Title
Sub-Index for Critical Points of Distance Functions

Permalink
https://escholarship.org/uc/item/4wb790kk

Author
Herzog, Barbara

Publication Date
2012

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA
RIVERSIDE

Sub-Index for Critical Points of Distance Functions

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Mathematics

by

Barbara Christine Herzog

June 2012

Dissertation Committee:

Dr. Fred Wilhelm, Chairperson
Dr. Reinhard Schultz
Dr. Stefano Vidussi
The Dissertation of Barbara Christine Herzog is approved:

______________________________
Committee Chairperson

______________________________

University of California, Riverside
Acknowledgments

First and foremost, I would like to thank my advisor, Dr. Fred Wilhelm, for his incredible support. I truly appreciate his tireless dedication, endless array of ideas, and clear focus on the bigger picture and overall timeline. During every step of this important journey, it was clear to me that he had my best interests in mind. Without a doubt, his encouragement and patience made my degree possible. Dr. Wilhelm, thank you so much!

I am also grateful for my family. Knowing that they are always there for me is a tremendous comfort. I appreciate my mother, Christine Herzog, for constantly reminding me that she believes I can do anything I set my mind to. She has always been a strong female role model for me, and her unconditional love and support continues to mean the world to me. I appreciate my sister, Elizabeth Johnston, for the way her independent and hard working spirit reminds me to never give up. She has always been one of the most important people in my life, and I cherish our close relationship. My nieces, Kayla and Alyssa Johnston, are the most amazing young women I know. Kayla’s commitment to academic pursuits and Alyssa’s dedication to athletics are both equally inspirational. Their incredible talent and intelligence remind me of the importance of being the best that I can be each and every day. I encourage you both to pursue your goals with enthusiasm. Working hard to achieve big dreams is a hallmark of our family. Our strength and unity will continue to carry us through whatever life has in store.

I would also like to thank Tony Alo. Words cannot express how much I have appreciated his support and encouragement every single solitary day over the last six years. There is no way I would have made it through this incredibly challenging journey without him. Being able to vent about my worries and fears as well as celebrate each and
every accomplishment along the way made this undertaking possible. Tony’s confidence, extreme patience, and wonderful advice kept me on track and focused at every stage of this process. Certainly, his unwavering belief that I would achieve this goal was the biggest inspiration of all. If Tony had a dollar for every time he told me that things were going to be okay, he would clearly be a millionaire by now. Tony, you are truly amazing and will continue to hold a very special place in my heart. Without a doubt, I am a better person because of the time I have spent with you.

Last but certainly not least, there are several friends who have significantly contributed to my accomplishment. I am grateful for Dr. Tiff Troutman for encouraging me to work with Dr. Wilhelm and being incredibly supportive every step of the way. Thanks, Tiff, for understanding my kind of crazy and being such a wonderful friend! I also appreciate my friendship with Dennis and Danaca Gumaer. Welcoming me into your home with open arms provided me with a necessary balance between work and play. Thank you so much for your tremendous generosity and kindness! I don’t know what I would have done without you! You are all very special to me, and I am fortunate to have you in my life.
To Christine, Elizabeth, Kayla, Alyssa, and Tony,

for all the inspiration and joy that you bring to my life.
Morse theory is based on the idea that a smooth function on a manifold yields data about the topology of the manifold. In this way it provides a tool for visualizing the shape of a space. Specifically, Morse’s Isotopy Lemma tells us that the homotopy type of a manifold does not change in regions without critical points. The topology only changes in the presence of a critical point. Morse’s Theorem states that the specific topological change is determined by the index of the Hessian at each critical point. In Morse Theory a smooth function is essential so that the differential and Hessian exist.

In Riemannian geometry, the distance function is not smooth everywhere. This means the differential as well the Hessian do not exist and Morse Theory cannot be applied. In order to generalize Morse Theory to this non-smooth function, an alternate definition of critical point and index are required. Grove and Shiohama developed a definition of critical point for the Riemannian distance function and used it to generalize Morse’s Isotopy Lemma [9]. Their generalization had a profound impact on the study of Riemannian geometry. Since no definition of index currently exists, Morse’s Theorem has not been generalized.
The purpose of this dissertation is to define a new notion, called sub-index, for critical points of Riemannian distance functions. We show that Morse’s connectedness corollary holds for the distance function when index is replaced by sub-index.
# Contents

## List of Figures

x

## 1 Background

1.1 Introduction .................................................. 1
1.2 Morse Theory .................................................. 3
1.3 The Riemannian Distance Function ......................... 6

## 2 Results

2.1 Preliminaries .................................................. 13
2.2 The Definition of Sub-Index and Main Results ............. 16
2.3 Technical Lemmas for the Connectedness Theorem ........... 20
2.4 Proof of the Connectedness Theorem ....................... 24
2.5 Lemmas Related to Conjugate Points ....................... 29
2.6 Proof of the Relative $\pi_1$ Theorem ..................... 38
2.7 The Generalized Butterfly Lemma ........................... 39

## Bibliography

44
List of Figures

1.1 An interval that does not contain a critical value. .................. 3
1.2 An interval that contains no critical values other than $c$. .......... 5
1.3 An illustration of Morse’s Theorem ..................................... 6
1.4 The unit circle with parameterization $x(t) = (\cos t, \sin t)$. ....... 9
1.5 The graph of $\text{dist}_p x(t) = \pi - |t - \pi|$ .............................. 10
1.6 Sample configuration of $\hat{\mathcal{P}}_x$ when $x$ is a regular point. .... 11
1.7 Sample configuration of $\hat{\mathcal{P}}_x$ when $x$ is a critical point. ....... 12

2.1 $A(\hat{\mathcal{P}}_{p_0})$ is the intersection of hemispheres .................. 16
2.2 The flat 2-torus and its unit tangent space at $x$, for $i = 1, 2, 3$ ..... 17
2.3 The sets involved in Lemma 30 ........................................... 22
Chapter 1

Background

1.1 Introduction

In order to determine the shape of a function or a space, it is of monumental importance to find and classify critical points. In Euclidean space, calculus provides us with the necessary tools for visualizing the shape of a smooth function. In this case, shape can refer to such things as minimums and maximums. The critical points occur where the gradient, given by the matrix of first partial derivatives, is the zero matrix. To classify the critical points as minimums or maximums, we can use the Hessian, represented by the matrix of second partial derivatives. Since the minimums and maximums are determined using derivatives, a smooth function is required.

Morse Theory, which was created by Marston Morse in the 1920’s, allows us to visualize the shape of an $n$-dimensional manifold $M$ by analyzing the critical points of a smooth function defined on it. In this case, shape can refer to minimums, maximums, and even homotopy type. By definition, a point $p$ in $M$ is a critical point of a smooth function $h : M \rightarrow \mathbb{R}$ if the differential

$$h_* : T_p M \rightarrow T_{h(p)} \mathbb{R}$$
is zero. Morse Theory tells us, via the Isotopy Lemma, that the homotopy type of a manifold does not change in a region that consists entirely of regular points, points that are not critical [11]. The homotopy type only changes in the presence of a critical point. Morse’s Theorem delineates the specific type of change by considering the index of the critical point [11]. The index of a critical point is defined as follows but can be thought of as the number of independent directions of decrease from the critical point.

**Definition 1.** For a smooth function, the index of a critical point is the dimension of the largest subspace on which the Hessian is negative definite.

In order to utilize Morse Theory, a smooth function is required so that the differential as well as the Hessian exists.

In Riemannian geometry, the study of Riemannian manifolds, the distance between two points is an important function defined on a manifold. Unfortunately, the distance function is not smooth everywhere, which means that Morse Theory cannot be applied in order to analyze it. To extend Morse Theory to this non-smooth function, an alternate definition of critical point and index, not related to the differential, is necessary. In 1977, Grove and Shiohama developed a notion of critical point for the Riemannian distance function and generalized the isotopy lemma to this case [9]. Their generalization had a profound effect on the study of Riemannian manifolds. Applications of their result include the Diameter Sphere Theorem [9], Gromov’s Betti Number Theorem [7], and Grove and Petersen’s Homotopy Finiteness Theorem [8].

Currently, no definition of index exists for a critical point of the Riemannian distance function. The main goal of this work has been to develop a notion of sub-index and use it to generalize a consequence of Morse’s Theorem.
1.2 Morse Theory

Morse theory is based on the idea that a smooth function on a manifold yields data about the topology of the manifold. In this way it provides a tool for visualizing the shape of a space. The two main results discussed in this section are Morse’s Isotopy Lemma and Morse’s Theorem.

Throughout this chapter, let \( M \) be an \( n \)-dimensional manifold and \( h : M \to \mathbb{R} \) be a smooth function. Define a sublevel set of \( M \) as follows.

\[ M^a := h^{-1}(-\infty, a]. \]

Note that for \( a < b \), the set \( M^a \) is a subset of \( M^b \). In fact, if \( a \) is not a critical value for \( h \), then by the Implicit Function Theorem, the sublevel set \( M^a \) is a smooth submanifold with boundary.

Example 3. To motivate Morse’s Isotopy Lemma, let \( M \) be the 2-dimensional torus shown in Figure 1.1.

![Figure 1.1: An interval that does not contain a critical value.](image)

Let \( h : M \to \mathbb{R} \) be the height of each point on \( M \). The critical points for \( h \) occur at the points \( x, y, z \) and \( w \) on \( M \) since the tangent plane at each of these points...
is horizontal. The corresponding critical values occur at heights of $0, h_1, h_2$, and $h_3$ respectively. Note that there are no critical values in the interval $[a, b]$. Further, both $M^a$ and $M^b$ have the same homotopy type. In fact, $M^a$ is a deformation retract of $M^b$. This illustrates Morse’s Isotopy Lemma which relates sublevel sets from an interval without critical values.

**Lemma 4.** (Morse’s Isotopy Lemma [11]) Suppose there are no critical values in $[a, b]$. Then $M^a$ is diffeomorphic to $M^b$. Further, $M^a$ is a deformation retract of $M^b$, so that the inclusion map $M^a \hookrightarrow M^b$ is a homotopy equivalence.

Morse’s Theorem, which describes how the homotopy type of a manifold changes at a critical point, applies only to non-degenerate critical points. Non-degenerate critical points are guaranteed to be isolated, while degenerate critical points can be isolated or not isolated.

**Definition 5.** A critical point is called non-degenerate if the Hessian of $h$, represented by the matrix of second partial derivatives evaluated at that point, has an inverse.

The index of the critical point determines the specific type of change. Informally, the index gives the number of independent directions of decrease as we move away from the critical point. A critical point with index zero corresponds to a minimum, since none of the independent directions from the critical point corresponds to a decrease. Further, a critical point with index equal to the dimension of the manifold represents a maximum, given that every independent direction from the critical point corresponds to a decrease.

In Example 3, the index of $w$ is two, since moving in either of the two independent directions away from $w$ causes the height to decrease. Both $y$ and $z$ have index
one, since there is only one direction that causes a decrease in height. The index of $x$ is zero, since there are no directions of decrease possible.

**Theorem 6.** (Morse [11]) Let $p$ be a nondegenerate critical point with $h(p) = c$ and index $\lambda$. Suppose $h^{-1}[c - \epsilon, c + \epsilon]$ is compact and contains no critical points other than $p$. Then $M^{c+\epsilon}$ has the homotopy type of $M^{c-\epsilon}$ with a $\lambda$-cell attached.

By definition, a $\lambda$-cell, where $\lambda$ is a whole number, is a space homeomorphic to a closed $\lambda$-dimensional ball. So a 0-cell is a point, a 1-cell is an interval, and a 2-cell is a disk.

**Example 7.** Again consider the height function on the 2-dimensional torus.

The interval $[c - \epsilon, c + \epsilon]$ contains no critical values other than $c$. The set $M^{c-\epsilon}$ is a bowl, as shown in Figure 1.3, and has the same homotopy type as a point. The set $M^{c+\epsilon}$ is a curved tube and has the same homotopy type as a circle. For the set $M^{c-\epsilon}$ to have the same homotopy type as $M^{c+\epsilon}$, it is necessary to attach a 1-cell, since the index of $y$ is one. After attaching an interval to $M^{c-\epsilon}$, the resulting space, depicted in Figure 1.3, will have the same homotopy type as $M^{c+\epsilon}$. In order to visualize this, note that we
can shrink the $M^{c-\epsilon}$ part of the space down to a point. The resulting space will be an interval attached to a point, i.e. a circle, which has the same homotopy type as $M^{c+\epsilon}$.

\[ M^{c-\epsilon} \quad M^{c-\epsilon} \text{ with a 1-cell attached} \quad M^{c+\epsilon} \]

Figure 1.3: An illustration of Morse’s Theorem

Morse’s Theorem implies the following corollary.

