A Second Order Smooth Variational Principle on Riemannian Manifolds

Daniel Azagra and Robb Fry

Abstract. We establish a second order smooth variational principle valid for functions defined on (possibly infinite-dimensional) Riemannian manifolds which are uniformly locally convex and have a strictly positive injectivity radius and bounded sectional curvature.

1   Introduction and Main Result

It is well known that a continuous function defined on an infinite-dimensional manifold (or on a Banach space) does not generally attain a minimum in situations in which there would typically exist minimizers if the function were defined on a finite-dimensional manifold (for instance when the infimum of the function in the interior of a ball is strictly smaller than the infimum of the function on the boundary of the ball). In fact, as shown in [1], the smooth functions with no critical points are dense in the space of continuous functions on every Hilbert manifold (this result may be viewed as a 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 is quite inconvenient because many important problems of PDEs and of optimization admit equivalent formulations as minimization (or variational) problems, that is, one is given a continuous (or differentiable, or convex, or lower semicontinuous, *etc.*) functional defined on a function space, or on a (usually infinite-dimensional) manifold, and one is asked to find a minimizer of the functional, which will provide a solution of the initial problem. Under several classes of rather restrictive assumptions (for instance convexity of the functional and reflexivity of the space) one may be able to show existence of minimizers, but in many natural cases exact minimizers do not generally exist, and one is thus forced to look for *approximate minimizers*, which will provide *approximate solutions* of the problem. For example, on finite-dimensional, complete and connected Riemannian manifolds the Hopf–Rinow Theorem guarantees that there exists a minimizing geodesic connecting any two points. For infinite-dimensional manifolds this fails in general, and so one must consider approximate minimizing geodesics.

This is why perturbed minimization principles, or variational principles, are important. A variational principle typically asserts that, for a given lower semicontin-

241

uous function $f : \mathcal{X} \to \mathbb{R}$, defined on a given space $\mathcal{X}$, and such that f is bounded below, there exists a function $\varphi : \mathcal{X} \to \mathbb{R}$ belonging to a given class $\mathcal{C}$ such that $f - \varphi$ attains a global minimum and φ can be prescribed to be small in some sense. Historically, Ekeland's variational principle [11] was the first perturbed minimization principle discovered, and one of the most powerful and widely applicable, because it holds for any complete metric space $\mathcal{X}$; in this case the perturbing class $\mathcal{C}$ is the set of small scalar multiples of distance functions to points, that is, the perturbing function φ has the shape of a flat cone. One of the many striking applications of Ekeland's principle was to show the existence of almost minimizing geodesics on complete infinite-dimensional Riemannian manifolds. This gives the approximate version of the Hopf–Rinow theorem mentioned above.

Despite the generality and power of Ekeland's result, there are some limitations to the applicability of this minimization principle due to the fact that the perturbing functions φ are not differentiable. Indeed, there are many situations (for instance when one wants to build a theory of subdifferentiability, in nonsmooth analysis, and in the study of viscosity solutions to Hamilton–Jacobi equations; see [6, 9] and the comments preceding Corollary 1.2 below) in which one needs the perturbing function φ to be differentiable. In order to remedy this deficiency Borwein and Preiss [5] established a *smooth* variational principle in which the space $\mathcal{X}$ is a Banach space with a C^1 smooth norm, and the perturbing functions φ are smooth functions which can be taken to be arbitrarily small and with an arbitrarily small Lipschitz constant. Later on, Deville, Godefroy and Zizler [7, 8], by using a new method of proof based on the use of Baire's category theorem, were able to extend this smooth variational principle to the class of all Banach spaces with smooth bump functions, as well as for higher orders of smoothness.

Smooth variational principles on Riemannian manifolds were not studied until very recently. In [2] a Riemannian version of the Deville–Godefroy–Zizler smooth variational principle was established within the class of complete Riemannian manifolds which are *uniformly bumpable*. This was applied to developing a theory of nonsmooth analysis and to the study of viscosity solutions to Hamilton–Jacobi equations on Riemannian manifolds. Later, in [3], the authors showed that the assumption on uniform bumpability can be dispensed with at least in separable manifolds, by establishing that every Lipschitz function f defined on a separable Riemannian manifold can be uniformly approximated by a sequence of C^∞ smooth Lipschitz functions (f_n) in such a way that the Lipschitz constants of f_n converge to the Lipschitz constant of f (this result easily implies that all separable Riemannian manifolds are uniformly bumpable; however, the question remains unsettled in the nonseparable case).

In this paper we will prove a second order smooth variational principle valid for functions defined on every complete Riemannian manifold of bounded sectional curvature, uniformly locally convex, and with a positive injectivity radius. In fact, the result is valid (see Theorem 3.7 below) for every complete Riemannian manifold which is *second order uniformly bumpable* according to Definition 2.1 given below.

Section 2 of the paper is devoted to providing a sufficient condition for a manifold to be second order uniformly bumpable (namely, that the manifold be uniformly locally convex and have positive injectivity radius and bounded sectional curvature).

The question of whether or not every complete Riemannian manifold is second

order uniformly bumpable is open, and seems to be of a different nature from its first order analogue. Indeed, the first order case can be settled by uniformly approximating a function of the form $a_{x_0,\delta}(x) := \max\{1 - (1/\delta)d(x, x_0), 0\}$ (which is $1/\delta$-Lipschitz) with $C^{1,1}$ functions of Lipschitz constants less than R/δ for some fixed $R > 1$; according to [3] this can be achieved in any separable Riemannian manifold for any $R > 1$. However, the second order case would require one to construct C^2 bump functions with the shape of $a_{x_0,\delta}$ and whose first and second derivatives are bounded by R/δ and R/δ^2 respectively (for some fixed $R > 1$). But it looks doubtful, even in the (noncompact) finite-dimensional case, that this could be done. Rather, to the contrary it seems plausible that curvature effects may prevent the uniformity that we need in the bounds for the second derivatives of the approximations to $a_{x_0,\delta}$ (in a manner that regions $\mathcal{U}$ of M with bigger sectional curvature would require bigger R's in order that the second derivatives of a C^2 approximation to $a_{x_0,\delta}$ be bounded by R/δ^2 when $x_0 \in \mathcal{U}$).

Let us state our main result.

Theorem 1.1 *Let M be a (possibly infinite-dimensional) complete Riemannian manifold. Assume that the sectional curvature of M is bounded, and that M is uniformly locally convex and has a positive injectivity radius. Then for every lower semicontinuous function $f \colon M \to (-\infty, \infty]$ which is bounded below, with $f \not\equiv +\infty$, and for every $\varepsilon > 0$, there exists a C^2 smooth function $\varphi \colon M \to \mathbb{R}$ such that*

(i)   *$f - \varphi$ attains its strong minimum on M;*
(ii)   *$\|\varphi\|_\infty < \varepsilon$;*
(iii)   *$\|d\varphi\|_\infty < \varepsilon$;*
(iv)   *$\|d^2\varphi\|_\infty < \varepsilon$.*

Recall that a function h is said to attain a strong minimum at a point x_0 provided $h(x_0) = \inf_{x \in M} h(x)$ and $\lim_{n \to \infty} d(x_n, x_0) = 0$ whenever (x_n) is a minimizing sequence (that is, if $\lim_{n \to \infty} h(x_n) = h(x_0)$).

