

Inf-Convolution and Regularization of Convex Functions on Riemannian Manifolds of Nonpositive Curvature

Daniel AZAGRA and Juan FERRERA

Departamento de Análisis Matemático
Facultad de Matemáticas
Universidad Complutense
28040 Madrid — Spain
azagra@mat.ucm.es ferrera@mat.ucm.es

Received: December 12, 2005

Accepted: March 27, 2006

ABSTRACT

We show how an operation of inf-convolution can be used to approximate convex functions with C^1 smooth convex functions on Riemannian manifolds with nonpositive curvature (in a manner that not only is explicit but also preserves some other properties of the original functions, such as ordering, symmetries, infima and sets of minimizers), and we give some applications.

Key words: inf-convolution, convex function, regularization, Riemannian manifold.

2000 Mathematics Subject Classification: 52A99, 53C99, 49J52, 46T05.

1. Introduction and main results

Smooth approximation is an old subject. Its importance lies on the fact that most analytical tools for studying the properties of functions defined on a normed space or on a Riemannian manifold require some degree of differentiability of the considered functions. However, many functions which arise naturally from geometrical problems on manifolds are only continuous (or even merely lower semicontinuous). One is thus

D. Azagra was supported by a Marie Curie Intra-European Fellowship of the European Community, Human Resources and Mobility Programme under contract number MEIF CT2003-500927. J. Ferrera was supported by BFM2003-06420.

tempted to approximate those functions by smooth functions to which one can apply more powerful analytical methods and obtain some information about the behavior of the approximations which will hopefully be shared with the original (nonsmooth) functions.

The theory of convex functions is also an old subject which plays an important role in many areas of mathematics. In Riemannian geometry it has been used, for instance, in the investigation of the structure of noncompact manifolds of positive curvature by Cheeger, Gromoll, Greene, Meyer, Siohama, Wu, and others, see [9–11, 13–15]. (Recall that a function $f : M \rightarrow \mathbb{R}$ defined on a Riemannian manifold M is said to be *convex* provided the function $\mathbb{R} \ni t \mapsto f \circ \gamma(t) \in \mathbb{R}$ is convex for every geodesic γ on M .) The existence of global convex functions on a Riemannian manifold has strong geometrical and topological implications; for instance it is shown in [10] 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.

Along with the papers cited above and the references therein, we must mention the important work of Bangert's on convex sets and convex functions on Riemannian manifolds, see [3–5]; he showed in particular that Alexandroff's Theorem about almost everywhere twice differentiability of convex functions on \mathbb{R}^n can be extended to convex functions on finite-dimensional Riemannian manifolds (providing as well a smart proof of Alexandroff's theorem on \mathbb{R}^n).

The aim of the present paper is to study to what extent one of the most useful methods for regularizing convex functions on normed spaces, namely that of infimal convolution, can be successfully used in the setting of Riemannian manifolds.

Let us first have a quick look at the three main methods (that is, partitions of unity, integral convolution with a sequence of mollifiers, and inf and sup convolution formulae) that are used in normed spaces to approximate continuous functions by smooth functions, and see how they can be adapted to the case when one wants to regularize a convex function f defined on a Riemannian manifold M .

Partitions of unity are useless in this setting because, even in the case when $M = \mathbb{R}^n$, they do not preserve convexity of the function f .

The integral convolution with a sequence of mollifiers reveals itself as the perfect tool when $M = \mathbb{R}^n$, because the integral convolution of a convex function f with any integrable function g with compact support, that is,

$$f * g(x) = \int_{\mathbb{R}^n} f(x - y)g(y) dy,$$

is a convex function. In [12–14] Greene and Wu studied to what extent those integral convolutions with mollifiers can be used to regularize convex functions defined on finite-dimensional Riemannian manifolds M (and applied this method to prove several theorems about the structure of complete noncompact manifolds of positive curvature). It turns out that this method works out in Riemannian manifolds only

when the original function f is *strictly convex*. More precisely, let $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ be a nonnegative C^∞ function with support in $[-1, 1]$, constant on a neighborhood of 0 and satisfying $\int_{\mathbb{R}^n} \kappa(\|x\|) dx = 1$, and let us define the functions

$$\varphi_\varepsilon(p) = \frac{1}{\varepsilon^n} \int_{v \in TM_p} f(\exp_p(v)) \kappa\left(\frac{\|v\|_p}{\varepsilon}\right) d\mu_p,$$

where $d\mu_p$ is the measure on the tangent space TM_p obtained from the Riemannian metric of M . Greene and Wu showed that if $f : M \rightarrow \mathbb{R}$ is a convex function defined on an n -dimensional Riemannian manifold M and K is a compact subset of M , then there exists an open neighborhood of K and an $\varepsilon_0 > 0$ such that the functions $\varphi_\varepsilon : U \rightarrow \mathbb{R}$ defined above are C^∞ smooth, converge to f uniformly on K as $\varepsilon \rightarrow 0$, and are *approximately convex* in the sense that

$$\liminf_{\varepsilon \rightarrow 0} \left(\inf \frac{d^2}{dt^2} \varphi_\varepsilon(\gamma(t)) \Big|_{t=0} \right) \geq 0,$$

where the infimum is taken over all geodesics $\gamma(t)$ parameterized by arc length and with $\gamma(0) \in K$.

Now, a C^∞ function $\varphi : M \rightarrow \mathbb{R}$ is called *strictly convex* if its second derivative along any geodesic is strictly positive everywhere on the geodesic. A (not necessarily smooth) function $f : M \rightarrow \mathbb{R}$ is then said to be *strictly convex* provided that for every $p \in M$ and every C^∞ strictly convex function φ defined on a neighborhood of p there is some $\varepsilon > 0$ such that $f - \varepsilon\varphi$ is convex on the neighborhood. With this terminology, the above result implies that if $f : M \rightarrow \mathbb{R}$ is strictly convex then for every compact subset K of M there exists a sequence of strictly convex C^∞ functions $\varphi_n = \varphi_{\varepsilon_n}$ such that $f = \lim_{n \rightarrow \infty} \varphi_n$ uniformly on K .

However, as Greene and Wu pointed out, this method cannot be used when f is not strictly convex, and the problem of approximating (not necessarily strictly) convex functions by smooth convex functions on Riemannian manifolds is open. That is one main limitation of the integral convolution technique on Riemannian manifolds. Another drawback of this method is the fact that it does not apply to functions defined on infinite-dimensional manifolds (even in the case when M is the Hilbert space).

We are left with the third method: infimal convolution. It is well known that if $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous convex function defined in $X = \mathbb{R}^n$ or in any infinite-dimensional reflexive Banach space (such as, for instance, the separable Hilbert space), then the inf-convolution formula

$$f_k(x) := \inf_{y \in X} \{f(y) + k\|x - y\|^2\},$$

where $\|\cdot\|$ is any equivalent norm in X whose dual norm is LUR, defines a sequence of C^1 smooth convex functions f_k which converge to f as $k \rightarrow +\infty$ (uniformly on bounded sets if f is uniformly continuous on bounded sets). We refer to [19] for a

survey on the inf-convolution operation on Banach spaces. In fact, a clever combination of inf- and sup-convolutions allows one to show that if f is a (not necessarily convex) function which is uniformly continuous on bounded sets of a superreflexive Banach space then f can be approximated by C^1 smooth functions with uniformly continuous derivatives uniformly on bounded sets; this was shown first by Lasry and Lions [18] in the case when X is the Hilbert space and then by Cepedello-Boiso [7, 8] for any superreflexive Banach space X .

In this situation it is natural to ask whether infimal convolution formulae can be used to regularize convex functions defined on Riemannian manifolds (either finite or infinite-dimensional). That is the question we try to address in this paper. Let us describe the main results that we will show in the following sections. If M is a complete Riemannian manifold and d is the geodesic distance on M , for any function $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$, and for $\lambda > 0$ we define the function

$$f_\lambda(x) = \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\}$$

for every $x \in M$. In section 2 we collect some general properties of f_λ that do not depend on the geometry of M and that we will need to use later in our proofs. We show for instance how the inf defining $f_\lambda(x)$ can be restricted to a suitable ball $B(x, r_x)$, and then use the estimates on the radius r_x to show that $\lim_{\lambda \rightarrow 0^+} f_\lambda(x) = f(x)$ pointwise whenever $f(x) < +\infty$, and that if f is uniformly continuous on bounded sets then f_λ converges to f uniformly on bounded sets. We also see that f_λ has the same inf and the same set of minimizers as f does, and that f_λ has the same symmetry properties as f (that is, if $T : M \rightarrow M$ is an isometry and $f(Tz) = f(z)$ for all z , then $f_\lambda(Tz) = f_\lambda(z)$ for all z).

