CRITICAL POINT THEORY OF THE ENERGY FUNCTIONAL ON PATH SPACES

WERNER BALLMANN

Contents

1. Introduction

Standard critical point theory is concerned with smooth functions on finite dimensional manifolds. Its origins reach at least as far back as [Mö][1], where Möbius used it to classify compact surfaces in Euclidean space. Later, Birkhoff, Morse, Lusternik, and Schnirelmann were interested in the existence of periodic geodesics on Riemannian manifolds and studied this problem via critical point theory on loop spaces, compare with [B2], [Mo], and [LS].

Recall that, for a Riemannian manifold M and a piecewise smooth curve $c\colon [a,b] \to M$, *length* and *energy* of c are defined by

$$L(c) = \int_a^b \|c'\| \quad \text{and} \quad E(c) = \frac{1}{2}\int_a^b \|c'\|^2. \tag{1}$$

Length is more geometric than energy. However, since the length of a curve does not change under reparametrization, the analysis of the

---

Date: January 27, 2015.

1991 *Mathematics Subject Classification.* 49–01, 53C23, 58E10.

Key words and phrases. Curves, energy, boundary conditions, critical points, index, nullity.

Supported by Max Planck Institute for Mathematics (MPIM) and Hausdorff Center for Mathematics (HCM).

[1] I am grateful to Peter Dombrowski for this reference.

length as a functional on spaces of curves is somewhat unpleasant. The energy functional has better analytic properties, and that is the reason why we study the energy functional.

To set the stage, let $\Omega = \Omega([a,b], M)$ be the space of piecewise smooth curves $c\colon [a,b] \to M$, endowed with the compact-open topology, and consider E as a functional on Ω. Although Ω is neither a manifold nor complete and E is not even continuous, critical point theory of E on Ω and subspaces of Ω can be developed rigorously, via so-called finite dimensional approximations, as standard critical point theory of smooth functions on finite dimensional manifolds, and that is the topic of these notes.

There are a number of good sources on critical point theory of length and energy on path spaces, and I have borrowed freely from them; compare for example with [Bo], [Mi], [GKM], [CE], and the monographs mentioned further up. There is also a less elementary approach via infinite dimensional manifolds of path spaces, see [Pa] and [Kl]. However, for most purposes, the elementary approach pursued in these notes seems quite adequate.

I would like to thank Saskia Voß for her careful reading of the notes.

2. Prerequisites

We assume that the reader is familiar with elementary Riemannian geometry. For later reference, we recall a few basic concepts and results.

2.1. **Length and energy.** Let $c\colon [a,b] \to M$ be a piecewise smooth curve. Then

$$L(c|[a,b])^2 = \Big(\int_a^b \|c'\|\Big)^2 \le |b-a| \cdot \int \|c'\|^2 = 2|b-a|E(c), \qquad (2)$$

by the Cauchy-Schwarz inequality. Equality holds if and only if c has constant speed, $\|c'(t)\| = \text{const}$.

2.2. **Jacobi fields and conjugate points.** Let $c\colon [a,b] \to M$ be a geodesic. Then a vector field V along c is called a *Jacobi field* if it satisfies the *Jacobi equation*

$$V'' + R(V,c')c' = 0.$$

Jacobi fields are exactly the variation fields of geodesic variations of c, that is, of variations (c_s) of $c = c_0$ such that all the c_s are geodesics, compare with Definition 3.1. For $t \in (a,b]$, we say that $c(t)$ is a *conjugate point of* $c(a)$ *along* c if there is a nonzero Jacobi field J along c with $J(a) = J(t) = 0$. The *multiplicity of* $c(t)$ *as a conjugate point of* $c(a)$ *along* c is equal to the dimension of the space of such Jacobi fields (where the multiplicity of $c(t)$ is equal to 0 if it is not a conjugate point of $c(a)$ along c).

It follows easily from the characterization of Jacobi fields as variation fields of geodesic variations and the definition of the exponential map exp that conjugate points correspond to critical points of the latter:

2.1. **Proposition.** *Let $p \in M$ and $v \in T_pM$ be in the domain of definition of* $\exp_p$*. Then* $\dim\ker\exp_{p*v}$ *is equal to the multiplicity of* $\exp(v)$ *as a conjugate point of p along the geodesic* $\exp(tv)$*,* $0 \le t \le 1$*.*

2.3. **Injectivity radius.** The following fact follows immediately from the continuity of the injectivity radius function of M:

2.2. **Theorem.** *Let $U \subseteq M$ be a relatively compact open subset. Then there is a constant $\varepsilon > 0$ such that any two points $p, q \in U$ with distance $d(p,q) < \varepsilon$ can be joined by a unique geodesic $c_{pq}\colon [0,1] \to M$ of length $< \varepsilon$ and $c_{pq}(t)$ depends smoothly on p, q, and t.*

3. First and second variation of energy

3.1. **Definition.** An *(ℓ-parameter) variation* of a curve $c \in \Omega$ is a map $F\colon U \times [a,b] \to M$, where U is an open neighborhood of $0 \in \mathbb{R}^\ell$, such that $c_0 = c$, where $c_s := F(s,.)$, for all $s \in U$, and such that there is a subdivision $t_0 < \cdots < t_k$ of $[a,b]$ (where $t_0 = a$ and $t_k = b$) such that F restricted to $U \times [t_{i-1}, t_i]$ is smooth, for all $1 \le i \le k$. A variation F of c is called *proper* if $c_s(a) = c(a)$ and $c_s(b) = c(b)$, for all $s \in U$. For a variation F of c, the vector fields $(\partial_j F)(0,.)$ along c are called the *variation fields* associated to F.

It follows that the map $U \to \Omega$, $s \mapsto c_s$, associated to a variation F is continuous with respect to the compact-open topology on Ω. We think of variations as smooth maps into Ω, although Ω is not a manifold. Nevertheless, for any variation F of c, the composition $E(s) := E(c_s)$ is a smooth map of $s \in U$. The derivative of this map at $s = 0$ is called the *first variation of energy at c with respect to F.*

3.2. **First variation of energy.** *Let F be a 1-parameter variation of $c \in \Omega$ with associated subdivision $t_0 < \cdots < t_k$ of $[a,b]$ as in Definition 3.1. Then the first variation of energy at c with respect to F is given by*

$$E'(0) = -\int_a^b \langle c'', V\rangle\, dt + \langle c'(t), V(t)\rangle\Big|_a^b - \sum_{i=1}^{k-1} \langle \Delta c'(t_i), V(t_i)\rangle, \qquad (3)$$

where $V = (\partial_s F)(0,.)$ denotes the variation field associated to F and where $\Delta c'(t) := c'(t+) - c'(t-)$.

