\|_p \leqslant 1\}.$$

Since $(TM_p, \|\cdot\|_p)$ is a Hilbert space, we have a linear isometric identification between this space and its dual $(T^*M_p, \|\cdot\|_p)$ through the mapping $TM_p \ni x \mapsto f_x = x \in T^*M_p$, where $f_x(y) = \langle x, y \rangle$ for every $y \in TM_p$.

For every piecewise C^1 smooth path $\gamma : [a, b] \to M$ we define its length as

$$L(\gamma) = \int_a^b \left\| \frac{d\gamma}{dt}(s) \right\|_{\gamma(s)} ds.$$

This length depends only on the path $\gamma[a,b]$ itself, and not on the way the point $\gamma(t)$ moves along it: if $h : [0,1] \to [a,b]$ is a continuous monotone function then $L(\gamma \circ h) = L(\gamma)$. We can always assume that a path γ is parameterized by arc length, which means that $\gamma : [0,T] \to M$ satisfies $\|\frac{d\gamma}{dt}(s)\|_{\gamma(s)} = 1$ for all s, and therefore

$$L(\gamma_{|_{[0,r]}}) = \int_0^r \left\| \frac{d\gamma}{dt}(s) \right\|_{\gamma(s)} ds = r$$

for each $r \in [0,T]$. For any two points $p, q \in M$, let us define

$$d(p,q) = \inf\{L(\gamma) : \gamma \text{ is a } C^1 \text{ smooth path joining } p \text{ and } q \text{ in } M\}.$$

Then d is a metric on M (called the g-distance on M) which defines the same topology as the one M naturally has as a manifold. For this metric we define the closed ball of center p and radius $r > 0$ as

$$B_g(p,r) = \{q \in M : d(p,q) \leqslant r\}.$$

Let us recall that in every Riemannian manifold there is a unique natural covariant derivation, namely the Levi–Civita connection (see Theorem 1.8.11 of [39]); following Klingenberg we denote this derivation by $\nabla_X Y$ for any vector fields X, Y on M. We should also recall that a geodesic is a C^∞ smooth path γ whose tangent is parallel along the path γ, that is, γ satisfies the equation $\nabla_{d\gamma(t)/dt} d\gamma(t)/dt = 0$. A geodesic always minimizes the distance between points which are close enough to each other.

Any path γ joining p and q in M such that $L(\gamma) = d(p,q)$ is a geodesic, and it is called a minimal geodesic. In the sequel all geodesic paths will be assumed to be parameterized by arc length, unless otherwise stated.

Theorem 2.1 (*Hopf-Rinow*). *If M is a finite-dimensional Riemannian manifold which is complete and connected, then there is at least one minimal geodesic connecting any two points in M.*

On the other hand, for any given point p, the statement "q can be joined to p by a unique minimal geodesic" holds for almost every $q \in M$; see [42].

As is well known, the Hopf–Rinow theorem fails when M is infinite-dimensional, but Ekeland [29] proved (by using his celebrated variational principle) that, even in infinite dimensions, the set of points that can be joined by a minimal geodesic in M is dense.

Theorem 2.2 (*Ekeland*). *If M is an infinite-dimensional Riemannian manifold which is complete and connected then, for any given point p, the set* $\{q \in M : q$ can be joined to p by a unique minimal geodesic$\}$ *is residual in M.*

The existence theorem for ODEs implies that for every $V \in TM$ there is an open interval $J(V)$ containing 0 and a unique geodesic $\gamma_V : J(V) \to M$ with $d\gamma(0)/dt = V$. This in turn implies that there is an open neighborhood $\tilde{T}M$ of M in TM such that for every $V \in \tilde{T}M$, the geodesic $\gamma_V(t)$ is defined for $|t| < 2$. The exponential mapping $\exp : \tilde{T}M \to M$ is then defined as $\exp(V) = \gamma_V(1)$, and the restriction of exp to a fiber TM_x in $\tilde{T}M$ is denoted by $\exp_x$.

Let us now recall some useful properties of the exponential map. See [39,41], for instance, for a proof of the following theorem.

Theorem 2.3. *For every Riemannian manifold* (M, g) *and every* $x \in M$ *there exists a number* $r > 0$ *and a map* $\exp_x : B(0_x, r) \subset TM_x \to M$ *such that*

(1) $\exp_x : B(0_x, \delta) \to B_M(x, \delta)$ *is a bi-Lipschitz* C^∞ *diffeomorphism, for all* $\delta \in (0, r]$.
(2) $\exp_x$ *takes the segments passing through* 0_x *and contained in* $B(0_x, r) \subset TM_x$ *into geodesic paths in* $B_M(x, r)$.
(3) $d\exp_x(0_x) = \mathrm{id}_{TM_x}$.

In particular, taking into account condition (3), *for every* $C > 1$, *the radius r can be chosen to be small enough so that the mappings* $\exp_x : B(0_x, r) \to B_M(x, r)$ *and* $\exp_x^{-1} : B_M(x, r) \to B(0_x, r)$ *are C-Lipschitz.*

Recall that a Riemannian manifold M is said to be *geodesically complete* provided the maximal interval of definition of every geodesic in M is all of $\mathbb{R}$. This amounts to saying that for every $x \in M$, the exponential map $\exp_x$ is defined on all of the tangent space TM_x (though, of course, $\exp_x$ is not necessarily injective on all of TM_x). It is well known that every complete Riemannian manifold is geodesically complete. In fact we have the following result (see [41, p. 224] for a proof).

Proposition 2.4. *Let* (M, g) *be a Riemannian manifold. Consider the following conditions*:

(1) *M is complete* (*with respect to the g-distance*).
(2) *All geodesics in M are defined on* $\mathbb{R}$.
(3) *For every* $x \in M$, *the exponential map* $\exp_x$ *is defined on all of* TM_x.
(4) *There is some* $x \in M$ *such that the exponential map* $\exp_x$ *is defined on all of* TM_x.

Then, (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4). *Furthermore, if we assume that M is* finite-dimensional, *then all of the four conditions are equivalent to a fifth*:

(5) *Every closed and* d_g*-bounded subset of M is compact*.

Next, let us recall some results about convexity in Riemannian manifolds.

Definition 2.5. We say that a subset U of a Riemannian manifold is *convex* if given $x, y \in U$ there exists a unique geodesic in U joining x to y, and such that the length of the geodesic is $\mathrm{dist}(x, y)$.

Every Riemannian manifold is *locally convex*, in the following sense.

Theorem 2.6 (*Whitehead*). *Let M be a Riemannian manifold. For every $x \in M$, there exists $c > 0$ such that for all r with $0 < r < c$, the open ball $B(x, r) = \exp_x B(0_x, r)$ is convex.*

This theorem gives rise to the notion of *uniformly locally convex* manifold, which will be of interest when discussing smooth variational principles and Hamilton–Jacobi equations on Riemannian manifolds.

Definition 2.7. We say that a Riemannian manifold M is *uniformly locally convex* provided that there exists $c > 0$ such that for every $x \in M$ and every r with $0 < r < c$ the ball $B(x, r) = \exp_x B(0_x, r)$ is convex.

This amounts to saying that the global convexity radius of M (as defined below) is strictly positive.

Definition 2.8. The convexity radius of a point $x \in M$ in a Riemannian manifold M is defined as the supremum in $\overline{\mathbb{R}^+}$ of the numbers $r > 0$ such that the ball $B(x, r)$ is convex. We denote this supremum by $c(M, x)$. We define the global convexity radius of M as $c(M) := \inf\{c(M, x) : x \in M\}$.

Remark 2.9. By Whitehead's theorem we know that $c(x, M) > 0$ for every $x \in M$. On the other hand, the function $x \mapsto c(x, M)$ is continuous on M, see [39, Corollary 1.9.10]. Consequently, if M is compact, then $c(M) > 0$, that is, M is uniformly locally convex.

The notion of injectivity radius of a Riemannian manifold will also play a role in the study of variational principles and Hamilton–Jacobi equations. Let us recall its definition.

Definition 2.10. We define the injectivity radius of a Riemannian manifold M at a point $x \in M$ as the supremum in $\overline{\mathbb{R}^+}$ of the numbers $r > 0$ such that $\exp_x$ is a C^∞ diffeomorphism onto its image when restricted to the ball $B(0_x, r)$. We denote this supremum by $i(M, x)$. The injectivity radius of M is defined by $i(M) := \inf\{i(M, x) : x \in M\}$.

Remark 2.11. For a finite-dimensional manifold M, it can be seen that $i(M, x)$ equals the supremum of the numbers $r > 0$ such that $\exp_x$ is injective when restricted to the ball $B(0_x, r)$, see [39]. However, for infinite-dimensional manifolds it is not quite clear if this is always true.

Remark 2.12. By Theorem 2.3 we know that $i(x, M) > 0$ for every $x \in M$. On the other hand, it is well known that the function $x \mapsto i(x, M)$ is continuous on M [39, Proposition 2.1.10]. Therefore, if M is compact, then $i(M) > 0$.

We will also need to use the parallel translation of vectors along geodesics. Recall that, for a given curve $\gamma : I \to M$, a number $t_0, t_1 \in I$, and a vector $V_0 \in TM_{\gamma(t_0)}$, there exists a unique parallel vector field $V(t)$ along $\gamma(t)$ such that $V(t_0) = V_0$. Moreover, the mapping defined by $V_0 \mapsto V(t)$ is a linear isometry between the tangent spaces $TM_{\gamma(t_0)}$ and $TM_{\gamma(t)}$, for each $t \in I$. We denote this mapping by $P_{t_0}^{t} = P_{t_0,\gamma}^{t}$, and we call it the parallel translation from $TM_{\gamma(t_0)}$ to $TM_{\gamma(t)}$ along the curve γ.

The parallel translation will allow us to measure the length of the "difference" between vectors (or forms) which are in different tangent spaces (or in duals of tangent spaces, that is, fibers of the cotangent bundle), and do so in a natural way. Indeed, let γ be a minimizing geodesic connecting two points $x, y \in M$, say $\gamma(t_0) = x, \gamma(t_1) = y$. Take vectors $V \in TM_x$, $W \in TM_y$. Then we can define the distance between V and W as the number

$$\|W - P_{t_0,\gamma}^{t_1}(V)\|_y = \|V - P_{t_1,\gamma}^{t_0}(W)\|_x$$

(this equality holds because $P_{t_0}^{t_1}$ is a linear isometry between the two tangent spaces, with inverse $P_{t_1}^{t_0}$). Since the spaces T^*M_x and TM_x are isometrically identified by the formula $v = \langle v, \cdot \rangle$, we can obviously use the same method to measure distances between forms $\xi \in T^*M_x$ and $\eta \in T^*M_y$ lying in different fibers of the cotangent bundle.

Finally, let us consider some mean value theorems. The following two results are easily deduced from the mean value theorem for functions of one variable, but it will be convenient to state and prove them for future reference.

Theorem 2.13 (*Mean value theorem*). *Let (M, g) be a Riemannian manifold, and $f : M \to \mathbb{R}$ a Fréchet differentiable mapping. Then, for every pair of points $p, q \in M$ and every minimal geodesic path $\sigma : I \to M$ joining p and q, there exists $t_0 \in I$ such that*

$$f(p) - f(q) = d(p, q)\, df(\sigma(t_0))(\sigma'(t_0));$$

in particular $|f(p) - f(q)| \leqslant \|df(\sigma(t_0))\|_{\sigma(t_0)} d(p, q)$.

Proof. Since σ is a minimal geodesic we may assume $I = [0, d(p, q)]$, $\|\sigma'(t)\|_{\sigma(t)} = 1$ for all $t \in I$, $\sigma(0) = q$, $\sigma(d(p, q)) = p$. Consider the function $h : I \to \mathbb{R}$ defined by $h(t) = f(\sigma(t))$. By applying the mean value theorem to the function h we get a point $t_0 \in I$ such that

$$f(p) - f(q) = h(d(p, q)) - h(0) = h'(t_0)(d(p, q) - 0) = df(\sigma(t_0))(\sigma'(t_0))d(p, q)$$

and, since $\|\sigma'(t_0)\|_{\sigma(t_0)} = 1$ and $|df(\sigma(t_0))(\sigma'(t_0))| \leqslant \|df(\sigma(t_0))\|_{\sigma(t_0)}$, we also get that $|f(p) - f(q)| \leqslant \|df(\sigma(t_0))\|_{\sigma(t_0)} d(p, q)$. $\square$

When the points cannot be joined by a minimal geodesic we have a less accurate but quite useful result which tells us that every function with a bounded derivative is Lipschitz with respect to the g-distance on M. In fact this results holds even for functions which take values in other Riemannian manifolds. For a differentiable function between Riemannian manifolds $f : M \to N$, we define the norm of the derivative $df(p)$ at a point $p \in M$ by

$$\begin{aligned}\|df(p)\|_p &:= \sup\{\|df(p)(v)\|_{f(p)} : h \in TM_p, \|v\|_p \leqslant 1\} \\ &= \sup\{\zeta(df(p)(v)) : v \in TM_p, \zeta \in T^*N_{f(p)}, \|v\|_p = 1 = \|\zeta\|_{f(p)}\}.\end{aligned}$$

Theorem 2.14 (*Mean value inequality*). *Let M, N be Riemannian manifolds, and $f : M \to N$ a Fréchet differentiable mapping. Assume that f has a bounded derivative, say $\|df(x)\|_x \leqslant C$ for every $x \in M$. Then f is C-Lipschitz, that is*

$$d_N(f(p), f(q)) \leqslant C d_M(p, q)$$

for all $p, q \in M$.

Proof. Fix any two points $p, q \in M$. Take any $\varepsilon > 0$. By definition of $d(p, q)$, there exists a C^1 smooth path $\gamma : [0, T] \to M$ with $\gamma(0) = q$, $\gamma(T) = p$, and

$$L(\gamma) \leqslant d_M(p, q) + \frac{\varepsilon}{C};$$

as usual we may assume $\|\gamma'(t)\|_{\gamma(t)} = 1$ for all $t \in [0, T] = [0, L(\gamma)]$. By considering the path $\beta(t) := f(\gamma(t))$, which joins the points $f(p)$ and $f(q)$ in N, and bearing in mind the definitions of $d_N(f(p), f(q))$ and the fact that $\|d\gamma(t)\|_{\gamma(t)} = 1$ for all t, we get

$$\begin{aligned} d_N(f(p), f(q)) \leqslant L(\beta) &= \int_0^T \|d\beta(t)\|_{\beta(t)}\, dt = \int_0^T \|df(\gamma(t))(d\gamma(t))\|_{f(\gamma(t))}\, dt \\ &\leqslant \int_0^T \|df(\gamma(t))\|_{\gamma(t)}\, dt \leqslant \int_0^T C\, dt = CT \\ &\leqslant C(d_M(p, q) + \varepsilon/C) = C d_M(p, q) + \varepsilon. \end{aligned}$$

We have shown that $d_N(f(p), f(q)) \leqslant C d_M(p, q) + \varepsilon$ for every $\varepsilon > 0$, which means that $d_N(f(p), f(q)) \leqslant C d_M(p, q)$. □

In Section 4, we will generalize these mean value theorems for the case of subdifferentiable or superdifferentiable functions defined on Riemannian manifolds.

The preceding mean value theorem has a converse, which is immediate in the case when M and N are Hilbert spaces, but requires some justification in the setting of Riemannian manifolds.

Proposition 2.15. *Let M, N be Riemannian manifolds. If $f : M \to N$ is K-Lipschitz (that is, $d_N(f(x), f(y)) \leqslant K d_M(x, y)$ for all $x, y \in M$), then $\|df(x)\|_x \leqslant K$ for every $x \in M$.*

Proof. Consider first the case when $N = \mathbb{R}$. Suppose that there exists $x_0 \in M$ with $\|df(x_0)\|_{x_0} > K$. Take $v_0 \in TM_{x_0}$ so that $\|v_0\|_{x_0} = 1$ and $df(x_0)(v_0) > K$. Consider the geodesic $\gamma(t) = \exp_{x_0}(tv_0)$ defined for $|t| \leqslant r_0$ with $r_0 > 0$ small enough. Define $F : [-r_0, r_0] \to \mathbb{R}$ by $F(t) = f(\gamma(t))$. We have that $F'(0) = df(x_0)(v_0) > K$. By the definition of $F'(0)$ we can find some $\delta_0 \in (0, r_0)$ such that

$$\frac{F(t) - F(0)}{t} > K \quad \text{if } |t| \leqslant \delta_0.$$

Taking $t_1 = -\delta_0$, $t_2 = \delta_0$ we get $F(t_1) - F(0) < Kt_1$ and $F(t_2) - F(0) > Kt_2$, hence, by summing,

$$F(t_2) - F(t_1) > K(t_2 - t_1).$$

If we set $x_1 = \gamma(t_1)$, $x_2 = \gamma(t_2)$ this means that

$$f(x_2) - f(x_1) > K(t_2 - t_1) = Kd(x_2, x_1),$$

which contradicts the fact that f is K-Lipschitz.

Now let us consider the general case when the target space is a Riemannian manifold N. Suppose that $\|df(x_0)\|_{x_0} > K$ for some $x_0 \in M$. Then there are $\zeta_0 \in T^*N_{f(x_0)}$ and $v_0 \in TM_{x_0}$ with $\|v_0\|_{x_0} = 1 = \|\zeta_0\|_{f(x_0)}$ and such that $K < \|df(x_0)\|_{x_0} = \zeta_0(df(x_0)(v_0))$. Take $s_0 > 0$ and $\varepsilon > 0$ small enough so that $\exp^{-1}_{f(x_0)} : B(f(x_0), s_0) \to B(0_{f(x_0)}, s_0)$ is a $(1+\varepsilon)$-Lipschitz diffeomorphism and $K < (1+\varepsilon)K < \|df(x_0)\|_{x_0}$. Now take $r_0 > 0$ small enough so that $f(B(x_0, r_0)) \subset B(f(x_0), s_0)$, and define the composition

$$g : B(x_0, r_0) \to \mathbb{R}, \quad g(x) = \zeta_0\left(\exp^{-1}_{f(x_0)}(f(x))\right).$$

It is clear that g is $(1+\varepsilon)K$-Lipschitz. But, since $d\exp^{-1}_{f(x_0)}(f(x_0))$ is the identity, we have that

$$\begin{aligned} dg(x_0)(v_0) &= \zeta_0\left(d\exp^{-1}_{f(x_0)}(f(x_0))(df(x_0)(v_0))\right) \\ &= \zeta_0(df(x_0)(v_0)) = \|df(x_0)\|_{x_0} > (1+\varepsilon)K, \end{aligned}$$

and this contradicts the result we have just proved for the case $N = \mathbb{R}$. □

3. Almost-critical-point-spotting results

As said in the introduction, in infinite dimensions one cannot generally hope to find any critical point for a given smooth function, whatever its shape, so one has to make do with almost critical points.

3.1. An approximate Rolle's theorem

We begin with an approximate version of Rolle's theorem which holds in every Riemannian manifold (even though it is infinite-dimensional) and ensures that every differentiable function which has a small oscillation on the boundary of an open set whose closure is complete has an almost critical point.

Theorem 3.1 (*Approximate Rolle's theorem*). *Let (M, g) be a Riemannian manifold, U an open subset of M such that $\overline{U}$ is complete and bounded with respect to the g-distance, and $p_0 \in M$, $R > 0$ be such that $B_g(p_0, R) \subseteq \overline{U}$. Let $f : \overline{U} \longrightarrow \mathbb{R}$ be a continuous function which is differentiable on U. Then*:

(1) *If $\sup f(U) > \sup f(\partial U)$ then, for every $r > 0$ there exists $q \in U$ such that $\|df(q)\|_q \leqslant r$.*
(2) *If $\inf f(U) < \inf f(\partial U)$ then, for every $r > 0$ there exists $q \in U$ such that $\|df(q)\|_q \leqslant r$.*
(3) *If $f(\overline{U}) \subseteq [-\varepsilon, \varepsilon]$ for some $\varepsilon > 0$, then there exists $q \in U$ such that $\|df(q)\|_q \leqslant \varepsilon/R$.*

Corollary 3.2. *Let (M, g) be a complete Riemannian manifold, U a bounded open subset of M, and $p_0 \in M$, $R > 0$ be such that $B_g(p_0, R) \subseteq \overline{U}$, $\varepsilon > 0$. Suppose that $f(\partial U) \subseteq [-\varepsilon, \varepsilon]$. Then there exists some $q \in U$ such that $\|df(q)\|_q \leqslant \varepsilon/R$.*

To prove Theorem 3.1 we begin with a simple lemma.

Lemma 3.3. *Let (M, g) be a Riemannian manifold, and $f : M \longrightarrow \mathbb{R}$ be a differentiable function on M. Suppose that $\|df(p)\|_p > \varepsilon > 0$. Then there exist a number $\delta > 0$ and two C^1 paths $\alpha, \beta : [0, \delta] \to M$, parameterized by arc length, such that*

$$f(\alpha(t)) < f(p) - \varepsilon t \quad \textit{and} \quad f(\beta(t)) > f(p) + \varepsilon t$$

for all $t \in (0, \delta]$.

