universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Comparison Theorems in Riemannian Geometry

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Verfasser:	Klaus Kröncke
Matrikel-Nummer:	0404772
Studienrichtung:	A 405 Mathematik
Betreuer:	Ao. Univ.-Prof. Michael Kunzinger

Wien, am 12. Juli 2010

Abstract

In Chapter 1, we first introduce basic concepts of curvature. Then we summarize the most important results of the theory of coverings. Then we describe manifolds of constant curvature. Chapter 2 is about methods of variational calculus of curves. With these methods, we prove the Bonnet-Myers theorem, which gives a fundamental property of complete Riemannian manifolds whose sectional curvature is bounded from below by a positive constant. In Chapter 3, we first introduce Jacobi fields and conjugate points. The rest of the Chapter is spent devoted to the Theorem of Rauch and some of its important conclusions. Rauch's theorem describes the behavior of Jacobi fields depending on the sectional curvature. As a conclusion, we obtain, among other things, the Theorem of Cartan-Hadamard which is a fundamental result about the structure of complete Riemannian-manifolds of nonpositive sectional curvature. In Chapter 4, we prove the Morse Index theorem which states that the index of the Index form along a geodesic γ equals the number of conjugate points along γ, each counted with its multiplicity. Chapter 5 provides the foundations of Morse theory, which are needed in the proof of the Sphere theorem. The Sphere theorem is proven in Chapter 6.

Zusammenfassung

Im ersten Kapitel führen wir zunächst Grundkonzepte der Krümmung ein. Danach fassen wir die wichtigsten Resultate aus der Überlagerungstheorie zusammen. Zuletzt beschreiben wir Mannigfaltigkeiten konstanter Krümmung. In Kapitel 2 geht es um Techniken der Variationsrechnung. Mit diesen Methoden beweisen wir das Theorem von Bonnet-Myers, eine fundamentale Aussage über vollständige Riemann-Mannigfaltigkeiten, deren Schnittkrümmung durch eine positive Konstante nach unten beschränkt ist. In Kapitel 3 führen wir zunächst Jacobi-Feldern und konjugierte Punkte ein. Der Rest des Kapitels ist dem Theorem von Rauch und einigen wichtigen Folgerungen gewidmet. Rauch's Theorem beschreibt das Verhalten von Jacobi-Feldern in Abhängigkeit von der Schnittkrümmung. Als Folgerung erhält man unter anderem das Theorem von Cartan-Hadamard, eine fundamentale Aussage über die Struktur vollständiger Riemann-Mannigfaltigkeiten nichtpositiver Schnittkrümmung. In Kapitel 4 beweisen wir das Morse-Index Theorem, das aussagt, dass der Index der Index-Form entlang einer Geodäte γ die konjugierten Punkte entlang γ mit ihrer Vielfachheit zählt. Kapitel 5 ist den Grundlagen der Morse Theorie gewidmet, die für den Beweis des Sphärentheorems benötigt werden. Das Sphärentheorem wird in Kapitel 6 bewiesen.

Acknowledgements

I wish to thank my supervisor Michael Kunzinger for his enthusiasm and care during the development of my thesis. He provided me an insight into a wide and interesting research topic. I also want to thank him for great lecture courses about semi-Riemannian geometry and Singularity theory.

Many thanks also to James Grant, Stefan Haller and Andreas Kriegl for useful comments and inspiring discussions which helped me to resolve some problems.

Above all, I wish to thank my parents, Lone and Heinrich Kröncke, for their backing and encouragement to pursue my interests.

Contents

Preface

Riemannian geometry is the branch of differential geometry that studies Riemannian manifolds, smooth manifolds with a Riemannian metric, i.e. with an inner product on the tangent space at each point which varies smoothly from point to point. This gives, in particular, local notions of angle, length of curves, surface area, and volume. Riemannian geometry originated with the vision of Bernhard Riemann expressed in his lecture "Über die Hypothesen, welche der Geometrie zu Grunde liegen" in 1854. During the next hundred years, Riemannian geometry grew steadily but inconspicuously as a branch of mathematics. Inspired by work of H. Rauch and M. Berger, this field of mathematics has exploded with activity in the last fifty years of the twentieth century.

My interest on Riemannian geometry grew during a lecture course about Global Semi-Riemannian geometry which was given by M. Kunzinger in the summer term 2009. There, we studied the even more complicated structure of Semi-Riemannian manifolds which are (in dimension four and with index one) the main objects of the theory of general relativity.

The main subject of this thesis are results of Global Riemannian geometry. In all of these theorems we assume some local property of the space (usually formulated using curvature assumption) to derive some information about the global structure of the space, including some information on the topological type of the manifold. We will now describe some important results which we are going to prove in this thesis.

A classical result is known as the Bonnet-Myers theorem, which gives important information about manifolds of positive sectional curvature. We are going to prove the following version of the theorem: If the sectional curvature of a complete Riemannian manifold M is bounded from below by a positive constant, then M is compact and its fundamental group is finite. Moreover, we get an explicit estimate for the diameter of M. There exists a stronger form of the theorem (where the assumption on the sectional curvature is replaced by an assumption on the Ricci curvature) which was proven in 1941 by S. B. Myers [Mye41]. The proof uses techniques of variational calculus.

A fundamental result in Riemannian geometry was proven in 1951 by H. Rauch [Rau51] and is known as the Rauch comparison theorem. Rauch's theorem describes the behavior of Jacobi fields depending on the sectional curvature. It is an essential tool to study Riemannian manifolds on which certain bounds on the sectional curvature are given.

From Rauch's theorem, we are going to derive the theorem of Cartan-Hadamard, which states the following: If the sectional curve of a complete Riemannian manifold M is nonpositive, then $\mathbb{R}^n$ is the universal covering of M via the exponential map $\exp_p : T_pM \to M$. In particular, if M is simply-connected, this map is a diffeomorphism.

Another interesting question is whether a compact manifold whose sectional curvature varies in a sufficiently small interval is topologically a sphere. The answer was given satisfactorily around 1960 by M. Berger and W. Klingenberg who proved the following: If the sectional curvature of a compact Riemannian manifold M lies in the interval $(\frac{1}{4}, 1]$ (we say that the sectional curvature is 1/4-pinched) then M is homeomorphic to a sphere. This result is known as the Topological Sphere theorem. A far reaching generalization of this theorem was recently obtained by S. Brendle and R. Schoen, who classified all Riemannian manifolds, which are pointwise 1/4-pinched (i.e. for each $p \in M$, the sectional curvature K_p lies in the interval $(\frac{1}{4}, 1]$), up to diffeomorphism. This is known as the Differentiable Sphere theorem.

A crucial step in the proof of the Topological Sphere theorem is to get an estimate of the injectivity radius of the manifold M. The proof of this estimate needs some elements of Morse theory, which we are going to develop in Chapter 5. Another essential tool is the concept of the cut locus which describes when a geodesic, starting from a fixed point, is not minimizing anymore if it passes a certain point.

This thesis is mostly based on the book *Riemannian geometry* by do Carmo [Car92]. The section about the Cartan-Hadamard theorem is based on [Lee97]. The Chapter about Morse theory is based on a book of Milnor [Mil63].

Chapter 1

Curvature

1.1 Concepts of curvature

In this section, we introduce some basic notions of in a Riemannian manifold and state basic properties. The proofs can be found in many books, e.g. in Chapter 4 of [Car92]. Throughout, we assume all manifolds to be of class C^∞ and the manifold topology should be hausdorff and satisfy the second countable axiom.

Definition 1.1.1. Let M be a Riemannian manifold. The Riemannian curvature tensor is the map $R : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, defined by

$$R(X,Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z,$$

where ∇ denotes the Levi-Civita connection in M.

Remark 1.1.2. Various definitions of the Riemannian curvature tensor may differ by a sign. The definition used here is also used in [Car92] and [O'N83].

Proposition 1.1.3. *The map R is $C^\infty(M)$ multilinear, hence a $(1,3)$-tensor field and satisfies*

$$R(X,Y)Z = -R(X,Y)Z,$$

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0.$$

The second equation is called the First Bianchi Identity.

We want to give a local expression of R. Let $(\varphi = (x^1, \ldots, x^n), U)$ be a local coordinate system of M. We write $\partial_i = \frac{\partial}{\partial x^i}$. Let the functions g_{ij} and Γ_{ij}^k on U be defined by

$$g_{ij} = g(\partial_i, \partial_j), \qquad \nabla_{\partial_i} \partial_j = \sum_k \Gamma_{ij}^k \partial_k.$$

The coefficients g_{ij} are called the local expression of the Riemannian metric. The Γ_{ij}^k are called the Christoffel symbols of the Levi-Civita connection of M. Using the properties of the Levi-Civita connection, it is not hard to see that

$$\sum_l \Gamma_{ij}^l g_{lk} = \frac{1}{2} \left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right), \tag{1.1}$$

cf. [Car92, p. 55]. Let

$$R(\partial_i, \partial_j)\partial_k = \sum_l R^l_{ijk}\partial_l.$$

The R^l_{ijk} are the components of R in the coordinate system (φ, U). Since $[\partial_i, \partial_j] = 0$,

$$R(\partial_i, \partial_j)\partial_k = \nabla_{\partial_j}\nabla_{\partial_i}\partial_k - \nabla_{\partial_i}\nabla_{\partial_j}\partial_k = \nabla_{\partial_j}(\sum_l \Gamma^l_{ik}\partial_l) - \nabla_{\partial_i}(\sum_l \Gamma^l_{jk}\partial_l),$$

which by a direct calculation yields

$$R^s_{ijk} = \sum_l \Gamma^l_{ik}\Gamma^s_{jl} - \sum_l \Gamma^l_{jk}\Gamma^s_{il} + \frac{\partial}{\partial x^j}\Gamma^s_{ik} - \frac{\partial}{\partial x^i}\Gamma^s_{jk}. \tag{1.2}$$

Remark 1.1.4. If M is 1-dimensional, then $R = 0$: Let $X, Y, Z \in \mathfrak{X}(M)$ and let $\{E\}$ be a local frame field. Then we can locally write $X = f \cdot E$, $Y = g \cdot E$ for smooth functions f and g and therefore,

$$R(X, Y)Z = R(f \cdot E, g \cdot E)Z = fg \cdot R(E, E)Z = 0.$$

Consider the expression

$$(X, Y, Z, W) \mapsto \langle R(X, Y)Z, W\rangle, \qquad X, Y, Z, W \in \mathfrak{X}(M).$$

By Proposition 1.1.3, it is a $(0, 4)$-tensor field. It satisfies the following symmetries:

Proposition 1.1.5. *For all $X, Y, Z, W \in \mathfrak{X}(M)$, we have*

$$\begin{aligned}
\langle R(X,Y)Z,W\rangle + \langle R(Y,Z)X,W\rangle + \langle R(Z,X)Y,W\rangle &= 0\\
\langle R(X,Y)Z,W\rangle &= -\langle R(Y,X)Z,W\rangle\\
\langle R(X,Y)Z,W\rangle &= -\langle R(X,Y)W,Z\rangle\\
\langle R(X,Y)Z,W\rangle &= \langle R(Z,W)X,Y\rangle
\end{aligned}$$

Definition 1.1.6. Let M^n, $n \geq 2$ be a Riemannian manifold, $p \in M$ and $\sigma \subset T_pM$ be a two-dimensional subspace of T_pM. The sectional curvature of σ is defined as

$$K_p(\sigma) := \frac{\langle R_p(x, y)x, y\rangle_p}{\langle x, x\rangle_p\langle y, y\rangle_p - \langle x, y\rangle_p^2},$$

where $\{x, y\}$ is a basis of σ.

By straightforward calculation, it is shown, that $K_p(\sigma)$ is well defined, i.e. it is independent of the particular choice of the basis $\{x, y\}$. We write $K_p(\sigma)$ or $K_p(x, y)$ for a basis $\{x, y\}$ of σ.

Lemma 1.1.7. *Let M be a Riemannian manifold. If the sectional curvature K of M is constant for all $p \in M$ and all two dimensional subspaces $\sigma \subset T_pM$, then*

$$\langle R(X, Y)Z, W\rangle = K[\langle X, Z\rangle\langle Y, W\rangle - \langle Y, Z\rangle\langle X, W\rangle] \qquad \forall X, Y, Z, W \in \mathfrak{X}(M)$$

Lemma 1.1.8. *Let (M, g) be a Riemannian manifold, $\lambda > 0$ and $\bar{g} = \lambda \cdot g$. Let $K, \bar{K}$ be the sectional curvature with respect to the metric $g, \bar{g}$, respectively. Then*

$$\bar{K} = \frac{1}{\lambda} K.$$

Proof. Let $p \in M$ and $(\varphi = (x^1, \ldots, x^n), U)$ local coordinates around p. We will distinguish the various objects corresponding to $g, \bar{g}$ by a bar. By (1.1), $\Gamma_{ij}^k = \bar{\Gamma}_{ij}^k$, and therefore, by (1.2), $R_{ijk}^l = \bar{R}_{ijk}^l$. It follows that the curvature tensors R and $\bar{R}$ are equal. Thus, for $x, y \in T_pM$ linearly independent,

$$\begin{aligned} \bar{K}_p(x, y) &= \frac{\bar{g}_p(\bar{R}_p(x, y)x, y)}{\bar{g}_p(x, x)\bar{g}_p(y, y) - \bar{g}_p(x, y)^2} \\ &= \frac{\lambda \cdot g_p(R_p(x, y)x, y)}{\lambda^2[g_p(x, x)g_p(y, y) - g_p(x, y)^2]} = \frac{1}{\lambda} K_p(x, y). \end{aligned}$$

□

1.2 Covering spaces

Next we recall some basics of the theory of covering spaces. For more details, see e.g [Hal09] and [Hat02].

Definition 1.2.1. Let X be a topological space. A covering of X is a continuous map $\pi : \tilde{X} \to X$ such that for each $x \in X$ there exists an open neighborhood U with the following properties: There exists an index set Λ and disjoint open sets $\tilde{U}_\lambda \subset \tilde{X}$, $\lambda \in \Lambda$ such that $\pi^{-1}(U) = \bigcup_{\lambda \in \Lambda} \tilde{U}_\lambda$ and for each $\lambda \in \Lambda$, $\pi|_{\tilde{U}_\lambda} : \tilde{U}_\lambda \to U$ is a homeomorphism. We say that U is evenly covered.

In this regard, we call X the basis, $\tilde{X}$ the covering space and π the covering map. For $x \in X$, the discrete set $\pi^{-1}(x) \subset \tilde{X}$ is called the fiber over x.

Definition 1.2.2. Two coverings $\pi_1 : \tilde{X}_1 \to X$ and $\pi_2 : \tilde{X}_2 \to X$ are called isomorphic if there exists a homeomorphism $\varphi : \tilde{X}_1 \to \tilde{X}_2$ such that $\pi_2 \circ \varphi = \pi_1$. If $\tilde{X}_1 = \tilde{X}_2$ and $\pi_1 = \pi_2$, φ is called a covering transformation. The covering transformations of $\tilde{X}$ form a subgroup of the homeomorphism group of $\tilde{X}$, which we denote by $\mathcal{C}(\tilde{X})$.

Example 1.2.3. The map $\pi : \mathbb{R} \to S^1$, defined by $\pi(t) = e^{2\pi i t}$ is a covering map. For each $k \in \mathbb{Z}$, $\varphi_k : t \mapsto t + k$ is a covering transformation.

Proposition 1.2.4 (Lifting property). *Let $\varphi : \tilde{X} \to X$ be a covering and Y be locally path-connected and simply-connected. Fix a point $y \in Y$ and let $f : Y \to X$ be continuous. Then for each $\tilde{x} \in \pi^{-1}(f(y))$, there exists a unique continuous map $\tilde{f} : Y \to \tilde{X}$ such that $\pi \circ \tilde{f} = f$ and $\tilde{f}(y) = \tilde{x}$.*

Proof. See [Hat02], Proposition 1.33 and Proposition 1.34 □

Such a map $\tilde{f}$ is called a lift of f over π. In particular, if $\alpha : [0, a] \to X$ is a path in X and a point $\tilde{px} \in \pi^{-1}(\alpha(0))$ is fixed, there exists a unique path $\tilde{\alpha} : [0, a] \to \tilde{X}$ such that $\pi \circ \tilde{\alpha} = \alpha$ and $\tilde{\alpha}(0) = \tilde{px}$.

Definition 1.2.5. A covering $\pi : \tilde{X} \to X$ is called a universal covering if $\tilde{X}$ is simply connected.

Remark 1.2.6. A universal covering $\pi : \tilde{X} \to X$ is the largest covering of X in the following sense: If X is connected and locally path-connected and $\pi_1 : \tilde{X}_1 \to X$ is another covering with $\tilde{X}_1$ connected, there exists a covering $\pi_0 : \tilde{X} \to \tilde{X}_1$ such that $\pi_1 \circ \pi_0 = \pi$. (cf. [Hal09], Proposition II.6.5)

A universal covering is uniquely determined up to isomorphy: Let $\pi_1 : \tilde{X}_1 \to X$ and $\pi_2 : \tilde{X}_2 \to X$ be two universal coverings. Fix $x \in X$ and $\tilde{x}_i \in \pi_i^{-1}(x)$, $i = 1, 2$. By Proposition 1.2.4, there exists a map $\varphi : \tilde{X}_1 \to \tilde{X}_2$ such that $\pi_2 \circ \varphi = \pi_1$ and $\varphi(\tilde{x}_1) = \tilde{x}_2$ and a map $\psi : \tilde{X}_2 \to \tilde{X}_1$ satisfying $\pi_1 \circ \psi = \pi_2$ and $\psi(\tilde{x}_2) = \tilde{x}_1$. By uniqueness in Proposition 1.2.4, $\psi \circ \varphi = \mathrm{id}_{\tilde{X}_1}$ and $\varphi \circ \psi = \mathrm{id}_{\tilde{X}_2}$.

In particular, if X is simply connected, each universal covering of X is isomorphic to the trivial covering $\mathrm{id} : X \to X$.

A universal covering does not always exist. Some connectedness conditions are required:

Theorem 1.2.7. *If X is connected, locally path-connected and semilocally simply-connected, there exists a universal covering of X.*

Proof. See [Hal09], Satz II.6.9. □

Proposition 1.2.8. *Let X satisfy the conditions of Theorem 1.2.7 and let $\pi : \tilde{X} \to X$ be the universal covering of X. Then*

(i) The group $\mathcal{C}(\tilde{X})$ is isomorphic to the fundamental group $\pi_1(X)$.

(ii) For each $x \in X$, $\mathcal{C}(\tilde{X})$ is acting sharply transitive on $\pi^{-1}(x)$, i.e. for $\tilde{x}_1, \tilde{x}_2 \in \pi^{-1}(x)$, there exists a unique element $\varphi \in \mathcal{C}(\tilde{X})$ such that $\varphi(\tilde{x}_1) = \tilde{x}_2$.

Proof. For a proof of (i), see [Hal09], Korollar II.4.12. By Proposition 1.2.4 there exist unique maps $\varphi, \psi : \tilde{X} \to X$ satisfying $\varphi(\tilde{x}_1) = \tilde{x}_2$, $\psi(\tilde{x}_2) = \tilde{x}_1$ and $\pi \circ \varphi = \pi \circ \psi = \pi$. By uniqueness in Proposition 1.2.4, $\varphi \circ \psi = \psi \circ \varphi = \mathrm{id}_{\tilde{M}}$, so $\varphi, \psi \in \mathcal{C}(X)$ which proves (ii). □

Example 1.2.9. Let $\mathbb{RP}^n$, $n \geq 2$ be the real projective space, obtained from identifying the antipodal points of S^n. Then the canonical projection $\pi : S^n \to \mathbb{RP}^n$ is the universal covering of $\mathbb{RP}^n$ and the covering transformation group $\mathcal{C}(S^n)$ consists precisely of the maps $\pm\mathrm{id} : p \mapsto \pm p$. By Proposition 1.2.8 (i), the fundamental group $\pi_1(\mathbb{RP}^n)$ is isomorphic to $\mathbb{Z}_2$.

If M is a manifold and $\pi : \tilde{M} \to M$ is a covering, then there exists a unique smooth structure on $\tilde{M}$ such that π is a local diffeomorphism. Then all covering transformations are diffeomorphisms on $\tilde{M}$. Since M is locally diffeomorphic to $\mathbb{R}^n$, there exists an universal covering of M if M is connected. If two smooth maps $\pi_i : \tilde{M}_i \to M$, $i = 1, 2$ are universal coverings, there exists a diffeomorphism $\varphi : \tilde{M}_1 \to \tilde{M}_2$, such that $\pi_2 \circ \varphi = \pi_1$.

Let (M, g) be a connected Riemannian manifold and $\pi : \tilde{M} \to M$ be the universal covering of M. Then we define a Riemannian metric $\tilde{g}$ on $\tilde{M}$ by $\tilde{g} = \pi^* g$, i.e. $\tilde{g}_{\tilde{p}}(v, w) := g_{\pi(\tilde{p})}(T_p\pi(v), T_p\pi(w))$. This metric on $\tilde{M}$ is called the covering metric. It is the unique metric on $\tilde{M}$ such that $\pi : \tilde{M} \to M$ is a local isometry.

Proposition 1.2.10. *Let M be a connected Riemannian manifold and $\pi : \tilde{M} \to M$ the universal covering of M with the covering metric. To each vector field $X \in \mathfrak{X}(M)$ we associate a vector field $\pi^*(X) \in \mathfrak{X}(\tilde{M})$, defined by $\pi^*(X)_{\tilde{p}} = (T_{\tilde{p}}\pi)^{-1}(X_{\pi(\tilde{p})})$. Then we have:*

(i) $\nabla^{\tilde{g}}_{\pi^*(X)}\pi^*(Y) = \nabla^g_X Y$ *for all* $X, Y \in \mathfrak{X}(M)$

(ii) If γ is a geodesic in M, then each lift of γ over π is a geodesic in $\tilde{M}$.

(iii) $R^{\tilde{g}}(\pi^*(X), \pi^*(Y))\pi^*(Z) = R^g(X,Y)Z$ *for all* $X, Y, Z \in \mathfrak{X}(M)$

(iv) $K^{\tilde{g}}_{\tilde{p}}(\sigma) = K^g_{\pi(\tilde{p})}(T_{\tilde{p}}\pi(\sigma))$ *for any $\tilde{p} \in \tilde{M}$ and any two-dimensional subspace $\sigma \subset T_{\tilde{p}}\tilde{M}$*

(v) If M is complete (cf. Definition 2.2.2 below), $\tilde{M}$ is also complete.

Proof. Properties (i)-(iv) are general properties of local isometries. To show completeness of $\tilde{M}$, let $\tilde{\gamma}_0$ be a geodesic in $\tilde{M}$, starting at $\tilde{p}$. By completeness of M, the geodesic $\gamma_0 = \pi \circ \tilde{\gamma}_0$ can be extended to a geodesic γ, which is defined for all time. Its lift starting at $\tilde{p}$ is an extension of $\tilde{\gamma}_0$ which is defined for all time. This proves (v). □

1.3 Space forms

We now want to describe manifolds of constant sectional curvature K. We may assume, by similarity, that the sectional curvature $K = -1, 0, 1$, cf. Lemma 1.1.8. Let $\mathbb{R}^n$ be equipped with the usual metric, that is $g_{ij} = \delta_{ij}$ in the natural coordinate system. By (1.1) and (1.2), it is easy to see that $R \equiv 0$ and therefore $K \equiv 0$. In Chapter 6 of [Car92], Example 2.8, it is shown that the unit sphere $S^n \subset \mathbb{R}^{n+1}$ with the usual metric has constant sectional curvature $K = 1$. Both spaces are complete (cf. Definition 2.2.2 below) and simply connected.

We also want to give an example of a complete and simply connected manifold with constant sectional curvature $K = -1$. Consider the half-space of $\mathbb{R}^n$ given by

$$H^n = \{(x_1, \ldots, x_n) | x_n > 0\}$$

and introduce on H^n the metric

$$g_x(v, w) = \frac{1}{x_n^2}(v_1 w_1 + \ldots + v_n w_n).$$

The space H^n together with this metric is called the Poincaré half-plane model. Clearly, H^n is simply connected.

Proposition 1.3.1. *The Riemannian manifold H^n is complete and its sectional curvature K satisfies $K \equiv -1$.*

Proof. See [Car92, p. 160-162]. □

Theorem 1.3.2. *Let M^n be a complete Riemannian manifold with constant sectional curvature K. Then the universal covering $\tilde{M}$ of M is isometric to:*

(i) H^n, *if* $K = -1$,

(ii) $\mathbb{R}^n$, *if* $K = 0$,

(iii) S^n, *if* $K = 1$.

In particular, if M is simply-connected, it is itself isometric to one of the spaces above.

Proof. See Chapter 8 of [Car92], Theorem 4.1. □

Remark 1.3.3. Let M^n as above and $|K| \neq 1$. From Lemma 1.1.8, we obtain the following

(i) If $K > 0$, then $\tilde{M}$ is isometric to $S^n(K)$, the sphere with radius $\frac{1}{\sqrt{K}}$.

(ii) If $K < 0$, $\tilde{M}$ is isometric to $H^n(K)$, the half-space of $\mathbb{R}^n$ with the metric

$$g_x(v, w) = \frac{1}{K \cdot x_n^2}(v_1 w_1 + \ldots + v_n w_n).$$

.

Definition 1.3.4. The complete simply-connected Riemannian manifolds with constant sectional curvature are called space forms.

Chapter 2

Calculus of variation

2.1 Variation of arc length and energy

In this section we define for a given curve α the arc length and energy of α and investigate their behavior under small perturbations of α. The assertions concerning the length of a curve are taken from [Car92], Chapter 9. Those concerning the energy are taken from [O'N83], Chapter 10.

Definition 2.1.1. Let $\alpha : [0, a] \to M$ be a piecewise smooth curve in a Riemannian manifold M. We define the arc length of α by

$$L(\alpha) = \int_0^a |\alpha'(t)| dt$$

and the energy of α by

$$E(\alpha) = \int_0^a |\alpha'(t)|^2 dt.$$

Remark 2.1.2. The arc length is invariant under a monotonic and orientation preserving reparametrization of a curve. In fact, let $\varphi : [0, a] \to \mathbb{R}$ so that $\varphi' > 0$. Set $\psi = \varphi^{-1}$. Then $\tilde{\alpha} : [\varphi(0), \varphi(a)] \to M$, $\tilde{\alpha}(t') = \alpha \circ \psi(t')$, is a monotonic and orientation preserving reparametrization of α. For the length of $\tilde{\alpha}$, we have

$$L(\tilde{\alpha}) = \int_{\varphi(0)}^{\varphi(a)} |\tilde{\alpha}'(t')| dt' = \int_{\varphi(0)}^{\varphi(a)} |\alpha'(\psi(t'))| \psi'(t') dt' = \int_0^a |\alpha'(t)| dt = L(\alpha)$$

which proves the claim. In contrast, the energy of a curve is not invariant under reparametrizations. Choose e.g. $\varphi(t) = 2t$ and a curve $\alpha : [0, a] \to M$ with unit speed, i.e. $|\alpha'(t)| = 1$. Then

$$E(\tilde{\alpha}) = \int_0^{2a} |\tilde{\alpha}'(t')|^2 dt' = \int_0^{2a} |\alpha'(\psi(t'))|^2 |\psi'(t')|^2 dt' = \frac{1}{2} \int_0^a |\alpha'(t)|^2 dt = \frac{a}{2}$$

but

$$E(\alpha) = \int_0^a |\alpha'(t)|^2 dt = a.$$

Lemma 2.1.3. *Let $\alpha : [0,a] \to M$ be a piecewise smooth curve in M. Then*

$$L(\alpha)^2 \leq aE(\alpha)$$

and equality occurs if and only if the velocity of α is constant.

Proof. By the Cauchy-Schwarz inequality,

$$L(\alpha)^2 = \left(\int_0^a 1 \cdot |\alpha'(t)|dt\right)^2 \leq \int_0^a 1dt \cdot \int_0^a |\alpha'(t)|^2 dt = aE(\alpha)$$

and equality occurs if and only if $|\alpha'(t)|$ is constant. □

Definition 2.1.4. Let $\alpha : [0,a] \to M$ be a piecewise smooth curve and $0 = t_0 < \ldots < t_k = a$ be a subdivision of $[0,a]$ such that α is smooth on each subinterval $[t_{i-1}, t_i]$. A piecewise smooth variation of α is a map $x : (-\epsilon, \epsilon) \times [0,a] \to M$, such that $x(0,t) = \alpha(t)$ for $t \in [0,a]$ and such that x is smooth on $(-\epsilon,\epsilon) \times [t_{i-1}, t_i]$ for $i = 1, \ldots k$. For each $s \in (-\epsilon, \epsilon)$, the map $x(s,.) : t \mapsto x(s,t)$ is called a longitudinal curve of the variation. For a fixed $t \in [0,a]$, $s \mapsto x(s,t)$ is called a transversal curve of the variation.
The piecewise smooth V along α, given by $V(t) = \frac{\partial}{\partial s} x(s,t)|_{s=0} = T_{x(0,t)}(\partial_s)$ is called the variational vector field of the variation x. Each $V(t)$ is the velocity of the transversal curve $s \mapsto x(s,t)$ at $s = 0$. A variation is called proper, if it keeps the endpoints fixed, i.e. $x(s,0) = \alpha(0)$ and $x(s,a) = \alpha(a)$.

Proposition 2.1.5. *Let $V(t)$ be a piecewise smooth vector field along a piecewise smooth curve $\alpha : [0,a] \to M$. Then, there exists a variation $x : (-\epsilon,\epsilon) \times [0,a] \to M$ of α, such that V is the variational vector field of x. In addition, if $|V(0)| = |V(a)| = 0$, it is possible to choose x as a proper variation.*

Proof. See Chapter 9 of [Car92], Proposition 2.2. □

For a given variation x, we define two real-valued functions $L_x, E_x : (-\epsilon, \epsilon) \to \mathbb{R}$ by

$$L_x(s) = L(x(s,.)) = \int_0^a |\frac{\partial}{\partial t} x(s,t)|dt, \qquad E_x(s) = E(x(s,.)) = \int_0^a |\frac{\partial}{\partial t} x(s,t)|^2 dt.$$

In Remark 2.1.2, we have seen that the arc length of a curve is independent of the parametrization. Therefore, if we investigate the arc length of a variation of the curve α, we may assume that the velocity of α is constant. Throughout this section, we will fix a decomposition $0 = t_0 < \ldots t_k = a$ of $[0,a]$ and assume that α and V are smooth on each subinterval $[t_{i-1}, t_i]$.

Proposition 2.1.6 (Formula for the first variation)**.** *Let $\alpha : [0,a] \to M$ be a piecewise smooth curve and let $x : (-\epsilon,\epsilon) \times [0,a] \to M$ be a variation of α. Then*

$$\begin{aligned} E'_x(0) = &- 2\int_0^a \langle V(t), \alpha''(t)\rangle dt \\ &- 2\sum_{i=1}^{k-1} \langle V(t_i), \Delta\alpha'(t_i)\rangle - 2\langle V(0), \alpha'(0)\rangle + 2\langle V(a), \alpha'(a)\rangle \end{aligned} \tag{2.1}$$

where $V(t)$ is the variational vector field of x and

$$\Delta\alpha'(t_i) = \lim_{t \searrow t_i} \alpha'(t) - \lim_{t \nearrow t_i} \alpha'(t).$$

If in addition, α has constant velocity $c > 0$ then we have the following formula for the first variation of arc length:

$$\begin{aligned} L'_x(0) = &-\frac{1}{c}\int_0^a \langle V(t), \alpha''(t)\rangle dt \\ &-\frac{1}{c}\sum_{i=1}^{k-1}\langle V(t_i), \Delta\alpha'(t_i)\rangle - \frac{1}{c}\langle V(0), \alpha'(0)\rangle + \frac{1}{c}\langle V(a), \alpha'(a)\rangle \end{aligned} \tag{2.2}$$

Proof. For a proof of (2.1),see Chapter 9 of [Car92], Proposition 2.4. A proof of (2.2) is given in [O'N83], Proposition 10.2. □

The following results show that the critical points of arc length and energy are precisely the unbroken geodesics.

Proposition 2.1.7. *Let $\alpha : [0, a] \to M$ be a piecewise smooth curve. Then, α is an unbroken geodesic if and only if $E'_x(0) = 0$ for each proper variation x of α. If α has constant velocity $c > 0$, then α is an unbroken geodesic if and only if $L'_x(0) = 0$ for each proper variation x of α.*

Proof. A proof of the fist statement is given in Chapter 9 of [Car92], Proposition 2.5. For a proof of the second statement, see [O'N83], Corollary 10.3. □

Every vector field V along a smooth curve α splits into a component which is parallel to α' and a component, which is orthogonal to α'. We denote these components by $tan\, V$ and $nor\, V$. They are given by

$$tan\, V = \langle V, \alpha'\rangle\alpha', \qquad nor\, V = V - \langle V, \alpha'\rangle\alpha'.$$

If α is a geodesic, we have, since $|\alpha''| = 0$,

$$(tan\, V)' = (\langle V, \alpha'\rangle\alpha')' = \langle V', \alpha'\rangle\alpha' = tan(V')$$

and

$$(nor\, V)' = (V - tan\, V)' = V' - tan(V') = nor(V').$$

In this case, we just write $tan\, V'$ and $nor\, V'$, respectively.

Proposition 2.1.8 (Formula for the second variation of arc length)**.** *Let $\gamma : [0, a] \to M$ be a geodesic with constant velocity $c > 0$ and let x be a proper variation of γ with variational vector field V. Then*

$$L''_x(0) = \frac{1}{c}\int_0^a [\langle nor\, V', nor\, V'\rangle - \langle R(V, \gamma')V, \gamma'\rangle]dt, \tag{2.3}$$

$$E''_x(0) = 2\int_0^a [\langle V', V'\rangle - \langle R(V, \gamma')V, \gamma'\rangle]dt. \tag{2.4}$$

Proof. For a proof of (2.2), see [O'N83], Theorem 10.4. A proof of (2.4) is given in Chapter 9 of [Car92], Proposition 2.8. □

Remark 2.1.9. Using integration by parts leads to the following:

$$
\begin{aligned}
L_x''(0) = &- \frac{1}{c} \int_0^a [\langle nor\, V'' - R(nor\, V, \gamma')\gamma', nor\, V\rangle] dt \\
&- \frac{1}{c} \sum_{i=1}^{k-1} \langle \Delta nor\, V', nor\, V\rangle(t_k)
\end{aligned} \tag{2.5}
$$

$$
E_x''(0) = - 2 \int_0^a [\langle V'' - R(V, \gamma')\gamma', V\rangle] dt - 2 \sum_{i=1}^{k-1} \langle \Delta V', V\rangle(t_k) \tag{2.6}
$$

For calculations of the formulas (2.5) and (2.6), see [O'N83], Corollary 10.8 and Chapter 9 of[Car92], Remark 2.9, respectively.

2.2 The Bonnet-Myers Theorem

Let M be a Riemannian manifold. Consider the map $d : M \times M \to \mathbb{R}$, given by

$$d(p, q) = \inf \{L(\alpha)|\alpha : [0, 1] \to M \text{ piecewise smooth, } \alpha(0) = p,\, \alpha(1) = q\}\,.$$

The map d defines a metric on M, called the Riemannian distance function. The diameter of M is defined by $\operatorname{diam}(M) = \sup \{d(p, q)|p, q \in M\}$.

Theorem 2.2.1 (Hopf-Rinow)**.** *Let M be a connected Riemannian manifold. Then the following assertions are equivalent:*

(i) The metric space (M, d) is complete.

(ii) There exists a point $p \in M$ such that $\exp_p$ is defined for all $v \in T_pM$.

(iii) For all $p \in M$, $\exp_p$ is defined for all $v \in T_pM$. (M is geodesically complete)

(iv) All closed and bounded subsets of M are compact.

These statements imply the following: For any $p, q \in M$, there exists a geodesic γ joining p to q such that $L(\gamma) = d(p, q)$, i.e. γ is minimizing.

Proof. See [O'N83], Proposition 5.21 and Proposition 5.22. □

Definition 2.2.2. A Riemannian manifold satisfying the assertions of Theorem 2.2.1 is called complete.

Remark 2.2.3. Let M be a complete Riemannian manifold and let $\operatorname{diam}(M)$ be finite. Then M itself is a closed and bounded subset of M, hence by Theorem 2.2.1, M is compact.

The following theorem states that a complete Riemannian manifold M is compact if the sectional curvature of M is bounded below by a positive constant. Moreover, we get an explicit estimate for the diameter. The proof is taken from [Lee97], Theorem 11.7.

Theorem 2.2.4 (Bonnet-Myers). *Let M be a complete Riemannian manifold whose sectional curvature K satisfies*

$$K \geq \delta > 0$$

for a real number δ. Then M is compact and its diameter satisfies $\mathrm{diam}(M) \leq \frac{\pi}{\sqrt{\delta}}$. *Moreover, the fundamental group $\pi_1(M)$ is finite.*

Proof. We first prove that $\mathrm{diam}(M) \leq \frac{\pi}{\sqrt{\delta}}$. Suppose the contrary: then there exist $p, q \in M$ such that $d(p,q) > \frac{\pi}{\sqrt{\delta}}$. By Theorem 2.2.1, there exists a minimizing geodesic γ joining p to q. After a reparametrization, we may assume that γ is normalized, i.e. $|\gamma'| = 1$. Then, γ is defined on the interval $[0, L]$, where $L = L(\gamma) = d(p,q)$. Let W be a parallel unit vector field along γ such that $\langle W, \gamma' \rangle = 0$ and let

$$V(t) = \sin\left(\frac{\pi t}{L}\right) W(t).$$

Observe that V vanishes at $t = 0$ and $t = L$ and that $V = nor\, V$. By straightforward calculation, we get

$$V'(t) = \frac{\pi}{L}\cos\left(\frac{\pi t}{L}\right) W(t), \qquad V''(t) = -\left(\frac{\pi}{L}\right)^2 \sin\left(\frac{\pi t}{L}\right) W(t).$$

By Proposition 2.1.5, we can choose a proper variation x of γ such that V is the variational vector field of x. Since γ is a geodesic, $L_x'(0) = 0$ by Proposition 2.1.7. By the formula for the second variation of arc length (2.5),

$$\begin{aligned}
L_x''(0) &= -\int_0^L \langle V'' - R(V,\gamma')\gamma', V\rangle dt \\
&= \int_0^L \left\langle \frac{\pi^2}{L^2}\left(\sin\frac{\pi t}{L}\right) W + \left(sin\frac{\pi t}{L}\right) R(W,\gamma')\gamma', \left(\sin\frac{\pi t}{L}\right) W \right\rangle dt \\
&= \int_0^L \left(\sin^2\frac{\pi t}{L}\right)\left(\frac{\pi^2}{L^2}\langle W, W\rangle - \langle R(W,\gamma')W, \gamma'\rangle\right) dt \\
&= \int_0^L \left(\sin^2\frac{\pi t}{L}\right)\left(\frac{\pi^2}{L^2} - K(W,\gamma')\right) dt \\
&\leq \int_0^L \left(\sin^2\frac{\pi t}{L}\right)\left(\frac{\pi^2}{L^2} - \delta\right) dt < 0.
\end{aligned}$$

Therefore, if $s \neq 0$ is small enough, $L(x(s,.)) = L_x(s) < L_x(0) = L(\gamma)$, so $x(s,.)$ is a curve of shorter length than γ which joins p to q. This contradicts the assumption that $L(\gamma) = d(p,q)$. Hence, the diameter of M is at most $\frac{\pi}{\sqrt{\delta}}$. By Remark 2.2.3, M is compact.