**Corollary 8.** Given the hypotheses in the theorem, for all $i = 1, 2, \ldots, (\lambda - 1)$

\[ \pi_i(M^{c+\epsilon}, M^{c-\epsilon}) = 0, \]

i.e. the pair $(M^{c+\epsilon}, M^{c-\epsilon})$ is $(\lambda - 1)$-connected.

This means a cell of dimension $1, 2, \ldots, \text{ or } (\lambda - 1)$, within $M^{c+\epsilon}$ whose boundary lies in $M^{c-\epsilon}$, can be deformed into $M^{c-\epsilon}$. Currently, Morse’s Theorem has not been generalized to the Riemannian distance function. However, our main result in Section 2.2 generalizes Corollary 8 to the Riemannian distance function given our new notion of sub-index.

### 1.3 The Riemannian Distance Function

A Riemannian manifold is a smooth manifold $M$ equipped with a metric $g$, defined on the tangent space $TM$, that varies smoothly from point to point. The metric
is a family of inner products

\[ g_p : T_pM \times T_pM \rightarrow \mathbb{R} \]

depending on the point \( p \) in \( M \). Varying smoothly means that for all vector fields \( X \) and \( Y \) on \( M \), the map

\[ p \mapsto g_p \left( X(p), Y(p) \right) \]

is smooth.

The simplest example of a Riemannian manifold is Euclidean space whose inner product is given by the dot product from multivariable calculus. Some formulas involving the dot product are

\[ \|v\| = (v \cdot v)^{\frac{1}{2}} \quad \text{and} \quad v \cdot w = \|v\|\|w\| \cos \theta \]

where \( v \) and \( w \) are vectors and \( \theta \) is the angle between them. In Riemannian geometry, the corresponding formulas are given by

\[ \|v\| = g(v,v)^{\frac{1}{2}} \quad \text{and} \quad g(v,w) = \|v\|\|w\| \cos \theta. \]

For a Riemannian manifold, the distance between two points is defined by considering the length of all curves that connect the points.

**Definition 9.** The length of a curve \( \gamma : [0,1] \rightarrow M \) is given by

\[ \text{Len}(\gamma) = \int_0^1 g\left( \gamma'(t), \gamma'(t) \right)^{\frac{1}{2}} dt. \]

**Definition 10.** A curve \( \gamma \) is parametrized by arc length if \( \|\gamma'(t)\| = 1 \), i.e. the curve has unit speed.

**Definition 11.** A geodesic is a curve \( \gamma_v : [0,1] \rightarrow M \), with \( \gamma_v'(0) = v \), that satisfies

\[ \gamma_v''(t) = 0. \]
In Riemannian geometry, a geodesic is the generalized notion of a line since it is a constant speed curve. Geodesics are uniquely determined by their initial point \( \gamma_v(0) \) and their initial direction \( v \in T_{\gamma_v(0)} M \). Unlike \( \mathbb{R}^n \), it is possible for two points on a manifold to be connected by multiple geodesics.

We can also define a geodesic using the exponential map.

**Definition 12.** Given a geodesic \( \gamma_v \) with \( \gamma_v(0) = p \), the exponential map

\[
\exp_p : T_p M \to M
\]

is defined as \( \exp_p(v) = \gamma_v(1) \).

In general, a geodesic \( \gamma_v : [0, 1] \to M \) can be written as \( \gamma_v(t) = \exp_p(tv) \).

When \( r \) is smaller than the injectivity radius, the exponential map takes lines in \( B(0, r) \subset T_p M \) starting at \( 0_p \) to geodesics in \( B(p, r) \) starting at \( p \).

**Definition 13.** The injectivity radius, given by \( \text{inj} M \), is the largest radius \( r \) such that

\[
\exp_p : B(0, r) \to B(p, r)
\]

is a diffeomorphism for all \( p \in M \).

**Definition 14.** For a fixed point, \( p \), the Riemannian distance function, \( \text{dist}_p : M \to \mathbb{R} \), is defined as

\[
\text{dist}_p x := \inf \left\{ \text{Len}(\gamma) \mid \gamma \text{ is a unit speed curve from } x \text{ to } p \right\}.
\]

**Definition 15.** A segment is a geodesic between two points whose length equals its distance.

So, a segment is the shortest geodesic between two points.
Example 16. To illustrate geodesics compared to segments as well as length compared to distance, let $M$ be the unit circle centered at the origin $O$ with $p = (1, 0)$ and parameterization
\[ x(t) = (\cos t, \sin t). \]

The curves in Figure 1.4 from $p$ to $x(\frac{3\pi}{2})$ moving either clockwise or counterclockwise are both geodesics. However, the clockwise one is a segment since it represents the shortest geodesic between the points.

![Figure 1.4: The unit circle with parameterization $x(t) = (\cos t, \sin t)$.

The length of each geodesic from $p$ to $x(t)$ can be determined using the arc length formula. On a circle, arc length is given by $s = r\theta$ where $r$ is the radius and $\theta$ is the non-negative angle formed by $O$ and $p$ and $Ox$. Since the radius of this circle is one, the length from $p$ to $x(t)$ is equal to the angle $\theta$. So, the length of the geodesic from $p$ to $x(\frac{3\pi}{2})$ moving counterclockwise is $\frac{3\pi}{2}$, while the length is $\frac{\pi}{2}$ when traveling clockwise. Thus, the distance is $\frac{\pi}{2}$, and only the geodesic associated with the clockwise path can be called a segment.

On the other hand, the length of the semi-circle from $p$ to $x(\pi)$ is $\pi$ regardless of traveling on a geodesic counterclockwise or clockwise from $p$. The distance is also $\pi$, and both of the geodesics described can be called segments from $p$ to $x(\pi)$. 
In this example, the distance from \( p \) to \( x(t) \) can be written explicitly as

\[
\text{dist}_p x(t) = \pi - |t - \pi|
\]

on the interval \([0, 2\pi]\), and its graph is given in Figure 1.5. Note that \( \text{dist}_p \) is not smooth at the point \( x(\pi) \).

In general, the Riemannian distance function is not smooth everywhere. It is smooth, however, at all points before the cut locus, the set of points where geodesics emanating from \( p \) stop being segments. For a geodesic \( \gamma_v : [0, \infty) \rightarrow M \), define \( t_v \) to be the largest parameter time in \([0, \infty)\) such that \( \gamma_v : [0, t_v] \rightarrow M \) is a segment.

In Example 16, a geodesic traveling from \( p \) counterclockwise will only be a segment until the point \( t_v = \pi \). Beyond \( x(\pi) \), it would be shorter to travel along a geodesic clockwise from \( p \).

The set of points \( \gamma_v(t_v) \) in \( M \) corresponding to the parameter times given by \( t_v \), for all geodesics emanating from \( p \), is called the cut locus.

**Definition 17.** For a point \( p \), the **cut locus** is the set of points in \( M \) given by

\[
\text{Cut}(p) := \left\{ \gamma_v(t_v) \mid v \in T_pM \right\}.
\]

In Example 16, \( x(\pi) \) is the only point in the cut locus. The point \( x(\pi) \) is also important because it represents the maximum distance from \( p \), meaning that it is a
critical point. In general, the set of critical points (excluding $p$) forms a subset of the cut locus. Although the Riemannian distance function $\text{dist}_p$ is not smooth on the cut locus, it is smooth on the set of points before the cut locus. This is given by

$$\left\{ \gamma_v(t) \mid t < t_v, \ v \in T_pM \right\}.$$

Even though the Riemannian distance function is not smooth everywhere, it is directionally differentiable. For $x \in M$, let $S_x$ be the unit tangent sphere at $x$, i.e. $S_x \subset T_xM$ and for all $w \in S_x$ we have $\|w\| = 1$.

**Definition 18.** For each $x \in M$, define the set

$$\hat{H}_x^p := \left\{ w \in S_x \mid w \text{ is tangent at } x \text{ to a segment from } x \text{ to } p \right\}.$$

By [13], the directional derivative of $\text{dist}_p$ in the direction of $v \in T_xM$ is given by

$$D_v(\text{dist}_p x) = -\cos \angle(v, \hat{H}_x^p).$$

Using this directional derivative, Grove and Shiohama created the definition of regular point and critical point for $\text{dist}_p$ as follows [9].

**Definition 19.** A point $x$ in $M$ is a **regular point** for $\text{dist}_p$ if there exists a $v$ in $T_xM$ such that $\angle(v, \hat{H}_x^p) > \frac{\pi}{2}$.

![Figure 1.6: Sample configuration of $\hat{H}_x^p$ when $x$ is a regular point.](image)
Since the vectors in $\hat{p}_x^p$ form an angle greater than $\frac{\pi}{2}$ with $v$, the vectors in $\hat{p}_x^p$ point in the same general direction from $x$ as shown in Figure 1.6. When $x$ is a regular point, the directional derivative will be

$$D_v(dist_p x) > 0.$$ 

This means moving in the direction of $v$ causes $dist_p$ to increase. Thus, the vector $v$ is “gradient-like” for $dist_p$.

A critical point is defined to be a point that is not regular.

**Definition 20.** A point $x$ in $M$ is a **critical point** for $dist_p$ if for all $v$ in $T_x M$ we have

$$\angle(v, \hat{p}_x^p) \leq \frac{\pi}{2}.$$ 

![Figure 1.7: Sample configuration of $\hat{p}_x^p$ when $x$ is a critical point.](image)

For a critical point, all vectors in the tangent space form an angle less than or equal to $\frac{\pi}{2}$ with $\hat{p}_x^p$, as illustrated in Figure 1.7. In this case, the vectors in $\hat{p}_x^p$ are fairly spread out in the unit tangent sphere, meaning that no such $v$ is possible. For a critical point, $D_v(dist_p x) \leq 0$ for all $v$ in $T_x M$.

Given these definitions, the generalized isotopy lemma is as follows.

**Lemma 21.** (Grove, Shiohama [9]) Suppose $dist_p$ has no critical values in $[a, b]$. Then

$$M^a := dist_p^{-1}(-\infty, a]$$

is homeomorphic to

$$M^b := dist_p^{-1}(-\infty, b].$$

Since $M^a$ is not smooth, homeomorphism is the strongest condition possible.
Chapter 2

Results

In this chapter, we present our definition of sub-index for a critical point of the Riemannian distance function. We also give our main results based on this definition.

2.1 Preliminaries

Throughout this chapter, let $M$ be an $n$-dimensional compact Riemannian manifold. For a fixed $p$ in $M$, we assume the critical points for the Riemannian distance function $dist_p$ are isolated. This is a reasonable assumption since the non-degeneracy requirement in Morse Theory implies isolated critical points.

The critical values can be made distinct by adding a smooth function $f$ to $dist_p$. Although $dist_p + f$ is not a distance function, the definitions of critical and regular point for $dist_p$ can be extended to $dist_p + f$ in a natural way. First, define

$$\hat{\nabla}_y^p := \{\hat{\nabla}_y^p + \nabla f\}.$$ 

A point $y$ will be critical for $dist_p + f$ if for all $v$ in $T_y M$ we have

$$\angle (v, \hat{\nabla}_y^p) \leq \frac{\pi}{2}.$$
On the other hand, a point \( y \) will be defined to be regular for \( \text{dist}_p + f \) if there exists a \( v \) in \( T_y M \) such that
\[
\left< v, \frac{\hat{\gamma}_y}{\| \hat{\gamma}_y \|} \right> > \frac{\pi}{2}.
\]

Using these definitions, the following lemma shows that \( \text{dist}_p \) can be altered so that its critical points remain the same but the critical values become distinct.