Proof. Let us show the existence of such a path α (a required path β can be obtained in a similar manner). Since $\|df(p)\|_p > \varepsilon$, there exists $h \in TM_p$ so that $\|h\|_p = 1$ and $df(p)(h) < -\varepsilon$. Then (by the characterization of the tangent space TM_p as the set of derivatives of all smooth paths passing through p) we can choose a C^1 path $\alpha : [0, r] \to M$, parameterized by arc length, such that

$$\frac{d\alpha}{dt}(0) = h \quad \text{and} \quad \alpha(0) = p.$$

Define the function $F : [0, r] \to \mathbb{R}$ by $F(t) = f(\alpha(t))$. We have that

$$F'(s) = df(\alpha(s))\left(\frac{d\alpha}{dt}(s)\right)$$

for all $s \in [0, r]$. In particular, for $s = 0$, we have that $F'(0) = df(p)(h) < -\varepsilon$, and therefore there exists some $\delta > 0$ such that

$$\frac{F(t) - F(0)}{t} < -\varepsilon$$

for all $t \in (0, \delta]$. This means that $f(\alpha(t)) < f(p) - \varepsilon t$ for all $t \in (0, \delta]$. $\square$

We will also make use of the following version of Ekeland's variational principle (see [30] for a proof).

Theorem 3.4 (*Ekeland's variational principle*). *Let X be a complete metric space, and let $f : X \longrightarrow [-\infty, \infty)$ be a proper upper semicontinuous function which is bounded above. Let $\varepsilon > 0$ and $x_0 \in X$ such that $f(x_0) > \sup\{f(x) : x \in X\} - \varepsilon$. Then for every λ with $0 < \lambda < 1$ there exists a point $z \in \mathrm{Dom}(f)$ such that*

(i) $\lambda d(z, x_0) \leqslant f(z) - f(x_0)$
(ii) $d(z, x_0) < \varepsilon/\lambda$
(iii) $\lambda d(x, z) + f(z) > f(x)$ *whenever* $x \neq z$.

3.2. Proof of Theorem 3.1

Case 1: Let $\eta = \sup f(U) - \sup f(\partial U) > 0$. Define $X = (\overline{U}, d_g)$, which is a complete metric space. Let $n > 1$ be large enough so that $\overline{U} \subset B_g(p_0, n)$, and set $\lambda = \min\{\eta/8n, r\} > 0$. Observe that, since the diameter of U is less than or equal to $2n$, we have that $\lambda d(x, y) \leqslant \eta/4$ for all $x, y \in \overline{U}$. Now, according to Ekeland's variational principle 3.4, there exists $q \in \overline{U}$ such that

$$f(y) \leqslant f(q) + \lambda d(y, q) \quad \text{for all } y \in X. \tag{3.1}$$

In fact, it must be $q \in U$: if $q \in \partial U$ then, taking a such that $f(a) \geqslant \sup f(U) - \eta/4$ we would get

$$\sup f(U) - \eta/2 = (\sup f(U) - \eta/4) - \eta/4 \leqslant f(a) - \lambda d(a, q) \leqslant f(q) \leqslant \sup f(\partial U)$$

a contradiction.

We claim that $\|df(q)\|_q \leqslant \lambda \leqslant r$. Indeed, assume that $\|df(q)\|_q > \lambda$. Then, according to Lemma 3.3, there would exist a C^1 path β, parameterized by arc length, such that $\beta(0) = q$ and

$$f(\beta(t)) > f(q) + \lambda t \tag{3.2}$$

for all $t > 0$ small enough. By combining (3.1) and (3.2), we would get that

$$f(q) + \lambda t < f(\beta(t)) \leqslant f(q) + \lambda d(\beta(t), q) \leqslant f(q) + \lambda L(\beta_{|[0,t]}) = f(q) + \lambda t$$

if $t > 0$ is small enough; but this is a contradiction.

Case 2: consider the function $-f$ and apply Case (1).

Case 3: We will consider two situations.

Case 3.1: Suppose that $f(p_0) \neq 0$. We may assume that $f(p_0) < 0$ (the case $f(p_0) > 0$ is analogous). Define $\lambda = \varepsilon/R$. According to Ekeland's variational principle, there exists $q \in \overline{U}$ such that

(i) $d(p_0, q) \leqslant \frac{1}{\lambda}(f(p_0) - f(q)) \leqslant \frac{1}{\lambda}(f(p_0) + \varepsilon) < R$, and
(ii) $f(q) < f(y) + \lambda d(y, q)$ if $y \neq q$.

The first property tells us that $q \in \operatorname{int} B_g(p_0, R) \subseteq U$. And, by using Lemma 3.3 as in Case 1, it is immediately seen that the second property implies that $\|df(q)\|_q \leqslant \lambda = \varepsilon/R$.

Case 3.2: Suppose finally that $f(p_0) = 0$. We may assume that $\|df(p_0)\|_{p_0} > \varepsilon/R$ (otherwise we are done). By Lemma 3.3, there exist $\delta > 0$ and a C^1 path α in U such that

$$f(\alpha(t)) < f(p_0) - \frac{\varepsilon}{R}\, t$$

if $0 < t \leqslant \delta$. Define $x_0 = \alpha(\delta) \in B_g(p_0, \delta)$. We have that

$$f(x_0) < f(p_0) - \frac{\varepsilon}{R}\, \delta = -\frac{\varepsilon}{R}\, \delta < 0.$$

By applying again Ekeland's variational principle with $\lambda = \varepsilon/R$ we get a point $q \in \overline{U}$ such that

(i) $d(q, x_0) \leqslant \frac{f(x_0)+\varepsilon}{\varepsilon} < \frac{-\varepsilon\delta/R+\varepsilon}{\varepsilon/R} = R - \delta$ and
(ii) $f(q) < f(y) + \frac{\varepsilon}{R} d(y, q)$ for all $y \neq q$.

Now, (i) implies that $d(q, p_0) \leqslant d(q, x_0) + d(x_0, p_0) < R - \delta + \delta = R$, that is, $q \in \operatorname{int} B_g(p_0, R) \subseteq U$. And, as above, bearing in mind Lemma 3.3, (ii) implies that $\|df(q)\|_q \leqslant \varepsilon/R$. □

Remark 3.5. If $\overline{U}$ is not complete the result is obviously false: consider for instance $M = (-1, 1) \subset \mathbb{R}$, $U = (0, 1)$, $\partial U = \{0\}$, $f(x) = x$. On the other hand, the estimate ε/R is sharp, as this example shows: $M = \mathbb{R}$, $U = (-1, 1)$, $f(x) = x$, $R = 1$, $p_0 = 0$, $\varepsilon = 1$.

3.3. A smooth variational principle

Now we turn our attention to perturbed minimization principles on Riemannian manifolds. Of course, since every Riemannian manifold is a metric space, Ekeland's variational principle quoted above holds true and is very useful in this setting: every lower semicontinuous function can be perturbed with a function whose shape is that of an almost flat cone in such a way that the difference attains a global minimum. But sometimes, especially when one wants to build a good theory of subdifferentiability, one needs results ensuring that the perturbation of the function is smooth, that is, one needs to replace that cone with a smooth function which is arbitrarily small and has an arbitrarily small Lipschitz constant. This is just what the Deville–Godefroy–Zizler smooth variational principle does in those Banach spaces having C^1 smooth Lipschitz bump functions; see [27].

Unfortunately, the main ideas behind the proof of this variational principle in the case of Banach spaces cannot be transferred to the setting of Riemannian manifolds in full generality. One has to impose some restriction on the structure of the manifold in order that those ideas work. That is why we need the following definition.

Definition 3.6. We will say that a Riemannian manifold M is *uniformly bumpable* provided there exist numbers $R > 1$ (possibly large) and $r > 0$ (small) such that for every $p \in M$, $\delta \in (0, r)$ there exists a C^1 smooth function $b : M \to [0, 1]$ such that:

(1) $b(p) = 1$.
(2) $b(x) = 0$ if $d(x, p) \geqslant \delta$.
(3) $\sup_{x \in M} \|db(x)\|_x \leqslant R/\delta$.

Remark 3.7. It is easy to see that every Riemannian manifold M is *bumpable*, in the sense that for every $p \in M$, $\delta > 0$, there exists a smooth bump function $b : M \to [0, 1]$ with $b(p) = 1$, $b(x) = 0$ for $x \notin B(p, \delta)$, and b is Lipschitz, that is $\sup_{x \in M} \|db(x)\|_x < \infty$. However it is not quite clear which Riemannian manifolds are uniformly bumpable. Of course every Hilbert space is uniformly bumpable, and there are many other natural examples of uniformly bumpable Riemannian manifolds. In fact we do not know of any Riemannian manifold which is not uniformly bumpable.

Open Problem 3.8. Is every Riemannian manifold uniformly bumpable? If not, provide useful characterizations of those Riemannian manifolds which are uniformly bumpable.

The following proposition provides some sufficient conditions for a Riemannian manifold to be uniformly bumpable: it is enough that $\exp_x$ is a diffeomorphism and preserves radial distances when restricted to balls of a fixed radius $r > 0$. This is always true when M is uniformly locally convex and has a strictly positive injectivity radius.

Proposition 3.9. *Let M be a Riemannian manifold. Consider the following six conditions*:

(1) *M is compact*.

(2) *M is finite-dimensional, complete, and has a strictly positive injectivity radius* $i(M)$.
(3) *M is uniformly locally convex and has a strictly positive injectivity radius.*
(4) *There is a constant* $r > 0$ *such that for every* $x \in M$ *the mapping* $\exp_x$ *is defined on* $B(0_x, r) \subset TM_x$ *and provides a* C^∞ *diffeomorphism*

$$\exp_x : B(0_x, r) \to B(x, r)$$

and the distance function is given here by the expression

$$d(y, x) = \| \exp_x^{-1}(y) \|_x \quad for\ all\ \ y \in B(x, r).$$

(5) *There is a constant* $r > 0$ *such that for every* $x \in M$ *the distance function to* x, $y \in M \mapsto d(y, x)$, *is* C^∞ *smooth on the punctured ball* $B(x, r) \setminus \{x\}$.
(6) *M is uniformly bumpable.*

Then (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\iff$ (4) $\Longrightarrow$ (5) $\Longrightarrow$ (6).

Proof. (1) $\Longrightarrow$ (2) is a trivial consequence of Remark 2.12.

(2) $\Longrightarrow$ (3): In [39, Chapter 2], the injectivity radius of a point $x \in M$ is characterized as the distance from x to the cut locus $C(x)$ of x. Hence, for every $r > 0$ with $r < i(M)$ and every $x \in M$ it is clear that $\exp_x$ is a diffeomorphism and preserves radial distances when restricted to balls of a fixed radius $r > 0$ in the tangent space TM_x, and M is uniformly locally convex. See Theorems 2.1.14 and 2.1.12 of [39].

(3) $\Longrightarrow$ (4): Since $i(M) > 0$, we know that there is some $r_1 > 0$ such that $\exp_x$ is a diffeomorphism onto its image when restricted to the ball $B(0_x, r_1)$, for all $x \in M$. The fact that M is uniformly locally convex clearly implies that there is some $r_2 > 0$ such that

$$d(y, x) = \| \exp_x^{-1}(y) \|_x \quad \text{for all}\ y \in B(x, r_2).$$

We may obviously assume that $r_1 = r_2 := r$. In particular $\exp_x$ maps $B(0_x, r)$ onto $B(x, r)$.

(4) $\Longrightarrow$ (3) is obvious.

(4) $\Longrightarrow$ (5) is trivial, since $\exp_x^{-1}$ is a C^∞ diffeomorphism between those balls, $\| \cdot \|_x$ is C^∞ smooth on $TM_x \setminus \{0_x\}$, and $d(y, x) = \| \exp_x^{-1}(y) \|_x$ for all $y \in B(x, r)$.

(5) $\Longrightarrow$ (6): Assume that the distance function $y \mapsto d(y, x)$ is C^∞ smooth on $B(x, r) \setminus \{x\}$. Let $\theta : \mathbb{R} \to [0, 1]$ be a C^∞ smooth Lipschitz function such that $\theta^{-1}(1) = (-\infty, 1/3]$ and $\theta^{-1}(0) = [1, \infty)$. For a given point $x \in M$ and a number $\delta \in (0, r)$, define $b : M \to [0, 1]$ by

$$b(y) = \theta\left(\frac{1}{\delta} d(y, x)\right).$$

Taking into account the fact that the distance function $y \mapsto d(y, x)$ is 1-Lipschitz and therefore the norm of its derivative is everywhere bounded by 1 (see Proposition

2.15), it is easy to check that b satisfies conditions 1-2-3 of Definition 3.6, for a constant $R = \|\theta'\|_\infty > 1$ that only depends on the real function θ, but not on the point $x \in M$. $\square$

Remark 3.10. The condition that M has a strictly positive injectivity radius is not necessary in order that M is uniformly bumpable, as the following example shows. Let M be the surface of $\mathbb{R}^3$ defined by the equation $z = 1/(x^2 + y^2)$, $(x, y) \neq (0, 0)$, with the natural Riemannian structure inherited from $\mathbb{R}^3$. Then $i(M) = 0$, but, as is not difficult to see, M is uniformly bumpable.

The following theorem is the natural extension of the Deville–Godefroy–Zizler smooth variational principle to Riemannian manifolds which are uniformly bumpable. Recall that a function $F : M \to \mathbb{R} \cup \{+\infty\}$ is said to attain a strong minimum at p provided $F(p) = \inf_{x\in M} F(x)$ and $\lim_{n\to\infty} d(p_n, p) = 0$ whenever (p_n) is a minimizing sequence (that is, if $\lim_{n\to\infty} F(p_n) = F(p)$).

Theorem 3.11 (*DGZ smooth variational principle*). *Let (M, g) be a complete Riemannian manifold which is uniformly bumpable, and let $F : M \longrightarrow (-\infty, +\infty]$ be a lower semicontinuous function that is bounded below, $F \not\equiv +\infty$. Then, for every $\delta > 0$ there exists a bounded C^1 smooth function $\varphi : M \longrightarrow \mathbb{R}$ such that*

(1) *$F - \varphi$ attains its strong minimum in M,*
(2) *$\|\varphi\|_\infty := \sup_{p\in M} |\varphi(p)| < \delta$, and $\|d\varphi\|_\infty := \sup_{p\in M} \|d\varphi(p)\|_p < \delta$.*

Remark 3.12. The assumption that M is complete is necessary here, as the following trivial example shows: $M = (-1, 1) \subset \mathbb{R}$, $f(x) = x$.

We will split the proof of Theorem 3.11 into three lemmas. In the sequel $B(x, r)$ denotes the open ball of center x and radius r in the metric space M, and $B(\varphi, r)$ stands for the open ball of center φ and radius r in the Banach space Y.

Lemma 3.13. *Let M be a complete metric space, and $(Y, \|\cdot\|)$ be a Banach space of real-valued bounded and continuous functions on M satisfying the following conditions*:

(1) *$\|\varphi\| \geqslant \|\varphi\|_\infty = \sup\{|\varphi(x)| : x \in M\}$ for every $\varphi \in Y$.*
(2) *There are numbers $C > 1, r > 0$ such that for every $p \in M$, $\varepsilon > 0$ and $\delta \in (0, r)$ there exists a function $b \in Y$ such that $b(p) = \varepsilon$, $\|b\|_Y \leqslant C\varepsilon(1+1/\delta)$, and $b(x) = 0$ if $x \notin B(p, \delta)$.*

Let $f : M \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function which is bounded below and such that $Dom(f) = \{x \in M | f(x) < +\infty\} \neq \emptyset$. Then, the set G of all the functions $\varphi \in Y$ such that $f + \varphi$ attains a strong minimum in M contains a G_δ dense subset of Y.

Proof. Take a number $N \in \mathbb{N}$ such that $N \geqslant 1/r$, and for every $n \in \mathbb{N}$ with $n \geqslant N$, consider the set

$$U_n = \left\{\varphi \in Y \,|\, \exists\, x_0 \in M \,:\, (f+\varphi)(x_0) < \inf\left\{(f+\varphi)(x) | x \in M \backslash B\left(x_0, \frac{1}{n}\right)\right\}\right\}.$$

Let us see that U_n is an open dense subset of Y. Indeed,

- U_n is open. Take $\varphi \in U_n$. By the definition of U_n there exists $x_0 \in M$ such that $(f+\varphi)(x_0) < \inf\{(f+\varphi)(x)|x \in M\backslash B(x_0, \frac{1}{n})\}$. Set $2\rho = \inf\{(f+\varphi)(x)|x \in M\backslash B(x_0, \frac{1}{n})\} - (f+\varphi)(x_0) > 0$. Then, since $\|\cdot\|_Y \geqslant \|\cdot\|_\infty$, we get that $B_Y(\varphi, \rho) \subset B_\infty(\varphi, \rho) \subset U_n$.
- U_n is dense in Y. Take $\varphi \in Y$ and $\varepsilon > 0$. Since $f+\varphi$ is bounded below there exists $x_0 \in M$ such that $(f+\varphi)(x_0) < \inf\{(f+\varphi)(x)|x \in M\} + \varepsilon$. Set now $\delta = 1/n < r$, and use condition (2) to find a function $b \in Y$ such that $b(x_0) = \varepsilon$, $\|b\|_Y \leqslant C(n+1)\varepsilon$, and $b(x) = 0$ for $x \notin B(x_0, \frac{1}{n})$. Then $(f+\varphi)(x_0) - b(x_0) < \inf\{(f+\varphi)(x)|x \in M\}$ and, if we define $h = -b$, we have

$$(f+\varphi+h)(x_0) < \inf\{(f+\varphi)(x)|x \in M\} \leqslant \inf\left\{(f+\varphi)(x)|x \notin B\left(x_0, \frac{1}{n}\right)\right\}.$$

Since $\inf\{(f+\varphi)(x)|x \notin B(x_0, \frac{1}{n})\} = \inf\{(f+\varphi+h)(x)|x \notin B(x_0, \frac{1}{n})\}$, it is obvious that the above inequality implies that $\varphi+h \in U_n$. On the other hand, we have $\|h\|_Y \leqslant C(n+1)\varepsilon$. Since C and n are fixed and ε can be taken to be arbitrarily small, this shows that $\varphi \in \overline{U_n}$, and U_n is dense in Y.

Therefore we can apply Baire's theorem to conclude that the set $G = \bigcap_{n=N}^{\infty} U_n$ is a G_δ dense subset of Y. Now we must show that if $\varphi \in G$ then $f+\varphi$ attains a strong minimum in M. For each $n \geqslant N$, take $x_n \in M$ such that $(f+\varphi)(x_n) < \inf\{(f+\varphi)(x)|x \notin B(x_n, \frac{1}{n})\}$. Clearly, $x_k \in B(x_n, \frac{1}{n})$ if $k \geqslant n$, which implies that $(x_n)_{n=N}^{\infty}$ is a Cauchy sequence in M and therefore converges to some $x_0 \in M$. Since f is lower semicontinuous and $\bigcap_{n=N}^{\infty} B(x_0, 1/n) = \{x_0\}$, we get

$$\begin{aligned}(f+\varphi)(x_0) &\leqslant \liminf (f+\varphi)(x_n) \leqslant \liminf \left[\inf\left\{(f+\varphi)(x)|x \in M \setminus B\left(x_0, \frac{1}{n}\right)\right\}\right] \\ &= \inf\left\{\inf\left\{(f+\varphi)(x)|x \in M \setminus B\left(x_0, \frac{1}{n}\right)\right\} : n \in \mathbb{N}, n \geqslant N\right\} \\ &= \inf\{(f+\varphi)(x)|x \in M \setminus \{x_0\}\},\end{aligned}$$

which means that $f+\varphi$ attains a global minimum at $x_0 \in M$.

Finally, let us check that in fact $f+\varphi$ attains a strong minimum at the point x_0. Suppose $\{y_n\}$ is a sequence in M such that $(f+g)(y_n) \to (f+g)(x_0)$ and (y_n) does not converge to x_0. We may assume $d(y_n, x_0) \geqslant \varepsilon$ for all n. Bearing in mind this inequality and the fact that $x_0 = \lim x_n$, we can take $k \in \mathbb{N}$ such that $d(x_k, y_n) > \frac{1}{k}$

for all n, and therefore

$$(f+\varphi)(x_0) \leqslant (f+\varphi)(x_k) < \inf\left\{(f+\varphi)(x) | x \notin B\left(x_k, \frac{1}{k}\right)\right\} \leqslant (f+\varphi)(y_n)$$

for all n, which contradicts the fact that $(f+\varphi)(y_n) \to (f+\varphi)(x_0)$. □

Lemma 3.14. *Let M be a uniformly bumpable Riemannian manifold. Then there are numbers $C > 1, r > 0$ such that for every $p \in M$, $\varepsilon > 0$ and $\delta \in (0, r)$ there exists a C^1 smooth function $b : M \to [0, \varepsilon]$ such that*:

(1) $b(p) = \varepsilon = \|b\|_\infty := \sup_{x\in M} |b(x)|$.
(2) $\|db\|_\infty := \sup_{x\in M} \|db(x)\|_x \leqslant C\varepsilon/\delta$.
(3) $b(x) = 0$ if $x \notin B(p, \delta)$.

In particular, $\max\{\|b\|_\infty, \|db\|_\infty\} \leqslant C\varepsilon(1 + 1/\delta)$.

Proof. The definition of uniformly bumpable manifold provides such b in the case when $\varepsilon = 1$. If $\varepsilon \neq 1$, it is enough to consider $b_\varepsilon = \varepsilon b$. □

Lemma 3.15. *Let (M, g) be a complete Riemannian manifold. Then the vector space* $Y = \{\varphi : M \to \mathbb{R} \,|\, \varphi \text{is } C^1 \text{ smooth, bounded and Lipschitz}\}$, *endowed with the norm* $\|\varphi\|_Y = \max\{\|\varphi\|_\infty, \|d\varphi\|_\infty\}$, *is a Banach space*.