To show finiteness of the fundamental group $\pi_1(M)$, let $\pi : \tilde{M} \to M$ be the universal covering of M with the covering metric. Then by Proposition 1.2.10, $\tilde{M}$ is a complete manifold whose sectional curvature $\tilde{K}$ also satisfies $\tilde{K} \geq \delta > 0$, so $\tilde{M}$ is compact by the argument above. By Proposition 1.2.8, there exists a one-to-one correspondence between the fiber $\pi^{-1}(p)$ over each point $p \in M$ and the fundamental group $\pi_1(M)$. Since $\pi^{-1}(p)$ is a discrete set in the compact manifold $\tilde{M}$, $\pi^{-1}(p)$ is finite. Thus, $\pi_1(M)$ is also finite. □

Remark 2.2.5. The existence of a $\delta > 0$ satisfying $K \geq \delta > 0$ is essential. Consider the paraboloid in $\mathbb{R}^3$, given by

$$M = \left\{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = z\right\}$$

with the metric induced from the standard Euclidean metric of $\mathbb{R}^3$. The sectional curvature of M satisfies $K > 0$ but M is not compact.

Chapter 3

The Rauch comparison theorem

3.1 Jacobi fields

Definition 3.1.1. Let $\gamma : [0, a] \to M$ be a geodesic. A Jacobi field along γ is a vector field $J \in \mathfrak{X}(\gamma)$ which satisfies

$$J'' = R(J, \gamma')\gamma' \tag{3.1}$$

This differential equation is called the Jacobi equation.

Definition 3.1.2. A variation of the geodesic γ is called a geodesic variation if all longitudinal curves of the variation are geodesic.

The following proposition states that the variational vector fields of geodesic variations of γ are precisely the Jacobi fields of γ.

Proposition 3.1.3. *If x is a geodesic variation of γ, then the variational vector field J of x is a Jacobi field. Conversely, if J is a Jacobi field along γ, there exists a geodesic variation of γ with variational vector field J.*

Proof. See [O'N83], Lemma 8.3 and [Bär06], Proposition 3.4.2. □

Proposition 3.1.4. *Each Jacobi field is smooth. For arbitrary $v, w \in T_pM$, there exists a unique Jacobi field J such that $J(0) = v$ and $J'(0) = w$. Thus, the set of all Jacobi fields along γ forms a vector space of dimension $2n$, where $n = \dim(M)$.*

Proof. See [O'N83], Lemma 8.5. □

Lemma 3.1.5. *Let $\gamma : [0, a] \to M$ be a geodesic and let J be a Jacobi field along γ. Then,*

$$\langle J, \gamma' \rangle(t) = \langle J'(0), \gamma'(0) \rangle \cdot t + \langle J(0), \gamma'(0) \rangle.$$

Proof. We obtain from the Jacobi equation

$$\langle J, \gamma' \rangle'' = \langle J'', \gamma' \rangle = \langle R(J, \gamma')\gamma', \gamma' \rangle = 0,$$

hence the function $f : [0,a] \to \mathbb{R}$, $f(t) = \langle J, \gamma' \rangle(t)$ is linear. Therefore,

$$f(t) = f'(0) \cdot t + f(0) = \langle J'(0), \gamma'(0) \rangle t + \langle J(0), \gamma'(0) \rangle.$$

□

For a fixed geodesic γ, we denote the space of Jacobi fields along γ by $\mathcal{J}$. The subspaces of $\mathcal{J}$ containing the Jacobi fields which are parallel and orthogonal to γ' are denoted by $tan\mathcal{J}$ and $nor\mathcal{J}$, respectively.

Proposition 3.1.6. *A vector field $J \in \mathfrak{X}(\gamma)$ is a Jacobi field if and only if $tanJ$ and $norJ$ are Jacobi fields. We obtain a splitting of $\mathcal{J}$ into a direct sum $\mathcal{J} = tan\mathcal{J} \oplus nor\mathcal{J}$.*

Proof. By linearity, $J = tanJ + norJ$ is a Jacobi field, if $tanJ$ and $norJ$ are. Conversely, suppose that J is a Jacobi field. We first show that $tanJ$ satisfies the Jacobi equation. By Lemma 3.1.5,

$$(tanJ)'' = \langle J, \gamma' \rangle'' \gamma' = 0 \cdot \gamma' = R(\langle J, \gamma' \rangle \gamma', \gamma')\gamma' = R(tanJ, \gamma')\gamma',$$

hence $tanJ$ is a Jacobi field. By linearity, $norJ = J - tanJ$ is also a Jacobi field. It follows that the map $\phi : \mathcal{J} \to tan\mathcal{J} \oplus nor\mathcal{J}$, defined by $\phi(J) = (tanJ, norJ)$, is well defined. It is a linear isomorphism, since its inverse is given by $\psi : tan\mathcal{J} \oplus nor\mathcal{J} \to \mathcal{J}$, $\psi(J_1, J_2) = J_1 + J_2$. □

Proposition 3.1.7 (Properties of tangential and orthogonal Jacobi fields)**.**

(i) The elements of $tan\mathcal{J}$ are precisely those of the form

$$J(t) = (c_1 + c_2 t)\gamma' \qquad c_1, c_2 \in \mathbb{R}.$$

(ii) A Jacobi field J is parallel to γ', if and only if $J(0)$ and $J'(0)$ are parallel to $\gamma'(0)$. Similarly, J is orthogonal to γ' if and only if $J(0)$ and $J'(0)$ are orthogonal to $\gamma'(0)$.

(iii) The dimension of $tan\mathcal{J}$ equals 2 and the dimension of $nor\mathcal{J}$ equals $2n-2$.

Proof. If $J \in tan\mathcal{J}$, then $J = \langle J, \gamma' \rangle \gamma'$. With $c_1 = \langle J(0), \gamma'(0) \rangle$ and $c_2 = \langle J'(0), \gamma'(0) \rangle$, the expression $J(t) = (c_1 + c_2 t)\gamma'$ follows from Lemma 3.1.5. Conversely, it is easy to check that each vector field of this form is a Jacobi field. This proves (i).

If J is parallel to γ', it is of the form $J(t) = (c_1 + c_2 t)\gamma'$, so $J(0) = c_1\gamma'(0)$ and $J'(0) = c_2\gamma'(0)$ are parallel to $\gamma'(0)$. Conversely, let $v, w \in T_pM$ be parallel to $\gamma'(0)$, i.e. $v = c_1\gamma'(0)$ and $w = c_2\gamma'(0)$. Then $J(t) = (c_1 + c_2 t)\gamma'$ is the unique Jacobi field which satisfies $J(0) = c_1\gamma'(0)$ and $J'(0) = c_2\gamma'(0)$ and by (i), J is parallel to γ. The analogous statement for $J \in nor\mathcal{J}$ is an immediate consequence of Lemma 3.1.5.

At last, we prove (iii). By (i), $\dim(tan\mathcal{J}) = 2$ and by Proposition 3.1.6, $\dim(nor\mathcal{J}) = \dim(\mathcal{J}) - \dim(tan\mathcal{J}) = 2n - 2$. □

Proposition 3.1.8 (Jacobi fields on manifolds with constant curvature)**.** *Let M be a Riemannian manifold with constant sectional curvature K and let γ :*

$[0,a] \to M$ be a normalized geodesic. Let J be a Jacobi field along γ which satisfies $J(0) = 0$ and $\langle J, \gamma' \rangle = 0$. Then

$$J(t) = \begin{cases} \frac{\sin(\sqrt{K}t)}{\sqrt{K}} W(t) & \text{if } K > 0, \\ tW(t) & \text{if } K = 0, \\ \frac{\sinh(\sqrt{|K|}t)}{\sqrt{|K|}} W(t) & \text{if } K < 0, \end{cases}$$

where W denotes the parallel vector field along γ with $W(0) = J'(0)$.

Proof. Let $\{E_1, \dots, E_n\}$ be a frame field along γ such that $E_1 = \gamma'$ and $E_2 = \frac{W}{|W|}$. This is possible since by Proposition 3.1.7 (ii), $J'(0)$ is orthogonal to $\gamma'(0)$. Then we can write

$$J(t) = \sum_{i=1}^{n} f_i(t) E_i(t).$$

Note that f_i is given by $\langle J, E_i \rangle$. Since $\langle J, \gamma' \rangle = 0$, $f_1 = 0$. For $i \geq 2$, the functions f_i satisfy

$$\begin{aligned} f_i'' &= \langle J'', E_i \rangle = \langle R(J, \gamma')\gamma', E_i \rangle \\ &= \langle R(\sum_{j=2}^{n} f_j E_j, \gamma')\gamma', E_i \rangle = \sum_{j=2}^{n} f_j \langle R(E_j, \gamma')\gamma', E_i \rangle \\ &\overset{1.1.7}{=} K \sum_{j=2}^{n} f_j [\langle E_j, \gamma' \rangle \langle \gamma', E_i \rangle - \langle \gamma', \gamma' \rangle \langle E_j, E_i \rangle] = -K f_i \end{aligned} \tag{3.2}$$

Now we want to calculate the initial conditions of (3.2). Since $|J(0)| = 0$, $f_i(0) = 0$ for $i \geq 2$. Since

$$|J'(0)| E_2(0) = |W(0)| E_2(0) = W(0) = J'(0) = \sum_{i=2}^{n} f_i'(0) E_i(0),$$

$f_2'(0) = |J'(0)|$ and $f_i'(0) = 0$ for $i \geq 3$. Therefore, $f_i = 0$ for $i \geq 3$. Thus, J is given by

$$J(t) = f_2(t) E_2(t) = \frac{f_2(t)}{|W(t)|} W(t) = \frac{f_2(t)}{|J'(0)|} W(t)$$

where f_2 is the unique solution of the initial value problem

$$f_2'' = -K f_2, \qquad f_2(0) = 0, \qquad f_2'(0) = |J'(0)|.$$

Solving this differential equation finishes the proof. □

3.2 Conjugate points

Definition 3.2.1. Let $\gamma : [0,a] \to M$ be a geodesic. Two points $\gamma(t_1), \gamma(t_2)$, $t_1 \neq t_2$ are called conjugate along γ, if there exists a Jacobi field $J \neq 0$ along γ such that $|J(t_1)| = 0$ and $|J(t_2)| = 0$.

Remark 3.2.2. Such a Jacobi field is orthogonal to γ': By Lemma 3.1.5, $t \mapsto \langle J, \gamma' \rangle(t)$ is a linear function and $\langle J, \gamma' \rangle(t_1) = \langle J, \gamma' \rangle(t_2) = 0$. Therefore, $\langle J, \gamma' \rangle(t) = 0$ for all $t \in [0,a]$.

Mostly, we will fix one of the two points at $t_1 = 0$. Let $\mathcal{J}_t$ be the space of Jacobi fields along γ' which vanish at $\gamma(0)$ and $\gamma(t)$. Note that $\mathcal{J}_t \neq \{0\}$ if and only if $\gamma(0)$ and $\gamma(t)$ are conjugate along γ'.

Definition 3.2.3. The dimension of the vector space $\mathcal{J}_t$ is called the multiplicity of the conjugate point $\gamma(t)$.

By Remark 3.2.2, $\mathcal{J}_t \subset nor\mathcal{J}$. Hence, an element $J \in \mathcal{J}_t$ is uniquely determined by the vector $J'(0)$, which is orthogonal to $\gamma'(0)$ by Proposition 3.1.7 (ii). Therefore, the multiplicity of a conjugate point $\gamma(t)$ can be at most $n-1$.

Example 3.2.4 (Conjugate points on manifolds with constant curvature)**.** Let M be a manifold of constant sectional curvature K and $\gamma : [0, a] \to M$ be a normalized geodesic.

- If $K > 0$, each point $\gamma(\frac{\pi}{\sqrt{K}}m)$, $m \in \mathbb{N}$ is conjugate to $\gamma(0)$ with multiplicity $n-1$ and all conjugate points are of this form: By Remark 3.2.2 it suffices to consider all $J \in nor\mathcal{J}$ with $J(0) = 0$. By Proposition 3.1.8, $J(t) = \sin(\sqrt{K}t)W(t)$ for a parallel vector field W, so J vanishes exactly at the points $\frac{\pi}{\sqrt{K}}m$.
- If $K \leq 0$, no conjugate points occur on γ. By Proposition 3.1.8, J is of the form $J(t) = tW(t)$ or $J(t) = \sinh(\sqrt{|K|}t)W(t)$ for a parallel vector field W. In both cases, $|J(t)| > 0$ for all $t > 0$.

Proposition 3.2.5. *Let $\gamma : [0, a] \to M$ be a geodesic and $p = \gamma(0)$. The following assertions are equivalent:*

(i) The point $\gamma(t)$ is conjugate to $\gamma(0)$ along γ.

(ii) The exponential map $\exp_p$ is singular at $t\gamma'(0)$, i.e. there exists a vector $v \in T_{t\gamma'(0)}(T_pM)$ such that $T_{t\gamma'(0)}exp_p(v) = 0$.

(iii) There exists a nontrivial geodesic variation x of γ, such that $x(s, 0) = \gamma(0)$ for all s and $\frac{\partial}{\partial s}x(0, t) = 0$. In this regard, nontrivial means that not all longitudinal curves of x equal γ.

Proof. See [O'N83], Proposition 10.10. □

Corollary 3.2.6. *Let $p \in M$ and $U \subset T_pM$ be a neighborhood of $0 \in T_pM$ such that $\exp_p : U \to M$ is a local diffeomorphism at each $v \in U$. Let $\gamma : [0, a] \to \exp_p(U)$ be a geodesic with $\gamma(0) = p$. Then, no point of γ is conjugate to $\gamma(0)$ along γ.*

Proof. This follows from the implicit function theorem and Proposition 3.2.5. □

Remark 3.2.7. If the geodesic $\gamma : [0, a] \to M$ is sufficiently short, there do not exist conjugate points. Indeed, if $\gamma([0, a])$ is contained in a normal neighborhood of p, then by Corollary 3.2.6 no point $\gamma(t)$ is conjugate to $\gamma(0)$ along γ.

Lemma 3.2.8. *Let $\gamma : [0, a] \to M$ be a geodesic, $t_1 \neq t_2$, $v \in T_{\gamma(t_1)}M$ and $w \in T_{\gamma(t_1)}M$. If $\gamma(t_1)$ and $\gamma(t_2)$ are not conjugate along γ, there exists a unique Jacobi field J along γ such that $J(t_1) = v$ and $J(t_2) = w$.*

Proof. Consider the linear map $\phi : \mathcal{J} \to T_{\gamma(t_1)}M \oplus T_{\gamma(t_2)}M$, given by $\phi(J) = (J(t_1), J(t_2))$. By definition of conjugate points, ϕ is injective. Since $\mathcal{J}$ and $T_{\gamma(t_1)}M \oplus T_{\gamma(t_2)}M$ are both vector spaces of dimension $2n$, ϕ is surjective. □

3.3 The Rauch Comparison Theorem

In Proposition 3.1.3, we have seen that Jacobi fields describe the difference between a geodesic and an infinitesimally close geodesic. Based on this interpretation, Rauch's theorem relates the sectional curvature of a Riemannian manifold to the rate at which its geodesics spread apart.

In dimension two, Rauch's theorem is an easy consequence of the classical theorem of Sturm on ordinary differential equations, cf. [Car92, p. 211]. In dimension higher than two, the proof is much less simple. A presentation of the theorem was given for the first time in 1951 by H. Rauch [Rau51].

Lemma 3.3.1.

(i) Let $f : [0, a] \to \mathbb{R}$ be a piecewise smooth function with $f(0) = 0$. Then there exists a piecewise smooth function $g : [0, a] \to \mathbb{R}$, such that $g(0) = f'(0)$ and $f(t) = t \cdot g(t)$ on $[0, a]$.

(ii) Let M be a Riemannian manifold, $\gamma : [0, a] \to M$ be a geodesic and V be a piecewise smooth vector field along γ with $|V(0)| = 0$. Then, there exists a piecewise smooth vector field A with $A(0) = V'(0)$ and $A(t) = t \cdot V(t)$ on $[0, a]$.

Proof. We have $f(t) = f(t) - f(0) = \int_0^t f'(s)ds = t \cdot \int_0^1 f'(ts)ds$. The function $g : [0, a] \to \mathbb{R}$, defined by $g(t) := \int_0^1 f'(ts)ds$ satisfies the conditions of (i).

To prove (ii), let V be as above. By choosing a frame field $\{E_1, \ldots, E_n\}$ along γ, we can write

$$V(t) = \sum_{i=1}^{n} f_i(t)E_i(t)$$

for piecewise smooth functions f_i. Clearly, $f_i(0) = 0$ for $1 \leq i \leq n$. By (i), there exist piecewise smooth g_i such that $g_i(t) = t \cdot f_i(t)$ and $g_i(0) = f_i'(0)$. The vector field A, defined by

$$A(t) = \sum_{i=1}^{n} g_i(t)E_i(t)$$

complies with the requirements of (ii). □

Let now $\gamma : [0, a] \to M$ be a fixed geodesic and V be a piecewise smooth vector field along γ. For $t_0 \in [0, a]$, we call

$$I_{t_0}(V, V) = \int_0^{t_0} [\langle V', V' \rangle - \langle R(\gamma', V)\gamma', V \rangle] \, dt \tag{3.3}$$

the index form of V along γ.

Lemma 3.3.2 (Index-Lemma). *Let $\gamma : [0, a] \to M$ be a geodesic without conjugate points to $\gamma(0)$ along γ on $(0, a]$. Let V be a piecewise smooth vector field along γ and let J be a Jacobi field along γ with $\langle V, \gamma' \rangle = 0$. Furthermore, let $|J(0)| = |V(0)| = 0$ and $J(t_0) = V(t_0)$ for a t_0 in $[0, a]$. Then,*

$$I_{t_0}(J, J) \leq I_{t_0}(V, V)$$

and equality occurs if and only if $J \equiv V$ on $[0, t_0]$.

Proof. Let $\mathcal{J}^{\perp}$ be the vector space of Jacobi fields along γ with $J(0) = 0$ and $\langle J, \gamma' \rangle = 0$. By Proposition 3.1.7, the dimension of $\mathcal{J}^{\perp}$ equals $n-1$ where $n = \dim(M)$. Let $J_1, \ldots, J_{n-1}$ be a basis of $\mathcal{J}^{\perp}$. Then we can write $J = \sum_i \alpha_i J_i$ for constants α_i, $1 \leq i \leq n$. Furthermore, we denote the orthogonal complement of $\gamma'(t)$ in $T_{\gamma(t)}M$ by $T_{\gamma(t)}M^{\perp}$. Since there are no conjugate points to $\gamma(0)$ on γ, $J_i(t) \neq 0$ for $(0, a]$ and the same holds for every nontrivial linear combination of the J_i. Therefore, the vectors $J_1(t), \ldots, J_{n-1}(t)$ are linearly independent for $t \in (0, a]$ and they form a basis of $T_{\gamma(t)}M^{\perp}$. For $t \neq 0$ we can write

$$V(t) = \sum_i f_i(t) J_i(t).$$

Since the Jacobi fields J_i are smooth by Proposition 3.1.4, the functions f_i are piecewise smooth functions on $(0, a]$. Next, we want to extend the f_i to piecewise smooth functions on $[0, a]$.

By Lemma 3.3.1 (ii), $J_i(t) = tA_i(t)$ for smooth vector fields A_i. It is clear that $J_i'(0) \neq 0$, otherwise $J_i \equiv 0$. The same holds for every nontrivial linear combination of the vectors $J_i(0)$, so the vectors $A_i(0) = J_i'(0)$ form a basis of $T_{\gamma(0)}M^{\perp}$. Since $J_i(t) = tA_i(t)$ and the $J_i(t)$ are a basis of $T_{\gamma(t)}M^{\perp}$, the $A_i(t)$ also form a basis of $T_{\gamma(t)}M^{\perp}$ for $t \in (0, a]$. Therefore we can write

$$V(t) = \sum_i g_i(t) A_i(t)$$

for piecewise smooth functions g_i on $[0, a]$.

Since $V(0) = 0$, $g_i(0) = 0$ for $i \in \{1, \ldots, n-1\}$. By Lemma 3.3.1 (i), $g_i(t) = t \cdot h_i(t)$ for piecewise smooth functions h_i on $[0, a]$. We have

$$V(t) = \sum_i f_i(t) J_i(t) = \sum_i g_i(t) A_i(t) = \sum_i t \cdot h_i(t) A_i(t) = \sum_i h_i(t) J_i(t)$$

and therefore $f_i(t) = h_i(t)$ for $t \neq 0$, which proves the claim.

Let now $0 = s_0 < s_1 < \ldots < s_k = t_0$ be a subdivision of $[0, t_0]$ so that the f_i are smooth on each subinterval $[s_{l-1}, s_l]$, $1 \leq l \leq k$. We show that on these subintervals

$$\langle V', V' \rangle - \langle R(\gamma', V)\gamma', V \rangle = \langle \sum_i f_i' J_i, \sum_j f_j' J_j \rangle + \frac{d}{dt} \langle \sum_i f_i J_i, \sum_j f_j J_j' \rangle \quad (3.4)$$

By the Jacobi equation, we have

$$R(\gamma', V)\gamma' = R(\gamma', \sum_i f_i J_i)\gamma' = \sum_i f_i R(\gamma', J_i)\gamma' = -\sum_i f_i J_i''.$$

Consider the left hand side of (3.4):

$$\begin{aligned}
&\langle V', V' \rangle - \langle R(\gamma', V)\gamma', V \rangle \\
&= \langle \sum_i f_i' J_i + \sum_i f_i J_i', \sum_j f_j' J_j + \sum_j f_j J_j' \rangle - \langle R(\gamma', V)\gamma', V \rangle \\
&= \langle \sum_i f_i' J_i, \sum_j f_j' J_j \rangle + \langle \sum_i f_i' J_i, \sum_j f_j J_j' \rangle + \langle \sum_i f_i J_i', \sum_j f_j' J_j \rangle \\
&+ \langle \sum_i f_i J_i', \sum_j f_j J_j' \rangle + \langle \sum_i f_i J_i'', \sum_j f_j J_j \rangle
\end{aligned}$$

Expansion of the right hand side leads to:

$$\begin{aligned}
&\frac{d}{dt}\langle\sum_i f_iJ_i, \sum_j f_jJ_j'\rangle\\
&= \langle\sum_i f_i'J_i + \sum_i f_iJ_i', \sum_j f_jJ_j'\rangle + \langle\sum_i f_iJ_i, \sum_j f_j'J_j' + \sum_j f_jJ_j''\rangle\\
&= \langle\sum_i f_i'J_i, \sum_j f_jJ_j'\rangle + \langle\sum_i f_iJ_i', \sum_j f_jJ_j'\rangle\\
&+ \langle\sum_i f_iJ_i, \sum_j f_jJ_j''\rangle + \langle\sum_i f_iJ_i, \sum_j f_j'J_j'\rangle.
\end{aligned}$$

Therefore, (3.4) is equivalent to

$$\langle\sum_i f_iJ_i', \sum_j f_j'J_j\rangle = \langle\sum_i f_iJ_i, \sum_j f_j'J_j'\rangle. \tag{3.5}$$

To show (3.5), we write

$$h(t) = \langle J_i', J_j\rangle - \langle J_i, J_j'\rangle.$$

Since $h(0) = 0$ and

$$\begin{aligned}
h'(t) &= \langle J_i'', J_j\rangle + \langle J_i', J_j'\rangle - \langle J_i', J_j'\rangle - \langle J_i, J_j''\rangle\\
&= -\langle R(\gamma', J_i)\gamma', J_j\rangle + \langle J_i, R(\gamma', J_j)\gamma'\rangle = 0
\end{aligned}$$

we conclude $h \equiv 0$. Using bilinearity yields (3.5), which proves (3.4).

Now we can apply (3.4) to V and we obtain

$$\begin{aligned}
I_{t_0}(V,V) &= \sum_{l=1}^{k}\int_{s_{l-1}}^{s_l}[\langle V', V'\rangle - \langle R(\gamma', V)\gamma', V\rangle]dt\\
&= \sum_{l=1}^{k}\int_{s_{l-1}}^{s_l}[\langle\sum_i f_i'J_i, \sum_j f_j'J_j\rangle + \frac{d}{dt}\langle\sum_i f_iJ_i, \sum_j f_jJ_j'\rangle]dt \qquad (3.6)\\
&= \int_0^{t_0}\langle\sum_i f_i'J_i, \sum_j f_j'J_j\rangle dt + \langle\sum_i f_iJ_i, \sum_j f_jJ_j'\rangle(t_0)
\end{aligned}$$

For $J(t) = \sum_i \alpha_iJ_i(t)$, we obtain from (3.6)

$$I_{t_0}(J,J) = \langle\sum_i \alpha_iJ_i, \sum_j \alpha_jJ_j'\rangle(t_0)$$

Since $J(t_0) = V(t_0)$, we have $f_i(t_0) = \alpha_i$ and therefore

$$I_{t_0}(V,V) = I_{t_0}(J,J) + \int_0^{t_0}|\sum_i f_i'J_i|^2dt. \tag{3.7}$$

It follows from (3.7) that $I_{t_0}(V,V) \geq I_{t_0}(J,J)$, which proves the first part of the proposition.

If $I_{t_0}(V,V) = I_{t_0}(J,J)$, then by (3.7), $\sum_i f_i'J_i = 0$ on each interval (s_{l-1}, s_l). Since the $J_i(t)$ are linearly independent for $t \neq 0$, $f_i'(t) = 0$ on (s_{l-1}, s_l). Therefore, the f_i are piecewise constant and continuous, hence constant on $[0, t_0]$. Since $f_i(t_0) = \alpha_i$, we have $f_i(t) = \alpha_i$ and therefore $V \equiv J$ on $[0, t_0]$, as claimed. □

We are now in a position to prove Rauch's theorem. We denote by M^n a manifold of dimension n.

Theorem 3.3.3 (Rauch). *Let $\gamma : [0,a] \to M^n$ and $\tilde{\gamma} : [0,a] \to \tilde{M}^{n+k}$, $k \geq 0$ be two geodesics in M and $\tilde{M}$, respectively, with $|\gamma'(t)| = |\tilde{\gamma}'(t)|$. Let J and $\tilde{J}$ be Jacobi fields along γ and $\tilde{\gamma}$, respectively, which satisfy*

$$|J(0)| = |\tilde{J}(0)| = 0, \quad |J'(0)| = |\tilde{J}'(0)|, \quad \langle J'(0), \gamma'(0)\rangle = \langle \tilde{J}'(0), \tilde{\gamma}'(0)\rangle. \tag{3.8}$$

Suppose that $\tilde{\gamma}$ does not contain points which are conjugate to $\tilde{\gamma}(0)$ along $\tilde{\gamma}$ and for all $t \in [0,a]$, $x \in T_{\gamma(t)}M$ and $\tilde{x} \in T_{\tilde{\gamma}(t)}\tilde{M}$, we have

$$\tilde{K}_{\tilde{\gamma}(t)}(\tilde{x}, \tilde{\gamma}'(t)) \geq K_{\gamma(t)}(x, \gamma'(t)), \tag{3.9}$$

if defined. (By $K_p(x,y)$ and $\tilde{K}_{\tilde{p}}(\tilde{x},\tilde{y})$, we denote the sectional curvature of M and $\tilde{M}$ at points $p \in M$ and $\tilde{p} \in \tilde{M}$, respectively.) Then,

$$|\tilde{J}(t)| \leq |J(t)| \qquad \forall t \in [0,a] \tag{3.10}$$

In addition, if there exits $t_0 \in (0,a]$ such that $|\tilde{J}(t_0)| = |J(t_0)|$, then

$$\tilde{K}_{\tilde{\gamma}(t)}(\tilde{J}(t), \tilde{\gamma}'(t)) = K_{\gamma(t)}(J(t), \gamma'(t)) \qquad \forall t \in (0, t_0], \tag{3.11}$$

if defined.

Proof. From Lemma 3.1.5 and (3.8), we obtain

$$|tan J| = |\langle J, \gamma'\rangle\gamma'| = |\langle J'(0), \gamma'(0)\rangle| t = |\langle \tilde{J}'(0), \tilde{\gamma}'(0)\rangle| t = |\langle \tilde{J}, \tilde{\gamma}'\rangle\tilde{\gamma}'| = |tan \tilde{J}|.$$

Therefore, to prove $|\tilde{J}(t)| \leq |J(t)|$, it suffices to show that $|nor\,\tilde{J}(t)| \leq |nor\,J(t)|$. By Proposition 3.1.6, $nor\,J$ and $nor\,\tilde{J}$ are Jacobi fields. So to prove the theorem, it suffices to prove the theorem for all Jacobi fields orthogonal to γ' and $\tilde{\gamma}'$, respectively. Thus, we suppose that

$$\langle J, \gamma'\rangle = \langle \tilde{J}, \tilde{\gamma}'\rangle = 0.$$

If $|J'(0)| = |\tilde{J}'(0)| = 0$, then $|J| = |\tilde{J}| = 0$ and the inequality $|\tilde{J}(t)| \leq |J(t)|$ is satisfied trivially. Otherwise, we define two real-valued functions $v(t) = |J(t)|^2$ and $\tilde{v}(t) = |\tilde{J}(t)|^2$. Since no point of $\tilde{\gamma}$ is conjugate to $\tilde{\gamma}(0)$ along $\tilde{\gamma}$, $\tilde{v}(t) \neq 0$ for $t \neq 0$ and the expression $\frac{v(t)}{\tilde{v}(t)}$ is well defined on $(0,a]$. We have $v(0) = 0 = \tilde{v}(0)$. A straightforward calculation shows

$$v'(0) = \tilde{v}'(0) = 0, \qquad v''(0) = 2\langle J'(0), J'(0)\rangle, \qquad \tilde{v}''(0) = 2\langle \tilde{J}'(0), \tilde{J}'(0)\rangle.$$

From L'Hospitals rule,

$$\lim_{t\to 0} \frac{v(t)}{\tilde{v}(t)} = \lim_{t\to 0} \frac{v''(t)}{\tilde{v}''(t)} = \frac{|J'(0)|^2}{|\tilde{J}'(0)|^2} \overset{(3.8)}{=} 1.$$

To show $|\tilde{J}(t)| \leq |J(t)|$, it therefore suffices to prove $\frac{d}{dt}(\frac{v(t)}{\tilde{v}(t)}) \geq 0$ for $t \in (0,a]$. By the quotient rule, this is equivalent to proving

$$v'(t)\tilde{v}(t) \geq v(t)\tilde{v}'(t). \tag{3.12}$$

Let $t_0 \in (0, a]$. If $v(t_0) = 0$, then

$$v'(t_0) = 2\langle J'(t_0), J(t_0)\rangle = 0,$$

and (3.12) is satisfied trivially. Suppose, therefore, that $v(t_0) \neq 0$. We already know that $\tilde{v}(t_0) \neq 0$. Define

$$U(t) = \frac{1}{\sqrt{v(t_0)}} J(t), \qquad \tilde{U}(t) = \frac{1}{\sqrt{\tilde{v}(t_0)}} \tilde{J}(t),$$

and observe that $|U(t_0))| = |\tilde{U}(t_0)| = 1$. Then by the Jacobi equation,

$$\begin{aligned}
\frac{v'(t_0)}{v(t_0)} &= \frac{2\langle J'(t_0), J(t_0)\rangle}{\langle J(t_0), J(t_0)\rangle} = 2\langle U'(t_0), U(t_0)\rangle = \langle U, U\rangle'(t_0) \\
&= \int_0^{t_0} \langle U, U\rangle'' \, dt = 2\int_0^{t_0} [\langle U', U'\rangle + \langle U'', U\rangle] dt \\
&= 2\int_0^{t_0} [\langle U', U'\rangle - \langle R(\gamma', U)\gamma', U\rangle] \, dt = 2I_{t_0}(U, U).
\end{aligned}$$

Analogously, one shows that

$$\frac{\tilde{v}'(t_0)}{\tilde{v}(t_0)} = 2I_{t_0}(\tilde{U}, \tilde{U}).$$

Therefore, to prove (3.12), it suffices to show $I_{t_0}(\tilde{U}, \tilde{U}) \leq I_{t_0}(U, U)$.

Let $\{E_1, \ldots, E_n\}$ and $\{\tilde{E}_1, \ldots, \tilde{E}_{n+k}\}$ be frame fields along γ and $\tilde{\gamma}$, respectively, such that

$$\begin{aligned}
E_1(t) &= \frac{\gamma'(t)}{|\gamma'(t)|}, \qquad E_2(t_0) = U(t_0) \\
\tilde{E}_1(t) &= \frac{\tilde{\gamma}'(t)}{|\tilde{\gamma}'(t)|}, \qquad \tilde{E}_2(t_0) = \tilde{U}(t_0).
\end{aligned}$$

We map each vector field $V(t) = \sum_i f_i(t)E_i(t)$ along γ to a vector field $\phi V(t)$ along $\tilde{\gamma}$, given by

$$\phi V(t) = \sum_{i=1}^{n} f_i(t)\tilde{E}_i(t).$$

The map ϕ satisfies

$$\begin{aligned}
\langle \phi V, \phi V\rangle &= \sum_i f_i^2 = \langle V, V\rangle, \\
\langle \phi V', \phi V'\rangle &= \sum_i (f_i')^2 = \langle V', V'\rangle, \\
\langle V, \gamma'\rangle &= f_1|\gamma'| = f_1|\tilde{\gamma}'| = \langle \phi V, \tilde{\gamma}\rangle.
\end{aligned}$$

With the inequality (3.9), we obtain

$$\begin{aligned}
&I_{t_0}(\phi(U), \phi(U)) \\
&\quad = \int_0^{t_0} \left[\langle \phi U', \phi U'\rangle - |\phi U(t)|^2 |\tilde{\gamma}'(t)|^2 \tilde{K}_{\tilde{\gamma}(t)}(\phi U(t), \tilde{\gamma}'(t))\right] dt \\
&\quad \leq \int_0^{t_0} \left[\langle U', U'\rangle - |U(t)|^2 |\gamma'(t)|^2 K_{\gamma(t)}(U(t), \gamma'(t))\right] dt = I_{t_0}(U, U)
\end{aligned} \tag{3.13}$$

The vector fields $\tilde{U}$ and ϕU along $\tilde{\gamma}$ are orthogonal to $\tilde{\gamma}'$ and satisfy $|\tilde{U}(0)| = |\phi U(0)| = 0$ and $\tilde{U}(t_0) = \phi U(t_0)$. Since $\tilde{U}$ is a Jacobi field, we obtain from Lemma 3.3.2, that

$$I_{t_0}(\tilde{U}, \tilde{U}) \leq I_{t_0}(\phi U, \phi U) \leq I_{t_0}(U, U), \tag{3.14}$$

so we have proven (3.12) and therefore the inequality (3.10) in the theorem.

Suppose that $|J(t_0)| = |\tilde{J}(t_0)|$ for some $t_0 \in (0, a]$. Then, $|tan J(t_0)| = |tan \tilde{J}(t_0)|$ by the calculation at the beginning of the proof and therefore, $|nor J(t_0)| = |nor \tilde{J}(t_0)|$. Note that

$$K_{\gamma(t)}(J(t), \gamma'(t)) = K_{\gamma(t)}(nor J(t), \gamma'(t))$$
$$\tilde{K}_{\tilde{\gamma}(t)}(\tilde{J}(t), \gamma'(t)) = \tilde{K}_{\tilde{\gamma}(t)}(nor \tilde{J}(t), \tilde{\gamma}'(t))$$

and the expressions are only defined if $|nor J|, |nor \tilde{J}| \neq 0$. Thus, it suffices to prove (3.11) for $J, \tilde{J}$ which satisfy

$$\langle J, \gamma' \rangle = \langle \tilde{J}, \tilde{\gamma}' \rangle = 0, \qquad J, \tilde{J} \neq 0.$$

By assumption, $\frac{v(0)}{\tilde{v}(0)} = \frac{v(t_0)}{\tilde{v}(t_0)}$. We have already shown that $\frac{d}{dt}\frac{v(t)}{\tilde{v}(t)} \geq 0$ for all $t \in (0, a]$. Therefore, $\frac{d}{dt}\frac{v(t)}{\tilde{v}(t)} = 0$ for $t \in (0, t_0]$ which is equivalent to

$$v'(t)\tilde{v}(t) = v(t)\tilde{v}'(t), \qquad t \in (0, t_0].$$

Since $\frac{v'(t_0)}{v(t_0)} = 2I_{t_0}(U, U)$ and $\frac{\tilde{v}'(t_0)}{\tilde{v}(t_0)} = 2I_{t_0}(\tilde{U}, \tilde{U})$, equality occurs in (3.14), so

$$I_{t_0}(\tilde{U}, \tilde{U}) = I_{t_0}(\phi U, \phi U) = I_{t_0}(U, U).$$

From the first equality, it follows from Lemma 3.3.2, that $\tilde{U} \equiv \phi(U)$ on $[0, t_0]$. By the second equality, the integrands of (3.13) are equal, hence for $t \in (0, t_0]$

$$\begin{aligned} K_{\gamma(t)}(J(t), \gamma'(t)) &= K_{\gamma(t)}(U(t), \gamma'(t)) \\ &= \tilde{K}_{\tilde{\gamma}(t)}(\phi U(t), \tilde{\gamma}'(t)) = \tilde{K}_{\tilde{\gamma}(t)}(\tilde{U}(t), \tilde{\gamma}'(t)) = \tilde{K}_{\tilde{\gamma}(t)}(\tilde{J}(t), \tilde{\gamma}'(t)), \end{aligned}$$

so we have proven Rauch's theorem. □

3.4 Conclusions from Rauch's Theorem

Proposition 3.4.1. *Let M be a Riemannian manifold whose sectional curvature K satisfies*

$$0 < L \leq K \leq H$$

for constants L and H. Let $\gamma : [0, a] \to M$ be a normalized geodesic. Let $\gamma(t_0)$ be the first point (with respect to the parameter t) which is conjugate to $\gamma(0)$ along γ. Then t_0 satisfies

$$\frac{\pi}{\sqrt{H}} \leq t_0 \leq \frac{\pi}{\sqrt{L}}.$$

Proof. First we show $t_0 \geq \frac{\pi}{\sqrt{H}}$. We compare M with a sphere of the same dimension and constant curvature H, which we denote by M_H. Let J be a Jacobi field along γ with $|J(0)| = |J(t_0)| = 0$. By Remark 3.2.2, $\langle J, \gamma' \rangle = 0$. Let $\gamma_1 :$

$[0,a] \to M_H$ be a normalized geodesic in M_H and $J_1(0)$ be a Jacobi field along γ_1 which satisfies $J_1(0)=0$, $|J_1'(0)| = |J'(0)|$ and $\langle J_1'(0), \gamma_1'(0)\rangle = \langle J'(0), \gamma'(0)\rangle$. We know from Example 3.2.4 that $\tilde{\gamma}$ does not contain conjugate points in the interval $(0, \frac{\pi}{\sqrt{H}})$. With Theorem 3.3.3, it follows that $|J(t)| \geq |J_1(t)| > 0$ for $t \in (0, \frac{\pi}{\sqrt{H}})$, so $t_0 \geq \frac{\pi}{\sqrt{H}}$ and the first inequality of the proposition is proven.