Proof. Using that the Levi-Civita connection is symmetric and metric, we get

$$\begin{aligned} E' &= \partial_s\Big(\frac{1}{2}\int_a^b \langle \partial_t F, \partial_t F\rangle\, dt\Big) = \frac{1}{2}\int_a^b \partial_s \langle \partial_t F, \partial_t F\rangle\, dt \\ &= \int_a^b \langle \partial_t F, D_s \partial_t F\rangle\, dt = \int_a^b \langle \partial_t F, D_t \partial_s F\rangle\, dt \\ &= \int_a^b \partial_t \langle \partial_t F, \partial_s F\rangle\, dt - \int_a^b \langle D_t \partial_t F, \partial_s F\rangle\, dt. \end{aligned}$$

The first term on the right is equal to

$$\sum_{i=1}^{k} \langle (\partial_t F)(.,t), (\partial_s F)(.,t)\rangle\Big|_{t=t_{i-1}}^{t=t_i}.$$

Now $(\partial_s F)(0,.) = V$, $(\partial_t F)(0,.) = c'$, and $(D_t\partial_t F)(0,.) = c''$, hence

$$E'(0) = -\int_a^b \langle c'', V\rangle \, dt + \sum_{i=1}^k \langle c'(t), V(t)\rangle\Big|_{t=t_{i-1}}^{t=t_i}. \qquad \square$$

We denote by $\mathcal{V}(c)$ the vector space of piecewise smooth vector fields along c. We think of $\mathcal{V}(c)$ as the tangent space of Ω at c. This point of view is justified by the following

3.3. **Technical lemma.** *Let $c \in \Omega$ and $V_1, \dots, V_\ell \in \mathcal{V}(c)$. Choose times $a = t_0 < \cdots < t_k = b$, and let $f_i \colon U \to M$ be smooth maps, where U is an open neighborhood of 0 in $\mathbb{R}^\ell$, such that*

$$f_i(0) = c(t_i) \quad \textit{and} \quad (\partial_j f_i)(0) = V_j(t_i),$$

for all $0 \le i \le k$ and $1 \le j \le \ell$. Then, by diminishing the size of U if necessary, there is an ℓ-parameter variation $F \colon U \times [a,b] \to M$ of c with variation fields $(\partial_j F)(0,.) = V_j$ such that $F(.,t_i) = f_i$, for all $0 \le i \le k$ and $1 \le j \le \ell$.

Proof. By adding further times t_i and maps f_i if necessary, we may assume that the restrictions $c|[t_{i-1}, t_i]$ and $V_j|[t_{i-1}, t_i]$ are smooth and that the segments $c([t_{i-1}, t_i])$ are contained in coordinate domains U_i of M. By diminishing the size of U if necessary, the images of f_{i-1} and f_i are contained in the domains U_i about $c([t_{i-1}, t_i])$. Choosing coordinates of M over U_i, we arrive at a simple extension problem for smooth maps $U \times [t_{i-1}, t_i] \to \mathbb{R}^m$. $\square$

Denote be $\mathcal{V}_0(c)$ the space of $V \in \mathcal{V}(c)$ with $V(a) = V(b) = 0$.

3.4. **Corollary.** *A vector field $V \in \mathcal{V}(c)$ is the variation field of a proper variation of c if and only if $V \in \mathcal{V}_0(c)$.*

3.5. **Regularity theorem.** *Let $c \in \Omega$. Then the first variation of energy at c vanishes, with respect to any proper variation of c, if and only if c is a geodesic.*

Proof. By (3), if c is a geodesic, then the first variation of energy at c vanishes, with respect to any proper variation of c.

Assume now that the first variation of energy at c vanishes, with respect to any proper variation of c. Choose a subdivision $t_0 < \cdots < t_k$ of $[a,b]$ such that the restrictions $c|[t_{i-1}, t_i]$ are smooth. Suppose first that $c''(t) \ne 0$, for some t in some (t_{i-1}, t_i). Let V be a piecewise smooth vector field along c with support in a neighborhood of t in (t_{i-1}, t_i) such that $\langle c'', V\rangle \ge 0$ and $\langle c''(t), V(t)\rangle > 0$. By Lemma 3.3, there is a proper variation F of c with variation field V. Then $E'(0) = 0$ with respect to F, by assumption. Now the second and third term on the right in

(3) vanish since V vanishes at the corresponding points. On the other hand, the contribution of the first term is nonzero, by the choice of V. This is a contradiction. It follows that $c''(t) = 0$ for all t in all (t_{i-1}, t_i). Hence c is a geodesic on each of the intervals $[t_{i-1}, t_i]$, by the continuity of c'' on these intervals.

It remains to show that $\Delta c'(t_i) = 0$, for all $1 \le i \le k-1$. To that end, suppose that there is such an i with $c'(t_i-) \neq c'(t_i+)$. Let V be a piecewise smooth vector field along c with support in (t_{i-1}, t_{i+1}) such that $\langle \Delta c'(t_i), V(t_i)\rangle \neq 0$. By Lemma 3.3, there is a proper variation F of c with variation field V. Then again $E'(0) = 0$ with respect to F, by assumption. Now the first and second term on the right in (3) vanish since $c'' = 0$ and $V(a) = V(b) = 0$. In the sum on the right, only the term with the given i survives and gives a nonzero contribution, by the choice of V. This is again a contradiction, hence c is a geodesic. $\square$

3.6. **Second variation of energy.** *Let $c \in \Omega$ be a geodesic and F be a 2-parameter variation of c. Then the second variation of energy at c with respect to F is given by*

$$\begin{aligned}(\partial_r\partial_s E)(0,0) = \int_a^b \{\langle V', W'\rangle - \langle R(V,c')c', W\rangle\}\,dt \\ + \langle c'(t), (D_r\partial_s F)(0,t)\rangle\Big|_a^b,\end{aligned} \tag{4}$$

where $V := (\partial_r F)(0,.)$ and $W := (\partial_s F)(0,.)$.

Proof. We have

$$\begin{aligned}(\partial_r\partial_s E) &= \int_a^b \partial_r\langle \partial_t F, D_s\partial_t F\rangle\,dt \\ &= \int_a^b \big\{\langle D_r\partial_t F, D_s\partial_t F\rangle\,dt + \langle \partial_t F, D_rD_s\partial_t F\rangle\big\}\,dt \\ &= \int_a^b \big\{\langle D_t\partial_r F, D_t\partial_s F\rangle\,dt + \langle \partial_t F, D_rD_t\partial_s F\rangle\big\}\,dt.\end{aligned}$$

At $r = s = 0$, the first term under the integral on the right hand side is equal to $\langle V', W'\rangle$. As for the second term, we have

$$D_rD_t\partial_s F = D_tD_r\partial_s F + R(\partial_r F, \partial_t F)\partial_s F.$$

On the other hand,

$$\langle \partial_t F, D_tD_r\partial_s F\rangle = \partial_t\langle \partial_t F, D_r\partial_s F\rangle - \langle D_t\partial_t F, D_r\partial_s F\rangle.$$

Now at $r = s = 0$, we have $\partial_t F = c'$ and $D_t \partial_t F = c'' = 0$. Collecting terms, we arrive at the asserted formula. □

For proper variations of a geodesic $c\colon [a,b] \to M$, the last term on the right of (4) vanishes. As a bilinear and symmetric form in $V, W \in \mathcal{V}(c)$, the remaining part is called the *index form* of c,

$$I(V,W) := \int_a^b \{\langle V', W'\rangle - \langle R(V,c')c', W\rangle\}\, dt. \tag{5}$$

3.7. **Proposition.** *We have*

$$\begin{aligned} I(V,W) = &-\int_a^b \{\langle V'' + R(V,c')c', W\rangle\}\, dt \\ &+ \langle V'(t), W(t)\rangle \Big|_a^b - \sum_{i=1}^{k-1} \langle \Delta V'(t_i), W(t_i)\rangle, \end{aligned} \tag{6}$$

where the subdivision $t_0 < \cdots < t_k$ of $[a,b]$ is chosen such that the restrictions $V|[t_{i-1},t_i]$ and $W|[t_{i-1},t_i]$ are smooth and where

$$\Delta V'(t) := V'(t+) - V'(t-). \tag{7}$$

Proof. This follows easily from $\langle V', W'\rangle = \langle V', W\rangle' - \langle V'', W\rangle$. □

As a first application of the second variation formula, we show that geodesics are not minimizing if I has negative directions in $\mathcal{V}_0(c)$.