Proof. It is obvious that $(Y, \|\cdot\|_Y)$ is a normed space. We only have to show that Y is complete. Let (φ_n) be a Cauchy sequence with respect to the norm $\|\cdot\|_Y$. Since the uniform limit of a sequence of continuous mappings between metric spaces is continuous, it is obvious that (φ_n) uniformly converges to a continuous function $\varphi : M \to \mathbb{R}$. Since T^*M_x is a complete normed space for each $x \in M$, it is also clear that $(d\varphi_n)$ converges to a function $\psi : M \to T^*M$ defined by

$$\psi(x) = \lim_{n\to\infty} d\varphi_n(x)$$

(where the limit is taken in TM_x for each $x \in M$). Let us see that $\psi = d\varphi$. Take $p \in M$. From Theorem 2.3 we know that there exists some $r > 0$ (depending on p) such that the exponential mapping is defined on $B(0_p, r) \subset TM_p$ and gives a diffeomorphism $\exp_p : B(0_p, r) \to B(p, r)$ such that the derivatives of $\exp_p$ and its inverse $(\exp_p)^{-1}$ are bounded by 2 on $B(0_p, r)$ and $B(p, r)$ respectively; in particular $\exp_p$ provides a bi-Lipschitz diffeomorphism between these balls. We denote $\widetilde{\varphi}(h) = (\varphi \circ \exp_p)(h)$, for $h \in B(0_p, r)$. We have

$$\left|\frac{\widetilde{\varphi}(h) - \widetilde{\varphi}(0) - \psi(p)(h)}{\|h\|}\right| = \left|\frac{\widetilde{\varphi}(h) - \widetilde{\varphi}(0)}{\|h\|} - \psi(p)(\frac{h}{\|h\|})\right|$$
$$\leqslant \left|\frac{\widetilde{\varphi}(h) - \widetilde{\varphi}(0) - (\widetilde{\varphi}_n(h) - \widetilde{\varphi}_n(0))}{\|h\|}\right|$$

$$+\left|\frac{\widetilde{\varphi}_n(h)-\widetilde{\varphi}_n(0)}{\|h\|}-d\widetilde{\varphi}_n(0)\left(\frac{h}{\|h\|}\right)\right|$$
$$+\left|(d\widetilde{\varphi}_n(0)-\psi(p))\left(\frac{h}{\|h\|}\right)\right|. \qquad (3.3)$$

Let us first consider the expression $|\frac{\widetilde{\varphi}(h)-\widetilde{\varphi}(0)-(\widetilde{\varphi}_n(h)-\widetilde{\varphi}_n(0))}{\|h\|}|$. By applying the mean value inequality theorem we get

$$|\widetilde{\varphi}_m(h)-\widetilde{\varphi}_m(0)-(\widetilde{\varphi}_n(h)-\widetilde{\varphi}_n(0))|\leqslant \sup_{x\in B(0_p,r)} \|d\widetilde{\varphi}_m(x)-d\widetilde{\varphi}_n(x)\|_p\|h\|_p$$
$$\leqslant 2\|d\varphi_m-d\varphi_n\|_\infty\|h\|_p.$$

Since (φ_n) is a Cauchy sequence in Y we deduce that for every $\varepsilon>0$ there exists $n_0\in\mathbb{N}$ such that $|\widetilde{\varphi}_m(h)-\widetilde{\varphi}_m(0)-(\widetilde{\varphi}_n(h)-\widetilde{\varphi}_n(0))|<(\varepsilon/3)\|h\|$ whenever $m,n\geqslant n_0$ so, by letting $m\to\infty$ we get that $|\widetilde{\varphi}(h)-\widetilde{\varphi}(0)-(\widetilde{\varphi}_n(h)-\widetilde{\varphi}_n(0))|<(\varepsilon/3)\|h\|$ if $n\geqslant n_0$.

On the other hand, the term $|(d\widetilde{\varphi}_n(0)-\widetilde{\psi}(p))(\frac{h}{\|h\|})|$ in the right-hand side of inequality (3.3) above is less than $\varepsilon/3$ when n is large enough; we may assume this happens if $n\geqslant n_0$.

Finally, if we fix $n=n_0$, the term $|\frac{\widetilde{\varphi}_{n_0}(h)-\widetilde{\varphi}_{n0}(0)}{\|h\|}-d\widetilde{\varphi}_{n_0}(0)(\frac{h}{\|h\|})|$ can be made to be less than $\varepsilon/3$ if $\|h\|$ is small enough, say $\|h\|\leqslant\delta$.

By combining these estimations we get that, for $n=n_0$, the left side of inequality (3.3) is less than ε if $\|h\|\leqslant\delta$. This shows that $\widetilde{\varphi}$ is differentiable at p, with $d\widetilde{\varphi}(0_p)=\psi(p)$. Hence φ is differentiable at p, with $d\varphi(p)=\psi(p)$.

To conclude that Y is a Banach space it only remains to check that $d\varphi=\psi$ is continuous and bounded. Take $\varepsilon>0$. Since (φ_n) is a Cauchy sequence in Y, there exists $n_0\in\mathbb{N}$ such that $\|d\varphi_n(y)-d\varphi_m(y)\|_y\leqslant\varepsilon$ for all $y\in M$ provided $n,m\geqslant n_0$. By letting $m\to\infty$ we deduce that $\|d\varphi_n(y)-\psi(y)\|_y\leqslant\varepsilon$ for all $y\in M$, if $n\geqslant n_0$. That is, we have

$$\lim_{n\to\infty}\|d\varphi_n-d\varphi\|_\infty=0.$$

In particular, this implies that $\|d\varphi\|_\infty<\infty$, that is, φ is Lipschitz. Now we can show $\psi=d\varphi$ is continuous. Take any $p\in M$. As above, there exists $r>0$ such that $\exp_p : B(0_p,r)\to B(p,r)$ is a 2-Lipschitz diffeomorphism, and so is the inverse $\exp_p^{-1}$. Define $\widetilde{\varphi}=\varphi\circ\exp_p : B(0_p,r)\to\mathbb{R}$. In order to see that $d\varphi$ is continuous at p it is enough to see that $d\widetilde{\varphi}$ is continuous at 0_p. By applying the mean value inequality we have that

$$\|d\widetilde{\varphi}(x)-d\widetilde{\varphi}(0)\|_p$$
$$\leqslant\|d\widetilde{\varphi}(x)-d\widetilde{\varphi}_n(x)\|_p+\|d\widetilde{\varphi}_n(x)-d\widetilde{\varphi}_n(0)\|_p+\|d\widetilde{\varphi}_n(0)-d\widetilde{\varphi}(0)\|_p$$

$$\leqslant 2\|d\varphi - d\varphi_n\|_\infty + \|d\widetilde{\varphi}_n(x) - d\widetilde{\varphi}_n(0)\|_p + 2\|d\varphi - d\varphi_n\|_\infty$$
$$= 4\|d\varphi - d\varphi_n\|_\infty + \|d\widetilde{\varphi}_n(x) - d\widetilde{\varphi}_n(0)\|_p \tag{3.4}$$

for all $n \in \mathbb{N}$, $x \in B(0_p, r) \subset TM_p$. Since $\|d\varphi - d\varphi_n\|_\infty \to 0$ as $n \to \infty$ we can find $n_0 \in \mathbb{N}$ so that

$$\|d\varphi - d\varphi_{n_0}\|_\infty \leqslant \varepsilon/8. \tag{3.5}$$

Finally, since $d\widetilde{\varphi}_{n_0}$ is continuous at 0_p, there exists $\delta \in (0, r)$ such that

$$\|d\widetilde{\varphi}_{n_0}(x) - d\widetilde{\varphi}_{n_0}(0)\|_p \leqslant \frac{\varepsilon}{2} \tag{3.6}$$

if $\|x\|_p \leqslant \delta$. By combining (3.4)–(3.6) we get that $\|d\widetilde{\varphi}(x) - d\widetilde{\varphi}(0)\|_p \leqslant \varepsilon$ if $\|x\|_p \leqslant \delta$. This shows that $d\widetilde{\varphi}$ is continuous at 0_p. □

Now the proof of Theorem 3.11 is an obvious combination of the above Lemmas.

Remark 3.16. It should be noted that Lemma 3.13 is quite a powerful statement from which a lot of other perturbed minimization principles can be obtained. For instance:

(1) When we take $M = X$, a complete metric space, and Y is the space of all the Lipschitz and bounded functions $f : X \to \mathbb{R}$, with the norm

$$\|f\|_Y = \|f\|_\infty + \mathrm{Lip}(f) = \|f\|_\infty + \sup\left\{\frac{|f(x) - f(y)|}{d(x,y)} : x, y \in X, x \neq y\right\}$$

(which satisfies (1) and (2) of Lemma 3.13 with $C = 1$ and any r), then we obtain a statement that is easily seen to imply Ekeland's variational principle.

(2) When we consider $M = X$, a Banach space having a C^1 smooth Lipschitz bump function, and we define Y as the Banach space of C^1 smooth Lipschitz functions $f : X \to \mathbb{R}$, with the norm

$$\|f\|_Y = \|f\|_\infty + \|f'\|_\infty,$$

then we recover the known DGZ smooth variational principle for Banach spaces.

(3) Let $M = X$ be any metric space in which some notion of *differentiability* has been defined, and Y be a Banach space of *differentiable* (whatever this word should mean in this context) and Lipschitz functions $f : X \to \mathbb{R}$, with the norm

$$\|f\|_Y = \|f\|_\infty + \mathrm{Lip}(f).$$

Suppose that X is uniformly bumpable in the sense that Y satisfies (2) of Lemma 3.13. Then we get a perturbed minimization principle with functions which are *differentiable* and Lipschitz.

Open Problem 3.17. Is Theorem 3.11 true if one drops the assumption that M is uniformly bumpable?

4. A notion of viscosity subdifferential for functions defined on Riemannian manifolds

4.1. Definitions and basic properties

Definition 4.1. Let (M, g) be a Riemannian manifold, and $f : M \longrightarrow (-\infty, \infty]$ be a proper function. We will say that f is subdifferentiable at a point $p \in \mathrm{dom}(f) = \{x \in M : f(x) < \infty\}$ provided there exists a C^1 function $\varphi : M \longrightarrow \mathbb{R}$ such that $f - \varphi$ attains a local minimum at the point p. In this case we will say that $\zeta = d\varphi(p) \in (TM_p)^* \simeq H^* = H$ is a subdifferential of f at p. We define the subdifferential set of f at p by

$$D^- f(p) = \{d\varphi(p) : \varphi \in C^1(M, \mathbb{R}),\ f - \varphi \text{ attains a local minimum at } p\}$$

a subset of T^*M_p. Similarly, we define

$$D^+ f(p) = \{d\varphi(p) : \varphi \in C^1(M, \mathbb{R}),\ f - \varphi \text{ attains a local maximum at } p\}$$

and we say that f is superdifferentiable at p provided $D^+ f(p) \neq \emptyset$.

For every $\zeta \in D^- f(p) \cup D^+ f(p)$, the norm of ζ is defined as

$$\|\zeta\|_p = \sup\{|\zeta(h)| : h \in TM_p, \|h\|_p = 1\}.$$

Remark 4.2. The following properties are obvious from the definition:

(1) f is subdifferentiable at p if and only if $-f$ is superdifferentiable at p, and

$$D^+(-f)(p) = -D^- f(p).$$

(2) If f has a local minimum at p then $0 \in D^- f(p)$.
(3) If h has a local maximum at p then $0 \in D^+ f(p)$.

Next, we give other useful equivalent definitions of subdifferentiability.

Theorem 4.3 (*Characterizations of subdifferentiability*). *Let $f : M \to (-\infty, \infty]$ be a function defined on a Riemannian manifold, $p \in M$, and $\eta \in T^*M_p$. The following statements are equivalent*:

(1) *$\eta \in D^- f(p)$, that is, there exists a C^1 smooth function $\varphi : M \to \mathbb{R}$ so that $f - \varphi$ attains a local minimum at p, and $\eta = d\varphi(p)$.*
(2) *There exists a function $\varphi : M \to \mathbb{R}$ so that $f - \varphi$ attains a local minimum at p, φ is Fréchet differentiable at p, and $\eta = d\varphi(p)$.*
(3) *For every chart $h : U \subset M \to H$ with $p \in U$, if we take $\zeta = \eta \circ dh^{-1}(h(p))$ then we have that*

$$\liminf_{v \to 0} \frac{(f \circ h^{-1})(h(p) + v) - f(p) - \langle \zeta, v \rangle}{\|v\|} \geqslant 0.$$

(4) *There exists a chart $h : U \subset M \to H$ with $p \in U$ and such that, for $\zeta = \eta \circ dh^{-1}(h(p))$, we have*

$$\liminf_{v \to 0} \frac{(f \circ h^{-1})(h(p) + v) - f(p) - \langle \zeta, v \rangle}{\|v\|} \geqslant 0.$$

Moreover, *if the function f is locally bounded below* (*that is, for every $x \in M$ there is a neighborhood U of x such that f is bounded below on U*), *then the above conditions are also equivalent to the following one*:

(5) *There exists a C^1 smooth function $\varphi : M \to \mathbb{R}$ so that $f - \varphi$ attains a* global *minimum at p, and $\eta = d\varphi(p)$.*

Consequently, any of these statements can be taken as a definition of $\eta \in D^- f(p)$. *Analogous statements are equivalent in the case of a superdifferentiable function; in particular $\zeta \in D^+ f(p)$ if and only if there exists a chart $h : U \subset M \to H$ with $p \in U$ and such that, for $\zeta = \eta \circ dh^{-1}(h(p))$,*

$$\limsup_{v \to 0} \frac{(f \circ h^{-1})(h(p) + v) - f(p) - \langle \zeta, v \rangle}{\|v\|} \leqslant 0.$$

Proof. (1) $\Longrightarrow$ (2) and (3) $\Longrightarrow$ (4) are obvious.

(2) $\Longrightarrow$ (3): If $f - \varphi$ has a local minimum at p then $g := f \circ h^{-1} - \varphi \circ h^{-1}$ has also a local minimum at $h(p)$, which implies

$$\liminf_{v \to 0} \frac{g(h(p) + v) - g(h(p))}{\|v\|} \geqslant 0$$

and, by combining this inequality with the fact that

$$\lim_{v\to 0} \frac{(\varphi \circ h^{-1})(h(p)+v) - (\varphi \circ h^{-1})(h(p)) - \langle \zeta, v\rangle}{\|v\|} = 0$$

(because $\zeta = d(\varphi \circ h^{-1})(h(p))$), it is easily deduced that

$$\liminf_{v\to 0} \frac{(f \circ h^{-1})(h(p)+v) - (f \circ h^{-1})(h(p)) - \langle \zeta, v\rangle}{\|v\|} \geqslant 0.$$

(4) $\Longrightarrow$ (1). In order to prove this we will use the following lemma, which is shown in [27] in a more general situation.

Lemma 4.4. *If V is an open set of a Hilbert space H, $x \in V$, and $F : V \to (-\infty, \infty]$ is a function satisfying*

$$\liminf_{v\to 0} \frac{F(x+v) - F(x) - \langle \tau, v\rangle}{\|v\|} \geqslant 0$$

for some $\tau \in H^*$, then there exists a C^1 smooth function $\psi : H \to \mathbb{R}$ such that $F - \psi$ has a local minimum at x, and $d\psi(x) = \tau$.

Take an open neighborhood V of p so that $\overline{V} \subset U$. Note that $F := f \circ h^{-1}$ is a function from the open subset $h(U)$ of the Hilbert space H into $(-\infty, \infty]$, and by the hypothesis we have that

$$\liminf_{v\to 0} \frac{F(h(p)+v) - F(h(p)) - \langle \zeta, v\rangle}{\|v\|} \geqslant 0.$$

By the preceding lemma, there exists a C^1 smooth function $\psi : h(U) \to \mathbb{R}$ such that $F - \psi$ has a local minimum at $h(p)$ and $\zeta = d\psi(h(p))$. Let us define $\phi := \psi \circ h : U \to \mathbb{R}$, which is a C^1 smooth function. It is clear that $F \circ h - \psi \circ h = f - \phi$ has a local minimum at p, and $d\phi(p) = d\psi(h(p)) \circ dh(p) = \zeta \circ dh(p) = \eta$. In order to finish the proof it is enough to extend ϕ to the complement of V by defining $\varphi = \theta\phi$, where θ is a C^1 smooth Uryshon-type function which is valued 1 on the set V and 0 outside U (such a function certainly exists because M has C^∞ smooth partitions of unity and $\overline{V} \subset U$). It is obvious that φ keeps the relevant properties of ϕ.

Finally, let us see that, when f is locally bounded below, (1) $\iff$ (5). Obviously, (5) $\Longrightarrow$ (1). To see that (1) $\Longrightarrow$ (5), let us assume that there exists a C^1 smooth function $\psi : M \to \mathbb{R}$ and some $r > 0$ such that $0 = f(p) - \psi(p) \leqslant f(x) - \psi(x)$ if $x \in B(p, r)$, and denote $\eta = d\psi(p)$. We have to see that there exists a C^1 smooth function $\varphi : M \to \mathbb{R}$ such that $f - \varphi$ attains a *global* minimum at p and $d\varphi(p) = \eta$.

Consider the open set $U = M \setminus \overline{B}(p, r/2)$. Since $f - \psi$ is locally bounded below, for each $x \in U$ there exist $\delta_x > 0$ and $m_x \in \mathbb{R}$ such that $B(x, \delta_x) \subset U$ and $m_x \leqslant f(y) - \psi(y)$ for all $y \in B(x, \delta_x)$. Consider the open covering

$$G := \{B(x, \delta_x) : x \in U\} \cup \{B(p, r)\}$$

of M. Since M has C^∞ smooth partitions of unity there exists a locally finite refinement $\{U_i\}_{i\in I}$ of the covering G and a family of functions $\{\psi_i\}_{i\in I} \subset C^\infty(M, [0, 1])$ so that $\mathrm{supp}(\psi_i) \subset U_i$ for each i and $\sum_{i\in I} \psi_i = 1$.

For each $i \in I$, if $U_i \subset B(p, r)$ then we define $\alpha_i = 0$. Otherwise we can pick some $x_i \in U = M \setminus \overline{B}(p, r/2)$ such that $U_i \subset B(x_i, \delta_{x_i})$, and in this case we define $\alpha_i = m_{x_i}$. Now we can define our function $\varphi : M \to \mathbb{R}$ by

$$\varphi(x) = \psi(x) + \sum_{i\in I} \alpha_i \psi_i(x).$$

It is clear that φ is a C^1 smooth function such that $\varphi = \psi$ on $\overline{B}(p, r/2)$ (indeed, take $x \in \overline{B}(p, r/2)$; if $x \in U_i$ then $U_i \subset B(p, r)$ because of the choice of the covering G and the δ_y, so $\alpha_i = 0$, while for all the rest of $j \in I$ we have $\psi_j(x) = 0$; therefore $\varphi(x) = \psi(x) + 0 = \psi(x)$). In particular, it follows that $\eta = d\psi(p) = d\varphi(p)$.

We claim that $f - \varphi$ attains a global minimum at p. Indeed, fix $x \in M$. If $x \in \overline{B}(p, r/2) = M \setminus U$ then, as we have just seen, $\varphi(x) = \psi(x)$, and $0 = (f - \varphi)(p) \leqslant (f - \psi)(x) = (f - \varphi)(x)$. If $x \in U$ then, for those $i \in I$ such that $x \in U_i$ we have $(f - \psi)(x) \geqslant m_{x_i} = \alpha_i$, while $\psi_j(x) = 0$ for those $j \in I$ with $x \notin U_j$. Therefore,

$$\begin{aligned} f(x) - \varphi(x) &= f(x) - \psi(x) - \sum_{i\in I} \alpha_i \psi_i(x) \\ &= f(x) - \psi(x) - \sum \{\alpha_i \psi_i(x) : i \in I, x \in U_i\} \\ &\geqslant \sup\{\alpha_i : i \in I, x \in U_i\} - \sum \{\alpha_i \psi_i(x) : i \in I, x \in U_i\} \\ &\geqslant 0 = f(p) - \varphi(p), \end{aligned}$$

and $f - \varphi$ has a global minimum at p. □

Corollary 4.5. *Let $f : M \to (-\infty, \infty]$ be a function defined on a Riemannian manifold, and let $h : U \subset M \to h(U) \subset H$ be a chart of M. Then,*

$$\begin{aligned} D^- f(p) &= \left\{ \zeta \circ dh(p) : \zeta \in H^*, \liminf_{v\to 0} \frac{(f \circ h^{-1})(h(p) + v) - f(p) - \langle \zeta, v\rangle}{\|v\|} \geqslant 0 \right\} \\ &= \{\zeta \circ dh(p) : \zeta \in D^-(f \circ h^{-1})(h(p))\}. \end{aligned}$$

Now we can show that subdifferentiable plus superdifferentiable equals differentiable.

Proposition 4.6. *A function f is differentiable at p if and only if f is both subdifferentiable and superdifferentiable at p. In this case,* $\{df(p)\} = D^- f(p) = D^+ f(p)$.

Proof. Assume first that f is both subdifferentiable and superdifferentiable at p. Then there exist C^1 functions $\varphi, \psi : M \to \mathbb{R}$ such that $f - \varphi$ and $f - \psi$ have a local minimum and a local maximum at p, respectively. We can obviously assume $f(p) = \varphi(p) = \psi(p)$. Then these conditions mean that $f(x) - \varphi(x) \geqslant 0$ and $f(x) - \psi(x) \leqslant 0$ for all $x \in U$, where U is an open neighborhood of p. On the other hand, $(f - \varphi) - (f - \psi) = \psi - \varphi$ has a local minimum at p, hence $0 = d(\psi - \varphi)(p) = d\psi(p) - d\varphi(p)$. That is, we have that

$$\varphi(x) \leqslant f(x) \leqslant \psi(x) \quad \text{for all } x \in U, \quad \varphi(p) = \psi(p) = f(p),$$
$$\text{and} \quad d\varphi(p) = d\psi(p).$$

By using charts, it is an easy exercise to check that these conditions imply that f is differentiable at p, with $df(p) = d\psi(p) = d\varphi(p)$; in particular this argument shows that $\{df(p)\} = D^- f(p) = D^+ f(p)$.