To show $t_0 \leq \frac{\pi}{\sqrt{L}}$, we compare M with a sphere of the same dimension and constant curvature L, denoted by M_L. Let now J_2 be a Jacobi field along a normalized geodesic $\gamma_2 : [0,a] \to M_L$ satisfying $|J_2(0)| = 0$, $|J_2'(0)| = |J'(0)|$ and $\langle J_2'(0), \gamma_2'(0)\rangle = \langle J'(0), \gamma'(0)\rangle$. By Proposition 3.1.7 (iii), $\langle J_2, \gamma_2'\rangle = 0$. Since γ does not contain conjugate points on $(0, t_0)$, we can apply Theorem 3.3.3 and we obtain $|J_2(t)| \geq |J(t)| > 0$ for $t \in (0, t_0)$. By Example 3.2.4, $|J_2(\frac{\pi}{\sqrt{L}})| = 0$ from which $t_0 \leq \frac{\pi}{\sqrt{L}}$ follows. This proves the second inequality. □

Lemma 3.4.2. *Let $\alpha : [0,a] \to M$ be a smooth curve and V be a vector field along α. If $\alpha(t) = p$ for all $t \in [0,a]$, then*

$$\frac{D}{dt}V(t) = \frac{\partial}{\partial t}V(t),$$

where $\frac{D}{dt}$ denotes the covariant derivative along the curve α and $\frac{\partial}{\partial t}$ is the usual derivative in the vector space T_pM.

Proof. Observe first that if $\alpha(t) = p$ for all t, then $V(t)$ is a curve in T_pM, so the derivative $\frac{\partial}{\partial t}V(t)$ makes sense. Let U be a neighborhood of p and $(\varphi = (x^1, \ldots, x^n), U)$ be local coordinates around p. In these coordinates, the covariant derivative can be written as

$$V'(t) = \sum_k \left(\frac{\partial V^k}{\partial t} + \sum_{i,j} \Gamma_{ij}^k(\alpha(t)) \frac{\partial (x^j \circ \alpha)}{\partial t}(t) V^i(t) \right) \partial_k|_{\alpha(t)},$$

cf. [O'N83, p. 66]. In the case $\alpha(t) = p$, the terms $\frac{\partial(x^j \circ \alpha)}{\partial t}(t)$ vanish, so

$$\frac{D}{dt}V(t) = V'(t) = \sum_k \frac{\partial V^k}{\partial t} \partial_k|_p = \frac{\partial}{\partial t}V(t)$$

which finishes the proof. □

Proposition 3.4.3. *Let M^n and $\tilde{M}^n$ be Riemannian manifolds with sectional curvature $K, \tilde{K}$, respectively. Suppose that $\inf \tilde{K} \geq \sup K$. Let $p \in M$, $\tilde{p} \in \tilde{M}$ and fix a linear isometry $i : T_pM \to T_{\tilde{p}}\tilde{M}$. Let $r > 0$ such that $\exp_p$ is defined on $B_r(0) \subset T_pM$ and the restriction $\exp_{\tilde{p}}|_{\tilde{B}_r(0)}$ is a local diffeomorphism. Let $\alpha : [0,a] \to \exp_p(B_r(0)) \subset M$ be a piecewise smooth curve which can be lifted to a curve in T_pM, i.e. there exists a piecewise smooth curve $\bar{\alpha}$ in T_pM such that $exp_p \circ \bar{\alpha} = \alpha$. Let $\tilde{\alpha} : [0,a] \to \exp_{\tilde{p}}(\tilde{B}_r(0)) \subset \tilde{M}$ be given by*

$$\tilde{\alpha}(s) = \exp_{\tilde{p}} \circ i \circ \bar{\alpha}(s), \qquad s \in [0,a].$$

Then, $L(\alpha) \geq L(\tilde{\alpha})$.

Proof. Since we can apply this Proposition to each smooth segment of α, we may assume α and $\bar{\alpha}$ to be smooth. Consider the geodesic variation

$$f:[0,1]\times[0,a]\to M, \qquad f(t,s)=\exp_p(t\bar{\alpha}(s)).$$

We set $\gamma_s(t)=f(t,s)$. By Proposition 3.1.3, the vector field $\frac{\partial f}{\partial s}(t,s)=J_s(t)$ is a Jacobi field along the geodesic γ_s which satisfies

$$J_s(0)=\frac{\partial f}{\partial s}(0,s)=T_{t\bar{\alpha}(s)}\exp_p(t\bar{\alpha}'(s))|_{t=0}=0$$

and

$$J_s(1)=\frac{\partial f}{\partial s}(1,s)=\frac{\partial}{\partial s}\alpha(s)=\alpha'(s).$$

Furthermore,

$$J_s'(0)=\frac{D}{dt}\frac{\partial f}{\partial s}(0,s)=\frac{D}{ds}\frac{\partial f}{\partial t}(0,s)=\frac{D}{ds}T_0\exp_p(\bar{\alpha}(s))=\frac{D}{ds}\bar{\alpha}(s)\overset{3.4.2}{=}\frac{\partial}{\partial s}\bar{\alpha}(s).$$

We now define a geodesic variation in $\tilde{M}$ by

$$\tilde{f}:[0,1]\times[0,a]\to M, \qquad \tilde{f}(t,s)=\exp_{\tilde{p}}(t\cdot i(\bar{\alpha}(s))).$$

Let $\tilde{\gamma}_s(t)=\tilde{f}(t,s)$ and consider the Jacobi field $\tilde{J}_s(t)=\frac{\partial\tilde{f}}{\partial s}(t,s)$ along the geodesic $\tilde{\gamma}_s$. As above, one shows that

$$|\tilde{J}_s(0)|=0, \qquad \tilde{J}_s(1)=\tilde{\alpha}'(s), \qquad \tilde{J}_s'(0)=i(\bar{\alpha}'(s)).$$

Because i is an isometry,

$$|J_s'(0)|=|\bar{\alpha}'(s)|=|i(\bar{\alpha}'(s))|=|\tilde{J}_s'(0)|$$

and

$$\langle\tilde{J}_s'(0),\tilde{\gamma}_s'(0)\rangle=\langle i(\bar{\alpha}'(s)),i(\gamma_s'(0))\rangle=\langle\bar{\alpha}'(s),\gamma_s'(0)\rangle=\langle J_s'(0),\gamma_s'(0)\rangle,$$

so the conditions (3.8) of Theorem 3.3.3 are satisfied. Since $\exp_{\tilde{p}}|_{\tilde{B}_r(0)}$ is a local diffeomorphism, the geodesics $\tilde{\gamma}_s$ do not contain conjugate points on $(0,1]$ by Corollary 3.2.6. Since $\inf\tilde{K}\geq\sup K$, condition (3.9) is satisfied. Moreover, observe that $|\gamma_s'|=|\tilde{\gamma}_s'|$, since

$$|\gamma_s'(0)|=|T_0\exp_p(\bar{\alpha}(s))|=|\bar{\alpha}(s)|=|i(\bar{\alpha}(s))|=|T_0\exp_{\tilde{p}}(i(\bar{\alpha}(s)))|=|\tilde{\gamma}_s'(0)|.$$

Therefore, we can apply Theorem 3.3.3 and we obtain $|\tilde{J}_s(t)|\leq|J_s(t)|$ for all $t\in[0,1]$. In particular,

$$|\tilde{\alpha}'(s)|=|\tilde{J}_s(1)|\leq|J_s(1)|=|\alpha'(s)|.$$

By this inequality,

$$L(\alpha)=\int_0^a|\alpha'(s)|ds\geq\int_0^a|\tilde{\alpha}'(s)|ds=L(\tilde{\alpha}).$$

□

A far reaching global generalization of Rauch's Theorem is the Theorem of Toponogov, which states the following:

Theorem 3.4.4 (Toponogov). *Let M be a complete Riemannian manifold whose sectional curvature K satisfies $K \geq H$ for a real number H. Let $\gamma_1 : [0, L_1] \to M$ and $\gamma_2 : [0, L_2] \to M$ be normalized geodesics in M with $\gamma_1(0) = \gamma_2(0)$. Assume that γ_1 is minimizing and, if $H > 0$, $L_2 \leq \frac{\pi}{\sqrt{H}}$. Let M_H^2 be the space form of dimension 2 and constant sectional curvature H and let $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ be two normalized geodesics in M_H^2 such that $\tilde{\gamma}_1(0) = \tilde{\gamma}_2(0)$, $L(\tilde{\gamma}_i) = L(\gamma_i)$, $i = 1, 2$ and $\measuredangle(\tilde{\gamma}_1'(0), \tilde{\gamma}_2'(0)) = \measuredangle(\gamma_1'(0), \gamma_2'(0))$. Then*

$$d(\gamma_1(L_1), \gamma_2(L_2)) \leq d(\tilde{\gamma}_1'(L_1), \tilde{\gamma}_2'(L_2)).$$

Proof. A proof using Rauch's theorem is given in [CE75]. A different approach, using the Hessian of the Riemannian distance function, is given in [Mey89] and [Pet06]. □

Toponogov's theorem turned out to be an extremely powerful tool in Riemannian geometry. Some of its applications can be found in [Mey89].

3.5 Manifolds of nonpositive curvature

In this section, we prove the Theorem of Cartan-Hadamard, which gives an important property of complete manifolds of nonpositive sectional curvature. The proofs are taken from Chapter 11 of [Lee97]. For simplicity, all geodesics in this section are assumed to be normalized.

Lemma 3.5.1. *Suppose $\tilde{M}$ and M are connected Riemannian manifolds, with $\tilde{M}$ complete and $\pi : \tilde{M} \to M$ a local isometry. Then M is complete and π is a covering map.*

Proof. We first show that M is complete. Let p be in the image of π and γ be a geodesic with initial point p and initial velocity $v \in T_pM$. Choose $\tilde{p}$ in $\pi^{-1}(p)$ and let $\tilde{v} = (T_{\tilde{p}}\pi)^{-1}(v)$. Let $\tilde{\gamma}$ be the geodesic with initial point $\tilde{p}$ and initial velocity $\tilde{v}$. Because $\tilde{M}$ is complete, $\tilde{\gamma}$ is defined for all time. Since π is a local isometry, $\pi \circ \tilde{\gamma}$ is a geodesic and since by construction $\pi(\tilde{p}) = p$ and $T_{\tilde{p}}\pi(\tilde{v}) = v$, we have $\pi \circ \tilde{\gamma} = \gamma$. Thus γ is defined for all time, so M is complete. For the rest of the proof, we call a geodesic $\tilde{\gamma}$, constructed as above, a lift of γ.

Next we show that π is surjective. Choose $\tilde{p} \in \tilde{M}$, let $p = \pi(\tilde{p})$ and let $q \in M$ be arbitrary. Because M is connected and complete, there exists a geodesic γ joining p to q. Let $\tilde{\gamma}$ be a lift of γ starting at $\tilde{p}$ and $r = d(p, q)$, then $\pi(\tilde{\gamma}(r)) = \gamma(r) = q$, so q is in the image of π.

To show that π is a covering map, we need to show that each point of M has a neighborhood U which is evenly covered. Let $p \in M$ and $\epsilon > 0$ so that $\exp_p : B_\epsilon(0) \to B_\epsilon(p)$ is a diffeomorphism. We show that $U := B_\epsilon(p)$ is evenly covered. Let $\pi^{-1}(p) = \{\tilde{p}_\lambda\}_{\lambda \in \Lambda}$ and $\tilde{U}_\lambda = B_\epsilon(\tilde{p}_\lambda)$. The first step is to show that the various sets $\tilde{U}_\lambda$ are disjoint. For any $\lambda \neq \mu$, there exists, by completeness of $\tilde{M}$, a minimizing geodesic $\tilde{\gamma}$ joining $\tilde{p}_\lambda$ to $\tilde{p}_\mu$. The projected curve $\gamma := \pi \circ \tilde{\gamma}$ is a geodesic from p to p. Since U is a normal neighborhood of p, γ must leave U and re-enter it, and thus must have length at least 2ϵ.

Therefore, $d(\tilde{p}_\lambda, \tilde{p}_\mu) = L(\tilde{\gamma}) = L(\gamma) \geq 2\epsilon$ and from the triangle inequality, it follows that $\tilde{U}_\lambda \cap \tilde{U}_\mu = \emptyset$.

The next step is to show that $\pi^{-1}(U) = \bigcup_{\lambda\in\Lambda} \tilde{U}_\lambda$. Since π is an isometry, it maps each $\tilde{U}_\lambda$ into U. Thus we need only to show that $\pi^{-1}(U) \subset \bigcup_{\lambda\in\Lambda} \tilde{U}_\lambda$. Let $\tilde{q} \in \pi^{-1}(U)$. This means that $q := \pi(\tilde{q}) \in U$ so there is a minimizing geodesic γ in U from q to p and $r = d(p,q) = L(\gamma) < \epsilon$. Let $\tilde{\gamma}$ be the lift of γ starting at $\tilde{q}$, then $\pi(\tilde{\gamma}(r)) = \gamma(r) = p$, so $\tilde{\gamma}(r) = \tilde{p}_\lambda$ for some $\lambda \in \Lambda$. Therefore, $d(\tilde{q}, \tilde{p}_\lambda) \leq L(\tilde{\gamma}) = r < \epsilon$, so $q \in \tilde{U}_\lambda$.

It remains only to show that $\pi : \tilde{U}_\lambda \to U$ is a diffeomorphism for each λ. It is certainly a local diffeomorphism (because π is). It is bijective since its inverse can be constructed explicitly: it is the map sending each radial geodesic starting at p to its lift starting at $\tilde{p}_\lambda$. This completes the proof. □

Theorem 3.5.2 (Cartan-Hadamard). *Let M^n be a complete Riemannian manifold whose sectional curvature K satisfies $K \leq 0$. Then, for each $p \in M$, $\exp_p : T_pM \to M$ is a covering map. In particular, the universal covering space of M is diffeomorphic to $\mathbb{R}^n$. If M is simply connected, then M itself is diffeomorphic to $\mathbb{R}^n$.*

Proof. We first show that for each $p \in M$, $\exp_p : T_pM \to M$ is a local diffeomorphism. Let γ be a geodesic starting from p. Since $K \leq \frac{1}{n}$ for all $n \in \mathbb{N}$, no conjugate points occur on γ, since by Proposition 3.4.1, such a point $\gamma(t_0)$ should satisfy $\pi\sqrt{n} \leq t_0$ for all $n \in \mathbb{N}$, which is not possible. (One should note that in the proof of Proposition 3.4.1, we only used the condition $K \leq H$ to show $\frac{\pi}{\sqrt{H}} \leq t_0$.) Therefore, by Proposition 3.2.5, $\exp_p$ is a local diffeomorphism at each $v \in T_pM$.

Thus, we can define a Riemannian metric $\tilde{g}$ on T_pM as $\tilde{g} = \exp_p^* g$, so $\exp_p : (T_pM, \tilde{g}) \to (M, g)$ is a local isometry. Each straight line $\tilde{\gamma} : t \mapsto tv$, $t \in \mathbb{R}$, $v \in T_pM$ is a geodesic in $(T_pM, \tilde{g})$, since $\exp_p \circ \tilde{\gamma}$ is a geodesic in (M, g). Therefore, $(T_pM, \tilde{g})$ is complete, since the geodesics starting from the origin are defined for all time. From Lemma 3.5.1, it follows that $\exp_p$ is a covering map. The remaining statements of the theorem follow immediately from uniqueness of the universal covering space. □

Definition 3.5.3. A complete, simply-connected Riemannian manifold with nonpositive sectional curvature is called a Cartan-Hadamard manifold.

Theorem 3.5.2 shows that there are topological restrictions on which manifolds can carry metrics such that the sectional curvature is nonpositive. It shows for example that the sectional curvature of each metric on the unit sphere S^n is positive somewhere. Similarly, the same holds for the real projective space $\mathbb{RP}^n$, since S^n is the universal covering space of $\mathbb{RP}^n$.

Remark 3.5.4. With some more algebraic topology, it is possible to prove Preissman's theorem, which states the following: If the sectional curvature of a compact Riemannian manifold satisfies $K < 0$, then each nontrivial abelian subgroup of the fundamental group $\pi_1(M)$ is infinite cyclic. This shows, for example, that the torus, whose fundamental group is $\mathbb{Z} \times \mathbb{Z}$, can not carry a metric of negative curvature. More generally, if two compact manifolds M_1 and M_2 have nontrivial abelian fundamental groups, then each metric on the product manifold $M_1 \times M_2$ gives a sectional curvature which is nonnegative somewhere. For more details, see Chapter 12 of [Car92].

Chapter 4

The Morse index theorem

In this chapter, we define for any geodesic γ a symmetric bilinear form, which is closely related to the formula for the second variation of energy of γ. The Index Theorem states that the index of this form is equal to the number of conjugate points along γ, each counted with its multiplicity. We follow Chapter 11 of [Car92].

4.1 Properties of the Index Form

Let $\gamma : [0, a] \to M$ be a geodesic. We denote by $\mathcal{V}(0, a) = \mathcal{V}$ the space of all piecewise smooth vector fields along γ that vanish at the endpoints of γ, i.e. $|V(0)| = |V(a)| = 0$.

The index form of γ is the quadratic form associated to the symmetric bilinear form I_a, defined by

$$I_a(V, W) = \int_0^a \langle V', W' \rangle - \langle R(\gamma', V)\gamma', W \rangle dt, \qquad V, W \in \mathcal{V},$$

c.f. (3.3).

Remark 4.1.1. By Equation (2.4), we have

$$E''_x(0) = 2I_a(V, V)$$

for each proper variation x of γ with variational vector field V.

Before we start to prove properties of the bilinear form I_a, we recall some basic definitions concerning symmetric bilinear forms. Let $B : V \times V \to \mathbb{R}$ be a symmetric bilinear form, defined on the real vector space V. The form B is called positive definite if $B(v, v) > 0$ for all $v \in V$, $v \neq 0$. Analogously, B is negative definite if $B(v, v) < 0$ for all $v \in V$, $v \neq 0$. The index of B is defined as

$$\text{ind}(B) := \max \{\dim(U) | U \text{ subspace of } V \text{ and } B|_U \text{ negative definite}\}.$$

If $B(v, w) = 0$ for all $w \in V$ implies $v = 0$, then B is called non-degenerate. Otherwise, B is degenerate. We call the vector space N, defined as

$$N := \{v \in V | B(v, w) = 0 \text{ for all } w \in V\},$$

the null space of B. The dimension of N is called the nullity of B.

Proposition 4.1.2. *A vector field $V \in \mathcal{V}$ belongs to the null space of I_a if and only if V is a Jacobi field along γ.*

Proof. Let $V, W \in \mathcal{V}$ and $0 = t_0 < \ldots < t_k = a$ be such that $V|_{[t_{i-1},t_i]}$ is smooth. By integration by parts, we obtain the following expression for I_a:

$$I_a(V,W) = -\int_0^a \langle V'' + R(\gamma', V)\gamma', W\rangle dt - \sum_{i=1}^{k-1} \langle \Delta V'(t_i), W(t_i)\rangle. \tag{4.1}$$

cf. [Car92], Chapter 9, Proposition 2.8. If V is a Jacobi field, it is unbroken and both terms of (4.1) vanish, so $I_a(V,W) = 0$ for all $W \in \mathcal{V}$. Suppose now that V belongs to the null space of I_a, so $I_a(V,W) = 0$ for all $W \in \mathcal{V}$. Let $f : [0,a] \to \mathbb{R}$ be a smooth function such that $f(t) > 0$ for $t \neq t_i$ and $f(t_i) = 0$, $i = 1, \ldots, k$. Define $W \in \mathcal{V}$ by

$$W(t) = f(t)(V'' + R(\gamma', V)\gamma').$$

Then

$$0 = I_a(V,W) = -\int_0^a f(t)|V'' + R(\gamma', V)\gamma'|^2 dt.$$

Therefore, the integrand is zero, hence $V|_{[t_{i-1},t_i]}$ is a Jacobi field for $i = 1, \ldots k$. Now we show that $\Delta V'(t_i) = 0$. Let $U \in \mathcal{V}$ be such that $U(t_i) = \Delta V'(t_i)$. Then

$$0 = I_a(V,U) = -\sum_{i=1}^{k-1} |\Delta V'(t_i)|^2,$$

which proves the claim. Since $V'' = R(V, \gamma')\gamma'$ on all points where V is smooth and by continuity of the right hand side, this equation is valid for all $t \in [0,a]$. Therefore, V is a Jacobi field. ☐

Corollary 4.1.3. *The form I_a is degenerate if and only if $\gamma(0)$ and $\gamma(a)$ are conjugate along γ. In this case, the nullity of I_a equals the multiplicity of $\gamma(a)$ as a conjugate point.*

Proof. By Proposition 4.1.2, the null space of I_a consists of the Jacobi fields which vanish in $\gamma(0)$ and $\gamma(a)$. Then the assertions follow by definition. ☐

Recall that a neighborhood $U \subset M$ is called totally normal if it is a normal neighborhood of each of its points, that is, for any points $p, q \in U$, there exists a unique minimizing geodesic joining p and q. In [Car92], it is shown that each point $p \in M$ has a totally normal neighborhood.

Since $\gamma([0,a])$ is compact we can choose a subdivision $0 = t_0 < t_1 < \cdots < t_k = a$ of $[0,a]$ such that each segment $\gamma|_{[t_{i-1},t_i]}$ is contained in a totally normal neighborhood. Therefore, by Remark 3.2.7, $\gamma|_{[t_{i-1},t_i]}$ is minimizing and does not contain conjugate points. We call such a subdivision normal and we will fix such a subdivision for the rest of this section.

We denote by $\mathcal{V}^-(0,a) = \mathcal{V}^-$ the subspace of the fields V such that $V|_{[t_{i-1},t_i]}$ is a Jacobi field for $i = 1, \ldots, k$. Since $\gamma(t_{i-1})$ and $\gamma(t_i)$ are not conjugate along γ it follows from Lemma 3.2.8, that $V|_{[t_{i-1},t_i]}$ is uniquely determined by the values $V(t_{i-1})$ and $V(t_i)$. Therefore, the map

$$\begin{aligned} \psi : \mathcal{V}^- &\to T_{\gamma(t_1)}M \oplus \cdots \oplus T_{\gamma(t_{k-1})}M, \\ V &\mapsto (V(t_1), \ldots, V(t_{k-1})) \end{aligned} \tag{4.2}$$

is a linear isomorphism. In particular, $\mathcal{V}^-$ is finite dimensional. Let $\mathcal{V}^+(0,a) = \mathcal{V}^+$ be the subspace of $\mathcal{V}$ consisting of the vector fields W such that $|W(t_1)| = \ldots = |W(t_{k-1})| = 0$. Observe that $\mathcal{V}^+ \cap \mathcal{V}^- = \{0\}$.

Proposition 4.1.4. *The space $\mathcal{V}$ is a direct sum $\mathcal{V} = \mathcal{V}^+ \oplus \mathcal{V}^-$. The subspaces $\mathcal{V}^+$ and $\mathcal{V}^-$ are orthogonal with respect to I_γ. In addition, $I_a|_{\mathcal{V}^+}$ is positive definite.*

Proof. Let $V \in \mathcal{V}$. We choose a vector field $W \in \mathcal{V}^-$, which is given by $W(t_j) = V(t_j)$, $i = 1, \ldots k$. By the isomorphism in (4.2), W is uniquely determined. Then $V - W \in \mathcal{V}^+$ and therefore $\mathcal{V} = \mathcal{V}^+ + \mathcal{V}^-$. Since $\mathcal{V}^+ \cap \mathcal{V}^- = \{0\}$, it follows that $\mathcal{V} = \mathcal{V}^+ \oplus \mathcal{V}^-$. In addition, if $X \in \mathcal{V}^-$ and $Y \in \mathcal{V}^+$, we have

$$\begin{aligned} I_a(X,Y) &= -\int_0^a \langle X'' + R(\gamma', X)\gamma', Y\rangle dt - \sum_{j=1}^{k-1} \langle \Delta X'(t_j), Y(t_j)\rangle \\ &= -\int_0^a \langle 0, Y\rangle dt - \sum_{j=1}^{k-1} \langle \Delta X'(t_j), 0\rangle = 0, \end{aligned}$$

which shows that $\mathcal{V}^+$ and $\mathcal{V}^-$ are orthogonal with respect to I_a.

It remains to show that $I_a|_{\mathcal{V}^+}$ is positive definite. Let $V \in \mathcal{V}^+$. Let $x : (-\epsilon, \epsilon) \times [0,a] \to M$ be a variation of γ with variational vector field V. Since $|V(t_i)| = 0$, it is possible, by Proposition 2.1.5, to choose x in such a way that it fixes the points $\gamma(t_i)$. We know that each segment $\gamma^i := \gamma|_{[t_{i-1}, t_i]}$ is minimizing. Therefore, by Lemma 2.1.3, $x^i(s,.) := x(s,.)|_{[t_{i-1},t_i]}$ satisfies

$$(t_i - t_{i-1})E(x^i(s,.)) \geq L(x^i(s,.))^2 \geq L(\gamma^i)^2 = (t_i - t_{i-1})E(\gamma^i),$$

which yields

$$E(x(s,.)) = \sum_i E(x^i(s,.)) \geq \sum_i E(\gamma^i) = E(\gamma).$$

Thus we obtain $I_a(V,V) = 2E_x''(0) \geq 0$. Suppose now that $I_a(V,V) = 0$ for a vector field $V \in \mathcal{V}^+$. We are going to show that $V = 0$. If $W \in \mathcal{V}^-$, $I_a(V,W) = 0$ by orthogonality . If $W \in \mathcal{V}^+$, consider the inequality

$$0 \leq I_a(V + cW, V + cW) = 2cI_a(V,W) + c^2 I_a(W,W),$$

valid for all $c \in \mathbb{R}$. This says that there exist real numbers $A \geq 0$ and B such that $Ac^2 + 2Bc \geq 0$ for all $c \in \mathbb{R}$. This is only possible if $B = 0$. Therefore, $I_a(V,W) = 0$ for all $W \in \mathcal{V}^+$. Thus, since $\mathcal{V} = \mathcal{V}^+ \oplus \mathcal{V}^-$, V belongs to the null space of I_a. By Proposition 4.1.2, V is a Jacobi field, hence $V \in \mathcal{V}^-$. Therefore, $V \in \mathcal{V}^+ \cap \mathcal{V}^- = \{0\}$. □

Corollary 4.1.5. *The index of I_a equals the index of I_a restricted to $\mathcal{V}^-$. In particular, the index of I_a is finite. The same is true for the nullity of I_a.*

Proof. By Proposition 4.1.4, $\operatorname{ind}(I_a) = \operatorname{ind}(I_a|_{\mathcal{V}^+}) + \operatorname{ind}(I_a|_{\mathcal{V}^-})$ and $\operatorname{ind}(I_a|_{\mathcal{V}^+}) = 0$. By Proposition 4.1.2, the null space of I_a is a subspace of $\mathcal{V}^-$ which proves the assertion about the nullity of I_a. □

4.2 The proof of the Index theorem

Theorem 4.2.1 (Morse index theorem). *indexMorse Index theoremThe index of I_a is finite and equals the number of points $\gamma(t)$, $0 < t < a$, conjugate to $\gamma(0)$ along γ, each counted with its multiplicity.*

Proof. We introduce the following notation. For $t \in [0, a]$ we denote by γ^t the restriction of γ to the interval $[0, t]$. We denote the corresponding index form, defined on $\mathcal{V}(0, t)$, by I_t. The index of I_t is denoted by $i(t)$. So we have defined a function $i : [0, a] \to \mathbb{N}$, whose behavior we wish to study. Recall that we have chosen a subdivision such that $\gamma|_{[t_{j-1},t_j]}$ is a minimizing geodesic. In particular, for $t \leq t_1$, each proper variation $x^t : (-\epsilon, \epsilon) \times [0, t] \to M$ of γ^t satisfies $L(x^t_s) \geq L(\gamma^t)$ for all $s \in (-\epsilon, \epsilon)$ and ϵ sufficiently small. Let $V \in \mathcal{V}(0, t)$ be arbitrary and x^t be a variation of γ^t with variational vector field V. Then, by Lemma 2.1.3, $tE(x^t_s) \geq L(x^t_s)^2 \geq L(\gamma^t)^2 = tE(\gamma^t)$, which shows that $I_t(V, V) = 2E''_x(0) \geq 0$.

From this we conclude that $i(t) = 0$ for small t. The function $i(t)$ is monotonically increasing. By definition of $i(t)$, there exists a subspace $\mathcal{U} \subset \mathcal{V}(0, t)$ of dimension $i(t)$ such that I_t is negative definite on $\mathcal{U}$. For $\bar{t} \geq t$, we can extend every element $V \in \mathcal{U}$ to an element $\bar{V} \in \mathcal{V}(0, \bar{t})$ by defining $\bar{V} = 0$ on $[t, \bar{t}]$. Clearly, $I_t(V, V) = I_{\bar{t}}(\bar{V}, \bar{V})$. From the definition of the index, it follows that $i(\bar{t}) \geq i(t)$.

Now we are going to explore other properties of $i(t)$. First, observe that $i(t)$ does not depend on the choice of normal subdivision of $[0, a]$. Thus, we can choose the subdivision in such a way that for a parameter $t \in (0, a)$ fixed, $t \in (t_{j-1}, t_j)$ for a $j \in \{1, \ldots, n\}$. By Corollary 4.1.5, $i(t)$ equals the index of I_a restricted to $\mathcal{V}^-(0, t)$. We denote this restriction again by I_t. By Lemma 3.2.8, the map

$$\begin{aligned} \psi_t : \mathcal{V}^-(0, t) &\to T_{\gamma(t_1)}M \oplus \cdots \oplus T_{\gamma(t_{j-1})}M =: S_j \\ V &\mapsto (V(t_1), \ldots, V(t_{j-1})) \end{aligned}$$

is an isomorphism for each $t \in (t_{j-1}, t_j)$, since $\gamma(t_{i-1})$ and $\gamma(t_i)$ are not conjugate along γ, $1 \leq i \leq j - 1$. Therefore, we may consider the forms I_t as a family of symmetric bilinear forms on a fixed space S_j: For $V, W \in S_j$ fixed, let $I_t(V, W) = I_t(\psi_t^{-1}(V), \psi_t^{-1}(W))$. By (4.1), we have

$$I_t(V, W) = \sum_{i=1}^{j-1} \langle \Delta\psi_t^{-1}(V)'(t_i), \psi_t^{-1}(W)(t_i) \rangle.$$

Varying $t \in (t_{j-1}, t_j)$ leaves the vectors $\Delta\psi_t^{-1}(V)'(t_i)$ fixed for $i \in \{1, \ldots, j-2\}$. Since $\psi_t^{-1}(V)|_{[t_{j-1},t]}$ is the solution of a second order ordinary differential equation, it depends differentiable on the boundary conditions. Therefore, $\Delta\psi_t^{-1}(V)'(t_{j-1})$ depends continuously on t, so I_t, considered as a bilinear form on S_j, depends continuously on t.

From the following two lemmas, we obtain the required information about $i(t)$.

Lemma 4.2.2. *If $\epsilon > 0$ is sufficiently small, $i(t - \epsilon) = i(t)$*

Proof of Lemma 4.2.2. Since $i(t)$ is monotonically increasing, $i(t - \epsilon) \leq i(t)$. On the other hand, let $\bar{S} \subset S_j$ be a subspace with $\dim \bar{S} = i(t)$ such that I_t is negative definite on $\bar{S}$. We claim the following: $I_{t-\epsilon}$ is still negative definite on $\bar{S}$ if ϵ is small enough. By bilinearity, it suffices to prove this for all $v \in \bar{S}$ with $|v| = 1$. Let E be the unit sphere in $\bar{S}$ and

$$\lambda_t = \max \{I_t(v, v) | v \in E\}.$$

The maximum exists, since E is compact. Clearly, $\lambda_t < 0$. By continuity of I_t in t, $\lambda_{t-\epsilon} < 0$ for $\epsilon > 0$ small enough. This proves that $I_{t-\epsilon}$ is negative definite on $\bar{S}$. Therefore $i(t - \epsilon) \geq i(t)$, which proves the Lemma. □

Let now $d(t)$ be the nullity of I_t. By Corollary 4.1.3, $d(t) > 0$ if and only if $\gamma(t)$ is conjugate to $\gamma(0)$ along γ.

Lemma 4.2.3. *If $\epsilon > 0$ is sufficiently small, $i(t + \epsilon) = i(t) + d(t)$.*

Proof of Lemma 4.2.3. We first show that $i(t + \epsilon) \leq i(t) + d(t)$. Since $\dim S_j = n(j - 1)$, the form I_t is positive definite on a subspace of dimension $n(j - 1) - i(t) - d(t)$. By continuity, $I_{t+\epsilon}$ is positive definite on this subspace for ϵ small enough. This follows from an analogous argument as in the proof of Lemma 4.2.2. Therefore,

$$i(t + \epsilon) \leq n(j - 1) - [n(j - 1) - i(t) - d(t)] = i(t) + d(t)$$

Now we are going to prove the reverse inequality. Let $\bar{S}$ be as in the proof of Lemma 4.2.2. Since I_t is negative definite on $\bar{S}$, $I_{t+\epsilon}$ is also negative definite on $\bar{S}$ for ϵ small. This follows again from the argument used in the proof of Lemma 4.2.2 We want to prove the following claim: If $I_{\tilde{t}}(V, V) = 0$ for an element $V \in S_j$, $V \neq 0$, then $I_{\tilde{t}+\epsilon}(V, V) < 0$.

By Proposition 4.1.2, the corresponding vector field $\psi_{\tilde{t}}^{-1}(V)$ is an unbroken Jacobi field on $[0, \tilde{t}]$. By Remark 3.2.2, $\langle \psi_{\tilde{t}}^{-1}(V), \gamma' \rangle = 0$. Let $t_0 \in (t_{j-1}, t_j)$ and $V_{t_0} := \psi_{t_0}^{-1}(V)$.

We are going to show that $\langle V_{t_0}, \gamma' \rangle = 0$. Observe that the segments $V_{t_0}|_{[0,t_{j-1}]}$ and $V_{t_0}|_{[t_{j-1},t_0]}$ are Jacobi fields. The vector field V_{t_0} does not vanish at t_{j-1} since otherwise $|V_{t_0}(t_{j-1})| = |V_{t_0}(t_0)| = 0$. In this case, $\psi_{\tilde{t}}^{-1}(V) \equiv 0$, since $\gamma(t_{j-1})$ and $\gamma(t_0)$ are not conjugate along γ. By Lemma 3.1.5,

$$\langle V_{t_0}, \gamma' \rangle(t) = \langle V_{t_0}'(0), \gamma'(0) \rangle t + \langle V_{t_0}(0), \gamma'(0) \rangle, \qquad t \in [0, t_{j-1}]$$

and

$$\langle V_{t_0}, \gamma' \rangle(t) = \langle V_{t_0}'(t_{j-1}), \gamma'(t_{j-1}) \rangle (t - t_{j-1}) + \langle V_{t_0}(t_{j-1}), \gamma'(t_{j-1}) \rangle, \; t \in [t_{j-1}, t_0].$$

Therefore, the function $f : [0, t_0] \to \mathbb{R}$, $f(t) = \langle V_{t_0}, \gamma' \rangle(t)$ is a piecewise linear function with t_{j-1} as its only breaking point. Since V_{t_0} vanishes at its endpoints, $f(0) = f(t_0) = 0$. Since $\psi_{\tilde{t}}^{-1}(V)$ is orthogonal to γ', $f(t_{j-1}) = \langle V_{t_0}, \gamma' \rangle(t_{j-1}) = \langle \psi_{\tilde{t}}^{-1}(V), \gamma' \rangle(t_{j-1}) = 0$ and therefore, $\langle V_{t_0}, \gamma' \rangle = f = 0$.

To prove the claim, it suffices to show that

$$I_{t_0}(V_{t_0}, V_{t_0}) > I_{t_0+\epsilon}(V_{t_0+\epsilon}, V_{t_0+\epsilon}),$$

where $V_{t_0+\epsilon} = \psi_{t_0+\epsilon}^{-1}(V)$. By definition, V_{t_0} and $V_{t_0+\epsilon}$ correspond to the same vector $V \in S_j$. We define a vector field W_{t_0} along $\gamma|_{[0,t_0+\epsilon]}$ by

$$W_{t_0}(t) = V_{t_0}(t), \qquad t \in [0, t_0],$$
$$W_{t_0}(t) = 0, \qquad t \in [t_0, t_0 + \epsilon].$$

Since V_{t_0} is orthogonal to γ', the same holds for W_{t_0}. By the same argument as used for V_{t_0}, one shows that $V_{t_0+\epsilon}$ is also orthogonal to γ'.