3.8. **Proposition.** *Let p and q be points in M and $c \in \Omega$ be a geodesic from p to q. Let $V \in \mathcal{V}_0(c)$ be a vector field along c with $I(V,V) < 0$ and F be a proper variation of c with variation field V. Then*

$$E(c_s) < E(c) \quad \textit{and} \quad L(c_s) < L(c),$$

for all sufficiently small nonzero s. In particular, c is not a shortest curve from p to q.

Proof. Let $F\colon (-\varepsilon, \varepsilon) \to M$ be a proper variation of c with variation field V. Then $E'(0) = 0$ and $E''(0) = I(V,V) < 0$, by Theorem 3.5 and (4), hence $E(c_s) < E(c)$, for all sufficiently small nonzero s. By (2), we also have

$$L^2(c_s) \le 2|b-a|E(c_s) \quad \text{and} \quad L^2(c) = 2|b-a|E(c). \qquad □$$

4. Boundary conditions and critical points

A *boundary conditon* on Ω is a closed submanifold $B \subseteq M \times M$. Given such a submanifold, we say that a curve $c \in \Omega$ is a *B-curve* if its end points $(c(a), c(b)) \in B$ and set

$$\Omega_B := \{c \in \Omega \mid c \text{ is a } B\text{-curve}\}. \tag{8}$$

We say that an ℓ-parameter variation F of $c \in \Omega_B$ is an *(ℓ-parameter) B-variation* if c_s is a B-curve, for all $s \in U$.

4.1. **Examples.** 0) $B = M \times M$ is called the *empty boundary condition.*
1) Given any pair of points $p, q \in M$, we may consider $B = \{(p,q)\}$. Then $\Omega_{pq} := \Omega_B$ is the space of piecewise smooth curves $c\colon [a,b] \to M$ from p to q. In the case $p = q$, we get the space $\Lambda_p := \Omega_{pp}$ of piecewise smooth loops $c\colon [a,b] \to M$ at p.
2) More generally, for closed submanifolds $P, Q \subseteq M$, we obtain the space Ω_{PQ} of piecewise smooth curves $c\colon [a,b] \to M$ from P to Q.
3) For $B = \{(p,p) \mid p \in M\}$, the diagonal in $M \times M$, we obtain the *free loop space of M*, that is, the space Λ of piecewise smooth closed curves $c\colon [a,b] \to M$.

A piecewise smooth vector field V along $c \in \Omega_B$ is called a *B-vector field* if it is the variation field of a B-variation. The vector space of B-vector fields is denoted by $\mathcal{V}_B(c)$.

4.2. **Lemma.** *Let $V_1, \dots, V_\ell$ be piecewise smooth vector fields along a curve $c \in \Omega_B$. Then there is an ℓ-parameter B-variation F of c with variation fields $(\partial_j F)(0,.) = V_j$ if and only if $(V_j(a), V_j(b))$ is tangent to B at $(c(a), c(b))$, for all $1 \le j \le \ell$. In particular,*

$$\mathcal{V}_B(c) = \{V \in \mathcal{V}(c) \mid (V(a), V(b)) \in T_{(c(a),c(b))}B\}.$$

Proof. Clearly, if F is a B-variation of c with associated variation fields $V_j = (\partial_j F)(0,.)$, then $(V_j(a), V_j(b))$ is tangent to B at $(c(a), c(b))$, for all $1 \le j \le \ell$. Conversely, suppose that $V_1, \dots, V_\ell$ are piecewise smooth vector fields along c such that $(V_j(a), V_j(b))$ is tangent to B at $(c(a), c(b))$, for all $1 \le j \le \ell$. Choose a smooth map $f\colon U \to B$, where U is an open neighborhood of 0 in $\mathbb{R}^\ell$, such that $f(0) = (c(a), c(b))$ and $(\partial_j f)(0) = (V_j(a), V_j(b))$, and denote first and second component of f by f_0 and f_1. Let F be a variation of c associated to the vector fields $V_1, \dots, V_\ell$ and maps f_0 and f_1 as in Lemma 3.3, where $t_0 = a$ and $t_1 = b$. Since f maps into B, we conclude that F is a B-variation of c with variation fields V_j. $\square$

We view Ω_B as a submanifold of Ω and $\mathcal{V}_B(c)$ as tangent space of Ω_B at c, for any $c \in \Omega_B$.

4.3. **Definition.** We say that $c \in \Omega_B$ is a *critical point* of the energy functional on Ω_B if the first variation of c vanishes, with respect to any B-variation F of c.

4.4. **Theorem.** *A curve $c \in \Omega_B$ is a critical point of the energy functional on Ω_B if and only if c is a geodesic such that*

$$(-c'(a), c'(b)) \perp T_{(c(a),c(b))}B \tag{9}$$

with respect to the product metric on $M \times M$.

4.5. **Definition.** We say that a geodesic $c \in \Omega_B$ is a *B-geodesic* if it is a critical point of E on Ω_B, that is, if it satisfies (9).

Proof of Theorem 4.4. If $c \in \Omega_B$ is a critical point of E on Ω_B, then c is a geodesic, by Lemma 3.5, and then (9) follows from (3). The other direction is clear. □

4.6. **Examples.** Critical points of E on Ω_B are
0) constant geodesics if $B = M \times M$;
1) geodesics $c\colon [a,b] \to M$ from p to q if $B = \{(p,q)\}$;
2) geodesics $c\colon [a,b] \to M$ from P to Q which hit P and Q perpendicularly at $t = a$ and $t = b$, respectively, if $B = P \times Q$;
3) periodic geodesics, that is, geodesics $c\colon [a,b] \to M$ with $c(a) = c(b)$ and $c'(a) = c'(b)$, if B is the diagonal in $M \times M$.

Recall that the second derivative of a smooth function at a point of a manifold, as a symmetric bilinear form on the tangent space at that point, is only well defined if the point is a critical point of the function. In our interpretation of critical points of the energy functional on Ω_B, this should mean that the second variation of energy is a symmetric bilinear form on $\mathcal{V}_B(c)$, for any B-geodesic c. To check this, let F be a 2-parameter B-variation of c with variation fields $(\partial_r F)(0,.) := V$ and $(\partial_s F(0,.) := W$. Then the last term in (4) turns into

$$\begin{aligned}\langle c'(b), (D_r\partial_s F)(0,b)\rangle &- \langle c'(a), (D_r\partial_s F)(0,a)\rangle \\ &= \langle(-c'(a), c'(b)), ((D_r\partial_s F)(a,t), (D_r\partial_s F)(b,t))\rangle \\ &= \langle(-c'(a), c'(b)), II((V(a), V(b)), (W(a), W(b))\rangle, \\ &=: II_n((V(a), V(b)), (W(a), W(b)),\end{aligned} \tag{10}$$

where $n := (-c'(a), c'(b))$ and II denotes the second fundamental form of B at $(c(a), c(b))$ with respect to the product metric on $M \times M$. Recall that n is normal to B, that is, $n \perp T_{(c(a),c(b))}B$, cf. (9).