**Lemma 22.** Suppose \( x \) is a critical point with \( \text{dist}_p x = c \). Let \( N_1 \subset N_2 \) be neighborhoods of \( x \) such that \( \overline{N}_1 \subset N_2 \) and \( x \) is the only critical point in \( N_2 \). Then for any \( \epsilon > 0 \) there is a function
\[
\text{Dist}_p : M \rightarrow \mathbb{R}
\]
with the following properties.

1) The set of critical points for \( \text{Dist}_p \) is the same as the set of critical points for \( \text{dist}_p \).

2) The point \( x \) is the only critical point of \( \text{Dist}_p \) with critical value \( \text{Dist}_p x \).

3) The difference function given by \( f \equiv \text{Dist}_p - \text{dist}_p \) is smooth, constant on \( N_1 \), and supported on \( N_2 \).

4) The function \( f \) satisfies \( \| f \|_{C^1} < \epsilon \), i.e. all directional derivatives are smaller than \( \epsilon \).

**Proof.** By Urysohn’s Lemma, there is a function \( \chi : M \rightarrow \mathbb{R} \) so that
\[
\chi := \begin{cases} 
1 & \text{on } \overline{N}_1 \\
0 & \text{on } M \setminus N_2
\end{cases}
\]
is smooth. The desired function \( \text{Dist}_p \) is obtained by setting
\[
\text{Dist}_p := \text{dist}_p + \delta \cdot \chi
\]
for \( \delta > 0 \) sufficiently small. It remains to show that \( \text{Dist}_p \) and \( \text{dist}_p \) have the same critical and regular points throughout \( M \).
On the set $\mathcal{N}_1$, we have $f \equiv \delta$, which means that $\nabla f \equiv 0$ and $\nabla y = y$. Similarly, $f \equiv 0$ on the set $M \setminus \mathcal{N}_2$. So on $\mathcal{N}_1 \cup (M \setminus \mathcal{N}_2)$, the critical and regular points for $\text{Dist}_p$ and $\text{dist}_p$ will be the same.

Now consider the set $\mathcal{N}_2 \setminus \mathcal{N}_1$. Since $x$ is the only critical point in $\mathcal{N}_2$, each point $y$ in $\mathcal{N}_2 \setminus \mathcal{N}_1$ is a regular point for $\text{dist}_p$. This means there exists a $v$ in $T_y M$ such that $\langle v, \nabla^p y \rangle > \frac{\pi}{2}$. Choose $\delta > 0$ small enough so that there exists $\hat{v} \in T_y M$ such that $\langle \hat{v}, \nabla^p x \rangle = \langle \hat{v}, \nabla^p y + \delta \cdot \nabla \chi \rangle > \frac{\pi}{2}$.

In fact, choose it small enough so that it is independent of the choice of $y$. As long as $\delta < \delta_0$, the points in $\mathcal{N}_2 \setminus \mathcal{N}_1$ will be regular for $\text{Dist}_p$. Thus, $\text{Dist}_p$ and $\text{dist}_p$ will have the same critical and regular points throughout $M$.

Finally, we ensure that the critical point $x$ has a unique critical value $\text{Dist}_p x$. Note that since $x$ is in $\mathcal{N}_1$, its critical value is given by

$$\text{Dist}_p x = \text{dist}_p x + \delta.$$  

Choose $\delta < \min \{\delta_0, \epsilon\}$ so that $\text{Dist}_p x$ is distinct from other critical values of $\text{Dist}_p$. 

Throughout the remainder of the chapter, let the critical points for $\text{dist}_p$ be denoted by $x_i$ where $i = a, \ldots, -1, 0, 1, \ldots, b$. For simplicity, we assume that the corresponding critical values $\text{dist}_p x_i = c_i$ are distinct and ordered by their subscripts. This means there is no need for $\nabla^p x_i$, and we can work exclusively with $\nabla^p x_i$. Define the sublevel sets as

$$M^{c_i} := \text{dist}_p^{-1}[0, c_i].$$
2.2 The Definition of Sub-Index and Main Results

The set $\uparrow^{p}_{x_0}$ plays an important role in defining $x_0$ as a critical point. Specifically, $x_0$ is a critical point for $\text{dist}_p$ if for all $v$ in $T_{x_0}M$, we have $\langle v, \uparrow^{p}_{x_0} \rangle \leq \frac{\pi}{2}$.

Unfortunately, the set $\uparrow^{p}_{x_0}$ is too unwieldy to be of further use on its own. Instead, we consider the set given by

$$A(\uparrow^{p}_{x_0}) := \left\{ v \in S_{x_0} \mid \langle v, \uparrow^{p}_{x_0} \rangle \geq \frac{\pi}{2} \right\}$$

where $S_{x_0}$ is the unit tangent sphere at $x_0$. The set $A(\uparrow^{p}_{x_0})$ is at maximal distance from $\uparrow^{p}_{x_0}$ in the unit tangent sphere, and its structure is well understood.

If $A(\uparrow^{p}_{x_0})$ is not empty, then for each $w$ in $\uparrow^{p}_{x_0}$ the set of vectors $v$ in the unit tangent sphere at $x_0$ such that $\langle v, w \rangle \geq \frac{\pi}{2}$ form a hemisphere. This means $A(\uparrow^{p}_{x_0})$ is the intersection of overlapping hemispheres in $S_{x_0}$. So, $A(\uparrow^{p}_{x_0})$ is a convex, totally geodesic submanifold of the unit tangent sphere. Being able to identify the structure of $A(\uparrow^{p}_{x_0})$ is a crucial component in our definition of sub-index and the results that follow.

**Example 23.** Suppose $\uparrow^{p}_{x_0}$ consists of two antipodal points $w_1$ and $w_2$ on the 2-sphere.

![Figure 2.1](image-url)
In Figure 2.1, the shaded area of the first sphere represents the set of vectors in the unit tangent sphere of \( x_0 \) that are at least \( \frac{\pi}{2} \) away from \( w_1 \). Similarly, the shaded area of the second sphere represents the vectors that are at least \( \frac{\pi}{2} \) away from \( w_2 \). The third sphere illustrates the set of vectors that are at least \( \frac{\pi}{2} \) away from both \( w_1 \) and \( w_2 \), i.e. \( A(\mathcal{U}_{x_0}) \).

Based on this framework, we present our definition of sub-index.

**Definition 24.** If \( x_0 \) is an isolated critical point of \( \text{dist}_p \), its sub-index is given by

\[
\lambda := \begin{cases} 
    n & \text{if } A(\mathcal{U}_{x_0}) = \emptyset \\
    n - 1 - \dim A(\mathcal{U}_{x_0}) & \text{if } A(\mathcal{U}_{x_0}) \neq \emptyset \text{ but } \partial A(\mathcal{U}_{x_0}) = \emptyset \\
    n & \text{if } \partial A(\mathcal{U}_{x_0}) \neq \emptyset 
\end{cases}
\]

**Example 25.** Let \( M \) be the flat 2-torus given by the rectangle in Figure 2.2 with opposite sides identified.

\[
M: \quad S_{x_i}:
\]

Figure 2.2: The flat 2-torus and its unit tangent space at \( x_i \) for \( i = 1, 2, 3 \)

With \( p \) as the center, we show that the points \( p, x_1, x_2, \) and \( x_3 \) are critical for \( \text{dist}_p \). They correspond to the points \( x, y, z, \) and \( w \), respectively, as shown in Example 7. We also show that the sub-index for critical points in Figure 2.2 is the same as the index of the corresponding critical points in Example 7.
First, consider $p$. Since there are no segments connecting $p$ to itself, $\mathbf{\uparrow}_p^p = \emptyset$. This implies that for all vectors $v \in T_pM$ we have $\lang (v, \mathbf{\uparrow}_p^p) \rang \leq \frac{\pi}{2}$, i.e. $p$ itself is a critical point for $\text{dist}_p$. In this case, $A(\mathbf{\uparrow}_p^p) = S_p$, which is a unit circle, and $\partial A(\mathbf{\uparrow}_p^p) = \emptyset$. Thus, $p$ is a minimum since its sub-index is

$$\lambda = n - 1 - \dim A(\mathbf{\uparrow}_p^p) = 2 - 1 - 1 = 0.$$ 

For the point $x_1$, the set $\mathbf{\uparrow}_{x_1}^p = \{-u, u\}$ as depicted on the unit tangent circle in Figure 2.2. So, for all vectors $v \in T_{x_1}M$ we have $\lang (v, \mathbf{\uparrow}_{x_1}^p) \rang \leq \frac{\pi}{2}$, which implies $x_1$ is a critical point. Note that $A(\mathbf{\uparrow}_{x_1}^p) = \{-w, w\}$, which means $\partial A(\mathbf{\uparrow}_{x_1}^p) = \emptyset$. Thus, the index is

$$\lambda = n - 1 - \dim A(\mathbf{\uparrow}_{x_1}^p) = 2 - 1 - 0 = 1.$$ 

The situation for $x_2$ is completely analogous to that of $x_1$.

The point $x_3$ is critical for $\text{dist}_p$ since $\mathbf{\uparrow}_{x_3}^p = \{\pm w, \pm u\}$ means that for all vectors $v \in T_{x_3}M$ we have $\lang (v, \mathbf{\uparrow}_{x_3}^p) \rang \leq \frac{\pi}{2}$. Then $A(\mathbf{\uparrow}_{x_3}^p) = \emptyset$ since there are no vectors in $S_{x_3}$ at least $\frac{\pi}{2}$ away from $\mathbf{\uparrow}_{x_3}^p$. Thus, the index is $\lambda = n = 2$, and $x_3$ is a maximum.

Therefore, the sub-index for critical points in Figure 2.2 matches the index of the corresponding critical points in Example 7.

Our main theorems are as follows.

**Theorem 26.** (Connectedness Theorem) Let $x_0$ be an isolated critical point for $\text{dist}_p$ with $\text{dist}_p(x_0) = c_0$ and sub-index $\lambda$. Then the inclusion $M^{c_0-\epsilon} \rightarrow M^{c_0+\delta}$ is $(\lambda - 1)$-connected, where $\epsilon < c_0 - c_{-1}$ and $\delta < c_1 - c_0$. In other words,

$$\pi_i(M^{c_0+\delta}, M^{c_0-\epsilon}) = 0$$

for $i = 0, 1, \ldots, (\lambda - 1)$.
Theorem 27. (Relative $\pi_1$ Theorem) Let $x_0$ be an isolated critical point for $\text{dist}_p$ with $\text{dist}_p(x_0) = c_0$. Suppose

$$\pi_1(M^{c_0 + \delta}, M^{c_0 - \epsilon}) \neq 0,$$

for $\epsilon < c_0 - c_1$ and $\delta < c_1 - c_0$. Then $\hat{\pi}^p_{x_0}$ is a pair of antipodal points, i.e. there are only two segments from $p$ to $x_0$ and they make angle $\pi$ at $x_0$. Moreover, the ends of these segments are not conjugate along the segments.

The proof of Theorem 26 is divided into three cases based on the definition of sub-index, meaning that the structure of $A(\hat{\pi}^p_{x_0})$ plays a key role. The necessary technical lemmas are presented in Section 2.3, and the proof is given in Section 2.4. For the general idea of the proof note that if $A(\hat{\pi}^p_{x_0})$ is empty, all vectors along segments emanating from $x_0$ point in a direction of decrease. This means $x_0$ is a local maximum. So, any cell of dimension less than $n$ can be deformed into $M^{c_0}$.

For the other two cases, $A(\hat{\pi}^p_{x_0})$ is not empty, and we consider a $k$-dimensional cell $E^k$, which is a subset of $\text{int}(M^{c_1} \setminus M^{c - 1})$ with its boundary in $\text{int} M^{c_0}$. To prove the theorem, we show that a flow can be created to move $E^k$, with $k = 1, \ldots, (\lambda - 1)$, into $\text{int} M^{c_0}$ while leaving the boundary of $E^k$ fixed. When the boundary of $A(\hat{\pi}^p_{x_0})$ is empty, $A(\hat{\pi}^p_{x_0})$ is a great subsphere. In this case, the key idea is that transversality allows $E^k$ to be moved away from $A(\hat{\pi}^p_{x_0})$.