We define a geodesic $\tilde{\gamma} : [0, t_0 + \epsilon - t_{j-1}] \to M$ by $\tilde{\gamma}(t) = \gamma(t_0 + \epsilon - t)$ and denote by $\tilde{I}_t$ the index form corresponding to $\tilde{\gamma}|_{[0,t]}$ for $t \in [0, t_0 + \epsilon - t_{j-1}]$. To each vector field U along γ, we associate a vector field $\tilde{U}$ along $\tilde{\gamma}$, which is given by $\tilde{U}(t) = U(t_0 + \epsilon - t)$, $t \in [0, t_0 + \epsilon - t_{j-1}]$. For $U \in \mathcal{V}(0, t_0 + \epsilon)$, we have

$$\begin{aligned}
I_{t_0+\epsilon}(U, U) &= \int_0^{t_0+\epsilon} \langle U', U' \rangle - \langle R(\gamma', U)\gamma', U \rangle dt \\
&= I_{t_{j-1}}(U, U) + \int_{t_{j-1}}^{t_0+\epsilon} \langle U', U' \rangle - \langle R(\gamma', U)\gamma', U \rangle dt \\
&= I_{t_{j-1}}(U, U) + \int_0^{t_0+\epsilon-t_{j-1}} \langle \tilde{U}', \tilde{U}' \rangle - \langle R(\tilde{\gamma}', \tilde{U})\tilde{\gamma}', \tilde{U} \rangle dt \\
&= I_{t_{j-1}}(U, U) + \tilde{I}_{t_0+\epsilon-t_{j-1}}(\tilde{U}, \tilde{U}).
\end{aligned}$$

Since $V_{t_0+\epsilon}$ and W_{t_0} coincide on $[0, t_{j-1}]$,

$$I_{t_{j-1}}(V_{t_0+\epsilon}, V_{t_0+\epsilon}) = I_{t_{j-1}}(W_{t_0}, W_{t_0}).$$

We have $\tilde{V}_{t_0+\epsilon}(0) = \tilde{W}_{t_0}(0) = 0$, $\tilde{V}_{t_0+\epsilon}(t_0 + \epsilon - t_{j-1}) = \tilde{W}_{t_0}(t_0 + \epsilon - t_{j-1})$ and $V_{t_0+\epsilon} \neq W_{t_0}$. Since the restriction of $V_{t_0+\epsilon}$ onto $[t_{j-1}, t_0 + \epsilon]$ is a Jacobi field along γ, $\tilde{V}_{t_0+\epsilon}$ is a Jacobi field along $\tilde{\gamma}$. Moreover, $\tilde{V}_{t_0+\epsilon}$ and $\tilde{W}_{t_0}$ are orthogonal to $\tilde{\gamma}'$, since $V_{t_0+\epsilon}$ and W_{t_0} are orthogonal to γ'. It follows from Lemma 3.3.2 that

$$\tilde{I}_{t_0+\epsilon-t_{j-1}}(\tilde{V}_{t_0+\epsilon}, \tilde{V}_{t_0+\epsilon}) < \tilde{I}_{t_0+\epsilon-t_{j-1}}(\tilde{W}_{t_0}, \tilde{W}_{t_0}).$$

Together with the equalities above, we obtain

$$I_{t_0}(V_{t_0}, V_{t_0}) = I_{t_0+\epsilon}(W_{t_0}, W_{t_0}) > I_{t_0+\epsilon}(V_{t_0+\epsilon}, V_{t_0+\epsilon}),$$

which proves the claim.

We obtain the following: if I_t is negative definite on a subspace $\bar{S} \subset S_j$, then $I_{t+\epsilon}$ is negative definite on the direct sum of $\bar{S}$ with the null space of I_t: Let $\bar{S}_0$ be the null space of I_t and let $v = v_0 + v_1 \in \bar{S}_0 \oplus \bar{S}$, $v \neq 0$. Then $I_t(v, v) = I_t(v_1, v_1) \leq 0$ ($<$, if $v_1 \neq 0$). If $v_1 \neq 0$, $I_{t+\epsilon}(v, v) < 0$ by continuity. If $v_1 = 0$, $I_{t+\epsilon}(v, v) = I_{t+\epsilon}(v_0, v_0) < 0$ by our claim. Since $\bar{S}_0 \oplus \bar{S}$ is finite dimensional, we can choose ϵ independent of all $v \in \bar{S}_0 \oplus \bar{S}$. This follows from the same argument as used in the proof of Lemma 4.2.2. Therefore,

$$i(t + \epsilon) \geq i(t) + d(t),$$

which proves the lemma. □

From Lemmas 4.2.2 and 4.2.3, we can describe the function $i(t)$ as follows: It equals zero in a neighborhood of 0, it is continuous from the left and it has a jump discontinuity at each point $\gamma(t)$ which is conjugate to $\gamma(0)$ along γ. By Proposition 4.1.2 and Corollary 4.1.3, the step size equals the multiplicity of $\gamma(t)$ as a conjugate point.

Let γ be extended to an open interval $U \supset [0, a]$. By Lemma 4.2.2 and Lemma 4.2.3, there exists for each $\tilde{t} \in [0, a]$ an open interval $U_{\tilde{t}} \subset U$ such that

$i(t) = i(\tilde{t})$ for $t \in U_{\tilde{t}}$, $t \leq \tilde{t}$ and $i(t) = i(\tilde{t}) + d(\tilde{t})$ for $t \in U_{\tilde{t}}$, $t > \tilde{t}$. Choose values $\tilde{t}_1, \ldots, \tilde{t}_l \in [0, a]$ such that $U \subset \bigcup_{j=1}^{l} U_j$ for the corresponding intervals U_j. Then $\mathrm{ind} I_a = i(a) = \sum_{j=1}^{l} d(\tilde{t}_j)$ which shows that the index of I_a is finite. This finishes the proof of the Index theorem. □

Corollary 4.2.4. *The set of points $\gamma(t)$ which are conjugate to $\gamma(0)$ along γ is discrete.*

Corollary 4.2.5. *A geodesic γ is not minimizing after its first conjugate point.*

Proof. By Theorem 4.2.1, $\mathrm{ind} I_a > 0$ if there occurs at least one point which is conjugate to $\gamma(0)$ along γ. Choose a vector field $V \in \mathcal{V}$ such that $I_a(V, V) < 0$ and a proper variation $x : (-\epsilon, \epsilon) \times [0, a] \to M$ of γ with variational vector field V. Since γ is a geodesic, $E_x'(0) = 0$ by Proposition 2.1.7. By the formula for the second variation of energy (2.4), $E_x''(0) = 2I_a(V, V) < 0$. Therefore by Lemma 2.1.3,

$$L(x(s, .))^2 \leq aE(x(s, .)) = aE_x(s) < aE_x(0) = aE(\gamma) = L(\gamma)^2$$

for $s \neq 0$ small enough, hence γ is not minimizing. □

Chapter 5

Morse theory

In this chapter we establish some elementary results concerning the theory of smooth real-valued functions on a manifold M. We well apply these results in the next chapter. For more details, see [Mil63] and [Mat02]. For an elaboration of the topics in Sections 5.1 and 5.2, see [Hag05].

5.1 Non degenerate critical points

In what follows, we will extend the definitions of critical points and the Hessian of functions $f \in C^\infty(\mathbb{R}^n)$ to functions on manifolds. By the Morse lemma, we will see that the form of functions $f \in C^\infty(M)$ near non degenerate critical points of f is very simple. This section mainly follows §2 of [Mil63].

Definition 5.1.1. Let M be a manifold and $f : M \to \mathbb{R}$ a smooth function. A point $p \in M$ is called a critical point of f if $T_p f = 0$. A real number c is called a critical value of f if there exists a critical point in $f^{-1}(c)$. We denote the set of all critical points of f in M by C_f. If $p \in M \setminus C_f$, we call p a regular point of f.

Remark 5.1.2. Let $(\varphi = (x^1, \ldots, x^n), U)$ be a chart of M around p. Then p is a critical point of f if and only if

$$\frac{\partial f}{\partial x^1}(p) = \ldots = \frac{\partial f}{\partial x^n}(p) = 0$$

and if and only if for all smooth curves $\alpha : (-\epsilon, \epsilon) \to M$ with $\alpha(0) = p$, $(f \circ \alpha)'(0) = 0$.

Definition 5.1.3. Let f and M be as above and let p be a critical point of f. The Hessian of f at p is defined as

$$\mathcal{H}_f^p : T_pM \times T_pM \to \mathbb{R}, \qquad (v, w) \mapsto \tilde{v}_p(\tilde{w}(f)),$$

where $\tilde{v}$ and $\tilde{w}$ are local extensions of v and w to vector fields, respectively, so $\tilde{v}_p = v$ and $\tilde{w}_p = w$.

We check that $\mathcal{H}_f^p$ is well defined. By definition, this form is independent of the chosen extension of v. Since p is a critical point of f,

$$\tilde{v}_p(\tilde{w}(f)) = \tilde{w}_p(\tilde{v}(f)) + [\tilde{v}, \tilde{w}]_p(f) = \tilde{w}_p(\tilde{v}(f)),$$

which shows that $\mathcal{H}_f^p$ does not depend on the chosen extension of w either. It also follows that $\mathcal{H}_f^p$ is symmetric. Hence we obtain:

Lemma 5.1.4. *The Hessian of f at p is a symmetric bilinear form on T_pM.*

Lemma 5.1.5. *Let $(\varphi = (x^1, \ldots, x^n), U)$ be a chart of M around p. With respect to the basis $\{\frac{\partial}{\partial x^i}|_p\}$ of T_pM, the Hessian of f at p is represented by the matrix $(\frac{\partial^2 f}{\partial x^i \partial x^j}(p))$.*

Proof. Let $v = \frac{\partial}{\partial x^i}|_p$ and $w = \frac{\partial}{\partial x^j}|_p$. The vector fields $\tilde{v} = \frac{\partial}{\partial x^i}$ and $\tilde{w} = \frac{\partial}{\partial x^j}$ are local extensions of v and w, respectively and we have

$$\mathcal{H}_f^p(\frac{\partial}{\partial x^i}|_p, \frac{\partial}{\partial x^j}|_p) = v(\tilde{w}(f)) = \frac{\partial}{\partial x^i}|_p(\frac{\partial f}{\partial x^j}) = \frac{\partial^2 f}{\partial x^i \partial x^j}(p).$$

□

Lemma 5.1.6. *Let $\alpha : (-\epsilon, \epsilon) \to M$ be a smooth curve such that $\alpha(0) = p$. Then $(f \circ \alpha)''(0) = \mathcal{H}_f^p(\alpha'(0), \alpha'(0))$.*

Proof. Let $(\varphi = (x^1, \ldots, x^n), U)$ be a chart of M around p. Then $\alpha'(t) = \sum_{i=1}^n \frac{(x^i \circ \alpha)}{dt} \frac{\partial}{\partial x^i}$. A direct calculation yields

$$(f \circ \alpha)'(t) = (f \circ \varphi^{-1} \circ \varphi \circ \alpha)'(t) = \sum_{i=1}^n \frac{(x^i \circ \alpha)}{dt}(t) \frac{\partial f}{\partial x^i}(\alpha(t)).$$

Since p is critical, $\frac{\partial f}{\partial x^i}(p) = 0$. By differentiating once again and substituting $t = 0$, we get

$$\begin{aligned}(f \circ \alpha)''(0) &= (f \circ \varphi^{-1} \circ \varphi \circ \alpha)''(0) \\ &= \sum_{i,j=1}^n \frac{(x^i \circ \alpha)}{dt}|_{t=0} \frac{(x^j \circ \alpha)}{dt}|_{t=0} \frac{\partial^2 f}{\partial x^i \partial x^j}(p) \stackrel{5.1.5}{=} \mathcal{H}_f^p(\alpha'(0), \alpha'(0))\end{aligned}$$

which proves the Lemma. □

Definition 5.1.7. A critical point p of f is called non-degenerate if the Hessian of f at p is non-degenerate. Otherwise we call p degenerate. If p is a non-degenerate critical-point, we denote by $\mathrm{ind}_f(p)$ the index of the bilinear form $\mathcal{H}_f^p$ and call it the index of f at p.

Definition 5.1.8. A smooth function $f : M \to \mathbb{R}$ is called a Morse function if all critical points of f are non-degenerate.

The proof of the following theorem ist taken from [MT97], Theorem 12.6.

Theorem 5.1.9 (Morse-Lemma)**.** *Let $f \in C^\infty(M)$ and $p \in M$ be a non-degenerate critical point of f with $\mathrm{ind}_f(p) = \lambda$. Then there exists a local coordinate system $(\chi = (u^1, \ldots, u^n), W)$ around p with $\chi(p) = 0$, such that*

$$f \circ \chi^{-1}(u) = f(p) - (u^1)^2 - \ldots - (u^\lambda)^2 + (u^{\lambda+1})^2 + \ldots + (u^n)^2.$$

Proof. After replacing f by $f - f(p)$, we may assume that $f(p) = 0$. We choose local coordinates $(\varphi = (x^1, \dots, x^n), U)$ such that $\varphi(p) = 0$ and $\varphi(U)$ is a convex neighborhood of 0 in $\mathbb{R}^n$. Consider the function $\tilde{f} := f \circ \varphi^{-1} \in C^\infty(\varphi(U))$. We know that $\tilde{f}(0) = f(p) = 0$. We calculate

$$\tilde{f}(x) = \int_0^1 \frac{\partial}{\partial t}\tilde{f}(t \cdot x)dt = \int_0^1 \sum_{i=1}^n x^i \frac{\partial \tilde{f}}{\partial x^i}(t \cdot x)dt = \sum_{i=1}^n x^i \int_0^1 \frac{\partial \tilde{f}}{\partial x^i}(t \cdot x)dt.$$

Therefore, we can write $\tilde{f}$ in the form

$$\tilde{f}(x) = \sum_{i=1}^n x^i g_i(x), \qquad g_i(x) = \int_0^1 \frac{\partial \tilde{f}}{\partial x^i}(t \cdot x)dt.$$

Since $g_i(0) = \frac{\partial \tilde{f}}{\partial x^i}(0) = 0$ we may repeat this calculation to get

$$g_i(x) = \sum_{j=1}^n x^j g_{ij}(x), \qquad g_{ij}(x) = \int_0^1 \frac{\partial g_i}{\partial x^j}(t \cdot x)dt.$$

Summing up, we can write $\tilde{f}$ in the form $\tilde{f}(x) = \sum_{i,j=1}^n x^i x^j g_{ij}(x)$ where g_{ij} is smooth on $\varphi(U)$. By introducing $h_{ij} = \frac{1}{2}(g_{ij} + g_{ji})$, (h_{ij}) becomes a symmetric matrix of smooth functions on $\varphi(U)$, and

$$\tilde{f}(x) = \sum_{i,j=1}^n x_i x_j h_{ij}(x). \tag{5.1}$$

By differentiating (5.1) twice and substituting 0, we get

$$\frac{\partial^2 \tilde{f}}{\partial x^i \partial x^j}(0) = 2h_{ij}(0).$$

In particular, since p is non-degenerate, the matrix $(h_{ij}(0))$ is invertible.

Now we want to transform this expression into diagonal form. Suppose inductively that there exists a chart $(\varphi = (x^1, \dots, x^n), U)$ such that f can be written as

$$f \circ \varphi^{-1}(x) = \sum_{i=1}^{k-1} \delta_i (x^i)^2 + \sum_{i,j=k}^n x^i x^j h_{ij}(x), \qquad \delta_i = \pm 1 \tag{5.2}$$

for smooth functions h_{ij}. The matrix $(h_{ij}(0))$ is of the form

$$\begin{pmatrix} D & 0 \\ 0 & E \end{pmatrix},$$

where D is a $k \times k$ matrix of the form $\operatorname{diag}(\pm 1, \dots, \pm 1)$, and E a symmetric $(n - k + 1) \times (n - k + 1)$ matrix of smooth functions, which is invertible.

Before we proceed with the next step, we have to make sure that $h_{kk}(0) \neq 0$. If $h_{kk}(0) = 0$, we perform a linear change of variables in $x^k, \dots, x^n$ by a matrix A and denote the new variables by $\tilde{x}^k, \dots, \tilde{x}^n$. Then (5.2) transforms to

$$f \circ \varphi^{-1} \circ B^{-1}(\tilde{x}) = \sum_{i=1}^{k-1} \delta_i (x^i)^2 + \sum_{i,j=k}^n \tilde{x}^i \tilde{x}^j \tilde{h}_{ij}(\tilde{x}), \qquad \delta_i \pm 1, \tag{5.3}$$

for smooth functions $\tilde{h}_{ij}$ where

$$B = \begin{pmatrix} \mathrm{id}_{\mathbb{R}^k} & 0 \\ 0 & A \end{pmatrix}$$

and $\tilde{x} = (x^1, \ldots, x^k, \tilde{x}^k, \ldots, \tilde{x}^n)$. Consider the matrix $(\tilde{h}_{ij})$. Comparing (5.2) and (5.3), we get

$$\tilde{h}_{ij}(\tilde{x}) = (A^{-1})^t h_{ij}(B^{-1}(\tilde{x}))A^{-1}.$$

By Sylvesters Law of inertia, we can choose A in such a way that $\tilde{h}_{kk}(0) \neq 0$.

After such a linear transformation, we may assume that $h_{kk}(0) \neq 0$ and after restricting $\varphi(U)$, if necessary, we may assume that h_{kk} has constant sign $\delta_k = \pm 1$ on $\varphi(U)$. Set

$$q = \sqrt{|h_{kk}|} \in C^\infty(\varphi(U), \mathbb{R})$$

and introduce new coordinates:

$$\begin{aligned} y^k &= q(x)\left(x^k + \sum_{i=k+1}^{n} x^i \frac{h_{ik}(x)}{h_{kk}(x)}\right), \\ y^j &= x^j \qquad \text{for } j \neq k,\ 1 \leq i \leq n. \end{aligned}$$

The Jacobi determinant for y as a function of x at $x = 0$ is easily seen to be $\frac{\partial y^k}{\partial x^k}(0) = q(0) \neq 0$. The change of coordinates thus defines a local diffeomorphism ψ around 0. Consider the new coordinates $(\psi \circ \varphi = (y^1, \ldots, y^n), V)$ for a suitable subset $V \subseteq U$. Then we have for $y = \psi(x)$:

$$\begin{aligned} (f \circ \varphi^{-1} \circ \psi^{-1})(y) &= (f \circ \varphi^{-1})(x) \\ &= \sum_{i=1}^{k-1} \delta_i (x^i)^2 + (x^k)^2 h_{kk}(x) + 2x^k \sum_{j=k+1}^{n} x^j h_{jk}(x) + \sum_{i,j=k+1}^{n} x^i x^j h_{ij}(x) \\ &= \sum_{i=1}^{k-1} \delta_i (x^i)^2 + h_{kk}(x)\left(x^k + \sum_{i=k+1}^{n} x^i \frac{h_{ik}(x)}{h_{kk}(x)}\right)^2 \\ &\quad - h_{kk}(x)\left(\sum_{i=k+1}^{n} x^i \frac{h_{ik}(x)}{h_{kk}(x)}\right)^2 + \sum_{i,j=k+1}^{n} x^i x^j h_{ij}(x) \\ &= \sum_{i=1}^{k-1} \delta_i (y^i)^2 + h_{kk}(x)(\frac{y^k}{q(x)})^2 + \sum_{i,j=k+1}^{n} x^i x^j \bar{h}_{ij}(x) \\ &= \sum_{i=1}^{k} \delta_i (y^i)^2 + \sum_{i,j=k+1}^{n} y^i y^j \bar{h}_{ij}(\psi^{-1}(y)), \end{aligned}$$

where $\bar{h}_{ij} \in C^\infty(\psi(\varphi(V)))$. Repeating this procedure a finite number of times, we obtain a diagonal form:

$$(f \circ \chi^{-1})(u) = \sum_{i=1}^{n} \delta_i (u^i)^2, \qquad \delta_i = \pm 1.$$

Permutating coordinates leads to

$$f \circ \chi^{-1}(u) = -(u^1)^2 - \ldots - (u^\mu)^2 + (u^{\mu+1})^2 + \ldots + (u^n)^2$$

for an integer μ. Consider the matrix $\frac{\partial^2 f}{\partial u^i \partial u^j}(0) = \operatorname{diag}(-2, \ldots, -2, 2 \ldots, 2)$. Since by Lemma 5.1.5, this matrix is a representation of the bilinear form $\mathcal{H}_f^p$, we get that μ equals the index of $\mathcal{H}_f^p$, so $\mu = \lambda$ and we have finished the proof. □

Corollary 5.1.10. *Let $f \in C^\infty(M)$ be a Morse function. Then the set C_f of all critical points of f is discrete in M.*

Proof. Let p be a critical point of f. By definition, p is non-degenerate. By Theorem 5.1.9, there exists a chart $(\chi = (u^1, \ldots, u^n), W)$ around p with $\chi(p) = 0$ such that f takes the form

$$f \circ \chi^{-1}(u) = f(p) - (u^1)^2 - \ldots - (u^\lambda)^2 + (u^{\lambda+1})^2 + \ldots + (u^n)^2.$$

We obtain $\frac{\partial f}{\partial u^i}(q) = \pm 2u^i(q)$. Since q is a critical point if and only if $\frac{\partial f}{\partial u^i}(q) = 0$ for all i, we obtain that $\chi^{-1}(0) = p$ is the only critical point of f in W. Hence, p is isolated in C_f. □

Corollary 5.1.11. *Let M be a compact Riemannian manifold and $f \in C^\infty(M)$ be a Morse function. Then C_f is finite.*

Proof. This follows immediately from Corollary 5.1.10. □

5.2 Homotopy type and critical values

In this section we study subsets of a manifold M of the following form: For a smooth function $f : M \to \mathbb{R}$, $a, b \in \mathbb{R}$ and $a < b$, let

$$M_f^a = f^{-1}(-\infty, a], \qquad M_f^{[a,b]} = f^{-1}[a, b].$$

If it is clear which function we consider, we write $M^a = M_f^a$ and $M^{[a,b]} = M_f^{[a,b]}$.

Remark 5.2.1. If a is not a critical value of f, then M^a is a smooth manifold with boundary and the boundary is given by $f^{-1}(a)$.

We are going to investigate the behavior of M^a while varying a. It will turn out that the behavior of M^b for b near a critical value a is closely related to the indexes of f at the critical points in $f^{-1}(a)$. We follow §3 of [Mil63].

Recall the definition of the flow of a smooth vector field. Let $X \in \mathfrak{X}(M)$ and consider the initial value problem

$$X \circ F(t, p) = \frac{\partial}{\partial t} F(t, p) \qquad F(0, p) = p.$$

For each $(t, p) \in \mathbb{R} \times M$, there exists a unique solution of this initial value problem, defined on an open neighborhood U of (t, p) in $\mathbb{R} \times M$. The corresponding map is called the flow of X. We write $Fl_t^X(p) = Fl^X(t, p)$. For a fixed $p \in M$, $t \mapsto Fl_t^X(p)$ is defined on a maximal open interval (t_-^p, t_+^p). A smooth vector field X is called complete if for each $p \in M$, $t \mapsto Fl_t^X(p)$ is defined on $\mathbb{R}$.

Lemma 5.2.2. *Let* $X \in \mathfrak{X}(M)$ *and* $\mathrm{supp}(X)$ *be compact. Then* X *is complete.*

Proof. Suppose that X is not complete. Then there exists a $p \in M$ such that the curve $c_p : t \mapsto Fl_t^X(p)$ is not globally defined, so $(t_-^p, t_+^p) \subsetneq \mathbb{R}$. We may assume that $t_+^p < \infty$, the other case works analogously. Consider a sequence $t_n < t_+^p$, converging to t_+^p. We know that the sequence $c_p(t_n)$ leaves every compact subset of M, in particular $\mathrm{supp}(X)$. (cf. [Kun08], 2.5.17 (ii)) Consider the smallest $n_0 \in \mathbb{N}$, such that $c_p(t_{n_0}) \notin \mathrm{supp}(X)$. Then, $X \circ c_p(t_{n_0}) = 0$, therefore the curve c_p stays at $c_p(t_{n_0})$ for all $t \geq t_0$. Thus, the sequence $c_p(t_n)$ is contained in the compact set $\mathrm{supp}(X) \cup c_p(t_{n_0})$ which is a contradiction to the fact stated above. □

Definition 5.2.3. Let X be a topological space. A subspace A of X is a deformation retract of X if there exists a continuous map $r : [0,1] \times X \to X$ such that

$$r_0 = \mathrm{id}_X, \qquad r_t|_A = \mathrm{id}_A \ \forall t \in [0,1], \qquad r_1(X) \subseteq A,$$

where $r_t(x) = r(t,x)$. Such a map is called a deformation retraction of X onto A.

Example 5.2.4. The unit sphere S^{n-1} is a deformation retract of $\mathbb{R}^n \setminus \{0\}$. A deformation retraction is given by

$$r(t,x) = (1-t)x + t\frac{x}{|x|}.$$

Remark 5.2.5. If A is a deformation retract of X, then A and X are of the same homotopy type (cf. [Hal09], Definition 1.3.5 and Definition 1.3.10). This shows that A inherits many properties of X; e.g., the fundamental groups $\pi_1(A, r_1(x))$ and $\pi_1(X,x)$ are isomorphic for each $x \in X$ (cf. [Hal09], Satz 1.3.27).

Theorem 5.2.6. *Let* $f \in C^\infty(M)$ *and* $a < b$ *such that* $M^{[a,b]}$ *is compact and does not contain critical points of* f*. Then* M^a *is diffeomorphic to* M^b*. Furthermore,* M^a *is a deformation retract of* M^b*.*

Proof. Choose a Riemannian metric on M. Then there exist a unique vector field $\mathrm{grad}(f)$, called the gradient of f, such that $\langle X, \mathrm{grad}(f)\rangle = X(f)$. It vanishes exactly at the critical points of f. Let $\rho : M \to \mathbb{R}$, $\rho \geq 0$ be smooth such that

$$\rho|_{M^{[a,b]}} \equiv \frac{1}{\langle \mathrm{grad}(f), \mathrm{grad}(f)\rangle}|_{M^{[a,b]}}$$

and $\mathrm{supp}(\rho)$ is compact. Since $M^{[a,b]}$ does not contain any critical points such a ρ exists. Then the vector field $X := \rho \cdot \mathrm{grad}(f)$ is smooth and has compact support. By Lemma 5.2.2, X is complete. Consider the map $Fl^X : \mathbb{R} \times M \to M$ and, for a fixed point $q \in M$, the real-valued function $t \mapsto f(Fl_t^X(q))$. We have

$$\frac{\partial}{\partial t}(f \circ Fl_t^X(q)) = X(f)_{Fl_t^X(q)} = \langle X, \mathrm{grad}(f)\rangle_{Fl_t^X(q)} = \rho|\mathrm{grad}(f)|^2 \geq 0.$$

Moreover, if $Fl_t^X(q) \in M^{[a,b]}$, we obtain $\frac{\partial}{\partial t}(f \circ Fl_t^X(q)) = 1$. Thus, the function $t \mapsto f(Fl_t^X(q))$ is monotonically increasing. For $q \in M^{(a,b)}$, we have $Fl_t^X(q) \in M^{(a,b)}$ for small values of t. It follows that $\frac{\partial}{\partial t}(f \circ Fl_t^X(q)) = 1$ for small t which yields $f(Fl_t^X(q)) = f(q)+t$ for small t. Both sides coincide as long as $f(q)+t \in [a,b]$. If $f(q) = a$ or $f(q) = b$, we have, by monotonicity, $(f(Fl_t^X(q))) = f(q)+t$ for small $t > 0$ or $t < 0$, respectively. Again, this equation is valid as long as $f(q)+t \in [a,b]$. In summary, we have shown that, if $q \in M^{[a,b]}$,

$$f(Fl_t^X(q))|_{[a-f(q),b-f(q)]} = (f(q)+t)|_{[a-f(q),b-f(q)]}.$$

Thus, if $a \leq f(q) \leq b$, then

$$f(Fl_{a-b}^X(q)) \leq f(Fl_{a-f(q)}^X(q)) = f(q)+a-f(q) = a. \tag{5.4}$$

If $f(q) \leq a$,

$$f(Fl_{a-b}^X(q)) \leq f(Fl_0^X(q)) = f(q) \leq a. \tag{5.5}$$

Combining (5.4) and (5.5), we conclude $Fl_{a-b}^X(M^b) \subseteq M^a$. Analogously, we obtain $Fl_{b-a}^X(M^a) \subseteq M^b$. Since $(Fl_{b-a}^X)^{-1} = Fl_{a-b}^X$, M^a is diffeomorphic to M^b and we have proven the first assertion.

To prove that M^a is a deformation retract of M^b, consider the map $r : [0,1] \times M^b \to M^b$, defined as

$$r_t(q) = \begin{cases} q & f(q) \leq a, \\ Fl_{t(a-f(q))}^X(q) & a \leq f(q) \leq b. \end{cases}.$$

We show that r is a deformation retraction. This map is well defined in the case $f(q) = a$ and obviously continuous in both variables. It is clear that $r_0 = \mathrm{id}_{M^b}$ and $r_t|_{M^a} = \mathrm{id}_{M^a}$. It remains to show that $r_1(M^b) \subseteq M^a$. Let $q \in M^b$. If $f(q) \leq a$, this is clear, if $a \leq f(q) \leq b$, this follows from (5.4). Hence we have finished the proof. □

Theorem 5.2.7. *Let $f \in C^\infty(M)$ and $p \in M$ be a critical point of f with $\mathrm{ind}_f(p) = \lambda$. Assume there exists $\epsilon > 0$ such that $M^{[f(p)-\epsilon,f(p)+\epsilon]}$ is compact and contains no critical points except p. Then there exists a subset $e^\lambda \subseteq M$ containing p, which is diffeomorphic to a λ-dimensional disk, such that $M^{f(p)-\epsilon} \cap e^\lambda = \partial e^\lambda$ and $M^{f(p)-\epsilon} \cup e^\lambda$ is a deformation retract of $M^{f(p)+\epsilon}$.*

Proof. By Theorem 5.1.9, there exist local coordinates $(\chi = (u^1,\ldots,u^n), W)$ around p with $\chi(p) = 0$, such that f takes the form

$$f \circ \chi^{-1}(u) = f(p) - (u^1)^2 - \ldots - (u^\lambda)^2 + (u^{\lambda+1})^2 + \ldots + (u^n)^2.$$

Now we choose an $\epsilon > 0$ that satisfies the following conditions:

(I) $M^{[f(p)-\epsilon,f(p)+\epsilon]}$ is compact and $M^{[f(p)-\epsilon,f(p)+\epsilon]} \cap (C_f \setminus \{p\}) = \emptyset$;

(II) $\bar{B}_{\sqrt{2\epsilon}} \subsetneq \chi(U) \subseteq \mathbb{R}^n$, where $\bar{B}_{\sqrt{2\epsilon}} := \{(u^1,\cdots,u^n) | \sum_{i=1}^n (u^i)^2 \leq 2\epsilon\}$.

Let $e^\lambda \subseteq M$ be defined as

$$e^\lambda := \left\{ q \in U | \sum_{i=1}^{\lambda} (u^i(q))^2 \leq \epsilon \wedge u^{\lambda+1}(q) = \ldots = u^n(q) = 0 \right\}. \tag{5.6}$$

By definition, $p \in e^\lambda$ and e^λ is diffeomorphic to a λ-dimensional disk. We want to show that $M^{f(p)-\epsilon} \cap e^\lambda = \partial e^\lambda$. Consider the expression

$$f(q) = f(p) - (u^1(q))^2 - \ldots - (u^\lambda(q))^2 + (u^{\lambda+1}(q))^2 + \ldots + (u^n(q))^2.$$

It is not hard to see that a point $q \in e^\lambda$ satisfies $f(q) \leq f(p) - \epsilon$ if and only if $q \in \partial e^\lambda$. Now we deform f to a smooth function $F : M \to \mathbb{R}$. First, consider a smooth function $\mu : \mathbb{R} \to \mathbb{R}$, satisfying the following conditions:

$$\begin{aligned} \mu(0) &> \epsilon \\ \mu(r) &= 0 \qquad \text{if } r \geq 2\epsilon \\ -1 \leq \mu'(r) &\leq 0 \qquad \text{for all } r \in \mathbb{R} \end{aligned} \tag{5.7}$$

Note that these conditions are compatible, i.e. such a function exists. Then we define the function $F : M \to \mathbb{R}$ as

$$F(q) := \begin{cases} f(q), & \text{if } q \in M \setminus U \\ f(q) - \mu\left(\sum_{i=1}^{\lambda}(u^i(q))^2 + 2\sum_{j=\lambda+1}^{n}(u^j(q))^2\right), & \text{if } q \in U. \end{cases}$$

By condition (II) and (5.7), $\operatorname{supp}(f - F) \subseteq \chi^{-1}(\bar{B}_{\sqrt{2\epsilon}}) \subsetneq U$, so F is smooth on M. Furthermore, $F \leq f$ on M.

By defining two functions $\xi, \eta : U \to \mathbb{R}$ on U as

$$\begin{aligned} \xi(q) &= (u^1(q))^2 + \ldots + (u^\lambda(q))^2, \\ \eta(q) &= (u^{\lambda+1}(q))^2 + \ldots + (u^n(q))^2, \end{aligned}$$

we can express f and F on U in the following way

$$\begin{aligned} f &= f(p) - \xi + \eta \\ F &= f(p) - \xi + \eta - \mu(\xi + 2\eta). \end{aligned} \tag{5.8}$$

After these preparations, we can perform the proof in six steps.

Step (A) **The sets $M_F^{f(p)+\epsilon}$ and $M_f^{f(p)+\epsilon}$ are equal.**

Proof of Step (A). Since $F \leq f$, the inclusion $M_f^{f(p)+\epsilon} \subseteq M_F^{f(p)+\epsilon}$ is trivial. Let $q \in M_F^{f(p)+\epsilon}$. To prove the other inclusion it suffices to consider the case where $f(q) \neq F(q)$. By definition of F, this can only happen if $q \in U$ and $\xi(q) + 2\eta(q) \leq 2\epsilon$. In this case, we obtain

$$F(q) \leq f(q) = f(p) - \xi(q) + \eta(q) \leq f(p) + \frac{1}{2}\xi(q) + \eta(q) \leq f(p) + \epsilon,$$

so we have proven the step. ☐

Step (B) **The functions F and f have the same critical points.**

Proof of Step (B). Since $f = F$ on $M \setminus U$, it suffices to show that $U \cap C_F = U \cap C_f = \{p\}$. On U, we can conclude from (5.7) and (5.8) that

$$\begin{aligned} \frac{\partial F}{\partial \xi} &= -1 - \mu'(\xi + 2\eta) < 0 \\ \frac{\partial F}{\partial \eta} &= 1 - 2\mu'(\xi + 2\eta) \geq 1. \end{aligned}$$

Since

$$\frac{\partial F}{\partial u^i} = \frac{\partial F}{\partial \xi}\frac{\partial \xi}{\partial u^i} + \frac{\partial F}{\partial \eta}\frac{\partial \eta}{\partial u^i} = \begin{cases} 2\frac{\partial F}{\partial \xi}u^i & i \leq \lambda \\ 2\frac{\partial F}{\partial \eta}u^i & i > \lambda \end{cases},$$

the partial derivatives $\frac{\partial F}{\partial u^i} = 0$ for all $i \in \{1, \cdots, n\}$ if and only if $u^i = 0$ for all i. Therefore, p is the only critical point of F in U. $\square$

Step (C) $\boxed{\text{The set } M_F^{[f(p)-\epsilon, f(p)+\epsilon]} \text{ is compact and } M_F^{[f(p)-\epsilon, f(p)+\epsilon]} \cap C_F = \emptyset.}$

Proof of Step (C). From (A) and the inequality $F \leq f$, it follows that

$$M_F^{[f(p)-\epsilon, f(p)+\epsilon]} \subset M_f^{[f(p)-\epsilon, f(p)+\epsilon]}.$$

By assumption, $M_f^{[f(p)-\epsilon, f(p)+\epsilon]}$ is compact, which implies the compactness of $M_F^{[f(p)-\epsilon, f(p)+\epsilon]}$. Since by (B), $C_F = C_f$ and the assumption that p is the only critical point of f in $M_f^{[f(p)-\epsilon, f(p)+\epsilon]}$, we obtain

$$M_F^{[f(p)-\epsilon, f(p)+\epsilon]} \cap C_F \subset M_f^{[f(p)-\epsilon, f(p)+\epsilon]} \cap C_f = \{p\}\,.$$

Therefore, it suffices to show that $M_F^{[f(p)-\epsilon, f(p)+\epsilon]} \cap \{p\} = \emptyset$. Since

$$F(p) = f(p) - \underbrace{\xi(p)}_{=0} + \underbrace{\eta(p)}_{=0} - \mu(\underbrace{\xi(p) + 2\eta(p)}_{=0}) = f(p) - \mu(0) \overset{(5.7)}{<} f(p) - \epsilon, \quad (5.9)$$

we conclude that $p \notin M_F^{[f(p)-\epsilon, f(p)+\epsilon]}$ which proves Step (C). $\square$

Step (D) $\boxed{\text{The set } M_F^{f(p)-\epsilon} \text{ is a deformation retract of } M_f^{f(p)+\epsilon}.}$

Proof of Step (D). By (C), $M_F^{[f(p)-\epsilon, f(p)+\epsilon]}$ is compact and does not contain critical points of F. From Theorem 5.2.6 it follows that $M_F^{f(p)-\epsilon}$ is a deformation retract of $M_F^{f(p)+\epsilon} \overset{(A)}{=} M_f^{f(p)+\epsilon}$. $\square$

Now we introduce the set $H \subset M$, given by

$$H = \overline{M_F^{f(p)-\epsilon} \setminus M_f^{f(p)-\epsilon}}.$$

By (5.9), $F(p) < f(p) - \epsilon$, so $p \in H$. In particular, H is not empty.

Step (E) $\boxed{\text{The set } H \text{ satisfies } M_F^{f(p)-\epsilon} = M_f^{f(p)-\epsilon} \cup H \text{ and } e^\lambda \subset H \subset U.}$

Proof of Step (E). First, we prove $M_F^{f(p)-\epsilon} = M_f^{f(p)-\epsilon} \cup H$. Obviously,

$$M_F^{f(p)-\epsilon} \subset M_f^{f(p)-\epsilon} \cup \overline{M_F^{f(p)-\epsilon} \setminus M_f^{f(p)-\epsilon}} = M_f^{f(p)-\epsilon} \cup H.$$

On the other hand, since $M_F^{f(p)-\epsilon}$ is closed, we have

$$M_f^{f(p)-\epsilon} \cup H \subset \overline{M_f^{f(p)-\epsilon} \cup M_F^{f(p)-\epsilon} \setminus M_f^{f(p)-\epsilon}} = \overline{M_F^{f(p)-\epsilon}} = M_F^{f(p)-\epsilon}.$$

Now we are going to show that $e^\lambda \subset H$. Let $q \in e^\lambda$. Then, by definition of e^λ, $0 = \xi(p) \leq \xi(q) \leq \epsilon$ and $\eta(q) = 0 = \eta(p)$. Since $\frac{\partial F}{\partial \xi} < 0$,

$$F(q) = F(\xi(q), \eta(q)) \leq F(\xi(p), \eta(p)) = F(p) \overset{(5.9)}{<} f(p) - \epsilon,$$

so $q \in M_F^{f(p)-\epsilon}$. Furthermore,

$$f(q) = f(p) - \underbrace{\xi(q)}_{\leq \epsilon} + \underbrace{\eta(q)}_{=0} \geq f(p) - \epsilon.$$

If $f(q) > f(p) - \epsilon$, then $q \in M_F^{f(p)-\epsilon} \setminus M_f^{f(p)-\epsilon} \subset H$. If $f(q) = \epsilon$, then $q \in \partial e^\lambda$ and there exists a sequence $\{q_n\}$ of points in the interior of e^λ converging to q. Since $f(q_n) > f(p) - \epsilon$, $q_n \in H$ and, by closedness of H, $q \in H$.