Now, if f is differentiable at p then, by the chain rule, so is $f \circ h^{-1}$ at $h(p)$ for any chart $h : U \subset M \to H$; in particular, putting $\zeta = d(f \circ h^{-1})(h(p))$, we have

$$\lim_{v \to 0} \frac{(f \circ h^{-1})(h(p) + v) - f(p) - \langle \zeta, v \rangle}{\|v\|} = 0,$$

which, thanks to Theorem 4.3, yields $df(p) = \zeta \circ dh(p) \in D^- f(p) \cap D^+ f(p)$. □

What the above proof really shows is the (not completely obvious) following result: a function f is differentiable at a point p if and only if its graph is trapped between the graphs of two C^1 smooth functions which have the same derivative at p and touch the graph of f at p.

Corollary 4.7 (*Criterion for differentiability*). *A function* $f : M \to \mathbb{R}$ *is Fréchet differentiable at a point p if and only if there are* C^1 *smooth functions* $\varphi, \psi : M \to \mathbb{R}$ *such that*

$$\varphi(x) \leqslant f(x) \leqslant \psi(x) \quad \textit{for all } x \in M, \ \varphi(p) = \psi(p) = f(p),$$
$$\textit{and} \quad d\varphi(p) = d\psi(p).$$

Let us say a few words about the relationship between subdifferentiability and continuity. In general, a subdifferentiable function need not be continuous. For instance, the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = 0$ if $x \in [0, 1]$, and 1 elsewhere, is Fréchet subdifferentiable everywhere in $\mathbb{R}$, and yet f is not continuous at 0 and 1. However, it is easy to see that subdifferentiability implies lower semicontinuity.

Proposition 4.8. *If f is subdifferentiable at p then f is lower semicontinuous at p. In the same way, superdifferentiability implies upper semicontinuity.*

Proof. The result is immediate in the case of a function $g : V \subset H \to (-\infty, \infty]$; indeed, if

$$\liminf_{v\to 0} \frac{g(x+v) - g(x) - \langle \tau, v\rangle}{\|v\|} \geqslant 0$$

then $\liminf_{y\to x} g(y) \geqslant g(x)$. The general case follows by applying Theorem 4.3. □

4.2. Some rules and fuzzy rules

Next, we study some properties of the subdifferentials related to composition, sum and product of subdifferentiable and differentiable functions. Of course, all the statements hold for superdifferentials as well, with obvious modifications.

Proposition 4.9 (*Chain rule*). *Let M, N be Riemannian manifolds, $g : M \to N$, and $f : N \to (-\infty, \infty]$. Assume that the function f is subdifferentiable at $g(p)$, and that g is Fréchet differentiable at p. Then the composition $f \circ g : M \to (-\infty, \infty]$ is subdifferentiable at p, and*

$$\{\zeta \circ dg(p) : \zeta \in D^- f(g(p))\} \subseteq D^-(f \circ g)(p).$$

Proof. Take $\zeta \in D^- f(g(p))$, then there exists a function $\varphi : N \to \mathbb{R}$ so that $f - \varphi$ has a local minimum at $g(p)$, φ is Fréchet differentiable at $g(p)$, and $\zeta = d\varphi(g(p))$. In particular there exists $\varepsilon > 0$ such that $f(y) - \varphi(y) \geqslant f(g(p)) - \varphi(g(p))$ whenever $d(y, g(p)) < \varepsilon$. Define $\psi = \varphi \circ g$. Since g is differentiable at p and φ is differentiable at $g(p)$, by the chain rule it follows that ψ is a function from M into $\mathbb{R}$ which is Fréchet differentiable at p, with $d\psi(p) = d\varphi(g(p)) \circ dg(p)$. Since g is continuous at p, there exists $\delta > 0$ such that $d(g(x), g(p)) < \varepsilon$ for all x with $d(x, p) < \delta$. Then we get $f(g(x)) - \varphi(g(x)) \geqslant f(g(p)) - \varphi(g(p))$ if $d(x, p) < \delta$, that is, $f \circ g - \psi$ has a local minimum at p. By Theorem 4.3 [(1) $\Longleftrightarrow$ (2)], this ensures that $f \circ g$ is subdifferentiable at p, with $\zeta \circ dg(p) = d\varphi(g(p)) \circ dg(p) = d\psi(p) \in D^- (f \circ g)(p)$. □

The following example shows that the inclusion provided by Proposition 4.9 is strict, in general.

Example 4.10. Let $M = N = \mathbb{R}$, $g(x) = |x|^{3/2}$, $f(y) = |y|^{1/2}$; $f \circ g(x) = |x|^{3/4}$. Then g is C^1 smooth on $\mathbb{R}$, and we have $dg(0) = 0$, $D^- f(g(0)) = D^- f(0) = (-\infty, \infty)$, $D^-(f \circ g)(0) = (-\infty, \infty)$. Therefore $\zeta \circ dg(0) = 0$ for every $\zeta \in D^- f(g(0))$.

Corollary 4.11. *Let M, N be Riemannian manifolds, $h : M \to N$ a C^1 diffeomorphism. Then, $f : M \to (-\infty, \infty]$ is subdifferentiable at p if and only if $f \circ h^{-1}$ is* *subdifferentiable at* $h(p)$, *and*

$$D^- f(p) = \{\zeta \circ dh(p) : \zeta \in D^-(f \circ h^{-1})(h(p))\}.$$

Proof. If $f : M \to (-\infty, \infty]$ is subdifferentiable at p then, by the preceding Proposition, $f \circ h^{-1} : N \to (-\infty, \infty]$ is subdifferentiable at $h(p) \in N$ and, moreover, we know that if $T \in D^- f(p)$ then $\zeta := T \circ dh^{-1}(h(p)) \in D^-(f \circ h^{-1})(h(p))$. Then $T = \zeta \circ dh(p)$, with $\zeta \in D^-(f \circ h^{-1})(h(p))$.

Conversely, if $f \circ h^{-1}$ is subdifferentiable at $h(p)$ then, again by the preceding result, $f = (f \circ h^{-1}) \circ h$ is subdifferentiable at p and, for any $\zeta \in D^-(f \circ h^{-1})(h(p))$, we have $\zeta \circ dh(p) \in D^{-1}\big((f \circ h^{-1}) \circ h\big)(p) = D^- f(p)$. $\square$

Proposition 4.12 (*Sum rule*). *For all functions* $f_1, f_2 : M \longrightarrow (-\infty, \infty]$, $p \in M$, *we have*

$$D^- f_1(p) + D^- f_2(p) \subseteq D^-(f_1 + f_2)(p)$$

and analogous inclusions hold for superdifferentials.

Proof. Take $\zeta_i \in D^- f_i(p)$, $i = 1, 2$. There are C^1 smooth functions $\varphi_i : M \to \mathbb{R}$ such that $f_i - \varphi_i$ have a minimum at p and $\zeta_i = d\varphi_i(p)$ for $i = 1, 2$. Then $(f_1 + f_2) - (\varphi_1 + \varphi_2) = (f_1 - \varphi_1) + (f_2 - \varphi_2)$ clearly has a minimum at p, hence $\zeta_1 + \zeta_2 = d(\varphi_1 + \varphi_2)(p)$ belongs to $D^-(f_1 + f_2)(p)$. $\square$

When one of the functions involved in the sum is uniformly continuous the inclusion provided by this statement can be reversed in a fuzzy way. This assumption is necessary in general, as a counterexample (in the Hilbert space) of Deville and Ivanov shows; see [28].

Theorem 4.13 (*Fuzzy rule for the subdifferential of the sum*). *Let* (M, g) *be a Riemannian manifold. Let* $f_1, f_2 : M \longrightarrow \mathbb{R}$ *be such that* f_1 *is lower semicontinuous and* f_2 *is uniformly continuous. Take* $p \in M$, *a chart* (U, φ) *with* $p \in U$, $\zeta \in D^-(f_1 + f_2)(p)$, $\varepsilon > 0$, *and a neighborhood* V *of* (p, ζ) *in the cotangent bundle* T^*M. *Then there exist* $p_1, p_2 \in U$, $\zeta_1 \in D^- f_1(p_1)$, $\zeta_2 \in D^- f_2(p_2)$ *such that*: $(p_i, \zeta_1 \circ d\varphi(p_1)^{-1} \circ d\varphi(p_i) + \zeta_2 \circ d\varphi(p_2)^{-1} \circ d\varphi(p_i)) \in V$ *for* $i = 1, 2$; *and* $|f_i(p_i) - f_i(p)| < \varepsilon$ *for* $i = 1, 2$.

Proof. Fix a chart (U, φ) such that $p \in U$ and T^*U is diffeomorphic to $U \times H^*$ through the canonical diffeomorphism $L : T^*U \to U \times H^*$ defined by $L(q, \xi) = (q, \xi \circ d\varphi(q)^{-1})$. The theorem can be reformulated as follows: *for every* $p \in U$, $\zeta \in D^-(f_1 + f_2)(p)$, *and* $\varepsilon > 0$, *there exist* $p_1, p_2 \in U$, $\zeta_1 \in D^- f_1(p_1)$, $\zeta_2 \in D^- f_2(p_2)$ *such that:* $d(p_1, p_2) < \varepsilon$, $||\zeta_1 \circ d\varphi(p_1)^{-1} + \zeta_2 \circ d\varphi(p_2)^{-1} - \zeta \circ d\varphi(p)^{-1}|| < \varepsilon$, *and* $|f_i(p_i) - f_i(p)| < \varepsilon$ *for* $i = 1, 2$. But this statement follows immediately from Deville and El Haddad's fuzzy rule for Banach spaces [24] applied to the functions $f_1 \circ \varphi^{-1}$ and $f_2 \circ \varphi^{-1}$. $\square$

Proposition 4.14 (*Product rule*). *Suppose* $f_1, f_2 : M \longrightarrow [0, \infty)$ *are functions subdifferentiable at* $p \in M$. *Then* $f_1 f_2$ *is subdifferentiable at p, and*

$$f_1(p)D^- f_2(p) + f_2(p)D^- f_1(p) \subseteq D^-(f_1 f_2)(p).$$

Proof. If $f_1(p) = f_2(p) = 0$ the result is obvious, so we may assume, for instance, that $f_1(p) > 0$.

Pick $\zeta_i \in D^- f_i(p)$, and find C^1 smooth functions $\varphi_i : M \to \mathbb{R}$ such that $f_i - \varphi_i$ have a local minimum at p and $\zeta_i = d\varphi_i(p)$ for $i = 1, 2$. As usual we may assume that $\varphi_i(p) = f_i(p)$, so that $f_i - \varphi_i \geqslant 0$. Since $\varphi_1(p) = f_1(p) > 0$ and φ_1 is continuous, there exists a neighborhood V of p such that $\varphi_1 \geqslant 0$ on V. We may assume that V is small enough so that the restriction of $\varphi_1 - f_1$ to V has a global minimum at p. Then we deduce that $f_1 f_2 \geqslant \varphi_1 f_2 \geqslant \varphi_1 \varphi_2$ on V, that is,

$$(f_1 f_2 - \varphi_1 \varphi_2)(x) \geqslant 0 = (f_1 f_2 - \varphi_1 \varphi_2)(p) \quad \text{for all } x \in V,$$

which means that $f_1 f_2 - \varphi_1 \varphi_2$ has a local minimum at p, and therefore

$$\begin{aligned} f_1(p)\zeta_2 + f_2(p)\zeta_1 &= \varphi_1(p)d\varphi_2(p) + \varphi_2(p)d\varphi_1(p) \\ &= d(\varphi_1\varphi_2)(p) \in D^-(f_1 f_2)(p). \qquad \square \end{aligned}$$

Remark 4.15. If the functions are not positive, the result is not necessarily true, as the following example shows: $M = \mathbb{R}$, $f_1(x) = |x|$, $f_2(x) = -1$, $p = 0$ (note that the function $(f_1 f_2)(x) = -|x|$ is not subdifferentiable at 0).

4.3. Topological and geometrical properties of the subdifferential sets

Proposition 4.16. $D^- f(p)$ *and* $D^+ f(p)$ *are closed and convex subsets of* T^*M_p. *In particular, if f is locally Lipschitz then these sets are* w^**-compact as well.*

Proof. Let us first check that $D^- f(p)$ is convex. Pick $\zeta_1, \zeta_2 \in D^- f(p)$, and find C^1 smooth functions $\varphi_1, \varphi_2 : M \to \mathbb{R}$ such that $d\varphi_i(p) = \zeta_i$, and $(f - \varphi_i)(x) \geqslant 0 = (f - \varphi_i)(p)$ for all x in a neighborhood of p. Take $t \in [0, 1]$, and define the function $\varphi_t : M \to \mathbb{R}$ by $\varphi_t(x) = (1 - t)\varphi_1(x) + t\varphi_2(x)$. It is immediately seen that φ_t is a C^1 smooth function such that $f - \varphi_t$ attains a local minimum at p, and therefore

$$(1 - t)\zeta_1 + t\zeta_2 = d\varphi_t(p) \in D^- f(p).$$

Now let us see that $D^- f(p)$ is closed. Take a chart $h : U \subset M \to H$ with $p \in U$. Since $dh(p) : TM_p \to H$ is a linear isomorphism and $(dh(p))^* : H^* \to (TM_p)^*$ (defined by $(dh(p))^*(\zeta) = \zeta \circ dh(p)$) is a linear isomorphism as well, and, by

Corollary 4.5, we know that

$$D^- f(p) = \{\zeta \circ dh(p) : \zeta \in D^-(f \circ h^{-1})(h(p))\} = (dh(p))^* \left(D^-(f \circ h^{-1})(h(p))\right)$$

it is enough to show that $D^-(f \circ h^{-1})(h(p))$ is closed in H^*. That is, we have to show the result in the case of a function $g : V \subset H \to (-\infty, \infty]$ which is subdifferentiable at a point x. So let us prove that $D^- g(x)$ is closed in $(H^*, \| \cdot \|)$. Let $(p_n) \subset D^- g(x)$ be such that $\|p_n - p\| \to 0$, and let us check that $p \in D^- g(x)$. We have

$$\liminf_{v \to 0} \frac{g(x+v) - g(x) - \langle p_n, v \rangle}{\|v\|} \geqslant 0$$

for all n, and therefore

$$\begin{aligned}
&\liminf_{v \to 0} \frac{g(x+v) - g(x) - \langle p, v \rangle}{\|v\|} \\
&\quad = \liminf_{v \to 0} \left[\frac{1}{\|v\|} (g(x+v) - g(x) - \langle p_n, v \rangle) + \frac{1}{\|v\|} \langle p_n - p, v \rangle \right] \\
&\quad \geqslant \liminf_{v \to 0} \frac{1}{\|v\|} (g(x+v) - g(x) - \langle p_n, v \rangle) + \liminf_{v \to 0} \frac{1}{\|v\|} \langle p_n - p, v \rangle \\
&\quad \geqslant 0 + \liminf_{v \to 0} \frac{1}{\|v\|} \langle p_n - p, v \rangle = -\|p_n - p\|
\end{aligned}$$

for all n, that is,

$$\liminf_{v \to 0} \frac{g(x+v) - g(x) - \langle p, v \rangle}{\|v\|} \geqslant -\|p_n - p\|$$

for all $n \in \mathbb{N}$, and since $\|p_n - p\| \to 0$ we deduce that

$$\liminf_{v \to 0} \frac{g(x+v) - g(x) - \langle p, v \rangle}{\|v\|} \geqslant 0,$$

which means $p \in D^- g(x)$.

Finally, when f is locally Lipschitz, by composing with the inverse of the exponential map (which provides a Lipschitz chart on a neighborhood of each point) and using Corollary 4.5, it is easily seen that $D^- f(p)$ and $D^+ f(p)$ are bounded. Then, by the Alaoglu–Bourbaki theorem it follows that these sets are w^*-compact. $\square$

4.4. Density of the points of subdifferentiability

As a consequence of the smooth variational principle, every lower semicontinuous function is subdifferentiable on a dense subset of its domain.

Proposition 4.17. *Let M be a Riemannian manifold. If* $f : M \longrightarrow (-\infty, \infty]$ *is lower semicontinuous and proper then* $\{p \in \mathrm{dom}(f) : D^- f(p) \neq \emptyset\}$ *is dense in* $\mathrm{dom(f)} := \{x \in M : f(x) < \infty\}$.

Proof. Assume first that M is complete and uniformly bumpable (such is, for instance, the case when M is a Hilbert space H). In this case we can deduce the result directly by applying the smooth variational principle 3.11 as follows. Pick any point p_0 with $f(p_0) < \infty$, and any open neighborhood U of p_0. We must show that there is some $p \in U$ such that $D^- f(p) \neq \emptyset$. Since M has smooth partitions of unity, there is a C^∞ smooth function $b : M \to [0, \infty)$ such that $b(y) > 0$ if and only if $y \in U$. Consider the function $g : M \to (-\infty, \infty]$ defined by

$$g(y) = \frac{1}{b(y)} \text{ if } y \in U \quad \text{and} \quad g(y) = \infty \text{ if } y \notin U.$$

The function g is lower semicontinuous on M, and C^∞ smooth on U. Then the sum $f + g$ is lower semicontinuous, and $(f + g)(p_0) < +\infty$. According to the smooth variational principle, there exists a C^1 smooth function $\varphi : M \to \mathbb{R}$ such that $(f + g) - \varphi$ attains a strong minimum at some point $p \in M$. In fact we have $p \in U$, because this function is valued $+\infty$ outside U. But, since the function $\varphi - g$ is C^1 smooth on U, and $f - (\varphi - g)$ attains its minimum at p, we conclude that

$$d(\varphi - g)(p) \in D^- f(p) \neq \emptyset.$$

Now let us consider the case when M is not necessarily complete or uniformly bumpable. Pick a point $p_0 \in \mathrm{dom}(f)$ and an open set U containing p_0. We may assume that U is small enough so that there is a chart $h : \overline{U} \subset V \to H$. By Corollary 4.5 we know that, for any $p \in M$, we have $D^- f(p) \neq \emptyset$ if and only if $D^-(f \circ h^{-1})(p) \neq \emptyset$, so it is enough to see that there is some $x \in h(U)$ with $D^-(f \circ h^{-1})(x) \neq \emptyset$. Define $F(x) = f \circ h^{-1}(x)$ if $x \in h(\overline{U})$, and $F(x) = +\infty$ otherwise. The function F is lower semicontinuous on H, and $F = f \circ h^{-1}$ on $h(U)$. Since the Hilbert space H is certainly complete and uniformly bumpable, we can apply the first part of the argument to find some $x \in h(U)$ so that $\emptyset \neq D^- F(x) = D^-(f \circ h^{-1})(x)$. □

4.5. Mean value inequalities

There are many subdifferential mean value inequality theorems for functions defined on Banach spaces. Here, we will only consider two of them, which complement each other. The first one is due to Deville [22] and holds for all lower semicontinuous functions f defined on an open convex set of a Banach space, even if they are not required to be everywhere subdifferentiable, but it demands a bound for *all* of the subgradients of the function at all the points where it is subdifferentiable. The second one is due to Godefroy (who improved a similar previous result of Azagra and Deville), see [5,33], and only demands the existence of *one* subdifferential or superdifferential which is bounded (by the same constant) at each point, but it requires the function to satisfy $D^- f(x) \cup D^+ f(x) \neq \emptyset$ for *all* the points x in the domain of f (an open convex subset of a Banach space).

Next, we extend these mean value inequality theorems to the setting of Riemannian manifolds. The main ideas of the proofs of these results could be adapted to obtain direct proofs which would be valid for the case of manifolds, but for shortness we choose here to deduce them from the Hilbert space case.

Theorem 4.18 (*Deville's mean value inequality*). *Let (M, g) be a Riemannian manifold, and let $f : M \longrightarrow \mathbb{R}$ be a lower semicontinuous function. Assume that there exists a constant $K > 0$ such that $\|\zeta\|_p \leqslant K$ for all $\zeta \in D^- f(p)$ and $p \in M$. Then,*

$$|f(p) - f(q)| \leqslant K d_M(p, q) \quad for\ all\ p, q \in M.$$

Proof. The result is true in the case when $M = H$ is a Hilbert space [22]. For completeness we give a hint of Deville's argument, which is an instructive application of the smooth variational principle. By standard arguments it suffices to show the result locally (see the proof of the general case below). Fix $x_0 \in H$. Since f is locally bounded below there are $N, \delta > 0$ so that $f(x) - f(x_0) \geqslant -N$ whenever $x \in B(x_0, 2\delta)$. For fixed $y \in B(x_0, \delta/4)$, $\varepsilon > 0$, consider the function defined by $F(x) = f(x) - f(y) - \alpha(\|x - y\|)$ for $\|x - y\| \leqslant \delta$, and $F(x) = +\infty$ elsewhere, where $\alpha : [0, \infty) \to [0, \infty)$ is C^1 smooth and satisfies $\alpha(t) = (K + \varepsilon)t$ if $t \leqslant \delta/2$, $\alpha(\delta) \geqslant N$, and $\alpha'(t) \geqslant K + \varepsilon$ for all $t > 0$. If $\inf F < 0$, by applying the Smooth Variational Principle one can get a point $x_1 \in B(y, \delta) \setminus \{y\}$ and a subgradient $\zeta \in D^- f(x_1)$ so that $\|\zeta\| > K$, a contradiction. Hence $F \geqslant 0$, and by letting $\varepsilon \to 0$ the local result follows. See [22] for the details.