In the statement of Theorem 1.1 we used the following notation:

$$\|\varphi\|_\infty = \sup_{x \in M} |\varphi(x)|$$

$$\|d\varphi\|_\infty = \sup_{x \in M} \|d\varphi(x)\|_x$$

$$\|d^2\varphi\|_\infty = \sup_{x \in M} \|d^2\varphi(x)\|_x,$$

where

$$\|d\varphi(x)\| = \sup_{v \in TM_x, \|v\|_x = 1} d\varphi(x)(v) = \sup_{v \in TM_x, \|v\|_x = 1} \langle \nabla\varphi(x), v \rangle_x,$$

and

$$\|d^2\varphi(x)\|_x = \sup_{v \in TM_x, \|v\|_x = 1} |d^2\varphi(x)(v, v)|.$$

Recall also that the Hessian $D^2\varphi$ of a C^2 smooth function φ on M is defined by

$$D^2\varphi(X, Y) = \langle \nabla_X \nabla\varphi, Y \rangle,$$

where $\nabla\varphi$ is the gradient of φ; X, Y are vector fields on M, and $\nabla_X Z$ is the covariant derivative of a vector field Z with respect to a vector field X. The Hessian is a symmetric tensor field of type $(0, 2)$, and for a point $p \in M$ the value $D^2\varphi(X, Y)(p)$ only depends on φ and the vectors $X(p), Y(p) \in TM_p$. So we can define the second derivative of φ at p as the symmetric bilinear form $d^2\varphi(p)\colon TM_p \times TM_p \to \mathbb{R}$

$$(v, w) \mapsto d^2\varphi(p)(v, w) := D^2\varphi(X, Y)(p),$$

where X, Y are any vector fields such that $X(p) = v, Y(p) = w$. A very useful way to compute $d^2\varphi(p)(v, v)$ is to take a geodesic γ with $\gamma'(0) = v$ and calculate

$$\frac{d^2}{dt^2}\varphi\big(\gamma(t)\big)\big|_{t=0},$$

which equals $d^2\varphi(p)(v, v)$. We will often write $d^2\varphi(p)(v)^2$ instead of $d^2\varphi(p)(v, v)$.

Let us fix some notation and terminology that we will be using throughout the paper. $M = (M, g)$ will always be a (possibly infinite-dimensional) complete Riemannian manifold, and $g(x) = \langle \cdot, \cdot \rangle_x$ its Riemannian metric. We refer to [13] for a basic introduction to infinite-dimensional Riemannian manifolds. However, instead of Lang's notation we will employ a more classical notation, such as that of Do Carmo's book [10]. The letters X, Y, Z, V, W will stand for smooth vector fields on M, and $\nabla_Y X$ will always denote the covariant derivative of X with respect to Y. The Riemannian curvature of M will be denoted by R. Recall that the value of $R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]}Z$ at a point $p \in M$ only depends on the values of X, Y, Z at p, and that the sectional curvature of M at a point p with respect to a plane spanned by two vectors $v, w \in TM_p$ is defined by

$$K(p; v, w) = \frac{\langle R(v, w)v, v \rangle}{|v \wedge w|^2},$$

where $|v \wedge w|$ is the area of the parallelogram defined by u, v in TM_p.

Geodesics in M will be denoted by γ, σ, and their velocity fields by γ', σ'. If X is a vector field along γ we will often denote $X'(t) = \frac{D}{dt}X(t) = \nabla_{\gamma'(t)}X(t)$. Recall that X is said to be parallel along γ if $X'(t) = 0$ for all t. The Riemannian distance in M will always be denoted by $d(x, y)$ (defined as the infimum of the lengths of all curves joining x to y in M).

We will often identify the tangent space of M at a point x, denoted by TM_x, with the cotangent space at x, denoted by TM_x^*. The space of bilinear forms on TM_x (respectively symmetric bilinear forms) will be denoted by $\mathcal{L}^2(TM_x)$ or $\mathcal{L}^2(TM_x, \mathbb{R})$ (resp. $\mathcal{L}_s^2(TM_x)$ or $\mathcal{L}_s^2(TM_x, \mathbb{R})$). Also, we will denote by $T_{2,s}(M)$ the tensor bundle of symmetric bilinear forms, that is

$$T_{2,s}(M) = \bigcup_{x \in M} \mathcal{L}_s^2(TM_x, \mathbb{R}),$$

and $T_{2,s}(M)_x = \mathcal{L}_s^2(TM_x, \mathbb{R})$.

We will make extensive use of the exponential mapping $\exp_x$ throughout the paper. Recall that for every $x \in M$ there exists a mapping $\exp_x$, defined on a neighborhood of 0 in the tangent space TM_x, and taking values in M, which is a local diffeomorphism and maps straight line segments passing through 0 onto geodesic segments in M passing through x.

By $i_M(x)$ we will denote the *injectivity radius of M at x*, that is, the supremum of the radius r of all balls $B(0_x, r)$ in TM_x for which $\exp_x$ is a diffeomorphism from $B(0_x, r)$ onto $B(x, r)$. The number $i_M(x)$ is strictly positive for every x. Similarly, $i(M)$ will denote the *global injectivity radius of M*, specifically $i(M) = \inf\{i_M(x) : x \in M\}$.

We will also need to recall some results about convexity in Riemannian manifolds. 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 $d(x, y)$. Every Riemannian manifold is *locally convex*, in the sense that 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 is Whitehead's Theorem). 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 $c_M(x)$ (by the result we have just mentioned, $c_M(x)$ is strictly positive for every x).

Whitehead's theorem gives rise to the notion of *uniformly locally convex* manifold: 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* (defined as $c(M) := \inf\{c_M(x) : x \in M\}$) is strictly positive.

It is also worth noting that the functions $x \mapsto i_M(x)$ and $x \mapsto c_M(x)$ are continuous, see [12].

As noted above, the Hopf–Rinow theorem fails in infinite dimensions, and hence generally geodesics minimizing the distance between two given points do not exist in infinite-dimensional Riemannian manifolds. But, if M is locally uniformly convex and $i(M) > 0$, they do always exist locally. In fact they exist in a *uniformly local* way: for any $r > 0$ with $r < \min\{i(M), c(M)\}$ and for all points x, y with $d(x, y) \leq r$, there is a unique (up to reparameterizations) minimizing geodesic γ connecting x and y (γ is defined by $\gamma(t) = \exp_x(tw_y)$, where $w_y = \exp_x^{-1}(y)$). Moreover, $d(x, y)$ is given by $d(x, y) = \| \exp_x^{-1}(y)\|_x$, see [2, Proposition 3.9]. This will be extensively used in Section 2.

Let us also recall that for a minimizing geodesic $\gamma\colon [0, \ell] \to M$ connecting x to y in M, and for a vector $v \in TM_x$ there is a unique parallel vector field P along γ such that $P(0) = v$. The mapping $TM_x \ni v \mapsto P(\ell) \in TM_y$ is called the *parallel translation* of v along γ and is a linear isometry from TM_x onto TM_y which we will denote by L_{xy}.

For a basic theory of Jacobi fields on infinite-dimensional Riemannian manifolds (and for any other unexplained terms of Riemannian geometry used in Section 2), we refer the reader to [13].

Let us finish this introduction by providing a typical application of the smooth variational principle, namely to show the existence and density of the set of points of subdifferentiability (of second order, in our case), a result which is essential in order to be able to use viscosity solutions to partial differential equations of second order.

For a lower semicontinuous function $f\colon M \to (-\infty, +\infty]$ we define the *second order subjet* of f at a point $x \in M$ as the set

$$J^{2,-} f(x) = \left\{ (d\varphi(x), d^2\varphi(x)) : \varphi \in C^2(M, \mathbb{R}), f - \varphi \text{ attains a local minimum at } x \right\}.$$

For an upper semicontinuous function f on M a similar definition is given for $J^{2,+} f(x)$, where "minimum" is replaced with "maximum".