It remains to show that $H \subset U$. By definition of F, we know that $F(q) \neq f(q)$, only if $q \in U$ and $\xi(q) + 2\eta(q) \leq 2\epsilon$. Therefore, $q \in M_F^{f(p)-\epsilon} \setminus M_f^{f(p)-\epsilon}$ only if $\xi(q)+2\eta(q) \leq 2\epsilon$, and by closedness of this set, the same holds for $q \in H$. So, H is contained in this set and in particular, $H \subset U$. □

Step (F) The set $M_f^{f(p)-\epsilon} \cup e^\lambda$ is a deformation retract of $M_f^{f(p)-\epsilon} \cup H$.

Proof of Step (F). We construct a deformation retract as follows. Let $r : [0,1] \times M_f^{f(p)-\epsilon} \cup H \to M_f^{f(p)-\epsilon} \cup H$ be defined as

$$r_t(q) = q, \qquad \text{if } q \in M_f^{f(p)-\epsilon}.$$

If $q \in H$, we have to distinguish three cases: If $\xi(q) \leq \epsilon$,

$$r_t(q) = \chi^{-1}\left(u^1(q), \ldots, u^\lambda(q), (1-t)u^{\lambda+1}(q), \ldots, (1-t)u^n(q)\right).$$

If $\epsilon \leq \xi(q) \leq \eta(q) + \epsilon$,

$$r_t(q) = \chi^{-1}\left(u^1(q), \ldots, u^\lambda(q), s_t u^{\lambda+1}(q), s_t \ldots, u^n(q)\right),$$

where s_t is given by

$$s_t = \begin{cases} (1-t) + t\sqrt{\frac{\xi(q)-\epsilon}{\eta(q)}}, & \text{if } \eta(q) \neq 0, \\ 0, & \text{if } \eta(q) = 0. \end{cases}$$

We check that the function $(t,q) \mapsto s_t u^i(q)$ is continuous for $\lambda < i \leq n$. Suppose that $\eta(q) \to 0$, then $u^i(q) \to 0$. Since $0 \leq \frac{\xi(q)-\epsilon}{\eta(q)} \leq 1$, s_t is bounded for all t. Thus, $s_t u^i(q) \to 0$, which proves continuity.

If and only if $\eta(q) + \epsilon \leq \xi(q)$, we have

$$f(q) = f(p) - \xi(q) + \eta(q) \leq f(p) - \epsilon, \tag{5.10}$$

so $q \in M_f^{f(p)-\epsilon}$. Therefore, we have to set $r_t(q) = q$. Note that r_t is well defined in the cases $\xi(q) = \epsilon$ and $\xi(q) = \eta(q) + \epsilon$. Thus, r is well defined and continuous in both variables. It is easy to see that $r_0 = \mathrm{id}|_{M_f^{f(p)-\epsilon} \cup H}$ and, by definition of e^λ, $r_t|_{e^\lambda} = \mathrm{id}|_{e^\lambda}$.

It remains to show that $r_1(M_f^{f(p)-\epsilon} \cup H) \subset M_f^{f(p)-\epsilon} \cup e^\lambda$. For $\eta(q) + \epsilon \leq \xi(q)$, this is clear by (5.10). If $\xi(q) \leq \epsilon$, $\xi(r_1(q)) = \xi(q)$ and $\eta(r_1(q)) = 0$, so $r_1(q) \in e^\lambda$. For $\epsilon \leq \xi(q) \leq \eta(q) + \epsilon$ and $\eta(q) = 0$, then $\xi(q) = \epsilon$ and the previous case applies. Finally, if $\eta(q) + \epsilon \leq \xi(q)$ and $\eta(q) \neq 0$, we have $\xi(r_1(q)) = \xi(q)$ and

$$\eta(r_1(q)) = \frac{\xi(q) - \epsilon}{\eta(q)} \eta(q) = \xi(q) - \epsilon,$$

so

$$f(r_1(q)) = f(p) - \xi(r_1(q)) + \eta(r_1(q)) = f(p) - \epsilon.$$

Hence we have proven step (F). □

To finish the proof, we just have to combine the last three steps. By (D), $M_F^{f(p)-\epsilon}$ is a deformation retract of $M_f^{f(p)+\epsilon}$ and by (F), $M_f^{f(p)-\epsilon} \cup e^\lambda$ is a deformation retract of $M_f^{f(p)-\epsilon} \cup H$. By (E), $M_f^{f(p)-\epsilon} \cup H$ equals $M_F^{f(p)-\epsilon}$. We obtain from Lemma 5.2.8 below that $M_f^{f(p)-\epsilon} \cup e^\lambda$ is a deformation retract of $M_f^{f(p)+\epsilon}$. □

Lemma 5.2.8. *Let X be a topological space and A, B be subspaces of X with $B \subset A$. If A is a deformation retract of X and B is a deformation retract of A, then B is a deformation retract of X.*

Proof. Let $r_1 : [0,1] \times X \to X$ be a deformation retraction from X onto A and $r_2 : [0,1] \times A \to A$ be a deformation retraction from A onto B. Then the map $r : [0,1] \times X \to X$, defined by

$$r(t,x) = \begin{cases} r_1(2t, x), & \text{if } t \in [0, \frac{1}{2}], \\ r_2(2t-1, r_1(1,x)), & \text{if } t \in [\frac{1}{2}, 1], \end{cases}$$

is a deformation retraction from X onto B. □

Now we state the general version of Theorem 5.2.7.

Theorem 5.2.9. *Let $f \in C^\infty$, $c \in \mathbb{R}$ and $p_1, \ldots, p_l \in f^{-1}(c)$ be non-degenerate critical points of f with $\operatorname{ind}_f(p_i) = \lambda_i$ for $1 \leq i \leq l$. Assume that there exists an $\epsilon > 0$ such that $M^{[c-\epsilon, c+\epsilon]}$ is compact and contains no critical points except $p_1, \ldots, p_l$. Then for $1 \leq i \leq l$, there exist pairwise disjoint subsets $e^{\lambda_i} \subseteq M$ containing p_i, each diffeomorphic to a λ_i-dimensional disk, such that $M^{c-\epsilon} \cap e^{\lambda_i} = \partial e^{\lambda_i}$ and $M^{c-\epsilon} \cup e^{\lambda_1} \cup \ldots \cup e^{\lambda_l}$ is a deformation retract of $M^{c+\epsilon}$.*

Sketch of proof. Choose pairwise disjoint neighborhoods U_i of p_i, $1 \leq i \leq l$ such that f can be written in the form

$$(f \circ \varphi_i^{-1})(u_i) = f(p_i) - (u_i^1)^2 - \ldots - (u_i^{\lambda_i})^2 + (u_i^{\lambda_i+1})^2 + \ldots + (u_i^n)^2,$$

for local coordinates $(U_i, \varphi_i = (u_i^1, \ldots, u_i^n))$ around p_i, $1 \leq i \leq l$. To do this proof, one has to generalize the construction of the proof of Theorem 5.2.7 □

Remark 5.2.10 (Cell Decomposition). Theorem 5.2.9 shows that $M^{c+\epsilon}$ is of the same homotopy type as $M^{c-\epsilon}$ with the λ_i-cells e^{λ_i}, $1 \leq i \leq l$ attached. For more details concerning cell decomposition, see [Mat02].

Lemma 5.2.11. *Let X be a topological space and A be a deformation retract of X. Let $\alpha : [0,1] \to X$ be a path in X, i.e. a continuous map from $[0,1]$ into X. If $\alpha(0), \alpha(1) \in A$, then α is homotopic, keeping endpoints fixed, to a curve in A.*

Proof. Let $r : [0,1] \times X \to X$ be a deformation retraction from X onto A. Consider the map $H : [0,1] \times [0,1] \to X$, given by

$$H_t(s) = r_t(\alpha(s)).$$

Then $H_0(\alpha) = \alpha$, $H_1(\alpha) \subset A$ and $H_t(\alpha(0)) = \alpha(0)$, $H_t(\alpha(1)) = \alpha(1)$, since $\alpha(0), \alpha(1) \in A$ for all $t \in [0,1]$. Therefore, H is a homotopy from α to the curve $H_1(\alpha)$ in A which leaves the endpoints fixed. □

The following corollary is stated in [Car92], Chapter 13, Lemma 3.3. It will play an important role later. The proof consists in repeatedly applying the previous results.

Corollary 5.2.12. *Let $f \in C^\infty(M)$ be a Morse function, $p, q \in M$ and $\alpha : [0,1] \to M$ be a path with $\alpha(0) = p$ and $\alpha(1) = q$. Let $a = \max\{f(p), f(q)\}$, $b = \max_{t\in[0,1]}(f \circ \alpha(t))$ and let c be the largest critical value in $[a,b]$ such that there exists a critical point in $f^{-1}(c)$ of index zero or one. If such a value does not exist, we set $c = a$. Assume that there exists $\epsilon > 0$ such that $M^{[a,b+\epsilon]}$ is compact. Then for all $\delta > 0$, α is homotopic, keeping endpoints fixed, to a curve β such that $\beta([0,1]) \subset M^{c+\delta}$.*

Proof. Note that $a \leq c \leq b$. Suppose that $b > a$, otherwise, there is nothing to prove. Since $M^{[c,b]}$ is compact (being a subset of the compact set $M^{[a,b+\epsilon]}$) and C_f is discrete, there exist only a finite number of critical points in $M^{[c,b]}$. Let $c_1 > \ldots > c_k$ be the critical values of f in $M^{[c,b]} \setminus f^{-1}(c)$ and $p_{ij} \in f^{-1}(c_i)$, $1 \leq i \leq k$, $1 \leq j \leq l_i$ be the corresponding critical points. By definition of c, $\lambda_{ij} := \mathrm{ind}_f(p_{ij}) \geq 2$. By choosing a smaller $\epsilon > 0$, if necessary, we may assume that c_i is the only critical value in $[c_i - \epsilon, c_i + \epsilon]$ for $i \in \{1, \ldots, k\}$.

Assume that b is not a critical value of f, so $b > c_1$. Then $M^{[c_1+\epsilon, b+\epsilon]}$ is compact and does not contain any critical points of f. By Theorem 5.2.6, $M^{c_1+\epsilon}$ is a deformation retract of $M^{b+\epsilon}$. By Lemma 5.2.11, there exists a homotopy between α and a new curve $\tilde{\alpha}$ with $\tilde{\alpha}([0,1]) \subset M^{c_1+\epsilon}$, which keeps the endpoints fixed. If $b = c_1$, we define $\tilde{\alpha} = \alpha$.

The set $M^{[c_1-\epsilon, c_1+\epsilon]}$ is compact and does not contain critical points except p_{1j}, $1 \leq j \leq l_1$. Therefore, we can apply Theorem 5.2.9. So, $M^{c_1-\epsilon} \cup e^{\lambda_{11}} \cup \ldots \cup e^{\lambda_{1l_1}}$ is a deformation retract of $M^{c_1+\epsilon}$ for pairwise disjoint sets $e^{\lambda_{1j}}$ which are diffeomorphic to λ_{1j}-dimensional disks and satisfy $M^{c_1-\epsilon} \cap e^{\lambda_{1j}} = \partial e^{\lambda_{1j}}$. Again by Lemma 5.2.11, $\tilde{\alpha}$ is homotopic to a path $\tilde{\tilde{\alpha}}$ in $M^{c_1-\epsilon} \cup e^{\lambda_{11}} \cup \ldots \cup e^{\lambda_{1l_1}}$. Since $\dim(e^{\lambda_{1j}}) \geq 2$, each path passing through $e^{\lambda_{1j}}$ can be deformed to a path which lies only in $\partial e^{\lambda_{1j}}$, but not in the interior of $e^{\lambda_{1j}}$. Note that the endpoints p and q are contained in $M^{c_1-\epsilon}$, since $a = \max\{f(p), f(q)\} < c_1 - \epsilon$. Therefore, by Theorem 5.2.9, p and q are not contained in the interior of any $e^{\lambda_{1j}}$. Thus, we can deform $\tilde{\tilde{\alpha}}$ to a curve α_1 in $M^{c_1-\epsilon} \cup \partial e^{\lambda_{11}} \cup \ldots \cup \partial e^{\lambda_{1l_1}} = M^{c_1-\epsilon}$.

Since $M^{[c_2+\epsilon, c_1-\epsilon]}$ is compact and does not contain critical points of f, we obtain from Theorem 5.2.6 and Lemma 5.2.11 a homotopy between α_1 and a curve $\tilde{\alpha_1}$ in $M^{c_2+\epsilon}$.

We repeat this procedure a finite number of times, obtaining a sequence of curves α_i with $\alpha_i([0,1]) \subset M^{c_i-\epsilon}$ which are homotopic to each other by an endpoint-fixing homotopy. Consider the curve α_k in $M^{c_k-\epsilon}$, constructed as above. Since $M^{[c+\delta, c_k-\epsilon]}$ is compact and does not contain critical points of f, we can again deform, by Theorem 5.2.6 and Lemma 5.2.11, α_k to a new curve in $M^{[c+\delta]}$, which we call β. This finishes the proof □

5.3 The path space

In this section, the object of our interest is the space of curves joining two fixed points in a Riemannian manifold. We will construct an approximation of this set, which itself is a finite dimensional manifold. We follow §16 of [Mil63].

Let M be a complete and connected Riemannian manifold and $p, q \in M$. We denote by $\Omega_{p,q}(M)$ the set of all piecewise smooth curves $\alpha : [0,1] \to M$ with $\alpha(0) = p$ and $\alpha(1) = q$. For the rest of this section, we assume p, q and M to be fixed, so we write $\Omega = \Omega_{p,q}(M)$. To equip Ω with a topology, we introduce a metric on this set. For $\alpha, \beta \in \Omega$, we define a distance function $D : \Omega \times \Omega \to \mathbb{R}$ as follows:

$$D(\alpha,\beta) := \max_{t\in[0,1]} d(\alpha(t),\beta(t)) + \sqrt{\int_0^1 \left(|\alpha'(t)| - |\beta'(t)|\right)^2 dt},$$

where $d : M \times M \to \mathbb{R}$ denotes the Riemannian distance function on M. It is not hard to see that D really defines a metric on Ω. This metric induces the required topology on M. The Energy function $E : \Omega \to \mathbb{R}$ is defined as

$$E(\alpha) = \int_0^1 |\alpha'(t)|^2 \, dt.$$

With respect to the metric on Ω, the Energy function is continuous, since

$$\begin{aligned}
|E(\alpha) - E(\beta)| &= \left| \int_0^1 (|\alpha'(t)|^2 - |\beta'(t)|^2) dt \right| \\
&\leq \left[\int_0^1 \left(|\alpha'(t)| - |\beta'(t)|\right)^2 dt \right]^{\frac{1}{2}} \left[\int_0^1 \left(|\alpha'(t)| + |\beta'(t)|\right)^2 dt \right]^{\frac{1}{2}} \\
&\leq D(\alpha,\beta) \left[\int_0^1 \left(|\alpha'(t)| - |\beta'(t)| + 2|\beta'(t)|\right)^2 dt \right]^{\frac{1}{2}} \\
&\leq D(\alpha,\beta) \left(\left[\int_0^1 \left(|\alpha'(t)| - |\beta'(t)|\right)^2 dt \right]^{\frac{1}{2}} + 2 \left[\int_0^1 |\beta'(t)|^2 dt \right]^{\frac{1}{2}} \right) \\
&\leq D(\alpha,\beta)(D(\alpha,\beta) + 2E(\beta)).
\end{aligned}$$

Therefore, $E(\alpha) \to E(\beta)$, whenever $\alpha \to \beta$ with respect to D.

We introduce some further notation. For $c > 0$, we set $\Omega^c = E^{-1}([0,c)) \subset \Omega$. Let $0 = t_0 < \ldots < t_k = 1$ be a subdivision of the unit interval. Then we define

$$\Omega(t_0, \ldots, t_k) = \left\{ \alpha \in \Omega | \alpha|_{[t_{i-1},t_i]} \text{ is an unbroken geodesic for } 1 \leq i \leq k \right\}.$$

Moreover, we set

$$\Omega(t_0 \ldots, t_k)^c = \Omega^c \cap \Omega(t_0, \ldots, t_k).$$

Proposition 5.3.1. *Let M be a complete Riemannian manifold, $p, q \in M$ and $c > 0$ such that $\Omega^c \neq \emptyset$. If the subdivision $0 = t_0 < \ldots < t_k = 1$ is sufficiently fine, the set $\Omega(t_0 \ldots, t_k)^c$ can be given the structure of a smooth finite dimensional manifold.*

Proof. The idea is to identify an element $\alpha \in \Omega(t_0 \dots, t_k)^c$ with the tuple $(\alpha(t_1), \dots, \alpha(t_{k-1})) \in M \times \dots \times M$. In this way, we can associate $\Omega(t_0 \dots, t_k)^c$ with a certain subset $V \subset M \times \dots \times M$ and adapt the manifold structure of the product. To associate each tuple $(r_1, \dots, r_{k-1}) \in V$ with a unique element in $\alpha \in \Omega(t_0 \dots, t_k)^c$, we need to make sure that r_i is in a normal neighborhood of r_{i-1} for $1 \leq i \leq k$. For this reason, we have to choose the subdivision fine enough.

Consider the compact set $K := \overline{B_{\sqrt{c}}(p)} = \{r \in M | d(p,r) \leq \sqrt{c}\} \subset M$. For $\alpha \in \Omega^c$, we have $\alpha([0,1]) \subset K$, since by Lemma 2.1.3, $L(\alpha)^2 \leq E(\alpha) < c$. We know that the map $\mathrm{Ex} : TM \to M \times M$, defined by $\mathrm{Ex}(r,v) = (r, \exp_r(v))$, is a diffeomorphism from a neighborhood W of the zero section $TM_0 = \{(x,0)|x \in M\} \subset TM$ onto a neighborhood of the diagonal $\Delta_M = \{(x,x)|x \in M\}$
$\subset M \times M$(cf. [Kun09], TH 2.4.6). Thus, for each $r \in M$, there exists a neighborhood U_r of r in M and a number $\epsilon_r > 0$ such that the set

$$\{(x,v) \in TM | x \in U_r, |v| < \epsilon_r\} \subset W.$$

Choose $r_1, \dots, r_n$ such that $K \subset U := U_{r_1} \cup \dots \cup U_{r_n}$ and let $\epsilon = \min\{\epsilon_{r_1}, \dots, \epsilon_{r_n}\}$. Then, Ex is a diffeomorphism on the set

$$\{(x,v) \in TM | x \in U, |v| < \epsilon\} \subset W.$$

In particular, for each $r \in K$, $\exp_r$ is a diffeomorphism on $B_\epsilon(0) \subset T_rM$. Therefore, there exists a unique minimizing geodesic joining two points $r_1, r_2 \in K$, whenever $d(r_1, r_2) < \epsilon$.

Now we choose a subdivision $0 = t_0 < \dots < t_k = 1$ of the unit interval such that $t_i - t_{i-1} < \frac{\epsilon^2}{c}$ for all $i \in \{1, \dots, k\}$. Then for each broken geodesic $\alpha \in \Omega(t_0 \dots, t_k)^c$, we obtain from Lemma 2.1.3,

$$\begin{aligned} d(\alpha(t_{i-1}), \alpha(t_i))^2 &\leq L(\alpha|_{[t_{i-1},t_i]})^2 \\ &= (t_i - t_{i-1}) E(\alpha|_{[t_{i-1},t_i]}) \\ &\leq (t_i - t_{i-1}) E(\alpha) < (t_i - t_{i-1})c < \epsilon^2. \end{aligned}$$

Therefore, $\alpha|_{[t_{i-1},t_i]}$ is the unique minimizing geodesic joining $\alpha(t_{i-1})$ to $\alpha(t_i)$. We now define a map

$$\begin{aligned} \varphi : \Omega(t_0 \dots, t_k)^c &\to M \times \dots \times M \\ \alpha &\mapsto (\alpha(t_1), \dots, \alpha(t_{k-1})). \end{aligned} \tag{5.11}$$

We want to show that φ is a homeomorphism onto its image. By the remark above, φ is injective. To prove continuity of φ, let $\{\alpha_n\}$ be a sequence in $\Omega(t_0 \dots, t_k)^c$ converging to $\alpha \in \Omega(t_0 \dots, t_k)^c$. From $D(\alpha_n, \alpha) \to 0$, it follows in particular that $(\alpha_n(t_1), \dots, \alpha_n(t_{k-1})) \to (\alpha(t_1), \dots, \alpha(t_{k-1}))$, which proves our claim. We denote by V the set $\varphi(\Omega(t_0 \dots, t_k)^c) \subset M \times \dots \times M$. By Lemma 2.1.3, the energy of a curve $\alpha \in \Omega(t_0 \dots, t_k)^c$ is given by

$$E(\alpha) = \sum_{i=1}^{k} \frac{d(\alpha(t_i), \alpha(t_{i-1}))^2}{t_i - t_{i-1}}.$$

It follows, that V is given explicitly by

$$V = \left\{ (r_1, \ldots, r_{k-1}) \in M \times \ldots \times M | \sum_{i=1}^{k} \frac{d(r_i, r_{i-1})^2}{t_i - t_{i-1}} < c \right\},$$

in which we set $r_0 = p$ and $r_k = q$. Therefore, V is an open set in $M \times \ldots \times M$, hence itself a manifold. It remains to show that the inverse of φ is continuous. Let $\left\{ (r_{n_1}, \ldots, r_{n_{k-1}}) \right\}$ be a sequence in V, converging to $(r_1, \ldots, r_{k-1}) \in V$, and denote the corresponding curves by α_n, α, respectively. Then

$$\begin{aligned} \alpha_n|_{[t_{i-1},t_i]}(t) &= \exp_{r_{n_{i-1}}} \left(\frac{t - t_{i-1}}{t_i - t_{i-1}} (\exp_{r_{n_{i-1}}} |_{B_\epsilon(0)})^{-1} (r_{n_i}) \right) \\ &\underset{n \to \infty}{\longrightarrow} \exp_{r_{i-1}} \left(\frac{t - t_{i-1}}{t_i - t_{i-1}} (\exp_{r_{i-1}} |_{B_\epsilon(0)})^{-1} (r_i) \right) = \alpha|_{[t_{i-1},t_i]}(t), \end{aligned}$$

so α_n converges pointwise to α. Since $[0,1]$ is compact, α_n converges uniformly to α, i.e. $\max_{t\in[0,1]}(\alpha_n(t), \alpha(t)) \to 0$. It remains to estimate the integral term. By the proof of Lemma 2.1.3 again,

$$\sqrt{\int_0^1 (|\alpha_n'(t)| - |\alpha'(t)|)^2 \, dt} = \sqrt{\sum_{i=1}^{k} \frac{\left(d(r_{n_i}, r_{n_{i-1}}) - d(r_i, r_{i-1})\right)^2}{t_i - t_{i-1}}} \underset{n \to \infty}{\longrightarrow} 0.$$

Together, $D(\alpha_n, \alpha) \to 0$, so φ^{-1} is continuous. Hence, $\Omega(t_0 \ldots, t_k)^c$ is homeomorphic to the manifold V. This homeomorphism induces a smooth structure on $\Omega(t_0 \ldots, t_k)^c$, such that $\varphi : \Omega(t_0 \ldots, t_k)^c \to V$ is a diffeomorphism. □

Theorem 5.3.2. *The topological space $\Omega(t_0 \ldots, t_k)^c$ is a deformation retract of Ω^c.*

Proof. We show that a deformation retraction is given by $r : [0,1] \times \Omega^c \to \Omega^c$, which is, for $s \in [t_{i-1}, t_i]$, defined as

$$r_s(\alpha) = \begin{cases} r_s(\alpha)|_{[t_{j-1},t_j]} & \text{min. geod. joining } \alpha(t_{j-1}) \text{ to } \alpha(t_j),\ j < i, \\ r_s(\alpha)|_{[t_i,s]} & \text{min. geod. joining } \alpha(t_{i-1}) \text{ to } \alpha(s), \\ r_s(\alpha)|_{[s,1]} = \alpha|_{[s,1]}. & \end{cases}$$

To begin with, note that r_s is well defined: There exist unique minimizing geodesics joining $\alpha(t_{j-1})$ to $\alpha(t_j)$ for $j < i$ and $\alpha(t_{i-1})$ to $\alpha(s)$. Next, we prove that $E(r_s(\alpha))$ is monotonically decreasing in s, so if $\alpha \in \Omega^c$, $r_s(\alpha) \in \Omega^c$ for all $s \in [0,1]$. It suffices to prove this for each subinterval $[t_{i-1}, t_i]$, so let $s_1 < s_2 \in [t_{i-1}, t_i]$. Since $r_{s_1}(\alpha)$ and $r_{s_2}(\alpha)$ coincide on $[0, t_{i-1}]$ and $[s_2, 1]$ it is enough to prove this for the restriction of the curves to the interval $[t_{i-1}, s_2]$. By Lemma 2.1.3, we have

$$\begin{aligned} (s_2 - t_{i-1}) E(r_{s_1}(\alpha)|_{[t_{i-1},s_2]}) &\geq L(r_{s_1}(\alpha)|_{[t_{i-1},s_2]})^2 \\ &\geq L(r_{s_2}(\alpha)|_{[t_{i-1},s_2]})^2 \\ &= (s_2 - t_{i-1}) E(r_{s_2}(\alpha)|_{[t_{i-1},s_2]}), \end{aligned}$$

since $r_{s_2}(\alpha)|_{[t_{i-1},s_2]}$ is a minimizing geodesic. This proves our claim.

We clearly have $r_0 = \mathrm{id}|_{\Omega^c}$, $r_s|_{\Omega(t_0\ldots,t_k)^c} = \mathrm{id}|_{\Omega(t_0\ldots,t_k)^c}$ for $s \in [0,1]$ and $r_1(\Omega^c) \subset \Omega(t_0\ldots,t_k)^c$. Now we show that r is continuous in both variables. It suffices to prove continuity on each subset $[t_{i-1},t_i] \times \Omega^c$.

Let $\{s_m\} \subset [t_{i-1},t_i]$, $\{\alpha_n\} \subset \Omega^c$ be sequences converging to s and α, respectively. We want to show that $D(r_{s_m}(\alpha_n), r_s(\alpha)) \underset{m,n\to\infty}{\longrightarrow} 0$. By the triangle inequality it suffices to show that $D(r_{s_m}(\alpha_n), r_{s_m}(\alpha)) \underset{n\to\infty}{\longrightarrow} 0$ and $D(r_{s_m}(\alpha), r_s(\alpha)) \underset{m\to\infty}{\longrightarrow} 0$. It is clear that $(\alpha_n(t_1),\ldots,\alpha_n(t_{i-1}),\alpha_n(s_m)) \underset{n\to\infty}{\longrightarrow} (\alpha(t_1),\ldots,\alpha(t_{i-1}),\alpha(s_m))$ and by a similar argument as in the proof of Theorem 5.3.1, we obtain that the sequence of broken geodesics $r_{s_m}(\alpha_n)|_{[0,s_m]}$ converges uniformly to $r_{s_m}(\alpha)|_{[0,s_m]}$ and

$$\sqrt{\int_0^{s_m} \left(|(r_{s_m}(\alpha_n))'(t)| - |(r_{s_m}(\alpha))'(t)|\right)^2 dt} \underset{n\to\infty}{\longrightarrow} 0.$$

Since $r_{s_m}(\alpha_n)|_{[s_m,1]} = \alpha_n|_{[s_m,1]}$ and $r_{s_m}(\alpha)|_{[s_m,1]} = \alpha|_{[s_m,1]}$, it follows that $r_{s_m}(\alpha_n)$ converges uniformly to $r_{s_m}(\alpha)$ and

$$\sqrt{\int_{s_m}^{1} \left(|(r_{s_m}(\alpha_n))'(t)| - |(r_{s_m}(\alpha))'(t)|\right)^2 dt} \underset{n\to\infty}{\longrightarrow} 0.$$

Therefore, $D(r_{s_m}(\alpha_n), r_{s_m}(\alpha)) \underset{n\to\infty}{\longrightarrow} 0$. Now we show $D(r_{s_m}(\alpha), r_s(\alpha)) \underset{m\to\infty}{\longrightarrow} 0$. It is clear that $r_{s_m}(\alpha)$ converges pointwise to $r_s(\alpha)$. To estimate the integral, we assume that $s_m \leq s$ for all $m \in \mathbb{N}$. The case $s_m \geq s$ works similarly. Then

$$\begin{aligned}
&\int_0^1 \left(|(r_{s_m}(\alpha))'(t)| - |(r_s(\alpha))'(t)|\right)^2 dt \\
&\qquad = \int_{t_{i-1}}^{s_m} \left(\frac{d(\alpha(s_m),\alpha(t_{i-1}))}{s_m - t_{i-1}} - \frac{d(\alpha(s),\alpha(t_{i-1}))}{s - t_{i-1}}\right)^2 dt \\
&\qquad\quad + \int_{s_m}^{s} \left(|(r_{s_m}(\alpha))'(t)| - \frac{d(\alpha(s),\alpha(t_{i-1}))}{s - t_{i-1}}\right)^2 dt \underset{m\to\infty}{\longrightarrow} 0,
\end{aligned}$$

which proves that $D(r_{s_m}(\alpha), r_s(\alpha)) \underset{m\to\infty}{\longrightarrow} 0$, so $r : [0,1]\times\Omega^c \to \Omega^c$ is continuous. This proves that $\Omega(t_0\ldots,t_k)^c$ is a deformation retract of Ω^c. ☐

Now consider Ω^c as an infinite dimensional manifold. For $\alpha \in \Omega^c$, we denote by $\mathcal{V}_\alpha$ the set of all piecewise smooth vector fields along α which vanish at the endpoints of α. Let $x : (-\epsilon,\epsilon) \times [0,1] \to M$ of α be a proper variation of α. Such a variation can be viewed as a smooth (in the sense that all transversal curves of x are smooth) curve $\tilde{x} : (-\epsilon,\epsilon) \to \Omega^c$ and $\tilde{x}'(0)$ corresponds to the variational vector field of x. Since this variation is proper, the variational vector field vanishes at the endpoints of α, so $\tilde{x}'(0)$ can be viewed as an element in $\mathcal{V}_\alpha$. On the other hand, by Proposition 2.1.5, every element $V \in \mathcal{V}_\alpha$ is the variational vector field of a proper variation of α. Hence, $\mathcal{V}_\alpha$ is the tangent space of Ω^c at α.

We say that an element $\alpha \in \Omega^c$ is a critical point of E if $(E \circ \tilde{x})'(0) = 0$ for all smooth curves $\tilde{x} : (-\epsilon,\epsilon) \to \Omega^c$ with $\tilde{x}(0) = \alpha$. These points are, by

Proposition 2.1.7, precisely the unbroken geodesics in Ω^c. Recall the definition of the index form of an unbroken geodesic α:

$$I_\alpha(V,W) = \int_0^1 \langle V', W' \rangle - \langle R(\gamma', V)\gamma', W \rangle dt, \qquad V, W \in \mathcal{V}_\alpha$$

Since $(E \circ \tilde{x})''(0) = 2I_\alpha(V,V)$ for each proper variation x of α with variational vector field V, we may view $2I_\alpha$ as the Hessian of E in α.

For a broken geodesic $\alpha \in \Omega(t_0 \ldots, t_k)^c$, let $\mathcal{V}_\alpha^-$ be the subspace of all $V \in \mathcal{V}_\alpha$ which are Jacobi fields on each subinterval $[t_{i-1}, t_i]$. By similar arguments as above, we see that $\mathcal{V}_\alpha^-$ is the tangent space of α in the submanifold $\Omega(t_0 \ldots, t_k)^c \subset \Omega^c$.

Lemma 5.3.3. *Let $\alpha \in \Omega(t_0 \ldots, t_k)^c$. Then α is an unbroken geodesic if and only if $E_x'(0) = 0$ for all proper variations $x : (-\epsilon, \epsilon) \times [0,1] \to M$ whose longitudinal curves lie in $\Omega(t_0 \ldots, t_k)^c$.*

Proof. If α is a geodesic, then, by Proposition 2.1.7, $E_x'(0) = 0$ for all proper variations x of α. Let now $\alpha \in \Omega(t_0 \ldots, t_k)^c$ and suppose that $E_x'(0) = 0$ for all proper variations $x : (-\epsilon, \epsilon) \times [0,1] \to M$ whose corresponding curve $\tilde{x}$ lies in $\Omega(t_0 \ldots, t_k)^c$. Then by the formula for the first variation (2.1),

$$0 = E_x'(0) = -2 \sum_{i=1}^{k-1} \langle V(t_i), \Delta\alpha'(t_i) \rangle$$

where V is the variational vector field of x. We have $V \in \mathcal{V}_\alpha^-$. By the construction of the subdivision $0 = t_0 < \ldots t_k = 1$ at the beginning of the proof of Proposition 5.3.1, $\alpha(t_i)$ is contained in a normal neighborhood of $\alpha(t_{i-1})$. By Remark 3.2.7, $\alpha(t_i)$ and $\alpha(t_{i-1})$ are not conjugate along $\alpha|_{[t_{i-1}, t_i]}$. Therefore, by Lemma 3.2.8, we can choose $V \in \mathcal{V}_\alpha^-$ such that $V(t_i) = \Delta\alpha'(t_i)$ for $1 \leq i \leq k-1$. Then

$$0 = E_x'(0) = -2 \sum_{i=1}^{k-1} |\Delta\alpha'(t_i)|^2,$$

which proves that $|\Delta\alpha'(t_i)| = 0$ for all i, so α is an unbroken geodesic. □

The next proposition shows why $\Omega(t_0 \ldots, t_k)^c$ is a good approximation of Ω^c. This is because the energy function, restricted to $\Omega(t_0 \ldots, t_k)^c$, is smooth and preserves many properties of the energy function on the full path space Ω^c.

Proposition 5.3.4 (The energy function on the manifold $\Omega(t_0 \ldots, t_k)^c$).

(i) The energy function $\tilde{E} := E|_{\Omega(t_0 \ldots, t_k)^c} : \Omega(t_0 \ldots, t_k)^c \to \mathbb{R}$ is smooth

(ii) The critical points of $\tilde{E}$ in $\Omega(t_0 \ldots, t_k)^c$ are the same as the critical points of E in Ω^c, namely the unbroken geodesics.

(iii) A critical point γ of $\tilde{E}$ is degenerate if and only if p and q are conjugate along γ.

(iv) The index of $\tilde{E}$ at γ in $\Omega(t_0 \ldots, t_k)^c$ equals the index of E at γ in Ω^c.

Proof. To prove (i), recall the map $\varphi : \Omega(t_0 \dots, t_k)^c \to M \times \dots \times M$, defined in (5.11). By Proposition 5.3.1, φ is a diffeomorphism onto its image, which we denote by V. On V, the energy function is given by

$$\tilde{E} \circ \varphi^{-1}(r_1, \dots, r_{k-1}) = \sum_{i=1}^{k} \frac{d(r_i, r_{i-1})^2}{t_i - t_{i-1}},$$

with $r_0 = p$, $r_k = q$. This expression is smooth on V (cf. [Bur09], Theorem 4.2.11). Therefore $\tilde{E}$ is smooth on $\Omega(t_0 \dots, t_k)^c$.

Assertion (ii) is precisely the statement of Lemma 5.3.3.

Before we prove (iii) and (iv), we investigate the correlation between the form I_γ and the Hessian of $\tilde{E}$ at γ. Let $V \in \mathcal{V}_\gamma^-$ and $\tilde{x} : (-\epsilon, \epsilon) \to \Omega(t_0 \dots, t_k)^c$ be a variation of the geodesic $\gamma = \tilde{x}(0)$ such that $\tilde{x}'(0) = V$. Using the formula for the second variation (2.4), it follows that $(\tilde{E} \circ \tilde{x})''(0) = 2I_\gamma(V, V)$. By Lemma 5.1.6, $\mathcal{H}_\gamma^{\tilde{E}} = 2I_\gamma|_{\mathcal{V}_\gamma^-}$.

By Proposition 4.1.4, I_γ is degenerate if its restriction to $\mathcal{V}_\gamma^-$ is. By Corollary 4.1.3, this is exactly the case if p and q are conjugate along γ. This proves (iii). Assertion (iv) follows directly from Corollary 4.1.5. □

Chapter 6

The Sphere theorem

6.1 Introduction

The aim of this chapter is to prove the sphere theorem, which states the following:

Theorem 6.1.1. *Let M^n be a complete and simply connected Riemannian manifold, whose sectional curvature satisfies*

$$0 < hK_{max} < K_{min} \leq K \leq K_{max}. \tag{6.1}$$

If $h = \frac{1}{4}$, then M is homeomorphic to a sphere.

The number h is called the pinching of M. Multiplying the metric with a constant (cf. Lemma 1.1.8), we can suppose that $K_{max} = 1$ and condition (6.1) can be replaced by

$$0 < h < K_{min} \leq K \leq 1. \tag{6.2}$$

Remark 6.1.2. In the case of even dimension, the theorem is false, if we replace (6.2) by

$$0 < \frac{1}{4} \leq K \leq 1. \tag{6.3}$$

The complex projective space $\mathbb{C}\mathrm{P}^n$, $n > 1$, with the so-called Fubini Study metric (see [Car92], Exercise 12 of Chapter 8), is a complete and simply connected Riemannian manifold of real dimension $2n$. Its sectional curvature satisfies (6.3) but $\mathbb{C}\mathrm{P}^n$ is not homeomorphic to the $2n$-dimensional sphere. However, M. Berger proved that each complete and simply connected Riemannian manifold satisfying (6.3) is either homeomorphic to a sphere or isometric to a symmetric space. This result is known as Berger's Rigidity theorem (see [CE75]).

In the odd dimensional case, it is known that the sphere theorem is still true if we replace (6.2) by (6.3), but it is not known whether the result can be improved in the sense that it is possible to choose $h < \frac{1}{4}$.

