Title
Sub-Index for Critical Points of Distance Functions

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'a 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.

ABSTRACT OF THE DISSERTATION

Sub-Index for Critical Points of Distance Functions

by

Barbara Christine Herzog

Doctor of Philosophy, Graduate Program in Mathematics
University of California, Riverside, June 2012
Dr. Fred Wilhelm, Chairperson

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

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 $h : M \longrightarrow \mathbb{R}$ if the differential

$$h_* : T_pM \longrightarrow T_{h(p)}\mathbb{R}$$

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* <u>*index*</u> *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 $h : M \longrightarrow \mathbb{R}$ be a smooth function. Define a sublevel set of M as follows.

Definition 2. *For a real number a, the* <u>*sublevel set*</u> *is given by $M^a := h^{-1}(-\infty, a]$.*

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.*

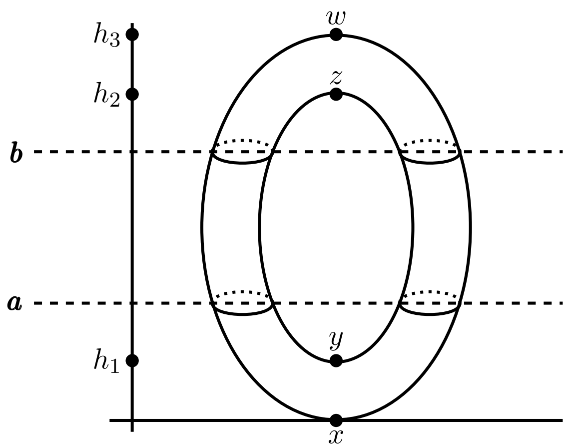

Figure 1.1: An interval that does not contain a critical value.

Let $h : M \longrightarrow \mathbb{R}$ 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* <u>*non-degenerate*</u> *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 λ. 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 λ-cell attached.*

By definition, a λ-*cell*, where λ is a whole number, is a space homeomorphic to a closed λ-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.*

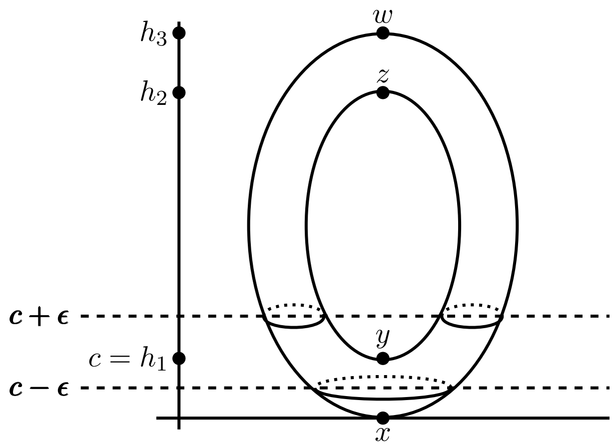

Figure 1.2: An interval that contains no critical values other than c.

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 $M^{c+\epsilon}$.*

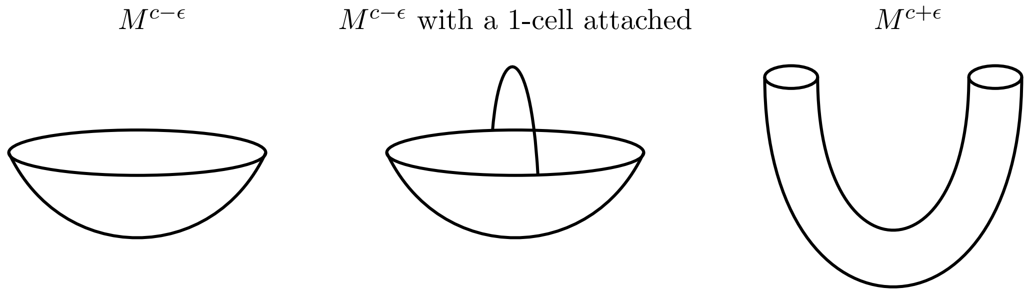

Figure 1.3: An illustration of Morse's Theorem

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 $1, 2, \ldots,$ 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

$$g_p : T_pM \times T_pM \longrightarrow \mathbb{R}$$

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

$$p \longmapsto g_p\Big(X(p), Y(p)\Big)$$

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

$$\|v\| = (v \cdot v)^{\frac{1}{2}} \qquad \text{and} \qquad v \cdot w = \|v\|\|w\| \cos\theta$$

