Global Analysis & the Banach Manifold of Class H^1 Curves

Thiago Brevidelli Garcia

July 2022

Contents

1 Introduction

Known as *global analysis*, or sometimes *non-linear functional analysis*, the field of study dedicated to the understanding of infinite-dimensional manifolds has seen remarkable progress in the past several decades. Among numerous discoveries, perhaps the greatest achievement in global analysis in the last century was the recognition of the fact that many interesting function spaces possess natural differentiable structures – which are usually infinite-dimensional.

As it turns out, many local problems regarding maps between finite-dimensional manifolds can be translated to global questions about the geometry of function spaces – hence the name "*global* analysis". More specifically, a remarkable number of interesting geometric objects can be characterized as "critical points" of functionals in functions spaces. The usual suspects are, of course, geodesics and minimal submanifolds in general, but there are many other interesting examples: harmonic functions, Einstein metrics, periodic solutions to Hamiltonian vector fields, etc. [5, ch. 11].

Such objects are the domain of the so called *calculus of variations*, which is generally concerned with finding functions that minimize or maximize a given functional, known as the *action functional*, by subjecting such functions to "small variations" – which is known as *the variational method*. The meaning of "small variations" have historically been a very dependent on the context of the problem at hand. Only recently, with the introduction of the tools of global analysis, the numerous ad-hoc methods under the umbrella of "variational method" have been unified into a coherent theory, which we describe in the following.

By viewing the class of functions we're interested in as a – most likely infinite-dimensional – manifold $\mathscr{F}$ and the action functional as a smooth functional $f : \mathscr{F} \longrightarrow \mathbb{R}$ we can find minimizing and maximizing functions by studying the critical points of f. More generally, modern calculus of variations is concerned with the study of critical points of smooth functionals $\Gamma(E) \longrightarrow \mathbb{R}$, where $E \longrightarrow M$ is a smooth fiber bundle over a finite-dimensional manifold M and Γ is a given section functor, such as smooth sections, continuous sections or Sobolev sections – notice that by taking $E = M \times N$ the manifold $\Gamma(E)$ is naturally identified with a space of functions $M \longrightarrow N$, which gets us back to the original case.

In these notes we hope to provide a very brief introduction to modern theory the calculus of variations by exploring one of the simplest concrete examples of the previously described program. We study the differential structure of the Banach manifold $H^1(I, M)$ of class H^1 curves in a finite-dimensional Riemannian manifold M, which encodes the solution to the *classic* variational problem: that of geodesics. Hence the particular action functional we are interested is the infamous *energy functional*

$$\begin{aligned} E : H^1(I, M) &\longrightarrow \mathbb{R} \\ \gamma &\longmapsto \frac{1}{2} \int_0^1 \|\dot{\gamma}(t)\|^2 \, \mathrm{d}t, \end{aligned}$$

as well as the *length functional*

$$\begin{aligned} L : H^1(I, M) &\longrightarrow \mathbb{R} \\ \gamma &\longmapsto \int_0^1 \|\dot{\gamma}(t)\| \, \mathrm{d}t \end{aligned}$$

In section 2 we will describe the differential structure of $H^1(I, M)$ and its canonical Riemannian metric. In section 3 we study the critical points of the energy functional E and describe how the fundamental results of the classical theory of the calculus of variations in the context of Riemannian manifolds can be reproduced in our new setting. Other examples of function spaces are explored in detail in [2, sec. 6]. The 11th chapter of [5] is also a great reference for the general theory of spaces of sections of fiber bundles.

We should point out that we will primarily focus on the broad strokes of the theory ahead and that we will leave many results unproved. The reasoning behind this is twofold. First, we don't want to bore the reader with the numerous technical details of some of the constructions we'll discuss in the following. Secondly, and this is more important, these notes are meant to be concise. Hence we do not have the necessary space to discuss neither technicalities nor more involved applications of the theory we will develop.

In particular, we leave the intricacies of Palais' and Smale's discussion of condition (C) – which can be seen as a substitute for the failure of a proper Hilbert space to be locally compact [3, ch. 2] – and its applications to the study of closed geodesics out of these notes. As previously stated, many results are left unproved, but we will include references to other materials containing proofs. We'll assume basic knowledge of differential and Riemannian geometry, as well as some familiarity with the classical theory of the calculus of variations – see [1, ch. 5] for the classical approach.

Before moving to the next section we would like to review the basics of the theory of real Banach manifolds.

1.1 Banach Manifolds

While it is certainly true that Banach spaces can look alien to someone who has never ventured outside of the realms of Euclidean space, Banach manifolds are surprisingly similar to their finite-dimensional counterparts. As we'll soon see, most of the usual tools of differential geometry can be quite easily translated to the context of Banach manifolds[1]. The reason behind this is simple: it turns out that calculus has nothing to do with $\mathbb{R}^n$.

What we mean by this last statement is that none of the fundamental ingredients of calculus – the ones necessary to define differentiable functions in $\mathbb{R}^n$, namely the fact that $\mathbb{R}^n$ is a complete normed space – are specific to $\mathbb{R}^n$. In fact, these ingredients are precisely the features of a Banach space. Thus we may naturally generalize calculus to arbitrary Banach spaces, and consequently generalize smooth manifolds to spaces modeled after Banach spaces. We begin by the former.

[1]The real difficulties with Banach manifolds only show up while proving certain results, and are mainly due to complications regarding the fact that not all closed subspaces of a Banach space have a closed complement.

Definition 1.1. Let V and W be Banach spaces and $U \subseteq V$ be an open subset. A continuous map $f : U \longrightarrow W$ is called *differentiable at* $p \in U$ if there exists a continuous linear operator $df_p \in \mathscr{L}(V, W)$ such that

$$\frac{\|f(p+h) - f(p) - df_p h\|}{\|h\|} \longrightarrow 0$$

as $h \longrightarrow 0$ in V.

Definition 1.2. Given Banach spaces V and W and an open subset $U \subseteq V$, a continuous map $f : U \longrightarrow W$ is called *differentiable of class* C^1 if f is differentiable at p for all $p \in U$ and the *derivative map*

$$\begin{aligned} df : U &\longrightarrow \mathscr{L}(V, W) \\ p &\longmapsto df_p \end{aligned}$$

is continuous. Since $\mathscr{L}(V, W)$ is a Banach space under the operator norm, we may recursively define functions of class C^n for $n > 1$: a function $f : U \longrightarrow W$ of class C^{n-1} is called *differentiable of class* C^n if the map[2]

$$d^{n-1} f : U \longrightarrow \mathscr{L}(V, \mathscr{L}(V, \cdots \mathscr{L}(V, W))) \cong \mathscr{L}(V^{\otimes n}, W)$$

is of class C^1. Finally, a map $f : U \longrightarrow W$ is called *differentiable of class* C^∞ or *smooth* if f is of class C^n for all $n > 0$.

The following lemma is also of huge importance, and it is known as *the chain rule*.

Lemma 1.1. *Given Banach spaces* V_1*,* V_2 *and* V_3*, open subsets* $U_1 \subseteq V_1$ *and* $U_2 \subseteq V_2$ *and two smooth maps* $f : U_1 \longrightarrow U_2$ *and* $g : U_2 \longrightarrow V_3$*, the composition map* $g \circ f : U_1 \longrightarrow V_3$ *is smooth and its derivative is given by*

$$d(g \circ f)_p = dg_{f(p)} \circ df_p$$

As promised, these simple definitions allows us to expand the usual tools of differential geometry to the infinite-dimensional setting. In fact, in most cases it suffices to simply copy the definition of the finite-dimensional case. For instance, as in the finite-dimensional case we may call a map between Banach manifolds M and N *smooth* if it can be locally expressed as a smooth function between open subsets of the model spaces. As such, we will only provide the most important definitions: those of a Banach manifold and its tangent space at a given point. Complete accounts of the subject can be found in [3, ch. 1] and [4, ch. 2].

Definition 1.3. A Banach manifold M is a Hausdorff topological space endowed with a maximal atlas $\{(U_i, \varphi_i)\}_i$, i.e. an open cover $\{U_i\}_i$ of M and homeomorphisms $\varphi_i : U_i \longrightarrow \varphi_i(U_i) \subseteq V_i$ – known as *charts* – where

- **(i)** Each V_i is a Banach space
- **(ii)** For each i and j, $\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \subseteq V_j \longrightarrow \varphi_i(U_i \cap U_j) \subseteq V_i$ is a smooth map
- **(iii)** $\{(U_i, \varphi_i)\}_i$ is maximal with respect to the items above

Definition 1.4. Given a Banach manifold M with maximal atlas $\{(U_i, \varphi_i)\}_i$ and $p \in M$, the tangent space T_pM of M at p is the quotient of the space $\{\gamma : (-\epsilon, \epsilon) \longrightarrow M \mid \gamma \text{ is smooth}, \gamma(0) = p\}$ by the equivalence relation that identifies two curves γ and η such that $(\varphi_i \circ \gamma)'(0) = (\varphi_i \circ \eta)'(0)$ for all i with $p \in U_i$.