In section 3 we assume that $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ is a convex function and we study under what conditions on M the functions f_λ are convex and C^1 smooth. It turns out that some assumptions on the geometry of M are necessary in order that the f_λ be convex and C^1 smooth (see Example 5.4 below); in particular we must require that the distance function $d : M \times M \rightarrow \mathbb{R}$, $(x, y) \mapsto d(x, y)$, be *uniformly locally convex on bounded sets near the diagonal* (see Definition 3.3 below). Under this assumption we prove that the functions f_λ are convex and C^1 smooth on any given bounded subset B of M , for all λ small enough. Moreover, if the distance d is convex on all of $M \times M$ then the f_λ are convex and C^1 smooth on all of M for all $\lambda > 0$.

In section 4 we study the question as to which manifolds satisfy the above technical assumption that d is uniformly locally convex on bounded sets near the diagonal (resp. convex on $M \times M$). First, we show that for every Riemannian manifold with nonpositive sectional curvature, and with the property that the convexity radius function of M is strictly positive on bounded sets (such is the case, for instance, of all complete finite-dimensional Riemannian manifolds), the distance $d : M \times M \rightarrow \mathbb{R}$ is uniformly locally convex on bounded sets near the diagonal. Secondly, we note that, for every Cartan-Hadamard manifold (that is, every simply connected complete

Riemannian manifold of nonpositive sectional curvature), the distance d is convex on all of $M \times M$. By combining these facts with the results of sections 2 and 3 we obtain, in the finite-dimensional case (see Corollary 4.4 below):

If M is a complete finite-dimensional Riemannian manifold with sectional curvature $K \leq 0$ and $f : M \rightarrow \mathbb{R}$ is a convex function, then, for every bounded open convex set U with compact closure \bar{U} , there exists $\lambda_0 > 0$ such that the functions $f_\lambda : M \rightarrow \mathbb{R}$ are convex and C^1 smooth on U for all $\lambda \in (0, \lambda_0)$. Moreover,

- (i) f_λ converges to f uniformly on \bar{U} .
- (ii) $f_\lambda \leq f$ for all $\lambda > 0$.
- (iii) f_λ has the same inf and the same set of minimizers as f .
- (iv) f_λ has the same symmetries as f (that is, if f is invariant with respect to an isometry $T : M \rightarrow M$, then so is f_λ).

And, in the case of a Cartan-Hadamard manifold (either finite-dimensional or infinite-dimensional, see Corollary 4.5 below):

If M is a Cartan-Hadamard manifold and $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ is a lower-semicontinuous convex function, then the functions $f_\lambda : M \rightarrow \mathbb{R}$ are convex and C^1 smooth on all of M for all $\lambda > 0$. Moreover,

- (i) *If f is uniformly continuous on bounded sets then f_λ converges to f uniformly on bounded sets.*
- (ii) $f_\lambda \leq f$ for all $\lambda > 0$.
- (iii) f_λ has the same inf and the same set of minimizers as f .
- (iv) f_λ has the same symmetries as f .

Finally, in section 5 we consider some corollaries and applications of the above results. We show that if C is a closed convex subset of a Cartan-Hadamard manifold then the distance function to C , $x \mapsto d(x, C) = \inf\{d(x, y) : y \in C\}$, is C^1 smooth on $M \setminus C$ and the function $x \mapsto d(x, C)^2$ is C^1 smooth and convex on all of M . We also note that this result is not true for Riemannian manifolds of positive curvature such as the 2-sphere, and therefore the results of section 3 cannot be extended to manifolds of positive curvature. Another consequence is that every closed convex subset of a Cartan-Hadamard manifold can be approximated by C^1 smooth convex bodies of M . Lastly, we note that if M is a Cartan-Hadamard manifold and $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ is convex and lower-semicontinuous, then the function

$$u(t, x) := \inf_{y \in M} \left\{ f(y) + \frac{1}{2t} d(x, y)^2 \right\} \text{ for } t > 0, \quad u(0, x) = f(x)$$

is the unique viscosity solution to the following Hamilton-Jacobi partial differential equation:

$$\begin{cases} \frac{\partial u(t,x)}{\partial t} + \frac{1}{2} \left\| \frac{\partial u(t,x)}{\partial x} \right\|_x^2 = 0, \\ u(0, x) = f(x), \end{cases}$$

where $u : [0, \infty) \times M \rightarrow \mathbb{R}$.

2. General properties

Throughout the paper, for a function $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$, we define

$$f_\lambda(x) = \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\}.$$

The following Proposition shows how, under certain conditions, the infimum defining f_λ can be localized on a neighborhood of the point x .

Proposition 2.1 (Localization). *Let M be a Riemannian manifold, $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ a function satisfying that $f(x) \geq -\frac{c}{2}(1 + d(x, x_0)^2)$ for some $c > 0$, $x_0 \in M$. Let $x \in M$ be such that $f(x) < +\infty$. Then, for all $\lambda \in (0, \frac{1}{2c})$ and for all $\rho > \bar{\rho}$, where*

$$\bar{\rho} = \bar{\rho}(x, \lambda, c) := \left(\lambda \frac{2f(x) + c(2d(x, x_0)^2 + 1)}{1 - 2\lambda c} \right)^{1/2},$$

we have that

$$f_\lambda(x) = \inf_{y \in B(x, \rho)} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\}.$$

Proof. Since

$$d(y, x_0)^2 \leq (d(y, x) + d(x, x_0))^2 \leq 2(d(x, y)^2 + d(x, x_0)^2),$$

we have that

$$c(d(x, y)^2 + d(x, x_0)^2) \geq \frac{c}{2} d(y, x_0)^2,$$

hence

$$-\frac{c}{2} - \frac{c}{2} d(y, x_0)^2 + \frac{1}{2\lambda} d(x, y)^2 \geq \frac{1}{2\lambda} d(x, y)^2 - cd(x, y)^2 - cd(x, x_0)^2 - \frac{c}{2},$$

that is

$$-\frac{c}{2} (1 + d(y, x_0)^2) + \frac{1}{2\lambda} d(x, y)^2 \geq \left(\frac{1}{2\lambda} - c \right) d(x, y)^2 - cd(x, x_0)^2 - \frac{c}{2}. \quad (1)$$

Now, for any given $\eta > 0$, if we set

$$r = r(x, \lambda, c, \eta) := \left(\lambda \frac{2f(x) + 2\eta + c(2d(x, x_0)^2 + 1)}{1 - 2\lambda c} \right)^{1/2},$$

by using (1) we obtain that, for every $y \in M$ with $d(y, x) > r$,

$$\begin{aligned} f(y) + \frac{1}{2\lambda}d(x, y)^2 &\geq -\frac{c}{2}(1 + d(y, x_0)^2) + \frac{1}{2\lambda}d(x, y)^2 \\ &\geq \left(\frac{1}{2\lambda} - c\right)d(x, y)^2 - cd(x, x_0)^2 - \frac{c}{2} \\ &\geq \left(\frac{1}{2\lambda} - c\right)\left(\lambda \frac{2f(x) + 2\eta + c(2d(x, x_0)^2 + 1)}{1 - 2\lambda c}\right) - cd(x, x_0)^2 - \frac{c}{2} \\ &= f(x) + \eta > f_\lambda(x), \end{aligned}$$

which implies that

$$\inf_{d(y,x)>r} \left\{ f(y) + \frac{1}{2\lambda}d(x, y)^2 \right\} > f_\lambda(x),$$

hence

$$f_\lambda(x) = \inf_{d(y,x)\leq r} \left\{ f(y) + \frac{1}{2\lambda}d(x, y)^2 \right\}.$$

Finally, since

$$\lim_{\eta \rightarrow 0} r(x, \lambda, c, \eta) = \bar{\rho}(x, \lambda, c),$$

it is clear that for every $\rho > \bar{\rho}$ we can find $\eta > 0$ small enough so that

$$\rho = \rho(x, \lambda, c) > r(x, \lambda, c, \eta) > \bar{\rho}(x, \lambda, c),$$

and therefore, from the above argument we deduce that

$$f_\lambda(x) = \inf_{d(y,x)\leq\rho} \left\{ f(y) + \frac{1}{2\lambda}d(x, y)^2 \right\}. \quad \square$$

Next we state several interesting properties of this inf-convolution operation, such as preservation of order and symmetry properties of the original function. We put off studying the conditions under which convexity is preserved until the next section.