We are going to prove the sphere theorem for dimension $n \geq 3$. For $n = 2, 3$ the result is even true for any $h \geq 0$. For $n = 2$, this follows from the Gauss-Bonnet theorem (See [Lee97], Theorem 9.7 and Corollary 9.9). For $n = 3$, it follows from a Theorem of R. Hamilton [Ham82].

The Sphere theorem has a long history, going back to the 1920s. At that time, H. Hopf first posed the question of whether a compact, simply connected manifold with suitably pinched curvature is homeomorphic to a sphere. In 1951, H. Rauch [Rau51] proved the Sphere theorem for $h = \frac{3}{4}$. Furthermore, he posed the question what the optimal pinching constant would be. This question was answered around 1960 by M. Berger [Ber60] and W. Klingenberg [Kli61] who proved Theorem 6.1.1.

6.2 The cut locus

In this section, M always denotes a complete Riemannian manifold. Furthermore, all geodesics are assumed to be normalized. We follow Section 13.2 of [Car92].

Let $p \in M$ and let $\gamma : [0, \infty) \to M$ be a geodesic with $\gamma(0) = p$. We know that for small $t > 0$, γ realizes the distance between $\gamma(0)$ and $\gamma(t)$, that is, $d(\gamma(0), \gamma(t)) = t$. Moreover, the set $T := \{t \in [0, \infty) | d(\gamma(0), \gamma(t)) = t\}$ is closed, since it is defined by a continuous equation.

On the other side, if $t_1 \notin T$, that is, if $d(\gamma(0), \gamma(t_1)) < t_1$, then for every $t_2 > t_1$, $t_2 \notin T$ either, since

$$d(\gamma(0), \gamma(t_2)) \leq d(\gamma(0), \gamma(t_1)) + d(\gamma(t_1), \gamma(t_2)) < t_1 + (t_2 - t_1) = t_2.$$

Together we have that T is of the form $[0, t_0]$ or $[0, \infty)$.

Definition 6.2.1. If T, defined as above, is of the form $[0, t_0]$, we call $\gamma(t_0)$ the cut point of p along γ. If the case $[0, \infty)$ occurs, we say that a cut point does not exist. We define the cut locus, denoted by $C_m(p)$, as the union of the cut points along all normalized geodesics starting from p.

Example 6.2.2. (i) If $M = \mathbb{R}^n$, there do not exist cut points, since each geodesic is minimizing. Therefore, $C_m(p) = \emptyset$ for all $p \in \mathbb{R}^n$.

(ii) If M is a sphere S^n, then the cut locus of each point p consists of its antipodal point $-p$.

(iii) If M is the real projective space $\mathbb{R}\mathrm{P}^n$, then the cut locus of $[p, -p] \in \mathbb{R}P^n$ is the subset $\mathbb{R}P^{n-1} \subset \mathbb{R}P^n$, obtained by identifying the antipodal points of the equator of p in S^n.

(iv) If M is a product manifold $M_1 \times M_2$, then the cut locus of (p, q) is given by $C_m(p) \times M_2 \cup M_1 \times C_m(q)$, where $C_m(p)$ is the cut locus of p in M_1 and $C_m(q)$ is the cut locus of q in M_2. This shows, together with the examples above, that the cut locus of a point (p, t) on the cylinder $S^1 \times \mathbb{R}$ is given by $\{-p\} \times \mathbb{R}$. On the torus, considered as $S^1 \times S^1$, the cut locus of a point (p, q) is given by $\{-p\} \times S^1 \cup S^1 \times \{-q\}$.

Proposition 6.2.3. *Let γ be a geodesic in M. Suppose that $\gamma(t_0)$ is the cut point of $p = \gamma(0)$ along γ. Then either*

(a) $\gamma(t_0)$ is the first conjugate point of $\gamma(0)$ along γ, or

(b) there exists a geodesic $\sigma \neq \gamma$ from p to $\gamma(t_0)$ such that $\sigma(t_0) = \gamma(t_0)$.

Conversely, if (a) or (b) is satisfied, then there exists $\tilde{t}$ in $(0, t_0]$ such that $\gamma(\tilde{t})$ is the cut point of p along γ.

Proof. Let t_0 satisfy the condition asserted and let $\{t_0 + \epsilon_i\}$, $\epsilon_i > 0$ be a sequence converging to t_0. Consider a sequence of minimizing geodesics $\sigma_i : [0, L_i] \to M$ joining p to $\gamma(t_0 + \epsilon_i)$ and let $\{\sigma_i'(0)\} \in S^{n-1} \subset T_pM$ be its corresponding sequence of tangent vectors at p. Note that $L_i = d(p, \gamma(t_0 + \epsilon_i))$. Since S^{n-1} is compact, there exists a convergent subsequence, which we will again denote by $\{\sigma_i'(0)\}$ such that $\{\sigma_i'(0)\} \to v$ for a $v \in T_pM$ with $|v| = 1$. We have

$$\exp_p(d(p, \gamma(t_0 + \epsilon_i))\sigma_i'(0)) = \gamma(t_0 + \epsilon_i)$$

and, by continuity,

$$\exp_p(d(p, \gamma(t_0))v) = \gamma(t_0). \tag{6.4}$$

Consider the geodesic $\sigma : [0, \infty) \to M$, defined by $\sigma(t) := \exp_p(tv)$. By (6.4), $\sigma(t_0) = \gamma(t_0)$. If $\sigma \neq \gamma$, assertion (b) is verified. If $\sigma = \gamma$, we are going to show that (a) holds. We show that $\exp_p$ is singular at $t_0\gamma'(0)$, which is, by Proposition 3.2.5, equivalent to the claim that $\gamma(t_0)$ is conjugate to $\gamma(0)$ along γ.

Suppose now that $\exp_p$ is not singular at $t_0\gamma'(0)$. Then, $\exp_p$ is a diffeomorphism from a neighborhood $U \subset T_pM$ of $t_0\gamma'(0)$ onto a neighborhood $V \subset M$ of $\gamma(t_0)$. Since σ_i is minimizing between p and $\gamma(t_0 + \epsilon_i)$ and γ is not, we have $\gamma(t_0 + \epsilon_i) = \sigma_i(t_0 + \epsilon_i')$ for some ϵ_i' satisfying $\epsilon_i' < \epsilon_i$ for every i. Then,

$$\exp_p((t_0 + \epsilon_i)\gamma'(0)) = \gamma(t_0 + \epsilon_i) = \sigma_i(t_0 + \epsilon_i') = \exp_p((t_0 + \epsilon_i')\sigma_i'(0)).$$

For ϵ_i small enough, $\gamma(t_0 + \epsilon_i) \in V$ and we have $(t_0 + \epsilon_i)\gamma'(0) = (t_0 + \epsilon_i')\sigma_i'(0)$. Since $t_0 + \epsilon_i > 0$ and $t_0 + \epsilon_i' > 0$ and $|\gamma'(0)| = |\sigma_i'(0)| = 1$, we have $\gamma'(0) = \sigma_i'(0)$ and therefore $\epsilon_i = \epsilon_i'$. This contradicts our assumption $\epsilon_i > \epsilon_i'$ and proves that $\gamma(0)$ is conjugate to $\gamma(t_0)$ along γ. Since γ is minimizing up to $\gamma(t_0)$, by Corollary 4.2.5 no conjugate point occurs before $\gamma(t_0)$.

Now we suppose that (a) holds. We know by Corollary 4.2.5 that a geodesic does not minimize distance after its first conjugate point, so $d(0, \gamma(t)) < t$ for all $t > t_0$. By definition, the cut point of p along γ occurs at $\gamma(\tilde{t})$, $\tilde{t} \leq t_0$.

If (b) holds, choose $\epsilon > 0$ small enough so that $\sigma(t_0 - \epsilon)$ and $\gamma(t_0 + \epsilon)$ are both contained in a totally normal neighborhood of $\sigma(t_0) = \gamma(t_0)$. We have

$$d(\sigma(t_0 - \epsilon), \gamma(t_0 + \epsilon)) \leq d(\sigma(t_0 - \epsilon), \sigma(t_0)) + d(\gamma(t_0), \gamma(t_0 + \epsilon)) = 2\epsilon.$$

If equality occurs, the curve that joins $\sigma(t_0 - \epsilon)$ to $\sigma(t_0)$ via σ and $\sigma(t_0) = \gamma(t_0)$ to $\gamma(t_0 + \epsilon)$ via γ, is minimizing, hence an unbroken geodesic. This implies $\sigma'(t_0) = \gamma'(t_0)$, which contradicts $\sigma \neq \gamma$. Therefore, $d(\sigma(t_0 - \epsilon), \gamma(t_0 + \epsilon)) < 2\epsilon$ which leads to

$$d(\gamma(0), \gamma(t_0 + \epsilon)) \leq d(\sigma(0), \sigma(t_0 - \epsilon)) + d(\sigma(t_0 - \epsilon), \gamma(t_0 + \epsilon)) < t_0 - \epsilon + 2\epsilon = t_0 + \epsilon.$$

Hence, the cut point of p along γ occurs at $\gamma(\tilde{t})$, $\tilde{t} \leq t_0$. □

Corollary 6.2.4. *Let γ be a geodesic. If $q = \gamma(t_0)$ is the cut point of $p = \gamma(0)$ along γ, then p is the cut point of q along the geodesic $\tilde{\gamma}$, which is defined as $\tilde{\gamma}(t) := \gamma(t_0 - t)$. In particular, $q \in C_m(p)$ if and only if $p \in C_m(q)$.*

Proof. By Proposition 6.2.3, either $\gamma(t_0)$ is the first conjugate point of $\gamma(0)$ along γ, or there exists a geodesic $\sigma \neq \gamma$ from p to $\gamma(t_0)$ such that $\sigma(t_0) = \gamma(t_0)$. If the first case occurs, we have a nontrivial Jacobi field V along γ, which vanishes at $\gamma(0)$ and $\gamma(t_0)$. Then, $\tilde{V}(t) := V(t_0 - t)$ is a Jacobi field along the geodesic $\tilde{\gamma}(t) = \gamma(t_0 - t)$, which vanishes at $\tilde{\gamma}(0) = q$ and $\tilde{\gamma}(t_0) = p$, that is, p is conjugate to q along $\tilde{\gamma}$. If the other case occurs, we have two geodesics $\tilde{\gamma} \neq \tilde{\sigma}$ joining q to p such that $\tilde{\gamma}(t_0) = \tilde{\sigma}(t_0) = q$.

From both cases, it follows, by Proposition 6.2.3, that there exists $\tilde{t} \in (0, t_0]$ such that $\tilde{\gamma}(\tilde{t})$ is the cut point of q along $\tilde{\gamma}$. Since γ is minimizing between p and q, the same holds for $\tilde{\gamma}$, so we have $\tilde{t} = t_0$. □

Corollary 6.2.5. *Let $p \in M$. If $q \in M \setminus C_m(p)$, then there exists a unique minimizing geodesic joining p to q.*

Proof. Since M is complete, there exists at least one minimizing geodesic γ joining p to $q = \gamma(t_0)$. If there is another minimizing geodesics $\sigma \neq \gamma$ joining p to q, we conclude from Proposition 6.2.3 that there exists a $\tilde{t} \in (0, t_0]$, such that $\gamma(\tilde{t})$ is the cut point of p along γ. If $\tilde{t} < t_0$, γ is not minimizing between p and q, if $\tilde{t} = t_0$, $q = \gamma(t_0) \in C_m(p)$. Both cases lead to a contradiction. □

For the next Corollary, we denote by $B_r(0)$ the open ball of radius $r > 0$ in T_pM, centered at zero.

Corollary 6.2.6. *The map $\exp_p : B_r(0) \to M$ is injective if and only if $r \leq d(p, C_m(p))$*

Proof. We set $\tilde{B}_r(p) := \exp_p(B_r(0))$. First, we assume $r \leq d(p, C_m(p))$. For an arbitrary $q \in \tilde{B}_r(p)$, we have $d(p,q) < r \leq d(p, C_m(p))$ and therefore $q \notin C_m(p)$. Hence, $\tilde{B}_r(p) \cap C_m(p) = \emptyset$ and the injectivity follows from Corollary 6.2.5.

If $r > d(p, C_m(p))$, there exists $q \in C_m(p)$ with $d(p,q) < r$. Consider a geodesic $\gamma : [0, \infty) \to M$, starting from p whose cut point is $q = \gamma(t_0)$. We know that $t_0 = d(p,q) < r$. Let $\epsilon > 0$ such that $t_0 + \epsilon < r$. Then γ is not minimizing between $p = \gamma(0)$ and $\gamma(t_0 + \epsilon)$. Let $\sigma \neq \gamma$ be the minimizing geodesic joining p to $\gamma(t_0 + \epsilon)$. It follows that $\sigma(t_0 + \epsilon') = \gamma(t_0 + \epsilon)$ with $\epsilon' < \epsilon$. We obtain $\exp_p((t_0 + \epsilon')\sigma'(0)) = \exp_p((t_0 + \epsilon)\gamma'(0))$, that is, $\exp_p$ is not injective on $B_r(0)$. □

Definition 6.2.7. We call

$$i(M) := \inf_{p \in M} d(p, C_m(p))$$

the injectivity radius of M. By Corollary 6.2.6, it is the largest number r, such that $\exp_p$ is injective on $B_r(0)$ for all $p \in M$.

Now we want to prove that the distance of p to its cut point along γ depends continuously on the initial direction of γ. Consider the unit tangent bundle T_1M of M, which is defined as

$$T_1M := \bigcup_{p \in M} p \times (S^{n-1})_p$$

where $(S^{n-1})_p$ denotes the unit sphere in T_pM. We define a function $f : T_1M \to \mathbb{R} \cup \{\infty\}$ by :

$$f(p,v) = \begin{cases} t_0, & \text{if } \gamma(t_0) \text{ is the cut point of } p \text{ along } \gamma(t) = \exp_p(tv), \\ \infty, & \text{if there is no cut point along } \gamma. \end{cases}$$

To prove continuity of f, we introduce a topology on $\mathbb{R} \cup \infty$. Remember that the set of all open intervals of $\mathbb{R}$ form a base of the Euclidean topology on $\mathbb{R}$. We define a base of our topology on $\mathbb{R} \cup \infty$ by adding all subsets of the form $(a, \infty] = (a, \infty) \cup \infty$ to the open intervals of $\mathbb{R}$. Observe that the set $[a, \infty]$ is compact in this topology, and a sequence $t_n \to \infty$ in this topology when $\lim_{n\to\infty} t_n = \infty$ in the usual sense.

Proposition 6.2.8. *The function f, defined above, is continuous.*

Proof. Let $\{(p_i, v_i)\}$ be a sequence in T_1M, converging to (p, v). We denote the geodesics with initial points p_i, p and initial velocities v_i, v by γ_i and γ, respectively. Let $\gamma(t_0^i)$ and $\gamma(t_0)$ be the cut points of $\gamma_i(0)$ and $\gamma(0)$ along γ_i and γ, respectively, where $t_0^i, t_0 \in \mathbb{R} \cup \infty$. First, we prove that $\limsup t_0^i \le t_0$. If $t_0 = \infty$, there is nothing to prove. Let $t_0 < \infty$ and $\epsilon > 0$. Assume that there exist infinitely many indices j such that $t_0 + \epsilon < t_0^j$. Consider the subsequence corresponding to these indices. Then

$$d(\gamma_j(0), \gamma_j(t_0 + \epsilon)) = t_0 + \epsilon$$

and by continuity of d,

$$\begin{aligned} d(\gamma(0), \gamma(t_0 + \epsilon)) &= d(p, \exp_p((t_0 + \epsilon)v)) \\ &= \lim d(p_j, \exp_{p_j}((t_0 + \epsilon)v_j)) \\ &= \lim d(\gamma_j(0), \gamma_j(t_0 + \epsilon)) = t_0 + \epsilon, \end{aligned}$$

which contradicts the fact that $\gamma(t_0)$ is a cut point of $\gamma(0)$ along γ. Therefore, $\limsup t_0^i \le t_0 + \epsilon$. Since ϵ was arbitrary, we have proven the claim.

Now let $\bar{t} = \liminf t_0^i$. Since

$$\bar{t} = \liminf t_0^i \le \limsup t_0^i \le t_0, \tag{6.5}$$

it suffices to show that $\bar{t} \ge t_0$ to complete the proof. If $\bar{t} = \infty$, again, nothing needs to be proven. Suppose that $\bar{t} < \infty$ and consider a subsequence of the sequence t_0^i, denoted by t_0^j, which converges to $\bar{t}$. Then, by Proposition 6.2.3, for each $j \in \mathbb{N}$ either

(a) $\gamma_j(t_0^j)$ is conjugate to $\gamma_j(0)$ along γ_j, or

(b) there exists a geodesic $\sigma_j \ne \gamma_j$, starting from p_j such that $\sigma_j(t_0^j) = \gamma_j(t_0^j)$.

We are going to show that γ satisfies (a) or (b) with $\gamma(\bar{t})$. By Proposition 6.2.3, the cut point $\gamma(t_0)$ of $\gamma(0)$ along γ satisfies $t_0 \le \bar{t}$, so the proof will be finished.

We consider two cases: First, assume that there exist infinitely many indices j that satisfy (a). Consider a subsequence, consisting of these indices, again denoted by γ_j. By Proposition 3.2.5, $\exp_{p_j}$ is singular at $t_0^j \gamma_j'(0)$, so $T_{t_0^j \gamma_j'(0)} \exp_{p_j}$ is not an isomorphism. We have

$$\gamma(\bar{t}) = \exp_p(\bar{t}v) = \lim \exp_{p_j}(t_0^j v_j) = \lim \gamma_j(t_0^j). \tag{6.6}$$

and by continuity, $T_{\bar{t}\gamma'(0)} \exp_p$ is not an isomorphism either. Again by Proposition 3.2.5, $\gamma(\bar{t})$ is conjugate to $\gamma(0)$ along γ, so (a) is satisfied.

If the first case does not occur, then (b) is satisfied for infinitely many j. By passing to a subsequence, we may assume that (b) is satisfied for all $j \in \mathbb{N}$. Consider geodesics σ_j and the corresponding sequence of tangent vectors $\{\sigma_j'(0)\} \subset T_1M$. Since all initial points $\sigma_j(0)$ are contained in a compact subset K of M,

$$\{\sigma_j'(0)\} \subset T_1K := T_1M|_K := \bigcup_{q\in K} q \times (S^{n-1})_q$$

and T_1K is obviously compact. Therefore, there exists a subsequence, again denoted by $\{\sigma_j'(0)\}$, converging to a unit vector w in T_pM. The geodesic $\sigma : [0,\infty) \to M$, $\sigma(t) = \exp_p(tw)$ joins p to $\gamma(\bar{t})$, since

$$\sigma(\bar{t}) = \exp_p(\bar{t}w) = \lim \exp_{p_j}(t_0^j \sigma_j'(0)) = \lim \sigma_j(t_0^j) = \lim \gamma_j(t_0^j) \overset{(6.6)}{=} \gamma(\bar{t})$$

If $\sigma \neq \gamma$, (b) is satisfied.

If $\sigma = \gamma$, we are going to show that $\gamma(\bar{t})$ is conjugate to $\gamma(0)$ along γ. Consider the map $\mathrm{Ex} : TM \to M \times M$, $\mathrm{Ex}(q,u) = (q, \exp_q(u))$. The tangent map

$$T_{(q,u)}\mathrm{Ex} : T_{(q,u)}(TM) \cong T_qM \oplus T_u(T_qM) \to T_qM \oplus T_{\exp_q(u)}M$$

is given by

$$T_{(q,u)}\mathrm{Ex} = \begin{pmatrix} \mathrm{id}_{T_qM} & 0 \\ * & T_u \exp_q \end{pmatrix}.$$

Suppose now that $\gamma(\bar{t})$ is not conjugate to $\gamma(0)$ along γ. Then, by Proposition 3.2.5, $T_{\bar{t}v} \exp_p$ is an isomorphism. By the above, it follows that $T_{(p,\bar{t}v)}\mathrm{Ex}$ is an isomorphism. Choose open neighborhoods U of $(p, \bar{t}v)$ in TM and V of $(p, \gamma(\bar{t}))$ in $M \times M$, such that $\mathrm{Ex} : U \to V$ is a diffeomorphism. We have

$$\mathrm{Ex}(p_j, t_0^j \sigma_j'(0)) = (p_j, \sigma_j(t_0^j)) = (p_j, \gamma_j(t_0^j)) = \mathrm{Ex}(p_j, t_0^j \gamma_j'(0)). \tag{6.7}$$

For j large enough, $(p_j, t_0^j\sigma_j'(0))$ and $(p_j, t_0^j\gamma_j'(0))$ are in U, since both sequences converge to $(p, \bar{t}v)$. This shows, by applying $\mathrm{Ex}|_U^{-1}$ to (6.7), that $\sigma_j'(0) = \gamma_j'(0)$, which contradicts $\sigma_j \neq \gamma_j$. Therefore, $\gamma(\bar{t})$ is conjugate to $\gamma(0)$ along γ. □

Lemma 6.2.9. *Let $p \in M$ and f be defined as above. Then*

(i) $M \setminus C_m(p) = \{\exp_p(tv) | t < f(p,v), v \in S^{n-1} \subset T_pM\}$

(ii) $C_m(p) = \{\exp_p(tv) | t = f(p,v), v \in S^{n-1} \subset T_pM\}$,

(iii) $M = \{\exp_p(tv) | t \le f(p,v), v \in S^{n-1} \subset T_pM\}$.

Proof. First, we prove (i). If $q \in M \setminus C_m(p)$, then by Corollary 6.2.5, there exists a unique minimizing geodesic $\gamma : [0,\infty) \to M$ with $\gamma(\tilde{t}) = q$. Since γ is minimizing, we have $\tilde{t} \le f(p, \gamma'(0))$ and, since $q \notin C_m(p)$, it follows that $\tilde{t} < f(p,\gamma'(0))$. Conversely, if a point $q = \exp_p(\tilde{t}v) =: \gamma(\tilde{t})$, $\tilde{t} < f(p,v)$ is given, then q is not the cut point of p along γ. Furthermore, γ is the unique minimizing geodesic joining p to $q = \gamma(\tilde{t})$: if there is a geodesic σ joining p to $q = \sigma(\tilde{t})$, then by Proposition 6.2.3, $\tilde{t} \geq f(p, \gamma'(0))$, which is a contradiction. Therefore, $q \notin C_m(p)$.

Next, we prove (ii). If $q \in C_m(p)$, there exists a geodesic γ, such that $q = \gamma(t_0)$ is the cut point of p along γ. By definition, $t_0 = f(p, \gamma'(0))$. Conversely, if γ is starting from p, the point $\gamma(f(p, \gamma'(0)))$ is the cut point of p along γ.

The proof of (iii) follows immediately from (i) and (ii). □

Remark 6.2.10. For all $p \in M$, the set $M \setminus C_m(p)$ is homeomorphic to an open ball of Euclidean space. From the representation of $M \setminus C_m(p)$ in Lemma 6.2.9 (i), we construct such a homeomorphism explicitly. It is not hard to see that the map

$$\begin{aligned} \varphi : T_pM \supset B_1(0) &\to M \setminus C_m(p) \\ tv &\mapsto \exp_p(\tan(\arctan(f(p, v))t)v), \end{aligned}$$

where $t \in [0, 1)$, $|v| = 1$ and $\arctan(\infty) := \frac{\pi}{2}$, is a homeomorphism. In this sense, the topology of M is contained in its cut locus.

Corollary 6.2.11. *For all $p \in M$, $C_m(p)$ is closed.*

Proof. By Lemma 6.2.9 (ii), we have

$$C_m(p) = \left\{\exp_p(tv) | t = f(p, v), v \in S^{n-1} \subset T_pM\right\},$$

where f is the function from Proposition 6.2.8. Therefore, if q is an accumulation point of $C_m(p)$, there exists a sequence of geodesics γ_j, such that $\gamma_j(t_j^0) \to q$, where $t_j^0 = f(p, \gamma_j'(0))$. Passing to a subsequence, if necessary, the sequence $\gamma_j'(0)$ converges to a unit vector $v \in T_pM$. Let γ be the geodesic with $\gamma(0) = p$, $\gamma'(0) = v$. Then, since f is continuous,

$$\begin{aligned} q &= \lim \gamma_j(t_j^0) = \lim \gamma_j(f(p, \gamma_j'(0))) = \lim \exp_p(f(p, \gamma_j'(0))\gamma_j'(0)) \\ &= \exp_p(f(p, \gamma'(0))\gamma'(0)) = \gamma(f(p, \gamma'(0))) \in C_m(p), \end{aligned}$$

which shows that $C_m(p)$ is closed. □

Corollary 6.2.12. *The following are equivalent:*

(i) M is compact

(ii) For all $p \in M$, $C_m(p)$ is compact.

(iii) There exists a point $p \in M$ such that $C_m(p)$ is compact.

(iv) Each geodesic $\gamma : [0, \infty) \to M$ has a cut point.

Proof. We know from Corollary 6.2.11 that $C_m(p)$ is closed for every $p \in M$. If (i) is satisfied, $C_m(p)$ is a closed subset of a compact set, hence also compact, which proves (ii). From (ii), (iii) follows trivially. If (iii) holds,

$$C_m(p) \overset{6.2.9(ii)}{=} \left\{\exp_p(tv) | t = f(p, v), v \in S^{n-1} \subset T_pM\right\}$$

is bounded, so $f(p, v) < \infty$ for all $v \in S^{n-1} \subset T_pM$. Therefore,

$$M \overset{6.2.9(iii)}{=} \left\{\exp_p(tv) | t \leq f(p, v), v \in S^{n-1} \subset T_pM\right\}$$

is bounded and (i) follows from the Hopf-Rinow theorem. To finish the proof, we show that (i) and (iv) are equivalent. If M is compact, then d is bounded, so the equality $d(\gamma(0), \gamma(t)) = t$ can not be valid for all $t > 0$. Thus, there exists a cut point on each geodesic in M. Conversely, if (iv) holds, then f is bounded. As above, one shows that M is bounded and therefore compact. □

Proposition 6.2.13. *Let $p \in M$. Suppose that there exists a point $q \in C_m(p)$ that realizes the distance from p to $C_m(p)$. Then either*

(a) there exists a minimizing geodesic γ from p to q along which q is conjugate to p, or

(b) there exist exactly two minimizing geodesics γ and σ from p to q; in addition, $\gamma'(L) = -\sigma'(L)$, $L = d(p,q)$.

Proof. Let γ be a minimizing geodesic joining p to q. By Proposition 6.2.3, either q is conjugate to p along γ and (a) holds or there exists another minimizing geodesic $\sigma \neq \gamma$, joining p to q with $\sigma(L) = q = \gamma(L)$. Suppose then that q is not conjugate to p along γ and σ, and that $\gamma'(L) \neq -\sigma'(L)$, from which we are going to derive a contradiction. In particular, this shows that there can only be two such geodesics.

Since $\gamma'(L) \neq -\sigma'(L)$, there exists $w \in T_qM$ such that

$$\langle w, \gamma'(L)\rangle < 0, \qquad \langle w, \sigma'(L)\rangle < 0.$$

Let $\tau : (-\epsilon, \epsilon) \to M$ be a curve with $\tau(0) = q$ and $\tau'(0) = w$. Since q is not conjugate to p along γ, it follows from Proposition 3.2.5 that $T_{L\gamma'(0)}exp_p : T_{L\gamma'(0)}(T_pM) \to T_qM$ is an isomorphism. Choose a neighborhood $U \subset T_pM$ of $L\gamma'(0)$ where $\exp_p$ is a diffeomorphism. We may assume that $\tau((-\epsilon, \epsilon)) \subset \exp_p(U)$.

Let $\tilde{\tau} : (-\epsilon, \epsilon) \to U$ be the curve such that $\exp_p(\tilde{\tau}(s)) = \tau(s)$, $s \in (-\epsilon, \epsilon)$, and let $x : (-\epsilon, \epsilon) \times [0, L] \to M$ be the variation $x(s,t) = \exp_p(\frac{t}{L}\tilde{\tau}(s))$ of γ. Observe that $\gamma_s(L) = \tau(s)$ and that the variational vector field V of x along γ satisfies $V(0) = 0$ and $V(L) = \tau'(0) = w$. Let $\gamma_s(t) = x(s,t)$. By the formula for the first variation of arc length (2.2),

$$\frac{d}{ds}L(\gamma_s)|_{s=0} = \langle w, \gamma'(L)\rangle < 0.$$

Because q is not conjugate to p along σ, we obtain in an analogous manner a variation of σ such that

$$\frac{d}{ds}L(\sigma_s)|_{s=0} = \langle w, \sigma'(L)\rangle < 0,$$
$$\sigma_s(L) = \tau(s),$$

where the geodesics σ_s are the longitudinal curves of the variation.

Assume that the geodesics γ_s, σ_s are normalized. If $s > 0$ is sufficiently small,

$$L(\gamma_s) < L(\gamma) = d(p,q) = d(p, C_m(p)) = \inf_{v \in (S^{n-1})_p} f(p,v) \leq f(p, \gamma_s'(0))$$

and analogously,

$$L(\sigma_s) < f(p, \sigma_s'(0)).$$

It follows that the geodesics γ_s and σ_s are both minimizing between p and $\tau(s)$, hence $L(\gamma_s) = L(\sigma_s) = d(p, \tau(s))$. On the other hand, by the above, $d(p, \tau(s)) \leq L(\gamma_s) < d(p, C_m(p))$, so $\tau(s) \notin C_m(p)$. Observe that for small $s > 0$, $\gamma_s \neq \sigma_s$ since γ_s is a variation of γ, σ_s is a variation of σ and $\gamma \neq \sigma$. By Corollary 6.2.5, there can be only one minimizing geodesic between p and $\tau(s)$, so we have deduced a contradiction. □

Definition 6.2.14. A curve $\alpha : [0, a] \to M$ is called closed, if $\alpha(0) = \alpha(a)$. A geodesic $\gamma : [0, a] \to M$ which is closed is called a closed geodesic if $\gamma'(0) = \gamma'(a)$.

Proposition 6.2.15. *If the sectional curvature K of a complete Riemannian manifold M satisfies*

$$0 < K_{min} \leq K \leq K_{max},$$

then either

(a) $i(M) \geq \frac{\pi}{K_{max}}$, or

(b) there exists a closed geodesic γ in M, whose length is less then any other closed geodesic in M, and which is such that

$$i(M) = \frac{1}{2} L(\gamma).$$

Proof. By the theorem of Bonnet-Myers (Theorem 2.2.4), M is compact. By Corollary 6.2.12, $C_m(r)$ is nonempty and compact for all points $r \in M$. Since M is compact, $T_1 M$ is compact and by Proposition 6.2.8 there exists a point $p \in M$ such that

$$d(p, C_m(p)) = \inf_{r \in M} d(r, C_m(r)) = \inf_{(r,v) \in T_1 M} f(r, v).$$

Because $C_m(p)$ is compact, there exists a point $q \in M$ such that q realizes the distance from p to $C_m(p)$. Let σ be a minimizing geodesic joining p and q.

Now we apply Proposition 6.2.3. If q is the first conjugate point to p along σ, it follows from Proposition 3.4.1 that

$$i(M) = \inf_{r \in M} d(r, C_m(r)) = d(p, q) \geq \frac{\pi}{\sqrt{K_{max}}}.$$

If q is not conjugate to p along σ, there exists another minimizing geodesic μ from p to q such that $\sigma(L) = \mu(L) = q$, where $L = d(p, q) = i(M)$. By Proposition 6.2.13, we have $\sigma'(L) = -\mu'(L)$. Furthermore,

$$d(q, C_m(q)) \geq \inf_{r \in M} d(r, C_m(r)) = d(p, q). \tag{6.8}$$

Since $q \in C_m(p)$, it follows from Corollary 6.2.4 that $p \in C_m(q)$ and therefore equality in (6.8) occurs. Thus, p realizes the distance from q to $C_m(q)$. We apply Proposition 6.2.13 again. Since q is not conjugate to p along σ, then p is not conjugate to q along $\tilde{\sigma}$ either, where $\tilde{\sigma}(t) = \sigma(L - t)$. The geodesic $\tilde{\mu}$, given by $\tilde{\mu}(t) = \mu(L - t)$ is another minimizing geodesic joining q to p. By Proposition 6.2.13, $\tilde{\sigma}'(L) = -\tilde{\mu}'(L)$. We obtain a curve $\gamma : [0, 2L] \to M$, defined by

$$\gamma(t) = \begin{cases} \sigma(t), & \text{if } 0 \leq t \leq L, \\ \tilde{\mu}(t - L), & \text{if } L \leq t \leq 2L. \end{cases}$$

Observe that γ is well defined and closed. From $\sigma'(L) = -\mu'(L) = \tilde{\mu}'(0)$ and $\tilde{\mu}'(L) = -\mu'(0) = \sigma'(0)$, it follows that γ is a closed geodesic with length $L(\gamma) = 2L = 2i(M)$.

Suppose now that there exists a closed geodesic $\tilde{\gamma}$ with $L(\tilde{\gamma}) < L(\gamma)$. Let $\tilde{p} = \tilde{\gamma}(0)$. Then the cut point of $\tilde{p}$ along $\tilde{\gamma}$ occurs at $\tilde{t} \leq \frac{L(\tilde{\gamma})}{2}$, since $\tilde{\gamma}$ is clearly not minimizing after $\tilde{q} = \tilde{\gamma}(\frac{L(\tilde{\gamma})}{2})$. So,

$$d(\tilde{p}, C_m(\tilde{p})) \leq \tilde{t} \leq \frac{L(\tilde{\gamma})}{2} < \frac{L(\gamma)}{2} = i(M) = \inf_{r \in M} d(r, C_m(r)),$$

which is not possible. Thus, γ is the closed geodesic of shortest length in M. □

6.3 The estimate of the injectivity radius

In what follows, we are going to improve, under stronger conditions on the curvature, the last result from the previous section. We will show that if the sectional curvature K of a complete Riemannian manifold satisfies $\frac{1}{4} < K_{min} \leq K \leq 1$, then case (a) of Proposition 6.2.15 occurs. The proof uses results of Morse Theory we obtained in Chapter 5. We follow Section 13.3 of [Car92].

Lemma 6.3.1. *Let M^n and $\tilde{M}^n$ be two Riemannian manifolds of the same dimension such that their sectional curvatures K and $\tilde{K}$, respectively, satisfy $\sup \tilde{K} \leq \inf K$. Let $\gamma : [0, a] \to M$ and $\tilde{\gamma} : [0, a] \to \tilde{M}$ be two geodesics of the same length. Then $\mathrm{ind} I_\gamma \geq \mathrm{ind} I_{\tilde{\gamma}}$ where $I_\gamma, I_{\tilde{\gamma}}$ denote the Index forms of the geodesics γ, $\tilde{\gamma}$, respectively.*

Proof. We denote by $\mathcal{V}$ and $\tilde{\mathcal{V}}$ the set of all piecewise smooth vector fields on γ and $\tilde{\gamma}$, respectively, which vanish at the endpoints. Choose frame fields $\{E_1, \ldots, E_n\}$ and $\{\tilde{E}_1, \ldots, \tilde{E}_n\}$ along γ and $\tilde{\gamma}$, respectively, such that $E_1(t) = \frac{\gamma'(t)}{|\gamma'(t)|}$ and $\tilde{E}_1(t) = \frac{\tilde{\gamma}'(t)}{|\tilde{\gamma}'(t)|}$. Let $W(t) = \sum_{i=1}^n f_i(t) E_i(t) \in \mathcal{V}$. We map W to $\phi W \in \tilde{\mathcal{V}}$, given by $\phi W(t) = \sum_{i=1}^n f_i(t) \tilde{E}_i(t)$. Obviously, the map $\phi : \mathcal{V} \to \tilde{\mathcal{V}}$ is bijective and satisfies

$$\begin{aligned} \langle W, \gamma' \rangle &= f_1 \, |\gamma'(t)| = f_1 \, |\tilde{\gamma}'(t)| = \langle \phi W, \tilde{\gamma}' \rangle, \\ \langle W, W \rangle &= \sum_{i=1}^n f_i(t)^2 = \langle \phi W, \phi W \rangle, \\ \langle W', W' \rangle &= \sum_{i=1}^n f_i'(t)^2 = \langle \phi W', \phi W' \rangle. \end{aligned}$$

We want to show that

$$\langle R(\gamma', W)\gamma', W \rangle(t) \geq \langle \tilde{R}(\tilde{\gamma}', \phi W)\tilde{\gamma}', \phi W \rangle(t) \qquad \forall t \in [0, a]. \tag{6.9}$$

Note that $W(t)$ and $\gamma'(t)$ are linearly dependent if and only if $\phi W(t)$ and $\tilde{\gamma}'(t)$ are linearly dependent. In this case, both sides of (6.9) vanish and the inequality is satisfied trivially. In the other case, since $\sup \tilde{K} \leq \inf K$,

$$\begin{aligned} \langle R(\gamma', W)\gamma', W \rangle(t) &= \left(|\gamma'|^2 |W|^2 - \langle \gamma', W \rangle^2 \right) K(\gamma', W)(t) \\ &\geq \left(|\tilde{\gamma}'|^2 |\phi W|^2 - \langle \tilde{\gamma}', \phi W \rangle^2 \right) \tilde{K}(\tilde{\gamma}', \phi W)(t) \\ &= \tilde{R}(\tilde{\gamma}', \phi W)\tilde{\gamma}', \phi W \rangle(t), \end{aligned}$$

which proves (6.9). It follows that

$$\begin{aligned}I_\gamma(W,W) &= \int_0^a \{\langle W', W'\rangle - \langle R(\gamma', W)\gamma', W\rangle\}\,dt \\ &\leq \int_0^a \{\langle \phi W', \phi W'\rangle - \langle \tilde{R}(\tilde{\gamma}', \phi W)\tilde{\gamma}', \phi W\rangle\}dt \\ &= I_{\tilde{\gamma}}(\phi W, \phi W).\end{aligned}$$

Therefore, if $I_{\tilde{\gamma}}$ is negative definite on a subspace $\tilde{\mathcal{W}} \subset \tilde{\mathcal{V}}$, I_γ is negative definite on $\mathcal{W} := \phi^{-1}(\tilde{\mathcal{W}}) \subset \mathcal{V}$, which proves the lemma. $\square$

Let $\alpha, \beta : [0,1] \to M$ be piecewise smooth curves such that $\alpha(1) = \beta(0)$. For the next Lemma, we introduce the following notation:

$$\alpha^-(t) = \alpha(1-t), \qquad \alpha * \beta(t) = \begin{cases} \alpha(2t) & t \in [0, \frac{1}{2}] \\ \beta(2t-1) & t \in [\frac{1}{2}, 1]. \end{cases}$$

Lemma 6.3.2 (Klingenberg)**.** *Let M be a complete Riemannian manifold whose sectional curvature satisfies $K \leq K_0$, where K_0 is a positive constant. Let $p, q \in M$ and $\gamma_0 \neq \gamma_1$ be two geodesics joining p and q. Assume that there exists a homotopy α_s, $s \in [0,1]$ with $\gamma_0 = \alpha_0$ and $\gamma_1 = \alpha_1$ such that α_s is piecewise smooth for all $s \in [0,1]$ and the homotopy fixes the endpoints p and q. Then there exists $s_0 \in [0,1]$ such that $L(\gamma_0) + L(\alpha_{s_0}) \geq \frac{2\pi}{\sqrt{K_0}}$.*

Proof. We may assume that $L(\gamma_0) < \frac{\pi}{\sqrt{K_0}}$, otherwise we set $s_0 = 0$ and the proof is finished. Furthermore, we assume γ_0 to be defined on the unit interval and, without loss of generality, we may assume that $\gamma_0(\tilde{t}) \neq q$ for any $\tilde{t} < 1$. Otherwise, we can apply the lemma to the restricted curve $\gamma_0|_{[0,\tilde{t}]}$ and the proof is again finished. Let $\sigma : [0,a] \to M$ be a normalized geodesic with $\sigma(0) = p$. By Proposition 3.4.1, a value t_0, where $\sigma(t_0)$ is conjugate to p along σ, satisfies $t_0 \geq \frac{\pi}{\sqrt{K_0}}$. (Note that in the proof of this proposition, we have only used the estimate $K \leq H$ to show $\frac{\pi}{\sqrt{H}} \leq t_0$.) Therefore by Proposition 3.2.5, $\exp_p$ is regular at each $v \in T_pM$ which satisfies $|v| < \frac{\pi}{\sqrt{K_0}}$. It follows that $\exp_p : B_{\frac{\pi}{\sqrt{K_0}}}(0) \to B_{\frac{\pi}{\sqrt{K_0}}}(p) \subset M$ is a local diffeomorphism.