If both $A(\hat{\pi}^p_{x_0})$ and its boundary are not empty, $A(\hat{\pi}^p_{x_0})$ contains a vector $w_s$ such that

$$A(\hat{\pi}^p_{x_0}) \subset B(w_s, \frac{\pi}{2}).$$

Extending $-w_s$ to a vector field near $x_0$ produces a local flow. The local flow can be glued to a global flow that will ultimately move $E^k$, with $k = 1, \ldots, (n-1)$, into $\text{int} M^{c_0}$.
2.3 Technical Lemmas for the Connectedness Theorem

For the critical point $x_0$, the following lemma shows that for a short time the distance along a geodesic in any direction from $x_0$ has a linear approximation. It will be used to determine when points of $M$ are in a particular sublevel set.

**Lemma 28.** Given $\epsilon > 0$, there exists $\rho > 0$ such that for all $v \in S_{x_0}$

$$c_0 - t \cdot \cos \langle (v, \hat{\mathbf{p}}_{\mathbf{x}_0}) \rangle - \epsilon \cdot t \leq \text{dist}_p \left( \exp_{x_0}(tv) \right) \leq c_0 - t \cdot \cos \langle (v, \hat{\mathbf{p}}_{\mathbf{x}_0}) \rangle + \epsilon \cdot t$$

for all $t \in [0, \rho]$.

**Proof.** Let $\epsilon > 0$. Suppose $v \in S_{x_0}$ and $\gamma_v(t)$ is the segment from $x_0$ to $\exp_{x_0}(tv)$ such that $\gamma_v'(0) = v$. Since $\text{dist}_p$ is directionally differentiable,

$$D_v(\text{dist}_p x_0) = -\cos \langle (v, \hat{\mathbf{p}}_{\mathbf{x}_0}) \rangle.$$

So, the Taylor polynomial representation of $\text{dist}_p \left( \exp_{x_0}(tv) \right)$ is given by

$$\text{dist}_p \left( \exp_{x_0}(tv) \right) = c_0 - t \cdot \cos \langle (v, \hat{\mathbf{p}}_{\mathbf{x}_0}) \rangle + o(t).$$

Choose $\rho_v > 0$, depending on $v$, such that for all $t \in [0, \rho_v]$ we have

$$c_0 - t \cdot \cos \langle (v, \hat{\mathbf{p}}_{\mathbf{x}_0}) \rangle - \epsilon \cdot t \leq \text{dist}_p \left( \exp_{x_0}(tv) \right) \leq c_0 - t \cdot \cos \langle (v, \hat{\mathbf{p}}_{\mathbf{x}_0}) \rangle + \epsilon \cdot t. \quad (2.1)$$

By continuity there exists a neighborhood $W_v$ of $v$, on which the inequalities in (2.1) are valid. In fact, we can find such a neighborhood for each $v$ in $S_{x_0}$. So, the set of such neighborhoods forms an open cover of $S_{x_0}$. Since $S_{x_0}$ is compact, there exists a finite subcover, say $\{W_i\}_{i=1}^k$. Define $\rho$ to be the minimum of $\{\rho_v\}_{i=1}^k$. Thus, for all $v$ in $S_{x_0}$ the inequalities (2.1) will hold on the interval $[0, \rho]$.

The next lemma establishes a set of points, given by $N$, that lie in $\text{int}M^{\circ}$.

Only part (1) is used in the remainder of this work.
Lemma 29. Given $\delta$ sufficiently small and $U_\delta$, the $\delta$-neighborhood of $A(\hat{\|}_x)$, there exists $R > 0$ such that:

1) for $N := \exp_{x_0} \left( t(S_{x_0} \setminus U_\delta) \right)$ with $t \in (0, 2R]$ we have $N \subset \text{int} M^{\infty}$ and

2) $\left( B(x_0, 2r) \setminus B(x_0, r) \right) \cap M^{\infty} \cap \frac{2r}{3} \subset N$ for $r \in (0, R]$.

Proof. Choose $\delta \in (0, \frac{\pi}{2})$ such that $9\frac{20}{3} - \cos(\frac{\pi}{2} - \delta) > 0$. Since the distance from $\hat{\|}_x$ is a continuous function on the compact set $S_{x_0} \setminus U_\delta$, a maximum angle exists, say $\alpha_1$. Choose $\epsilon_1 < \cos \alpha_1$. Then by Lemma 28 there exists $\rho_1 > 0$ such that

$$\text{dist}_p \left( \exp_{x_0} (tv) \right) \leq c_0 - t \cdot \cos \angle(v, \hat{\|}_{x_0}) + \epsilon_1 \cdot t$$

on the set $N_1 := \left\{ \exp_{x_0} (tv) \mid v \in S_{x_0} \setminus U_\delta, \ t \in [0, \rho_1] \right\}$. Now we show that $N_1 \subset \text{int} M^{\infty}$.

Since $\alpha_1 = \max \left\{ \angle(v, \hat{\|}_{x_0}) \right\}$ over all $v$ in $S_{x_0} \setminus U_\delta$ and $\epsilon_1 < \cos \alpha_1$ on $N_1$, we have

$$\text{dist}_p \left( \exp_{x_0} (tv) \right) \leq c_0 - t \cdot \cos \angle(v, \hat{\|}_{x_0}) + \epsilon_1 \cdot t$$

$$\leq c_0 - t \cdot \cos \alpha_1 + \epsilon_1 \cdot t$$

$$< c_0$$

Thus, $N_1 \subset \text{int} M^{\infty}$.

Similarly, since the distance from $\hat{\|}_{x_0}$ is a continuous function on the compact set $\overline{U}_\delta$, a minimum angle exists, say $\alpha_2$. Choose $\epsilon_2 < \frac{9}{20} - \cos \alpha_2$. Then by Lemma 28 there exists $\rho_2 > 0$ such that on the set $N_2 := \left\{ \exp_{x_0} (tv) \mid v \in \overline{U}_\delta, \ t \in [0, \rho_2] \right\}$

$$c_0 - t \cdot \cos \angle(v, \hat{\|}_{x_0}) - \epsilon_2 \cdot t \leq \text{dist}_p \left( \exp_{x_0} (tv) \right).$$

Using $\alpha_2$ and $\epsilon_2$, on $N_2$ we have

$$\text{dist}_p \left( \exp_{x_0} (tv) \right) \geq c_0 - t \cdot \cos \alpha_2 - \epsilon_2 \cdot t$$

$$> c_0 - t \cdot \cos \alpha_2 - t \left( \frac{9}{20} - \cos \alpha_2 \right)$$

$$= c_0 - \frac{9}{20} t.$$
Define $R := \frac{1}{2} \min\{\rho_1, \rho_2\}$ and for $r \leq R$

\[
N := \left\{ \exp_{x_0}(tv) \mid v \in S_{x_0} \setminus U_\delta, \; t \in (0, 2r) \right\}.
\]

Since $N \subset N_1$, the first part of the lemma has been satisfied. It remains to show the second part.

Consider the points $\exp_{x_0}(tv)$ that are in the annulus $\left( B(x_0, 2r) \setminus B(x_0, r) \right)$ but are not in $N$, i.e. the points in

\[
N_3 := \left\{ \exp_{x_0}(tv) \mid v \in U_\delta, \; t \in [r, 2r) \right\}.
\]

Using contrapositive, we need to show that $N_3$ is not in $M^{c_0-\frac{2}{10}R}$. Since $t < 2r \leq \rho_2$, we have $N_3 \subset N_2$. So on $N_3$,

\[
\text{dist}_p\left( \exp_{x_0}(tv) \right) > c_0 - \frac{9}{20} t > c_0 - \frac{9}{10} r.
\]

Thus, the points in $N_3$ are not in $M^{c_0-\frac{2}{10}R}$. \hfill \Box

**Lemma 30.** (Local Reduction Lemma) Suppose for $R > 0$, $\overline{B}(x_0, 2R)$ is contained in $\text{int}(M^{c_1} \setminus M^{c^{-1}})$. Then $M^{c_0-\frac{2}{10}R} \cup \overline{B}(x_0, R)$ is a strong deformation retract of $\text{int}(M^{c_1})$.

---

*Figure 2.3: The sets involved in Lemma 30*
Proof. Since \( \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}) \) consists only of regular points, we can define a negative gradient-like vector field \( X \) on it. Then \( X \) defines a local flow \( \psi(y, t) \). In order to create a deformation retraction, we need to consider how long it takes each point \( y \) in \( \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}) \) to end up in \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \). Define the function

\[
\tau : \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}) \rightarrow \mathbb{R}
\]

to be the minimum amount of time that it takes \( y \) to arrive in \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \) as it flows with \( \psi \). Since each \( y \) in \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \) is already in the desired set, we have \( \tau(y) = 0 \).

Now we create a strong deformation retraction of the set \( \text{int}M^c \) into the set \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \). Define \( \phi : \text{int}M^c \times [0, 1] \rightarrow \text{int}M^c \) by

\[
\phi(y, t) := \begin{cases} 
\psi \left(y, \tau(y) \cdot t\right) & \text{if } y \in \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}) \\
y & \text{if } y \in M^{c-1} \cup \{x_0\}
\end{cases}
\]

At \( t = 0 \), the points in \( M^{c-1} \cup \{x_0\} \) remain fixed, and for

\[
y \in \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}),
\]

we have \( \phi(y, 0) = \psi(y, 0) = y \). For \( t = 1 \), the points in \( M^{c-1} \cup \{x_0\} \) remain fixed in \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \), and for the points \( y \) in \( \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}) \), we have

\[
\phi(y, 1) = \psi(y, \tau(y)) \in M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R).
\]

Further, for all \( y \) in \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \), we either have \( \phi(y, t) = y \) when \( y \) is in \( M^{c-1} \cup \{x_0\} \) or for \( y \in \text{int}M^c \setminus (M^{c-1} \cup \{x_0\}) \) we know \( \tau(y) = 0 \) and

\[
\phi(y, t) = \psi \left(y, \tau(y) \cdot t\right) = \psi(y, 0) = y.
\]

Therefore, \( \phi \) is a strong deformation retraction of \( \text{int}M^c \) into \( M^{\operatorname{co}} \frac{\partial}{\partial R} \cup \overline{B}(x_0, R) \). ∎
2.4 Proof of the Connectedness Theorem

In this section, we restate Theorem 26 and present its proof.

**Theorem.** Let $x_0$ be an isolated critical point for $\text{dist}_p$ with $\text{dist}_p(x_0) = c_0$ and sub-index $\lambda$. Then the inclusion $M^{c_0-\epsilon} \hookrightarrow M^{c_0+\delta}$ is $(\lambda - 1)$-connected, where $\epsilon < c_0 - c_1 - 1$ and $\delta < c_1 - c_0$. In other words,

$$\pi_i(M^{c_0+\delta}, M^{c_0-\epsilon}) = 0$$

for $i = 0, 1, \ldots, (\lambda - 1)$.

**Proof. Case 1:** Suppose $A(\hat{\eta}_x^{P_{x_0}}) = \emptyset$. Then there are no vectors $v \in S_{x_0}$ such that

$$\langle v, \hat{\eta}_x^{P_{x_0}} \rangle \geq \frac{\pi}{2}.$$ 

Since $x_0$ is a critical point, we know that $\langle v, \hat{\eta}_x^{P_{x_0}} \rangle \leq \frac{\pi}{2}$ for all tangent vectors $v$. So for all $v \in S_{x_0}$, we must have $\langle v, \hat{\eta}_x^{P_{x_0}} \rangle < \frac{\pi}{2}$. From this, the directional derivative tells us that $D_v(\text{dist}_p x_0) < 0$ for all $v \in T_{x_0}M$, meaning that the distance between $x_0$ and $p$ decreases regardless of the direction we travel away from $x_0$. Thus, the point $x_0$ must be a maximum. This means a cell of dimension $1, 2, \ldots, (n-1)$ within $\text{int}(M^{c_1} \setminus M^{c_1-1})$ with boundary in $\text{int}M^{c_0}$ can be deformed into $\text{int}M^{c_0}$. Therefore, $\pi_i(M^{c_1}, M^{c_0}) = 0$ for $i = 0, 1, \ldots, (n-1)$.