4.7. **Proposition.** *In the above situation, we have*

$$\begin{aligned}H_B(V,W) &:= (\partial_r\partial_s E)(0,0) \\ &= I(V,W) + II_n((V(a), V(b)), (W(a), W(b))),\end{aligned} \tag{11}$$

where II_n denotes the second fundamental form of B at $(c(a), c(b))$ with respect to the normal vector $n = (-c'(a), c'(b))$. □

We see that the second variation of the energy on Ω_B at a B-geodesic c, the *Hessian* $H_B = H_B(V,W)$ of E at c, is a symmetric bilinear form on the space $\mathcal{V}_B(c)$ of B-vector fields along c.

5. Finite dimensional approximations

Our aim is to study critical point theory of the energy functional E on the spaces Ω_B. In standard critical point theory, a first requirement on the given function is that its sublevels are relatively compact. For a given boundary condition B, we replace this requirement by the following

5.1. **Compactness Condition.** *Given a constant $\kappa > 0$, there is a relatively compact open subset $U = U_{B,\kappa}$ such that*

$$c \in \Omega_B \text{ and } E(c) < \kappa \Longrightarrow \operatorname{im} c \subseteq U. \tag{C}$$

5.2. **Remarks.** 1) If M is compact, any boundary condition B satisfies Condition C. More generally, if B is compact, then Condition C is satisfied. To see this, let B_a be the projection of B onto the first factor in $M \times M$. Then any curve $c \in \Omega_B$ has initial point $c(a) \in B_a$. Hence, by (2), the image of c is contained in the open neighborhood U of radius $\sqrt{|b-a|\kappa/2}$ about B_a if $E(c) < \kappa$. Since B_a is compact and M is complete, U is relatively compact.
2) Condition C corresponds to the well known Palais-Smale condition in the calculus of variations, compare with [Pa] and [Kl].
3) It is reasonable to discuss critical point theory of E on path components of Ω_B, that is, on *B-homotopy classes* of curves in Ω_B. Then it suffices to require (C) for curves c in the corresponding B-homotopy class. This refinement of Condition C occurs in [Th], where Thorbergsson uses it to discuss the existence of periodic geodesics in free homotopy classes of closed curves on non-compact manifolds.

Throughout this part, let B be a boundary condition satisfying Condition C above. Let $\kappa > 0$, choose $U = U_{B,\kappa}$ according to Condition C, and set

$$\Omega_B^\kappa = \{c \in \Omega_B \mid E(c) < \kappa\}. \tag{12}$$

Condition C says that the images of curves in Ω_B^κ are contained in U. Since U is relatively compact, its injectivity radius is positive, compare with Theorem 2.2

For any $c \in \Omega_B^\kappa$ and $a \le s < t \le b$, we have $L(c|[s,t])^2 < 2|s-t|\kappa$, by (2), and hence

$$2|s-t|\kappa < \varepsilon^2 \Longrightarrow d(c(s), c(t)) < \varepsilon. \tag{13}$$

Fix a subdivison $t_0 < t_1 < \cdots < t_k$ of $[a,b]$ such that

$$t_i - t_{i-1} < \varepsilon^2/2\kappa, \quad \text{for all } 1 \le i \le k. \tag{14}$$

Then, by (13),

$$d(c(s), c(t)) < \varepsilon, \quad \text{for all } 1 \le i \le k \text{ and } s,t \in [t_{i-1}, t_i]. \tag{15}$$

Let $\Omega_B^\kappa(t_0,\dots,t_k) \subseteq \Omega_B^\kappa$ be the subspace of geodesic polygons, that is, of curves c in Ω_B^κ such that $c|[t_{i-1},t_i]$ is a geodesic, for all $1 \le i \le k$.

5.3. **Theorem.** *With respect to the compact-open topology, there is a deformation retraction of Ω_B^κ onto $\Omega_B^\kappa(t_0,\dots,t_k)$. Moreover, the sublevels $\Omega_B^\lambda(t_0,\dots,t_k)$ of E are relatively compact in $\Omega_B^\kappa(t_0,\dots,t_k)$, for all $\lambda < \kappa$.*

Proof. Define a deformation retraction $F\colon [0,1]\times\Omega_B^\kappa \to \Omega_B^\kappa$ of Ω_B^κ onto $\Omega_B^\kappa(t_0,\dots,t_k)$ by replacing the pieces of $c \in \Omega_B^\kappa$ between $c(t_{i-1})$ and $c(st_i+(1-s)t_{i-1})$ by the unique geodesic of length $< \varepsilon$ with these end points, appropriately parametrized, leaving the rest of c untouched, for all $0 \le s \le 1$ and $1 \le i \le k$. That is, for such s and i and $t_{i-1} \le t \le st_i + (1-s)t_{i-1}$, we set

$$F(s,c)(t) = c_{c(t_{i-1})c(st_i+(1-s)t_{i-1})}\big(\frac{t-t_{i-1}}{s(t_i-t_{i-1})}\big)$$

and $F(s,c)(t) = c(t)$ otherwise. Then $F(0,c) = c$ and $F(1,c)$ is the geodesic polygon which connects the points $c(t_0), c(t_1), \dots, c(t_k)$ consecutively by the unique geodesic segments of length $< \varepsilon$.

Since U is relatively compact, it follows easily that sublevels of E, for $\lambda < \kappa$, are relatively compact. $\square$

For any $c \in \Omega_B^\kappa(t_0,\dots,t_k)$, the points $p_i := c(t_i) \in U$ determine c, by (15). Furthermore,

$$E(c) = \sum_{1\le i\le k} \frac{d^2(c(t_{i-1}),c(t_i))}{t_i - t_{i-1}} = \sum_{1\le i\le k} \frac{d^2(p_{i-1},p_i)}{t_i-t_{i-1}}. \tag{16}$$

Thus $\Omega_B^\kappa(t_0,\dots,t_k)$ is canonically homoeomorphic to

$$\begin{aligned} N &:= M_B^\kappa(t_0,\dots,t_k) \\ &:= \{c = (p_0,p_1,\dots,p_k) \in M^{k+1} \mid (p_0,p_k) \in B \text{ and } E(c) < \kappa\}, \end{aligned} \tag{17}$$

where $E(c)$ is given by the right hand sum in (16). Note that N is an open subset of $B \times M^{k-1}$ and that the energy functional E is a smooth function on N. Thus we have arrived at the aimed for finite dimensional approximation N of Ω_B, given the upper bound κ on E.

We want to show that, up to homotopy equivalence, the topology of Ω_B is caught by the spaces $\Omega_B^\kappa(t_0,\dots,t_k)$.