Proposition 2.2. *Let M be a Riemannian manifold, $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ a function. We have that:*

- (i) $f_\lambda \leq f$ for all $\lambda > 0$.
- (ii) If $0 < \lambda_1 < \lambda_2$ then $f_{\lambda_2} \leq f_{\lambda_1}$.
- (iii) $\inf f_\lambda = \inf f$ and, moreover, if f is lower semicontinuous then every minimizer of f_λ is a minimizer of f , and conversely.
- (iv) If T is an isometry of M onto M , and f is invariant under T (that is, $f(Tz) = f(z)$ for all $z \in M$), then f_λ is also invariant under T , for all $\lambda > 0$.

Proof. (i) and (ii) are obvious.

(iii) Note that

$$\begin{aligned} \inf_{x \in M} f_\lambda(x) &= \inf_{x \in M} \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\} \\ &= \inf_{y \in M} \inf_{x \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\} = \inf_{y \in M} f(y). \end{aligned}$$

Since $f_\lambda \leq f$ it is then obvious that every minimizer of f is a minimizer of f_λ as well. Conversely, assume that f is lower semicontinuous, let x_0 be a minimizer of f_λ , that is, $f_\lambda(x_0) = \inf_{y \in M} f_\lambda(y) = \inf_{y \in M} f(y)$, and let us see that $f(x_0) = \inf_{y \in M} f(y)$. Choose $\{x_n\}_{n \in \mathbb{N}} \subset M$ so that $f(x_n) + \frac{1}{2\lambda} d(x_0, x_n)^2 \leq f_\lambda(x_0) + \frac{1}{n}$; then we have

$$0 \leq \frac{1}{2\lambda} d(x_0, x_n)^2 \leq f_\lambda(x_0) + \frac{1}{n} - f(x_n) \leq f_\lambda(x_0) + \frac{1}{n} - \inf_{y \in M} f(y) = \frac{1}{n} \rightarrow 0,$$

hence (x_n) converges to x_0 , and since f is lower semicontinuous we get

$$f(x_0) \leq \lim_n f(x_n) \leq \lim_n \left(f_\lambda(x_0) + \frac{1}{n} \right) = f_\lambda(x_0) = \inf_{y \in M} f_\lambda(y) = \inf_{y \in M} f(y)$$

and x_0 is a minimizer of f .

(iv) We have that

$$\begin{aligned} f_\lambda(Tx) &= \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(Tx, y)^2 \right\} = \inf_{y \in M} \left\{ f(Ty) + \frac{1}{2\lambda} d(Tx, Ty)^2 \right\} \\ &= \inf_{y \in M} \left\{ f(Ty) + \frac{1}{2\lambda} d(x, y)^2 \right\} = \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\} = f_\lambda(x). \quad \square \end{aligned}$$

Now we apply Proposition 2.1 to show that, under natural continuity assumptions on f , the regularizations f_λ converge to the original function f as λ goes to 0.

Proposition 2.3 (Convergence). *Let M be a Riemannian manifold, $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ a function satisfying that $f(x) \geq -\frac{c}{2}(1 + d(x, x_0)^2)$ for some $c > 0$, $x_0 \in M$, and consider*

$$f_\lambda(x) = \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\}$$

for $0 < \lambda < 1/2c$.

- (i) *Assume that f is uniformly continuous on bounded subsets of M . Then $\lim_{\lambda \rightarrow 0} f_\lambda = f$ uniformly on each bounded subset of M .*
- (ii) *Assume that f is continuous on M . Then $\lim_{\lambda \rightarrow 0} f_\lambda = f$ uniformly on compact subsets of M .*

- (iii) Assume that f is uniformly continuous and bounded on all of M . Then $\lim_{\lambda \rightarrow 0} f_\lambda = f$ uniformly on M .
- (iv) In general, if f is only assumed to be lower semicontinuous, we have that $\lim_{\lambda \rightarrow 0} f_\lambda(x) = f(x)$ for every $x \in M$ with $f(x) < +\infty$.

Proof. (i) According to Proposition 2.1, for every $x \in M$, $\lambda \in (0, 1/2c)$, $\rho > \bar{\rho}(x, \lambda, c)$, we have that

$$f_\lambda(x) = \inf_{d(x,y) \leq \rho} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\}.$$

Fix $R > 0$. As is easily shown, a uniformly continuous function on a Riemannian manifold is bounded on bounded sets, hence we can find $k > 0$ so that $|f(x)| \leq k$ for all $x \in B(x_0, 2R)$.

For any given $\varepsilon > 0$, by uniform continuity of f , there exists $\delta > 0$ such that if $y, x \in B(x_0, 2R)$ and $d(x, y) \leq \delta$ then $|f(x) - f(y)| \leq \varepsilon/3$. We can assume $\delta < R$. Now, since

$$\lim_{\lambda \rightarrow 0^+} \left(\lambda \frac{2k + c(2R^2 + 1)}{1 - 2c\lambda} \right)^{1/2} = 0,$$

there is $\lambda_\varepsilon > 0$ such that if $0 < \lambda < \lambda_\varepsilon$ then

$$0 < \rho(x, \lambda, c) \leq \left(\lambda \frac{2k + c(2R^2 + 1)}{1 - 2c\lambda} \right)^{1/2} < \delta$$

for all $x \in B(x_0, R)$, and therefore

$$f_\lambda(x) = \inf_{d(y,x) \leq \delta} \left\{ f(y) + \frac{1}{2\lambda} d(y, x)^2 \right\}$$

for all $x \in B(x_0, R)$, $\lambda \in (0, \lambda_\varepsilon)$. But, since $f_\lambda \leq f$ for all λ , this really means that

$$f_\lambda(x) = \inf_{y \in A_x} \left\{ f(y) + \frac{1}{2\lambda} d(y, x)^2 \right\},$$

where

$$A_x := \left\{ y \in B(x, \delta) : f(y) + \frac{1}{2\lambda} d(y, x)^2 \leq f(x) \right\}.$$

By the definition of inf, we can take $y_x \in A_x$ such that

$$f_\lambda(x) + \frac{\varepsilon}{3} \geq f(y_x) + \frac{1}{2\lambda} d(y_x, x)^2.$$

Then, bearing in mind that $y_x \in A_x \subseteq B(x, \delta) \subseteq B(x_0, 2R)$ when $x \in B(x_0, R)$, we get

$$\begin{aligned} |f(x) - f_\lambda(x)| &= f(x) - f_\lambda(x) \leq f(x) - f(y_x) - \frac{1}{2\lambda} d(y_x, x)^2 + \frac{\varepsilon}{3} \\ &\leq |f(x) - f(y_x)| + \frac{\varepsilon}{3} \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon \end{aligned}$$

for all $x \in B(x_0, R)$, $\lambda \in (0, \lambda_\varepsilon)$. This shows that $\lim_{\lambda \rightarrow 0^+} f_\lambda = f$ uniformly on $B(x_0, R)$.

(ii) Let K be a compact subset of M . By compactness, it is easily seen that for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $x \in K$, $y \in M$, and $d(x, y) \leq \delta$ then $|f(x) - f(y)| \leq \varepsilon/3$. One can now repeat the above argument with the precaution of always taking $x \in K \subset B(x_0, R)$, $y \in M$, $d(x, y) \leq \delta$.