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

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

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* <u>*length of a curve*</u> $\gamma : [0, 1] \longrightarrow M$ *is given by*

$$Len(\gamma) = \int_0^1 g\Big(\gamma'(t), \gamma'(t)\Big)^{\frac{1}{2}} dt.$$

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

Definition 11. *A* <u>*geodesic*</u> *is a curve* $\gamma_v : [0, 1] \longrightarrow M$*, 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 γ_v with $\gamma_v(0) = p$, the* <u>*exponential map*</u>

$$exp_p : T_pM \longrightarrow M$$

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

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

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

Definition 13. *The* <u>*injectivity radius*</u>*, given by $injM$, is the largest radius r such that*

$$exp_p : B(0_p, r) \longrightarrow B(p, r)$$

is a diffeomorphism for all $p \in M$.

Definition 14. *For a fixed point, p, the* <u>*Riemannian distance function*</u>*, $dist_p : M \to \mathbb{R}$, is defined as*

$$dist_p x := \inf \Big\{ Len(\gamma) \;\Big|\; \gamma \text{ is a unit speed curve from } x \text{ to } p \Big\}.$$

Definition 15. *A* <u>*segment*</u> *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.

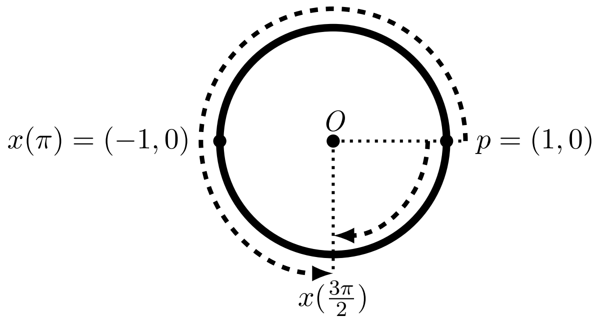

Figure 1.4: The unit circle with paramterization $x(t) = (\cos t, \sin t)$.

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 θ is the non-negative angle formed by $\overline{\mathrm{O}p}$ and $\overline{\mathrm{O}x}$. Since the radius of this circle is one, the length from p to $x(t)$ is equal to the angle θ. 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 π regardless of traveling on a geodesic counterclockwise or clockwise from p. The distance is also π, 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

$$dist_p x(t) = \pi - |t - \pi|$$

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

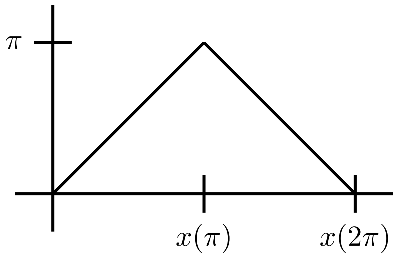

Figure 1.5: The graph of $dist_p x(t) = \pi - |t - \pi|$

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 $\gamma_v : [0, \infty) \longrightarrow M$, define t_v to be the largest parameter time in $[0, \infty)$ such that $\gamma_v : [0, t_v] \longrightarrow M$ 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*

$$Cut(p) := \Big\{ \gamma_v(t_v) \;\Big|\; v \in T_pM \Big\}.$$

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 $dist_p$ is not smooth on the cut locus, it is smooth on the set of points before the cut locus. This is given by

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

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*

$$\Uparrow_x^p := \left\{w \in S_x \;\middle|\; w \textit{ is tangent at } x \textit{ to a segment from } x \textit{ to } p\right\}.$$

By [13], the directional derivative of $dist_p$ in the direction of $v \in T_xM$ is given by

$$D_v(dist_p x) = -\cos \sphericalangle(v, \Uparrow_x^p).$$

Using this directional derivative, Grove and Shiohama created the definition of regular point and critical point for $dist_p$ as follows [9].