**Set up for cases 2 and 3:** For the remaining two cases choose $\delta \in (0, \frac{\pi}{2})$ such that

$$\frac{9}{20} - \cos \left( \frac{\pi}{2} - \delta \right) > 0. \tag{2.2}$$

Let $U_\delta$ be the $\delta$-neighborhood of $A(\hat{\eta}_x^{P_{x_0}})$. Then by Lemma 29 there exists $R > 0$ such that $N \subset \text{int}M^{c_0}$ where

$$N := \exp_{x_0} \left( t(S_{x_0} \setminus U_\delta) \right) \quad \text{for} \quad t \in (0, 2R).$$
Let $E^k$ be a $k$-cell with $k = 0, 1, \ldots, (\lambda - 1)$ such that $E^k \subset \text{int}(M^{c_1} \setminus M^{c - 1})$ and $\partial E^k \subset \text{int}M^{c_0}$. Since $c_0 - \text{dist}_p$ is a continuous function on the compact set $\partial E^k$, there exists a minimum, say $m$. Then

$$\text{dist}_p(\partial E^k) \leq c_0 - m.$$  

Choose $r < \min\left\{2R, \frac{1}{2}m\right\}$ so that $B(x_0, 2r)$ is contained in $\text{int}(M^{c_1} \setminus M^{c - 1})$ and $2r$ is smaller than the injectivity radius at $x_0$. Then $\overline{B}(x_0, r)$ is a ball around $x_0$ in $\text{int}(M^{c_1} \setminus M^{c - 1})$, and by the Local Reduction Lemma $M^{c_0} - \frac{\mu}{2}r \cup \overline{B}(x_0, \frac{1}{2}r)$ is a strong deformation retract of $\text{int}M^{c_1}$. Since $r < \frac{1}{2}m$ we have

$$\text{dist}_pB(x_0, 2r) > c_0 - 2r > c_0 - m.$$  

This means $\partial E^k$ is outside the $2r$-ball. So the strong deformation retract moves $E^k$ into $M^{c_0} - \frac{\mu}{2}r \cup \overline{B}(x_0, \frac{1}{2}r)$ while keeping $\partial E^k$ fixed. Since $r < 2R$, we know from Lemma 29 that

$$N_r := \left\{ \exp_{x_0}(t(S_{x_0} \setminus U_\delta)) \mid t \in [0, r] \right\} \subset \text{int}M^{c_0}.$$  

It remains to show that we can create a homotopy that fixes $\partial E^k$ and moves $E^k \cap \overline{B}(x_0, \frac{1}{2}r)$ into $N_r$, which we know is a subset of $\text{int}M^{c_0}$.

**Case 2:** Suppose $A(\uparrow_{x_0}^p) \neq \emptyset$ but $\partial A(\uparrow_{x_0}^p) = \emptyset$. Define

$$C_r A(\uparrow_{x_0}^p) := \left\{ \exp_{x_0}(tA(\uparrow_{x_0}^p)) \mid t \in [0, r] \right\}.$$  

Note that the sum of the dimension of the cell and the dimension of $C_r A(\uparrow_{x_0}^p)$ yields:

$$\dim E^k + \dim C_r A(\uparrow_{x_0}^p) \leq \left( \lambda - 1 \right) + \left( \dim A(\uparrow_{x_0}^p) + 1 \right) = \lambda + \dim A(\uparrow_{x_0}^p) = \left( n - 1 - \dim A(\uparrow_{x_0}^p) \right) + \dim A(\uparrow_{x_0}^p) = n - 1 < n.$$
This means by transversality we can apply a small homotopy so that

$$\left\{ E_k \cap B(x_0, \frac{1}{2}r) \right\} \cap C_r A(\uparrow x_0) = \emptyset$$

and the points outside $B(x_0, \frac{1}{2}r)$ remained fixed. Since

$$\left\{ E_k \cap B(x_0, \frac{1}{2}r) \right\} \subset B(x_0, r) \setminus C_r A(\uparrow x_0),$$

the following lemma shows that $E_k \cap B(x_0, \frac{1}{2}r)$ can be moved into $N_r \setminus B(x_0, \frac{1}{2}r)$, a subset of $\text{int} M^{co}$, while keeping the boundary of the cell fixed. This will complete the proof.

**Lemma 31.** There exists an isotopy of $M \setminus C_r A(\uparrow x_0)$ that fixes $M \setminus B(x_0, r)$ and restricts to a strong deformation retract of $B(x_0, r) \setminus C_r A(\uparrow x_0)$ onto $N_r \setminus B(x_0, \frac{1}{2}r)$.

**Proof.** First, use radial geodesics from $x_0$ to deform $B(x_0, r) \setminus \{x_0\}$ onto

$$B(x_0, r) \setminus B(x_0, \frac{1}{2}r).$$

Since $C_r A(\uparrow x_0)$ is a union of these radial geodesics, this restricts to an isotopy of $B(x_0, r) \setminus C_r A(\uparrow x_0)$ to

$$B(x_0, r) \setminus \left\{ B(x_0, \frac{1}{2}r) \cup C_r A(\uparrow x_0) \right\}.$$

Now we move the points into $N_r \setminus B(x_0, \frac{1}{2}r)$. For any $\delta \in (0, \frac{\pi}{2})$ the set $S_{x_0} \setminus U_\delta$ is a strong deformation retract of $S_{x_0} \setminus A(\uparrow x_0)$. Exponentiating this retract gives an isotopy of $B(x_0, r) \setminus C_r A(\uparrow x_0)$ that leaves the metric spheres around $x_0$ invariant. Thus, it carries $B(x_0, r) \setminus \left\{ B(x_0, \frac{1}{2}r) \cup C_r A(\uparrow x_0) \right\}$ into $N_r \setminus B(x_0, \frac{1}{2}r)$ without moving points in $N_r \setminus B(x_0, \frac{1}{2}r)$. Since the strong deformation retract can be given by a vector field, we can use a partition of unity to glue it to the zero vector field on $M \setminus B(x_0, r)$ so that the result will be an isotopy of $M \setminus C_r A(\uparrow x_0)$ that fixes $M \setminus B(x_0, r)$. □
**Case 3:** Suppose that both $A(\hat{p}_x^P)$ and its boundary are not empty. By definition, $A(\hat{p}_x^P)$ consists of unit tangent vectors at least $\frac{\pi}{2}$ away from $\hat{p}_x^P$. So, $A(\hat{p}_x^P)$ is a subset of a $\frac{\pi}{2}$-ball in the unit tangent sphere. This means we can choose a vector $w_s$ in $A(\hat{p}_x^P)$ such that
\[
A(\hat{p}_x^P) \subset \overline{B}(w_s, \frac{\pi}{2}).
\] (2.3)

Given that $A(\hat{p}_x^P)$ is a positively curved manifold with boundary, it can be shown that the soul satisfies this condition.

Using a partition of unity, define a vector field $W$ on $M$ to be $dexp_{x_0}(-w_s)$ on $B\left(x_0, \frac{3}{2}r\right)$ and supported on $B(x_0, 2r)$. Let $\psi(y, t)$ be the flow defined from $W$. Define
\[
\tau : M \longrightarrow \mathbb{R}
\]
to be the shortest time it takes for a point to either arrive in $N_r$ or leave $B(x_0, 2r)$ as it flows with $\psi$. Then the desired homotopy $\Psi : M \times [0, 1] \longrightarrow M$ is given by
\[
\Psi(y, t) = \psi\left(y, \tau(y) \cdot t\right).
\]

Note that for $y \in E_k \setminus B(x_0, 2r)$ we have $\tau(y) = 0$ so $\Psi(y, t) = y$ for all $t$. This means the boundary of the cell remains fixed during the homotopy.

Since the cell is in $M^{\alpha - \frac{3}{2}r} \cup \overline{B}\left(x_0, \frac{1}{2}r\right)$, it is enough to show that points in $\overline{B}\left(x_0, \frac{1}{2}r\right)$ flow into $N_r$ before they leave $B(x_0, 2r)$. In order to verify this, it is more convenient to work in $S_{x_0}$. However, the field $dexp_{x_0}(-w_s)$ may not be of unit length, thereby causing a distortion. To compensate for this, we further restrict $r$ so that the flow will take at least $\frac{7}{5}r$ to move each $y \in \overline{B}\left(x_0, \frac{1}{2}r\right)$ out of $B(x_0, 2r)$. By the triangle inequality, we have
\[
\left\|exp_{x_0}^{-1}(y) - rw_s\right\| \leq \frac{7}{5}r < \frac{3}{2}r.
\] (2.4)
For \( y \in \overline{B}(x_0, \frac{1}{2}r) \) we now claim that

\[
\theta := \angle \left( \exp_{x_0}^{-1}(y) - rw_s, \ w_s \right) > \frac{\pi}{2} + \delta.
\]

By (2.3), the claim implies that \( \exp_{x_0}^{-1}(y) - rw_s \in S_{x_0} \setminus U_\delta \), so

\[
\Psi(y, r) \subset \left( N_r \cap B(x_0, 2r) \right).
\]

To prove the claim, first note that

\[
\cos \theta = \frac{g \left( \exp_{x_0}^{-1}(y) - rw_s, \ w_s \right)}{\| \exp_{x_0}^{-1}(y) - rw_s \|} = \frac{g \left( \exp_{x_0}^{-1}(y), \ w_s \right) - r}{\| \exp_{x_0}^{-1}(y) - rw_s \|} \leq \frac{\| \exp_{x_0}^{-1}(y) \| - r}{\| \exp_{x_0}^{-1}(y) - rw_s \|} \text{ by the Cauchy-Schwarz inequality}
\]

\[
\leq -\frac{r}{2\| \exp_{x_0}^{-1}(y) - rw_s \|} \text{ since } y \in \overline{B}(x_0, \frac{1}{2}r)
\]

\[
< -\frac{1}{3}. \tag{2.5}
\]

The last inequality is due to (2.4). By (2.2), \( \delta \in (0, \frac{\pi}{2}) \) must satisfy \( \frac{9}{20} - \cos(\frac{\pi}{2} - \delta) > 0 \),

which means \( \frac{1}{3} - \cos(\frac{\pi}{2} - \delta) > 0 \). This implies

\[
-\frac{1}{3} < -\cos \left( \frac{\pi}{2} - \delta \right) = -\sin \delta = \cos \left( \frac{\pi}{2} + \delta \right). \tag{2.6}
\]

Combining inequalities (2.5) and (2.6), we have

\[
\cos \theta < -\frac{1}{3} < \cos \left( \frac{\pi}{2} + \delta \right)
\]

which means \( \theta > \frac{\pi}{2} + \delta \) as claimed. \( \square \)
2.5 Lemmas Related to Conjugate Points

For the Relative $\pi_1$ Theorem, we use a proof by contradiction to show that $p$ is not conjugate to $x_0$ along the two segments between them. In this section, we give the definition of conjugate point as well as lemmas related to both the presence of conjugate points and the lack of conjugate points.

Throughout this section, suppose $x_0$ is a critical point for $\text{dist}_p$ and $v \in \pi_{x_0}^p$. Let $\gamma_v : [0, 1] \to M$ be the geodesic from $\gamma_v(0) = x_0$ to $\gamma_v(1) = p$ with $\gamma_v'(0) = v$.

**Definition 32.** The point $p$ is conjugate to $x_0$ along $\gamma_v(t) = \exp_p(tv)$ if there exists a nonzero Jacobi field $J$ along $\gamma_v$ with $J(0) = 0$ and $J(1) = 0$.

Note that if $p$ is conjugate to $x_0$ along $\gamma_v$, then $\ker (d\exp_{x_0})_v$ is not zero. The next lemma is a special case of a result in [4] but is presented with a different proof.

**Lemma 33.** Suppose $w \in S_{x_0}$ is orthogonal to $\ker (d\exp_{x_0})_v$. Then there is a unique Jacobi field $J_w$ along $\gamma_v$ so that $J_w(0) = w$ and $J_w(1) = 0$.

**Proof.** Let $\mathcal{N}$ be the family of nonzero Jacobi fields $N$ so that $N(0) = N(1) = 0$.

Let $\mathcal{P}$ be the family of Jacobi fields $P$ so that $P(1) = 0$ and $P'(1) \perp N'(1)$ for all $N \in \mathcal{N}$.