5.4. **Theorem.** *For all $\kappa > 0$ and subdivisions $t_0 < \cdots < t_k$ of $[a,b]$, consider the inclusions $i\colon \Omega_B^\kappa(t_0,\dots,t_k) \to \Omega_B$. Let K be a compact metric space. Then we have:*

(1) *If $f\colon K \to \Omega_B$ is a continuous map, then there exists a continuous map $g\colon K \to \Omega_B^\kappa(t_0,\dots,t_k)$ such that $i\circ g$ is homotopic to f, for all sufficiently large $\kappa > 0$ and all sufficiently fine subdivisions $t_0 < \cdots < t_k$ of $[a,b]$ as above.*

(2) *Let $f_0, f_1\colon K \to \Omega_B$ and $g_0, g_1\colon K \to \Omega_B^\kappa(t_0,\dots,t_k)$ be continuous maps such $i\circ g_0$ is homotopic to f_0 and $i\circ g_1$ is homotopic to f_1. Then, if f_0 is homotopic to f_1, then g_0 and g_1 are homotopic in $\Omega_B^\lambda(s_0,\dots,s_\ell)$, for all sufficiently large $\lambda \geq \kappa$ and all sufficiently fine subdivisions $s_0 < \cdots < s_\ell$ of $[a,b]$.*

Proof. Recall that $f\colon K \to \Omega_B$ is continuous if and only if

$$F\colon K\times[a,b] \to M, \quad F(x,t) := f(x)(t), \tag{18}$$

is continuous. Since $K\times[a,b]$ is compact, there is a lower bound $\varepsilon > 0$ on the injectivity radius of M in the image of F as in Theorem 2.2. By the continuity of F and the compactness of $K\times[a,b]$, there is a $\delta > 0$ such that $d(F(x,s),F(y,t)) < \varepsilon$ if $d(x,y)+|s-t| < \delta$. As in the proof of Theorem 5.3, we can now deform f into a map $g\colon K \to \Omega_B^\kappa(t_0,\dots,t_k)$, for all sufficiently large $\kappa > 0$ and all sufficiently fine subdivisions of $[a,b]$. This proves (1), and the proof of (2) is similar. $\square$

We discuss now two nontrivial applications of what we discussed so far, assuming a basic fact from algebraic topology: For a compact manifold M of dimension m and a point $p\in M$, there is a non-trivial homotopy group $\pi_k(M,p)$, for some $k \geq 1$.

The main idea in the proof of Theorems 5.5 and 5.6 is described in [B1] in the case $M = S^2$ and in [B2] for $M = S^m$. However, higher homotopy groups had not been defined yet when [B1] and [B2] were written so that the convincing arguments of Birkhoff remain somewhat informal. Later, Lusternik and Fet observed the extension to general compact manifolds, using the fact about homotopy groups of compact manifolds mentioned above.

5.5. **Theorem.** *Let M be a compact Riemannian manifold and p be a point in M. Then there is a nontrivial geodesic loop $c\colon [0,1] \to M$ at p, that is, c is not constant and $c(0) = c(1) = p$.*

Proof. There is a first $1 \leq k \leq m$ such that $\pi_k(M,p)$ is nontrivial. Thus there is a continuous map $F\colon [0,1]^k \to M$ which sends the boundary $\partial[0,1]^k$ of $[0,1]^k$ to p such that F not homotopic to a constant map in the space $\mathcal{F}$ of all continuous maps $[0,1]^k \to M$ which send $\partial[0,1]^k$ to p. It is easy to see that F is homotopic, in $\mathcal{F}$, to a smooth map, and therefore we may assume that F is smooth.

The correspondence from (18) defines an isomorphism $\pi_k(M,p) = \pi_{k-1}(\Lambda_p, p)$, where (here) $\Lambda_p = \Omega_{pp}$ denotes the space of continuous loops at p (and p the constant loop at p), endowed with the compact-open topology. It follows that $f\colon [0,1]^{k-1} \to \Lambda_p$ is not homotopically trivial in the space $\mathcal{G}$ of continuous maps $[0,1]^{k-1} \to \Lambda_p$ which send the boundary $\partial[0,1]^{k-1}$ to the constant loop p.

Since M is compact, Condition C is satisfied. Furthermore, since F is smooth, there is a $\kappa > 0$ such that $E(f(x)) < \kappa$, for all $x \in [0,1]^{k-1}$. Hence, chosing a sufficiently fine subdivision $t_0 < \cdots < t_k$ of $[0,1]$, the deformation retraction in the proof of Theorem 5.3 applies and deforms f into a map f_1 with values in the finite dimensional manifold $\Lambda_p^\kappa(t_0,\dots,t_k)$. Note that this deformation leaves constant curves invariant. Hence it defines a homotopy of f to f_1 inside $\mathcal{G}$. If there would be no nontrivial geodesic loop $c\colon [0,1] \to M$ at p with $E(c) < \kappa$, then the constant loop p would be the only critical point of E on $\Lambda_p^\kappa(t_0,\dots,t_k)$. But then the gradient flow of E on $\Lambda_p^\kappa(t_0,\dots,t_k)$ would deform f_1 into a neighborhood of the constant loop p, where we note again that the gradient flow leaves the latter invariant. Then f_1, and hence also f, would be homotopic, in $\mathcal{G}$, to the constant map with value p, a contradiction. □

The above proof of Theorem 5.5 also applies in the more general case of complete and non-contractible Riemannian manifolds, except that the first non-trivial homotopy group $\pi_k(M,p)$ might occur only for some $k > m$.

5.6. **Theorem** (Lusternik and Fet [LF]). *For any compact Riemannian manifold M, there is a nontrivial periodic geodesic $c\colon [0,1] \to M$.*

Proof. Without loss of generality, we may assume that M is connected. We denote by Λ the space of continuous closed curves $c\colon [0,1] \to M$, endowed with the compact-open topology. Then the projection $P\colon \Lambda \to M$, $P(c) = c(0)$, is a Serre fibration with fibers the loop spaces Ω_q, $q \in M$. Hence there is a long exact sequence of homotopy groups,

$$\cdots \to \pi_{k+1}(M,p) \to \pi_k(\Omega_p, p) \to \pi_k(\Lambda, p) \to \pi_k(M,p) \to \cdots$$

This sequence splits since the canonical embedding $I\colon M \to \Lambda$, which identifies each point $q \in M$ with the corresponding constant loop at q, is a right inverse of P. It follows that $\pi_{k-1}(\Lambda, M) \cong \pi_k(M,p)$, where the isomorphism is induced by the correspondence from (18). Mutato mutandis, the rest of the argument is now the same as that in the proof of Theorem 5.5. □

6. Morse index theorem and ramifications

Let $c \in \Omega$ be a geodesic with end points $p = c(a)$ and $q = c(b)$. The Morse index theorem determines index and nullity of c as a critical point of the energy on Ω_{pq} in terms of conjugate points along c. The Hessian of E at c, denoted by H_0, is equal to the restriction of the index form I to $\mathcal{V}_0(c)$.

6.1. **Proposition.** *A vector field $V \in \mathcal{V}(c)$ belongs to the annihilator of $\mathcal{V}_0(c)$ with respect to I if and only if V is a Jacobi field along c.*

Proof. If V is a Jacobi field along c, then V belongs to the annihilator of $\mathcal{V}_0(c)$ with respect to I, by (6).