Now consider the general case of a Riemannian manifold. Fix any two points $p, q \in M$, and consider a continuous and piecewise C^1 smooth path $\gamma : [0, T] \to M$, parameterized by arc length, with $\gamma(0) = p, \gamma(T) = q$. Take $\varepsilon > 0$. According to Theorem 2.3, for each $x \in \gamma([0, T])$ there exists $r_x > 0$ so that $\exp_x : B(0_x, 2r_x) \subset TM_x \to B(x, 2r_x) \subset M$ is a C^∞ diffeomorphism so that the derivatives of $\exp_x$ and $\exp_x^{-1}$ are bounded by $1 + \varepsilon$ on these balls. Since $\gamma([0, T])$ is compact, there are a finite collection of points $x_1 = p, x_2, \ldots, x_n = q \in \gamma[0, T]$ so that

$$\gamma([0, T]) \subset \bigcup_{j=1}^{n} B(x_j, r_j),$$

where we denote $r_j = r_{x_j}$ for short. Set $r = \min\{r_1, \ldots, r_n\}$, and pick an $m \in \mathbb{N}$ big enough so that $T/m < r/2$. Define $t_0 = 0 < t_1 = T/m < \cdots < t_j = jT/m < \cdots <$

$T = t_m$, and consider the points a_j, b_j with $a_j = b_{j-1} = \gamma(t_{j-1})$, $j = 1, \ldots, m$, and $b_m = \gamma(t_m)$.

For each $j \in \{1, \ldots, m-1\}$ we may choose an $i_j \in \{1, \ldots, n\}$ so that

$$\gamma[t_{j-1}, t_j] \cap B(x_{i_j}, r_{i_j}) \neq \emptyset,$$

and we also set $i_0 = 1$, $i_m = n$ (so that $x_{i_0} = p$ and $x_{i_m} = q$). Since the length of the restriction of γ to $[t_{j-1}, t_j]$, which we denote γ_j, is $t_j - t_{j-1} = T/m < r/2 \leqslant r_{i_j}/2$, this obviously means that

$$\gamma[t_{j-1}, t_j] \subset B(x_{i_j}, 2r_{i_j})$$

for each $j = 1, \ldots, m$. In order to avoid an unnecessary burden of notation, in the sequel we denote $y_j = x_{i_j}$, and $s_j = r_{i_j}$, for $j = 0, 1, \ldots, m$.

Consider the function $f_j : B(0_{y_j}, 2s_j) \to \mathbb{R}$ defined by $f_j = f \circ \exp_{y_j}$. By Corollary 4.11 we know that

$$D^- f_j(x) = \{\zeta \circ d\exp_{y_j}(x) : \zeta \in D^-(f)(\exp_{y_j}(x))\}$$

for all $x \in B(0_{y_j}, 2s_j)$. Since $\|\zeta\|_y \leqslant K$ for all $\zeta \in D^- f(y)$ with $y \in M$, and $\|d\exp_{y_j}(x)\| \leqslant (1+\varepsilon)$ for all $x \in B(0_{y_j}, 2s_j)$, we deduce that $\|\eta\|_{y_j} \leqslant (1+\varepsilon)K$ for all $\eta \in D^- f_j(x)$, $x \in B(0_{y_j}, 2s_j)$. Then we can apply the result for the case $H = TM_{x_j}$ and the function f_j to see that

$$\begin{aligned}|f(a_j) - f(b_j)| &= |f_j(\exp_{y_j}^{-1}(a_j)) - f(\exp_{y_j}^{-1}(b_j))| \\ &\leqslant (1+\varepsilon)K d_{TM_{y_j}}\left(\exp_{y_j}^{-1}(a_j)), \exp_{y_j}^{-1}(b_j)\right)\end{aligned}$$

for all $j = 1, 2, \ldots, m$. On the other hand, since $\exp_{x_j}^{-1}$ is $(1+\varepsilon)$-Lispchitz we also have

$$d_{TM_{y_j}}\left(\exp_{y_j}^{-1}(a_j)), \exp_{y_j}^{-1}(b_j)\right) \leqslant (1+\varepsilon) d_M(a_j, b_j)$$

for all $j = 1, 2, \ldots, m$. By combining these two last inequalities we deduce that

$$|f(a_j) - f(b_j)| \leqslant (1+\varepsilon)^2 K d_M(a_j, b_j) \leqslant (1+\varepsilon)^2 K \int_{t_{j-1}}^{t_j} \|d\gamma(t)\|\, dt$$

for all $j = 1, \ldots, m$. Therefore,

$$|f(p) - f(q)| = \left|\sum_{j=1}^{m} (f(a_j) - f(b_j))\right| \leqslant \sum_{j=1}^{m} |f(a_j) - f(b_j)|$$

$$\leqslant (1+\varepsilon)^2 K \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} \|d\gamma(t)\|\, dt = (1+\varepsilon)^2 K \int_0^T \|d\gamma(t)\|\, dt$$
$$= (1+\varepsilon)^2 K L(\gamma).$$

By taking the infimum over the set of continuous and piecewise C^1 paths γ joining p and q with length $L(\gamma)$, we get

$$|f(q) - f(p)| \leqslant (1+\varepsilon)^2 K d_M(q,p).$$

Finally, by letting ε go to 0 we obtain the desired inequality: $|f(q) - f(p)| \leqslant K d_M(q,p)$. $\square$

Corollary 4.19. *Let (M,g) be a Riemannian manifold, and let $f : M \longrightarrow \mathbb{R}$ be a continuous function. Then*

$$\sup\left\{\|\zeta\|_p : \zeta \in D^- f(p),\, p \in M\right\} = \sup\left\{\|\zeta\|_p : \zeta \in D^+ f(p),\, p \in M\right\}.$$

These quantities are finite if and only if f is Lipschitz on M, and in this case they are equal to the Lipschitz constant of f.

Theorem 4.20 (*Godefroy's mean value inequality*). *Let (M,g) be a Riemannian manifold, and let $f : M \longrightarrow \mathbb{R}$ be a Borel function such that*

$$D^- f(p) \cup D^+ f(p) \neq \emptyset$$

for every $p \in M$. Define $\Phi : M \longrightarrow \mathbb{R}$ by

$$\Phi(p) = \inf\{\|\zeta\|_p : \zeta \in D^- f(p) \cup D^+ f(p)\}.$$

Then, for every path $\gamma : I \to M$ parameterized by arc length, one has that

$$\mu\left(f(\gamma(I))\right) \leqslant \int_I \Phi(\gamma(t))\, dt.$$

Here μ is the Lebesgue measure in $\mathbb{R}$.

Proof. The result is already proved in the case when $M = H$ is a Hilbert space, see [33]. Let us see how the general case can be deduced. Let us denote $I = [0,T]$. For a given $\varepsilon > 0$, choose points $y_j = x_{i_j}, a_j, b_j$, and numbers $s_j = r_{i_j}, t_j$, exactly as in the proof of Theorem 4.18. Let us denote $f_j = f \circ \exp_{y_j} : B(0_{y_j}, 2s_j) \to \mathbb{R}$, $\gamma_j = \exp_{y_j}^{-1} \circ \gamma : I_j := [t_{j-1}, t_j] \to B(0_{y_j}, 2s_j) \subset TM_{y_j}$, and

$$\Phi_j(y) = \inf\{\|\zeta\|_{y_j} : \zeta \in D^- f_j(x) \cup D^+ f_j(x)\}$$

for each $x \in B(0_{y_j}, 2s_j)$. Since $D^- f_j(x) = \{\zeta \circ d\exp_{y_j}(x) : \zeta \in D^-(f)(\exp_{y_j}(x))\}$ for all $x \in B(0_{y_j}, 2s_j)$, and $\exp_{y_j}$ is $(1+\varepsilon)$-bi-Lipschitz on these balls, it is easy to see that $\Phi_j(x) \leqslant (1+\varepsilon)\Phi(\exp_{y_j}(x))$ for all $x \in B(0_{y_j}, 2s_j)$.

By applying the result for $H = TM_{y_j}$, the function f_j and the path γ_j, we get that

$$\mu\left(f(\gamma(I_j))\right) = \mu\left(f_j(\gamma_j(I_j))\right) \leqslant \int_{I_j} \Phi_j(\gamma_j(t))\, dt \tag{4.1}$$

for all $j = 1, 2, \ldots, m$. But we also have that

$$\int_{I_j} \Phi_j(\gamma_j(t))\, dt \leqslant \int_{I_j} (1+\varepsilon)\Phi(\exp_{y_j}(\gamma_j(t)))\, dt = (1+\varepsilon)\int_{I_j} \Phi(\gamma(t))\, dt. \tag{4.2}$$

By combining inequalities (4.1) and (4.2), and summing over $j = 1, \ldots, m$, we get

$$\begin{aligned}\mu(f(\gamma(I))) &\leqslant \sum_{j=1}^{m} \mu\left(f(\gamma(I_j))\right) \leqslant \sum_{j=1}^{m} (1+\varepsilon)\int_{I_j} \Phi(\gamma(t))\, dt \\ &\leqslant (1+\varepsilon)\sum_{j=1}^{m} \int_{I_j} \Phi(\gamma(t))\, dt = (1+\varepsilon)\int_I \Phi(\gamma(t))\, dt.\end{aligned}$$

Finally, by letting ε go to 0 we get $\mu(f(\gamma(I))) \leqslant \int_I \Phi(\gamma(t))\, dt$. □

Corollary 4.21. *Let (M, g) be a Riemannian manifold, $f : M \to \mathbb{R}$ a Borel function such that for every $p \in M$ there exists $\zeta \in D^- f(p) \cup D^+ f(p)$ with $\|\zeta\|_p \leqslant K$. Then,*

$$\mu(f(\gamma(I))) \leqslant K L(\gamma)$$

for every path $\gamma : I \to M$. In particular, when f is continuous it follows that $|f(p) - f(q)| \leqslant K d_M(p, q)$ for all $p, q \in M$.

5. (Sub)differentiability of convex functions on Riemannian manifolds

The aim of this section is to prove that every (continuous) convex function defined on a Riemannian manifold is everywhere subdifferentiable, and differentiable on a dense set.

Definition 5.1. Let M be a Riemannian manifold. A function $f : M \to \mathbb{R}$ is said to be convex provided that the function $f \circ \sigma : I \subseteq \mathbb{R} \to \mathbb{R}$ is convex for every geodesic $\sigma : I \to M$ (parameterized by arc length).

The following proposition is probably known, at least in the case when M is finite-dimensional, but we provide a short proof for the reader's convenience, as we have not been able to find an explicit reference.

Proposition 5.2. *Let M be a Riemannian manifold. If a function $f : M \to \mathbb{R}$ is convex and locally bounded, then f is locally Lipschitz. In particular, every continuous convex function is locally Lipschitz.*

Proof. Take $p \in M$. Since f is locally bounded there exists $R > 0$ such that f is bounded on the ball $B(p, R)$. According to Theorem 2.6, there exists $r > 0$ with $0 < r < R/2$ such that the open balls $B(p, 2r)$ and $B(p, r)$ are convex. Fix $C = \sup\{f(x) : x \in B(p, 2r)\}$, and $m = \inf\{f(x) : x \in B(p, 2r)\}$. We are going to see that f is K-Lispchitz on the ball $B(p, r)$, where $K = (C - m)/r$. Indeed, take $x_1, x_2 \in B(p, r)$. Since $B(p, r)$ is convex, there exists a unique geodesic $\gamma : [t_1, t_2] \to B(p, r)$, with length $d(x_1, x_2) = t_2 - t_1$, joining x_1 to x_2. Take $v_1 \in TM_{x_1}$ such that $\gamma(t) = \exp_{x_1}((t - t_1)v_1)$ for $t \geqslant t_1$ small enough. Since the ball $B(p, 2r)$ is still convex and $x_1 \in B(p, r)$, we may define a geodesic $\sigma_1 : [-r, r] \to B(p, 2r) \subset M$ through x_1 by

$$\sigma_1(t) = \exp_{x_1}(tv_1) \quad \text{for all } t \in [-r, r].$$

In the same way we may take $v_2 \in TM_{x_2}$ and define a geodesic $\sigma_2 : [-r, r] \to B(p, 2r) \subset M$ through x_2 by

$$\sigma_2(t) = \exp_{x_2}(tv_2) \quad \text{for all } t \in [-r, r],$$

in such a way that $\gamma(t) = \exp_{x_2}((t - t_2)v_2)$ for $t \leqslant t_2$ with $|t|$ small enough. Set $t_3 = t_1 - r$, $t_4 = t_2 + r$, $x_3 = \sigma_1(-r)$, $x_4 = \sigma_2(r)$, and $I = [t_3, t_4]$. Then, if we define $\sigma : I \to B(p, 2r)$ by

$$\sigma(t) = \begin{cases} \sigma_1(t - t_1) & \text{if } t \in [t_3, t_1]; \\ \gamma(t) & \text{if } t \in [t_1, t_2]; \\ \sigma_2(t - t_2) & \text{if } t \in [t_2, t_4], \end{cases}$$

it is clear that σ is a geodesic joining x_3 to x_4 in $B(p, 2r)$. Now, since f is convex, the function $g : [t_3, t_4] \subset \mathbb{R} \to \mathbb{R}$ defined as

$$g(t) = f(\sigma(t))$$

is convex. Therefore we have

$$\frac{g(t_1) - g(t_3)}{t_1 - t_3} \leqslant \frac{g(t_2) - g(t_1)}{t_2 - t_1} \leqslant \frac{g(t_4) - g(t_2)}{t_4 - t_2},$$

where $t_3 = t_1 - r < t_1 < t_2 < t_2 + r = t_4$. Bearing in mind that $x_3, x_4 \in B(p, 2r)$, and $t_2 - t_1 = d(x_1, x_2)$, it follows that:

$$-\frac{C-m}{r} \leqslant \frac{f(x_1) - f(x_3)}{r} \leqslant \frac{f(x_2) - f(x_1)}{d(x_1, x_2)} \leqslant \frac{f(x_4) - f(x_2)}{r} \leqslant \frac{C-m}{r}.$$

This shows that $|f(x_1) - f(x_2)| \leqslant K d(x_1, x_2)$ for all $x_1, x_2 \in B(p, r)$, where $K = (C-m)/r$. $\square$

Let us recall that for a locally Lipschitz function $F : H \to \mathbb{R}$ on a Hilbert space H, we may define the generalized directional derivative $F^0(x, v)$ as the

$$\limsup_{(y,t)\to(x,0^+)} \frac{F(y+tv) - F(y)}{t}.$$

For every $x \in H$, $F^0(x, v)$ is a subadditive positively homogeneous function of v, and the set $\{x^* \in H^* : x^*(v) \leqslant F^0(x, v) \text{ for all } v\}$ is called the generalized gradient of F at x, and is denoted by $\partial F(x)$. The generalized gradient is a nonempty, convex, w^*-compact subset of H^*; see [10] for more information.

Theorem 5.3. *Let $g : M \to \mathbb{R}$ be a continuous convex function on a Riemannian manifold. Then g is subdifferentiable at every point of M.*

Proof. Let $\phi_p : U_p \to H$ be an exponential chart at p. We have $\phi_p(p) = 0$. Given another point $q \in U_p$, take a $(\phi_p, v) \in TM_q$, and denote $\sigma_{q,v}(t) = \phi_q^{-1}(tw)$, where $(\phi_p, v) \sim (\phi_q, w)$, which is a geodesic passing through q with derivative (ϕ_p, v). Here, $(\phi_p, v) \sim (\phi_q, w)$ means that $w = d(\phi_q \circ \phi_p^{-1})(\phi_p(q))(v)$, or equivalently $v = d(\phi_p \circ \phi_q^{-1})(0_q)(w)$.

Let us define

$$f^0(p, v) = \limsup_{q\to p \ \ t\to 0^+} \frac{f(\sigma_{q,v}(t)) - f(q)}{t}.$$

Claim 5.4. *We have that $f^0(p, v) = \inf_{t>0} \frac{f(\sigma_{p,v}(t)) - f(p)}{t}$, and consequently*

$$f^0(p, v) = \inf_{t>0} \frac{(f \circ \phi_p^{-1})(tv) - (f \circ \phi_p^{-1})(0)}{t}.$$

Claim 5.5. *There exists $x^* \in H^*$ such that $x^*(v) \leqslant f^0(p, v)$ for every $v \in H$.*

From these facts it follows that:

$$(f \circ \phi_p^{-1})(tv) - (f \circ \phi_p^{-1})(0) - x^*(tv) \geqslant 0$$

for every $v \in S_H$, and every $t \in [0, r)$, provided that $B(0, r) \subset \phi_p(U_p)$. Hence $(f \circ \phi_p^{-1}) - x^*$ attains a minimum at 0 and therefore $x^* \in D^-((f \circ \phi_p^{-1}))(0)$. We then conclude that $D^- f(p) \neq \emptyset$ by Corollary 4.5. This shows the theorem. □

Proof of Claim 5.4. Fix a $\delta > 0$. Since $f \circ \sigma_{q,v}$ is convex we have that

$$f^0(p, v) = \lim_{\varepsilon \to 0^+} \sup_{d(p,q) \leqslant \varepsilon\delta} \sup_{0<t<\varepsilon} \frac{f(\sigma_{q,v}(t)) - f(q)}{t}$$
$$= \lim_{\varepsilon \to 0^+} \sup_{d(p,q) \leqslant \varepsilon\delta} \frac{f(\sigma_{q,v}(\varepsilon)) - f(\sigma_{q,v}(0))}{\varepsilon} = (*).$$

Next, we estimate $d(\sigma_{p,v}(\varepsilon), \sigma_{q,v}(\varepsilon))$. We have

$$\begin{aligned} d(\sigma_{p,v}(\varepsilon), \sigma_{q,v}(\varepsilon)) &\leqslant K_p ||\phi_p(\sigma_{p,v}(\varepsilon)) - \phi_p(\sigma_{q,v}(\varepsilon))|| = K_p ||\varepsilon v - \phi_p(\sigma_{q,v}(\varepsilon))|| \\ &= K_p ||\varepsilon v - (\phi_p \circ \phi_q^{-1})(\varepsilon w)|| \\ &= K_p ||\varepsilon v - (\phi_p \circ \phi_q^{-1})(0) - \varepsilon d(\phi_p \circ \phi_q^{-1})(0)(w) - o(\varepsilon)|| \\ &= K_p ||(\phi_p \circ \phi_q^{-1})(0) + o(\varepsilon)|| \leqslant K_p(||\phi_p(q)|| + ||o(\varepsilon)||) \\ &\leqslant K_p(L_p d(p, q) + ||o(\varepsilon)||) \leqslant K_p(L_p \varepsilon\delta + \varepsilon\delta) \leqslant C\varepsilon\delta, \end{aligned}$$

where L_p and K_p are the Lipschitz constants of ϕ_p and ϕ_p^{-1} respectively, $C = K_p(L_p + 1)$, and ε is small enough so that $\|o(\varepsilon)\| \leqslant \varepsilon\delta$ and $||D(\phi_p \circ \phi_q^{-1})(v) - v|| < \delta$.

Since f is locally Lipschitz there exists $K > 0$ so that f is K-Lispchitz on a neighborhood of p which may be assumed to be U_p. From the above estimates we get that, for $d(p, q) \leqslant \varepsilon\delta$,

$$\begin{aligned} &\left| \frac{f(\sigma_{q,v}(\varepsilon)) - f(\sigma_{q,v}(0))}{\varepsilon} - \frac{f(\sigma_{p,v}(\varepsilon)) - f(\sigma_{p,v}(0))}{\varepsilon} \right| \\ &\leqslant \frac{1}{\varepsilon} (|f(\sigma_{q,v}(\varepsilon)) - f(\sigma_{p,v}(\varepsilon))| + |f(p) - f(q)|) \\ &\leqslant \frac{K}{\varepsilon} (d(\sigma_{p,v}(\varepsilon), \sigma_{q,v}(\varepsilon)) + d(p, q)) \leqslant K(C + 1)\delta. \end{aligned}$$

Now we deduce that

$$(*) \leqslant \lim_{\varepsilon \to 0^+} \frac{f(\sigma_{p,v}(\varepsilon)) - f(\sigma_{p,v}(0))}{\varepsilon} + K(C + 1)\delta$$

and, by letting $\delta \to 0$, we get

$$f^0(p,v) \leqslant \lim_{t\to 0^+} \frac{f(\sigma_{p,v}(t)) - f(p)}{t} = \inf_{t>0} \frac{f(\sigma_{p,v}(t)) - f(p)}{t}.$$

Since the other inequality holds trivially the claim is proved. □

Proof of Claim 5.5. We have that

$$\begin{aligned}
&\limsup_{q\to p}\sup_{t\to 0^+} \frac{f(\sigma_{q,v}(t)) - f(\sigma_{q,v}(0))}{t} \\
&= \lim_{\varepsilon\to 0^+} \sup_{d(p,q)<\varepsilon\ 0<t<\varepsilon} \frac{f(\sigma_{q,v}(t)) - f(\sigma_{q,v}(0))}{t} \\
&= \lim_{\varepsilon\to 0^+} \sup_{d(p,q)<\varepsilon\ 0<t<\varepsilon} \frac{(f\circ\phi_p^{-1})(\phi_p(\sigma_{q,v}(t)) - (f\circ\phi_p^{-1})(\phi_p(q))}{t} \\
&= \lim_{\varepsilon\to 0^+} \sup_{d(p,q)<\varepsilon\ 0<t<\varepsilon} \frac{F(y+\lambda_y(t)) - F(y)}{t},
\end{aligned}$$

where $(f\circ\phi_p^{-1}) = F$, $y = \phi_p(q)$, and $\lambda_y(t) = \phi_p(\sigma_{q,v}(t)) - \phi_p(q)$. Next, we get

$$\lim_{\varepsilon\to 0^+} \sup_{d(p,q)<\varepsilon\ 0<t<\varepsilon} \frac{F(y+\lambda_y(t)) - F(y)}{t} = \lim_{\varepsilon\to 0^+} \sup_{||y||<\varepsilon\ 0<t<\varepsilon} \frac{F(y+\lambda_y(t)) - F(y)}{t},$$

because $L_p||y|| \leqslant d(p,q) \leqslant K_p||y||$ (recall that ϕ_p and $(\phi_p)^{-1}$ are Lipschitz).