Definition 1.5. Given $p \in M$ and a chart $\varphi_i : U_i \longrightarrow V_i$ with $p \in U_i$, let

$$\begin{aligned} \phi_{p,i} : T_pM &\longrightarrow V_i \\ [\gamma] &\longmapsto (\varphi_i \circ \gamma)'(0) \end{aligned}$$

[2] Here we consider the *projective tensor product* of Banach spaces. See [3, ch. 1].

Proposition 1.1. *Given $p \in M$ and a chart φ_i with $p \in U_i$, $\phi_{p,i}$ is a linear isomorphism. For any given charts φ_i, φ_j, the pullback of the norms of V_i and V_j by φ_i and φ_j respectively define equivalent norms in T_pM. In particular, any choice chart gives T_pM the structure of a topological vector space, and this topology is independent of this choice*[3].

Proof. The first statement about $\phi_{p,i}$ being a linear isomorphism should be clear from the definition of T_pM. The second statement about the equivalence of the norms is equivalent to checking that $\phi_{p,i} \circ \phi_{p,j}^{-1} : V_j \longrightarrow V_i$ is continuous for each i and j with $p \in U_i$ and $p \in U_j$.

But this follows immediately from the identity

$$\begin{aligned}(\phi_{p,i} \circ \phi_{p,j}^{-1})v &= (\varphi_i \circ \varphi_j^{-1} \circ \gamma_v)'(0) \\ \text{(chain rule)} &= d(\varphi_i \circ \varphi_j^{-1})_{\varphi_j(p)}\dot{\gamma}_v(0) \\ &= d(\varphi_i \circ \varphi_j^{-1})_{\varphi_j(p)}v\end{aligned}$$

where $v \in V_j$ and $\gamma_v : (-\epsilon, \epsilon) \longrightarrow V_j$ is any smooth curve with $\gamma_v(0) = \varphi_j(p)$ and $\dot{\gamma}_v(0) = v$: $\phi_{p,i} \circ \phi_{p,j}^{-1} = d(\varphi_i \circ \varphi_j^{-1})_{\varphi_j(p)}$ is continuous by definition. ∎

Notice that a single Banach manifold may be "modeled after" multiple Banach spaces, in the sense that the V_i's of definition 1.3 may vary with i. Lemma 1.1 implies that for each i and j with $p \in U_i \cap U_j$, $(d\varphi_i \circ \varphi_j^{-1})_{\varphi_j(p)} : V_j \longrightarrow V_i$ is a continuous linear isomorphism, so that we may assume that each connected component of M is modeled after a single Banach space V. It is sometimes convenient, however, to allow ourselves the more lenient notion of Banach manifold afforded by definition 1.3.

We should also note that some authors assume that both the V_i's and M itself are *separable*, in which case the assumption that M is Hausdorff is redundant. Although we are primarily interested in manifolds modeled after separable spaces, in the interest of affording ourselves a greater number of examples we will *not* assume separability – unless explicitly stated otherwise. Speaking of examples...

Example 1.1. Any Banach space V can be seen as a Banach manifold with atlas given by $\{(V, \mathrm{id} : V \longrightarrow V)\}$ – sometimes called *an affine Banach manifold*. In fact, any open subset $U \subseteq V$ of a Banach space V is a Banach manifold under a global chart $\mathrm{id} : U \longrightarrow V$.

Example 1.2. The group of units $A^\times$ of a Banach algebra A is an open subset, so that it constitutes a Banach manifold modeled after A [2, sec. 3]. In particular, given a Banach space V the group $\mathrm{GL}(V)$ of continuous linear isomorphisms $V \longrightarrow V$ is a – possibly non-separable – Banach manifold modeled after the space $\mathscr{L}(V) = \mathscr{L}(V, V)$ under the operator norm: $\mathrm{GL}(V) = \mathscr{L}(V)^\times$.

Example 1.3. Given a complex Hilbert space H, the space $\mathrm{U}(H)$ of unitary operators $H \longrightarrow H$ – endowed with the topology of the operator norm – is a Banach manifold modeled after the closed subspace $\mathfrak{u}(H) \subseteq \mathscr{L}(H)$ of continuous skew-symmetric operators $H \longrightarrow H$ [6, p. 4].

These last two examples are examples of Banach Lie groups – i.e. Banach manifolds endowed with a group structure whose group operations are smooth. Perhaps more interesting to us is the fact that these are both examples of function spaces. Having reviewed the basics of the theory of Banach manifolds we can proceed to our in-depth exploration of a particular example, that of the space $H^1(I, M)$.

2 The Structure of $H^1(I, M)$

Throughout this section let M be a finite-dimensional Riemannian manifold. As promised, in this section we will highlight the differential and Riemannian structures of the space $H^1(I, M)$ of class H^1 curves in a M. The first question we should ask ourselves is an obvious one: what is $H^1(I, M)$? Specifically, what is a class H^1 curve in M?

[3] In general T_pM is not a normed space, since the norms induced by two distinct choices of chard need not to coincide. Nevertheless, the topology induced by these norms is the same.

Given an interval I, recall that a continuous curve $\gamma : I \longrightarrow \mathbb{R}^n$ is called *a class* H^1 curve if γ is absolutely continuous, $\dot\gamma(t)$ exists for almost all $t \in I$ and $\dot\gamma \in H^0(I, \mathbb{R}^n) = L^2(I, \mathbb{R}^n)$. It is a well known fact that the so called *Sobolev space* $H^1([0,1], \mathbb{R}^n)$ of all class H^1 curves in $\mathbb{R}^n$ is a Hilbert space under the inner product given by

$$\langle \gamma, \eta \rangle_1 = \int_0^1 \gamma(t) \cdot \eta(t) + \dot\gamma(t) \cdot \dot\eta(t) \, \mathrm{d}t$$

Finally, we may define…

Definition 2.1. Given an n-dimensional manifold M, a continuous curve $\gamma : I \longrightarrow M$ is called *a class* H^1 curve if $\varphi_i \circ \gamma : J \longrightarrow \mathbb{R}^n$ is a class H^1 curve for any chart $\varphi_i : U_i \subseteq M \longrightarrow \mathbb{R}^n$ – i.e. if γ can be locally expressed as a class H^1 curve in terms of the charts of M. We'll denote by $H^1(I, M)$ the set of all class H^1 curves $I \longrightarrow M$.

Remark. From now on we fix $I = [0,1]$.

Notice in particular that every piece-wise smooth curve $\gamma : I \longrightarrow M$ is a class H^1 curve. This answer raises and additional question though: why class H^1 curves? The classical theory of the calculus of variations – as described in [1, ch. 5] for instance – is usually exclusively concerned with the study of piece-wise smooth curves, so the fact that we are now interested a larger class of curves – highly non-smooth curves, in fact – *should* come as a surprise to the reader.

To answer this second question we return to the case of $M = \mathbb{R}^n$. Denote by $C'^\infty(I, \mathbb{R}^n)$ the space of piece-wise curves in $\mathbb{R}^n$. As described in section 1, we would like $C'^\infty(I, \mathbb{R}^n)$ to be a Banach manifold under which both the energy functional and the length functional are smooth maps. As most function spaces, $C'^\infty(I, \mathbb{R}^n)$ admits several natural topologies. Some of the most obvious candidates are the uniform topology and the topology of the $\|\cdot\|_0$ norm, which are the topologies induced by the norms

$$\begin{aligned} \|\gamma\|_\infty &= \sup_t \|\gamma(t)\| \\ \|\gamma\|_0 &= \sqrt{\int_0^1 \|\gamma(t)\|^2 \, \mathrm{d}t} \end{aligned}$$

respectively.

The problem with the first candidate is that $L : C'^\infty(I, \mathbb{R}^n) \longrightarrow \mathbb{R}$ is not a continuous map under the uniform topology. This can be readily seen by approximating the curve

$$\begin{aligned} \gamma : I &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t, 1-t) \end{aligned}$$

with “staircase curves” $\gamma_n : I \longrightarrow \mathbb{R}^n$ for larger and larger values of n, as shown in figure 1: clearly $\gamma_n \longrightarrow \gamma$ in the uniform topology, but $L(\gamma_n) = 2$ does not approach $L(\gamma) = \sqrt{2}$ as n approaches ∞.