We have

$$\ker (d\exp_{x_0})_v = \left\{ N'(0) \mid N \in \mathcal{N} \right\}.$$
Since the Riccati operator on $N \oplus P$ is self adjoint we know that for all $P \in P$ and all $N \in N$

$$g \left( P, N' \right) \bigg|_0 = g \left( P', N \right) \bigg|_0 = 0.$$  

We conclude that the set $\{ P (0) \mid P \in P \}$ is precisely the orthogonal complement of $\ker (d \exp_{x_0})_v$. So given any $w \perp \ker (d \exp_{x_0})_v$, choose $J_w$ to be the unique $P \in P$ with $P (0) = w$. \hfill \Box

Given $w \in S_{x_0}$ let $c_w(t)$ be the geodesic such that $c'_w(0) = w$ and $c_w(0) = x_0$.

For $H \in \mathbb{R}$, define

$$T_{2,H}^{x_0,w}(t) := \text{dist}_p x_0 - t \cdot \cos \varangle (w, \bar{P}_{x_0}^p) + \frac{1}{2} H \cdot t^2.$$  

**Lemma 34.** Let $H$ be:

(1) $g \left( J'_w(0), J_w(0) \right)$ if $w \in S_{x_0}$ is orthogonal to $\ker (d \exp_{x_0})_v$,

(2) any number if $w \in S_{x_0}$ is not orthogonal to $\ker (d \exp_{x_0})_v$.

Then there exists an interval $[0, m]$, depending on $w$, for which

$$\text{dist}_p \left( c_w(t) \right) \leq T_{2,H}^{x_0,w}(t) + o(t^2).$$

**Proof.** Given $v \in \bar{P}_{x_0}^p$ and $w \in S_{x_0}$, it suffices to find a vector field $V$ along $\gamma_v$ with $V(0) = w$ and $V(1) = 0$ so that $I(V, V) \leq H$. This is because given a vector field $V$ along $\gamma_v$, there is a variation $\tilde{\gamma}$ whose variation field is $V$. By the first and second variation formulas, we know

$$\left. \frac{d\text{Len}(\tilde{\gamma})}{ds} \right|_{s=0} = -\cos \varangle (w, v) \quad \text{and} \quad \left. \frac{d^2\text{Len}(\tilde{\gamma})}{ds^2} \right|_{s=0} = I(V, V).$$

So the Taylor polynomial gives us

$$\text{Len}(\tilde{\gamma}) \leq \text{Len}(\gamma_v) - t \cdot \cos \varangle (w, v) + \frac{1}{2} t^2 \cdot I(V, V) + o(t^2).$$
Since distance is the minimum of length, there is an interval on which

\[ \text{dist}_p(c_w(t)) \leq \text{dist}_p(x_0 - t \cdot \cos \angle(w, \hat{p}_{x_0}) + \frac{1}{2}t^2 \cdot I(V, V) + o(t^2). \]

Suppose \( w \) is orthogonal to \( \text{ker}(d\text{exp}_{x_0})_v \). By Lemma 33, there is a Jacobi field \( J_w \) along \( \gamma_v \) with \( J_w(0) = w \) and \( J_w(1) = 0 \). So for \( H = I(J_w, J_w) = -g(J_w'(0), J_w(0)) \), the result holds.

Now suppose \( w \) is not orthogonal to \( \text{ker}(d\text{exp}_{x_0})_v \). First, we consider the special case when \( w \) is in \( \text{ker}(d\text{exp}_{x_0})_v \). By lemma 33, there exists a nonzero Jacobi field \( J_w \) along \( \gamma_v \) such that \( J_w(0) = 0 \) and \( J_w'(0) = w \). From this we create a vector field that does not vanish at both ends. Specifically, define a vector field \( V_\epsilon \) by

\[
V_\epsilon(t) := \begin{cases} 
J(t) \cdot \left( \|J(\epsilon)\| \right)^{-1} & \text{if } t \in (\epsilon, 1] \\
W_\epsilon(t) & \text{if } t \in [0, \epsilon]
\end{cases}
\]

where \( W_\epsilon \) is the Jacobi field with \( W_\epsilon(\epsilon) = \frac{J(\epsilon)}{\|J(\epsilon)\|} \) and \( W_\epsilon(0) = J'(0) = w \). Then the index form is given by

\[
I(V_\epsilon, V_\epsilon) = g\left( \frac{J'(\epsilon)}{\|J(\epsilon)\|}, \frac{J(\epsilon)}{\|J(\epsilon)\|} \right) - g\left( \frac{J'(\epsilon)}{\|J(\epsilon)\|}, \frac{J(\epsilon)}{\|J(\epsilon)\|} \right) + g\left( W_\epsilon'(\epsilon), W_\epsilon(\epsilon) \right) - g\left( W_\epsilon'(0), W_\epsilon(0) \right)
\]

\[
= -\frac{1}{\|J(\epsilon)\|^2}g\left( J'(\epsilon), J(\epsilon) \right) + g\left( W_\epsilon'(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|} \right) - g\left( W_\epsilon'(0), J'(0) \right).
\]

Note that the limit of the first term gives us

\[
\lim_{\epsilon \to 0} \frac{-g\left( J'(\epsilon), J(\epsilon) \right)}{\|J(\epsilon)\|^2} = \lim_{\epsilon \to 0} \frac{-g\left( J'(\epsilon), J(\epsilon) \right)}{g\left( J(\epsilon), J(\epsilon) \right)}
\]

\[
= -\lim_{\epsilon \to 0} \frac{g\left( J''(\epsilon), J(\epsilon) \right) + g\left( J'(\epsilon), J'(\epsilon) \right)}{2g\left( J'(\epsilon), J(\epsilon) \right)}
\]

\[
= -\infty
\]
since $J^\prime(0) \neq 0$. So for an upper bound on $I(V_\epsilon, V_\epsilon)$, it suffices to bound

$$g\left(W_\epsilon^\prime(0), J^\prime(0)\right) \text{ and } g\left(W_\epsilon^\prime(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\right)$$

independent of $\epsilon$.

Let $\{E_i\}_{i=1}^{n-1}$ be an orthonormal parallel frame for the normal space of $\gamma_\epsilon$ with $E_1(0) = J^\prime(0)$. Write $J = \sum_{i=1}^{n-1} f_i E_i$ where each $f_i$ is a smooth function. Now we approximate each $f_i$. Since $J(0) = 0$, $f_i(0) = 0$ for all $i$. Given that $E_1(0) = J^\prime(0)$, we know $f_1^\prime(0) = 1$ and $f_i^\prime(0) = 0$ for all $i = 2, \ldots, n-1$. Since $J$ is a Jacobi field with $J(0) = 0$,

$$J^{\prime\prime}(0) = \sum_{i=1}^{n-1} f_i^{\prime\prime}(0) E_i(0) = -R\left(J(0), \gamma^{\prime}(0)\right)\gamma^{\prime}(0) = 0.$$

So $f_i^{\prime\prime}(0) = 0$ for all $i$. Using Taylor’s Theorem, there exists an interval on which

$$f_1(t) = t + O(t^3) \text{ and } f_i(t) = O(t^3) \text{ for } i = 2, \ldots, n-1.$$

We use this to approximate $\frac{J(t)}{\|J(t)\|}$. First, note that

$$\|J(t)\|^2 = \sum_{i=1}^{n-1} f_i^2(t) = t^2 + O(t^4) = t^2 \left(1 + O(t^2)\right).$$

Taking the square root, we have

$$\|J(t)\| = \sqrt{t^2(1 + O(t^2))} = t \left(1 + O(t^2)\right) = t + O(t^3).$$

Combining gives us

$$\frac{J(t)}{\|J(t)\|} = \frac{\sum_{i=1}^{n-1} f_i(t) E_i(t)}{t + O(t^3)}, \quad (2.7)$$

$$\frac{f_1(t)}{t + O(t^3)} = \frac{t + O(t^3)}{t + O(t^3)} = 1 + O(t^2),$$

and for $i \geq 2$

$$\frac{f_i(t)}{t + O(t^3)} = \frac{O(t^3)}{t + O(t^3)} = O(t^2).$$
In order to approximate $W'_\epsilon$, we write $W_\epsilon = \sum_{i=1}^{n-1} g_{\epsilon,i} E_i$, where each $g_{\epsilon,i}$ is a smooth function depending on $\epsilon$. Since the space of Jacobi fields with bounded endpoints is compact, there is a bound $B$ on $[0, \epsilon]$, independent of $\epsilon$, so that

$$\|g_{\epsilon,i}''\| = \|R(W_\epsilon, \gamma', \gamma', E_i)\| \leq B.$$  

Then

$$\|W''_\epsilon\| = \left\| \sum_{i=1}^{n-1} g_{\epsilon,i}'' E_i \right\| \leq B.$$  

Since $W_\epsilon(0) = J'(0) = w$ and $E_1(0) = J'(0)$, we have $g_{\epsilon,1}(0) = 1$ and $g_{\epsilon,i}(0) = 0$ for $i = 2, \ldots, n - 1$. Using Taylor’s Theorem and (2.8), there exists an interval $[0, m]$, independent of $\epsilon$, on which

$$g_{\epsilon,1}(t) = 1 + g_{\epsilon,1}'(0) t + \mathcal{O}(t^2)$$  

and

$$g_{\epsilon,i}(t) = g_{\epsilon,i}'(0) t + \mathcal{O}(t^2)$$  

for $i = 2, \ldots, n - 1$.

Given that $W_\epsilon(\epsilon) = \frac{J(\epsilon)}{\|J(\epsilon)\|}$ and (2.7), we have

$$g_{\epsilon,1}(\epsilon) = 1 + g_{\epsilon,1}'(0) \epsilon + \mathcal{O}(\epsilon^2) = \frac{f_1(\epsilon)}{\epsilon + \mathcal{O}(\epsilon^3)} = 1 + \mathcal{O}(\epsilon^2)$$  

and

$$g_{\epsilon,i}(\epsilon) = g_{\epsilon,i}'(0) \epsilon + \mathcal{O}(\epsilon^2) = \frac{f_i(\epsilon)}{\epsilon + \mathcal{O}(\epsilon^3)} = \mathcal{O}(\epsilon^2)$$  

for $i \geq 2$.

So, $g_{\epsilon,i}'(0) = \mathcal{O}(\epsilon)$ for all $i$. Since $W'_\epsilon(0) = \sum_{i=1}^{n-1} g_{\epsilon,i}'(0) E_i(0)$, we have

$$\left\| g\left(W'_\epsilon(0), J'(0)\right) \right\| = \left\| g\left(W'_\epsilon(0), w\right) \right\| \leq \|\mathcal{O}(\epsilon)\|. \quad (2.9)$$

To estimate $W''_\epsilon(\epsilon) = \sum_{i=1}^{n-1} g_{\epsilon,i}'(\epsilon) E_i(\epsilon)$ we need to bound $g_{\epsilon,i}'(\epsilon)$. Note that by the Fundamental Theorem of Calculus and the fact that $g_{\epsilon,i}'(0) = \mathcal{O}(\epsilon)$, we have

$$\|g_{\epsilon,i}'(\epsilon)\| = \left\| g_{\epsilon,i}'(0) + \int_0^\epsilon g_{\epsilon,i}''(t) dt \right\| \leq \|\mathcal{O}(\epsilon)\|.$$  

33
This means by the Cauchy-Schwarz inequality
\[ \left\| g\left( W'_{\epsilon}(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\right) \right\| \leq \|W'_{\epsilon}(\epsilon)\| \leq \|O(\epsilon)\|. \] (2.10)

Thus from (2.9) and (2.10) we have
\[ I(V_{\epsilon}, V_{\epsilon}) = -\frac{1}{\|J(\epsilon)\|^2} g\left( J'(\epsilon), J(\epsilon) \right) + g\left( W'_{\epsilon}(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\right) - g\left( W'_{\epsilon}(0), J'(0) \right) \]
\[ \leq -\frac{1}{\|J(\epsilon)\|^2} g\left( J'(\epsilon), J(\epsilon) \right) + \|O(\epsilon)\| \rightarrow -\infty \quad \text{as } \epsilon \rightarrow 0. \]

So choose any number for \( H \). Then choose \( 0 < \epsilon < m \) such that
\[ -\frac{1}{\|J(\epsilon)\|^2} g\left( J'(\epsilon), J(\epsilon) \right) + \|O(\epsilon)\| \leq H. \]