Now, if we take $||y|| < \varepsilon$ and $0 < t < \varepsilon$, we have

$$\begin{aligned}
\left|\frac{F(y+\lambda_y(t)) - F(y)}{t} - \frac{F(y+tv) - F(y)}{t}\right| &= \left|\frac{F(y+\lambda_y(t)) - F(y+tv)}{t}\right| \\
&\leqslant K' \frac{||\lambda_y(t) - tv||}{t} = K'\varphi(t),
\end{aligned}$$

where K' is the Lipschitz constant of F and φ satisfies $\lim_{t\to 0^+} \varphi(t) = 0$, because

$$\lambda_y(t) - tv = \phi_p(\sigma_{q,v}(t)) - \phi_p(q) - tv = o(t).$$

Finally, we have

$$\left| \sup_{||y||<\varepsilon\ 0<t<\varepsilon} \frac{F(y+\lambda_y(t)) - F(y)}{t} - \sup_{||y||<\varepsilon\ 0<t<\varepsilon} \frac{F(y+tv) - F(y)}{t} \right|$$

$$\leqslant \sup_{||y||<\varepsilon \;\; 0<t<\varepsilon} \left| \frac{F(y+\lambda_y(t))-F(y)}{t} - \frac{F(y+tv)-F(y)}{t} \right|$$
$$\leqslant K' \sup_{0<t<\varepsilon} \varphi(t),$$

which goes to 0 as $\varepsilon \to 0^+$. Therefore $f^0(p,v) = F^0(0,v)$ and $x^*(v) \leqslant f^0(p,v)$ for every $v \in H$, provided that $x^* \in \partial F(0)$, the generalized gradient of F at 0. $\square$

Theorem 5.6. *Let $g : M \to \mathbb{R}$ be a continuous convex function on a complete Riemannian manifold. Then the set* $\text{Diff}(g) := \{x \in M : g \text{ is differentiable at } x\}$ *is dense in M.*

Proof. According to Proposition 4.17, $\text{Diff}^+(g) := \{p \in M : D^+g(p) \neq \emptyset\}$ is dense in M. On the other hand, by Theorem 5.3, we know that $\text{Diff}^-(g) := \{p \in M : D^-g(p) \neq \emptyset\} = M$. Then, by Proposition 4.6, we get that

$$\text{Diff}(g) = \text{Diff}^+(g) \cap \text{Diff}^-(g) = \text{Diff}^+(g) \text{ is dense in } M. \qquad \square$$

By using more sophisticated tools, this result can be extended to the category of locally Lipschitz functions, as we next show.

Theorem 5.7. *Let $g : M \to \mathbb{R}$ be a locally Lipschitz function. If M is finite-dimensional, then g is differentiable almost everywhere, that is, the set $M \setminus \text{Diff}(g)$ has measure zero. If M is infinite-dimensional, then the set of points of differentiability of g, $\text{Diff}(g)$, is dense in M.*

Proof. Since M is separable, it suffices to prove the result for any small enough open set $U \subset M$ so that g is Lipschitz on U. Take a point $p \in U$. Since the exponential mapping at p is locally almost an isometry, in particular Lipschitz, it provides us with a chart $h = \Phi_p : V \to H$ which is Lipschitz, for a suitably small open set $V \subset U$. Then the composition $g \circ h^{-1} : h(V) \subset H \to \mathbb{R}$ is a Lipschitz function from an open subset of a Hilbert space into $\mathbb{R}$.

In the case when H is finite-dimensional, the classic theorem of Rademacher tells us that $g \circ h^{-1}$ is differentiable almost everywhere in $h(V)$ (see [31]) and, since h is a C^1 diffeomorphism (so h preserves points of differentiability and sets of measure zero), it follows that g is differentiable almost everywhere in V.

If H is infinite-dimensional then we can apply a celebrated theorem of Preiss that ensures that every Lipschitz function from an open set of an Asplund Banach space (such as the Hilbert space) has at least one point of differentiability [44]. By this theorem, it immediately follows that $g \circ h^{-1}$ is differentiable on a dense subset of $h(V)$. Since again h is a C^1 diffeomorphism, we have that g is differentiable on a dense subset of V.

Finally, since M can be covered by a countable union of such open sets V on each of which g is Lipschitz, the result follows. □

Corollary 5.8. *Let M be a Riemannian manifold, and $f : M \to \mathbb{R}$ a convex and locally bounded function. Then f is differentiable on a dense subset of M (whose complement has measure zero if M is finite-dimensional).*

Proof. By Proposition 5.2 we know that f is locally Lipschitz. Then, by Theorem 5.7. it follows that f is differentiable on a dense subset of M. □

6. Hamilton–Jacobi equations in Riemannian manifolds

First-order Hamilton–Jacobi equations are of the form

$$F(x, u(x), du(x)) = 0$$

in the stationary case, and of the form

$$F(t, x, u(x, t), du(t, x)) = 0$$

in the evolution case. These equations arise, for instance, in optimal control theory, Lyapounov theory, and differential games.

Even in the simplest cases, such as the space $\mathbb{R}^n$, it is well known that very natural Hamilton–Jacobi equations do not always admit classical solutions. However, weaker solutions, such as the so-called viscosity solutions, do exist under very general assumptions. There is quite a large amount of literature about viscosity solutions to Hamilton–Jacobi equations, see [8,10–20,26,27] and the references cited therein, for instance. All these works deal with Hamilton–Jacobi equations in $\mathbb{R}^n$ or in infinite-dimensional Banach spaces.

Examples of Hamilton–Jacobi equations also arise naturally in the setting of Riemannian manifolds, see [1]. However, we do not know of any work that has studied nonsmooth solutions, in general, or viscosity solutions, in particular, to Hamilton–Jacobi equations defined on Riemannian manifolds (either finite-dimensional or infinite-dimensional). This may be due to the lack of a theory of nonsmooth calculus for functions defined on Riemannian manifolds.

In this final section, we will show how the subdifferential calculus we have developed can be applied to get results on existence and uniqueness of viscosity solutions to some Hamilton–Jacobi equations defined on Riemannian manifolds. We will also prove some results about "regularity" (meaning Lipschitzness) of viscosity solutions to some of these equations.

There are lots of Hamilton–Jacobi equations on Riemannian manifolds M for which the tools we have just developed could be used in one way or the other to get interesting results about viscosity solutions. For instance, one could get a maximum principle for stationary first-order Hamilton–Jacobi equations of the type

$$\begin{cases} u(x) + F(x, du(x)) = 0 \text{ for all } x \in \Omega, \\ u(x) = 0 \text{ for all } x \in \partial\Omega, \end{cases}$$

where Ω is an open submanifold of M with boundary $\partial\Omega$. One could also prove a maximum principle for parabolic Hamilton–Jacobi equations of the form

$$\begin{cases} u_t + F(x, u_x) = 0, \\ u(0, x) = u_0(x), \end{cases}$$

where $u : \mathbb{R}^+ \times M \to \mathbb{R}$, and $u_0 : M \to \mathbb{R}$ is an initial condition (assumed to be bounded and uniformly continuous), in the manner of [26, Section 6].

However, this final section is only intended to give a glimpse of the potential applications of nonsmooth calculus to the theory of Hamilton–Jacobi equations on Riemannian manifolds, and not to elaborate a comprehensive treatise on such equations. That is why we will restrict ourselves mainly to one of the most interesting examples of first-order Hamilton–Jacobi equations, namely equations of the form

$$(*) \begin{cases} u + G(du) = f, \\ u \text{ bounded}, \end{cases}$$

where $f : M \to \mathbb{R}$ is a bounded uniformly continuous function, and $G : T^*M \to \mathbb{R}$ is a function defined on the cotangent bundle of M. In fact these equations are really of the form

$$(*) \begin{cases} u + F(du) = 0, \\ u \text{ bounded}, \end{cases}$$

where $F : T^*M \to \mathbb{R}$, since we can always take a function F of the form $F(x, \xi_x) = G(x, \xi_x) - f(x)$.

A bounded Fréchet-differentiable function $u : M \to \mathbb{R}$ is a classical solution of the equation $(*)$ provided that

$$u(p) + F(p, du(p)) = 0 \quad \text{for every} \quad p \in M.$$

Let us now introduce the notion of viscosity solution.

Definition 6.1. An upper semicontinuous (usc) function $u : M \to \mathbb{R}$ is a *viscosity subsolution* of $u + F(du) = 0$ if $u(p) + F(p, \zeta) \leqslant 0$ for every $p \in M$ and $\zeta \in D^+u(p)$.

A lower semicontinuous (lsc) function $u : M \to \mathbb{R}$ is a *viscosity supersolution* of $u + F(du) = 0$ if $u(p) + F(p, \zeta) \geqslant 0$ for every $p \in M$ and $\zeta \in D^- u(p)$. A continuous function $u : M \to \mathbb{R}$ is a *viscosity solution* of $u + F(du) = 0$ if it is both a viscosity subsolution and a viscosity supersolution of $u + F(du) = 0$.

We can define viscosity solutions on a open set $\Omega \subset M$ in a natural way when the functions are defined on Ω.

Remark 6.2. Since for a Fréchet differentiable function u we have $D^+ u(p) = D^- u(p) = \{du(p)\}$, it is clear that every bounded Frechet differentiable viscosity solution of $u + F(du) = 0$ is a classical solution of $(*)$.

We are going to show the existence and uniqueness of viscosity solutions to Hamilton–Jacobi equations of the form $(*)$ provided that $F : T^*M \to \mathbb{R}$ is a function defined on the cotangent bundle of M which satisfies a certain uniform continuity condition, see Definition 6.10 below. The manifold M must also satisfy the following requirement.

Throughout the remainder of this section M will be a complete Riemannian manifold (either finite- or infinite-dimensional) such that M satisfies conditions (3) or (4) (which are both equivalent) of Proposition 3.9, that is, M is uniformly locally convex and has a strictly positive injectivity radius. Equivalently, there is a constant $r = r_M > 0$ such that for every $x \in M$ the mapping $\exp_x$ is defined on $B(0_x, r) \subset TM_x$ and provides a C^∞ diffeomorphism

$$\exp_x : B(0_x, r) \to B(x, r)$$

and the distance function is given by the expression

$$d(y, x) = \| \exp_x^{-1}(y) \|_x \quad \text{for all } \; y \in B(x, r).$$

In particular, all compact manifolds satisfy this property. In the remainder of this section the constant $r = r_M$ will be fixed.

Note also that if M satisfies condition (3) of Proposition 3.9 then M is uniformly bumpable and therefore the smooth variational principle 3.11 holds for M.

We begin with a simple observation that if M is uniformly bumpable then so is $M \times M$.

Lemma 6.3. *Let M be a Riemannian manifold. If M is uniformly bumpable then $M \times M$ is uniformly bumpable as well.*

Proof. The natural Riemannian structure in $M \times M$ induced by (M, g) is the one given by

$$(g \times g)_{(p_1, p_2)}\big((v_1, v_2),\ (w_1, w_2)\big) := g_{p_1}(v_1, w_1) + g_{p_2}(v_2, w_2).$$

Let $d_{M \times M}$ denote the Riemannian distance that this metric gives rise to in the product $M \times M$. It is obvious that if $\gamma(t) = \big(\alpha(t), \beta(t)\big)$ is a path in $M \times M$ then α and β are paths in M satisfying

$$\max\{L(\alpha), L(\beta)\} \leqslant L(\gamma) \leqslant L(\alpha) + L(\beta) \leqslant 2\max\{L(\alpha), L(\beta)\},$$

which implies that

$$\begin{aligned}\max\{d_M(x_1, y_1), d_M(x_2, y_2)\} &\leqslant d_{M\times M}((x_1, x_2), (y_1, y_2)) \\ &\leqslant d_M(x_1, y_1) + d_M(x_2, y_2) \\ &\leqslant 2\max\{d_M(x_1, y_1), d_M(x_2, y_2)\}\end{aligned}$$

for every $x = (x_1, x_2)$, $y = (y_1, y_2) \in M \times M$.

Since M is uniformly bumpable, there exist numbers $R = R_M > 1$ and $r_M > 0$ such that for every $p_0 \in M$, $\delta \in (0, r_M)$ there exists a C^1 smooth function $b : M \to [0, 1]$ such that $b(p_0) = 1$, $b(x) = 0$ if $d_M(x, p_0) \geqslant \delta$, and $\sup_{x\in M} \|db(x)\|_x \leqslant R/\delta$. Now take a point $p = (p_1, p_2) \in M \times M$. For any $\delta \in (0, r_M)$, there are C^1 smooth bumps b_1, b_2 on M such that $b_i(p_i) = 1$, $b_i(x_i) = 0$ whenever $d_M(x_i, p_i) \geqslant \delta$, and $\|db_i(x_i)\| \leqslant R/\delta$ for every $x_i \in M$; $i = 1, 2$. Define a C^1 smooth bump $b : M \times M \to \mathbb{R}$ by

$$b(x) = b(x_1, x_2) = b_1(x_1)b_2(x_2) \quad \text{for all } x = (x_1, x_2) \in M \times M.$$

It is obvious that $b(p_1, p_2) = 1$. If $d_{M\times M}(x, p) \geqslant 2\delta$ we have that

$$2\max\{d_M(x_1, p_1), d_M(x_2, p_2)\} \geqslant d_{M\times M}(x, p) \geqslant 2\delta,$$

so $d_M(x_i, p_i) \geqslant \delta$ for some $i \in \{1, 2\}$, hence $b_i(x_i) = 0$ for the same i, and $b(x) = 0$. Finally, we have that

$$\|db(x_1, x_2)\|^2_{(x_1,x_2)} = \|db_1(x_1)\|^2_{x_1} + \|db_2(x_2)\|^2_{x_2} \leqslant 2(R/\delta)^2,$$

which means that

$$\|db(x)\|_x \leqslant \frac{2\sqrt{2}R}{2\delta}$$

for every $x = (x_1, x_2) \in M \times M$. Therefore $M \times M$ satisfies the conditions in Definition 3.6, with $R_{M\times M} = 2\sqrt{2}R$, and $r_{M\times M} = 2r_M$. □

Since we are assuming that M is uniformly locally convex and $i(M) > r > 0$, hence that the distance function $y \mapsto d(y, x)$ is C^∞ smooth on $B(x, r) \setminus \{x\}$ for every $x \in M$, we can consider the distance function $d : M \times M \to \mathbb{R}$ and its partial derivatives $\partial d(x_0, y_0)/\partial x$ and $\partial d(x_0, y_0)/\partial y$. We next see that these partial derivatives satisfy a nice antisymmetry property. In order to compare them in a natural way we need to use the parallel translation from TM_{x_0} to TM_{y_0} along the geodesic joining x_0 to y_0 (note that there is a unique minimizing geodesic joining x_0 to y_0 because M is uniformly locally convex and $d(x_0, y_0) < r$).

Notation 6.4. Let $x_0, y_0 \in M$ be such that $d(x_0, y_0) < r$. Let $\gamma(t) = \exp_{x_0}(tv_0)$, $0 \leqslant t \leqslant 1$ be the unique minimizing geodesic joining these two points. For every vector $w \in TM_{x_0}$, we denote

$$L_{x_0 y_0}(w) = P^1_{0,\gamma}(w)$$

the parallel translation of w from x_0 to y_0 along γ. Recall that the mapping $L_{x_0 y_0} : TM_{x_0} \to TM_{y_0}$ is a linear isometry, with inverse $L_{y_0 x_0} : TM_{y_0} \to TM_{x_0}$. As we customarily identify TM_p with T^*M_p (via the linear isometry $v \mapsto \langle v, \cdot \rangle_p$), the isometry $L_{x_0 y_0}$ induces another linear isometry between the cotangent fibers $T^*M_{x_0}$ and $T^*M_{y_0}$. We will still denote this new isometry by $L_{x_0 y_0} : T^*M_{x_0} \to T^*M_{y_0}$.

Lemma 6.5. *Let $x_0, y_0 \in M$ be such that $0 < d(x_0, y_0) < r$. Then*

$$L_{y_0 x_0}\left(\frac{\partial d(y_0, x_0)}{\partial y}\right) = -\frac{\partial d(x_0, y_0)}{\partial x}.$$

Proof. Denote $r_0 = d(x_0, y_0) < r$. Consider the geodesic $\gamma(t) = \exp_{x_0}(tv_0)$, $0 \leqslant t \leqslant 1$, where $y_0 = \exp_{x_0}(v_0)$. By the definitions of parallel translation and geodesic it is clear that

$$L_{x_0 y_0}(v_0) = \gamma'(1) = d\exp_{x_0}(v_0)(v_0).$$

On the other hand, under the current assumptions on M, and by the Gauss lemma (see [39,41]), we know that $\gamma'(1)$ is orthogonal to the sphere $S(x_0, r_0) = \{y \in M : d(y, x_0) = r_0\} = \exp_{x_0}(S(0_{x_0}, r_0))$. Since this sphere $S(x_0, r_0)$ is a one-codimensional submanifold of M defined as the set of zeros of the smooth function $y \mapsto d(y, x_0) - r_0$ and (as is easily checked)

$$\frac{\partial d(y_0, x_0)}{\partial y} \neq 0,$$

we also have that this partial derivative is orthogonal to the sphere $S(x_0, r_0)$ at the point y_0. Therefore,

$$L_{x_0 y_0}(v_0) = \gamma'(1) = \lambda \frac{\partial d(y_0, x_0)}{\partial y}$$

for some $\lambda \neq 0$. Furthermore, since the function $t \mapsto d(\gamma(t), x_0)$ is increasing, we get that $\lambda > 0$. Finally, it is clear that $y \mapsto d(y, x_0)$ is 1-Lipschitz, and

$$\left\| \frac{\partial d(y_0, x_0)}{\partial y} \right\| = 1,$$

from which we deduce that $\lambda = \|L_{x_0 y_0}(v_0)\|_{y_0} = \|v_0\|_{x_0}$, and

$$L_{x_0 y_0}(v_0) = \|v_0\|_{x_0} \frac{\partial d(x_0, y_0)}{\partial y}. \tag{6.1}$$

Now consider the geodesic $\beta(t) = \exp_{y_0}(t w_0)$, $0 \leqslant t \leqslant 1$, where $\exp_{y_0}(w_0) = x_0$. By the definitions of parallel translation and geodesic we know that

$$L_{x_0 y_0}(v_0) = -w_0 \quad \text{and} \quad \|w_0\|_{y_0} = \|v_0\|_{x_0}. \tag{6.2}$$

A completely analogous argument to the one we used for γ above shows that

$$L_{y_0 x_0}(w_0) = \|w_0\| \frac{\partial d(x_0, y_0)}{\partial x}. \tag{6.3}$$

By combining (6.1)–(6.3) we immediately get that

$$L_{y_0 x_0}\left(\frac{\partial d(y_0, x_0)}{\partial y}\right) = \frac{v_0}{\|v_0\|_{x_0}} = -\frac{L_{y_0 x_0}(w_0)}{\|w_0\|_{y_0}} = -\frac{\partial d(x_0, y_0)}{\partial x}. \qquad \square$$

The following proposition can be viewed as a perturbed minimization principle for the sum or the difference of two functions. Its proof is a consequence of the smooth variational principle 3.11 and Lemma 6.5.

Proposition 6.6. *Let $u, v : M \to \mathbb{R}$ be two bounded functions which are upper semicontinuous (usc) and lower semicontinuous (lsc) respectively. Then, for every $\varepsilon > 0$, there exist $x_0, y_0 \in M$, and $\zeta \in D^+u(x_0)$, $\xi \in D^-v(y_0)$ such that*

(i) $d(x_0, y_0) < \varepsilon$,
(ii) $\|\zeta - L_{y_0 x_0}(\xi)\|_{x_0} < \varepsilon$
(iii) $v(z) - u(z) \geqslant v(y_0) - u(x_0) - \varepsilon$ *for every* $z \in M$.

Here $L_{y_0 x_0} : T^*M_{y_0} \to T^*M_{x_0}$ stands for the parallel translation.