Given a continuous function $F\colon \mathfrak{X} \to \mathbb{R}$, where $\mathfrak{X} := \{(x, r, \zeta, A) : x \in M,\ r \in \mathbb{R}, \zeta \in TM_x, A \in \mathcal{L}^2_s(TM_x)\}$, one says that an upper semicontinuous function $u\colon M \to \mathbb{R}$ is a *viscosity subsolution* of the equation $F = 0$ provided that

$$F(x, u(x), \zeta, A) \leq 0$$

for all $x \in M$ and $(\zeta, A) \in J^{2,+} u(x)$. Similarly, a *viscosity supersolution* of $F = 0$ on M is a lower semicontinuous function $u\colon M \to \mathbb{R}$ such that

$$F(x, u(x), \zeta, A) \geq 0$$

for every $x \in M$ and $(\zeta, A) \in J^{2,-} u(x)$. If u is both a viscosity subsolution and a viscosity supersolution of $F = 0$, we say that u is a *viscosity solution* of $F = 0$ on M.

In order that this definition is meaningful and can be used to get interesting results about weak solutions to second order PDE in a Riemannian manifold M, one should be able to show that continuous functions have points x, y at which the semijets $J^{2,+} u(x)$ and $J^{2,-} u(y)$ are nonempty (otherwise we would have the nonsensical situation that there exists a continuous function u which is a viscosity solution to *every* equation of the form $F\big(x, u(x), du(x), d^2u(x)\big) = 0$). In finite dimensions this easily follows by local compactness, but in the infinite-dimensional case one has to use a variational principle. The following corollary ensures existence and density of points x where $J^{2,-} u(x)$ is nonempty (obviously a similar result holds for $J^{2,+} u$ with u upper semicontinuous).

Corollary 1.2 *Let M be as in Theorem 1.1, and let $f\colon M \to (-\infty, +\infty]$ be a lower semicontinuous function. Then the set $\{z \in M : J^{2,-} f(z) \neq \varnothing\}$ is dense in the set $\{x \in M : f(x) < \infty\}$.*

Proof Pick any point x_0 with $f(x_0) < \infty$, and any open neighborhood U of x_0. Since M has smooth partitions of unity, there is a C^∞ smooth function $b\colon M \to [0, \infty)$ such that $b(y) > 0$ if and only if $y \in U$. Consider the function $h\colon M \to (-\infty, \infty]$ defined by

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

The function h is lower semicontinuous on M, and C^∞ smooth on U. Then the sum $f + h$ is lower semicontinuous, and $(f + h)(x_0) < +\infty$. According to Theorem 1.1, there exists a C^2 smooth function $\varphi\colon M \to \mathbb{R}$ such that $(f + h) - \varphi$ attains a strong minimum at some point $x \in M$. In fact we have $x \in U$, because this function is valued $+\infty$ outside U. As the function $\varphi - h$ is C^2 smooth on U, and $f - (\varphi - h)$ attains its minimum at x, we conclude that $(d(\varphi - h)(x), d^2(\varphi - h)(x)) \in J^{2,-} f(x) \neq \varnothing$. ∎

2   Second Order Uniformly Bumpable Manifolds

In this section we will show that all complete Riemannian manifolds that are uniformly locally convex and have bounded sectional curvature and a strictly positive injectivity radius are *second order uniformly bumpable*, meaning the following.

Definition 2.1   We will say that a Riemannian manifold M is second order uniformly bumpable provided there exist $R > 1$, $r > 0$ such that for every $z \in M$ and for every $\delta \in (0, r)$ there exists a C^2 smooth function $b \colon M \to [0, 1]$ such that

(i) $b(z) = 1$,
(ii) $b(x) = 0$ whenever $d(x, z) \geq \delta$,
(iii) $\|db\|_\infty \leq R/\delta$,
(iv) $\|d^2 b\|_\infty \leq R/\delta^2$.

One should compare Definition 2.1 with the definition of (first order) *uniformly bumpable manifold* given in [2]: we note that, in the latter one, condition (iv) is omitted and C^2 is replaced with C^1, these are the only differences.

The key to the proof of the main result of this section is the following estimation for the norm of the second derivative of the square of the distance function, for which we have found no reference (of course there are well known estimations in the finite-dimensional case, see for instance [14, Lemma 2.9 and Exercise 4, p. 153], but they all depend on results established exclusively in finite dimensions, and of which we have found no infinite-dimensional versions in the literature).

Proposition 2.2   *If M is a Riemannian manifold whose sectional curvature K is bounded, say $|K| \leq K_0$, and $c(M) > 0$, $i(M) > 0$, then for every r with $0 < r < \min\{i(M), c(M), \pi/2\sqrt{K_0}\}$ and every $z \in M$, the function $\varphi(x) := d(x, z)^2$ is C^∞ smooth on $B(z, r)$ and its second derivative satisfies*

$$\|d^2\varphi(x)\|_x \leq \left(2 + \frac{2}{3} K_0 d(x, z)^2 \right)$$

for all $x \in B(z, r)$.

In the proof of this proposition we will have to use some well known results about the second variation of the arc length and the energy functionals. Let us briefly review the facts that we will be using.

Fix a number r such that $0 < r < \min\{i(M), c(M), \pi/2\sqrt{K_0}\}$, and take two points $x, x_0 \in M$ with $d(x, x_0) < r$. Let γ be the unique minimizing geodesic, parameterized by arc-length, connecting x_0 to x. Denote $\ell = d(x, x_0)$, the length of γ. Consider $\alpha(t, s)$, a smooth variation of γ, that is a smooth mapping $\alpha \colon [0, \ell] \times [-\varepsilon, \varepsilon] \to M$ such that $\alpha(t, 0) = \gamma(t)$ for all $t \in [0, \ell]$. Consider the length and the energy functionals, defined by

$$L(s) = L(\alpha_s) = \int_0^\ell \|\alpha_s'(t)\| \, dt$$

and

$$E(s) = E(\alpha_s) = \int_0^\ell \|\alpha_s'\|^2 \, dt,$$

where α_s is the variation curve defined by $\alpha_s(t) = \alpha(t, s)$ for every $t \in [0, \ell]$. According to the Cauchy–Schwarz inequality (applied to the functions $f \equiv 1$ and $g(t) = \|\alpha_s'(t)\|$ on the interval $[0, \ell]$) we have that

$$(2.1) \qquad\qquad\qquad\qquad L(s)^2 \leq \ell E(s),$$

with equality if and only if $\|\alpha_s'(t)\|$ is constant. In particular we have

$$L(0)^2 = \ell E(0)$$

because $\alpha_0 = \gamma$ is a geodesic.

Therefore, if we furthermore assume that α_s is a geodesic for each s (that is, α is a variation of γ through geodesics) we have that

$$(2.2) \qquad\qquad\qquad\qquad L(s)^2 = \ell E(s)$$

for every $s \in [-\varepsilon, \varepsilon]$.