We assume also γ_1 to be defined on the unit interval. We denote the homotopy between γ_0 and γ_1 by H, so the map $H : [0,1] \times [0,1] \to M$ satisfies $\alpha_s(t) = H(s,t)$ and $H(s,0) = p$, $H(s,1) = q$ for all $s \in [0,1]$. The image of the map $\tilde{\gamma}_0 : t \mapsto t \cdot \gamma_0'(0)$ lies in $B_{\frac{\pi}{\sqrt{K_0}}}(0)$ and satisfies $\exp_p \circ \tilde{\gamma}_0 = \gamma_0$. We want to show that for small values of s, H can be lifted to a homotopy $\tilde{H}$ on T_pM.

Choose open sets $\tilde{U}_1, \ldots \tilde{U}_n \subset T_pM$ such that $\tilde{\gamma}_0([0,1]) \subset \tilde{U} := \tilde{U}_1 \cup \ldots \cup \tilde{U}_n$ and $\exp_p : \tilde{U}_i \to \exp_p(\tilde{U}_i)$ is a diffeomorphism for $i \in \{1, \ldots, n\}$. We have $\gamma_0([0,1]) \subset U := \exp_p(\tilde{U})$ and by continuity there exists $\delta > 0$, such that $H([0,\delta] \times [0,1]) \subset U$. We know that $\tilde{\gamma}_0([0,1]) \cap \exp_p^{-1}(q) = \{\gamma_0'(0)\}$, since $\gamma_0(t) \neq q$ for any $t < 1$. Therefore, it is possible to restrict $\tilde{U}$ in such a way that $\tilde{U} \cap \exp_p^{-1}(q) = \{\gamma_0'(0)\}$.

Define $\tilde{H} : [0,\delta] \times [0,1] \to T_pM$ by $\tilde{H}|_{\tilde{U}_i} = (\exp_p|_{\tilde{U}_i})^{-1} \circ H$. The map is well defined and satisfies $\exp_p \circ \tilde{H} = H|_{[0,\delta]\times[0,1]}$ and $\tilde{H}(s,0) = 0$ for all $s \in [0,\delta]$. Since $\exp_p(\tilde{H}(s,1)) = H(s,1) = q$ for all $s \in [0,\delta]$, we have $\tilde{H}([0,\delta] \times \{1\}) \subset$

$\tilde{U} \cap \exp_p^{-1}(q) = \{\gamma_0'(0)\}$. Therefore, we have $\tilde{H}(s,1) = \gamma_0'(0)$ for all $s \in [0,\delta]$, so the homotopy $\tilde{H}$ fixes the endpoints 0 and $\gamma_0'(0)$ in T_pM.

Let $S \subset [0,1]$ be the maximal interval containing 0 such that there exists a homotopy $\tilde{H}$ in T_pM with $\exp_p \circ \tilde{H} = H_{S\times[0,1]}$ which fixes the endpoints 0 and $\gamma_0'(0)$. We have $[0,\delta] \subset S$ and S is open in $[0,1]$: If $[0,s_1] \subset S$, we consider the curve $\tilde{\alpha}_{s_1} = \tilde{H}(s_1,.)$ in T_pM. In a manner analogous to the above, one shows that for s near s_1, α_s can be lifted to a path $\tilde{\alpha}_s$ in T_pM in such a way that $\tilde{\alpha}_s$ depends continuously on s. In other words, there exists $\delta_1 > 0$ such that $[0, s_1 + \delta_1] \subset S$.

Now we claim the following: For all $\epsilon > 0$, there exist $s_\epsilon \in S$ such that $\max_{t\in[0,1]} |\tilde{\alpha}_{s_\epsilon}(t)| > \frac{\pi}{\sqrt{K_0}} - \epsilon$. Suppose the contrary. Let $\{s_n\} \subset S$, $n \in \mathbb{N}$ be a sequence converging to $s \in [0,1]$. Then the corresponding curves $\tilde{\alpha}_{s_n}$ are contained in the compact set $\overline{B_{\frac{\pi}{\sqrt{K_0}}-\epsilon}(0)}$. There exists a subsequence, which we again denote by $\{s_n\}$, such that the sequence $\{\tilde{\alpha}_{s_n}\}$ converges uniformly to a curve $\tilde{\alpha}_s$. By continuity,

$$\exp_p(\tilde{\alpha}_s(t)) = \exp_p(\lim_{n\to\infty} \tilde{\alpha}_{s_n}(t)) = \lim_{n\to\infty} \exp_p(\tilde{\alpha}_{s_n}(t)) = \lim_{n\to\infty} \alpha_{s_n}(t) = \alpha_s(t).$$

which shows that $\tilde{\alpha}_s$ is a lift of the curve α_s. Therefore, $s \in S$, so S is closed. Since S is also open we have $S = [0,1]$. Thus, for all $s \in [0,1]$, we have that the corresponding curves $\tilde{\alpha}_s$ are in particular contained in $B_{\frac{\pi}{\sqrt{K_0}}}(0)$, in other words, $\tilde{H}([0,1]\times[0,1]) \subset B_{\frac{\pi}{\sqrt{K_0}}}(0)$.

Since $\exp_p \circ \tilde{H} = H$, we have that $\exp_p(\tilde{H}(1,t)) = H(1,t) = \gamma_1(t)$. Consider the straight line $\tilde{\gamma}_1 : t \mapsto t \cdot \gamma_1'(0)$ in T_pM. Obviously, we have $\exp_p(\tilde{\gamma}_1(t)) = \exp_p(\tilde{H}(1,t))$ for all $t \in [0,1]$. Since $\exp_p$ is a diffeomorphism around 0 and $\tilde{H}(1,0) = 0$, we have $\tilde{\gamma}_1(t) = \tilde{H}(1,t)$ for small values of t. Since $\tilde{H}(\{1\}\times[0,1]) \subset B_{\frac{\pi}{\sqrt{K_0}}}(0)$ and $\exp_p$ is a local diffeomorphism on $B_{\frac{\pi}{\sqrt{K_0}}}(0)$, this equality is even true for all $t \in [0,1]$. Therefore, $\gamma_0'(0) = \tilde{H}(0,1) = \tilde{H}(1,1) = \gamma_1'(0)$. This contradicts the assumption that $\gamma_0 \neq \gamma_1$. Hence we have proven the claim.

Next we show that

$$L(\gamma_0) + L(\alpha_{s_\epsilon}) \geq \frac{2\pi}{\sqrt{K_0}} - 2\epsilon. \tag{6.10}$$

Consider the curve $\beta = \alpha_{s_\epsilon} * \gamma_0^-$. It is clear that β is a loop, i.e. $\beta(0) = \beta(1) = p$ and that $L(\beta) = L(\gamma_0) + L(\alpha_{s_\epsilon})$. The curve $\tilde{\beta} = \tilde{\alpha}_{s_\epsilon} * \tilde{\gamma}_0^-$ satisfies $\exp_p \circ \tilde{\beta} = \beta$. Now consider M_{K_0}, the sphere of constant sectional curvature K_0 and $\dim(M_{K_0}) = \dim(M)$. Fix a point $\tilde{p} \in M_{K_0}$, a linear isometry $i : T_pM \to T_{\tilde{p}}M_{K_0}$ and consider the curve $\widehat{\beta} = \exp_{\tilde{p}}(i(\tilde{\beta}))$. Note that $\exp_{\tilde{p}}|_{B_{\frac{\pi}{\sqrt{K_0}}}(0)}$ is a diffeomorphism. Since $K \leq K_0$, we obtain from Proposition 3.4.3 that

$$L(\beta) \geq L(\widehat{\beta}). \tag{6.11}$$

Let $t_\epsilon \in [0,1]$ such that $d(\tilde{p}, \widehat{\beta}(t_\epsilon)) = \max_{t\in[0,1]} d(\tilde{p}, \widehat{\beta}(t))$. Then

$$L(\widehat{\beta}) \geq d(\widehat{\beta}(0), \widehat{\beta}(t_\epsilon)) + d(\widehat{\beta}(t_\epsilon), \widehat{\beta}(1)) = 2d(\tilde{p}, \widehat{\beta}(t_\epsilon)). \tag{6.12}$$

Since $\exp_{\tilde{p}}|_{B_{\frac{\pi}{\sqrt{K_0}}}(0)}$ is a diffeomorphism,

$$\begin{aligned} d(\tilde{p},\widehat{\beta}(t_\epsilon)) &= \max_{t\in[0,1]} d(\tilde{p},\widehat{\beta}(t)) \\ &= \max_{t\in[0,1]} |i(\tilde{\beta}(t))| \\ &= \max_{t\in[0,1]} |\tilde{\beta}(t)| \geq \max_{t\in[0,1]} |\tilde{\alpha}_{s_\epsilon}(t)| \geq \frac{\pi}{\sqrt{K_0}} - \epsilon. \end{aligned} \tag{6.13}$$

Putting the things above together, we obtain

$$L(\gamma_0)+L(\alpha_{s_\epsilon}) = L(\beta) \overset{(6.11)}{\geq} L(\widehat{\beta}) \overset{(6.12)}{\geq} 2d(\tilde{p},\widehat{\beta}(t_\epsilon)) \overset{(6.13)}{\geq} \frac{2\pi}{\sqrt{K_0}} - 2\epsilon,$$

which proves (6.10).

Now choose a sequence $\{\epsilon_n\}$, $\epsilon_n > 0$, $n \in \mathbb{N}$ converging to 0 and let $\{s_{\epsilon_n}\}$ be a sequence in $[0,1]$ such that the curve $\alpha_{s_{\epsilon_n}}$ satisfies (6.10) with $\epsilon = \epsilon_n$. By taking a subsequence of $\{s_{\epsilon_n}\}$ which converges to a value $s_0 \in [0,1]$, we obtain that

$$L(\gamma_0)+L(\alpha_{s_0}) \geq \frac{2\pi}{\sqrt{K_0}},$$

which concludes the proof. □

Theorem 6.3.3. *Let M^n, $n \geq 3$ be a simply connected Riemannian manifold whose sectional curvature K satisfies $\frac{1}{4} < K_{min} \leq K \leq 1$. Then $i(M) \geq \pi$.*

Proof. By the Theorem of Bonnet-Myers (Theorem 2.2.4), M is compact. Suppose that $i(M) < \pi$. Then, by Proposition 6.2.15, there exists a closed geodesic γ of length $L = L(\gamma) < 2\pi$. Consider γ to be normalized. By Corollary 4.2.4, the set of points which are conjugate to $p = \gamma(0)$ along γ is discrete. Choose an $\epsilon > 0$, satisfying the following conditions:

(i) $\gamma(L-\epsilon)$ is not conjugate to p along γ

(ii) $3\epsilon < 2\pi - \frac{\pi}{\sqrt{\tilde{K}}}$, where $\tilde{K} = K_{min}$.

(iii) $3\epsilon < 2\pi - L$

(iv) $5\epsilon < L$

By (i) and Proposition 3.2.5, $(L-\epsilon)\gamma'(0) \in T_pM$ is a regular value of $\exp_p$. Hence, there exists $\epsilon_1 > 0$ such that $\exp_p : B_{\epsilon_1}((L-\epsilon)\gamma'(0)) \to M$ is a diffeomorphism onto its image. We choose ϵ_1 so small that $\epsilon_1 \leq \epsilon$. Consider the set X of points v in T_pM where $T_v \exp_p$ is singular. By Sard's theorem (cf. [BG88], Theorem 4.3.1), $\exp_p(X) \subset M$ has Lebesgue measure 0. Therefore, there exists a regular value $q \in \exp_p(B_{\epsilon_1}((L-\epsilon)\gamma'(0)))$ with $d(\gamma(L-\epsilon),q) < \epsilon$. Let $w \in B_{\epsilon_1}((L-\epsilon)\gamma'(0))$ such that $\exp_p(w) = q$. Then,

$$|w| \leq |w-(L-\epsilon)\gamma'(0)| + |(L-\epsilon)\gamma'(0)| < \epsilon_1 + L - \epsilon \leq L$$

and by (iv),

$$3\epsilon < L-2\epsilon = |(L-\epsilon)\gamma'(0)|-\epsilon \leq |(L-\epsilon)\gamma'(0)-w|+|w|-\epsilon < \epsilon_1+|w|-\epsilon \leq |w|.$$

We obtain a geodesic $\gamma_1 : [0,1] \to M$, $\gamma_1(t) := \exp_p(tw)$ whose length $L(\gamma_1) = |w|$ satisfies $3\epsilon < L(\gamma_1) < L = L(\gamma)$. Let $\gamma_0 : [0,1] \to M$ be a minimizing geodesic joining p to q. Its length satisfies

$$\begin{aligned} L(\gamma_0) = d(p,q) &\leq d(p, \gamma(L-\epsilon)) + d(\gamma(L-\epsilon), q) \\ &= d(\gamma(L), \gamma(L-\epsilon)) + d(\gamma(L-\epsilon), q) < 2\epsilon, \end{aligned}$$

hence $\gamma_0 \neq \gamma_1$.

Recall from Section 5.3 the set of all curves in M which join p to q, denoted by $\Omega_{p,q} = \Omega$. Since M is simply connected, there exists a homotopy $\tilde{\alpha}_s$, $s \in [0,1]$ such that $\tilde{\alpha}_0 = \gamma_0$ and $\tilde{\alpha}_1 = \gamma_1$ which fixes the endpoints p and q. Since γ_0 and γ_1 are smooth, we may assume the homotopy to be piecewise smooth (c.f. [KN63, p. 284] and [MT97], Lemma 6.6). Therefore, the map $s \mapsto \tilde{\alpha}_s$ can be considered as a curve in Ω. Then for $c \in \mathbb{R}$ large enough, $\tilde{\alpha}_s \in \Omega^c$ for all $s \in [0,1]$. By Proposition 5.3.1, we can choose a decomposition $0 = t_0 < \ldots < t_k = 1$ of $[0,1]$, such that $\Omega(t_0, \ldots, t_k)^c$ can be given the structure of a finite dimensional manifold. We denote $V = \Omega(t_0, \ldots, t_k)^c$. By Proposition 5.3.2 V is a deformation retract of Ω^c. Thus, by Lemma 5.2.11, $\tilde{\alpha}_s$ can be deformed by a homotopy to a curve α_s in V, such that the homotopy fixes the endpoints $\gamma_0, \gamma_1 \in V$.

Consider the Energy function $E : V \to \mathbb{R}$. By Proposition 5.3.4 (i), $E \in C^\infty(V)$. Since q is a regular value of $\exp_p$, $\exp_p$ is regular at each $v \in \exp_p^{-1}(q)$. By Proposition 3.2.5, q is not conjugate to p on any geodesic which joins p to q. Therefore by Proposition 5.3.4 (iii), all critical points of E are non-degenerate, hence E is a Morse function. We now want to apply Corollary 5.2.12 to the curve α_s. Let

$$\begin{aligned} a &= \max\{E(\gamma_0), E(\gamma_1)\} \overset{2.1.3}{=} E(\gamma_1), \\ b &= \max_{s\in[0,1]} \{E(\alpha_s)\}, \end{aligned}$$

and let $\tilde{a}$ be the largest critical value in $[a,b]$ such that there exists a critical point of $E^{-1}(\tilde{a})$ of index zero or one. If such a value does not exist, we set $\tilde{a} = a$. Even in this case, $\tilde{a}$ is a critical point of E, since the unbroken geodesic γ_1 is, by Proposition 5.3.4 (ii), a critical point satisfying $E(\gamma_1) = a = \tilde{a}$.

Let $\epsilon_2 > 0$ so small that $b + \epsilon_2 < c$. We claim that $V^{[a,b+\epsilon_2]}$ is a compact subset of V. It is equivalent to show that $\varphi(V^{[a,b+\epsilon_2]})$ is a compact subset of $\varphi(V)$ where $\varphi : V \to \varphi(V) \subset M \times \ldots \times M$ denotes the homeomorphism from the proof of Proposition 5.3.1. In fact, we can express $\varphi(V^{[a,b+\epsilon]})$ explicitly as

$$\varphi(V^{[a,b+\epsilon]}) = \left\{ (r_1, \ldots, r_{k-1}) \in M \times \ldots \times M | a \leq \sum_{i=1}^{k} \frac{d(r_i, r_{i-1})^2}{t_i - t_{i-1}} \leq b + \epsilon \right\},$$

where $p = r_0$, $q = r_k$. Clearly, this set is bounded and closed, hence a compact subset of $M \times \ldots \times M$ and $\varphi(V^{[a,b+\epsilon]}) \subset \varphi(V)$. Therefore, $\varphi(V^{[a,b+\epsilon]})$ is a compact subset of $\varphi(V)$ which proves the claim.

Now we can use Corollary 5.2.12. For every $\delta > 0$, we can deform α_s to a curve β_s which satisfies $E(\beta_s) \leq \tilde{a} + \delta$ for all $s \in [0,1]$. Let $\sigma \in V$ be a critical point of E with $E(\sigma) = \tilde{a}$. By 5.3.4 (ii), σ is an unbroken geodesic. For $s \in [0,1]$, we have, by Lemma 2.1.3

$$L(\beta_s)^2 \leq E(\beta_s) \leq \tilde{a} + \delta = E(\sigma) + \delta = L(\sigma)^2 + \delta$$

and therefore,

$$L(\beta_s) \leq L(\sigma) + \tilde{\delta} \tag{6.14}$$

for a $\tilde{\delta} > 0$ satisfying $2L(\sigma)\tilde{\delta} + \tilde{\delta}^2 = \delta$. Choose $\delta > 0$ so small that $\tilde{\delta} \leq \epsilon$. We have already seen that $L(\gamma_0) \leq 2\epsilon$ and $3\epsilon < L(\gamma_1) < L$. To estimate $L(\sigma)$, we first consider the case where $\tilde{a} = a$. Then, $E(\sigma) = E(\gamma_1)$, and since both geodesics are defined on the interval $[0, 1]$, we obtain that

$$L(\sigma) \overset{2.1.3}{=} L(\gamma_1) < L \overset{(iii)}{<} 2\pi - 3\epsilon.$$

If $\tilde{a} > a$, $\text{ind}_E(\sigma) < 2$ and we compare σ with a curve $\tilde{\sigma}$ of the same length on the sphere $M_{\tilde{K}}$ of constant curvature $\tilde{K} = K_{min} > \frac{1}{4}$ and $\dim(M_{\tilde{K}}) = n = \dim(M)$. Then,

$$\text{ind} I_{\tilde{\sigma}} \overset{6.3.1}{\leq} \text{ind} I_{\sigma} \overset{5.3.4(iv)}{=} \text{ind}_E(\sigma) < 2. \tag{6.15}$$

If $L(\tilde{\sigma}) > \frac{\pi}{\sqrt{\tilde{K}}}$, $\tilde{\sigma}$ contains, by Example 3.2.4, points which are conjugate to $\tilde{p} = \tilde{\sigma}(0)$ along $\tilde{\sigma}$, and each conjugate point is of multiplicity $n - 1$. It follows from the Morse Index theorem (Theorem 4.2.1) that $\text{ind} I_{\tilde{\sigma}} \geq n - 1 \geq 2$ which contradicts (6.15).

Therefore by condition (ii), $L(\sigma) \leq \frac{\pi}{\sqrt{\tilde{K}}} < 2\pi - 3\epsilon$. Summing up, in both cases, we obtain

$$L(\beta_s) \overset{(6.14)}{\leq} L(\sigma) + \tilde{\delta} < 2\pi - 3\epsilon + \tilde{\delta} \leq 2\pi - 2\epsilon \tag{6.16}$$

for all $s \in [0, 1]$. On the other hand, by the Lemma of Klingenberg (Lemma 6.3.2), there exists a $s_0 \in [0, 1]$ such that $L(\gamma_0) + L(\beta_{s_0}) \geq 2\pi$. Therefore,

$$L(\beta_{s_0}) \geq 2\pi - L(\gamma_0) \geq 2\pi - 2\epsilon. \tag{6.17}$$

Since (6.16) and (6.17) are inconsistent, the existence of such a closed geodesic γ, which we deduced at the beginning of the proof, is not possible. Hence we have finished the proof of Theorem 6.3.3. □

6.4 The proof of the Sphere theorem

In this section, we will establish a sequence of lemmas before we are going to prove the Sphere theorem itself. In the proof, we will construct a homeomorphism explicitly. Throughout, all geodesics are assumed to be normalized. We follow Section 13.4 of [Car92].

Lemma 6.4.1. *Let M be a Riemannian manifold, $p \in M$ and $r > 0$ such that $\exp_p : B_r(0) \to B_r(p)$ is a diffeomorphism. Let $\alpha : (-\epsilon, \epsilon) \to M$ be a smooth curve in $B_r(p)$ such that $\alpha(0) \neq p$. Then*

$$\frac{d}{ds} d(p, \alpha(s))|_{s=0} = \langle \alpha'(0), \gamma'(L) \rangle,$$

where $\gamma : [0, L] \to M$ is the unique minimizing geodesic joining p to $\alpha(0)$.

Proof. Let $\bar{\alpha} = \exp_p^{-1} \circ \alpha$ and consider the variation of γ given by $x : (-\epsilon, \epsilon) \times [0, L] \to M$, $x(s,t) = \exp_p(\frac{t}{L}\bar{\alpha}(s))$. All longitudinal curves $x(s,.)$ are minimizing geodesics joining p to $\alpha(s)$. The variational vector field V of x along γ satisfies $V(0) = 0$ and $V(L) = \frac{\partial}{\partial s}x(0,L) = \alpha'(0)$. Thus, by the formula for the first variation (2.2),

$$\frac{d}{ds}d(p, \alpha(s))|_{s=0} = \frac{d}{ds}L(x(s,.))|_{s=0} = L_x'(0) = \langle V(L), \gamma'(L)\rangle = \langle \alpha'(0), \gamma'(L)\rangle.$$

□

Lemma 6.4.2 (Berger). *Let M be a compact Riemannian manifold and $p, q \in M$ such that $d(p,q) = \mathrm{diam}(M)$. Then for all $w \in T_pM$, there exists a minimizing geodesic γ joining $p = \gamma(0)$ to q such that $\langle \gamma'(0), w\rangle \geq 0$.*

Proof. Let $\lambda(t) := \exp_p(tw)$ and let $\gamma_t : [0, L_t] \to M$ be a minimizing geodesic joining $\lambda(t) = \gamma_t(0)$ to $q = \gamma_t(L_t)$. We claim that for all $n \in \mathbb{N}$, there exists a $t_n \in [0, \frac{1}{n}]$ such that $\langle \gamma_{t_n}'(0), \lambda'(t_n)\rangle \geq 0$.

Suppose to the contrary that there exists an $n \in \mathbb{N}$ such that for all $t \in [0, \frac{1}{n}]$, $\langle \gamma_t'(0), \lambda'(t)\rangle < 0$. Let $q_t \neq \lambda(t)$ be a point of γ_t such that $\lambda(t)$ is contained in a totally normal neighborhood of q_t. Then there exists $\epsilon > 0$, such that $\exp_{q_t} : B_\epsilon(0) \to B_\epsilon(q_t)$ is a diffeomorphism and $\lambda(t) \in B_\epsilon(p)$.

Since $q_t = \gamma_t(a)$ for a number $a \in [0, L_t]$, the geodesic $\tilde{\gamma}_t : [0, a] \to M$, $\tilde{\gamma}_t(r) = \gamma_t(a - r)$ is the unique minimizing geodesic joining q_t to $\lambda(t)$. By Lemma 6.4.1,

$$\frac{d}{ds}d(\lambda(s), q_t)|_{s=t} = \langle \tilde{\gamma}_t{}'(a), \lambda'(t)\rangle = -\langle \gamma_t'(0), \lambda'(t)\rangle > 0.$$

Therefore, there exists $\epsilon_t > 0$ such that for $s \in (t - \epsilon_t, t)$, $d(q_t, \lambda(s)) < d(q_t, \lambda(t))$ holds, hence

$$d(q, \lambda(s)) \leq d(q, q_t) + d(q_t, \lambda(s)) < d(q, q_t) + d(q_t, \lambda(t)) = d(q, \lambda(t)).$$

This shows that the function $s \mapsto d(q, \lambda(s))$, $s \in (t - \epsilon_t, t]$ is strictly monotonically increasing for every $t \in [0, \frac{1}{n}]$, so $s \mapsto d(q, \lambda(s))$ is strictly monotonically increasing on $[0, \frac{1}{n}]$. In particular, we have $d(q,p) = d(q, \gamma(0)) < d(q, \gamma(s))$ for any $s \in (0, \frac{1}{n}]$ which contradicts the assumption that $d(p,q) = \mathrm{diam}(M)$. Hence we have proven the claim.

Consider the sequence of tangent vectors $\gamma_{t_n}'(0) \in T_{\lambda(t_n)}M$ satisfying $t_n \in [0, \frac{1}{n}]$ and $\langle \gamma_{t_n}'(0), \lambda'(t_n)\rangle \geq 0$. Since the geodesics γ_t are normalized, $\{\gamma_{t_n}'(0)\}$ is contained in the compact set $\bigcup_{t\in[0,1]}(S^{n-1})_{\lambda(t)} \subset TM$, where $(S^{n-1})_{\lambda(t)}$ denotes the unit sphere in $T_{\lambda(t)}M$. By passing to a subsequence, we may assume that $\{\gamma_{t_n}'(0)\}$ converges to a vector $v \in (S^{n-1})_p \subset T_pM$. Set $L = d(p,q)$. The geodesic $\gamma : [0, L] \to M$, defined by $\gamma(t) = \exp_p(tv)$, satisfies

$$\gamma(L) = \exp_p(d(p,q)v) = \lim_{n\to\infty} \exp_{\lambda(t_n)}(d(\lambda(t_n), q)\gamma_{t_n}'(0)) = \lim_{n\to\infty} \gamma_{t_n}(L_{t_n}) = q,$$

so γ is a minimizing geodesic joining p to q. Moreover,

$$\langle \gamma'(0), w\rangle = \lim_{n\to\infty} \langle \gamma_{t_n}'(0), \lambda'(t_n)\rangle \geq 0,$$

which finishes the proof. □

Lemma 6.4.3 (Spherical law of cosines). *Let M_K be a sphere of constant curvature $K > 0$ and $\sigma_i : [0, L_i] \to M_K$, $i = 1, 2, 3$ be minimizing geodesics such that*

$$\sigma_1(0) = \sigma_3(L_3), \qquad \sigma_2(0) = \sigma_1(L_1), \qquad \sigma_3(0) = \sigma_2(L_2),$$

i.e. the geodesics form a triangle in M_K, which we denote by T. Let $a = L_1$, $b = L_2$, $c = L_3$ be the side lengths of T and

$$\alpha = \measuredangle(-\sigma_2'(L_2), \sigma_3'(0)), \qquad \beta = \measuredangle(-\sigma_3'(L_3), \sigma_1'(0)), \qquad \gamma = \measuredangle(-\sigma_1'(L_1), \sigma_2'(0))$$

be the interior angles of T. Then

$$\cos(\sqrt{K}c) = \cos(\sqrt{K}a)\cos(\sqrt{K}b) + \sin(\sqrt{K}c)\sin(\sqrt{K}c)\cos(\gamma). \tag{6.18}$$

Remark 6.4.4. If M_K is of constant curvature $K = 0$, $K < 0$, then the analogous expressions of (6.18) are

$$c^2 = a^2 + b^2 - 2ab\cos(\gamma),$$
$$\cosh(\sqrt{|K|}c) = \cosh(\sqrt{|K|}a)\cosh(\sqrt{|K|}b) + \sinh(\sqrt{|K|}c)\sinh(\sqrt{|K|}c)\cos(\gamma).$$

Proof. We prove (6.18) for $K = 1$, the general case follows from multiplying the metric with a positive constant. Consider M_1 as the unit sphere in $\mathbb{R}^{n+1}$ with the usual metric. Let $u = \sigma_3(0)$, $v = \sigma_1(0)$ and $w = \sigma_2(0)$ be the vertices of the triangle. Note that σ_1 is a segment of the unit circle in a plane which contains v and w. Since σ_1 joins the unit vectors v and w and $a = L(\sigma_1)$, we obtain $\cos(a) = \langle v, w\rangle$. Analogously,

$$\cos(b) = \langle u, w\rangle \qquad \cos(c) = \langle u, v\rangle.$$

Now we are going to show that

$$\sin(b)\sigma_2'(0) = u - w\cos(b). \tag{6.19}$$

If u and w are linearly dependent, then $u = \pm w$ and $b = 0$ or $b = \pi$. In both cases, (6.19) is satisfied trivially. Now we suppose that u and w are linearly independent. Then the tangent vector $\sigma_2'(0)$ at w is the unit vector perpendicular to w in the $u - w$ plane whose direction is given by the component of u perpendicular to w. Thus,

$$\sigma_2'(0) = \frac{u - w\langle u, w\rangle}{|u - w\langle u, w\rangle|} = \frac{u - w\cos(b)}{\sqrt{1 - \cos(b)^2}} = \frac{u - w\cos(b)}{\sin(b)}.$$

Similarly,

$$\sin(a)(-\sigma_1'(L_1)) = v - w\cos(a). \tag{6.20}$$

It follows that

$$\begin{aligned}\sin(a)\sin(b)\cos(\gamma) &= \sin(a)\sin(b)\langle -\sigma_1'(L_1), \sigma_2'(0)\rangle \\ &= \langle v - w\cos(a), u - w\cos(b)\rangle = \cos(c) - \cos(a)\cos(b),\end{aligned}$$

which finishes the proof. $\square$

Lemma 6.4.5. *Let M^n $(n \geq 3)$ be a simply connected Riemannian manifold whose sectional curvature K satisfies*

$$\frac{1}{4} < \delta \leq K \leq 1$$

and let $p, q \in M$ be such that $d(p,q) = \operatorname{diam}(M)$. Then for any ρ with $\frac{\pi}{2\sqrt{\delta}} < \rho < \pi$,

$$M = B_\rho(p) \cup B_\rho(q)$$

and $B_\rho(p), B_\rho(q)$ are both diffeomorphic to a Euclidean ball.

Proof. By the estimate of the injectivity radius (Theorem 6.3.3), $\rho < i(M)$. By Corollary 6.2.6, the map $\exp_r : B_\rho(0) \to B_\rho(r)$ is a bijection for each $r \in M$. By Proposition 3.4.1, none of the points $r_1 \in B_\rho(r)$ is conjugate to r along the minimizing geodesic joining r to r_1. Therefore, by Proposition 3.2.5, each point $v \in B_\rho(0) \subset T_r M$ is a regular point of $\exp_r$, so $\exp_r$ is a local diffeomorphism at v. This shows that $\exp_r : B_\rho(0) \to B_\rho(r)$ is a diffeomorphism for each $r \in M$, in particular for p and q. It is at this point that the estimate $i(M) \geq \pi$ enters, in a crucial manner, in the Sphere theorem.

Suppose there exists a point $r \in M$ such that $d(p,r) \geq \rho$ and $d(q,r) \geq \rho$. We may assume that $d(p,r) \geq d(q,r)$. Let $\mu_1 : [0, L_1] \to M$ be a minimizing geodesic joining $q = \mu_1(0)$ to r. Then for $q' = \mu_1(\rho) \in \partial B_\rho(q)$, $d(p,q') \geq \rho$, otherwise,

$$d(q,r) = d(q,q') + d(q',r) = \rho + d(q',r) > d(p,q') + d(q',r) \geq d(p,r),$$

which contradicts the assumption that $d(p,r) \geq d(q,r)$. On the other hand, by the Theorem of Bonnet-Myers (Theorem 2.2.4), M is compact and $\operatorname{diam}(M) \leq \frac{\pi}{\sqrt{\delta}} < 2\rho$. Let $\mu_2 : [0, L_2] \to M$ be a minimizing geodesic joining $q = \mu_2(0)$ to p and $q'' = \mu_2(\rho) \in \partial B_\rho(q)$, then

$$d(p,q'') = d(p,q) - d(q,q'') < 2\rho - \rho = \rho.$$

Since $B_\rho(q)$ is homeomorphic to a Euclidean ball, $\partial B_\rho(q)$ is path-connected. We have found points $q', q'' \in \partial B_\rho(q)$ satisfying $d(p,q') \geq \rho$ and $d(p,q'') < \rho$. By continuity, there exists a point $r_0 \in \partial B_\rho(q)$ such that

$$d(p,r_0) = \rho = d(q,r_0).$$

Consider a minimizing geodesic λ joining p to r_0. By Lemma 6.4.2, there exists a minimizing geodesic γ joining p to q such that $\langle \gamma'(0), \lambda'(0) \rangle \geq 0$. We denote by s the point on γ such that $d(p,s) = \rho$.

We now want to estimate the distance between r_0 and s. It would be possible to compare the triangle, spanned by λ and γ, with a triangle on M_δ^2, the sphere of dimension 2 and constant curvature δ. By the Spherical law of cosines, we would obtain an estimate of the distances on M_δ^2 and by Toponogov's theorem (Theorem 3.4.4), we would get an upper bound for $d(s, r_0)$. However, since we have not proven Toponogov's theorem in this thesis, we give a different approach here, using Proposition 3.4.3.

Consider the sphere of constant curvature δ and dimension n, denoted by M_δ. Choose a point $\tilde{p} \in M_\delta$ and fix a linear isometry $i : T_p M \to T_{\tilde{p}} M_\delta$. Choose

$\epsilon > 0$ so small that $\rho + \epsilon < \pi \leq i(M)$. Then, like in the first part of the proof, one shows that $\exp_p : B_{\rho+\epsilon}(0) \to B_{\rho+\epsilon}(p)$ is a diffeomorphism. Let

$$\begin{aligned}
\tilde{r}_0 &= \exp_{\tilde{p}} \circ i \circ (\exp_p |_{B_{\rho+\epsilon}(0)})^{-1}(r_0),\\
\tilde{s} &= \exp_{\tilde{p}} \circ i \circ (\exp_p |_{B_{\rho+\epsilon}(0)})^{-1}(s),\\
\bar{r}_0 &= i \circ (\exp_p |_{B_{\rho+\epsilon}(0)})^{-1}(r_0),\\
\bar{s} &= i \circ (\exp_p |_{B_{\rho+\epsilon}(0)})^{-1}(s),\\
\theta &= \measuredangle(\bar{r}_0, \bar{s}) = \measuredangle(\lambda'(0), \gamma'(0)) \leq \frac{\pi}{2}.
\end{aligned}$$

We will give an explicit construction of curves $\alpha, \tilde{\alpha}$ in M, M_δ, respectively, to which we can apply Proposition 3.4.3.

It is clear that $\bar{s}, \bar{r}_0 \in \partial B_\rho(0) \subset T_{\tilde{p}} M_\delta$. Let now $\bar{\alpha}$ be the shortest curve in $\partial B_\rho(0)$ joining $\bar{r}_0$ to $\bar{s}$. Then $\bar{\alpha}$ is a segment of the circle of radius ρ in the plane generated by $\bar{s}$ and $\bar{r_0}$. We assume $\bar{\alpha}$ to be parameterized by angle, so $\bar{\alpha} : [0, \theta] \to \partial B_\rho(0)$ and $|\bar{\alpha}'(u)| = \rho$ for all $u \in [0, \theta]$. Consider the curve $\tilde{\alpha} = \exp_{\tilde{p}} \circ \bar{\alpha}$ joining $\tilde{r}_0$ and $\tilde{s}$ and the geodesic variation $x : [0, \theta] \times [0, \rho] \to M_\delta$, defined as $x(u, t) = \exp_{\tilde{p}}(\frac{t}{\rho}\bar{\alpha}(u))$. Let $\tilde{\lambda}_u = x(u, .)$ and $\tilde{J}_u(t) = \frac{\partial x}{\partial u}(u, t)$. By Proposition 3.1.3, $\tilde{J}_u$ is a Jacobi field along $\tilde{\lambda}_u$ for all $u \in [0, \theta]$. We calculate

$$\tilde{J}_u(0) = \frac{\partial x}{\partial u}(u, t)|_{t=0} = T_{\frac{t}{\rho}\bar{\alpha}(u)} \exp_{\tilde{p}}(\frac{t}{\rho}\bar{\alpha}'(u))|_{t=0} = 0$$

and

$$\tilde{J}_u'(0) = \frac{D}{dt}\frac{\partial x}{\partial u}(u, 0) = \frac{D}{du}\frac{\partial x}{\partial t}(u, 0) = \frac{D}{du} T_0 \exp_{\tilde{p}}(\frac{\bar{\alpha}(u)}{\rho}) = \frac{D}{du}\frac{\bar{\alpha}(u)}{\rho} \overset{3.4.2}{=} \frac{\partial}{\partial u}\frac{\bar{\alpha}(u)}{\rho}.$$

Furthermore, we have

$$\tilde{J}_u(\rho) = \frac{\partial}{\partial u} x(u, \rho) = \tilde{\alpha}'(u)$$

and, by Lemma 6.4.1,

$$0 = \frac{d}{du} d(p, \tilde{\alpha}(u)) = \langle \tilde{\alpha}'(u), \tilde{\lambda}_u'(\rho) \rangle = \langle \tilde{J}_u(\rho), \tilde{\lambda}_u'(\rho) \rangle.$$

Since also $\langle \tilde{J}_u(0), \tilde{\lambda}_u'(0) = 0$, it follows from Lemma 3.1.5 that $\tilde{J}_u$ is orthogonal to $\tilde{\lambda}_u$ for each $u \in [0, \rho]$. Since each geodesic $\tilde{\lambda}_u$ is normalized, we obtain from Proposition 3.1.8

$$\tilde{J}_u(t) = \frac{\sin(\sqrt{\delta}t)}{\sqrt{\delta}} W_u(t)$$

where W_u is the parallel vector field along $\tilde{\lambda}_u$ such that $W_u(0) = \tilde{J}_u'(0)$. Therefore,

$$|\tilde{\alpha}'(u)| = |\tilde{J}_u(\rho)| \leq \frac{1}{\sqrt{\delta}} |\tilde{J}_u'(0)| = \frac{1}{\sqrt{\delta}\rho} |\bar{\alpha}'(u)| = \frac{1}{\sqrt{\delta}}.$$

By integrating, we get

$$L(\tilde{\alpha}) = \int_0^\theta |\tilde{\alpha}'(u)| du \leq \int_0^\theta \frac{1}{\sqrt{\delta}} du = \frac{\theta}{\sqrt{\delta}} \leq \frac{\pi}{2\sqrt{\delta}}.$$

Now let $\alpha = \exp_p \circ i^{-1} \circ \bar{\alpha}$. Since $\exp_p : B_{\rho+\epsilon}(0) \to B_{\rho+\epsilon}(p)$ is in particular a local diffeomorphism, we can apply Proposition 3.4.3 to get

$$d(r_0, s) \leq L(\alpha) \leq L(\tilde{\alpha}) \leq \frac{\pi}{2\sqrt{\delta}}$$

which gives the estimate we wanted.