Now suppose \( w \) is not orthogonal to \( ker(dexp_{x_0})_{v} \) and \( w \) is not in \( ker(dexp_{x_0})_{v} \). Write \( w = w_{\text{tang}} + w_{\perp} \) with respect to \( ker(dexp_{x_0})_{v} \). Then there exists a Jacobi field \( U_{w} \) along \( \gamma_v \) with \( U_{w}(0) = w_{\perp} \) and \( U_{w}(1) = 0 \), and there exists a Jacobi field \( J \) such that \( J(0) = J(1) = 0 \) and \( J'(0) = w_{\text{tang}} \). Define the vector field \( V_{\epsilon} \) as in the proof of the special case previously discussed, and let \( V_{\epsilon,1} := \frac{J}{\|J(\epsilon)\|} \). Then
\[ I(U_{w}, U_{w}) = -g\left( U'_{w}(0), U_{w}(0) \right) \] (2.11)
and
\[ I(V_{\epsilon}, V_{\epsilon}) \leq -\frac{1}{\|J(\epsilon)\|^2} g\left( J'(\epsilon), J(\epsilon) \right) + \|O(\epsilon)\| \rightarrow -\infty \quad \text{as } \epsilon \rightarrow 0. \] (2.12)
So \( U_{w} + V_{\epsilon} \) is a vector field along \( \gamma_v \) with \( (U_{w} + V_{\epsilon})(0) = w \) and \( (U_{w} + V_{\epsilon})(1) = 0 \). Now consider
\[ I(U_{w} + V_{\epsilon}, U_{w} + V_{\epsilon}) = I(U_{w}, U_{w}) + 2I(U_{w}, V_{\epsilon}) + I(V_{\epsilon}, V_{\epsilon}). \] (2.13)
Based on (2.11) and (2.12), it remains to show that we have a bound on
\[ I(U_{w}, V_{\epsilon}) = g\left( V'_{\epsilon,1}(1), U_{w}(1) \right) - g\left( V'_{\epsilon,1}(\epsilon), U_{w}(\epsilon) \right) + g\left( W'_{\epsilon}(\epsilon), U_{w}(\epsilon) \right) - g\left( W'_{\epsilon}(0), U_{w}(0) \right) \]
\[ = -g\left( V'_{\epsilon,1}(\epsilon), U_{w}(\epsilon) \right) + g\left( W'_{\epsilon}(\epsilon), U_{w}(\epsilon) \right) - g\left( W'_{\epsilon}(0), U_{w}(0) \right). \]
From the proof of the special case when \( w \) is in \( \ker(d\exp_{x_0}) \), we know that 
\[
\|W'(\epsilon)\| \leq \|O(\epsilon)\| \quad \text{and} \quad \|W'(0)\| \leq \|O(\epsilon)\|.
\]
So
\[
\left\| g\left(W'(\epsilon), U_w(\epsilon)\right) \right\| \leq \|W'(\epsilon)\| \cdot \|U_w(\epsilon)\| \leq \|O(\epsilon)\| \cdot \|U_w(\epsilon)\|
\]
and
\[
\left\| g\left(W'(0), w_\perp\right) \right\| \leq \|W'(0)\| \cdot \|w_\perp\| \leq \|O(\epsilon)\|.
\]
To estimate \( g\left(V_{\epsilon,1}(\epsilon), U_w(\epsilon)\right) \), we also write \( J = \sum_{i=1}^{n-1} f_i E_i \) as in the proof of the special case. Then there exists a uniform interval on which
\[
f_1(t) = \|w_{\text{tang}}\| \cdot t + O(t^3) \quad \text{and} \quad f_i(t) = O(t^3) \quad \text{for} \quad i \geq 2.
\]
Now write \( U_w = \sum_{i=1}^{n-1} h_i E_i \), where each \( h_i \) is a smooth function. Since \( U_w(0) \perp J'(0) \), we have \( h_1(0) = 0 \). So, \( h_1(t) = O(t) \) and \( h_i(t) = h_i(0) + O(t) \) for \( i \geq 2 \) on a uniform interval. Then
\[
\left\| g\left(V'_{\epsilon,1}(\epsilon), U_w(\epsilon)\right) \right\| = \frac{1}{\|J(\epsilon)\|} \sum_{i=1}^{n-1} f_i'(\epsilon) h_i(\epsilon)
\]
\[
= \frac{1}{\|J(\epsilon)\|} \left[ \left( \|w_{\text{tang}}\| + O(\epsilon^2) \right) O(\epsilon) + \sum_{i=2}^{n-1} O(\epsilon^3) \left( h_i(0) + O(\epsilon) \right) \right]
\]
\[
= \frac{O(\epsilon)}{\|J(\epsilon)\|}
\]
Therefore,
\[
I(U_w, V') = -g\left(J'(\epsilon) \|J(\epsilon)\|, U_w(\epsilon)\right) + \left(W'(\epsilon), U_w(\epsilon)\right) - g\left(W'(0), w_\perp\right)
\]
\[
\leq -\frac{O(\epsilon)}{\|J(\epsilon)\|} + \|O(\epsilon)\| \cdot \|U_w(\epsilon)\| - \|O(\epsilon)\| \quad (2.14)
\]
which is smaller than a constant \( K \) that is independent of \( \epsilon \).

Thus using (2.11), (2.12), and (2.14), equation (2.13) becomes
\[
I(U_w + V_\epsilon, U_w + V_\epsilon) \leq -g\left(U'_w(0), U_w(0)\right) + 2K - \frac{1}{\|J(\epsilon)\|^2} g\left(J'(\epsilon), J(\epsilon)\right) + \|O(\epsilon)\|.
\]

So for a given $H$, choose $\epsilon$ such that

$$-g\left(U_w'(0),U_w(0)\right) + 2K \frac{1}{\|J(\epsilon)\|^2} g\left(J'(\epsilon),J(\epsilon)\right) + \|O(\epsilon)\| \leq H.$$ 

\[ \square \]

The following lemma provides an explicit formula for $dist_p$ when $p$ and $x_0$ are not conjugate along any $\gamma_v$ for $v \in \mathbb{H}^p_{x_0}$. In this case, $\ker(dexp_{x_0})_v$ is zero for all $v \in \mathbb{H}^p_{x_0}$.

Although the lemma is not used in our current work, it may be useful in the future.

**Lemma 35.** Suppose $p$ and $x_0$ are not conjugate along any $\gamma_v$ for $v \in \mathbb{H}^p_{x_0}$. Given $w \in S_{x_0}$ we set

$$w(\mathbb{H}^p_{x_0}) := \left\{ v \in \mathbb{H}^p_{x_0} \mid \varangle(w,v) = \varangle(w,\mathbb{H}^p_{x_0}) \right\}.$$ 

Then there exists an interval $[0,m]$ on which

$$dist_p\left(c_w(t)\right) = \min_{v \in w(\mathbb{H}^p_{x_0})} \left\{ dist_p x_0 - t \cdot \cos \varangle(w,v) + \frac{1}{2} H \cdot t^2 \right\} + o(t^2)$$

where $H = -g\left(J'_w(0),J_w(0)\right)$.

**Proof.** Fix $w \in S_{x_0}$. Let $\{s_i\}$ be a sequence with $s_i \to 0$, and $\{\sigma_{s_i}\}$ be a sequence of segments from $c_w(s_i)$ to $p$. Then $\left\{(\sigma_{s_i})^{-1}(1)\right\}$, the sequence of tangent vectors at $p$, has a subsequence that converges to a vector $u$. Let $\sigma$ be the segment from $x_0$ to $p$ with $(\sigma^{-1})'(1) = u$.

By the Inverse Function Theorem there exists a neighborhood $U$ of $u$ such that $exp_p|_U$ is one-to-one. So there exists a lift $\tilde{c}_w$ of $c_w$ with $\tilde{c}_w(1) = u$ and

$$exp_p\left(\tilde{c}_w(t)\right) = c_w(t)$$

for all $t \in U$. Using $c_w(t)$, we can produce a variation of $\sigma$ by geodesics, called $\alpha$, so that the variation field $J_w$ is a Jacobi field with $J_w(0) = w$ and $J_w(1) = 0$. By the first
and second variation formulas,

\[
\left. \frac{d\text{Len}(\alpha)}{dt} \right|_{t=0} = -\cos \angle(w, \sigma'(0)) \quad \text{and} \quad \left. \frac{d^2\text{Len}(\alpha)}{dt^2} \right|_{t=0} = -g(J'_w(0), J_w(0)).
\]

So,

\[
dist_p(c_w(t)) = \text{Len}(\sigma_s) = dist_p x_0 - t \cdot \cos \angle(w, \sigma'(0)) - t^2 \cdot g(J'_w(0), J_w(0)) + o(t^2).
\]

On the other hand, \(dist_p\) is directionally differentiable, and each \(\sigma_s\) is a segment from \(x_0\) to \(p\). So, \(\sigma'(0) \in \mathcal{P}^{p}_{x_0}\), which means

\[
\angle(w, \sigma'(0)) \geq \angle(w, \mathcal{P}^{p}_{x_0}) = \min_{v \in w(\mathcal{P}^{p}_{x_0})} \angle(w, v).
\]

Thus, for \(H = -g(J'_w(0), J_w(0))\)

\[
dist_p(c_w(t)) \geq \min_{v \in w(\mathcal{P}^{p}_{x_0})} \left\{ dist_p x_0 - t \cdot \cos \angle(w, v) + \frac{1}{2} H \cdot t^2 \right\} + o(t^2). \tag{2.15}
\]

By Lemma 34, there exists an interval on which

\[
dist_p(c_w(t)) \leq T_{2, H}^{x_0, w}(t) + o(t^2)
\]

\[
= \min_{v \in w(\mathcal{P}^{p}_{x_0})} \left\{ dist_p x_0 - t \cdot \cos \angle(w, v) + \frac{1}{2} H \cdot t^2 \right\} + o(t^2) \tag{2.16}
\]

for \(H = -g(J'_w(0), J_w(0))\). Combining inequalities (2.15) and (2.16), we have

\[
dist_p(c_w(t)) = \min_{v \in w(\mathcal{P}^{p}_{x_0})} \left\{ dist_p x_0 - t \cdot \cos \angle(w, v) + \frac{1}{2} H \cdot t^2 \right\} + o(t^2)
\]

on an interval \([0, m]\) with \(H = -g(J'_w(0), J_w(0))\). \qed
2.6 Proof of the Relative $\pi_1$ Theorem

We now restate Theorem 27 and present its proof.

**Theorem.** Let $x_0$ be an isolated critical point for $\text{dist}_p$ with $\text{dist}_p(x_0) = c_0$. Suppose

$$\pi_1(M^{c_0+\delta}, M^{c_0-\epsilon}) \neq 0,$$

for $\epsilon < c_0 - c - 1$ and $\delta < c_1 - c_0$. Then $\hat{\gamma}^p_{x_0}$ is a pair of antipodal points, i.e. there are only two segments from $p$ to $x_0$ and they make angle $\pi$ at $x_0$. Moreover, the ends of these segments are not conjugate along the segments.

**Proof.** Since $\pi_1(M^{c_0+\delta}, M^{c_0-\epsilon}) \neq 0$, Theorem 26 implies that $\lambda = 1$. If $n = 1$, the manifold $M$ must be a circle since this is the only one-dimensional Riemannian manifold with critical points. Thus, $x_0$ and $p$ must be antipodal points, and the result holds.

For $n > 1$, the definition of sub-index implies that $A(\hat{\gamma}^p_{x_0})$ is not empty but its boundary is. This means $A(\hat{\gamma}^p_{x_0})$ must be a sub-sphere of $S_{x_0}$. Also,

$$\lambda = 1 = n - 1 - \dim A(\hat{\gamma}^p_{x_0})$$

which means $\dim A(\hat{\gamma}^p_{x_0}) = n - 2$. So, $\hat{\gamma}^p_{x_0}$ is the set of vectors in an $(n-1)$-dimensional space that make an angle greater than or equal to $\frac{\pi}{2}$ with an $(n-2)$-dimensional sub-sphere. Thus, $\hat{\gamma}^p_{x_0}$ consists of two antipodal points, say $v$ and $-v$.