Definition 19. *A point x in M is a regular point for $dist_p$ if there exists a v in T_xM such that $\sphericalangle(v, \Uparrow_x^p) > \frac{\pi}{2}$.*

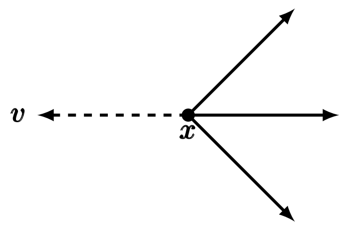

Figure 1.6: Sample configuration of $\Uparrow_x^p$ when x is a regular point.

Since the vectors in $\Uparrow_x^p$ form an angle greater than $\frac{\pi}{2}$ with v, the vectors in $\Uparrow_x^p$ 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* <u>*critical point*</u> *for $dist_p$ if for all v in T_xM we have $\sphericalangle(v, \Uparrow_x^p) \leq \frac{\pi}{2}$.*

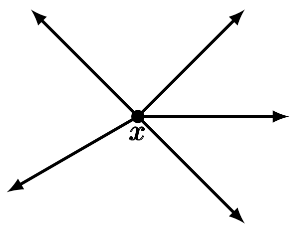

Figure 1.7: Sample configuration of $\Uparrow_x^p$ when x is a critical point.

For a critical point, all vectors in the tangent space form an angle less than or equal to $\frac{\pi}{2}$ with $\Uparrow_x^p$, as illustrated in Figure 1.7. In this case, the vectors in $\Uparrow_x^p$ 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_xM.

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

$$\widehat{\Uparrow_y^p} := \{\Uparrow_y^p + \nabla f\}.$$

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

$$\sphericalangle\left(v,\ \widehat{\Uparrow_y^p}\right) \leq \frac{\pi}{2}.$$

On the other hand, a point y will be defined to be regular for $dist_p + f$ if there exists a v in $T_y M$ such that

$$\sphericalangle\left(v,\ \widehat{\Uparrow_y^p}\right) > \frac{\pi}{2}.$$

Using these definitions, the following lemma shows that $dist_p$ 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 $dist_p x = c$. 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*

$$Dist_p : M \longrightarrow \mathbb{R}$$

with the following properties.

1) The set of critical points for $Dist_p$ is the same as the set of critical points for $dist_p$.

2) The point x is the only critical point of $Dist_p$ with critical value $Dist_p x$.

3) The difference function given by $f \equiv Dist_p -\ dist_p$ 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 ϵ.

Proof. By Urysohn's Lemma, there is a function $\chi : M \longrightarrow \mathbb{R}$ 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 $Dist_p$ is obtained by setting

$$Dist_p := dist_p + \delta \cdot \chi$$

for $\delta > 0$ sufficiently small. It remains to show that $Dist_p$ and $dist_p$ have the same critical and regular points throughout M.

On the set $\overline{N}_1$, we have $f \equiv \delta$, which means that $\nabla f \equiv 0$ and $\widehat{\Uparrow_y^p} = \Uparrow_y^p$. Similarly, $f \equiv 0$ on the set $M \setminus N_2$. So on $\overline{N}_1 \bigcup (M \setminus N_2)$, the critical and regular points for $Dist_p$ and $dist_p$ will be the same.