The issue with this particular example is that while $\gamma_n \longrightarrow \gamma$ uniformly, $\dot\gamma_n$ does not converge to $\dot\gamma$ in the uniform topology. This hints at the fact that in order for E and L to be continuous maps we need to control both γ and $\dot\gamma$. Hence a natural candidate for a norm in $C'^\infty(I, \mathbb{R}^n)$ is

$$\|\gamma\|_1^2 = \|\gamma\|_0^2 + \|\dot\gamma\|_0^2 \,,$$

which is, of course, the norm induced by the inner product $\langle \, , \rangle_1$ – here $\|\cdot\|_0$ denotes the norm of $H^0(I, \mathbb{R}^n) = L^2(I, \mathbb{R}^n)$.

The other issue we face is one of completeness. Since $\mathbb{R}^n$ has a global chart, we expect $C'^\infty(I, \mathbb{R}^n)$ to be affine too. In other words, it is natural to expect $C'^\infty(I, \mathbb{R}^n)$ to be Banach space. In particular, $C'^\infty(I, \mathbb{R}^n)$ must be complete. This is unfortunately not the case for $C'^\infty(I, \mathbb{R}^n)$ in the $\|\cdot\|_1$ norm, but we can consider its completion. Lo and behold, a classical result by Lebesgue establishes that this completion just so happens to coincide with $H^1(I, \mathbb{R}^n)$.

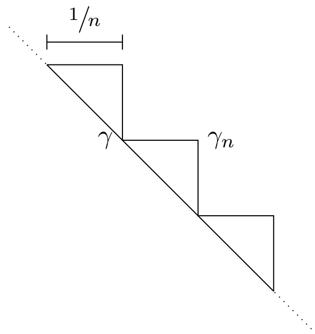

Figure 1: A diagonal line representing the curve γ overlaps a staircase-like curve γ_n, whose steps measure $1/n$ in width and height.

It's also interesting to note that the completion of $C'^{\infty}(I, \mathbb{R}^n)$ with respect to the norms $\|\cdot\|_\infty$ and $\|\cdot\|_0$ are $C^0(I, \mathbb{R}^n)$ and $H^0(I, \mathbb{R}^n)$, respectively, and that the natural inclusions

$$H^1(I, \mathbb{R}^n) \hookrightarrow C^0(I, \mathbb{R}^n) \hookrightarrow H^0(I, \mathbb{R}^n) \tag{1}$$

are continuous.

This can be seen as a particular case of a more general result regarding spaces of sections of vector bundles over the unit interval I. Explicitly, we find...

Proposition 2.1. *Given an Euclidean bundle $E \longrightarrow I$ – i.e. a vector bundle endowed with a Riemannian metric – the space $C^0(E)$ of all continuous sections of E is the completion of $C'^{\infty}(E)$ under the norm given by*

$$\|\xi\|_\infty = \sup_t \|\xi_t\|$$

Proposition 2.2. *Given an Euclidean bundle $E \longrightarrow I$, the space $H^0(E)$ of all square integrable sections of E is the completion of $C'^{\infty}(E)$ under the inner product given by*

$$\langle \xi, \eta \rangle_0 = \int_0^1 \langle \xi_t, \eta_t \rangle \, \mathrm{d}t$$

Proposition 2.3. *Given an Euclidean bundle $E \longrightarrow I$, the space $H^1(E)$ of all class H^1 sections of E is the completion of the space $C'^{\infty}(E)$ of piece-wise smooth sections of E under the inner product given by*

$$\langle \xi, \eta \rangle_1 = \langle \xi, \eta \rangle_0 + \left\langle \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \xi, \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \eta \right\rangle_0$$

Proposition 2.4. *Given an Euclidean bundle $E \longrightarrow I$, the inclusions*

$$H^1(E) \hookrightarrow C^0(E) \hookrightarrow H^0(E)$$

are continuous. More precisely, $\|\xi\|_\infty \leqslant \sqrt{2}\,\|\xi\|_1$ and $\|\xi\|_0 \leqslant \|\xi\|_\infty$.

Proof. Given $\xi \in H^0(E)$ we have

$$\|\xi\|_0^2 = \int_0^1 \|\xi_t\|^2 \, \mathrm{d}t \leqslant \int_0^1 \|\xi\|_\infty^2 \, \mathrm{d}t = \|\xi\|_\infty^2$$

Now given $\xi \in H^1(E)$ fix $t_0, t_1 \in I$ with $\|\xi\|_\infty = \|\xi_{t_1}\|$ and $\|\xi_{t_0}\| \leqslant \|\xi\|_0$. If $t_0 < t_1$ then

$$\begin{aligned}
\|\xi\|_\infty^2 &= \|\xi_{t_0}\|^2 + \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t} \|\xi_t\|^2 \, \mathrm{d}t \\
&\leqslant \|\xi\|_0^2 + \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t} \|\xi_t\|^2 \, \mathrm{d}t \\
(\nabla \text{ is compatible with the metric}) &= \|\xi\|_0^2 + \int_{t_0}^{t_1} 2 \left\langle \xi_t, \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \xi_t \right\rangle \, \mathrm{d}t \\
(\text{Cauchy-Schwarz}) &\leqslant \|\xi\|_0^2 + \int_0^1 2 \|\xi_t\| \cdot \left\| \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \xi_t \right\| \, \mathrm{d}t \\
&\leqslant \|\xi\|_0^2 + \int_0^1 \|\xi_t\|^2 + \left\| \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \xi_t \right\|^2 \, \mathrm{d}t \\
&= \|\xi\|_0^2 + \|\xi\|_0^2 + \left\| \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} \xi \right\|_0^2 \\
&\leqslant 2 \|\xi\|_1^2
\end{aligned}$$

Similarly, if $t_0 > t_1$ then

$$\|\xi\|_\infty^2 = \|\xi_{t_0}\|^2 + \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t} \|\xi_t\|^2 \, \mathrm{d}t = \|\xi_{t_0}\|^2 + \int_{1-t_0}^{1-t_1} \frac{\mathrm{d}}{\mathrm{d}t} \|\xi_{1-t}\|^2 \, \mathrm{d}t \leqslant 2 \|\xi\|_1^2$$

■

Remark. Apply proposition 2.4 to the trivial bundle $I \times \mathbb{R}^n \longrightarrow I$ to get the continuity of the maps in (1).

We are particularly interested in the case of the pullback bundle $E = \gamma^*TM \longrightarrow I$, where $\gamma : I \longrightarrow M$ is a piece-wise smooth curve.

$$\begin{tikzcd}
\gamma^*TM \arrow[r] \arrow[d, "\pi"'] & TM \arrow[d, "\pi"] \\
I \arrow[r, "\gamma"'] & M
\end{tikzcd}$$

We now have all the necessary tools to describe the differential structure of $H^1(I, M)$.

2.1 The Charts of $H^1(I, M)$

We begin with a technical lemma.

Lemma 2.1. *Let $W \subseteq TM$ be an open neighborhood of the zero section in TM. Given $\gamma \in C'^\infty(I, M)$, denote by $W_{\gamma,t}$ the set $W \cap T_{\gamma(t)}M$ and let $W_\gamma = \bigcup_t W_{\gamma,t}$. Then $H^1(W_\gamma) = \{X \in H^1(\gamma^*TM) : X_t \in W_{\gamma,t}\ \forall t\}$ is an open subset of $H^1(\gamma^*TM)$.*

Proof. Let $C^0(W_\gamma) = \{X \in C^0(\gamma^*TM) : X_t \in W_{\gamma,t}\ \forall t\}$. We claim $C^0(W_\gamma)$ is open in $C^0(\gamma^*TM)$. Indeed, given $X \in C^0(W_\gamma)$ there exists $\delta > 0$ such that

$$\begin{aligned}
\|X - Y\|_\infty < \delta &\implies \|X_t - Y_t\| < \delta\ \forall t \\
&\implies Y_t \in W_{\gamma,t}\ \forall t \\
&\implies Y \in C^0(W_\gamma)
\end{aligned}$$

Finally, notice that $H^1(W_\gamma)$ is the inverse image of $C^0(W_\gamma)$ under the continuous inclusion $H^1(\gamma^*TM) \hookrightarrow C^0(\gamma^*TM)$ and is therefore open. ■

Let $W \subseteq TM$ be an open neighborhood of the zero section in TM such that $\exp\restriction_W : W \longrightarrow \exp(W)$ is invertible – whose existence follows from the fact that the injectivity radius depends continuously on $p \in M$.

Definition 2.2. Given $\gamma \in C'^{\infty}(I, M)$ let $W_\gamma, W_{\gamma,t} \subseteq \gamma^* TM$ be as in lemma 2.1, define

$$\begin{aligned} \exp_\gamma : H^1(W_\gamma) &\longrightarrow H^1(I, M) \\ X &\longmapsto \exp \circ X : I \longrightarrow M \\ &\qquad\qquad t \longmapsto \exp_{\gamma(t)}(X_t) \end{aligned}$$

and let $U_\gamma = \exp_\gamma(H^1(W_\gamma))$.