Proof of Proposition 6.6. We can obviously assume that $\varepsilon < r(M)$. Let $b : \mathbb{R} \to \mathbb{R}$ be a C^∞ smooth function such that b is nonincreasing,

$$b(t) = b(0) > 2(\|v\|_\infty + \|u\|_\infty) + \varepsilon \quad \text{if } t \leqslant \varepsilon/4, \qquad \text{and } b(t) = 0 \quad \text{if } t \geqslant \varepsilon. \tag{6.4}$$

Define the function $w : M \times M \to \mathbb{R}$ by

$$w(x, y) = v(y) - u(x) - b(d(x, y)) \quad \text{for all } (x, y) \in M \times M.$$

The function w is lower semicontinuous and bounded. By Lemma 6.3 we know that $M \times M$ is uniformly bumpable, and $M \times M$ is obviously complete, so we can apply the smooth variational principle 3.11 to the function w to get a pair $(x_0, y_0) \in M \times M$ and a C^1 smooth function $g : M \times M \to \mathbb{R}$ such that

(a) $\|g\|_\infty < \varepsilon/2 > \|dg\|_\infty$

(b) $v(y) - u(x) - b(d(x, y)) - g(x, y) \geqslant v(y_0) - u(x_0) - b(d(x_0, y_0)) - g(x_0, y_0)$ for all $x, y \in M$.

If we take $x = x_0$ in (b) we get that v is subdifferentiable at the point y_0, and

$$\xi := \frac{\partial g(x_0, y_0)}{\partial y} + \frac{\partial (b \circ d)(x_0, y_0)}{\partial y} \in D^- v(y_0). \tag{6.5}$$

In a similar manner, by taking $y = y_0$ in (b) we get that

$$\zeta := -\left(\frac{\partial g(x_0, y_0)}{\partial x} + \frac{\partial (b \circ d)(x_0, y_0)}{\partial x} \right) \in D^+ u(x_0). \tag{6.6}$$

Let us note that

$$\begin{aligned} &L_{y_0 x_0}\left(\frac{\partial (b \circ d)(x_0, y_0)}{\partial y} \right) + \frac{\partial (b \circ d)(x_0, y_0))}{\partial x} \\ &= b'(d(x_0, y_0)) \left[L_{y_0 x_0}\left(\frac{\partial d(x_0, y_0)}{\partial y} \right) + \frac{\partial d(x_0, y_0)}{\partial x} \right] = 0, \end{aligned} \tag{6.7}$$

thanks to Lemma 6.5 when $x_0 \neq y_0$, and to the definition of b when $x_0 = y_0$. Therefore,

$$\begin{aligned} &\|L_{y_0 x_0}(\xi) - \zeta\|_{x_0} \\ &= \left\| L_{y_0 x_0}\left(\frac{\partial g(x_0, y_0)}{\partial y} \right) + L_{y_0 x_0}\left(\frac{\partial (b \circ d)(x_0, y_0)}{\partial y} \right) \right. \\ &\quad \left. + \frac{\partial g(x_0, y_0)}{\partial x} + \frac{\partial (b \circ d)(x_0, y_0)}{\partial x} \right\|_{x_0} \\ &= \|L_{y_0 x_0}\left(\frac{\partial g(x_0, y_0)}{\partial y} \right) + \frac{\partial g(x_0, y_0)}{\partial x} \|_{x_0} \\ &\leqslant \left\| \frac{\partial g(x_0, y_0)}{\partial y} \right\|_{y_0} + \left\| \frac{\partial g(x_0, y_0)}{\partial x} \right\|_{x_0} \leqslant \|dg\|_\infty + \|dg\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

which shows (ii).

On the other hand, if we had $d(x_0, y_0) \geqslant \varepsilon$ then, by taking $x = y = z$ in (b) we would get

$$b(0) \leqslant v(z) - u(z) - g(z,z) + g(x_0, y_0) - v(y_0) + u(x_0)$$
$$\leqslant 2(\|v\|_\infty + \|u\|_\infty) + \varepsilon,$$

which contradicts the definition of b, see (1) above. Therefore $d(x_0, y_0) < \varepsilon$ and (i) is proved.

Finally, if we take $z = x = y$ in (b) and we bear in mind that $\|g\|_\infty < \varepsilon/2$ and the function b is nonincreasing, we get that

$$v(z) - u(z) \geqslant v(y_0) - u(x_0) + b(0) - b(d(x_0, y_0)) + g(z,z) - g(x_0, y_0)$$
$$\geqslant v(y_0) - u(x_0) + 0 - \varepsilon/2 - \varepsilon/2 = v(y_0) - u(x_0) - \varepsilon,$$

which shows (iii) and finishes the proof. □

Remark 6.7. Let us observe that the preceding proposition is no longer true if the manifold is not complete. For example: $M = (0,1) \subset \mathbb{R}$, $g(x) = x$, $f(x) = 0$, and $\varepsilon > 0$ small.

Definition 6.8. For a given open set $\Omega \subset M$ and a function $u : \Omega \to \mathbb{R}$, we define the upper semicontinuous envelope of u, which we denote u^*, by

$$u^*(x) = \inf\{v(x) \mid v : \Omega \to \mathbb{R} \text{ is continuous and } u \leqslant v \text{ on } \Omega\} \quad \text{for any } x \in \Omega.$$

In a similar way we define the lower semicontinuous envelope, denoted by u_*.

Proposition 6.9. *Let Ω be an open subset of M. Let $\mathcal{F}$ be a uniformly bounded family of upper semicontinuous functions on Ω, and let $u = \sup\{v : v \in \mathcal{F}\}$. Then, for every $p \in \Omega$, and every $\zeta \in D^+u^*(p)$, there exist sequences $\{v_n\} \subset \mathcal{F}$, and $\{(p_n, \zeta_n)\} \subset T^*(\Omega)$, with $\zeta_n \in D^+v_n(p_n)$ and such that*

(1) $\lim_n v_n(p_n) = u^*(p)$
(2) $\lim_n (p_n, \zeta_n) = (p, \zeta)$.

Proof. Fix a chart (U, φ), with $p \in U$. Let us consider the family $\mathcal{F} \circ \varphi^{-1} = \{v \circ \varphi^{-1} : v \in \mathcal{F}\}$. The functions of this collection are upper semicontinuous on $\varphi(U \cap \Omega)$, and the family is uniformly bounded. On the other hand $u \circ \varphi^{-1} = sup\{v \circ \varphi^{-1} : v \in \mathcal{F}\}$, and $u^* \circ \varphi^{-1} = (u \circ \varphi^{-1})^*$. Now apply [27, Proposition VIII.1.6] (which is nothing but the result we want to prove in the case of a Banach space) to the Hilbert space, the open set $\varphi(U \cap \Omega)$, the family $\mathcal{F} \circ \varphi^{-1}$, the point $\varphi(p)$, and the superdifferential

$\zeta \circ d\varphi(p)^{-1} \in D^{+}(u \circ \varphi^{-1})(\varphi(p))$. We get sequences $\{\varphi(p_n)\}$ in $\varphi(U \cap \Omega)$, $\{v_n \circ \varphi^{-1}\}$ in $\mathcal{F} \circ \varphi^{-1}$, and $\zeta_n \circ d\varphi(p)^{-1} \in D^{+}(v_n \circ \varphi^{-1})(\varphi(p_n))$ such that $\lim_n \varphi(p_n) = \varphi(p)$, $\lim_n \zeta_n \circ d\varphi(p)^{-1} = \zeta \circ d\varphi(p)^{-1}$, and

$$\lim_n (v_n \circ \varphi^{-1})(\varphi(p_n)) = (u \circ \varphi^{-1})^*(\varphi(p)) = u^* \circ \varphi^{-1}(\varphi(p)).$$

Hence $\lim_n p_n = p$, $\lim_n v_n(p_n) = u^*(p)$, and

$$\lim_n \zeta_n \circ d\varphi(p_n)^{-1} = \lim_n \zeta_n \circ d\varphi(p_n)^{-1} \circ d\varphi(p) \circ d\varphi(p)^{-1} = \lim_n \zeta_n \circ d\varphi(p)^{-1}$$

because φ is C^1, so $\lim_n d\varphi(p_n)^{-1} \circ d\varphi(p) = \mathrm{id}$. The result follows trivially from the local representation of the cotangent bundle. □

Now we introduce the notion of uniform continuity that we have to require of $F : T^*M \to \mathbb{R}$ in order to prove the existence and uniqueness of viscosity solutions to the Hamilton–Jacobi equation $(*)$.

Definition 6.10. We will say that a function $F : T^*M \to \mathbb{R}$ is intrinsically uniformly continuous provided that for every $\varepsilon > 0$ there exists $\delta \in (0, r_M)$ such that

$$d(x, y) \leqslant \delta,\ \zeta \in T^*M_x,\ \xi \in T^*M_y,\ \|\zeta - L_{yx}(\xi)\|_x \leqslant \delta \Longrightarrow |F(x, \zeta) - F(y, \xi)| \leqslant \varepsilon.$$

Remark 6.11. It should be noted that if F satisfies the above definition then F is continuous. This is obvious once we notice that the mapping

$$(x, \zeta) \in T^*M_x \mapsto L_{xx_0}(\zeta)$$

is continuous at (x_0, ζ_0), that is, if $(x_n, \zeta_n) \to (x_0, \zeta_0)$ in T^*M then $L_{x_n x_0}(\zeta_n) \to \zeta_0$, for every $(x_0, \zeta_0) \in T^*M$. The fact that this mapping is continuous is an easy consequence of the definition of the parallel translation along a curve as a solution to an ordinary linear differential equation.

Remark 6.12. Consider a finite-dimensional manifold M embedded in $\mathbb{R}^n$, so $T^*M \subset \mathbb{R}^{2n}$. Assume that M satisfies the following condition:
$\forall \varepsilon\ \exists\ \delta > 0$: $v, h \in TM_x\ \ x \in M$, $\|v\|_x \leqslant \delta$ norm of $h \leqslant 1 \Longrightarrow \|d\exp_x(v)(h) - h\|_x \leqslant \varepsilon$ (note in particular that this condition is automatically met when M is compact, and in many other natural examples). Then every function $F : T^*M \to \mathbb{R}$ which is uniformly continuous with respect to the usual Euclidean metric in $\mathbb{R}^{2n}$ is intrinsically uniformly continuous as well, as is easily seen. Consequently there are lots of natural examples of intrinsically uniformly continuous functions $F : T^*M \to \mathbb{R}$.

Now we can prove the following *maximum principle* for Hamilton–Jacobi equations of the form $(*)$.

Theorem 6.13. *Let $f, g : M \to \mathbb{R}$ be bounded uniformly continuous functions, and $F : T^*M \to \mathbb{R}$ be intrinsically uniformly continuous. If u is a bounded viscosity subsolution of $u+F(du) = f$ and v is a bounded viscosity supersolution of $v+F(dv) = g$, then $v - u \geqslant \inf(g - f)$.*

Proof. If $\varepsilon > 0$ is given, then, by Proposition 6.6, there exist $p, q \in U$, and $\zeta \in D^+u(p)$, $\xi \in D^-v(q)$ such that

(i) $d(p, q) < \varepsilon$, $\|\zeta - L_{qp}(\xi)\|_p < \varepsilon$
(ii) $v(x) - u(x) \geqslant v(q) - u(p) - \varepsilon$ for every $x \in M$.

Since u and v are viscosity sub and super solutions respectively, we have $u(p) + F(p, \zeta) \leqslant f(p)$ and $v(q) + F(q, \xi) \geqslant g(q)$. Hence, for every $x \in M$,

$$\begin{aligned} v(x) - u(x) &\geqslant v(q) - u(p) - \varepsilon \geqslant g(q) - F(q, \xi) - f(p) + F(p, \zeta) - \varepsilon \\ &\geqslant \inf(g - f) + (f(q) - f(p)) + (F(p, \zeta) - F(q, \xi)) - \varepsilon. \end{aligned}$$

Now, if we let $\varepsilon \to 0^+$, we have that $f(q) - f(p)$ goes to 0 because f is uniformly continuous. On the other hand, the fact that F is intrinsically uniformly continuous implies that $F(p, \zeta) - F(q, \xi)$ goes to 0 as $\varepsilon \to 0^+$. Consequently we obtain $v - u \geqslant \inf(g - f)$. □

Remark 6.14. In fact, an obvious modification of the above proof yields the following result on continuous dependence of viscosity solutions of equations of the form (∗) with respect to the Hamiltonians F. Namely, let $F, G : T^*M \to \mathbb{R}$ be intrinsically uniformly continuous Hamiltonians, and define $\|F - G\|_\infty = \sup\{|F(p, \zeta) - G(p, \zeta)| : (p, \zeta) \in T^*M\}$. If u and v are viscosity solutions of $u+F(du) = 0$ and $v+G(dv) = 0$, respectively, then $|v(x) - u(x)| \leqslant \|F - G\|_\infty$ for every $x \in M$.

Proposition 6.15. *Let Ω be an open subset of M. Let $\mathcal{F}$ be a uniformly bounded family of functions on Ω, and let $u = \sup\{v : v \in \mathcal{F}\}$. If every v is a viscosity subsolution of $u + F(du) = 0$, then u^* is a viscosity subsolution of $u + F(du) = 0$.*

Proof. Let $p \in \Omega$ and $\zeta \in D^+u^*(p)$. According to Proposition 6.9, there exist sequences $\{v_n\} \subset \mathcal{F}$, and $\{(p_n, \zeta_n)\} \subset T^*(\Omega)$ with $\zeta_n \in D^+v_n(p_n)$ and such that

(i) $\lim_n v_n(p_n) = u^*(p)$
(ii) $\lim_n (p_n, \zeta_n) = (p, \zeta)$.

Since v_n are viscosity subsolutions of $u + F(du) = 0$, we have that $v_n(p_n) + F(p_n, \zeta_n) \leqslant 0$. Hence $u^*(p) + F(p, \zeta) \leqslant 0$. □

Corollary 6.16. *The supremum of two viscosity subsolutions is a viscosity subsolution.*

Theorem 6.17. *Let M be a complete Riemannian manifold which is uniformly locally convex and has a strictly positive injectivity radius. Let $F : T^*M \to \mathbb{R}$ be an intrin-*

sically uniformly continuous function (*see Definition* 6.10). *Assume also that there is a constant* $A > 0$ *so that* $-A \leqslant F(x, 0_x) \leqslant A$ *for every* $x \in M$. *Then, there exists a unique bounded viscosity solution of the equation* $u + F(du) = 0$.

Proof. Uniqueness follows from Theorem 6.13, by taking $f = g = 0$. In order to show existence, let us define $\mathcal{F}$ as the family of the viscosity subsolutions $w : M \to \mathbb{R}$ to $u + F(du) = 0$ that satisfy

$$-A \leqslant w(p) \leqslant A \quad \text{for every} \quad p \in M.$$

The family $\mathcal{F}$ is nonempty, as the function $w_0(p) = -A$ belongs to $\mathcal{F}$ (because $-A + F(p, 0_p) \leqslant 0$). Let u be the upper semicontinuous envelope of $\sup\{w : w \in \mathcal{F}\}$, and v be the lower semicontinuous envelope of u. By the definition, we have $v \leqslant u$. On the other hand, according to Proposition 6.15, u is a viscosity subsolution of $u + F(du) = 0$.

Claim 6.18. *v is a viscosity supersolution of $u + F(du) = 0$.*

Once we have proved the claim, we have that $u \leqslant v$ by Proposition 6.13, hence $u = v$ is a viscosity solution, and existence is established.

So let us prove the claim. If v is not a viscosity supersolution, there exist $p_0 \in M$ and $\zeta_0 \in D^- v(p)$ such that $v(p_0) + F(p_0, \zeta_0) < 0$. By Theorem 4.3(5), there exists a C^1 smooth function $h : M \to \mathbb{R}$ with $\zeta_0 = dh(p_0)$ and such that $v - h$ attains a global minimum at p_0. Hence we may assume that

$$v(p_0) + F(p_0, dh(p_0)) < 0, \quad v(p_0) = h(p_0), \quad \text{and} \quad h(p) \leqslant v(p) \quad \text{for all } p \in M. \tag{6.8}$$

From the inequality $h(p) \leqslant v(p) \leqslant u(p) \leqslant A$ we get $h(p_0) < A$: otherwise $A - h$ would have a local minimum at p_0, and consequently $dh(p_0) = 0$, which implies $v(p_0) + F(p_0, dh(p_0)) = h(p_0) + F(p_0, dh(p_0)) = A + F(p_0, 0) \geqslant A - A = 0$, a contradiction with (6.8).

Now we can take a number $\delta > 0$ and a C^1 smooth function $b : M \to [0, \infty)$ with support on $B(p_0, \delta)$, $b(p_0) > 0$, and such that $||b||_\infty$, $||db||_\infty$ are small enough so that

$$h(p) + b(p) + F(p, dh(p) + db(p)) < 0 \quad \text{for every} \quad p \in B(p_0, 2\delta) \tag{6.9}$$

and

$$h(p) + b(p) \leqslant A \quad \text{for every} \quad p \in M. \tag{6.10}$$

This is possible because of (6.8) and the fact that F is continuous.

Let us consider the following function:

$$w(p) = \begin{cases} \max\{h(p)+b(p), u(p)\} & \text{if} \quad p \in B(p_0, 2\delta), \\ u(p) & \text{otherwise.} \end{cases}$$

We have that $w(p) = u(p)$ if $p \in \Omega_1 := M \setminus \overline{B}(p_0, \delta)$, because $u(p) \geqslant v(p) \geqslant h(p) = h(p) + b(p)$ whenever $p \in B(p_0, 2\delta) \setminus \overline{B}(p_0, \delta)$. Therefore w is a viscosity subsolution of $u + F(du) = 0$ on Ω_1. On the other hand, bearing (6.9) in mind, it is clear that w is the maximum of two viscosity subsolutions on $\Omega_2 := B(p_0, 2\delta)$, and consequently w is a viscosity subsolution on Ω_2. Therefore w is a viscosity subsolution of $u + F(du) = 0$ on $M = \Omega_1 \cup \Omega_2$. This implies that $w \in \mathcal{F}$, since $-A \leqslant u \leqslant w$ and $w(p) \leqslant A$, by (6.10).

Finally, we have that $u \geqslant w$, because $u \geqslant \sup \mathcal{F}$. Therefore we have $u(p) \geqslant w(p) \geqslant h(p) + b(p)$ on $B(p_0, \delta)$ and in particular $v(p_0) = u_*(p_0) \geqslant h(p_0) + b(p_0) > h(p_0)$, which contradicts (6.8). □

When M is compact, the preceding Theorem takes on a simpler appearance.

Corollary 6.19. *Let M be a compact Riemannian manifold, $f : M \to \mathbb{R}$ a continuous function, and $F : T^*M \to \mathbb{R}$ an intrinsically uniformly continuous function. Then there exists a unique viscosity solution of the equation $u + F(du) = f$.*

Proof. This follows immediately from Theorem 6.17, taking into account the following facts: (1) if M is compact then M is uniformly locally convex and $i(M) > 0$ (see Remarks 2.9 and 2.12); (2) every viscosity solution u is continuous, hence u is bounded on the compact manifold M; and of course (3) f is uniformly continuous because f is continuous on M, compact. □

Remark 6.20. In particular, when a compact manifold M is regarded as embedded in $\mathbb{R}^n$, so $T^*M \subset \mathbb{R}^{2n}$, and $F : T^*M \to \mathbb{R}$ is uniformly continuous with respect to the usual Euclidean metric in $\mathbb{R}^{2n}$, then Corollary 6.19 and Remark 6.12 yield the existence of a unique viscosity solution to the equation $u + F(du) = f$.

However, the requirement that F is uniformly continuous cannot be relaxed in principle, because the cotangent bundle T^*M is never compact, so, even though F is continuous, we cannot ensure that F is uniformly continuous on T^*M.

Remark 6.21. It should be noted that one may pose a Hamilton–Jacobi problem such as

$$u + F(du) = 0, \quad u \text{ bounded} \tag{$*$}$$

on a manifold M without presupposing any Riemannian structure defined on M. Then one may consider the natural question whether it is possible to find a suitable Riemannian structure g which makes M uniformly locally convex and with a positive injectivity radius, and which makes F intrinsically uniformly continuous as well. In other words, one can seek for a Riemannian manifold N with positive convexity and injectivity radii and a diffeomorphism $\psi : N \to M$ so that the function $G : T^*N \to \mathbb{R}$ is intrinsically uniformly continuous, where $G = F \circ (T^*\psi)$, with $T^*\psi : T^*N \to T^*M$ defined by $T^*\psi(x,\eta) := \Big(\psi(x), \eta \circ (d\psi^{-1})(\psi(x))\Big)$. Then, by Theorem 6.17, the equation

$$v + G(dv) = 0, \quad v \text{ bounded} \tag{$**$}$$

has a unique viscosity solution. But it is obvious that v is a viscosity solution to $(**)$ if and only if the function $u = v \circ \psi^{-1}$ is a viscosity solution to $(*)$. Hence the equation $(*)$ has a unique viscosity solution as well. This means that Theorem 6.17 above is applicable to even more situations than one might think of at a single glance. The following example reveals the power of this scheme.

Example 6.22. Let M be the submanifold of $\mathbb{R}^3$ defined by

$$z = \frac{1}{x^2 + y^2},$$

let $F : T^*M \subset \mathbb{R}^6 \to \mathbb{R}$, and consider the Hamilton–Jacobi equation $u + F(du) = 0$. If we endow M with its natural Riemannian structure inherited from $\mathbb{R}^3$, M will not be uniformly locally convex, and besides $i(M) = 0$, so Theorem 6.17 is not directly applicable. Now let us define N by

$$z = \frac{1}{x^2 + y^2 - 1}, \quad z > 0,$$

with its natural Riemannian structure as a submanifold of $\mathbb{R}^3$. It is clear that N is uniformly locally convex and has a positive injectivity radius. The mapping $\psi : N \to M$ defined by

$$(x, y, z) = \left(\frac{\sqrt{x^2 + y^2 - 1}}{\sqrt{x^2 + y^2}}\, x,\ \frac{\sqrt{x^2 + y^2 - 1}}{\sqrt{x^2 + y^2}}\, y,\ z \right)$$

is a C^∞ diffeomorphism. Assume that the function $G = F \circ (T^*\psi) : T^*N \to \mathbb{R}$ is uniformly continuous with respect to the usual metric in $\mathbb{R}^6$. Since N satisfies the property of Remark 6.12, we have that G is intrinsically uniformly continuous. Therefore, by the preceding Remark 6.21, the equation $u + F(du) = 0$, u bounded on M, has a unique viscosity solution.