Now define $\varphi(x) = d(x, x_0)^2$. The function φ is C^∞ smooth on $B(x_0, r)$ because $r < \min\{i(M), c(M)\}$, so on this ball $\exp_{x_0}^{-1}$ is a C^∞ diffeomorphism, and

$$d(x, x_0)^2 = \| \exp_{x_0}^{-1}(x) \|_{x_0}^2.$$

Let us take a vector $v \in TM_x$. We want to calculate

$$d^2\varphi(x)(v)^2,$$

which is given by

$$\frac{d^2}{ds^2} \varphi\big(\sigma_x(s) \big) \big|_{s=0},$$

where $\sigma_x(s) = \exp_x(sv)$. To this end let us denote by $\alpha_s \colon [0, \ell] \to M$ the unique minimizing geodesic joining the point x_0 to the point $\sigma_x(s)$ (notice that now, for $s \neq 0$, α_s is not necessarily parameterized by arc-length), and let us define $\alpha \colon [0, \ell] \times [-\varepsilon, \varepsilon] \to M$ by $\alpha(t, s) = \alpha_s(t)$. Then α is a smooth variation through geodesics of $\gamma(t) = \alpha(t, 0)$ and we have

$$\varphi\big(\sigma_x(s) \big) = L(s)^2 = \ell E(s),$$

and therefore

$$(2.3) \qquad\qquad\qquad\qquad d^2\varphi(x)(v)^2 = \ell E''(0).$$

If we denote the variational field of α by $X(t) = \partial\alpha(t, 0)/\partial s$, then the formula for the second variation of energy (see [13, Chapter XI]) tells us that

$$(2.4) \quad \frac{1}{2}E''(0) = \int_0^\ell \big(\langle X', X' \rangle - \langle R(\gamma', X)\gamma', X \rangle \big) \, dt + \Big\langle \frac{D}{ds}\frac{\partial\alpha}{\partial s}(t, 0), \gamma'(t) \Big\rangle \Big|_{t=0}^{t=\ell},$$

or equivalently

(2.5) $\dfrac{1}{2} E''(0) = -\displaystyle\int_0^\ell \langle X, X'' + R(\gamma', X)\gamma' \rangle\, dt + \langle X(t), X'(t) \rangle\big|_{t=0}^{t=\ell}$

$$+ \left\langle \frac{D}{ds}\frac{\partial\alpha}{\partial s}(t,0), \gamma'(t) \right\rangle\bigg|_{t=0}^{t=\ell},$$

where we denote $X' = \nabla_{\gamma'(t)}X$, and $X'' = \nabla_{\gamma'(t)}X'$.

Note that, since the variation field of a variation through geodesics is always a Jacobi field, and since the points x and x_0 are not conjugate (recall that $r < i(M)$), X is the unique Jacobi field along γ satisfying that $X(0) = 0$, $X(\ell) = v$, that is, X is the unique vector field along γ satisfying

$$X''(t) + R\big(\gamma'(t), X(t)\big)\gamma'(t) = 0, \quad \text{and} \quad X(0) = 0,\ X(\ell) = v.$$

On the other hand, since the curves $s \to \alpha(\ell, s) = \sigma_x(s)$ and $s \to \alpha(0, s) \equiv x_0$ are geodesics, we have that

$$\left\langle \frac{D}{ds}\frac{\partial\alpha}{\partial s}(t,0), \gamma'(t) \right\rangle\bigg|_{t=0}^{t=\ell} = 0.$$

These observations allow us to simplify the formulas (2.4) and (2.5) by dropping the terms that vanish, thus obtaining

(2.6) $\dfrac{1}{2} E''(0) = \displaystyle\int_0^\ell \big(\langle X', X' \rangle - \langle R(\gamma', X)\gamma', X \rangle \big)\, dt,$

and also

(2.7) $\dfrac{1}{2} E''(0) = \langle X(\ell), X'(\ell) \rangle.$

Here we must recall that the right-hand side of (2.6) is called the index form and is denoted by $I(X, X)$. In fact, one can define for all smooth vector fields X, Y along the geodesic $\gamma\colon [0, \ell] \to M$,

$$I(X, Y) := \int_0^\ell \big(\langle X', Y' \rangle - \langle R(\gamma', X)\gamma', Y \rangle \big)\, dt$$

$$= -\int_0^\ell \langle X'' + R(\gamma', X)\gamma', Y \rangle\, dt + \langle X'(t), Y(t) \rangle\big|_{t=0}^{t=\ell}$$

(where the last equality is justified by integrating by parts, see also [13, Proposition XI.1.1]). In particular, when X is a Jacobi field, then

($\sharp$) $I(X, Y) = \langle X', Y \rangle\big|_0^\ell$

for all smooth vector fields Y along γ. It will also be useful to remark that the mapping $(X, Y) \mapsto I(X, Y)$ is bilinear on the smooth vector fields along γ.

By combining (2.3) and (2.6) we get

$$(2.8) \qquad d^2\varphi(x_0)(v)^2 = 2\ell \int_0^\ell \left(\langle X', X' \rangle - \langle R(\gamma', X)\gamma', X \rangle \right) dt,$$

or equivalently

$$(2.9) \qquad d^2\varphi(x_0)(v)^2 = 2\ell \langle X'(\ell), X(\ell) \rangle.$$

We are going to use these formulas to deduce the estimation in the statement of Proposition 2.2, but we will need to combine them with a couple of facts about Jacobi fields, and with the Rauch Comparison Theorem. First, we must use the following.

Lemma 2.3 *Let $\gamma \colon [0, \ell] \to M$ be a geodesic whose length, ℓ, is the distance between its end points, and let Y be a field along γ with $Y(0) = 0$ and $Y(\ell) = 0$. Then $I(Y, Y) \geq 0$.*

Proof This result is stated and proved in [13, Theorem 1.7 of Chapter XI] under the additional assumption that Y is orthogonal to γ, but it is true for any Y (and in fact almost the same proof holds, with an additional remark). We will write the whole proof for the reader's convenience. Define

$$\beta(t, s) = \exp_{\gamma(t)}(sY(t))$$

for $0 \leq t \leq \ell, 0 \leq s \leq \varepsilon$, and for a sufficiently small $\varepsilon > 0$. For each s, $\beta_s(t) := \beta(t, s)$ is a curve, not necessarily a geodesic, joining the end points of γ; that is,

$$\beta_s(0) = \gamma(0), \quad \text{and} \quad \beta_s(\ell) = \gamma(\ell)$$

(because of the assumption that $Y(0) = 0, Y(\ell) = 0$). Moreover, $\beta(t, 0) = \gamma(t)$, so β is a variation of γ which leaves the end points fixed, and the variation field of β is

$$\frac{\partial \beta(t, 0)}{\partial s} = Y(t).$$

Using the formula for the second variation of energy [13, Chapter XI] (and taking into account that the curves $s \mapsto \beta(0, s)$ and $s \mapsto \beta(\ell, s)$ are geodesics, in fact points), we have that

$$(2.10) \qquad \frac{1}{2} E''(0) = \int_0^\ell \left(\langle Y', Y' \rangle - \langle R(\gamma', Y)\gamma', Y \rangle \right) dt := I(Y, Y).$$

On the other hand, according to equation (2.1), and bearing in mind that γ is a minimizing geodesic, we have that

$$\ell E(0) = L(0)^2 \leq L(s)^2 \leq \ell E(s),$$

hence

$$E(0) \leq E(s).$$

That is, a geodesic which minimizes length also minimizes energy. Therefore we have

$$E''(0) \geq 0,$$

which, combined with equation (2.10) yields $I(Y, Y) \geq 0$. ∎

The above lemma allows us to improve Corollary 1.8 of [13, Chapter XI] as follows.

Proposition 2.4 *Let $\gamma\colon [0, \ell] \to M$ be a geodesic whose length, ℓ, is the distance between its end points. Let X be a Jacobi field along γ, and Z any smooth vector field along γ such that $X(0) = Z(0)$ and $X(\ell) = Z(\ell)$. Then*

$$I(X, X) \leq I(Z, Z).$$

In other words, among all vector fields along γ with the same boundary conditions, the unique Jacobi field along γ determined by those conditions minimizes the index form.