Conversely, assume that $V \in \mathcal{V}(c)$ belongs to the annihilator of $\mathcal{V}_0(c)$. Choose a subdivision $t_0 < \cdots < t_k$ of $[a, b]$ such that the restrictions $V|[t_{i-1}, t_i]$ are smooth. Suppose first that there is some t in some (t_{i-1}, t_i) such that $V''(t) + R(V(t), c'(t))c'(t) \neq 0$. Choose a smooth function $\varphi\colon \mathbb{R} \to \mathbb{R}$ with support in (t_{i-1}, t_i) such that $\varphi \geq 0$ and $\varphi(t) = 1$. Then $W := \varphi \cdot (V'' + R(V, c')c') \in \mathcal{V}_0(c)$ and $I_0(V, W) < 0$, a contradiction, since V belongs to the annihilator of $\mathcal{V}_0(c)$. It follows that the restrictions $V|[t_{i-1}, t_i]$ are Jacobi fields. It remains to check that $V'(t_i-) = V'(t_i+)$, for all $1 \leq i \leq k-1$. This follows easily from a corresponding choice of W. □

For a symmetric bilinear form H on a vector space V, *null space* and *nullity* of H are defined as

$$\begin{aligned} \ker H &:= \{v \in V \mid H(v, w) = 0 \text{ for all } w \in V\}, \\ \operatorname{null} H &:= \dim \ker H. \end{aligned} \tag{19}$$

Following this terminology, the null space $\mathcal{N}_0(c)$ and nullity $\operatorname{null}_0(c)$ of c as a critical point of E on Ω_{pq} are defined with respect to the Hessian H_0 of E at c,

$$\mathcal{N}_0(c) := \ker H_0 \quad \text{and} \quad \operatorname{null}_0(c) := \dim \mathcal{N}_0(c). \tag{20}$$

For $t \in (a, b]$, denote by $\mu(t)$ the multiplicity $\mu(t)$ of $c(t)$ as a conjugate point of $c(a)$ along c.

6.2. **Corollary.** *The nullity of c as a critical point of E on Ω_{pq} is equal to the multiplicity of $c(b)$ as a conjugate point of $c(a)$ along c,*

$$\operatorname{null}_0(c) = \mu(b) \leq \dim M - 1.$$

6.3. **Proposition.** *We have*

$$\mu(t) = 0, \textit{ for all } t \in (a, b] \quad \Longleftrightarrow \quad H_0 > 0.$$

6.4. **Remark.** Further on, we will only use the implication left to right and only under the stronger assumption that c is minimal beyond $c(b)$. On the other hand, the proof below only copies the usual proof of the fact that geodesics are minimal on some initial leg.

Proof of Proposition 6.3. Suppose first that $\mu(t) = 0$, for all $t \in (a, b]$. W.l.o.g., assume that $a = 0$ and $b = 1$. Then $q = \exp v$ with $v = c'(0)$.

Let $\mathcal{R}$ be the open subset of T_pM where the rank of exp is maximal, that is, where exp is a local diffeomorphism. Use exp to pull back the Riemannian metric of M to $\mathcal{R}$. Now a neighborhood of the line segment tv, $0 \le t \le 1$, belongs to $\mathcal{R}$, and hence the radial unit vector field in $\mathcal{R}$ can be used as usual to show that any proper variation of the line segment tv, $0 \le t \le 1$, has length at least $\|v\|$ with respect to the pulled back metric. Since exp is a local diffeomorphism along the line segment, we can lift any proper variation of c to a proper variation of the line segment. It follows that, for any proper variation of c, we have $L(c_s) \ge L(c)$, for all s sufficiently small. Hence $H_0 \ge 0$, by Proposition 3.8.

Suppose now that $V \in \mathcal{V}_0(c)$ satisfies $H_0(V, V) = 0$, and let $W \in \mathcal{V}_0(c)$. Then we have

$$0 \le H_0(V - sW, V - sW) = -2sH_0(V, W) + s^2 H_0(W, W).$$

It follows that $H_0(V, W) = 0$ and hence that $V \in \mathcal{N}_0(c)$. Hence V is a Jacobi field along c with $V(0) = V(1) = 0$, by Proposition 6.1. Since 1 is not conjugate to 0 along c, we conclude that $V = 0$. It follows that H_0 is strictly positive.

Conversely, suppose that $c(t)$, $t \in (a, b]$, is conjugate to $c(a)$ along c, and let J be a nonzero Jacobi field along c with $J(a) = J(t) = 0$. Define $V \in \mathcal{V}_0(c)$ to be equal to J on $[a, t]$ and to be zero on $[t, b]$. Then $V \ne 0$ and $H_0(V, V) = 0$, hence H_0 is not strictly positive. $\square$

6.5. **Corollary.** *Suppose that $\mu(t) = 0$, for all $t \in (a, b]$. Let $V \in \mathcal{V}(c)$, and let $W \in \mathcal{V}(c)$ be the unique Jacobi field along c with $W(a) = V(a)$ and $W(b) = V(b)$. Then*

$$I(W, W) < I(V, V) \text{ unless } W = V.$$

Proof. Suppose that $W \ne V$. Since $W - V \in \mathcal{V}_0(c)$ and $H_0 = I|\mathcal{V}_0(c)$, Proposition 6.3 applies and shows that

$$\begin{aligned} 0 &< I(V - W, V - W) \\ &= I(V, V) - 2I(W, V) + I(W, W) \\ &= I(V, V) - 2\langle W', V\rangle\big|_a^b + \langle W', V\rangle\big|_a^b \end{aligned}$$

$$= I(V,V) - \langle W', V\rangle\big|_a^b = I(V,V) - I(W,W),$$

where we use (6), that W is a Jacobi field, and that V and W coincide at the end points of $[a,b]$. □

Choose a subdivision $t_0 < t_1 < \cdots < t_k$ of $[a,b]$ such that no point in $(t_{i-1}, t_i]$ is conjugate to t_{i-1} along $c|[t_{i-1}, t_i]$, for all $1 \le i \le k$. Set

$$\mathcal{U}_0(c) = \{V \in \mathcal{V}_0(c) \mid V(t_i) = 0, \text{ for all } 0 \le i \le k\}. \tag{21}$$

$$\mathcal{P}_0(c) = \{V \in \mathcal{V}_0(c) \mid V|[t_{i-1}, t_i] \text{ is a Jacobi field, for all } 1 \le i \le k\}. \tag{22}$$

6.6. **Proposition.** *We have*

(1) $\mathcal{V}_0(c) = \mathcal{U}_0(c) \oplus \mathcal{P}_0(c)$ *as an* H_0*-orthogonal sum;*
(2) $H_0 > 0$ *on* $\mathcal{U}_0(c)$.
(3) $\dim \mathcal{P}_0(c) = (k-1)\dim M$.

Proof. The index form splits into the sum of the index forms of the restrictions $c|[t_{i-1}, t_i]$. By Propositions 6.1 and 6.3 applied to these restrictions, $\mathcal{P}_0(c)$ is the annihilator of $\mathcal{U}_0(c)$ with respect to $I = H_0$ and $H_0 > 0$ on $U_0(c)$, hence (1) and (2).