Now consider the set $N_2 \setminus \overline{N}_1$. Since x is the only critical point in N_2, each point y in $N_2 \setminus \overline{N}_1$ is a regular point for $dist_p$. This means there exists a v in $T_y M$ such that $\sphericalangle(v, \Uparrow_y^p) > \frac{\pi}{2}$. Choose $\delta_0 > 0$ small enough so that there exists $\widehat{v} \in T_y M$ such that

$$\sphericalangle\left(\widehat{v}, \widehat{\Uparrow_x^p}\right) = \sphericalangle\left(\widehat{v}, \ \Uparrow_y^p + \delta_0 \cdot \nabla \chi\right) > \frac{\pi}{2}.$$

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 $N_2 \setminus \overline{N}_1$ will be regular for $Dist_p$. Thus, $Dist_p$ and $dist_p$ will have the same critical and regular points throughout M.

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

$$Dist_p x = dist_p x + \delta.$$

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

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

$$M^{c_i} := dist_p^{-1}[0, c_i].$$

2.2 The Definition of Sub-Index and Main Results

The set $\Uparrow_{x_0}^{p}$ plays an important role in defining x_0 as a critical point. Specifically, x_0 is a critical point for $dist_p$ if for all v in $T_{x_0}M$, we have $\sphericalangle(v, \Uparrow_{x_0}^{p}) \leq \frac{\pi}{2}$. Unfortunately, the set $\Uparrow_{x_0}^{p}$ is too unwieldy to be of further use on its own. Instead, we consider the set given by

$$A(\Uparrow_{x_0}^{p}) := \left\{ v \in S_{x_0} \;\middle|\; \sphericalangle(v, \Uparrow_{x_0}^{p}) \geq \frac{\pi}{2} \right\}$$

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

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

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

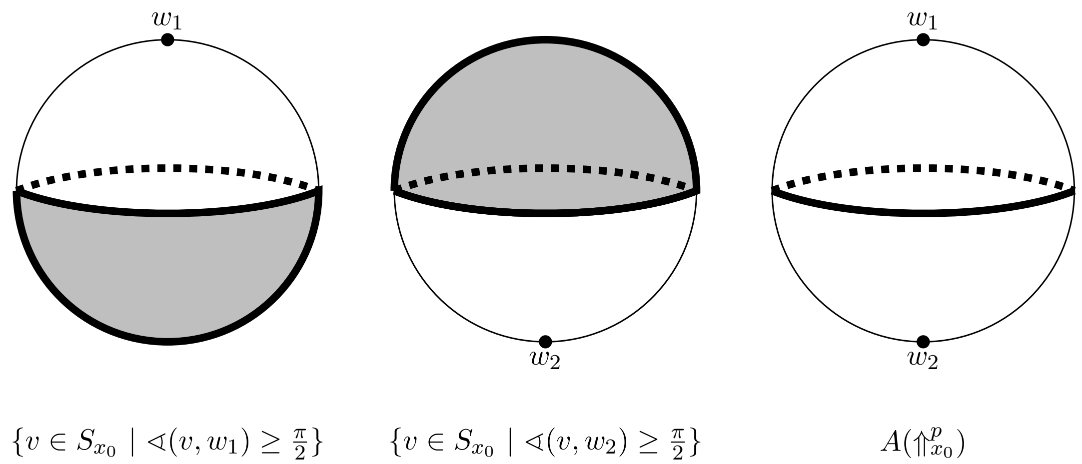

Figure 2.1: $A(\Uparrow_{x_0}^{p})$ is the intersection of hemispheres

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. $A(\Uparrow^p_{x_0})$.

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

Definition 24. *If x_0 is an isolated critical point of $dist_p$, its* <u>*sub-index*</u> *is given by*

$$\lambda := \begin{cases} n & \text{if } A(\Uparrow^p_{x_0}) = \emptyset \\ n - 1 - \dim A(\Uparrow^p_{x_0}) & \text{if } A(\Uparrow^p_{x_0}) \neq \emptyset \text{ but } \partial A(\Uparrow^p_{x_0}) = \emptyset \\ n & \text{if } \partial A(\Uparrow^p_{x_0}) \neq \emptyset \end{cases}$$