Now let us see how Deville's mean value Theorem 4.18 allows to deduce a result on the regularity of viscosity solutions (or even subsolutions) to Hamilton–Jacobi equations with a "coercive" structure.

Corollary 6.23. *Let M be a Riemannian manifold, and $F : \mathbb{R} \times T^*M \to \mathbb{R}$ a function. Consider the following Hamilton–Jacobi equation*:

$$F(u(x), du(x)) = 0. \tag{HJ3}$$

Assume that there exists a constant $K > 0$ such that $F(t, \zeta_x) > 0$ whenever $\|\zeta_x\|_x \geqslant K$ and $t \in \mathbb{R}$. Let $u : M \to \mathbb{R}$ be a viscosity subsolution of (HJ3). *Then*:

(1) *u is K-Lipschitz, that is, $|u(x) - u(y)| \leqslant K d(x, y)$ for every $x, y \in M$.*
(2) *If M is finite dimensional, u is Fréchet differentiable almost everywhere.*
(3) *If M is infinite-dimensional u is Fréchet differentiable on a dense subset of M.*

Proof. If u is a viscosity subsolution then $F(u(x), \zeta_x) \leqslant 0$ for every $x \in M$ and $\zeta_x \in D^- u(x)$. Hence $\|\zeta_x\| \leqslant K$ for every $\zeta_x \in D^- u(x)$ (otherwise $F(u(x), \zeta_x) > 0$, a contradiction). Then, by Theorem 4.18, u is K-Lipschitz.

On the other hand, (2) and (3) follow immediately from Theorem 5.7. □

Let us conclude with a brief study of a HJ equation which is not of the form (∗) above, but which is still very interesting because of the geometrical significance of its unique viscosity solution. Let M be a complete Riemannian manifold, Ω a bounded open subset of M, and let $\partial\Omega$ be the boundary of Ω. Consider the Hamilton–Jacobi equation

$$\text{(HJ4)} \quad \begin{cases} \|du(x)\|_x = 1 & \text{for all } x \in \Omega, \\ u(x) = 0 & \text{for all } x \in \partial\Omega. \end{cases}$$

There is no classical solution of (HJ4). Indeed, if we had a function $u : \overline{\Omega} \subset M \to \mathbb{R}$ which is differentiable on Ω and satisfies $\|du(x)\|_x = 1$ for $x \in \Omega$ and $u = 0$ on $\partial\Omega$, then we could apply Theorem 3.1 to find a point $x_0 \in \Omega$ so that $\|du(x_0)\|_{x_0} < \frac{1}{2}$, a contradiction.

Nevertheless, we are going to see that there is a unique viscosity solution to (HJ4), namely the distance function to the boundary $\partial\Omega$. By definition, a function u is a viscosity solution to (HJ4) if and only if u is continuous; $u = 0$ on $\partial\Omega$; $\|\zeta\|_x \geqslant 1$ for all $\zeta \in D^- u(x)$, $x \in \Omega$; and $\|\zeta\|_x \leqslant 1$ for all $\zeta \in D^+ u(x)$, $x \in \Omega$.

Theorem 6.24. *Let M be a complete Riemannian manifold, and Ω a bounded open subset of M with boundary $\partial\Omega$. Then the function $x \mapsto d(x, \partial\Omega) := \inf\{d(x, y) : y \in \partial\Omega\}$ is a viscosity solution of Eq.* (HJ4). *Moreover, if M is uniformly locally convex and has a positive injectivity radius, then $d(\cdot, \partial\Omega)$ is the* unique *viscosity solution of this equation.*

Proof. Let us first check uniqueness. Assume $u, v : \overline{\Omega} \to \mathbb{R}$ are viscosity solutions of (HJ4). Since u and v are continuous, and $u = v = 0$ on $\partial\Omega$, we can extend u and v with continuity to the whole of M by setting $u = 0 = v$ on $M \setminus \Omega$. It is enough to see that $u \leqslant v$ on Ω (in a similar way, or by symmetry, $v \leqslant u$, hence $u = v$). To this end we take any $\alpha \in (0, 1)$ and we check that $\alpha u \leqslant v$. Indeed, suppose we had that $\inf\{v(x) - \alpha u(x) : x \in \Omega\} < 0$. Pick ε with

$$0 < 2\varepsilon < \min \left\{ \frac{1-\alpha}{2}, \, -\inf\{v(x) - \alpha u(x) : x \in \Omega\} \right\}.$$

Note that, as u and v are viscosity solutions, we have $\|\zeta\|_x \leqslant 1$ for every $\zeta \in D^+u(x) \cup D^+v(x)$, $x \in \Omega$, so by the mean value Theorem 4.18 u and v are 1-Lipschitz. In particular, since Ω is bounded we have that u and v are bounded. Then, according to Proposition 6.6, there exist $x_0, y_0 \in M$, $\zeta \in D^+(\alpha u)(x_0)$, $\xi \in D^-v(y_0)$ with

(1) $d(x_0, y_0) < \varepsilon$
(2) $\|\zeta - L_{x_0 y_0}(\xi)\|_{x_0} < \varepsilon$
(3) $\inf(v - \alpha u) \geqslant v(y_0) - \alpha u(x_0) - \varepsilon$.

Taking into account the facts that u and v are 1-Lipschitz, and $u = v = 0$ on $M \setminus \Omega$, it is easy to see that (3) and the choice of ε imply that $x_0, y_0 \in \Omega$. Now, since u and v are viscosity solutions we have that

$$\frac{1}{\alpha}\zeta \in D^+u(x_0) \Longrightarrow \left\| \frac{1}{\alpha}\zeta \right\|_{x_0} \leqslant 1 \Longrightarrow \|\zeta\|_{x_0} \leqslant \alpha$$

and

$$\xi \in D^-v(y_0) \Longrightarrow \|\xi\|_{y_0} \geqslant 1.$$

Now, from (2), and bearing in mind that $L_{x_0 y_0}$ is a linear isometry, we get that

$$1 \leqslant \|\xi\|_{y_0} = \|L_{x_0 y_0}(\xi)\|_{x_0} \leqslant \|\zeta\|_{x_0} + \varepsilon \leqslant \alpha + \varepsilon < 1,$$

a contradiction.

Now let us prove that $u := d(\cdot, \partial\Omega)$ is a viscosity solution to (HJ4), hence the only one. The property $u = 0$ on $\partial\Omega$ is obvious from the definition, so we only have to check the conditions on the norms of the vectors of $D^-u(x)$ and $D^+u(x)$, for $x \in \Omega$.

Step 1: Take $\xi \in D^-u(x)$, $x \in \Omega$. We have to see that $\|\xi\|_x \geqslant 1$. By Theorem 4.3 we can pick a C^1 smooth function $\varphi : M \to \mathbb{R}$ so that $u(y) - \varphi(y) \geqslant u(x) - \varphi(x) = 0$ for all $y \in M$. Fix $0 < \varepsilon < 1$. Now, for every α with $0 < \alpha < d(x, \partial\Omega)$, by the definition of $d(x, \partial\Omega)$ we can take $x_\alpha \in \partial\Omega$ with

$$d(x, \partial\Omega) \geqslant d(x, x_\alpha) - \frac{\varepsilon\alpha}{4}.$$

Next, by making use of Ekeland's approximate Hopf–Rinow type Theorem 2.2, we can find a point $y_\alpha \in \Omega$ with

$$d(x_\alpha, y_\alpha) < \frac{\varepsilon\alpha}{4}$$

and a geodesic $\gamma_\alpha : [0, T_\alpha] \to \overline{\Omega} \subset M$ joining $x = \gamma_\alpha(0)$ to $y_\alpha = \gamma_\alpha(T_\alpha)$, and such that $L(\gamma_\alpha) = d(x, y_\alpha)$. Then we have

$$L(\gamma_\alpha) = d(x, y_\alpha) \leqslant d(x, x_\alpha) + d(x_\alpha, y_\alpha) \leqslant d(x, x_\alpha) + \frac{\varepsilon\alpha}{4} \leqslant d(x, \partial\Omega) + \frac{\varepsilon\alpha}{2},$$

that is

$$d(x, \partial\Omega) \geqslant L(\gamma_\alpha) - \frac{\varepsilon\alpha}{2}. \tag{6.11}$$

Set $v_\alpha = d\gamma_\alpha(0)/dt \in TM_x$, so that $\gamma_\alpha(t) = \exp_x(tv_\alpha)$ and $\|v_\alpha\|_x = 1$, and define $z_\alpha = \gamma_\alpha(\alpha)$. Then we have

$$\begin{aligned}\varphi(z_\alpha) - \varphi(x) &\leqslant u(z_\alpha) - u(x) = d(z_\alpha, \partial\Omega) - d(x, \partial\Omega)\\ &\leqslant d(z_\alpha, \partial\Omega) - L(\gamma_\alpha) + \frac{\varepsilon\alpha}{2} \leqslant d(z_\alpha, y_\alpha) + d(y_\alpha, x_\alpha) - L(\gamma_\alpha) + \frac{\varepsilon\alpha}{2}\\ &\leqslant L(\gamma_{\alpha|_{[\alpha, T_\alpha]}})) + \frac{\varepsilon\alpha}{2} - L(\gamma_\alpha) + \frac{\varepsilon\alpha}{2}\\ &= L(\gamma_\alpha) - \alpha + \frac{\varepsilon\alpha}{2} - L(\gamma_\alpha) + \frac{\varepsilon\alpha}{2} = \alpha(-1 + \varepsilon),\end{aligned}$$

hence

$$\frac{\varphi(z_\alpha) - \varphi(x)}{\alpha} \leqslant -1 + \varepsilon. \tag{6.12}$$

By the mean value theorem there is $s_\alpha \in [0, \alpha]$ such that

$$d\varphi(\gamma_\alpha(s_\alpha)) \left(\frac{d\gamma_\alpha(s_\alpha)}{dt} \right) = \frac{\varphi(z_\alpha) - \varphi(x)}{\alpha}. \tag{6.13}$$

By combining (6.12) and (6.13), and bearing in mind that $\|d\gamma_\alpha(s)/dt\|_{\gamma_\alpha(s)} = 1$ for all s, we get that

$$\|d\varphi(\gamma_\alpha(s_\alpha))\|_{\gamma_\alpha(s_\alpha)} \geqslant 1 - \varepsilon \tag{6.14}$$

for every $\alpha \in (0, u(x))$. Then, since the functions $y \to d\varphi(y)$ and $(y, \zeta) \to \|\zeta\|_y$ are continuous, and $\gamma_\alpha(s_\alpha) = \exp_x(s_\alpha v_\alpha) \to \exp_x(0) = x$ as $\alpha \to 0$, it follows that

$$\|\xi\|_x = \|d\varphi(x)\|_x = \lim_{\alpha \to 0^+} \|d\varphi(\gamma_\alpha(s_\alpha))\|_{\gamma_\alpha(s_\alpha)} \geqslant 1 - \varepsilon. \tag{6.15}$$

Finally, by letting $\varepsilon \to 0$ in (6.15), we deduce that $\|\xi\|_x \geqslant 1$.

Step 2: Now take $\zeta \in D^+u(x)$, $x \in \Omega$, and let us see that $\|\zeta\|_x \leqslant 1$. This is much easier. Pick a C^1 smooth function $\psi : M \to \mathbb{R}$ so that $d\psi(x) = \zeta$ and $u(y) - \psi(y) \leqslant u(x) - \psi(x) = 0$ for all $y \in M$. For each $v \in TM_x$ consider the geodesic $\gamma_v(t) = \exp_x(tv)$. Since $u = d(\cdot, \partial\Omega)$ is 1-Lipschitz we have that

$$\psi(\gamma_v(t)) - \psi(\gamma_v(0)) \geqslant u(\gamma_v(t)) - u(\gamma_v(0)) \geqslant -d(\gamma_v(t), \gamma_v(0)) = -t,$$

hence

$$\frac{\psi(\gamma_v(t)) - \psi(\gamma_v(0))}{t} \geqslant -1$$

for all $t > 0$ small enough, and

$$d\psi(x)(v) = \lim_{t \to 0^+} \frac{\psi(\gamma_v(t)) - \psi(\gamma_v(0))}{t} \geqslant -1. \tag{6.16}$$

As (6.16) holds for every $v \in TM_x$, we conclude that $\|\zeta\|_x = \|d\psi(x)\|_x \leqslant 1$. $\square$

Acknowledgments

We thank Luis Guijarro and Ildefonso Díaz for several valuable discussions concerning the results of this paper.

D.A. was supported by a Marie Curie Fellowship of the European Community Training and Mobility of Researchers Programme under Contract Number HPMF-CT-2001-01175. J.F. was partially supported by BFM Grant 2000/0609. F.L.-M. enjoyed a study license granted by the Consejería de Educación de la Comunidad de Madrid.

References

[1] R. Abraham, J.E. Marsden, Foundations of Mechanics, second ed., The Benjamin/Cummings Publishing Company, Inc., Reading, MA, 1980.

[2] D. Aussel, J.-N. Corvellec, M. Lassonde, Mean value property and subdifferential criteria for lower semicontinuous functions, Trans. Amer. Math. Soc. 347 (1995) 4147–4161.

[3] D. Azagra, Smooth negligibility and subdifferential calculus in Banach spaces, with applications, Ph.D. Thesis, Universidad Complutense de Madrid, October 1997.

[4] D. Azagra, M. Cepedello Boiso, Uniform approximation of mappings by smooth mappings with no critical points on Hilbert manifolds, Duke Math. J. 124 (1) (2004) 47–66.
[5] D. Azagra, R. Deville, Subdifferential Rolle's and mean value inequality theorems, Bull. Austral. Math. Soc. 56 (1997) 319–329.
[6] D. Azagra, J. Gómez, J.A. Jaramillo, Rolle's theorem and negligibility of points in infinite-dimensional Banach espaces, J. Math. Anal. Appl. 213 (1997) 487–495.
[7] D. Azagra, M. Jiménez-Sevilla, The failure of Rolle's theorem in infinite-dimensional Banach spaces, J. Funct. Anal. 182 (2001) 207–226.
[8] G. Barles, Solutions de viscosité des équations de Hamilton–Jacobi, Mathematics and Applications, Paris, Springer, Berlin, 1994.
[9] J. Bes, J. Ferrera, On a multidimensional version of Rolle's theorem, Publ. Dto. Análisis Mat. UCM, Ser. 1 no. 42, 1996–1997, pp. 21–27.
[10] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, P.R. Wolenski, Nonsmooth Analysis and Control Theory, Graduate Texts in Mathematics, vol. 178, 1998 Springer, Berlin.
[11] M.G. Crandall, L.C. Evans, P.L. Lions, Some properties of viscosity solutions to Hamilton–Jacobi equations, Trans. Amer. Math. Soc. 282 (1984) 487–502.
[12] M.G. Crandall, H. Ishii, P.L. Lions, Uniqueness of viscosity solutions revisited, J. Math. Soc. Japan 39 (1987) 581–596.
[13] M.G. Crandall, H. Ishii, P.L. Lions, User's guide to viscosity solutions of second order to fully nonlinear partial differential equations, Bull. Amer. Math. Soc. 27 (1992) 1–67.
[14] M.G. Crandall, P.L. Lions, Solutions de viscosité non bornées des équations de Hamilton–Jacobi du premier ordre, C. R. Acad. Sci. Paris 198 (1981) 217–220.
[15] M.G. Crandall, P.L. Lions, Viscosity solutions of Hamilton–Jacobi equations, Trans. Amer. Math. Soc. 277 (1983) 1–42.
[16] M.G. Crandall, P.L. Lions, Condition d'unicité pour les solutions généralisées des équations de Hamilton–Jacobi du premier ordre, C. R. Acad. Sci. Paris 292 (1984) 183–186.
[17] M.G. Crandall, P.L. Lions, Hamilton–Jacobi equations in infinite dimensions, Part I: uniqueness of viscosity solutions, J. Funct. Anal. 62 (1985) 379–396;
M.G. Crandall, P.L. Lions, Hamilton–Jacobi equations in infinite dimensions, Part II: existence of viscosity solutions, J. Funct. Anal. 65 (1986) 368–405;
M.G. Crandall, P.L. Lions, Hamilton–Jacobi equations in infinite dimensions, Part III, J. Funct. Anal. 68 (1986) 214–247;
M.G. Crandall, P.L. Lions, Hamilton–Jacobi equations in infinite dimensions, Part IV: ubounded linear terms, J. Funct. Anal. 90 (1990) 237–283;
M.G. Crandall, P.L. Lions, Hamilton–Jacobi equations in infinite dimensions, Part V: B-continuous solutions, J. Funct. Anal. 97 (1991) 417–465.
[18] M.G. Crandall, P.L. Lions, On existence and uniqueness of solutions of Hamilton–Jacobi equations, Non. Anal. Theor. Math. Appl. 10 (1986) 353–370.
[19] M.G. Crandall, P.L. Lions, Unbounded viscosity solutions of Hamilton–Jacobi equations, Illinois J. Math. 31 (1987) 665–688.
[20] M.G. Crandall, P.L. Lions, Solutions de viscosité pour les équations de Hamilton–Jacobi en dimension infinie intervenant dans le contrôle optimal de problèmes d'évolution, C. R. Acad. Sci. Paris 305 (1987) 233–236.
[21] R. Deville, Stability of subdifferentials of nonconvex functions in Banach spaces, Set-Valued Anal. 2 (1994) 141–157.
[22] R. Deville, A mean value theorem for the non differentiable mappings, Serdica Math. J. 21 (1995) 59–66.
[23] R. Deville, Smooth variational principles and nonsmooth analysis in Banach spaces, in: F.H. Clarke, R.J. Stern (Eds.), Nonlinear Analysis, Differential Equations and Control, NATO Sci. Ser. C Math. Phys. Sci., vol. 528, Kluwer Academic Publishers, Dordrecht, 1999, pp. 369–405.
[24] R. Deville, E.El. Haddad, The viscosity subdifferential of the sum of two functions in Banach spaces, I, First order case, J. Convex Anal. 3 (2) (1996) 295–308.
[25] R. Deville, E.M. El Haddad, The subdifferential of the sum of two functions in Banach spaces, II, Second order case, Bull. Austral. Math. Soc. 51 (1995) 235–248.

[26] R. Deville, N. Ghoussoub, Perturbed minimization principles and applications, in: W.B. Johnson, J. Lindenstrauss (Eds.), Handbook of the Geometry of Banach Spaces, vol. 1, North-Holland, Elsevier Science, Amsterdam, 2001, pp. 393–435 (Chapter 10).
[27] R. Deville, G. Godefroy, V. Zizler, Smoothness and renormings in Banach spaces, Pitman Monographies and Surveys in Pure and Applied Mathematics, vol. 64, Pitman, London, 1993.
[28] R. Deville, M. Ivanov, Smooth variational principles with constraints, Arch. Math. 69 (1997) 418–426.
[29] I. Ekeland, The Hopf–Rinow theorem in infinite dimension, J. Differential Geom. 13 (2) (1978) 287–301.
[30] I. Ekeland, Nonconvex minimization problems, Bull. Amer. Math. Soc. 1(3) (New series) (1979) 443-474.
[31] H. Federer, Geometric measure theory, Geometric measure theory, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer, New York, NY, 1969.
[33] G. Godefroy, Some remarks on subdifferential calculus, Rev. Mat. Complutense 11 (2) (1998) 269 –279.
[34] R.E. Greene, K. Shiohama, Convex functions on complete noncompact manifolds: topological structure, Invent. Math. 63 (1981) 129–157.
[35] R.E. Greene, K. Shiohama, Convex functions on complete noncompact manifolds: differentiable structure, Ann. Sci. École. Norm. Sup. 4 série 14 (1981) 357–367.
[36] R.E. Greene, H. Wu, C^∞ convex functions and manifolds of positive curvature, Acta Math. 137 (1976) 209–245.
[37] R.E. Greene, H. Wu, C^∞ approximations of convex, subharmonic, and plurisubharmonic functions, Ann. Sci. École. Norm. Sup. 4 série 12 (1979) 47–84.
[38] E.M. El Haddad, Calcul sous-differentiel et solutions de viscosité des équations de Hamilton–Jacobi, Thèse doctorale présentée à l'Université Bordeaux I en 1994.
[39] W. Klingenberg, Riemannian Geometry, Walter de Gruyter Studies in Mathematics, vol. 1, Walter de Gruyter, Berlin, New York, 1982.
[40] I. Kupka, Counterexample to the Morse–Sard theorem in the case of infinite-dimensional manifolds, Proc. Amer. Math. Soc. 16 (1965) 954–957.
[41] S. Lang, Fundamentals of Differential Geometry, Graduate Texts in Mathematics, vol. 191, Springer, New York, 1999.
[42] J. Milnor, Morse Theory, Ann. of Math. Studies, Princeton University Press, Princeton, 1963.
[43] R.R. Phelps, Convex functions, monotone operators and differentiability, Lecture Notes in Mathematics, vol. 1364, Springer, Berlin, New York, 1993.
[44] D. Preiss, Differentiability of Lipschitz functions on Banach spaces, J. Funct. Anal. 91 (1990) 312–345.