Finally, we find…

Theorem 2.1. *Given $\gamma \in C'^{\infty}(I, M)$, the map $\exp_\gamma : H^1(W_\gamma) \longrightarrow U_\gamma$ is bijective. The collection $\{(U_\gamma, \exp_\gamma^{-1} : U_\gamma \longrightarrow H^1(\gamma^* TM))\}_{\gamma \in C'^{\infty}(I,M)}$ is an atlas for $H^1(I, M)$ under the final topology of the maps $\exp_\gamma$ – i.e. the coarsest topology such that such maps are continuous. This atlas gives $H^1(I, M)$ the structure of a* separable *Banach manifold modeled after separable Hilbert spaces, with typical representatives*[4] $H^1(\gamma^* TM) \cong H^1(I, \mathbb{R}^n)$.

The fact that $\exp_\gamma$ is bijective should be clear from the definition of U_γ and W_γ. That each $\exp_\gamma^{-1}$ is a homeomorphism is also clear from the definition of the topology of $H^1(I, M)$. Moreover, since $C'^{\infty}(I, M)$ is dense, $\{U_\gamma\}_{\gamma \in C'^{\infty}(I,M)}$ is an open cover of $H^1(I, M)$. The real difficulty of this proof is showing that the transition maps

$$\exp_\eta^{-1} \circ \exp_\gamma : \exp_\gamma^{-1}(U_\gamma \cap U_\eta) \subseteq H^1(\gamma^* TM) \longrightarrow H^1(\eta^* TM)$$

are diffeomorphisms, as well as showing that $H^1(I, M)$ is separable. We leave these details as an exercise to the reader – see theorem 2.3.12 of [3] for a full proof.

It's interesting to note that this construction is functorial. More precisely…

Theorem 2.2. *Given finite-dimensional Riemannian manifolds M and N and a smooth map $f : M \longrightarrow N$, the map*

$$\begin{aligned} H^1(I, f) : H^1(I, M) &\longrightarrow H^1(I, N) \\ \gamma &\longmapsto f \circ \gamma \end{aligned}$$

is smooth. In addition, $H^1(I, f \circ g) = H^1(I, f) \circ H^1(I, g)$ and $H^1(I, \mathrm{id}) = \mathrm{id}$ for any composable smooth maps f and g. We thus have a functor $H^1(I, -) : \mathbf{Rnn} \longrightarrow \mathbf{BMnd}$ from the category **Rnn** *of finite-dimensional Riemannian manifolds and smooth maps onto the category* **BMnd** *of Banach manifolds and smooth maps.*

We would also like to point out that this is a particular case of a more general construction: that of the Banach manifold $H^1(E)$ of class H^1 sections of a smooth fiber bundle $E \longrightarrow I$ – not necessarily a vector bundle. Our construction of $H^1(I, M)$ is equivalent to that of the manifold $H^1(I \times M)$, in the sense that the canonical map

$$\begin{aligned} \tilde{\cdot} : H^1(I, M) &\longrightarrow H^1(I \times M) \\ \gamma &\longmapsto \tilde{\gamma} : I \longrightarrow I \times M \\ &\qquad\qquad t \longmapsto (t, \gamma(t)) \end{aligned}$$

can be easily checked to be a diffeomorphism.

The space $H^1(E)$ is modeled after the Hilbert spaces $H^1(F)$ of class H^1 sections of open subbundles $F \subseteq E$ which have the structure of a vector bundle – the so called *vector bundle neighborhoods of E*. This construction is highlighted in great detail and generality in the first section of [5, ch. 11], but unfortunately we cannot afford such a diversion in these short notes. Having said that, we are now finally ready to layout the Riemannian structure of $H^1(I, M)$.

[4] Any trivialization of $\gamma^* TM$ induces an isomorphism $H^1(\gamma^* TM) \xrightarrow{\sim} H^1(I \times \mathbb{R}^n) \cong H^1(I, \mathbb{R}^n)$.

2.2 The Metric of $H^1(I, M)$

We begin our discussion of the Riemannian structure of $H^1(I, M)$ by looking at its tangent bundle. Notice that for each $\gamma \in C'^{\infty}(I, M)$ the chart $\exp_\gamma^{-1} : U_\gamma \longrightarrow H^1(\gamma^* TM)$ induces a canonical isomorphism $\phi_\gamma = \phi_{\gamma,\gamma} : T_\gamma H^1(I, M) \xrightarrow{\sim} H^1(\gamma^* TM)$, as described in proposition 1.1. In fact, these isomorphisms may be extended to a canonical isomorphism of vector bundles, as seen in…

Lemma 2.2. *Given $i = 0, 1$, the collection $\{(\psi_{i,\gamma}(H^1(W_\gamma) \times H^i(\gamma^* TM)), \psi_{i,\gamma}^{-1})\}_{\gamma \in C'^{\infty}(I,M)}$ with*

$$\begin{aligned} \psi_{i,\gamma} : H^1(W_\gamma) \times H^i(\gamma^* TM) &\longrightarrow \textstyle\coprod_{\eta \in H^1(I,M)} H^i(\eta^* TM) \\ (X, Y) &\longmapsto \psi_{i,\gamma}(X) : I \longrightarrow \exp_\gamma(X)^* TM \\ &\qquad\qquad t \longmapsto (d\exp)_{X_t} Y_t \end{aligned}$$

gives $\coprod_{\gamma \in C'^{\infty}(I,M)} H^i(\gamma^* TM) \longrightarrow H^1(I, M)$ the structure of a smooth vector bundle[5].

Proposition 2.5. *There is a canonical isomorphism of vector bundles*

$$TH^1(I, M) \xrightarrow{\sim} \coprod_{\gamma \in H^1(I,M)} H^1(\gamma^* TM)$$

whose restriction $T_\gamma H^1(I, M) \xrightarrow{\sim} H^1(\gamma^* TM)$ is given by ϕ_γ for all $\gamma \in C'^{\infty}(I, M)$.

Proof. Note that the sets $H^1(W_\gamma) \times T_\gamma H^1(I, M)$ are precisely the images of the charts

$$\varphi_\gamma^{-1} : \varphi_\gamma(H^1(W_\gamma) \times T_\gamma H^1(I, M)) \subseteq TH^1(I, M) \longrightarrow H^1(W_\gamma) \times T_\gamma H^1(I, M)$$

of $TH^1(I, M)$ given by[6]

$$\begin{aligned} \varphi_\gamma : H^1(W_\gamma) \times T_\gamma H^1(I, M) &\longrightarrow TH^1(I, M) \\ (X, Y) &\longmapsto (d\exp_\gamma)_X \phi_\gamma(Y) \end{aligned}$$

By composing charts we get a fiber-preserving, fiber-wise linear diffeomorphism

$$\varphi_\gamma(H^1(W_\gamma) \times T_\gamma H^1(I, M)) \subseteq TH^1(I, M) \xrightarrow{\sim} \psi_{1,\gamma}(H^1(W_\gamma) \times H^1(\gamma^* TM)),$$

which takes $\varphi_\gamma(X, Y) \in T_{\exp_\gamma(X)} H^1(I, M)$ to $\psi_{1,\gamma}(X, \phi_\gamma(Y)) \in H^1(\exp_\gamma(X)^* TM)$. With enough patience, one can deduce from the fact that φ_γ^{-1} and $\psi_{1,\gamma}^{-1}$ are charts that these maps agree in the intersections of the open subsets $\varphi_\gamma(H^1(W_\gamma) \times T_\gamma H^1(I, M))$, so that they may be glued together into a global smooth map $\Phi : TH^1(I, M) \longrightarrow \coprod_{\eta \in H^1(I,M)} H^1(\eta^* TM)$.

Since this map is a fiber-preserving, fiber-wise linear local diffeomorphism, this is an isomorphism of vector bundles. Furthermore, by construction

$$\Phi(X)_t = \psi_{1,\gamma}(0, \phi_\gamma(X))_t = (d\exp)_{0_{\gamma(t)}} \phi_\gamma(X)_t = \phi_\gamma(X)_t$$

for each $\gamma \in C'^{\infty}(I, M)$ and $X \in T_\gamma H^1(I, M)$. In other words, $\Phi\restriction_{T_\gamma H^1(I,M)} = \phi_\gamma$ as required. ∎

At this point it may be tempting to think that we could now define the metric of $H^1(I, M)$ in a fiber-wise basis via the identification $T_\gamma H^1(I, M) \cong H^1(\gamma^* TM)$. In a very real sense this is what we are about to do, but unfortunately there are still technicalities in our way. The issue we face is that proposition 2.2 only applies for *smooth* vector bundles $E \longrightarrow I$, which may not be the case for $E = \gamma^* TM$ if $\gamma \in H^1(I, M)$ lies outside of $C'^{\infty}(I, M)$. In fact, neither $\langle X, Y \rangle_0$ nor $\langle\, , \rangle_1$ are defined *a priori* for $X, Y \in H^0(\gamma^* TM)$ with $\gamma \notin C'^{\infty}(I, M)$.