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

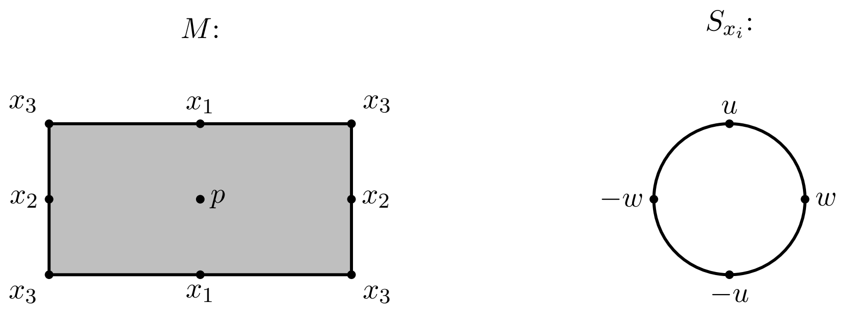

Figure 2.2: The flat 2-torus and its unit tangent space at x_i for $i = 1, 2, 3$

With p as the center, we show that the points p, x_1, x_2, and x_3 are critical for $dist_p$. 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, $\Uparrow_p^p = \emptyset$. This implies that for all vectors $v \in T_pM$ we have $\sphericalangle(v, \Uparrow_p^p) \leq \frac{\pi}{2}$, i.e. p itself is a critical point for $dist_p$. In this case, $A(\Uparrow_p^p) = S_p$, which is a unit circle, and $\partial A(\Uparrow_p^p) = \emptyset$. Thus, p is a minimum since its sub-index is

$$\lambda = n - 1 - \dim A(\Uparrow_p^p) = 2 - 1 - 1 = 0.$$

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

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

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

The point x_3 is critical for $dist_p$ since $\Uparrow_{x_3}^p = \{\pm w, \pm u\}$ means that for all vectors $v \in T_{x_3}M$ we have $\sphericalangle(v, \Uparrow_{x_3}^p) \leq \frac{\pi}{2}$. Then $A(\Uparrow_{x_3}^p) = \emptyset$ since there are no vectors in S_{x_3} at least $\frac{\pi}{2}$ away from $\Uparrow_{x_3}^p$. 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 $dist_p$ with $dist_p(x_0) = c_0$ and sub-index λ. Then the inclusion $M^{c_0-\epsilon} \hookrightarrow M^{c_0+\delta}$ 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 π_1 Theorem) Let x_0 be an isolated critical point for $dist_p$ with $dist_p(x_0) = c_0$. 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 $\Uparrow_{x_0}^p$ is a pair of antipodal points, i.e. there are only two segments from p to x_0 and they make angle π 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 $A(\Uparrow_{x_0}^p)$ 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 $A(\Uparrow_{x_0}^p)$ 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, $A(\Uparrow_{x_0}^p)$ is not empty, and we consider a k-dimensional cell E^k, which is a subset of $int(M^{c_1} \setminus M^{c_{-1}})$ with its boundary in $intM^{c_0}$. To prove the theorem, we show that a flow can be created to move E^k, with $k = 1, \ldots, (\lambda - 1)$, into $intM^{c_0}$ while leaving the boundary of E^k fixed. When the boundary of $A(\Uparrow_{x_0}^p)$ is empty, $A(\Uparrow_{x_0}^p)$ is a great subsphere. In this case, the key idea is that transversality allows E^k to be moved away from $A(\Uparrow_{x_0}^p)$.

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

$$A(\Uparrow_{x_0}^p) \subset \overline{B}\left(w_s, \frac{\pi}{2}\right).$$

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 $intM^{c_0}$.