(iii) Choose $k > 0$ such that $|f(x)| \leq k$ for all $x \in M$. Let us first observe that the infimum defining f_λ can be restricted to the set $\{y \in M : d(y, x) \leq 2\sqrt{k\lambda}\} = B(x, 2\sqrt{k\lambda})$. Indeed, if $d(y, x) > 2\sqrt{k\lambda}$ then

$$f(y) + \frac{1}{2\lambda}d(x, y)^2 > -k + 2k = k \geq f(x) \geq f_\lambda(x).$$

Next, for any given $\varepsilon > 0$, the uniform continuity of f provides us with $\delta > 0$ so that $|f(y) - f(x)| \leq \varepsilon/3$ whenever $d(x, y) \leq \delta$. On the other hand, since $\lim_{\lambda \rightarrow 0^+} 2\sqrt{k\lambda} = 0$, there exists $\lambda_\varepsilon > 0$ such that $0 < 2\sqrt{k\lambda} < \delta$ for $0 < \lambda < \lambda_\varepsilon$. Then, for any $x \in M$, $\lambda \in (0, \lambda_\varepsilon)$ the infimum defining $f_\lambda(x)$ can be restricted to the set $A_x := \{y \in B(x, \delta) : f(y) + \frac{1}{2\lambda}d(x, y)^2 \leq f(x)\}$. Now, the same estimations as in (i) above show that $|f_\lambda(x) - f(x)| \leq \varepsilon$.

(iv) Fix $x \in M$. Since f is lower semicontinuous we have $f(x) \leq \liminf_{y \rightarrow x} f(y)$, hence, for any given $\varepsilon > 0$, there exists $\delta > 0$ so that

$$f(x) - f(y) \leq \frac{\varepsilon}{3}$$

for all $y \in B(x, \delta)$. Now, since

$$\lim_{\lambda \rightarrow 0^+} \left(\lambda \frac{2f(x) + c(2d(x, x_0)^2 + 1)}{1 - 2c\lambda} \right)^{1/2} = 0,$$

there is $\lambda_\varepsilon > 0$ such that if $0 < \lambda < \lambda_\varepsilon$ then

$$0 < \left(\lambda \frac{2f(x) + c(2d(x, x_0)^2 + 1)}{1 - 2c\lambda} \right)^{1/2} < \delta$$

and therefore

$$f_\lambda(x) = \inf_{d(y, x) \leq \delta} \left\{ f(y) + \frac{1}{2\lambda}d(y, x)^2 \right\}$$

for all $\lambda \in (0, \lambda_\varepsilon)$. But, since $f_\lambda \leq f$ for all λ , this means

$$f_\lambda(x) = \inf_{y \in A_x} \left\{ f(y) + \frac{1}{2\lambda}d(y, x)^2 \right\},$$

where

$$A_x := \left\{ y \in B(x, \delta) : f(y) + \frac{1}{2\lambda}d(y, x)^2 \leq f(x) \right\}.$$

By the definition of infimum, we can take $y_x \in A_x$ such that

$$f_\lambda(x) + \frac{\varepsilon}{3} \geq f(y_x) + \frac{1}{2\lambda}d(y_x, x)^2.$$

Then we get

$$\begin{aligned} |f(x) - f_\lambda(x)| &= f(x) - f_\lambda(x) \leq f(x) - f(y_x) - \frac{1}{2\lambda}d(y_x, x)^2 + \frac{\varepsilon}{3} \\ &\leq f(x) - f(y_x) + \frac{\varepsilon}{3} \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon \end{aligned}$$

for all $\lambda \in (0, \lambda_\varepsilon)$. This shows that $\lim_{\lambda \rightarrow 0^+} f_\lambda(x) = f(x)$. \square

3. Regularization of convex functions

In order to see that the operation $f \rightarrow f_\lambda$ preserves convexity we will need to use the following Lemmas.

Lemma 3.1. *Let M be a Riemannian manifold, and $F : M \times M \rightarrow \mathbb{R} \cup \{+\infty\}$ a convex function (where $M \times M$ is endowed with its natural product Riemannian metric). Assume either that M has the property that every two points can be connected by a geodesic in M , or else that F is continuous and M is complete. Then, the function $\psi : M \rightarrow \mathbb{R}$ defined by*

$$\psi(x) = \inf_{y \in M} F(x, y)$$

is also convex.

Proof. Let $\gamma : I \rightarrow M$ be a geodesic. We have to see that the function $t \mapsto \psi(\gamma(t))$ is convex, that is, for any $t_0, t_1 \in I$, and for any $s \in [0, 1]$, $\psi(\gamma(st_1 + (1-s)t_0)) \leq s\psi(\gamma(t_1)) + (1-s)\psi(\gamma(t_0))$. We may assume (up to an affine change of parameters) that $t_0 = 0$ and $t_1 = 1$, so we have to show that $\psi(\gamma(t)) \leq t\psi(x_1) + (1-t)\psi(x_0)$, where $x_0 := \gamma(0)$ and $x_1 := \gamma(1)$.

Fix an arbitrary $t \in [0, 1]$. For any $\varepsilon > 0$, by the definition of ψ , we can pick $y_0, y_1 \in M$ so that

$$F(x_1, y_1) < \psi(x_1) + \varepsilon \quad \text{and} \quad F(x_0, y_0) < \psi(x_0) + \varepsilon. \quad (2)$$

Let $\sigma : J \rightarrow M$ be a geodesic connecting y_0 and y_1 . (If such σ does not exist then we can use continuity of F and Ekeland's approximate Hopf-Rinow type theorem to get points \bar{y}_0 and \bar{y}_1 close enough to y_0 and y_1 so that (2) remains true when replacing y_j with \bar{y}_j , and a geodesic $\bar{\sigma}$ connecting \bar{y}_0 to \bar{y}_1 ; the rest of the argument applies without changes.) We can also assume that $J = [0, 1]$, $y_0 = \sigma(0)$, $y_1 = \sigma(1)$.

It is clear that, because γ and σ are geodesics in M , the path $t \mapsto (\gamma(t), \sigma(t))$ is a geodesic joining the points (x_0, y_0) and (x_1, y_1) in the product manifold $M \times M$.

Now, since $t \mapsto F(\gamma(t), \sigma(t))$ is convex, we have that

$$\begin{aligned} \psi(\gamma(t)) &= \inf_{y \in M} F(\gamma(t), y) \leq F(\gamma(t), \sigma(t)) \leq tF(x_1, y_1) + (1-t)F(x_0, y_0) \\ &\leq t(\psi(x_1) + \varepsilon) + (1-t)(\psi(x_0) + \varepsilon) = t\psi(x_1) + (1-t)\psi(x_0) + \varepsilon \end{aligned}$$

and this holds for every $\varepsilon > 0$, hence we can conclude that $\psi(\gamma(t)) \leq t\psi(x_1) + (1-t)\psi(x_0)$. \square

Lemma 3.2. *Let M be a Riemannian manifold with the property that any two points of M can be joined by a minimizing geodesic, and let $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ be a convex function. Then, for every $x_0 \in M$ there exists a number $c \geq 0$ such that $f(x) \geq -\frac{c}{2}(1 + d(x, x_0)^2)$ for all $x \in M$.*

Proof. Choose $\zeta \in D^-f(x_0)$. For a given $x \in M$, let γ be a minimizing geodesic connecting x_0 to x , say $\gamma(t) = \exp_{x_0}(tv)$, $t \in [0, d(x, x_0)]$, for some $v \in TM_{x_0}$ with $\|v\|_{x_0} = 1$. Since f is convex we have $f(\exp_{x_0}(tv)) - f(x_0) \geq \langle \zeta, tv \rangle_{x_0}$ for every $t \in [0, d(x_0, x)]$, and in particular

$$\begin{aligned} f(x) - f(x_0) &\geq \langle \zeta, d(x_0, x)v \rangle_{x_0} \geq -\|\zeta\|_{x_0}d(x_0, x)\|v\|_{x_0} \\ &= -\|\zeta\|_{x_0}d(x_0, x) \geq -\|\zeta\|_{x_0}(1 + d(x_0, x)^2), \end{aligned}$$

and therefore

$$f(x) \geq f(x_0) - \|\zeta\|_{x_0}(1 + d(x_0, x)^2) \geq -\frac{c}{2}(1 + d(x_0, x)^2)$$

for all $x \in M$ if we put $c = 2(\|\zeta\|_{x_0} + |f(x_0)|)$. \square

It will be also useful to recall that every convex function $f : M \rightarrow \mathbb{R}$ which is locally bounded is continuous (in fact locally Lipschitz); a proof of this statement can be found in [2, Proposition 5.2].