Since no point in $(t_{i-1}, t_i]$ is conjugate to t_{i-1} along $c|[t_{i-1}, t_i]$, any pair of tangent vectors v_{i-1} and v_i of M at $c(t_{i-1})$ and $c(t_i)$, respectively, determines a unique Jacobi field J along $c|[t_{i-1}, t_i]$ such that $J(t_{i-1}) = v_{i-1}$ and $J(t_i) = v_i$. Thus evaluation at $t_1, \dots, t_{k-1}$,

$$\mathcal{P}_0(c) \to T_{c(t_1)}M \oplus \cdots \oplus T_{c(t_{k-1})}M,\ V \mapsto (V(t_1), \dots, V(t_{k-1})), \tag{23}$$

is an isomorphism, hence (3). □

For a symmetric bilinear form H on a vector space V, the *index* of H is defined to be the maximal dimension of a subspace of V on which H is strictly negative. Following this terminology, the index $\mathrm{ind}_0(c)$ of c as a critical point of E on Ω_{pq} is defined to be the index of the Hessian H_0 of E at c.

From Proposition 6.1 we already know that the null space $\mathcal{N}_0(c)$ of H_0 is contained in $\mathcal{J}_0(c)$. Now Proposition 6.6 implies that index and nullity of c are realized on $\mathcal{P}_0(c)$:

6.7. **Corollary.** *We have*

$$\mathrm{ind}_0(c) = \mathrm{ind}(H_0|\mathcal{P}_0(c)) \quad \textit{and} \quad \mathrm{null}_0(c) = \mathrm{null}(H_0|\mathcal{P}_0(c)).$$

6.8. **Morse index theorem.** *There are only finitely many conjugate points* $c(t)$, $t \in (a,b]$, *of* $c(a)$ *along* c *and*

$$\mathrm{ind}_0(c) = \sum_{t \in (a,b)} \mu(t) \quad \textit{and} \quad \mathrm{null}_0(c) = \mu(b).$$

Proof. For $a < t \leq b$, let $c_t := c|[a,t]$ and consider $i(t) = \mathrm{ind}_0(c_t)$. If c does not have conjugate points along $[a,t]$, then $i(t) = 0$, by Proposition 6.3, hence

$$i(t) = 0, \text{ for all } t > a \text{ sufficiently close to } a. \tag{24}$$

Obviously, i is monotonically increasing,

$$\text{if } s \leq t, \text{ then } i(s) \leq i(t). \tag{25}$$

Let now $t \in (a,b]$ and choose a subdivision $t_0 < t_1 < \cdots < t_k$ of $[a,b]$ as further up, that is, such that no point in $(t_{i-1}, t_i]$ is conjugate to t_{i-1} along $c|[t_{i-1}, t_i]$, for all $1 \leq i \leq k$, and such that $t \in (t_{i-1}, t_i)$, for some $1 \leq i \leq k$. Following (23), we get an isomorphism

$$\begin{aligned} \mathcal{E}_t \colon \mathcal{P}_0(c_t) &\to T_{c(t_1)}M \oplus \cdots \oplus T_{c(t_{i-1})}M, \\ \mathcal{E}_t(V) &:= (V(t_1), \dots, V(t_{i-1})). \end{aligned} \tag{26}$$

Note that the target space does not depend on t as long as $t \in (t_{i-1}, t_i)$. Thus the index form on $\mathcal{J}_0(c_t)$ corresponds to a symmetric endomorphism I_t on a fixed Euclidean vector space, for all $t_{i-1} < t < t_i$. The index of c_t corresponds to the number of negative eigenvalues of I_t. Clearly, I_t depends continuously on $t \in (t_{i-1}, t_i)$ and hence

$$i(s) = i(t), \text{ for all } t \in (a,b] \text{ and all } s < t \text{ close to } t. \tag{27}$$

For all s sufficiently close to $t \in (t_{i-1}, t_i)$, the number of positive eigenvalues of I_s is at least the number of positive eigenvalues of I_t, hence

$$i(s) \leq i(t) + \mu(t), \text{ for all } t \in (a,b) \text{ and all } s > t \text{ close to } t. \tag{28}$$

Let $s > t$ and onsider the subdivisons $a = t_0 < \cdots < t_{i-1} < t$ of $[a,t]$ and $a = t_0 < \cdots < t_{i-1} < s$ of $[a,s]$. Suppose that $V \in \mathcal{P}_0(c_t)$ and $\tilde{V} \in \mathcal{P}_0(c_s)$ correspond to the same tuple of vectors under the isomorphisms $\mathcal{E}_t$ and $\mathcal{E}_s$ as in (26). Then $V|[a, t_{i-1}] = \tilde{V}|[a, t_{i-1}]$. Furthermore, we may think of V as a vector field along $c|[a,s]$ by setting $V = 0$ on $[t,s]$. Hence, since there are no conjugate points of $c(t_{i-1})$ along $[t_{i-1}, t_i]$, Corollary 6.5 applies and shows that $I(\tilde{V}, \tilde{V}) \leq I(V,V)$, where the inequality is strict if $V(t_{i-1}) \neq 0$. This latter is the case for nonzero Jacobi fields V along c_t with $V(0) = V(t) = 0$, that is, for nonzero elements in $\mathcal{N}_0(c_t)$. Hence

$$i(s) \geq i(t) + \mu(t), \text{ for all } a < t < s \leq b. \tag{29}$$

Now the claim about conjugate points and the index of c are immediate consequences of (24) – (29) (minus (26)). The claim about the nullity of c repeats the statement of Corollary 6.2. $\square$

Now we come to the ramifications. Let $B \subseteq M \times M$ be a boundary condition, and consider the Hessian H_B of E at c as in (11),

$$H_B(V,W) = I(V,W) + II_n((V(a),V(b)),(W(a),W(b))).$$

Choose a subdivision $t_0 < t_1 < \cdots < t_k$ of $[a,b]$ such that no point in $(t_{i-1},t_i]$ is conjugate to t_{i-1} along $c|[t_{i-1},t_i]$, for all $1 \le i \le k$. Define $\mathcal{U}_0(c)$ as in (21) and set

$$\mathcal{P}_B(c) = \{V \in \mathcal{V}_B(c) \mid \\ V|[t_{i-1},t_i] \text{ is a Jacobi field, for all } 1 \le i \le k\}. \tag{30}$$

6.9. **Proposition.** *We have*

(1) $\mathcal{V}_B(c) = \mathcal{U}_0(c) \oplus \mathcal{P}_B(c)$ *as an* H_B*-orthogonal sum;*
(2) $H_B > 0$ *on* $\mathcal{U}_0(c)$*.*
(3) $\dim \mathcal{P}_B(c) = \dim B + (k-1)\dim M$. □

Thus index $\mathrm{ind}_B(c) := \mathrm{ind}\, H_B$ and nullity $\mathrm{null}_B(c) := \mathrm{null}\, H_B$ of c as a critical point of E on Ω_B are achieved on the finite dimensional space $\mathcal{P}_B(c)$. Now $\mathcal{P}_0(c)$ is contained in $\mathcal{P}_B(c)$ and $H_B|\mathcal{P}_0(c) = H_0$, and therefore we would like to invest our results about H_0 for the determination of index and nullity of c.