Let $s_0 \in \gamma$ be a point such that $d(r_0, s_0) = \min_{t\in[0,L(\gamma)]} d(r_0, \gamma(t))$. Then we have

$$d(r_0, s_0) \leq d(r_0, s) \leq \frac{\pi}{2\sqrt{\delta}} < \rho,$$

so $\gamma(t_0) = s_0$ for a $t_0 \in (0, L(\gamma))$ and s_0 is in the image of the diffeomorphic map $\exp_{r_0} : B_\rho(0) \to B_\rho(r_0)$. The unique minimizing geodesic $\sigma_1 : [0, L(\sigma_1)] \to M$ joining $\sigma_1(0) = s_0$ to $\sigma_1(L(\sigma_1)) = r_0$ is orthogonal to γ at s_0 since by Lemma 6.4.1,

$$0 = \frac{d}{dt} d(r_0, \gamma(t))|_{t=t_0} = \langle -\sigma_1'(0), \gamma'(t_0)\rangle.$$

Since $d(p, q) \leq \frac{\pi}{\sqrt{\delta}}$, we have either $d(p, s_0) \leq \frac{\pi}{2\sqrt{\delta}}$ or $d(q, s_0) \leq \frac{\pi}{2\sqrt{\delta}}$. In either of the two cases, we obtain the desired contradiction. Consider the case $d(p, s_0) \leq \frac{\pi}{2\sqrt{\delta}}$, the other case works analogously. Let $\sigma_2 : [0, L(\sigma_2)] \to M$ be defined as $\sigma_2(t) = \gamma(t_0 - t)$, then $\sigma_2(0) = \gamma(t_0) = s_0$, $\sigma_2(L(\sigma_2)) = \gamma(0) = p$ and σ_1 and σ_2 are orthogonal at s_0.

We now proceed in a similar manner as before to estimate the distance between p and r_0. Let M_δ be as above. Fix a point $\tilde{s}_0 \in M_\delta$, a linear isometry $i : T_{s_0}M \to T_{\tilde{s}_0}M_\delta$ and let

$$\begin{aligned}
\tilde{r}_0 &= \exp_{\tilde{s}_0} \circ i \circ (\exp_{s_0}|_{B_\rho(0)})^{-1}(r_0),\\
\tilde{p} &= \exp_{\tilde{s}_0} \circ i \circ (\exp_{s_0}|_{B_\rho(0)})^{-1}(p),\\
\tilde{\sigma}_1 &= \exp_{\tilde{s}_0} \circ i \circ (\exp_{s_0}|_{B_\rho(0)})^{-1} \circ \sigma_1,\\
\tilde{\sigma}_2 &= \exp_{\tilde{s}_0} \circ i \circ (\exp_{s_0}|_{B_\rho(0)})^{-1} \circ \sigma_2,\\
\theta &= \measuredangle(\tilde{\sigma}_1'(0), \tilde{\sigma}_2'(0)) = \measuredangle(\sigma_1'(0), \sigma_2'(0)) = \frac{\pi}{2}.
\end{aligned}$$

It is clear that $\tilde{\sigma}_1, \tilde{\sigma}_2$ are minimizing geodesics joining $\tilde{s}_0$ to $\tilde{r}_0, \tilde{p}$, respectively. Let $\tilde{\alpha}$ be a minimizing geodesic joining $\tilde{r}_0$ to $\tilde{p}$ and let $a = L(\tilde{\sigma}_1) = L(\sigma_1) \leq \frac{\pi}{2\sqrt{\delta}}$, $b = L(\tilde{\sigma}_2) = L(\sigma_2) \leq \frac{\pi}{2\sqrt{\delta}}$ and $c = L(\tilde{\alpha})$. By the Spherical law of cosines (Lemma 6.4.3),

$$\begin{aligned}
\cos(\sqrt{\delta}c) &= \cos(\sqrt{\delta}a)\cos(\sqrt{\delta}b) + \sin(\sqrt{\delta}a)\sin(\sqrt{\delta}b)\cos(\theta)\\
&= \cos(\sqrt{\delta}a)\cos(\sqrt{\delta}b) \geq 0
\end{aligned}$$

by the estimates of a and b. This shows that $L(\tilde{\alpha}) = c \leq \frac{\pi}{2\sqrt{\delta}}$. We claim that $\tilde{\alpha} \subset B_\rho(\tilde{p})$. Let β_t be a minimizing geodesic joining $\tilde{s}_0$ to $\tilde{\alpha}(t)$. Then $\tilde{\sigma}_1$, $\tilde{\alpha}|_{[0,t]}$ and β_t form a geodesic triangle, whose side lengths are given by a, t and $b_t = L(\beta_t)$. Again by Lemma 6.4.3,

$$\cos(\sqrt{\delta}b_t) = \cos(\sqrt{\delta}a)\cos(\sqrt{\delta}t) + \sin(\sqrt{\delta}a)\sin(\sqrt{\delta}t)\cos(\psi)$$

where $\psi = \measuredangle(-\sigma_1'(a), \tilde{\alpha}'(0))$. Thus, $f : [0, c] \to \mathbb{R}$, $f(t) = \cos(\sqrt{\delta}b_t)$ is, by the form of the right hand side, equal to zero or of the form $d_1\cos(\sqrt{\delta}t + d_2)$ for some constants d_1, d_2. Therefore, since $f(0) = \cos(\sqrt{\delta}a) \geq 0$, $f(c) = \cos(\sqrt{\delta}b) \geq 0$ and $c = L(\tilde{\alpha}) \leq \frac{\pi}{2\sqrt{\delta}}$, it follows that $f(t) \geq 0$ for all $t \in [0, c]$ and $b_t = L(\beta_t) = d(\tilde{s}_0, \tilde{\alpha}(t)) \leq \frac{\pi}{2\sqrt{\delta}} < \rho$. This proves the claim.

Again we compare $\tilde{\alpha}$ with a curve α in M, defined by

$$\alpha = \exp_p \circ i^{-1} \circ (\exp_{\tilde{p}} |_{B_\rho})^{-1} \circ \tilde{\alpha}.$$

By Proposition 3.4.3,

$$d(p, r_0) \leq L(\alpha) \leq L(\tilde{\alpha}) \leq \frac{\pi}{2\sqrt{\delta}} < \rho$$

which contradicts the fact that $r_0 \in \partial B_\rho(p)$. □

Remark 6.4.6. In what follows, we do not use the assumptions on the curvature anymore. In fact, if it is possible to cover a manifold M by two balls, like we did in the previous lemma, then M is homeomorphic to a sphere.

Lemma 6.4.7. *Under the assumptions of Lemma 6.4.5, on each geodesic of length ρ starting from p there exists a unique point $m \in \gamma((0, \rho))$ such that*

$$d(p, m) = d(q, m) < \rho.$$

Similarly, on each geodesic σ of length ρ, starting from q, there exists a unique point of σ, which is equidistant from p and q.

Proof. Let $\gamma : [0, \rho] \to M$ be a geodesic with $\gamma(0) = p$. Consider the continuous function

$$f_\gamma(s) = d(q, \gamma(s)) - d(p, \gamma(s)).$$

Since by Theorem 6.3.3, $i(M) \geq \pi > \rho$, γ is minimizing, i.e. $d(p, \gamma(\rho)) = \rho$. By Lemma 6.4.5, $\gamma(\rho) \in B_\rho(q)$, so $d(q, \gamma(\rho)) < \rho$. Together we have

$$f_\gamma(0) = d(q, p) > 0, \qquad f_\gamma(\rho) = d(q, \gamma(\rho)) - d(p, \gamma(\rho)) < 0$$

and by continuity, there exists an $s_0 \in (0, \rho)$ such that $f(s_0) = 0$. Thus, $m = \gamma(s_0)$ satisfies $d(p, m) = d(q, m)$. By Lemma 6.4.5, it is clear that $m \in B_\rho(p) \cap B_\rho(q)$.

To show uniqueness of such a point, we follow a slightly different approach than [Car92]. For $m = \gamma(s_0) \in B_\rho(p) \cap B_\rho(q)$, let $\sigma : [0, L] \to M$ be the unique minimizing geodesic joining q to $m = \sigma(L)$. By Lemma 6.4.1,

$$\frac{d}{ds} d(q, \gamma(s))|_{s=s_0} = \langle \gamma'(s_0), \sigma'(L) \rangle \leq 1, \tag{6.21}$$

since $|\gamma'(s_0)| = |\sigma'(L)| = 1$. We are going to show that $\langle \gamma'(s_0), \sigma'(L) \rangle < 1$, so (6.21) is even satisfied with strict inequality.

We have equality in (6.21) if and only if $\gamma'(s_0) = \sigma'(L)$. In this case, $\sigma(L + s) = \gamma(s_0 + s)$ which shows that the three points p, q and m all lie on the same minimizing geodesic. Suppose that $L \leq s_0$, then

$$\begin{aligned} d(p, q) = d(\gamma(0), \sigma(0)) &= d(\gamma(0), \gamma(s_0 - L)) \\ &= d(\gamma(s_0), \gamma(0)) - d(\gamma(s_0), \gamma(s_0 - L)) \\ &< d(\gamma(s_0), \gamma(0)) = d(m, p), \end{aligned}$$

which contradicts the fact that $d(p,q) = \mathrm{diam}(M)$. If $L \geq s_0$, one shows in an analogous manner that $d(p,q) < d(m,q)$, which again leads to a contradiction.

Therefore, by (6.21),

$$\begin{aligned} \frac{d}{ds} f_\gamma(s)|_{s=s_0} &= \left(\frac{d}{ds} d(q,\gamma(s)) - \frac{d}{ds} d(p,\gamma(s)) \right) |_{s=s_0} \\ &= \langle \gamma'(s_0), \sigma'_{s_0}(L) \rangle - 1 < 0. \end{aligned} \tag{6.22}$$

It follows that f_γ is strictly monotonically decreasing around s, whenever $\gamma(s)$ lies in $B_\rho(p) \cap B_\rho(q)$. This proves the uniqueness of a zero point s_0 of f_γ, since $\gamma(s_0) \in B_\rho(p) \cap B_\rho(q)$. □

We now define a function $h_1 : (S^{n-1})_p \to M$, which associates to each unit vector $v \in T_pM$ the point $m \in M$ along the geodesic with initial value p and initial velocity v, which is equidistant from p and q. By Lemma 6.4.7, h_1 is well defined. Analogously, we define a function $h_2 : (S^{n-1})_q \to M$.

Lemma 6.4.8. *The maps $h_1 : (S^{n-1})_p \to M$ and $h_2 : (S^{n-1})_q \to M$ are homeomorphisms onto their images. In addition, their images are equal, i.e.*

$$h_1((S^{n-1})_p) = h_2((S^{n-1})_q).$$

Proof. Consider the function $g : M \to \mathbb{R}$, given by $g(r) = d(q,r) - d(p,r)$. Note that g is smooth on $B_\rho(p) \cap B_\rho(q)$. We are going to show that

$$N = g^{-1}(0) = \{r \in M | d(p,r) = d(q,r)\}$$

is an $n-1$ dimensional compact submanifold of M. By Lemma 6.4.5, $N \subset B_\rho(p) \cap B_\rho(q)$. We show that g is regular on $B_\rho(p) \cap B_\rho(q)$. Let $r \in B_\rho(p) \cap B_\rho(q)$ and let γ be the unique minimizing geodesic joining p to $r = \gamma(s_0)$. Then

$$T_r g(\gamma'(s_0)) = (g \circ \gamma)'(s_0) = f'_\gamma(s_0) \overset{(6.22)}{\neq} 0$$

where f_γ denotes the function of Lemma 6.4.7. Therefore, $T_r g$ has the maximal rank 1, so g is regular at r. It follows that N is a closed submanifold of M with $\dim(N) = n-1$. Since M is compact, N is also compact. Now we show that

$$h_1((S^{n-1})_p) = h_2((S^{n-1})_q) = N.$$

It clear that $h_1((S^{n-1})_p) \subset N$, $h_2((S^{n-1})_q) \subset N$. Conversely, if $r \in N$, there exist unique minimizing geodesics γ, σ joining p and q to r, respectively. Clearly, $h_1(\gamma'(0)) = r = h_2(\sigma'(0))$. Therefore $h_1 : (S^{n-1})_p \to N$ and $h_2 : (S^{n-1})_q \to N$ are bijective. It remains to show that they are homeomorphisms.

We prove this for h_1, for h_2 the proof is completely analogous. Consider the compact submanifold $\tilde{N} = (\exp_p|_{B_\rho(0)})^{-1}(N) \subset T_pM$ and the continuous map $\tilde{h}_1 : \tilde{N} \to (S^{n-1})_p$, given by $\tilde{h}_1(v) = \frac{v}{|v|}$. We show that $\tilde{h}_1 \circ (\exp_p|_{\tilde{N}})^{-1} = h_1^{-1}$. Let $r \in N$. As above, let γ be the unique minimizing geodesic joining p to $r = \gamma(s_0)$. Then

$$\tilde{h}_1 \circ (\exp_p|_{\tilde{N}})^{-1}(r) = \tilde{h}_1(s_0\gamma'(0)) = \gamma'(0) = h_1^{-1}(r).$$

This shows that $\tilde{h}_1 = h_1^{-1} \circ \exp_p|_{\tilde{N}} : \tilde{N} \to (S^{n-1})_p$ is a continuous bijection. Since $\tilde{N}$ and $(S^{n-1})_p$ both are compact, $\tilde{h}_1$ is a homeomorphism. Therefore, $h_1 = \exp_p|_{\tilde{N}} \circ \tilde{h}_1^{-1}$ is also a homeomorphism, which finishes the proof. □

Remark 6.4.9. With some more work, one could show that the functions h_1 and h_2 are even diffeomorphisms but we do not need this fact.

Theorem 6.4.10 (Sphere theorem). *Let M^n, $n \geq 3$ be a simply connected, complete Riemannian manifold, whose sectional curvature K satisfies*

$$\frac{1}{4} < \delta \leq K \leq 1.$$

Then M is homeomorphic to a sphere.

Proof. Let the functions g, h_1 and h_2 be as above. Consider the closed sets $D_1 = g^{-1}([0,\infty))$ and $D_2 = g^{-1}((-\infty, 0])$. It is clear that $D_1 \cup D_2 = M$ and $D_1 \cap D_2 = N = g^{-1}(0)$. We claim that

$$D_1 = \left\{\exp_p(tv) | v \in (S^{n-1})_p, 0 \leq t \leq d(p, h_1(v))\right\}, \tag{6.23}$$

$$D_2 = \left\{\exp_q(tv) | v \in (S^{n-1})_q, 0 \leq t \leq d(q, h_2(v))\right\}. \tag{6.24}$$

We show the first equality, the other one works analogously. Let $r \in D_1$, so $d(p,r) \leq d(q,r)$. By Lemma 6.4.5, $r \in B_\rho(p)$ and there exists a unique minimizing geodesic γ joining p to $r = \gamma(s_0)$. Consider the function f_γ of Lemma 6.4.7, which is, by continuity, strictly positive before its unique zero point and strictly negative after its zero point. The zero point of f_γ is given by $d(p, h_1(w))$, where $w = \gamma'(0)$, since

$$f_\gamma(d(p, h_1(w))) = g(\gamma(d(p, h_1(w)))) = g(h_1(w)) = 0.$$

Since $r = \gamma(s_0) \in D_1$, $f_\gamma(s_0) \geq 0$, so $s_0 \leq d(p, h_1(w))$ which proves that r is contained in the right hand side of (6.23). Conversely, let $r = \gamma(s_0)$ be a point of the geodesic $\gamma(s) = \exp_p(sw)$ such that $s_0 \leq d(p, h_1(w))$. Then $g(r) = f_\gamma(s_0) \geq 0$, since f_γ is strictly positive before its zero point $d(p, h_1(w))$. This shows that $r \in D_1$.

Now we give an explicit construction of the homeomorphism $\varphi : S^n \to M$. Fix a point $\tilde{p}$ in S^n and a linear isometry $i : T_{\tilde{p}}S^n \to T_pM$. We denote the antipodal point of $\tilde{p}$ by $\tilde{q}$. Recall that the sphere consists of all minimizing geodesics joining $\tilde{p}$ to $\tilde{q}$, which are all of length π. Therefore, we have

$$S^n = \left\{\exp_{\tilde{p}}(tv) | v \in (S^{n-1})_{\tilde{p}}, 0 \leq t \leq \pi\right\}.$$

Let S_1 be northern hemisphere of S^n, relative to $\tilde{p}$, S_2 be the southern hemisphere of S^n and E be the equator of S^n. These sets are given by

$$S_1 = \left\{\exp_{\tilde{p}}(tv) | v \in (S^{n-1})_{\tilde{p}}, 0 \leq t \leq \frac{\pi}{2}\right\}, \tag{6.25}$$

$$S_2 = \left\{\exp_{\tilde{p}}(tv) | v \in (S^{n-1})_{\tilde{p}}, \frac{\pi}{2} \leq t \leq \pi\right\}, \tag{6.26}$$

$$E = \left\{\exp_{\tilde{p}}(tv) | v \in (S^{n-1})_{\tilde{p}}, t = \frac{\pi}{2}\right\}. \tag{6.27}$$

Note that for $\tilde{r} \in S^n$, $\tilde{r} \neq \tilde{q}$, there exists a unique vector $v \in (S^{n-1})_{\tilde{p}}$ and a unique value $t \in [0, \pi)$, such that $\tilde{r} = \exp_{\tilde{p}}(tv)$. If $\tilde{r} = \tilde{q}$, then for all $v \in (S^{n-1})_{\tilde{p}}$, $\exp_{\tilde{p}}(\pi v) = \tilde{q}$. We define φ by

$$\varphi(\exp_{\tilde{p}}(tv)) = \begin{cases} \exp_p(t\frac{2}{\pi}d(p, h_1(i(v)))i(v)) & 0 \leq t \leq \frac{\pi}{2}, \\ \exp_q((2 - \frac{2t}{\pi})d(q, h_1(i(v)))h_2^{-1}(h_1(i(v)))) & \frac{\pi}{2} \leq t \leq \pi. \end{cases}$$

Step (A) $\boxed{\varphi \text{ is well defined}}$

Since a point $\exp_{\tilde{p}}(tv)$ is uniquely determined by t, v unless $t = 0$ or $t = \pi$, we have to show that φ is well defined in these cases. In fact, $\varphi(\tilde{p}) = \varphi(\exp_{\tilde{p}}(0v)) = \exp_p(0) = p$ and $\varphi(\tilde{q}) = \varphi(\exp_{\tilde{p}}(\pi v)) = \exp_q(0) = q$ for all $v \in (S^{n-1})_{\tilde{p}}$. Next, we show that the two definitions of φ coincide if $t = \frac{\pi}{2}$. Let $v \in (S^{n-1})_{\tilde{p}}$ be arbitrary and $t = \frac{\pi}{2}$. Then

$$\begin{aligned}
\varphi(\exp_{\tilde{p}}(\frac{\pi}{2}v)) &= \exp_q(d(q, h_1(i(v)))h_2^{-1}(h_1(i(v)))) \\
&= \exp_q(d(q, h_2 \circ h_2^{-1} \circ h_1(i(v)))h_2^{-1}(h_1(i(v)))) \\
&= h_2 \circ h_2^{-1} \circ h_1(i(v)) \\
&= h_1(i(v)) = \exp_p(d(p, h_1(i(v)))i(v))
\end{aligned}$$

so the proof of Step (A) is finished.

Step (B) $\boxed{\varphi|_{S_1} : S_1 \to D_1 \text{ is bijective and } \varphi^{-1}(D_1) = S_1.}$

Note that $\varphi^{-1}(p) = \{\tilde{p}\} \subset S_1$. From the expressions (6.23) and (6.25), it is clear that $\varphi(S_1) \subset D_1$. Let $r = \exp_p(tv) \in D_1 \setminus \{p\}$, so $0 < t \leq d(p, h_1(v))$. By Lemma 6.4.5, $D_1 \subset B_\rho(p)$ and the minimizing geodesic joining p to r is unique. Therefore, t and v are uniquely determined by r. Then $\tilde{r} = \exp_{\tilde{p}}(\frac{t}{d(p,h_1(v))}\frac{\pi}{2}i^{-1}(v)) \in S_1 \setminus \{\tilde{p}\}$ and $\varphi(\tilde{r}) = r$, so $\varphi|_{S_1} : S_1 \to D_1$ is surjective. By uniqueness of t and v, $\tilde{r} \in S_1$ is the only point in $\varphi^{-1}(r)$, hence $\varphi|_{S_1}$ is injective and $\varphi^{-1}(D_1) = S_1$.

Step (C) $\boxed{\varphi|_{S_2} : S_2 \to D_2 \text{ is bijective and } \varphi^{-1}(D_2) = S_2.}$

By (6.24) and (6.26), $\varphi(S_2) \subset D_2$. Let $\varphi(\exp_{\tilde{p}}(tv)) = q$, then $t = \pi$, which shows that $\exp_{\tilde{p}}(tv) = \tilde{q}$. Therefore, $\varphi^{-1}(q) = \{\tilde{q}\} \subset S_2$. Let $r \in D_2 \setminus \{q\}$. There exists a unique minimizing geodesic σ joining q to r. Then $r = \exp_q(tv)$ where $v = \sigma'(0)$ and, by (6.24), $0 < t \leq d(q, h_2(v))$. Note that v and t are again unique (since $D_2 \subset B_\rho(q)$). Let $w \in (S^{n-1})_{\tilde{p}}$ be given by $i^{-1} \circ h_1^{-1} \circ h_2(v)$ and $s = \pi(1 - \frac{t}{2d(q,h_2(v))})$. Consider the point $\tilde{r} = \exp_{\tilde{p}}(sw)$. Since $\frac{\pi}{2} \leq s < \pi$, $\tilde{r} \subset S_2 \setminus \{\tilde{q}\}$ and $\varphi(\tilde{r}) = \varphi(\exp_{\tilde{p}}(sw)) = \exp_q(tv) = r$, so $\varphi|_{S_2} : S_2 \to D_2$ is surjective. By uniqueness of v and t, $\tilde{r} \in S_2$ is the unique point in $\varphi^{-1}(r)$, which shows injectivity of $\varphi|_{S_2}$ and that $\varphi^{-1}(D_2) = S_2$.

Step (D) $\boxed{\varphi : S^n \to M \text{ is a homeomorphism.}}$

By its definition, the function φ is continuous. Since $\varphi(S^n) = \varphi(S_1) \cup \varphi(S_2) = D_1 \cup D_2 = M$, φ is surjective. Let $r \in M$, then r is contained in D_1 or in D_2. If $r \in D_1$, then by (B) there exists a unique point $\tilde{r} \in \varphi^{-1}(r)$ and $\tilde{r} \in S_1$. Analogously, if $r \in D_2$, then by (C) there exists a unique point $\tilde{r} \in \varphi^{-1}(r)$ and $\tilde{r} \in S_2$. This shows that φ is injective. Therefore, $\varphi : S^n \to M$ is a continuous bijection. Since S^n and M both are compact, φ is a homeomorphism. □

Remark 6.4.11. Together with Remark 6.4.9, one sees that the homeomorphism φ maps S_1 diffeomorphically onto D_1 and S_2 diffeomorphically onto D_2. So it is only at the equator of S^n that φ fails to be diffeomorphic.

6.5 Further developments

The result we proved above is known as the Topological Sphere theorem. It gives rise to a large number of questions. We will present some results which have been obtained since the 1960's. Because of lack of space and time, it is impossible to mention all important developments that are related to the Topological Sphere theorem. For a more detailed overview, see e.g. [AM97] and [BS09a].

First, we consider some theorems where the conditions on the curvature are replaced or completemented by conditions on the diameter. The following result was shown by S. Y. Cheng [Che75]:

Theorem 6.5.1 (Maximal Diameter theorem). *Let M be a complete Riemannian manifold such that its sectional curvature K satisfies $K \geq H > 0$ for constant H and suppose that* $\operatorname{diam}(M) = \frac{\pi}{\sqrt{H}}$. *Then M^n is isometric to M^n_H, the sphere of constant curvature H.*

Proof. See [CE75], Theorem 6.5. □

Note that this is a stronger version of the Theorem of Bonnet-Myers (Theorem 2.2.4).

Berger's Rigidity theorem shows that a complete simply-connected Riemannian manifold whose sectional curvature K satisfies $1 \geq K \geq \frac{1}{4} > 0$ is either homeomorphic to a sphere or isometric to a symmetric space. M. Berger showed that the diameter of M is a distinguishing factor.

Theorem 6.5.2 (Minimal Diameter theorem). *Let M be a complete simply-connected Riemannian manifold such that its sectional curvature K satisfies* $1 \geq K \geq \frac{1}{4} > 0$.

(i) If $\operatorname{diam}(M) \geq \pi$, *$M$ is homeomorphic to S^n.*

(ii) If $\operatorname{diam}(M) = \pi$, *$M$ is isometric to a symmetric space.*

Proof. See [CE75], Theorem 6.6. □

In 1977, K. Grove and K. Shiohama [GS77] proved the following generalization of the Sphere theorem, where the upper bound on the curvature is replaced by a lower bound on the diameter:

Theorem 6.5.3. *Let M^n be a compact Riemannian manifold with sectional curvature greater than* 1. *If* $\operatorname{diam}(M) > \frac{\pi}{2}$, *$M$ is homeomorphic to S^n.*

The next natural question is whether it is possible to replace, under stronger conditions, homeomorphic by diffeomorphic in the statement of Theorem 6.1.1. Observe that the homeomorphism φ of the Sphere theorem is obtained by gluing two discs along their boundaries, so in general φ might not be a diffeomorphism. In fact, there exist manifolds which are homeomorphic, but not diffeomorphic to a sphere. Such manifolds are called exotic spheres. In 1956, Milnor [Mil56] proved that there exist at least seven smooth structures on S^7.

In 1966, D. Gromoll [Gro66] showed that any complete and simply-connected Riemannian manifold M^n whose sectional curvature K satisfies the inequality $\delta(n) < K \leq 1$ is diffeomorphic to S^n. The constant $\delta(n)$ depends only on the dimension of M and converges to 1 as $n \to \infty$.

This result was improved in the following years. In 1971, M. Sugimoto, K. Shiohama, and H. Karcher [SSK71] proved the Differentiable Sphere theorem with a pinching constant δ independent of n ($\delta = 0.87$). The pinching constant was subsequently improved by E. Ruh [Ruh73] ($\delta = 0.80$) and by K. Grove, H. Karcher, and E. Ruh [GKR74] ($\delta = 0.76$).

In 1982, R. Hamilton [Ham82] introduced new ideas to study this problem. Given a compact Riemannian manifold (M, g_0), he studied the following geometric evolution equation for the Riemannian metric:

$$\frac{\partial}{\partial t} g(t) = -2\mathrm{Ric}_{g(t)}, \qquad g(0) = g_0 \tag{6.28}$$

The solution of this initial value problem is called the Ricci flow. Hamilton proved that for any given metric g_0, there exists a $T > 0$ such that the solution of (6.28) is defined on the interval $[0, T)$.

Hamilton's idea was to define a kind of nonlinear diffusion equation which would tend to smooth out irregularities in the metric. In fact, he showed that in dimension 3, the Ricci flow deforms metrics with positive Ricci curvature to constant curvature metrics.

The Ricci flow turned out to be a powerful tool for studying Riemannian manifolds. Ricci flow techniques are an essential ingredient in the proof of the Poincaré Conjecture by G. Perelman. It is also a useful tool to study 1/4-pinched manifolds. In recent years, S. Brendle and R. Schoen achieved fundamental results. By using the Ricci flow, they classified all manifolds of 1/4-pinched curvature up to diffeomorphism:

Theorem 6.5.4 (S. Brendle, R. Schoen [BS09b]). *Let M be a compact Riemannian manifold of dimension $n \geq 4$ such that for all $p \in M$, the sectional curvature satisfies*

$$0 < K_p(\sigma_1) < 4K_p(\sigma_2)$$

for all two-dimensional planes $\sigma_1, \sigma_2 \subset T_pM$. Then M admits a metric of constant scalar curvature and therefore is diffeomorphic to a spherical space form.

Note that the conditions of Theorem 6.5.4 are even weaker than in Theorem 6.1.1, since the inequality $0 < \frac{1}{4} < K \leq 1$ is only satisfied at each point but not necessarily globally. In this regard, we say that M is pointwise 1/4-pinched. The result above is known as the Differentiable sphere theorem. Moreover, S. Brendle and R. Schoen also classified those manifolds whose curvature is only weakly pointwise 1/4-pinched (i.e. the sectional curvature at each point lies in the closed interval $[\frac{1}{4}, 1]$):

Theorem 6.5.5 (S. Brendle, R. Schoen [BS08]). *Let M be a compact Riemannian manifold of dimension $n \geq 4$ such that for all $p \in M$, the sectional curvature satisfies*

$$0 \leq K_p(\sigma_1) \leq 4K_p(\sigma_2)$$

for all two-dimensional planes $\sigma_1, \sigma_2 \subset T_pM$. Then M is either locally symmetric or diffeomorphic to a spherical space form.

Bibliography

[AM97] Abresch, Uwe ; Meyer, Wolfgang T. *Injectivity radius estimates and sphere theorems.* Grove, Karsten (ed.) et al., Comparison geometry. Cambridge: Cambridge University. Math. Sci. Res. Inst. Publ. 30, 1-47 (1997). 1997

[Ber60] Berger, M.: Les variétés riemanniennes $\frac{1}{4}$-pincées. In: *Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser.* 14 (1960), S. 161–170

[BG88] Berger, Marcel ; Gostiaux, Bernard: *Differential geometry: manifolds, curves, and surfaces. Transl. from the French by Silvio Levy.* Graduate Texts in Mathematics, 115. New York etc.: Springer-Verlag. XII, 474 p. , 1988

[Bär06] Bär, Christian: *Differentialgeometrie.* Lecture Notes, 182 p. 2006. – http://geometrie.math.uni-potsdam.de/documents/baer/skripte/skript-DiffGeo.pdf

[BS08] Brendle, Simon ; Schoen, Richard M.: Classification of manifolds with weakly 1/4-pinched curvatures. In: *Acta Math.* 200 (2008), Nr. 1, S. 1–13

[BS09a] Brendle, Simon ; Schoen, Richard. *Sphere theorems in geometry.* Cao, Huai-Dong (ed.) et al., Geometry, analysis, and algebraic geometry. Forty years of the Journal of Differential Geometry. Based on the 7th conference on geometry and topology, Cambridge, MA, USA, May 2008. In memory of Prof. C.-C. Hsiung. Somerville, MA: International Press. Surveys in Differential Geometry 13, 49-84 (2009). 2009

[BS09b] Brendle, Simon ; Schoen, Richard M.: Manifolds with 1/4-pinched curvature are space forms. In: *J. Amer. Math. Soc.* 22 (2009), S. 287–307

[Bur09] Burtscher, Annegret: *Isomorphisms of Algebras of Smooth and Generalized Functions*, University of Vienna, Diplomarbeit, 2009

[Car92] do Carmo, Manfredo P.: *Riemannian geometry. Translated from the Portuguese by Francis Flaherty.* Mathematics: Theory & Applications. Boston, MA etc.: Birkhäuser. xiii, 300 p. , 1992

[CE75] Cheeger, Jeff ; Ebin, David G.: *Comparison theorems in Riemannian geometry.* North-Holland Mathematical Library. Vol. 9. Amsterdam - oxford: North- Holland Publishing Company; New York: American Elsevier Publishing Company, Inc. VIII, 174 p. , 1975

[Che75] Cheng, Shiu-Yuen: Eigenvalue comparison theorems and its geometric applications. In: *Math. Z.* 143 (1975), S. 289–297

[GKR74] Grove, Karsten ; Karcher, Hermann ; Ruh, Ernst A.: Jacobi fields and Finsler metrics on compact Lie groups with an application to differentiable pinching problems. In: *Math. Ann.* 211 (1974), S. 7–21

[Gro66] Gromoll, D.: Differenzierbare Strukturen und Metriken positiver Krümmung auf Sphären. In: *Math. Ann.* 164 (1966), S. 353–371

[GS77] Grove, Karsten ; Shiohama, Katsuhiro: A generalized sphere theorem. In: *Math. Ann.* 106 (1977), S. 201–211

[Hag05] Hage, Sebastian: *Einführung in die Grundlagen der Morse Theorie*. Seminarvortrag. 2005. – http://www.uni-math.gwdg.de/schick/teach/Introductionpdf

[Hal09] Haller, Stefan: *Algebraische Topologie*. Lecture Notes. 2009. – http://www.mat.univie.ac.at/ stefan/files/AT09/AT09.pdf

[Ham82] Hamilton, Richard S.: Three-manifolds with positive Ricci curvature. In: *J. Differ. Geom.* 17 (1982), S. 255–306

[Hat02] Hatcher, Allen: *Algebraic topology.* Cambridge: Cambridge University Press. xii, 544 p., 2002

[Kli61] Klingenberg, W.: Über Riemannsche Mannigfaltigkeiten mit positiver Krümmung. In: *Comment. Math. Helv.* 35 (1961), S. 47–54

[KN63] Kobayashi, Shoshichi ; Nomizu, Katsumi: *Foundations of Differential geometry Vol. 1.* New York, London: Interscience Publishers, 329 p., 1963

[Kun08] Kunzinger, Michael: *Differentialgeometrie 1.* Lecture Notes. 2008. – http://www.mat.univie.ac.at/ mike/teaching/ss08/dg.pdf

[Kun09] Kunzinger, Michael. *Differentialgeometrie 2.* Lecture Notes. 2009

[Lee97] Lee, John M.: *Riemannian manifolds: an introduction to curvature.* Graduate Texts in Mathematics. 176. New York, NY: Springer. xv, 224 p., 1997

[Mat02] Matsumoto, Yukio: *An introduction to Morse theory. Transl. from the Japanese by Kiki Hudson and Masahico Saito.* Translations of Mathematical Monographs. Iwanami Series in Modern Mathematics. 208. Providence, RI: American Mathematical Society (AMS). xiv, 219 p., 2002

[Mey89] Meyer, Wolfgang: Topogonovs Theorem and Applications. In: *Lecture Notes, College on Differential Geometry, Trieste, 51 p.* (1989)

[Mil56] Milnor, John W.: On manifolds homeomorphic to the 7-sphere. In: *Math. Ann.* 64 (1956), S. 399–405

[Mil63] Milnor, John W.: *Morse theory. Based on lecture notes by M. Spivak and R. Wells.* Annals of Mathematics Studies. No.51. Princeton, N.J.: Princeton University Press, VI, 153 p. , 1963

[MT97] Madsen, Ib ; Tornehave, Jørgen: *From calculus to cohomology: de Rham cohomology and characteristic classes.* Cambridge: Cambridge University Press. vii, 286 p., 1997

[Mye41] Myers, S.B.: Riemannian manifolds with positive mean curvature. In: *Duke Math. J.* 8 (1941), S. 401–404

[O'N83] O'Neill, Barrett: *Semi-Riemannian geometry. With applications to relativity.* Pure and Applied Mathematics, 103. New York-London etc.: Academic Press. XIII, 468 p., 1983

[Pet06] Petersen, Peter: *Riemannian geometry. 2nd ed.* Graduate Texts in Mathematics 171. New York, NY: Springer. xvi, 401 p. , 2006

[Rau51] Rauch, H.E.: A contribution to differential geometry in the large. In: *Math. Ann.* 54 (1951), S. 38–55

[Ruh73] Ruh, Ernst A.: Krümmung und differenzierbare Struktur auf Sphären. II. In: *Math. Ann.* 205 (1973), S. 113–129

[SSK71] Sugimoto, M. ; Shiohama, K. ; Karcher, H.: On the differentiable pinching problem. In: *Math. Ann.* 195 (1971), S. 1–16

Index

Curriculum Vitae

Personal information

Name	Klaus Kröncke
Date of Birth	1st of August, 1986
Place of Birth	Vienna, Austria
Nationality	German
email	klaus.kroencke@gmx.at

Education

Since 10/2007	Bachelor Studies in Physics at the University of Vienna
Since 10/2005	Diploma Studies in Mathematics at the University of Vienna
Since 10/2004	Diploma Studies in Geography at the University of Vienna
06/2004	Matura (school leaving examination) passed with honors
1996-2004	Bundesgymnasium Wien VI
1992-1996	Volksssschule Wien VI

Employment

10/2009-02/2010	Tutor at the Faculty of Mathematics, University of Vienna
03/2009	Co-worker at the exhibition IMAGINARY, Vienna
10/2008-02/2009	Tutor at the Institute of Mathematics, University of Natural Resources and Applied Life Sciences, Vienna

Stipends

2007	Leistungsstipendium (merit scholarship) by the University of Vienna for the academic year 2006/07

Other activities

Since 2006	Co-organizer of the Astronomical Summer camp (Astronomisches Sommerlager), Germany