In order that f_λ is convex whenever f is, we will have to require that the distance function $d : M \times M \rightarrow \mathbb{R}$ is convex on a band around the diagonal of $M \times M$. More precisely, we will use the following.

Definition 3.3. Let M be a Riemannian manifold. We say that the distance function $d : M \times M \rightarrow \mathbb{R}$ is uniformly locally convex on bounded sets near the diagonal if, for every bounded subset B of M , there exists $r > 0$ such that d is convex on $B(x, r) \times B(x, r)$, and the set $B(x, r)$ is convex in M , for all $x \in B$.

Examples of manifolds satisfying this definition are the cylinder $x^2 + y^2 = 1$ in \mathbb{R}^3 , the Poincaré half-plane, or the subsets of \mathbb{R}^3 defined by $z = 1/(x^2 + y^2)$ or $z = xy$. In general, every complete finite-dimensional Riemannian manifold of nonpositive sectional curvature meets this condition, as we will show in the next section.

Proposition 3.4. *Let M be a Riemannian manifold with the property that any two points of M can be joined by a geodesic, and let $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower-semicontinuous convex function.*

- (i) *Assume that f is bounded on bounded sets and that the distance function $d : M \times M \rightarrow \mathbb{R}$ is uniformly locally convex on bounded sets near the diagonal. Then, for every bounded subset B of M there exists $\lambda_0 > 0$ such that f_λ is convex on B for all $\lambda \in (0, \lambda_0)$.*
- (ii) *Assume that the distance function $d : M \times M \rightarrow \mathbb{R}$ is convex on all of $M \times M$. Then f_λ is convex on M for every $\lambda > 0$.*

Proof. (i) We may well assume $B = B(x_0, R)$ for some $x_0 \in M$, $R > 0$. Let $r > 0$ be small enough so that the function $(x, y) \mapsto d(y, x)$ is convex on $B(x, 2r) \times B(x, 2r)$ and $B(x, 2r)$ is convex for every $x \in B(x_0, R)$. Let k be a bound for f on $B(x_0, 2R)$. We may assume $2r < R$. We have that, for every $z \in B(x_0, 2R)$,

$$\bar{\rho}(z, \lambda, k) := \left(\lambda \frac{2f(z) + k(1 + 2R^2)}{1 - 2\lambda k} \right)^{1/2} \leq \left(\lambda \frac{2k + k(1 + 2R^2)}{1 - 2\lambda k} \right)^{1/2} \rightarrow 0 \quad \text{as } \lambda \rightarrow 0^+$$

hence we can choose $\lambda_0 > 0$ small enough so that $\bar{\rho}(z, \lambda_0, k) < r$ for all $z \in B(x_0, 2R)$ and therefore, according to Proposition 2.1, we have that, for every $\lambda \in (0, \lambda_0)$,

$$f_\lambda(z) = \inf_{y \in B(z, r)} \left\{ f(y) + \frac{1}{2\lambda} d(z, y)^2 \right\} = \inf_{y \in B(x, 2r)} \left\{ f(y) + \frac{1}{2\lambda} d(z, y)^2 \right\}$$

whenever $z \in B(x, r)$, $x \in B(x_0, R)$. Note that $B(x, 2r)$, as a convex subset of M , still has the property that any two of its points can be joined by a geodesic inside $B(x, 2r)$. Now, assuming $0 < \lambda < \lambda_0$, and fixing $x \in B(x_0, R)$, because the function $F(z, y) := f(y) + \frac{1}{2\lambda} d(z, y)^2$ is jointly convex on $B(x, 2r) \times B(x, 2r)$, we deduce from Lemma 3.1 that $z \mapsto f_\lambda(z) = \inf_{y \in B(x, 2r)} F(z, y)$ is convex on $B(x, r)$, for all $0 < \lambda < \lambda_0$. Since $x \in B(x_0, R)$ is arbitrary this implies that f_λ is convex on $B(x_0, R)$, for every $\lambda \in (0, \lambda_0)$.

(ii) Here we can use Lemma 3.1 on all of $M \times M$ with no need to localize the infimum defining $f_\lambda(x)$, so it follows that f_λ is convex for all $\lambda > 0$. □

Remark 3.5. Note that in Case (2) of the above proposition we do not require continuity of the function f , so f is permitted to take the value $+\infty$ at some points; in particular we are allowed to take f to be the indicator function of a closed convex subset C of M , that is

$$\delta_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

Remark 3.6. An examination of the above proof and the statement of Lemma 3.1 reveals that, if one assumes that f is continuous and M is complete, it is not necessary to require that every two points of M can be connected by a minimizing geodesic in M .

The following proposition shows that the functions f_λ are superdifferentiable at a point x if d is differentiable on a suitable ball around x . We refer the reader to [2] for the properties of viscosity subdifferentials on Riemannian manifolds; here we will only make use of the very definition of the subdifferential and the superdifferential sets of $f : M \rightarrow [-\infty, \infty]$, namely,

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

and

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

the fact that f is differentiable at x if and only if $D^- f(x) \neq \emptyset \neq D^+ f(x)$ (in which case $\{df(x)\} = D^- f(x) = D^+ f(x)$), and that a convex function $f : M \rightarrow \mathbb{R}$ is everywhere subdifferentiable [2, Theorem 5.3]

Proposition 3.7. *Suppose that the infimum defining $f_\lambda(x)$ can be restricted to a ball $B_x = B(x, r_x)$ of radius r_x small enough so that the function $y \mapsto d(y, x)^2$ is differentiable on B_x , and that this infimum is attained at a point $y_x \in B_x$. Then f_λ is superdifferentiable at x , and*

$$\frac{1}{\lambda} d(x, y_x) \frac{\partial}{\partial x} d(x, y_x) \in D^+ f_\lambda(x).$$

Proof. We have that

$$\begin{aligned} f_\lambda(z) - f_\lambda(x) &\leq f(y_x) + \frac{1}{2\lambda} d(z, y_x)^2 - f(y_x) - \frac{1}{2\lambda} d(x, y_x)^2 \\ &= \frac{1}{2\lambda} d(z, y_x)^2 - \frac{1}{2\lambda} d(x, y_x)^2, \end{aligned}$$

so

$$f_\lambda(z) - \frac{1}{2\lambda} d(z, y_x)^2 \leq f_\lambda(x) - \frac{1}{2\lambda} d(x, y_x)^2$$

for every $z \in B_x$, that is, $f_\lambda - \frac{1}{2\lambda} d(\cdot, y_x)^2$ attains a local maximum at x , hence f_λ is superdifferentiable at x , with $d(\frac{1}{2\lambda} d(\cdot, y_x)^2)(x) \in D^+ f_\lambda(x)$. \square

Next we show that convex differentiable functions are automatically of class C^1 . In this proof we will make use of the parallel transport of vectors along geodesics. Recall that, for a given curve $\gamma : I \rightarrow M$, numbers $t_0, t_1 \in I$, 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 $V_0 \mapsto V(t_1)$ is a linear isometry between the

tangent spaces $TM_{\gamma(t_0)}$ and $TM_{\gamma(t_1)}$, for each $t_1 \in I$. In the case when γ is a unique minimizing geodesic joining the points $\gamma(t_0) = x$ and $\gamma(t_1) = y$, we will denote this mapping by L_{xy} , and we call it the parallel transport from TM_x to TM_y along the geodesic γ . See [16] for general reference on these topics. The parallel transport allows 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

$$\|v - L_{yx}(w)\|_x = \|w - L_{xy}(v)\|_y.$$

(This equality holds because L_{xy} is a linear isometry between the two tangent spaces, with inverse L_{yx} .) 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 $\zeta \in T^*M_x$ and $\eta \in T^*M_y$ lying on different fibers of the cotangent bundle.