Proof Consider $Y := Z - X$ a smooth vector field along γ such that $Y(0) = Y(\ell) = 0$. According to the preceding lemma we have $I(Y, Y) \geq 0$, hence, bearing in mind the facts that the index form is bilinear and symmetric, that $X = Z$ at 0 and ℓ, and that $I(X, Z) = \langle X', Z \rangle |_0^\ell$ (because X is a Jacobi field, see ($\sharp$) above), we get

$$
\begin{aligned}
0 &\leq I(Z - X, Z - X) \\
&= I(Z, Z) + I(X, X) - 2I(X, Z) \\
&= I(Z, Z) + \langle X', X \rangle |_0^\ell - 2\langle X', Z \rangle |_0^\ell \\
&= I(Z, Z) - \langle X', X \rangle |_0^\ell \\
&= I(Z, Z) - I(X, X).
\end{aligned}
$$

Therefore $I(X, X) \leq I(Z, Z)$. ∎

As noted above, we will also need to use an infinite-dimensional version of the Rauch Comparison Theorem, which we next state for the reader's convenience.

Theorem 2.5 (Rauch Comparison Theorem) *Let M and $\widetilde{M}$ be Riemannian manifolds of the same dimension, which may be infinite. Let γ and $\widetilde{\gamma}$ be geodesics in M and $\widetilde{M}$, respectively, parameterized by arc-length, and defined on the same interval $[0, \ell]$. Let J and $\widetilde{J}$ be Jacobi fields along these geodesics, orthogonal to γ and $\widetilde{\gamma}$, respectively. Assume that*

(i)   *$J(0) = \widetilde{J}(0) = 0$, and $J(t), \widetilde{J}(t) \neq 0$ for $0 < t \leq \ell$;*
(ii) *$\|J'(0)\| = \|\widetilde{J}'(0)\|$;*
(iii) *the length of γ is the distance between its end points;*
(iv) *the sectional curvature of M is less than or equal to the sectional curvature of $\widetilde{M}$.*

Then for all $t \in (0, \ell]$, we have

$$\|\widetilde{J}(t)\| \leq \|J(t)\|,$$

and

$$\frac{\langle \widetilde{J}'(t), \widetilde{J}(t) \rangle}{\langle \widetilde{J}(t), \widetilde{J}(t) \rangle} \leq \frac{\langle J'(t), J(t) \rangle}{\langle J(t), J(t) \rangle}.$$

A proof of this result can be found in [13, p. 319]. The last inequality (which is usual in many versions of the R.C.T.) does not appear in Lang's statement of [13, Chapter XI, Theorem 5.1], but is established and used in his proof.

Now we are ready to provide the following.

Proof of Proposition 2.2   According to equations (2.8) and (2.9)

$$d^2\varphi(x_0)(v)^2 = 2\ell \int_0^\ell \left(\langle X', X' \rangle - \langle R(\gamma', X)\gamma', X \rangle \right) dt = 2\ell\langle X'(\ell), X(\ell) \rangle,$$

where X is the unique Jacobi field along γ such that $X(0) = 0$ and $X(\ell) = v$ (and $\gamma \colon [0, \ell] \to M$ is the unique minimizing geodesic, parameterized by arc length, such that $\gamma(0) = x_0$, $\gamma(\ell) = x$). Recall that we are assuming $d(x, x_0) \leq r < \min\{i(M), c(M), \pi/2\sqrt{K_0}\}$.

With the help of Proposition 2.4 and the Rauch Comparison Theorem, we are going to estimate $d^2\varphi(x_0)(v)^2$. Consider the vector field along γ defined by

$$Z(t) = \frac{t}{\ell}P(t),$$

where $P(t)$ is parallel along γ and $P(\ell) = v$. Since P is parallel (that is $P'(t) = \nabla_{\gamma'(t)}P(t) = 0$), we have

$$Z'(t) = \frac{1}{\ell}P(t),$$

and also

$$\langle Z'(t), Z'(t) \rangle = \frac{1}{\ell^2}\|v\|_x^2,$$

because parallel translation is a linear isometry.

On the other hand, by our assumption on the sectional curvature, we have

$$-\langle R\big(\gamma'(t), P(t)\big)\gamma'(t), P(t)\rangle \leq K_0|\gamma'(t) \wedge P(t)|^2 \leq K_0\|\gamma'(t)\|^2\|P(t)\|^2 = K_0\|v\|^2.$$

Therefore

$$
\begin{aligned}
I(Z, Z) &= \int_0^\ell \langle Z'(t), Z'(t) \rangle - \big\langle R\big(\gamma'(t), Z(t)\big) \gamma'(t), Z(t) \big\rangle \, dt \\
&= \int_0^\ell \frac{1}{\ell^2} \|v\|_x^2 - \Big\langle R\Big(\gamma'(t), \frac{t}{\ell} P(t)\Big) \gamma'(t), \frac{t}{\ell} P(t) \Big\rangle \, dt \\
&= \frac{1}{\ell^2} \int_0^\ell \|v\|_x^2 - t^2 \big\langle R\big(\gamma'(t), P(t)\big) \gamma'(t), P(t) \big\rangle \, dt \\
&\leq \frac{1}{\ell^2} \int_0^\ell \|v\|_x^2 + t^2 K_0 \|v\|_x^2 \, dt \\
&= \Big(\frac{1}{\ell} + \frac{\ell}{3} K_0\Big) \|v\|_x^2.
\end{aligned}
$$

Using Proposition 2.4 we deduce that

$$
I(X, X) \leq I(Z, Z) \leq \Big(\frac{1}{\ell} + \frac{\ell}{3} K_0\Big) \|v\|_x^2,
$$

hence that

$$
(2.11) \qquad d^2\varphi(x)(v)^2 = 2\ell I(X, X) \leq \Big(2 + \frac{2\ell^2}{3} K_0\Big) \|v\|_x^2.
$$

In order to conclude the proof we only need to make sure that the left-hand side of this inequality is nonnegative. To obtain a lower bound for $d^2\varphi(x)(v)^2$ we will make use of the Rauch Comparison Theorem stated above. We will compare our manifold M with a manifold $\widetilde{M}$ of constant curvature equal to $K_0 \geq 0$, modelled in the same Hilbert space as M is.

Assume first that X is orthogonal to γ. Take a geodesic $\widetilde{\gamma} \colon [0, \ell] \to \widetilde{M}$ of length ℓ, and a vector $\widetilde{v} \in TM_{\widetilde{\gamma}(0)}$ orthogonal to $\widetilde{\gamma}'(0)$ with $\|\widetilde{v}\| = \|X'(0)\|$. A Jacobi field $\widetilde{X}$ along $\widetilde{\gamma}$ with $\widetilde{X}(0) = 0$ and $\|\widetilde{X}'(0)\| = \|X'(0)\|$ is given by

$$
\widetilde{X}(t) = \frac{\sin(\sqrt{K_0} t)}{\sqrt{K_0}} \widetilde{P}(t),
$$

where $\widetilde{P}(t)$ is the parallel translation of the vector $\widetilde{v}$ along $\widetilde{\gamma}$, with $\widetilde{P}(0) = \widetilde{v}$, see [13, Chapter IX, Proposition 2.12]. Since the sectional curvature of M is less than or equal to K_0, $X(0) = 0 = \widetilde{X}(0)$ but $X(t) \neq 0 \neq \widetilde{X}(t)$ for $0 < t \leq \ell$ (recall that $\gamma, \widetilde{\gamma}$, having lengths less than $i(M)$, do not have conjugate points, see [13, Theorem 3.1 of Chapter IX]), $\|X'(0)\| = \|\widetilde{X}'(0)\|$, X is orthogonal to γ and $\widetilde{X}$ is orthogonal to $\widetilde{\gamma}$, we get from the Rauch Comparison Theorem that

$$
\frac{\langle \widetilde{X}'(t), \widetilde{X}(t) \rangle}{\langle \widetilde{X}(t), \widetilde{X}(t) \rangle} \leq \frac{\langle X'(t), X(t) \rangle}{\langle X(t), X(t) \rangle}
$$

for all $t \in (0, \ell]$. Since $\|\widetilde{P}(t)\| = \|v\|_x = \|X(\ell)\|$ and $\widetilde{X}'(t) = \cos(\sqrt{K_0}t)\widetilde{P}(t)$, we deduce, by taking $t = \ell$, that

$$\sqrt{K_0} \frac{\cos(\sqrt{K_0}\ell)}{\sin(\sqrt{K_0}\ell)} \leq \frac{\langle X'(\ell), X(\ell) \rangle}{\langle X(\ell), X(\ell) \rangle} = \frac{\langle X'(\ell), X(\ell) \rangle}{\|v\|_x^2},$$

hence

(2.12) $$\langle X'(\ell), X(\ell) \rangle \geq \sqrt{K_0} \frac{\cos(\sqrt{K_0}\ell)}{\sin(\sqrt{K_0}\ell)} \|v\|_x^2.$$