Nevertheless, we can get around this limitation by extending the metric $\langle\, , \rangle_0$ and the covariant derivative $\frac{\nabla}{\mathrm{dt}} = \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}}$ to those $H^0(\gamma^* TM)$ with $\gamma \notin C'^{\infty}(I, M)$. In other words, we'll show…

[5] Here we use the canonical identification $T_{\gamma(t)} M \cong T_{X_t} TM$ to apply the vector $Y_t \in T_{\gamma(t)} M$ to the map $(d\exp)_{X_t} : T_{X_t} TM \longrightarrow T_{\exp_{\gamma(t)}(X_t)} M$.

[6] Once more, we use the canonical identification $T_X H^1(W_\gamma) \cong H^1(\gamma^* TM)$ to apply the vector $\phi_\gamma(Y) \in H^1(\gamma^* TM)$ to $(d\exp_\gamma)_X : T_X H^1(W_\gamma) \longrightarrow T_{\exp_\gamma(X)} H^1(I, M)$.

Theorem 2.3. *The vector bundle $\coprod_{\gamma\in H^1(I,M)} H^0(\gamma^*TM) \longrightarrow H^1(I,M)$ admits a canonical Riemannian metric whose restriction to the fibers $H^0(\gamma^*TM) = \coprod_{\eta\in H^1(I,M)} H^0(\eta^*TM)\Big|_\gamma$ for $\gamma \in C'^\infty(I,M)$ is given by $\langle\,,\rangle_0$ as defined in proposition 2.2.*

Proof. Given $\gamma \in C'^\infty(I,M)$ and $X \in H^1(\gamma^*TM)$, let

$$\begin{aligned} g_X^\gamma : H^0(\gamma^*TM) \times H^0(\gamma^*TM) &\longrightarrow \mathbb{R} \\ (Y,Z) &\longmapsto \int_0^1 \langle (d\exp)_{X_t} Y_t, (d\exp)_{X_t} Z_t\rangle \,\mathrm{d}t \end{aligned}$$

This is clearly a Riemannian metric in the bundle $H^1(W_\gamma) \times H^0(\gamma^*TM) \longrightarrow H^1(W_\gamma)$. Now by composing with the chart $\psi_{0,\gamma}^{-1}$ as in lemma 2.2 we get a Riemannian metric g_γ in the bundle $\coprod_{\eta\in U_\gamma} H^0(\eta^*TM) = \coprod_{\eta\in H^1(I,M)} H^0(\eta^*TM)\Big|_{U_\gamma} \longrightarrow U_\gamma$. One can then quickly verify that the g^γ's agree in the intersection of the U_γ's, so that they define a global Riemannian metric g in $\coprod_{\eta\in H^1(I,M)} H^0(\eta^*TM)$.

Furthermore, given $\gamma \in C'^\infty(I,M)$ and $X, Y \in H^0(\gamma^*TM)$ by construction we have

$$g_\gamma(X,Y) = g_0^\gamma(X,Y) = \int_0^1 \langle (d\exp)_{0_{\gamma(t)}} X_t, (d\exp)_{0_{\gamma(t)}} Y_t\rangle \,\mathrm{d}t = \int_0^1 \langle X_t, Y_t\rangle \,\mathrm{d}t = \langle X, Y\rangle_0$$

■

Proposition 2.6. *The map*

$$\begin{aligned} \partial : H^1(I,M) &\longrightarrow \coprod_\gamma H^0(\gamma^*TM) \\ \gamma &\longmapsto \dot\gamma \end{aligned}$$

is a smooth section of $\coprod_{\gamma\in H^1(I,M)} H^0(\gamma^*TM) \longrightarrow H^1(I,M)$.

Proposition 2.7. *Denote by $\nabla^0 : \mathfrak{X}(H^1(I,M)) \times \Gamma\left(\coprod_\gamma H^0(\gamma^*TM)\right) \longrightarrow \Gamma\left(\coprod_\gamma H^0(\gamma^*TM)\right)$ the Levi-Civita connection of $\coprod_\gamma H^0(\gamma^*TM)$. The map*

$$\begin{aligned} \mathfrak{X}(H^1(I,M)) &\longrightarrow \Gamma\left(\coprod_\gamma H^0(\gamma^*TM)\right) \\ \tilde X &\longmapsto \nabla^0_{\tilde X}\partial \end{aligned}$$

is such that

$$(\nabla^0_X \partial)_\gamma = \nabla_{\frac{\mathrm{d}}{\mathrm{d}t}} X = \frac{\nabla}{\mathrm{d}t} X$$

for all $\gamma \in C'^\infty(I,M)$ and $X \in H^1(\gamma^*TM) \cong T_\gamma H^1(I,M)$. Given some arbitrary $\gamma \in H^1(I,M)$ and $X \in H^1(\gamma^*TM)$ we therefore denote $(\nabla^0_X\partial)_\gamma$ simply by $\frac{\nabla}{\mathrm{d}t}X$.

The proofs of these last two propositions were deemed too technical to be included in here, but see proposition 2.3.16 and 2.3.18 of [3]. We may now finally describe the canonical Riemannian metric of $H^1(I,M)$.

Definition 2.3. Given $\gamma \in H^1(I,M)$ and $X, Y \in H^1(\gamma^*TM)$, let

$$\langle X, Y\rangle_1 = \langle X, Y\rangle_0 + \left\langle \frac{\nabla}{\mathrm{d}t}X, \frac{\nabla}{\mathrm{d}t}Y \right\rangle_0$$

At this point it should be obvious that definition 2.3 does indeed endow $H^1(I, M)$ with the structure of a Riemannian manifold: the inner products $\langle\,,\rangle_1 : H^1(\gamma^*TM) \times H^1(\gamma^*TM) \longrightarrow \mathbb{R}$ may be glued together into a single positive-definite section $\langle\,,\rangle_1 \in \Gamma\left(\mathrm{Sym}^2 \coprod_\gamma H^1(\gamma^*TM)\right)$ – whose smoothness follows from theorem 2.3, proposition 2.6 and proposition 2.7 – which is then mapped to a positive-definite section of $\mathrm{Sym}^2 TH^1(I, M)$ by the induced isomorphism

$$\Gamma\left(\mathrm{Sym}^2 \coprod_\gamma H^1(\gamma^*TM)\right) \xrightarrow{\sim} \Gamma(\mathrm{Sym}^2 TH^1(I, M))$$

We are finally ready to discuss some applications.

3 Applications to the Calculus of Variations

As promised, in this section we will apply our understanding of the structure of $H^1(I, M)$ to the calculus of variations, and in particular to the geodesics problem. We also describe some further applications, such as the Morse index theorem and the Jacobi-Darboux theorem. We start by defining…

Definition 3.1. Given $\gamma \in H^1(I, M)$, a variation $\{\gamma_t\}_t$ of γ is a smooth curve $\gamma_\cdot : (-\epsilon, \epsilon) \longrightarrow H^1(I, M)$ with $\gamma_0 = \gamma$. We call the vector $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \gamma_t \in H^1(\gamma^*TM)$ *the variational vector field of* $\{\gamma_t\}_t$.

We should note that the previous definition encompasses the classical definition of a variation of a curve, as defined in [1, ch. 5] for instance: any piece-wise smooth function $H : I \times (-\epsilon, \epsilon) \longrightarrow M$ determines a variation $\{\gamma_t\}_t$ given by $\gamma_t(s) = H(s, t)$. This is representative of the theory that lies ahead, in the sense that most of the results we'll discuss in the following are minor refinements of the classical theory. Instead, the value of the theory we will develop in here lies in its conceptual simplicity: instead of relying in ad-hoc methods we can now use the standard tools of calculus to study the critical points of the energy functional E.

What we mean by this last statement is that by look at the energy functional as a smooth function $E \in C^\infty(H^1(I, M))$ we can study its classical "critical points" – i.e. curves γ with a variation $\{\gamma_t\}_t$ such that $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} E(\gamma_t) = 0$ – by looking at its derivative. The first variation of energy thus becomes a particular case of a formula for dE, and the second variation of energy becomes a particular case of a formula for the Hessian of E at a critical point. Without further ado, we prove…

Theorem 3.1. *The energy functional*

$$\begin{aligned} E : H^1(I, M) &\longrightarrow \mathbb{R} \\ \gamma &\longmapsto \frac{1}{2}\|\partial\gamma\|_0 = \frac{1}{2}\int_0^1 \|\dot{\gamma}(t)\|^2 \,\mathrm{d}t \end{aligned}$$

is smooth and $dE_\gamma X = \left\langle \partial\gamma, \frac{\nabla}{\mathrm{d}t} X\right\rangle_0$.