It is also well known that the mapping $y \mapsto L_{xy}$ is well defined and continuous on a neighborhood of each $x \in M$, in the following sense: if (x_n) converges to x in M then $\exp_{x_n}(L_{xx_n}(v))$ converges to $\exp_x(v)$ uniformly on the set $\{v \in TM_x : \|v\|_x \leq \delta\}$ for some $\delta > 0$ (a fact which we use at the end of the proof of the following proposition).

Proposition 3.8. *Let M be a Riemannian manifold, and let $f : M \rightarrow \mathbb{R}$ be a differentiable convex function. Then f is of class C^1 on M .*

Proof. Assume that f is not C^1 , then there are $\varepsilon > 0$, a point $x \in M$ and a sequence $(x_n) \subset M$ converging to x such that

$$\|L_{x_n x}[df(x_n)] - df(x)\|_x > 2\varepsilon$$

for all $n \in \mathbb{N}$. Therefore, for every $n \in \mathbb{N}$ we can pick $h_n \in TM_x$ with $\|h_n\|_x = 1$ such that

$$\langle L_{x_n x}[df(x_n)] - df(x), h_n \rangle_x > 2\varepsilon \quad \text{for all } n \in \mathbb{N}.$$

Since f is differentiable at x , there exists $\delta > 0$ so that

$$f(\exp_x(tv)) - f(x) - \langle df(x), tv \rangle_x \leq \varepsilon t$$

for all $v \in TM_x$ with $\|v_x\| = 1$ and $|t| \leq \delta$. On the other hand, by convexity of f , we have

$$\langle df(x_n), tv \rangle_{x_n} \leq f(\exp_{x_n}(tv)) - f(x_n)$$

for all $w \in TM_{x_n}$ with $\|w\|_{x_n} = 1$ and $|t| \leq \delta$. By combining these inequalities we get

$$\begin{aligned} 2\varepsilon\delta &\leq \langle L_{x_n x}[df(x_n)] - df(x), h_n \rangle_x \delta = \langle L_{x_n x}[df(x_n)], \delta h_n \rangle_x - \langle df(x), \delta h_n \rangle_x \\ &= \langle df(x_n), \delta L_{x_n x} h_n \rangle_{x_n} - \langle df(x), \delta h_n \rangle_x \\ &\leq f(\exp_{x_n}(\delta L_{x_n x} h_n)) - f(x_n) + f(x) - f(\exp_x(\delta h_n)) + \varepsilon\delta \\ &= f(\exp_{x_n}(\delta L_{x_n x} h_n)) - f(\exp_x(\delta h_n)) + f(x) - f(x_n) + \varepsilon\delta \rightarrow 0 + 0 + \varepsilon\delta \end{aligned}$$

(by continuity of f , \exp , and the parallel translation L), so we get that $2\varepsilon\delta \leq \varepsilon\delta$, which is not possible. \square

Now we can prove that, under the same assumptions as in Proposition 3.4, if f is convex then f_λ is of class C^1 for $\lambda > 0$ small enough.

Theorem 3.9. *Let M be a Riemannian manifold and let $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous and convex function. Assume that every two points of M can be connected by a minimizing geodesic in M .*

- (i) *Suppose that f is bounded on bounded sets and that the distance function $d : M \times M \rightarrow \mathbb{R}$ is uniformly locally convex on bounded sets near the diagonal. Then, for every bounded open convex subset B of M there exists $\lambda_0 > 0$ such that f_λ is a C^1 smooth convex function on B , for all $\lambda \in (0, \lambda_0)$.*
- (ii) *Suppose that the distance function $d : M \times M \rightarrow \mathbb{R}$ is convex on all of $M \times M$. Then f_λ is a C^1 smooth convex function on M , for every $\lambda > 0$.*

Proof. (i) We can give an almost self-contained proof of this in the finite-dimensional case, so let us first assume that $\dim M < +\infty$. We may also assume $B = B(x_0, R)$ for some $x_0 \in M$, $R > 0$. Since the injectivity radius $x \mapsto i(x)$ is a continuous positive function, it is bounded below by a positive number on the compact subset $\bar{B}(x_0, R)$ of M . This implies that there exists $r_0 > 0$ such that the function $y \mapsto d(x, y)^2$ is C^1 smooth on $B(x, r_0)$ for every $x \in B(x_0, R)$. We can obviously assume that $r < r_0$ and repeat the argument of the proof of Proposition 3.4 to get a $\lambda_0 > 0$ such that f_λ is convex on $B(x_0, R)$ for all $\lambda \in (0, \lambda_0)$ and, moreover, that, for every $x \in B(x_0, R)$, the infimum defining $f_\lambda(x)$ can be restricted to the ball $\bar{B}(x, r)$, which is contained in $B(x, r_0)$ (so that, in particular, $y \mapsto d(y, x)^2$ is C^1 smooth on $B(x, r)$). Besides, this infimum is attained, because the involved functions are continuous and the ball $\bar{B}(x, r)$ is compact. According to Proposition 3.7, we then get that f_λ is superdifferentiable at x , for every $x \in B(x_0, R)$, $\lambda \in (0, \lambda_0)$.

On the other hand, since f_λ is convex on $B(x_0, R)$, we know that f_λ is subdifferentiable on $B(x_0, R)$ (see [2, Theorem 5.3]). That is, f_λ is both subdifferentiable and superdifferentiable at each point of $B(x_0, R)$, hence f_λ is differentiable on $B(x_0, R)$ (see [2, Proposition 4.6]). Since f_λ is convex, Proposition 3.8 allows to conclude that f_λ is C^1 smooth on $B(x_0, R)$ for each $\lambda \in (0, \lambda_0)$.

Let us now consider the case when $\dim M = +\infty$. Here we will use the following result from [1]:

Theorem ([1, Theorem 3.3]). *Let M be a connected, complete Riemannian manifold, and let $f : M \rightarrow \mathbb{R}$ be a continuous function, bounded below by a constant c . Then we have that for every $\alpha > 0$ the function*

$$f_\alpha(x) = \inf_{y \in M} \{f(y) + \alpha d(x, y)^2\}$$

is bounded below by c , is Lipschitz on bounded sets, and satisfies

$$\lim_{\alpha \rightarrow +\infty} f_\alpha(x) = f(x).$$

Moreover, for every $x_0 \in M$ with $\partial_P f_\alpha(x_0) \neq \emptyset$, there is a $y_0 \in M$ such that:

- (a) Every minimizing sequence $\{y_n\}$ in the definition of $f_\alpha(x_0)$ converges to y_0 , and consequently the infimum is a strong minimum.
- (b) There is a minimizing geodesic γ joining x_0 and y_0 .
- (c) f_α is differentiable at x_0 .
- (d) $L_{x_0 y_0}[df_\alpha(x_0)] \in \partial_P f(y_0)$.

Finally, if we assume that M is finite-dimensional (or, more generally, if we assume that M can be infinite-dimensional but still has the property that every two points of M are connected by a minimizing geodesic) then the same remains true of every lower semicontinuous function $f : M \rightarrow (-\infty, +\infty]$ which is bounded below by c .

Here ∂_P denotes the proximal subdifferential introduced in [1], which has the property that $\partial_P g(x) \subseteq D^-g(x)$ and, for a convex g , in fact $\partial_P g(x) = D^-g(x) \neq \emptyset$ for all x .

Since f_λ , being convex on B for all $\lambda \in (0, \lambda_0)$, satisfies $\partial_P f_\lambda(x) = D^-f_\lambda(x) \neq \emptyset$ for all $x \in B$, we can apply the above Theorem to get that f_λ is differentiable at every point $x \in B$, hence (by Proposition 3.8) of class C^1 on B .