In particular, since $\ell \leq r < \pi/2\sqrt{K_0}$, we obtain that

$$\langle X'(\ell), X(\ell) \rangle \geq 0.$$

On the other hand, if X is tangent to γ then

$$X(t) = \pm \frac{t}{\ell} \|v\|_x \gamma'(t),$$

and

$$X'(t) = \pm \frac{1}{\ell} \|v\|_x \gamma'(t),$$

hence

$$\langle X'(\ell), X(\ell) \rangle = \frac{1}{\ell} \|v\|_x^2$$

and we also have

$$\langle X'(\ell), X(\ell) \rangle \geq 0.$$

Now, from Propositions 2.3 and 2.4 of Chapter IX of [13], we know that every Jacobi field X along γ with $X(0) = 0$ can be written in the form

$$X = X^\top + X^\perp,$$

where $X^\top$ and $X^\perp$ are Jacobi fields along γ, $X^\top$ and $(X^\top)'$ are tangent to γ, and $X^\perp$ and $(X^\perp)'$ are orthogonal to γ. In particular $\langle X^\top, (X^\perp)' \rangle = 0$ and $\langle X^\perp, (X^\top)' \rangle = 0$. This implies that

$$\langle X'(t), X(t) \rangle = \langle (X^\top)'(t), X^\top(t) \rangle + \langle (X^\perp)'(t), X^\perp(t) \rangle \geq 0,$$

and hence

(2.13) $$d^2\varphi(x)(v)^2 = 2\ell \langle X'(\ell), X(\ell) \rangle \geq 0,$$

in any case.

By combining (2.13) with equation (2.11) above we finally get

$$0 \leq d^2\varphi(x)(v)^2 \leq \left(2 + \frac{2\ell^2}{3}K_0 \right) \|v\|_x^2$$

for all $v \in TM_x$, which implies

$$\|d^2\varphi(x)\|_x \leq \left(2 + \frac{2\ell^2}{3}K_0 d(x, x_0)^2 \right),$$

and the proof of Proposition 2.2 finishes.   ∎

Corollary 2.6 *If $i(M) > 0$, $c(M) > 0$, and the sectional curvature of M is bounded, then M is second order uniformly bumpable.*

Proof Fix any $r > 0$ with $r < \min\{i(M), c(M)\}$, and define $R = 46 + 2K_0 r^2$, where K_0 is a bound for the sectional curvature of M. For every $\delta \in (0, r)$ find a C^∞ function $\theta \colon \mathbb{R} \to [0, 1]$ such that

(i) $\theta(t) = 1$ for $t \leq 0$,
(ii) $\theta(t) = 0$ for $t \geq \delta^2$,
(iii) $\|\theta'\|_\infty \leq 3/\delta^2$,
(iv) $\|\theta''\|_\infty \leq 10/\delta^4$.

Now, for a given $z \in M$, consider the function $\varphi(x) = d(x, z)^2$, and define $b \colon M \to [0, 1]$ by

$$b(x) = \begin{cases} \theta(\varphi(x)) & \text{if } d(x, z) \leq \delta, \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that b is C^∞ smooth on M and b satisfies conditions (i) and (ii) of Definition 2.1.

In order to estimate $\|db\|_\infty$, first note that, since $x \mapsto d(x, z)$ is 1-Lipschitz, we have

$$(2.14) \qquad\qquad \|d\varphi(x)\| \leq 2d(x, z)$$

for all $x \in B(z, r)$. Then for all $x \in B(z, \delta)$,

$$\|db(x)\| = \left|\theta'\big(\varphi(x)\big)\right| \|d\varphi(x)\| \leq \frac{3}{\delta^2} 2\delta = \frac{6}{\delta},$$

and therefore

$$\|db\|_\infty \leq \frac{6}{\delta} \leq \frac{R}{\delta},$$

so condition (iii) of the definition is met as well. On the other hand, by using (2.14) again, as well as Proposition 2.2, we can estimate $\|d^2 b\|_\infty$ as follows. For every $x \in B(z, \delta)$ we have

$$|d^2 b(x)(v)^2| = |\theta''(\varphi(x))(d\varphi(x)(v))^2 + \theta'(\varphi(x))d^2\varphi(x)(v)^2|$$

$$\leq |\theta''(\varphi(x))(d\varphi(x)(v))^2| + |\theta'(\varphi(x))d^2\varphi(x)(v)^2|,$$

which implies

$$\|d^2\varphi\|_\infty \leq \frac{46 + 2K_0 r^2}{\delta^2} = \frac{R}{\delta^2},$$

that is, condition (iv) of Definition 2.1 is also satisfied.   ∎

3　The Rest of the Proof

In this section we will show that every complete Riemannian manifold which is second-order uniformly bumpable satisfies the natural translation to the Riemannian setting of the second-order DGZ smooth variational principle (established in [7] for all C^2 smooth Banach spaces). The result will be a consequence of several auxiliary lemmas.

Lemma 3.1 *Let M be a second order 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^2 smooth function $b \colon M \to [0, \varepsilon]$ such that:*

(i)　$b(p) = \varepsilon = \|b\|_\infty := \sup_{x \in M} |b(x)|$.
(ii)　$\|db\|_\infty := \sup_{x \in M} \|db(x)\|_x \le C\varepsilon/\delta$.
(iii)　$\|d^2 b\|_\infty := \sup_{x \in M} \|d^2 b(x)\|_x \le C\varepsilon/\delta^2$.
(iv)　$b(x) = 0$ *if* $x \notin B(p, \delta)$.

In particular, $\max\{\|b\|_\infty, \|db\|_\infty, \|d^2 b\|_\infty\} \le C\varepsilon(1 + 1/\delta + 1/\delta^2)$.

Proof In the case when $\varepsilon = 1$ we get a required b from the definition of second order uniform bumpability. If $\varepsilon \neq 1$, it is enough to multiply b by ε. ∎

Lemma 3.2 *The space $Y = \{\varphi \in C^2(M, \mathbb{R}) : \varphi, \|d\varphi\|, \|d^2\varphi\|$ are bounded on $M\}$, endowed with the norm*

$$\|\varphi\|_Y := \max\{\|\varphi\|_\infty, \|d\varphi\|_\infty, \|d^2\varphi\|_\infty\},$$

is a Banach space.

Proof The space $(Y, \|\cdot\|_Y)$ is clearly a normed space. Let us check that it is complete.