Proof. The fact that E is smooth should be clear from the smoothness of ∂ and $\|\cdot\|_0$. Furthermore, from the definition of $\frac{\nabla}{\mathrm{d}t}$ we have

$$\begin{aligned} \left\langle \partial\gamma, \frac{\nabla}{\mathrm{d}t} X\right\rangle_0 &= \langle \partial\gamma, (\nabla^0_X \partial)_\gamma\rangle_0 \\ &= (\tilde{X}\langle\partial,\partial\rangle_0)(\gamma) - \langle(\nabla^0_X\partial)_\gamma, \partial\gamma\rangle_0 \\ &= 2\tilde{X}E(\gamma) - \left\langle \frac{\nabla}{\mathrm{d}t} X, \partial\gamma\right\rangle_0 \\ &= 2dE_\gamma X - \left\langle \partial\gamma, \frac{\nabla}{\mathrm{d}t} X\right\rangle_0 \end{aligned}$$

where $\tilde{X} \in \mathfrak{X}(H^1(I, M))$ is any vector field with $\tilde{X}_\gamma = X$. ■

As promised, by applying the chain rule and using the compatibility of ∇ with the metric we arrive at the classical formula for the first variation of energy E.

Corollary 3.1. *Given a piece-wise smooth curve $\gamma : I \longrightarrow M$ with $\gamma \restriction_{[t_i,t_{i+1}]}$ smooth and a variation $\{\gamma_t\}_t$ of γ with variational vector field X we have*

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} E(\gamma_t) = \sum_i \langle \dot\gamma(t), X_t \rangle |_{t=t_i}^{t_{i+1}} - \int_0^1 \left\langle \frac{\nabla}{\mathrm{d}t}\dot\gamma(t), X_t \right\rangle \mathrm{d}t$$

Another interesting consequence of theorem 3.1 is…

Corollary 3.2. *The only critical points of E in $H^1(I, M)$ are the constant curves.*

Proof. Clearly every constant curve is a critical point. On the other hand, if $\gamma \in H^1(I, M)$ is such that $\left\langle \partial\gamma, \frac{\nabla}{\mathrm{d}t} X \right\rangle_0 = dE_\gamma X = 0$ for all $X \in H^1(\gamma^*TM)$ then $\partial\gamma = 0$ and therefore γ is constant. ■

Another way to put is to say that the problem of characterizing the critical points of E in $H^1(I, M)$ is not interesting at all. This shouldn't really come as a surprise, as most interesting results from the classical theory are concerned with particular classes of variations of a curves, such as variations with fixed endpoints or variations through loops. In the next section we introduce two submanifolds of $H^1(I, M)$, corresponding to the classes of variations previously described, and classify the critical points of the restrictions of E to such submanifolds.

3.1 The Critical Points of E

We begin with a technical lemma.

Lemma 3.1. *The maps $\sigma, \tau : H^1(I, M) \longrightarrow M$ with $\sigma(\gamma) = \gamma(0)$ and $\tau(\gamma) = \gamma(1)$ are submersions.*

Proof. To see that σ and τ are smooth it suffices to observe that their local representation in U_γ for $\gamma \in C'^\infty(I, M)$ is given by the maps

$$\begin{aligned} U \subseteq H^1(W_\gamma) &\longrightarrow T_{\gamma(0)}M \\ X &\longmapsto X_0 \end{aligned} \qquad\qquad \begin{aligned} U \subseteq H^1(W_\gamma) &\longrightarrow T_{\gamma(1)}M \\ X &\longmapsto X_1 \end{aligned}$$

which are indeed smooth functions. This local representation also shows that

$$\begin{aligned} d\sigma_\gamma : H^1(\gamma^*TM) &\longrightarrow T_{\gamma(0)}M \\ X &\longmapsto X_0 \end{aligned} \qquad\qquad \begin{aligned} d\tau_\gamma : H^1(\gamma^*TM) &\longrightarrow T_{\gamma(1)}M \\ X &\longmapsto X_1 \end{aligned}$$

are surjective maps for all $\gamma \in H^1(I, M)$. ■

We can now show…

Theorem 3.2. *The subspace $\Omega_{pq}M \subseteq H^1(I, M)$ of curves joining $p, q \in M$ is a submanifold whose tangent space $T_\gamma\Omega_{pq}M$ is the subspace of $H^1(\gamma^*TM)$ consisting of class H^1 vector fields X along γ with $X_0 = X_1 = 0$. Likewise, the space $\Lambda M \subseteq H^1(I, M)$ of free loops is a submanifold whose tangent at γ is given by all $X \in H^1(\gamma^*TM)$ with $X_0 = X_1$.*

Proof. To see that these are submanifolds, it suffices to note that $\Omega_{pq}M$ and ΛM are the inverse images of the closed submanifolds $\{(p, q)\}, \{(p, p) : p \in M\} \subseteq M \times M$ under the submersion $(\sigma, \tau) : H^1(I, M) \longrightarrow M \times M$.

The characterization of their tangent bundles should also be clear: any curve $(-\epsilon, \epsilon) \longrightarrow H^1(I, M)$ passing through $\gamma \in \Omega_{pq}M$ whose image is contained in $\Omega_{pq}M$ is a variation of γ with fixed endpoints, so its variational vector field X satisfies $X_0 = X_1 = 0$. Likewise, any variation of a loop $\gamma \in \Lambda M$ trough loops – i.e. a curve $(-\epsilon, \epsilon) \longrightarrow \Lambda M$ passing through γ – satisfies $X_0 = X_1$. ■

Finally, as promised we will provide a characterization of the critical points of $E \restriction_{\Omega_{pq}M}$ and $E \restriction_{\Lambda M}$.

Theorem 3.3. *The critical points of $E\restriction_{\Omega_{pq}M}$ are precisely the geodesics of M joining p and q. The critical points of $E\restriction_{\Lambda M}$ are the closed geodesics of M – including the constant maps.*

Proof. We start by supposing that γ is a geodesic. Since γ is smooth,

$$dE_\gamma X = \int_0^1 \left\langle \dot\gamma(t), \frac{\nabla}{\mathrm{d}t} X \right\rangle \,\mathrm{d}t = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\langle \dot\gamma(t), X_t\rangle - \left\langle \frac{\nabla}{\mathrm{d}t}\dot\gamma(t), X \right\rangle \,\mathrm{d}t = \langle \dot\gamma(1), X_1\rangle - \langle \dot\gamma(0), X_0\rangle$$

Now if $\gamma \in \Omega_{pq}M$ and $X \in T_\gamma\Omega_{pq}M$ then $dE_\gamma X = \langle \dot\gamma(1), 0\rangle - \langle \dot\gamma(0), 0\rangle = 0$. Likewise, if γ is a closed geodesic and $X \in T_\gamma \Lambda M$ we find $dE_\gamma X = 0$ since $\dot\gamma(0) = \dot\gamma(1)$ and $X_0 = X_1$. This establishes that the geodesics are indeed critical points of the restrictions of E.

Suppose $\gamma \in \Omega_{pq}M$ is a critical point and let $Y, Z \in H^1(\gamma^* TM)$ be such that

$$\frac{\nabla}{\mathrm{d}t} Y = \partial\gamma \qquad\qquad Y_0 = 0 \qquad\qquad \frac{\nabla}{\mathrm{d}t} Z = 0 \qquad\qquad Z_1 = Y_1$$

Let $X_t = Y_t - tZ_t$. Then $X_0 = X_1 = 0$ and $\frac{\nabla}{\mathrm{d}t} X = \partial\gamma - Z$. Furthermore,

$$\langle Z, \partial\gamma - Z\rangle_0 = \left\langle Z, \frac{\nabla}{\mathrm{d}t} X \right\rangle_0 = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\langle Z_t, X_t\rangle \,\mathrm{d}t = \langle Z_1, X_1\rangle - \langle Z_0, X_0\rangle = 0$$

and

$$\langle \partial\gamma, \partial\gamma - Z\rangle_0 = \left\langle \partial\gamma, \frac{\nabla}{\mathrm{d}t} X \right\rangle_0 = dE_\gamma X = 0,$$

which implies $\|\partial\gamma - Z\|_0^2 = 0$. In other words, $\partial\gamma = Z \in H^1(\gamma^* TM)$ and therefore $\frac{\nabla}{\mathrm{d}t}\dot\gamma(t) = \frac{\nabla}{\mathrm{d}t} Z = 0$ – i.e. γ is a geodesic.