It remains to show that $x_0$ and $p$ are not conjugate along the segments $\gamma_v$ and $\gamma_{-v}$. Assume there is a non-zero Jacobi field along $\gamma_v$ that vanishes at $x_0$ and $p$. This means $\ker(d\exp_{x_0})_v$ is not zero. Let $K := \ker(d\exp_{x_0})_v$ and $K^\perp$ be the orthogonal complement of $K$. Since $K$ is a non-zero subspace, its dimension must be greater than or equal to one. So, $\dim K^\perp < n - 2$.

In order to obtain a contradiction, we show that a 1-cell $E^1$, in $\text{int}(M^{c_1} \setminus M^{c - 1})$ with its boundary in $\text{int}M^{c_0}$, can be moved into $\text{int}M^{c_0}$. Using Lemma 30, we can move
the cell into the union of a small $r$-ball around $x_0$ and a sublevel set in $intM^c$, while keeping the boundary of the cell fixed outside of the ball. Define

$$C_r \mathcal{K}^\perp := \left\{ \exp_{x_0}(t\mathcal{K}^\perp) \mid t \in [0,r] \right\}.$$ 

Note that the sum of the dimension of the cell and the dimension of $C_r \mathcal{K}^\perp$ gives us

$$\dim E^1 + \dim C_r \mathcal{K}^\perp = 1 + (\dim \mathcal{K}^\perp + 1)$$
$$= \dim \mathcal{K}^\perp + 2$$
$$< (n - 2) + 2$$
$$= n.$$ 

By transversality, we can apply a small homotopy so that $E^1 \cap C_r \mathcal{K}^\perp = \emptyset$ inside the $r$-ball. Using Lemma 34, there exists an interval on which $\text{dist}_p(c_w(t))$ decreases as long as $w$ is in a small neighborhood of $C_r \mathcal{K}^\perp$. Thus, a small homotopy can be used to move $E^1$ into $intM^c$ while keeping the boundary fixed. This contradicts the fact that $\pi_1(M^c, M^c) \neq 0$. 

\[ \square \]

### 2.7 The Generalized Butterfly Lemma

The next two lemmas are called butterfly lemmas because of the technique used to prove them. Essentially, the proof entails considering the union of a ball and two cones that give the appearance of a butterfly. The lemmas are used to prove the finiteness theorems stated after them.
Lemma 36. (Cheeger [2]) Given \( n \in \mathbb{N}, \ v > 0, \ D, K \in \mathbb{R}, \) and an \( n\)-dimensional manifold \( M \) with

\[
diam M \leq D, \ \vol M \geq v \quad \text{and} \quad \sec M \geq K, \]

there exist \( c_n > 0 \) (depending on \( v, D \) and \( K \)) such that every smooth closed geodesic on \( M \) has length greater than \( c_n \).

Theorem 37. (Cheeger [12]) Given \( n \geq 2, \ v, D, K \in (0, \infty) \), the class of closed Riemannian \( n\)-manifolds with

\[
diam M \leq D, \ \vol M \geq v, \ \text{and} \ |\sec M| \leq K
\]

contains only finitely many diffeomorphism types.

Lemma 38. (Grove, Petersen [12]) Given \( n \geq 2, \ v, D \in (0, \infty), \) and \( K \in \mathbb{R} \), let \( M \) be an \( n\)-manifold with

\[
diam M \leq D, \ \vol M \geq v, \ \text{and} \ \sec M \geq -K^2
\]

Then there exists \( \alpha \in (0, \frac{\pi}{2}) \) and \( \delta > 0 \) (both depending on \( n, v, D, K \)) such that if \( p, q \in M \) satisfy \( \text{dist}_p q \leq \delta \), then either \( p \) is \( \alpha \)-regular for \( q \) or \( q \) is \( \alpha \)-regular for \( p \).

By definition, a point \( x \in M \) is \( \alpha \)-regular for \( \text{dist}_p \), with \( \alpha \in [0, \frac{\pi}{2}] \), if there exists a \( v \in T_x M \) such that

\[
\langle (v, \bar{v}_x^p) \rangle > \pi - \alpha.
\]

A regular point for \( \text{dist}_p \) can be called \( \frac{\pi}{2} \)-regular.

Theorem 39. (Grove, Petersen [8]) Given \( n \geq 2, \ v, D \in (0, \infty), \) and \( K \in \mathbb{R} \), the class of Riemannian \( n\)-manifolds with

\[
diam M \leq D, \ \vol M \geq v, \ \text{and} \ \sec M \geq -K^2
\]

contains only finitely many homotopy types.
Our generalized butterfly lemma is presented below. We consider a sequence \( \{M_i\} \) of \( n \)-manifolds that are said to collapse. This means when using Gromov-Hausdorff convergence, the limit space \( X \) has dimension strictly smaller than \( n \). For the proof, we show that the pre-limit spaces \( M_i \) consist of two overlapping balls and two wings.

**Lemma 40. (Generalized Butterfly Lemma)** Given \( k \geq 0 \) and \( D > 0 \), let \( \{M_i\} \) be a sequence of closed Riemannian \( n \)-manifolds with

\[
\text{diam } M_i \leq D \quad \text{and} \quad \text{sec } M_i \geq k.
\]

Suppose

\[
M_i \xrightarrow{G-H} X
\]

with \( \text{dim } X = n - m \), where \( 1 \leq m \leq n - 1 \). Let \( p_i, q_i \in M_i \) be mutually critical, i.e. \( p_i \) is critical for \( \text{dist}_{q_i} \) and \( q_i \) is critical for \( \text{dist}_{p_i} \). Define

\[
w(p_i, \theta_i) := \exp_{p_i} \left( D \cdot B \left( A(\hat{p}_i), \theta_i \right) \right) \quad \text{and} \quad w(q_i, \theta_i) := \exp_{q_i} \left( D \cdot B \left( A(\hat{q}_i), \theta_i \right) \right).
\]

If

\[
dist_{p_i q_i} \rightarrow 0,
\]

then either \( \text{dim } w(p_i, 0) \) or \( \text{dim } w(q_i, 0) \) is greater than or equal to \( n - m \).

Note that \( \text{dim } w(p_i, 0) = \text{dim } A(\hat{p}_i) + 1 \).

**Proof.** First we claim there exist sequences \( \{r_i\} \) and \( \{\theta_i\} \) converging to zero such that

\[
M_i = B(p_i, r_i) \cup B(q_i, r_i) \cup w(p_i, \theta_i) \cup w(q_i, \theta_i).
\]

(2.17)

Without loss of generality say \( \text{dist}_{p_i q_i} = \frac{1}{i} \). Set

\[
r_i := \left( \frac{1}{i} \right)^{\frac{1}{4}}.
\]
Then for each \( m_i \in M_i \) not contained in \( B(p_i, r_i) \cup B(q_i, r_i) \), consider the triangle formed by \( p_i, q_i, \) and \( m_i \). Define \( a_i := \text{dist}_q m_i, \ b_i := \text{dist}_p m_i, \ c_i := \text{dist}_p q_i = \frac{1}{i}, \) and \( \alpha_i \) to be the interior angle at \( p_i \). Note that \( a_i, b_i \geq r_i = \left( \frac{1}{i} \right)^{\frac{1}{4}}. \)

Without loss of generality, suppose \( a_i \geq b_i \). To prove the claim we need to produce a sequence \( \{ \theta_i \} \) converging to zero such that \( \alpha_i \geq \frac{\pi}{2} - \theta_i \), since this means \( m_i \) is in \( w(p_i, \theta_i) \). Since \( p_i \) is critical for \( \text{dist}_q \), we can choose a segment from \( p_i \) to \( q_i \) such that \( \alpha_i \leq \frac{\pi}{2} \). By the triangle version of Toponogov’s Theorem, \( \alpha_i \geq \pi_k \) where \( \pi_k \) is the angle at \( p_i \) in the space form \( S^n_k \). By the Law of Cosines with \( k = 0 \), we have

\[
 a_i^2 = b_i^2 + \frac{1}{i^2} - \frac{2}{i} b_i \cos \pi_i, \quad (2.18)
\]

Since \( a_i \geq b_i \geq \left( \frac{1}{i} \right)^{\frac{1}{4}} \), we know

\[
0 \geq b_i^2 - a_i^2 \geq \frac{2}{i} b_i \cos \pi_i - \left( \frac{1}{i} \right)^{\frac{3}{4}} \quad \text{by equation (2.18)}
\]

\[
\geq \frac{2}{i} \left( \frac{1}{i} \right)^{\frac{1}{4}} \cos \pi_i - \left( \frac{1}{i} \right)^{\frac{3}{4}}
\]

\[
= \frac{2}{i^{\frac{7}{4}}} \cos \pi_i - \frac{1}{i^{\frac{3}{2}}}
\]

Since \( \frac{\pi}{2} \geq \alpha_i \geq \pi_k \),

\[
0 \leq \cos \alpha_i \leq \cos \pi_k \leq \frac{1}{2} \left( \frac{1}{i} \right)^{\frac{1}{4}}.
\]

Thus, as \( i \) approaches infinity

\[
\left( \frac{1}{i} \right)^{\frac{1}{4}} \longrightarrow 0
\]

which means \( \cos \alpha_i \) converges to zero. Since \( \alpha_i \leq \frac{\pi}{2} \), there must exist \( \theta_i \longrightarrow 0 \) such that \( \alpha_i \geq \frac{\pi}{2} - \theta_i \). Therefore, \( m_i \) is in \( w(p_i, \theta_i) \) and the claim has been established.

Assume \( \dim w(p_i, 0) \) and \( \dim w(q_i, 0) \) are both less than or equal to \( n - m - 1 \). Let

\[
j := \max \left\{ \dim w(p_i, 0), \dim w(q_i, 0) \right\}.
\]
Then $j < n - m$. Let $\beta_X(\epsilon)$ be the maximal number of $\epsilon$-separated points in $X$. Based on Theorem 5.4 in [1], there exists a constant $c_0$ and $0 < \epsilon_0 < 1$ so that for $\epsilon \in (0, \epsilon_0]$ we have

$$\beta_X(\epsilon) \geq c_0 \epsilon^{-(n-m)}.$$  

For any fixed $\epsilon < \epsilon_0$, we can choose a natural number $N$ so that for all $i \geq N$, we have $r_i, \theta_i < \frac{\epsilon}{2}$. Then for $i \geq N$ the maximal number of $\epsilon$-separated points in the set $B(p_i, r_i) \cup B(q_i, r_i)$ is one. By Corollary 8.4 in [1], there exists a constant $c_1$ and $0 < \epsilon_1 < 1$ so that for $\epsilon \in (0, \epsilon_1]$ the number of $\epsilon$-separated points in $w(p_i, \theta_i) \cup w(q_i, \theta_i)$ is less than or equal to $c_1 \epsilon^{-j}$. So, using (2.17), the maximal number of $\epsilon$-separated points in $M_i$ is

$$\beta_{M_i}(\epsilon) \leq 1 + c_1 \epsilon^{-j}.$$  

Now choose $\epsilon_2 \leq \min\{\epsilon_0, \epsilon_1\}$ so that

$$1 + c_1 \left(\frac{\epsilon_2}{2}\right)^{-j} < c_0 \epsilon_2^{-(n-m)}.$$  

(2.19)

By Gromov-Hausdorff convergence, we know

$$\beta_{M_i}\left(\frac{\epsilon_2}{2}\right) \geq \beta_X(\epsilon_2)$$

for all $i$ sufficiently large. Thus,

$$1 + c_1 \left(\frac{\epsilon_2}{2}\right)^{-j} \geq \beta_{M_i}\left(\frac{\epsilon_2}{2}\right) \geq \beta_X(\epsilon_2) \geq c_0 \epsilon_2^{-(n-m)}$$

which contradicts (2.19).  

□
Bibliography


[3] Jeff Cheeger and Detlef Gromoll. On the structure of complete manifolds of non-

second Portuguese edition by Francis Flaherty.

[5] Fuquan Fang and Xiaochun Rong. Curvature, diameter, homotopy groups, and


Based on the 1981 French original [ MR0682063 (85e:53051)], With appendices by
M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael
Bates.


[12] Peter Petersen. Riemannian geometry, volume 171 of Graduate Texts in Mathe-

[13] Conrad Plaut. Metric spaces of curvature \( \geq k \). In Handbook of geometric topology,

44