Let (φ_n) be a Cauchy sequence in $(Y, \|\cdot\|_Y)$. Since $\|\varphi_n\|_Y \ge \|\varphi_n\|_\infty$ and the space of continuous bounded functions on M with the norm $\|\cdot\|_\infty$ is complete, we know that there exists a continuous bounded function $\varphi \colon M \to \mathbb{R}$ such that

$$\|\varphi_n - \varphi\|_\infty \to 0.$$

We have to see that $\varphi \in C^2(M)$, $d\varphi$ and $d^2\varphi$ are bounded, and $\|d\varphi_n - d\varphi\|_\infty \to 0$, and $\|d^2\varphi_n - d^2\varphi\|_\infty \to 0$. To this end, fix a point $x \in M$ and a number r with $0 < r < i_M(x)$ and such that both $\exp_x \colon B(0_x, r) \subset TM_x \to B(x, r) \subset M$ and its inverse $\exp_x^{-1}$ are 2-Lipschitz diffeomorphisms, and consider the functions $\psi_n(y) = \varphi_n \circ \exp_x$ and $\psi(y) = \varphi \circ \exp_x$, defined on the ball $B(0_x, r)$ in TM_x. We have

$$(3.1) \qquad d\psi_n(w_y)(v) = d\varphi_n(y)\big(d\exp_x(w_y)(v)\big),$$

and

$$(3.2) \qquad d\psi(w_y)(v) = d\varphi(y)\big(d\exp_x(w_y)(v)\big),$$

where we denote $w_y = \exp_x^{-1}(y)$. Therefore

$$\sup_{w_y \in B(0_x, r)} \|d\psi_m(w_y) - d\psi_n(w_y)\| \le 2 \sup_{y \in B(x, r)} \|d\varphi_m(y) - d\varphi_n(y)\|$$

$$\le 2\|d\varphi_m - d\varphi_n\|_\infty$$

and, because (φ_n) is a Cauchy sequence in Y, the definition of $\|\cdot\|_Y$ implies that the right-hand side of the above inequality goes to 0 as $m, n \to \infty$, which shows that (ψ_n) is a Cauchy sequence in the space $\{f \in C^1(B(0,r)) : f \text{ and } df \text{ are bounded}\}$ with the norm $\|f\| = \max\{\|f\|_\infty, \|df\|_\infty\}$. Since this space is complete it follows that ψ_n converges to some $\widetilde{\psi} \in C^1(B(0,r))$, and $d\psi_n$ converges to $d\widetilde{\psi}$, in the norm $\|\cdot\|_\infty$. On the other hand we already know that $\|\varphi_n - \varphi\|_\infty \to 0$, which implies $\|\psi_n - \psi\|_\infty \to 0$, so $\psi = \widetilde{\psi}$ by the uniqueness of the limit. Therefore $\varphi = \psi \circ \exp_x^{-1}$ is C^1 on $B(x,r)$, and since x is arbitrary it follows that $\varphi \in C^1(M)$.

To see that $d\varphi_n$ converges to $d\varphi$ in the norm $\|\cdot\|_\infty$, let us first observe that equations (3.1) and (3.2), together with the facts that $d\exp_x(w_y)$ is a linear isomorphism and $d\psi_n \to d\psi$, imply that

$$\text{(3.3)} \qquad \|d\varphi_n(y) - d\varphi(y)\|_y \to 0,$$

that is $d\varphi_n \to d\varphi$ pointwise on TM^*. Now, since $d\varphi_n$ is a Cauchy sequence in the norm $\|\cdot\|_\infty$, for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$\text{(3.4)} \qquad \|d\varphi_n(y) - d\varphi_m(y)\|_y \le \varepsilon$$

for all $y \in M$, whenever $n, m \ge n_0$. By taking limits as $m \to \infty$ in (3.4), and using (3.3) and continuity of $\|\cdot\|_y$, we deduce that

$$\|d\varphi_n(y) - d\varphi(y)\|_y \le \varepsilon$$

for all $y \in M$, whenever $n \ge n_0$. This shows that $\|d\varphi_n - d\varphi\|_\infty \to 0$.

In order to check that $\varphi \in C^2(M)$ and $d^2\varphi_n \to d^2\varphi$, we need to use the following fact, which relates the second derivatives of ψ_n and φ_n.

Fact 3.3 *Let $f \colon M \to \mathbb{R}$ be a C^2 smooth function, and define $h = f \circ \exp_x$ on a neighborhood of a point $0 \in TM_x$. Let $\widetilde{V}$ be a vector field defined on a neighborhood of 0 in TM_x, and consider the vector field defined by $V(y) = d\exp_x(w_y)\big(\widetilde{V}(w_y)\big)$ on a neighborhood of x, where $w_y := \exp_x^{-1}(y)$, and let*

$$\sigma_y(t) = \exp_x\big(w_y + t\widetilde{V}(w_y)\big).$$

Then we have that

$$D^2 h(\widetilde{V}, \widetilde{V})(w_y) = D^2 f(V, V)(y) + \langle \nabla f(y), \sigma_y''(0)\rangle.$$

The proof of this fact is just a calculation, see [4, Lemma 2.7]. We apply this with $\widetilde{V}(w) = v$, a constant field on TM_x, $f = \varphi_n$, $h = \psi_n$ to obtain

$$\text{(3.5)} \quad d^2\psi_n(w_y)(v,v) = d^2\varphi_n(y)\big(d\exp_x(w_y)(v), d\exp_x(w_y)(v)\big)$$
$$+ \langle \nabla\varphi_n(y), \sigma_y''(0)\rangle,$$

where

$$\sigma_y(t) = \exp_x(w_y + tv).$$

Since the vector field $M \ni y \mapsto \sigma_y''(0) \in TM$ is continuous and $\sigma_x''(0) = 0$, we can assume without loss of generality that $r > 0$ is small enough so that $\|\sigma_y''(0)\|_y \leq 1$ for all $y \in B(x, r)$. Then, again using the fact that $\exp_x$ is 2-biLipschitz on $B(x, r)$, we get from (3.5) that

$$\sup_{w_y \in B(0,r)} \|d^2\psi_n(w_y) - d^2\psi_m(w_y)\|$$

$$\leq 2 \sup_{y \in B(x,r)} \|d^2\varphi_n(y) - d^2\varphi_n(y)\| + \sup_{y \in B(x,r)} \|d\varphi_m(y) - d\varphi_n(y)\|$$

$$\leq 2\|d^2\varphi_n - d^2\varphi_m\|_\infty + \|d\varphi_n - d\varphi_m\|_\infty,$$

which implies that (ψ_n) is a Cauchy sequence in the space $\{\, f \in C^2(B(0,r)) : f, df,$ and d^2f are bounded$\}$ with the norm $\|f\| = \max\{\|f\|_\infty, \|df\|_\infty, \|d^2f\|_\infty\}$. This space is well known to be complete, hence there exists some $\widetilde{\psi} \in C^2(B(0,r))$ such that ψ_n converges to $\widetilde{\psi}$, and $d^2\psi_n$ converges to $d^2\widetilde{\psi}$, in the norm $\|\cdot\|_\infty$. Since we already know that $\|\psi_n - \psi\|_\infty \to 0$, we get that $\psi = \widetilde{\psi}$, and therefore $\varphi = \psi \circ \exp_x^{-1}$ is C^2 on $B(x, r)$. It follows that $\varphi \in C^2(M)$.

Moreover, equation (3.5) (and the same equation replacing φ_n with φ and ψ_n with ψ), together with the facts that $d\exp_x(w_y)$ is a linear isomorphism, and that $d\varphi_n \to \varphi$, imply that

$$\|d^2\varphi_n(y) - d^2\varphi(y)\|_y \to 0,$$

for each $y \in B(x, r)$, that is $d^2\varphi_n \to d^2\varphi$ pointwise on $T_{2,s}(M)$.

By combining this with the fact that $(d^2\varphi_n)$ is a Cauchy sequence in the norm $\|\cdot\|_\infty$, one can easily deduce (as in the case of $(d\varphi_n)$) that $\|d^2\varphi_n - d^2\varphi\|_\infty \to 0$. ∎

In the sequel $B(\varphi, r)$ stands for the open ball of center φ and radius r in the Banach space Y.