(ii) As in case (i), let us first give a self-contained proof for the finite-dimensional case. If the distance function $d : M \times M \rightarrow \mathbb{R}$ is convex on all of $M \times M$ then $y \mapsto d(y, x)$ is convex on M for all $x \in M$, which implies that the cut locus of x is empty for every $x \in M$ and that the function $y \mapsto d(y, x)^2$ is differentiable on all of M for every $x \in M$. On the other hand, we claim that the infimum defining $f_\lambda(x)$ is attained for every $x \in M$. Indeed, fix $x \in M$ with $f(x) < +\infty$. From the proof of Lemma 3.2 we know that there exists $c = c_x \geq 0$ such that $f(y) \geq -cd(x, y)$ for every $y \in M$. Then we have

$$f(y) + \frac{1}{2\lambda}d(x, y)^2 \geq -cd(x, y) + \frac{1}{2\lambda}d(x, y)^2 \rightarrow +\infty$$

if $d(x, y) \rightarrow +\infty$, so there exists $R > 0$ large enough so that if $d(x, y) \geq R$ then

$$f(y) + \frac{1}{2\lambda}d(x, y)^2 \geq f(x) \geq f_\lambda(x),$$

hence

$$f_\lambda(x) = \inf_{y \in \bar{B}(x, R)} \left\{ f(y) + \frac{1}{2\lambda}d(x, y)^2 \right\},$$

and now it is clear that this infimum is attained because $f + \frac{1}{2\lambda}d(\cdot, x)^2$ is lower semicontinuous and $\bar{B}(x, R)$ is compact.

Therefore, according to Proposition 3.7, f_λ is superdifferentiable. Because f_λ is convex, this means that f_λ is differentiable, hence C^1 smooth on M , for all $\lambda > 0$.

In the infinite-dimensional case we only have to bear in mind that, according to Proposition 3.4(2), f_λ is convex on all of M , so the same proof as in case (i) applies. \square

Remark 3.10. If one assumes that f is continuous and M is complete, it is not necessary to require that every two points of M can be connected by a minimizing geodesic in M .

4. Which manifolds do the above results apply to?

Let us say a few words about the manifolds satisfying the assumptions of Theorem 3.9. The following theorem is a restatement of [17, Theorem IX.4.3, p. 257]

Theorem 4.1. *Let M be a Riemannian manifold with seminegative sectional curvature $K \leq 0$, and U a convex open set. Let β_1, β_2 be disjoint geodesics in U , defined on the same interval. Let $\alpha_t : [a, b] \rightarrow U$ be the unique geodesic joining $\beta_1(t)$ with $\beta_2(t)$, and let $\ell(t) = \text{length}(\alpha_t)$, that is, $\ell(t) = d(\beta_1(t), \beta_2(t))$. Then $\ell''(t) \geq 0$ for all t , and in particular $\ell(t)$ is a convex function.*

From this theorem it is immediate to deduce that the above results on regularization of convex functions apply to manifolds of seminegative sectional curvature, as we next see.

Corollary 4.2. *Let M be a Riemannian manifold with sectional curvature $K \leq 0$.*

- (i) *Suppose that M has a convexity radius function which is strictly positive on bounded subsets of M . (Such is the case, for instance, of a complete finite-dimensional Riemannian manifold M .) Then the distance function d is uniformly locally convex on bounded sets near the diagonal of $M \times M$.*
- (ii) *Suppose that M is simply connected (which, together with the curvature assumption, amounts to saying that M is a Cartan-Hadamard manifold). Then the distance function d is convex on all of M .*

Proof. (i) Let B be a bounded subset of M . Since the convexity radius function $x \mapsto c(x)$ is bounded below on B by a number $r > 0$, we have that the open ball $B(x, r)$ is convex for every $x \in B$. Therefore, for every $x \in B$ and for every pair of disjoint geodesic segments $\beta_1, \beta_2 : I := (a, b) \rightarrow M$ contained in $B(x, r)$, Theorem 4.1 tells us that the function $t \mapsto \ell(t) := d(\beta_1(t), \beta_2(t))$ is convex. If β_1, β_2 are not disjoint and neither of them is constant (in which case the result would be trivial) then we

can only have the equality $d(\beta_1(t), \beta_2(t)) = 0$ for a unique $t = t_0$, at which point the function $\ell(t)$ attains an absolute minimum, and Theorem 4.1 shows that $\ell(t)$ is convex on (a, t_0) and on (t_0, b) . But a real function which is convex on (a, t_0) and on (t_0, b) , and which attains its minimum at t_0 , must in fact be convex on all of $I = (a, b)$.

This proves that the distance function d is convex on $B(x, r) \times B(x, r)$, for every $x \in B$, which in turn means that d is uniformly locally convex near the diagonal.

(ii) In a Cartan-Hadamard manifold M every ball is convex, and two distinct geodesics in M can intersect in only one point (see [17, p. 259-261]), so the above argument applies globally. \square

Remark 4.3. The assumption on curvature is necessary in order that d be uniformly locally convex near the diagonal: it is easy to see that, for many disjoint nonconstant geodesic segments β_1 and β_2 in the sphere S^2 , the function $t \rightarrow d(\beta_1(t), \beta_2(t))$ is *not* convex. Indeed, the geodesics in the sphere are just the great circles, so we can take for instance two distinct geodesic segments β_1 and $\beta_2 : [-\varepsilon, \varepsilon] \rightarrow S^2$ lying on two distinct parallel meridians near the equator, parameterized in such a way that $\beta_1(0)$ and $\beta_2(0)$ lie on the equator line and with the property that $d(\beta_1(t), \beta_2([-\varepsilon, \varepsilon])) = d(\beta_1(t), \beta_2(t))$ for all $t \in [-\varepsilon, \varepsilon]$. Then it is clear that the function

$$(-\varepsilon, \varepsilon) \ni t \mapsto d(\beta_1(t), \beta_2(t))$$

attains a strict maximum at $t = 0$ and therefore it cannot be convex.

Furthermore, as we will see in the next section, an important consequence of Theorem 3.9 fails in the sphere S^2 , so the assumption on the jointly convexity of the distance function $d : M \times M \rightarrow \mathbb{R}$ near the diagonal seems to be much more than a mere technical requirement and is probably a necessary condition for the functions f_λ to be convex whenever f is.

We conclude with a corollary that sums up what the results we have shown tell us in the case of a Riemannian manifold of nonpositive curvature.

When M is a complete finite-dimensional Riemannian manifold of nonpositive curvature we have the following result. Recall that a convex function f on a finite-dimensional Riemannian manifold M that only takes finite values is automatically continuous (see [3]).

Corollary 4.4. *Let M be a complete finite-dimensional Riemannian manifold with sectional curvature $K \leq 0$. Let $f : M \rightarrow \mathbb{R}$ be a convex function. Then, for every bounded open convex set U with compact closure \bar{U} , there exists $\lambda_0 > 0$ such that the functions $f_\lambda : M \rightarrow \mathbb{R}$, defined by*

$$f_\lambda(x) = \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\},$$

are convex and C^1 smooth on U for all $\lambda \in (0, \lambda_0)$. Moreover,

- (i) f_λ converges to f uniformly on \bar{U} .
- (ii) $f_\lambda \leq f$ for all $\lambda > 0$.
- (iii) f_λ has the same infimum and the same set of minimizers as f .
- (iv) f_λ has the same symmetries as f (that is, if f is invariant with respect to an isometry $T : M \rightarrow M$, then so is f_λ).

Note that in this result we do not allow f to take infinite values. We are able to deal with functions $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ when we furthermore assume that M is a Cartan-Hadamard manifold (that is, a simply connected Riemannian manifold of nonpositive curvature), either finite or infinite-dimensional. Also recall that Cartan-Hadamard manifolds enjoy the property that every two points can be connected by a minimizing geodesic (see [17]).

Corollary 4.5. *Let M be a Cartan-Hadamard manifold (either finite-dimensional or infinite-dimensional). Let $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower-semicontinuous convex function. Then the functions $f_\lambda : M \rightarrow \mathbb{R}$, defined by*

$$f_\lambda(x) = \inf_{y \in M} \left\{ f(y) + \frac{1}{2\lambda} d(x, y)^2 \right\},$$

are convex and C^1 smooth on all of M for all $\lambda > 0$. Moreover,

- (i) $f_\lambda \leq f$ for all $\lambda > 0$.
- (ii) f_λ has the same infimum and the same set of minimizers as f .
- (iii) f_λ has the same symmetries as f does.
- (iv) If f is uniformly continuous on bounded sets f_λ converges to f uniformly on bounded sets.