Finally, if $\gamma \in \Lambda M$ with $\gamma(0) = \gamma(1) = p$ we may apply the argument above to conclude that γ is a geodesic joining p to $q = p$. To see that γ is a closed geodesic apply the same argument again for $\eta(t) = \gamma(1 + 1/2)$ to conclude that $\dot\gamma(0) = \dot\eta(1/2) = \dot\gamma(1)$. ■

We should point out that the first part of theorem 3.3 is a particular case of a result regarding critical points of the restriction of E to the submanifold $H^1_{N_0,N_1}(I, M) \subseteq H^1(I, M)$ of curves joining submanifolds $N_0, N_1 \subseteq M$: the critical points of $E\restriction_{H^1_{N_0,N_1}(I,M)}$ are the geodesics γ joining N_0 to N_1 with $\dot\gamma(0) \in T_{\gamma(0)}N_0^\perp$ and $\dot\gamma(1) \in T_{\gamma(1)}N_1^\perp$. The proof of this result is essentially the same as that of theorem 3.3, given that $T_\gamma H^1_{N_0,N_1}(I, M)$ is subspace of H^1 vector fields X along γ with $X_0 \in T_{\gamma(0)}N_0$ and $X_1 \in T_{\gamma(1)}N_1$.

3.2 Second Order Derivatives of E

Having establish a clear connection between geodesics and critical points of E, the only thing we're missing to complete our goal of providing a modern account of the classical theory is a refurnishing of the formula for second variation of energy. Intuitively speaking, the second variation of energy should be a particular case of a formula for the second derivative of E. The issue we face is, of course, that in general there is no such thing as "the second derivative" of a smooth function between manifolds.

Nevertheless, the metric of $H^1(I, M)$ allow us to discuss "the second derivative" of E in a meaningful sense by looking at the Hessian form, which we define in the following.

Definition 3.2. Given a – possibly infinite-dimensional – Riemannian manifold N and a smooth functional $f : N \longrightarrow \mathbb{R}$, we call the symmetric tensor

$$d^2 f(X, Y) = \nabla df(X, Y) = XYf - df\nabla_X Y$$

the Hessian of f.

We can now apply the classical formula for the second variation of energy to compute the Hessian of E at a critical point.

Theorem 3.4. *If γ is a critical point of $E\restriction_{\Omega_{pq}M}$ then*

$$(d^2E\restriction_{\Omega_{pq}M})_\gamma(X,Y) = \left\langle \frac{\nabla}{\mathrm{d}t}X, \frac{\nabla}{\mathrm{d}t}Y \right\rangle_0 - \langle R_\gamma X, Y\rangle_0, \tag{2}$$

where $R_\gamma : H^1(\gamma^*TM) \longrightarrow H^1(\gamma^*TM)$ is given by $(R_\gamma X)_t = R(X_t, \dot\gamma(t))\dot\gamma(t)$. Formula (2) also holds for critical points of E in ΛM.

Proof. Given the symmetry of d^2E, it suffices to take $X \in T_\gamma\Omega_{pq}M$ and show

$$d^2E_\gamma(X,X) = \left\| \frac{\nabla}{\mathrm{d}t}X \right\|_0^2 - \langle R_\gamma X, X\rangle_0$$

To that end, we fix a variation $\{\gamma_t\}_t$ of γ with fixed endpoints and variational field X and compute

$$\begin{aligned} d^2E_\gamma(X,X) &= \left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_{t=0} E(\gamma_t) \\ \text{(second variation of energy)} &= \int_0^1 \left\| \frac{\nabla}{\mathrm{d}t}X \right\|^2 - \langle R(X_t, \dot\gamma(t))\dot\gamma(t), X_t\rangle \,\mathrm{d}t \\ &= \left\| \frac{\nabla}{\mathrm{d}t}X \right\|_0^2 - \langle R_\gamma X, X\rangle_0 \end{aligned}$$

■

Next we discuss some further applications of the theory we've developed so far. In particular, we will work towards Morse's index theorem and and describe how one can apply it to establish the Jacobi-Darboux theorem. We begin with a technical lemma, whose proof amounts to an uninspiring exercise in analysis – see lemma 2.4.6 of [3].

Lemma 3.2. *Let $\Omega^0_{pq}M \subseteq C^0(I,M)$ be the space of continuous curves joining p to q. Then the inclusion $\Omega_{pq}M \hookrightarrow \Omega^0_{pq}M$ is continuous and compact. Likewise, if M is compact and $\Lambda^0M \subseteq C^0(I,M)$ is the space of continuous free loops then the inclusion $\Lambda M \hookrightarrow \Lambda^0 M$ is continuous and compact.*

As a first consequence, we prove…

Proposition 3.1. *Given a critical point γ of E in $\Omega_{pq}M$, the self-adjoint operator $A_\gamma : T_\gamma\Omega_{pq}M \longrightarrow T_\gamma\Omega_{pq}M$ given by*

$$\langle A_\gamma X, Y\rangle_1 = \langle X, A_\gamma Y\rangle_1 = d^2E_\gamma(X,Y)$$

has the form $A_\gamma = \mathrm{Id} + K_\gamma$ where $K_\gamma : T_\gamma\Omega_{pq}M \longrightarrow T_\gamma\Omega_{pq}M$ is a compact operator. The same holds for ΛM if M is compact.

Proof. Consider $K_\gamma = -\left(\mathrm{Id} - \frac{\nabla^2}{\mathrm{d}t^2}\right)^{-1} \circ (\mathrm{Id} + R_\gamma)$. We will show that K_γ is compact and that $A_\gamma = \mathrm{Id} + K_\gamma$ for γ in both $\Omega_{pq}M$ and ΛM – in which case assume M is compact.

Let $\gamma \in \Omega_{pq}M$ be a critical point. By theorem 3.3 we know that γ is a geodesic. Let $X, Y \in \Gamma(\gamma^*TM)$ with $X_0 = X_1 = Y_0 = Y_1 = 0$. Then

$$\begin{aligned} \langle X, Y\rangle_1 &= \langle X, Y\rangle_0 + \left\langle \frac{\nabla}{\mathrm{d}t}X, \frac{\nabla}{\mathrm{d}t}Y \right\rangle \\ &= \langle X, Y\rangle_0 + \int_0^1 \left\langle \frac{\nabla}{\mathrm{d}t}X, \frac{\nabla}{\mathrm{d}t}Y \right\rangle \,\mathrm{d}t \\ &= \langle X, Y\rangle_0 + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\langle X_t, Y_t\rangle - \left\langle \frac{\nabla^2}{\mathrm{d}t^2}X, Y \right\rangle \,\mathrm{d}t \\ &= \langle X, Y\rangle_0 - \left\langle \frac{\nabla^2}{\mathrm{d}t^2}X, Y \right\rangle_0 + \langle X_t, Y_t\rangle|_{t=0}^1 \\ &= \left\langle \left(\mathrm{Id} - \frac{\nabla^2}{\mathrm{d}t^2}\right)X, Y \right\rangle_0 \end{aligned} \tag{3}$$

Since $\Gamma(\gamma^*TM) \subseteq H^1(\gamma^*TM)$ is dense, (3) extends to all of $T_\gamma \Omega_{pq} M$. Hence given $X, Y \in T_\gamma \Omega_{pq} M$ we have

$$\begin{aligned}
\langle A_\gamma X, Y \rangle_1 &= \left\langle \frac{\nabla}{\mathrm{d}t} X, \frac{\nabla}{\mathrm{d}t} Y \right\rangle_0 - \langle R_\gamma X, Y \rangle_0 \\
&= \langle X, Y \rangle_1 - \langle X, Y \rangle_0 - \langle R_\gamma X, Y \rangle_0 \\
&= \langle X, Y \rangle_1 - \langle (\mathrm{Id} + R_\gamma) X, Y \rangle_0 \\
&= \langle X, Y \rangle_1 - \left\langle \left(\mathrm{Id} - \frac{\nabla^2}{\mathrm{d}t^2} \right)^{-1} \circ (\mathrm{Id} + R_\gamma) X, Y \right\rangle_1 \\
&= \langle X, Y \rangle_1 + \langle K_\gamma X, Y \rangle_1
\end{aligned}$$

Now consider a critical point $\gamma \in \Lambda M$ – i.e. a closed geodesic. Equation (3) also holds for $X, Y \in \Gamma(\gamma^*TM)$ with $X_0 = X_1$ and $Y_0 = Y_1$, so it holds for all $X, Y \in T_\gamma \Lambda M$. Hence by applying the same argument we get $\langle A_\gamma X, Y \rangle_1 = \langle (\mathrm{Id} + K_\gamma) X, Y \rangle_1$.