Lemma 3.4 *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:*

(i)   *$\|\varphi\| \geq \|\varphi\|_\infty = \sup\{|\varphi(x)| : x \in M\}$ for every $\varphi \in Y$.*
(ii)   *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 $b(p) = \varepsilon$, $\|b\|_Y \leq C\varepsilon(1 + 1/\delta + 1/\delta^2)$, 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 $\mathrm{Dom}(f) = \{x \in M \mid f(x) < +\infty\} \neq \varnothing$. 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 The proof follows the lines of that of Lemma 3.13 in [2] (which in turn is very similar to the original proof of [8]), with some small changes. We write the complete proof (rather than just indicating the changes) for the reader's convenience and for completeness.

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

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

Claim 3.5 U_n *is open.*

Take $\varphi \in U_n$. By the definition of U_n there exists $x_0 \in M$ such that $(f + \varphi)(x_0) < \inf\{(f + \varphi)(x) \mid x \in M \backslash B(x_0, \frac{1}{n})\}$. Set $2\rho = \inf\{(f + \varphi)(x) \mid x \in M \backslash B(x_0, \frac{1}{n})\} - (f + \varphi)(x_0) > 0$. Then since $\| \cdot \|_Y \geq \| \cdot \|_\infty$, we get that $B_Y(\varphi, \rho) \subset B_\infty(\varphi, \rho) \subset U_n$.

Claim 3.6 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 $(f + \varphi)(x_0) < \inf\{(f + \varphi)(x) \mid x \in M\} + \varepsilon$. Set now $\delta = 1/n < r$, and use condition (2) to find a function $b \in Y$ such that $b(x_0) = \varepsilon$, $\|b\|_Y \leq C(n^2 + n + 1)\varepsilon$, and $b(x) = 0$ for $x \notin B(x_0, \frac{1}{n})$. Then $(f + \varphi)(x_0) - b(x_0) < \inf\{(f + \varphi)(x) \mid x \in M\}$ and, if we define $h = -b$, we have

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

Since $\inf\{(f + \varphi)(x) \mid x \notin B(x_0, \frac{1}{n})\} = \inf\{(f + \varphi + h)(x) \mid x \notin B(x_0, \frac{1}{n})\}$, it is obvious that the above inequality implies that $\varphi + h \in U_n$. On the other hand, we have $\|h\|_Y \leq C(n^2 + n + 1)\varepsilon$. 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 $n \geq N$, take $x_n \in M$ such that $(f + \varphi)(x_n) < \inf\{(f + \varphi)(x) \mid x \notin B(x_n, \frac{1}{n})\}$. Clearly, $x_k \in B(x_n, \frac{1}{n})$ if $k \geq n$, 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

$$(f + \varphi)(x_0) \leq \liminf(f + \varphi)(x_n)$$

$$\leq \liminf \left[\inf \left\{ (f + \varphi)(x) \ \middle| \ x \in M \setminus B\left(x_0, \frac{1}{n}\right) \right\} \right]$$

$$= \inf \left\{ \inf \left\{ (f + \varphi)(x) \ \middle| \ x \in M \setminus B\left(x_0, \frac{1}{n}\right) \right\} : n \in \mathbb{N}, n \geq N \right\}$$

$$= \inf\{ (f + \varphi)(x) \mid x \in M \setminus \{x_0\}\},$$

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 $(f + g)(y_n) \rightarrow (f + g)(x_0)$ and (y_n) does not converge to x_0. We may assume $d(y_n, x_0) \geq \varepsilon$ 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

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

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

By combining Lemmas 3.1, 3.2 and 3.4 we immediately deduce the following.

Theorem 3.7 *Let M be a second order uniformly bumpable complete Riemannian manifold. Then for every lower semicontinuous function $f\colon M \to (-\infty,\infty]$ that is bounded below with $f \not\equiv +\infty$, and for every $\varepsilon > 0$, there exists a C^2 smooth function $\varphi\colon M \to \mathbb{R}$ such that*

(i)   *$f - \varphi$ attains its strong minimum on M,*
(ii) *$\|\varphi\|_\infty < \varepsilon$,*
(iii) *$\|d\varphi\|_\infty < \varepsilon$,*
(iv) *$\|d^2\varphi\|_\infty < \varepsilon$.*

Finally, Theorem 1.1 follows from Theorem 3.7 and Corollary 2.6.

References

[1]   D. Azagra and M. Cepedello Boiso, *Uniform approximation of continuous mappings by smooth mappings with no critical points on Hilbert manifolds.* Duke Math. J. **124**(2004), 47–66. doi:10.1215/S0012-7094-04-12412-1

[2]   D. Azagra, J. Ferrera, and F. López-Mesas, *Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds.* J. Funct. Anal. **220**(2005), 304–361.   doi:10.1016/j.jfa.2004.10.008

[3]   D. Azagra, J. Ferrera, F. López-Mesas, and Y. Rangel, *Smooth approximation of Lipschitz functions on Riemannian manifolds.* J. Math. Anal. Appl. **326**(2007), 1370–1378. doi:10.1016/j.jmaa.2006.03.088

[4]   D. Azagra, J. Ferrera, and B. Sanz, *Viscosity solutions to second order partial differential equations on Riemannian manifolds.* J. Differential Equations **245**(2008), no. 2, 307–336. doi:10.1016/j.jde.2008.03.030

[5]   J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions.* Trans. Amer. Math. Soc. **303**(1987), 517–527. doi:10.2307/2000681

[6]   Jonathan M. Borwein and Qiji J. Zhu, *Viscosity solutions and viscosity subderivatives in smooth Banach spaces with applications to metric regularity.* SIAM J. Control Optim. **34**(1996), 1568–1591. doi:10.1137/S0363012994268801

[7]   R. Deville, G. Godefroy, and V. Zizler, *Smoothness and renormings in Banach spaces.* Pitman Monographs and Surveys in Pure and Applied Mathematics 64. Longman, 1993.

[8]   _____, *A smooth variational principle with applications to Hamilton–Jacobi equations in infinite dimensions.* J. Funct. Anal. **111**(1993), 197–212.   doi:10.1006/jfan.1993.1009

[9]   R. Deville and N. Ghoussoub, *Perturbed minimization principles and applications.* In: Handbook of the Geometry of Banach Spaces, Vol. I, North-Holland, Amsterdam, 2001, 393–435.

[10] M. P. do Carmo, *Riemannian Geometry.* Mathematics: Theory and Applications, Birkhäuser Boston, 1992.

[11] I. Ekeland, *On the variational principle.* J. Math. Anal. Appl. **47**(1974), 324–353. doi:10.1016/0022-247X(74)90025-0

[12] W. Klingenberg, *Riemannian Geometry.* Walter de Gruyter Studies in Mathematics 1. Berlin–New York, 1982.

[13] S. Lang, *Fundamentals of Differential Geometry.* Graduate Texts in Math. 191. Springer-Verlag, New York, 1999.

[14] T. Sakai, *Riemannian Geometry.* Transl. Math. Monogr.149. American Mathematical Society, Providence, RI, 1992

[15] M. Spivak, *A Comprehensive Introduction to Differential Geometry, Vol. IV.* Second edition. Publish or Perish, Wilmington, DE, 1979.

ICMAT(CSIC-UAM-UC3-UCM), Departamento de Análisis Matemático, Facultad Ciencias Matemáticas, Universidad Complutense, 28040 Madrid, Spain
e-mail: azagra@mat.ucm.es

Department of Mathematics and Statistics, School of Advanced Technologies and Mathematics, Thompson Rivers University, Kamloops, BC V2C 2N5
e-mail: rfry@tru.ca