5. Some applications

If X is a Hilbert space (or more generally a reflexive Banach space), it is well known that for every closed convex subset C of X the distance function to C , that is, $x \mapsto d(x, C)$ is convex and C^1 smooth away from C (even though C might not have a smooth boundary), and, as a consequence, every such C can be approximated by C^1 smooth convex bodies. We next show how the results proved above allow us to extend these two theorems to the class of Cartan-Hadamard manifolds (either finite or infinite-dimensional), and we also note that this result completely fails in the sphere S^2 : there are closed convex sets C of arbitrarily small diameter in S^2 such that $x \mapsto d(x, C)$ is not convex on any neighborhood of C .

Corollary 5.1. *Let C be a closed convex subset of a Cartan-Hadamard manifold. Then the distance function to C , $x \mapsto d(x, C) = \inf\{d(x, y) : y \in C\}$ is C^1 smooth on $M \setminus C$ and, moreover, the function $x \mapsto d(x, C)^2$ is C^1 smooth and convex on all of M .*

Proof. Define $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

The function f is lower semicontinuous and convex on M . According to Theorem 3.9, the function $f_\lambda : M \rightarrow \mathbb{R}$,

$$f_\lambda(x) = \inf\left\{f(y) + \frac{1}{2\lambda}d(x, y)^2\right\} = \inf\left\{\frac{1}{2\lambda}d(x, y)^2 : y \in C\right\} = \frac{1}{2\lambda}d(x, C)^2,$$

is C^1 smooth and convex on M for all $\lambda > 0$. By taking $\lambda = 1/2$ we get that the squared distance function to C is C^1 smooth and convex on M . \square

Definition 5.2. We say that a subset C of a Riemannian manifold M is a C^1 smooth convex body of M provided C is closed, convex, has nonempty interior, and ∂C is a one-codimensional C^1 smooth submanifold of M .

Corollary 5.3. *Let C be a closed convex subset of a Cartan-Hadamard manifold, and let U be an open subset of M with $d(C, M \setminus U) > 0$. Then there exists a C^1 smooth convex body D of M such that $C \subset D \subset U$.*

Proof. Since $d(C, \partial U) > 0$ we can take $r = \frac{1}{2}d(C, \partial U)$ and define $D = \{x \in M : d(x, C) \leq r\}$. It is clear that $C \subset D \subset U$, and D happens to be a C^1 smooth convex body because $x \mapsto d(x, C)^2$ is C^1 smooth and convex, and the derivative of the function $d(\cdot, C)^2$ is nonzero at every point $x \in M \setminus C$. (A convex function has a null derivative only at the points, if any, where it attains its minimum.) \square

The following example shows that Theorem 3.9 and the above corollaries are false in general if we do not require that the manifold M has nonpositive sectional curvature.

Example 5.4. Let M be the sphere $x^2 + y^2 + z^2 = 1$ in \mathbb{R}^3 endowed with its usual Riemannian metric. Let C be a closed geodesic segment of diameter less than a number 2ε with $0 < \varepsilon < 1/2$. For instance, with the notation of Remark 4.3, we can take $C = \beta_2([-\varepsilon, \varepsilon])$. Then it is easy to see that the function $d(\cdot, C)^2$ defined on M by

$$d(x, C)^2 = \inf_{y \in C} d(x, y)^2,$$

is not convex on any open neighborhood of C . Indeed, if we consider the geodesic segment $\beta_1 : [-\varepsilon, \varepsilon] \rightarrow S^2$ as in Remark 4.3 then we have that

$$t \mapsto d(\beta_1(t), C)^2 = d(\beta_1(t), \beta_2(t))^2$$

has a strict maximum at $t = 0$, hence it is not convex. Since β_1 can be taken to be arbitrarily close to $C = \beta_2([- \varepsilon, \varepsilon])$ this implies that $d(x, C)^2$ is not convex on any open neighborhood of C .

This example shows that Corollary 5.1 fails in M (hence so does Theorem 3.9).

Finally, it should be noted that there is a strong link between the regularization method we have just presented and the following Hamilton-Jacobi partial differential equation:

$$(*) \begin{cases} \frac{\partial u(t,x)}{\partial t} + \frac{1}{2} \left\| \frac{\partial u(t,x)}{\partial x} \right\|_x^2 = 0, \\ u(0, x) = f(x), \end{cases}$$

where $u : [0, \infty) \times M \rightarrow \mathbb{R}$, $f : M \rightarrow \mathbb{R} \cup \{+\infty\}$. If we assume that M is a finite-dimensional Cartan-Hadamard manifold and f is convex and lower-semicontinuous, then the function

$$u(t, x) := \inf_{y \in M} \left\{ f(y) + \frac{1}{2t} d(x, y)^2 \right\} \quad \text{for } t > 0, \quad u(0, x) = f(x)$$

is the unique viscosity solution of $(*)$ (see [2] for the definition of viscosity solution to Hamilton-Jacobi equations on Riemannian manifolds). This is not very difficult to show directly. Alternatively, one can prove that Theorem 3.6 and section 7.2 of [6] remain true when \mathbb{R}^n is replaced with a finite-dimensional Cartan-Hadamard manifold.

References

- [1] D. Azagra and J. Ferrera, *Proximal calculus on Riemannian manifolds*, *Mediterr. J. Math.* **2** (2005), no. 4, 437–450.
- [2] D. Azagra, J. Ferrera, and F. López-Mesas, *Nonsmooth analysis and Hamilton-Jacobi equations on Riemannian manifolds*, *J. Funct. Anal.* **220** (2005), no. 2, 304–361.
- [3] V. Bangert, *Analytische Eigenschaften konvexer Funktionen auf Riemannschen Mannigfaltigkeiten*, *J. Reine Angew. Math.* **307/308** (1979), 309–324.
- [4] ———, *Über die Approximation von lokal konvexen Mengen*, *Manuscripta Math.* **25** (1978), no. 4, 397–420 (German, with English summary).
- [5] ———, *Totally convex sets in complete Riemannian manifolds*, *J. Differential Geom.* **16** (1981), no. 2, 333–345.
- [6] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, *Mathématiques & Applications*, vol. 17, Springer-Verlag, Berlin, 1994.
- [7] M. Cepedello Boiso, *Approximation of Lipschitz functions by Δ -convex functions in Banach spaces*, *Israel J. Math.* **106** (1998), 269–284.
- [8] ———, *On regularization in superreflexive Banach spaces by infimal convolution formulas*, *Studia Math.* **129** (1998), no. 3, 265–284.
- [9] J. Cheeger and D. Gromoll, *On the structure of complete manifolds of nonnegative curvature*, *Ann. of Math. (2)* **96** (1972), 413–443.

- [10] R. E. Greene and K. Shiohama, *Convex functions on complete noncompact manifolds: topological structure*, Invent. Math. **63** (1981), no. 1, 129–157.
- [11] ———, *Convex functions on complete noncompact manifolds: differentiable structure*, Ann. Sci. École Norm. Sup. (4) **14** (1981), no. 4, 357–367 (1982).
- [12] R. E. Greene and H. Wu, *On the subharmonicity and plurisubharmonicity of geodesically convex functions*, Indiana Univ. Math. J. **22** (1972/73), 641–653.
- [13] ———, *C^∞ convex functions and manifolds of positive curvature*, Acta Math. **137** (1976), no. 3-4, 209–245.
- [14] ———, *C^∞ approximations of convex, subharmonic, and plurisubharmonic functions*, Ann. Sci. École Norm. Sup. (4) **12** (1979), no. 1, 47–84.
- [15] D. Gromoll and W. Meyer, *On complete open manifolds of positive curvature*, Ann. of Math. (2) **90** (1969), 75–90.
- [16] W. Klingenberg, *Riemannian geometry*, de Gruyter Studies in Mathematics, vol. 1, Walter de Gruyter & Co., Berlin-New York, 1982.
- [17] S. Lang, *Fundamentals of differential geometry*, Graduate Texts in Mathematics, vol. 191, Springer-Verlag, New York, 1999.
- [18] J.-M. Lasry and P.-L. Lions, *A remark on regularization in Hilbert spaces*, Israel J. Math. **55** (1986), no. 3, 257–266.
- [19] T. Strömberg, *The operation of infimal convolution*, Dissertationes Math. (Rozprawy Mat.) **352** (1996).
- [20] C. Udriște, *Convex functions and optimization methods on Riemannian manifolds*, Mathematics and its Applications, vol. 297, Kluwer Academic Publishers Group, Dordrecht, 1994.