As for the compactness of K_γ in the case of $\Omega_{pq} M$, from (3) we get $\|K_\gamma X\|_1^2 = -\langle (\mathrm{Id} + R_\gamma) X, K_\gamma X \rangle_0$, so that proposition 2.4 implies

$$\|K_\gamma X\|_1^2 \leqslant \|\mathrm{Id} + R_\gamma\| \cdot \|K_\gamma X\|_\infty \cdot \|X\|_0 \leqslant \sqrt{2}\, \|\mathrm{Id} + R_\gamma\| \cdot \|K_\gamma X\|_1 \cdot \|X\|_0 \tag{4}$$

Given a bounded sequence $(X_n)_n \subseteq T_\gamma \Omega_{pq} M$, it follow from lemma 3.2 that $(X_n)_n$ is relatively compact as a C^0-sequence. From (4) we then get that $(K_\gamma X_n)_n$ is relatively compact as an H^1-sequence, as desired. The same argument holds for ΛM if M is compact – so that we can once more apply lemma 3.2. ■

Once again, the first part of this proposition is a particular case of a broader result regarding the space of curves joining submanifolds of M: if $N \subseteq M$ is a totally geodesic manifold of codimension 1 and $\gamma \in H^1_{N,\{q\}}(I, M)$ is a critical point of the restriction of E then $A_\gamma = \mathrm{Id} + K_\gamma$. These results aren't that appealing on their own, but they allow us to establish the following result, which is essential for stating Morse's index theorem.

Corollary 3.3. *Given a critical point γ of $E\restriction_{\Omega_{pq}M}$, there is an orthogonal decomposition*

$$T_\gamma \Omega_{pq} M = T_\gamma^- \Omega_{pq} M \oplus T_\gamma^0 \Omega_{pq} M \oplus T_\gamma^+ \Omega_{pq} M,$$

where $T_\gamma^- \Omega_{pq} M$ is the finite-dimensional subspace spanned by eigenvectors with negative eigenvalues, $T_\gamma^0 \Omega_{pq} M = \ker A_\gamma$ and $T_\gamma^+ \Omega_{pq} M$ is the proper Hilbert subspace given by the closure of the subspace spanned by eigenvectors with positive eigenvalues. The same holds for critical points γ of $E\restriction_{\Lambda M}$ and $T_\gamma \Lambda M$ if M is compact.

Definition 3.3. Given a critical point γ of $E\restriction_{\Omega_{pq}M}$ we call the number $\dim T_\gamma^- \Omega_{pq} M$ *the Ω-index of γ*. Likewise, we call $\dim T_\gamma^- \Lambda M$ for a critical point γ of $E\restriction_{\Lambda M}$ *the Λ-index of γ*. Whenever the submanifold γ lies in is clear from context we refer to the Ω-index or the Λ-index of γ simply by *the index of γ*.

This definition highlights one of the greatest strengths of our approach: while the index of a geodesic γ can be defined without the aid of the tools developed in here, by using of the Hessian form $d^2 E_\gamma$ we can place definition 3.3 in the broader context of Morse theory. In fact, the geodesics problems and the energy functional where among Morse's original proposed applications. Proposition 3.1 and definition 3.3 amount to a proof that the Morse index of E at a critical point γ is finite.

We are now ready to state Morse's index theorem.

Theorem 3.5 (Morse)**.** *Let $\gamma \in \Omega_{pq} M$ be a critical point of E. Then the index of γ is given of the sum of the multiplicities of the proper conjugate points*[7] *of γ in the interior of I.*

[7] By "conjugate points of a geodesic γ" we of course mean points conjugate to $\gamma(0) = p$ along γ.

Unfortunately we do not have the space to include the proof of Morse's theorem in here, but see theorem 2.5.9 of [3]. The index theorem can be generalized for $H^1_{N,\{q\}}(I,M)$ by replacing the notion of conjugate point with the notion of focal points of N – see theorem 7.5.4 of [1] for the classical approach. What we are really interested in, however, is the following consequence of Morse's theorem.

Theorem 3.6 (Jacobi-Darboux)**.** *Let $\gamma \in \Omega_{pq}M$ be a critical point of E.*

- **(i)** *If there are no conjugate points of γ then there exists a neighborhood $U \subseteq \Omega_{pq}M$ of γ such that $E(\eta) > E(\gamma)$ for all $\eta \in U$ with $\eta \neq \gamma$.*
- **(ii)** *Let $k > 0$ be the sum of the multiplicities of the conjugate points of γ in the interior of I. Then there exists an immersion*
$$i : B^k \longrightarrow \Omega_{pq}M$$
of the unit ball $B^k = \{v \in \mathbb{R}^k : \|v\| < 1\}$ with $i(0) = \gamma$, $E(i(v)) < E(\gamma)$ and $L(i(v)) < L(\gamma)$ for all nonzero $v \in B^k$.

Proof. First of all notice that given $\eta \in U_\gamma$ with $\eta = \exp_\gamma(X)$, $X \in H^1(W_\gamma)$, the Taylor series for $E(\eta)$ is given by $E(\eta) = E(\gamma) + \frac{1}{2}d^2E_\gamma(X,X) + \cdots$. More precisely,

$$\frac{\left|E(\exp_\gamma(X)) - E(\gamma) - \frac{1}{2}d^2E_\gamma(X,X)\right|}{\|X\|_1^2} \longrightarrow 0 \tag{5}$$

as $X \longrightarrow 0$.

Let γ be as in **(i)**. Since γ has no conjugate points, it follows from Morse's index theorem that $T^-_\gamma\Omega_{pq}M = 0$. Furthermore, by noticing that any piece-wise smooth $X \in \ker A_\gamma$ is Jacobi field vanishing at p and q one can also show $T^0_\gamma\Omega_{pq}M = 0$. Hence $T_\gamma\Omega_{pq}M = T^+_\gamma\Omega_{pq}M$ and therefore $d^2E\restriction_{\Omega_{pq}M}$ is positive-definite. Now given $\eta = \exp_\gamma(X)$ as before, (5) implies that $E(\eta) > E(\gamma)$, provided $\|X\|_1$ is sufficiently small.

As for part **(ii)**, fix an orthonormal basis $\{X_j : 1 \leqslant j \leqslant k\}$ of $T^-_\gamma\Omega_{pq}M$ consisting of eigenvectors of A_γ with negative eigenvalues $-\lambda_i$. Let $\delta > 0$ and define

$$\begin{aligned} i : B^k &\longrightarrow \Omega_{pq}M \\ v &\longmapsto \exp_\gamma(\delta(v_1 \cdot X_1 + \cdots + v_k \cdot X_k)) \end{aligned}$$

Clearly i is an immersion for small enough δ. Moreover, from (5) and

$$E(i(v)) = E(\gamma) - \frac{1}{2}\delta^2\sum_j \lambda_j \cdot v_j + \cdots$$

we find that $E(i(v)) < E(\gamma)$ for sufficiently small δ. In particular, $L(i(v))^2 \leqslant E(i(v)) < E(\gamma) = L(\gamma)^2$. ∎

We should point out that part **(i)** of theorem 3.6 is weaker than the classical formulation of the Jacobi-Darboux theorem – such as in theorem 5.5.3 of [1] for example – in two aspects. First, we do not compare the length of curves γ and $\eta \in U$. This could be amended by showing that the length functional $L : H^1(I,M) \longrightarrow \mathbb{R}$ is smooth and that its Hessian d^2L_γ is given by $C \cdot d^2E_\gamma$ for some $C > 0$. Secondly, unlike the classical formulation we only consider curves in an H^1-neighborhood of γ – instead of a neighborhood of γ in $\Omega_{pq}M$ in the uniform topology. On the other hand, part **(ii)** is definitively an improvement of the classical formulation: we can find curves $\eta = i(v)$ shorter than γ already in an H^1-neighborhood of γ.

This concludes our discussion of the applications of our theory to the geodesics problem. We hope that these short notes could provide the reader with a glimpse of the rich theory of the calculus of variations and global analysis at large. We once again refer the reader to [3, ch. 2], [5, ch. 11] and [2, sec. 6] for further insight on modern variational methods.

References

[1] Claudio Gorodski. *An introduction to Riemannian geometry*. Preliminary version 3. June 2022. URL: https://www.ime.usp.br/~gorodski/teaching/mat5771-2022/master07-05-2022.pdf.

[2] Jr. James Eells. “A setting for global analysis”. In: *Bulletin of the American Mathematical Society* 72.5 (1966), pp. 751–807.

[3] Wilhelm Klingenberg. *Riemannian Geometry*. De Gruyter Studies in Mathematics. De Gruyter, 2011. ISBN: 9783110905120.

[4] Serge Lang. *Fundamentals of Differential Geometry*. 1st ed. Graduate Texts in Mathematics. Springer, 1999. ISBN: 9780387985930.

[5] Chuu-lian Terng Richard Palais. *Critical Point Theory and Submanifold Geometry*. Lecture Notes in Mathematics. Springer, 1988. ISBN: 3540503994.

[6] Martin Schottenloher. *The Unitary Group In Its Strong Topology*. 2013. DOI: 10.48550/ARXIV.1309.5891.