6.10. **Proposition.** *Let H be a symmetric bilinear form on a real vector space V of finite dimension. Then, for any subspace $U \subseteq V$,*

$$\mathrm{null}(H) = \mathrm{null}(H|U^\perp) - \mathrm{def}\, U, \tag{31}$$

$$\mathrm{ind}(H) = \mathrm{ind}(H|U) + \mathrm{ind}(H|U^\perp) + \mathrm{def}\, U, \tag{32}$$

where $U^\perp$ denotes the annihilator with respect to H and

$$\mathrm{def}\, U := \dim(U \cap U^\perp) - \dim(U \cap \ker H).$$

Proof. We have $\ker H \subseteq U^\perp$ and $U^{\perp\perp} = U + \ker H$, hence

$$\ker(H|U^\perp) = U^\perp \cap U^{\perp\perp} = U \cap U^\perp + \ker H,$$

and hence (31). As for (32), let U_0 be a complement of $U \cap U^\perp$ in U and U_1 be a complement of $U^\perp \cap U^{\perp\perp}$ in $U^\perp$. Then U_0 and U_1 are maximal subspaces of U and $U^\perp$, respectively, on which H is non-degenerate. Hence $U_0 + U_1$ is a maximal subspace of $U + U^\perp$ on which H is non-degenerate. In particular, we have

$$V = U_0 + U_1 + (U_0 + U_1)^\perp$$

as a direct and H-orthogonal sum. Let W_0 be a complement of $U \cap \ker H$ in $U \cap U^\perp$ and W_1 be a complement of $U \cap U^\perp + \ker H$ in $(U_0 + U_1)^\perp$.

Then W_0+W_1 is a subspace of $(U_0+U_1)^\perp$ on which H is non-degenerate and

$$(U_0 + U_1)^\perp = \ker H + W_0 + W_1$$

as a direct sum. Since $\dim W_0 = \operatorname{def} U = \dim W_1$ and since H vanishes on W_0, it follows that

$$\operatorname{ind} H|(W_0 + W_1) = \frac{1}{2}\dim(W_0 + W_1) = \operatorname{def} U. \qquad \square$$

Now $\mathcal{P}_0(c) \subseteq \mathcal{P}_B(c)$, and the H_B-orthogonal complement of $\mathcal{P}_0(c)$ in $\mathcal{P}_B(c)$ is equal to

$$\mathcal{J}_B(c) = \{V \in \mathcal{P}_B(c) \mid V \text{ is a smooth Jacobi field}\}, \tag{33}$$

see (6) and (11). In particular,

$$\mathcal{P}_0(c) \cap \mathcal{J}_B(c) = \mathcal{J}_0(c), \tag{34}$$

the space of Jacobi fields J along c with $J(a) = J(b) = 0$ of dimension $\operatorname{null}_0(c)$. Furthermore,

$$\begin{aligned}\mathcal{P}_0(c) &\cap \ker(H_B|\mathcal{P}_B(c)) \\ &= \{V \in \mathcal{J}_0(c) \mid (-V'(a), V'(b)) \text{ is perpendicular to } B\},\end{aligned} \tag{35}$$

see again (6) and (11).

6.11. **Corollary.** *We have*

$$\operatorname{null}_B(c) = \operatorname{null}(H_B|\mathcal{J}_B(c)) - \operatorname{null}_0(c) + \nu, \tag{36}$$

$$\operatorname{ind}_B(c) = \operatorname{ind}(H_B|\mathcal{J}_B(c)) + \operatorname{ind}_0(c) + \operatorname{null}_0(c) - \nu, \tag{37}$$

where $\nu = \dim(\mathcal{P}_0(c) \cap \ker(H_B|\mathcal{P}_B(c)))$.

Now the challenge is the discussion of $\mathcal{J}_B(c)$ and H_B on $\mathcal{J}_B(c)$.

6.12. **Example.** As an example, consider the case where B is the diagonal in $M\times M$. Then $\Omega_B = \Lambda$, the free loop space of M, and B-geodesics are periodic geodesics. For any such geodesic $c\colon [a,b] \to M$,

(1) $\mathcal{J}_B(c)$ is the space of Jacobi fields V along c with $V(a) = V(b)$;
(2) $\ker(H_B|\mathcal{P}_B(c))$ is the space of periodic Jacobi fields along c, that is, Jacobi fields V along c with $V(a) = V(b)$ and $V'(a) = V'(b)$;
(3) $\mathcal{P}_0(c) \cap \ker(H_B|\mathcal{P}_B(c))$ is the space of periodic Jacobi fields V along c such that $V(a) = V(b) = 0$.

Since the diagonal is totally geodesic in $M \times M$, the second term in (11) vanishes. Furthermore, for $V, W \in \mathcal{J}_B(c)$ we have

$$H_B(V,W) = \langle V'(b) - V'(a), W(a)\rangle. \tag{38}$$

This formula connects H_B with the so-called Poincaré map of c; see [BTZ] for this, for a detailed study of index and nullity of periodic geodesics and their iterates, and for further references.

References

[BTZ] W. Ballmann, G. Thorbergsson, and W. Ziller: Closed geodesics on positively curved manifolds. *Annals of Math.* 116 (1982), 213–247.

[B1] G. D. Birkhoff: Dynamical systems with two degrees of freedom. *Transactions Amer. Math. Soc.* 18 (1917), 199–300.

[B2] G. D. Birkhoff: *Dynamical systems.* Amer. Math. Soc. Colloquium Publications, Vol. 9, 1927, viii+295 pp.

[Bo] R. Bott: *Morse theory and its applications to homotopy theory.* Lecture notes by A. van de Ven, Mathematisches Institut der Universität Bonn, 1960, 39 pp.

[CE] J. Cheeger and D. G. Ebin: *Comparison theorems in Riemannian geometry.* North-Holland Math. Library, Vol. 9, North-Holland Publishing Co., 1975, viii+174 pp.

[GKM] D. Gromoll, W. Klingenberg, and W. Meyer: *Riemannsche Geometrie im Grossen.* Lecture Notes in Mathematics, No. 55, Springer-Verlag, 1968, vi+287 pp.

[Kl] W. Klingenberg: *Lectures on closed geodesics.* Grundlehren der Mathematischen Wissenschaften, Vol. 230. Springer-Verlag, 1978, x+227 pp.

[LF] L. Lusternik and A. Fet: Variational problems on closed manifolds. *Dokl. Akad. Nauk SSSR (N. S.)* 81 (1951), 17–18 [Russian].

[LS] L. Lusternik and L. Schnirelmann: Sur le problème de trois géodésiques fermées sur les surfaces de genre 0. *C. R. Acad. Sci. Paris* 189 (1929), 209-271.

[Mi] J. Milnor: *Morse theory.* Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, 1963, vi+153 pp.

[Mö] A. F. Möbius: Theorie der elementaren Verwandschaft. *Berichte über die Verhandlungen der Königlichen Sächsischen Ges. Wiss., mathematisch-physikalische Klasse* 15 (1863), 18–57.

[Mo] M. Morse: *The calculus of variations in the large.* Amer. Math. Soc. Colloquium Publications, Vol. 18, 1934, ix+368 pp.

[Pa] R. S. Palais: Morse theory on Hilbert manifolds. *Topology* 2 (1963), 299–340.

[Th] G. Thorbergsson: Closed geodesics on non-compact Riemannian manifolds. *Math. Z.* 159 (1978), 249–258.

Mathematisches Institut, Universität Bonn, Endenicher Allee 60, 53115 Bonn, and Max-Planck-Institut für Mathematik, Vivatsgasse 7, 53111 Bonn, Germany

E-mail address: `ballmann@mpim-bonn.mpg.de`