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 $v \in S_{x_0}$*

$$c_0 - t \cdot \cos \sphericalangle(v, \Uparrow_{x_0}^p) - \epsilon \cdot t \leq dist_p\Big(exp_{x_0}(tv)\Big) \leq c_0 - t \cdot \cos \sphericalangle(v, \Uparrow_{x_0}^p) + \epsilon \cdot t$$

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 $dist_p$ is directionally differentiable,

$$D_v(dist_p x_0) = -\cos \sphericalangle(v, \Uparrow_{x_0}^p).$$

So, the Taylor polynomial representation of $dist_p\Big(exp_{x_0}(tv)\Big)$ is given by

$$dist_p\Big(exp_{x_0}(tv)\Big) = c_0 - t \cdot \cos \sphericalangle(v, \Uparrow_{x_0}^p) + o(t).$$

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

$$c_0 - t \cdot \cos \sphericalangle(v, \Uparrow_{x_0}^p) - \epsilon \cdot t \leq dist_p\Big(exp_{x_0}(tv)\Big) \leq c_0 - t \cdot \cos \sphericalangle(v, \Uparrow_{x_0}^p) + \epsilon \cdot t. \quad (2.1)$$

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 $\{W_{v_i}\}_{i=1}^k$. Define ρ to be the minimum of $\{\rho_{v_i}\}_{i=1}^k$. Thus, for all v in S_{x_0} the inequalities (2.1) will hold on the interval $[0, \rho]$. $\square$

The next lemma establishes a set of points, given by N, that lie in $intM^{c_0}$. Only part (1) is used in the remainder of this work.

Lemma 29. *Given δ sufficiently small and U_δ, the δ-neighborhood of $A(\Uparrow_{x_0}^p)$, there exists $R > 0$ such that:*

1) for $N := exp_{x_0}\Big(t(S_{x_0} \setminus U_\delta)\Big)$ with $t \in (0, 2R]$ we have $N \subset intM^{c_0}$ and

2) $\Big(B(x_0, 2r) \setminus B(x_0, r)\Big) \cap M^{c_0 - \frac{9}{10}r} \subset N$ for $r \in (0, R]$.

Proof. Choose $\delta \in (0, \frac{\pi}{2})$ such that $\frac{9}{20} - \cos(\frac{\pi}{2} - \delta) > 0$. Since the distance from $\Uparrow_{x_0}^p$ is a continuous function on the compact set $S_{x_0} \setminus U_\delta$, a maximum angle exists, say α_1. Choose $\epsilon_1 < \cos\alpha_1$. Then by Lemma 28 there exists $\rho_1 > 0$ such that

$$dist_p\Big(exp_{x_0}(tv)\Big) \leq c_0 - t \cdot \cos\sphericalangle(v, \Uparrow_{x_0}^p) + \epsilon_1 \cdot t$$

on the set $N_1 := \Big\{exp_{x_0}(tv) \;\Big|\; v \in S_{x_0} \setminus U_\delta,\ \ t \in [0, \rho_1]\Big\}$. Now we show that $N_1 \subset intM^{c_0}$. Since $\alpha_1 = \max\Big\{\sphericalangle(v, \Uparrow_{x_0}^p)\Big\}$ over all v in $S_{x_0} \setminus U_\delta$ and $\epsilon_1 < \cos\alpha_1$ on N_1, we have

$$\begin{aligned} dist_p\Big(exp_{x_0}(tv)\Big) &\leq c_0 - t \cdot \cos\sphericalangle(v, \Uparrow_{x_0}^p) + \epsilon_1 \cdot t \\ &\leq c_0 - t \cdot \cos\alpha_1 + \epsilon_1 \cdot t \\ &< c_0 \end{aligned}$$

Thus, $N_1 \subset intM^{c_0}$.

Similarly, since the distance from $\Uparrow_{x_0}^p$ is a continuous function on the compact set $\overline{U}_\delta$, a minimum angle exists, say α_2. Choose $\epsilon_2 < \frac{9}{20} - \cos\alpha_2$. Then by Lemma 28 there exists $\rho_2 > 0$ such that on the set $N_2 := \Big\{exp_{x_0}(tv) \;\Big|\; v \in \overline{U}_\delta,\ \ t \in [0, \rho_2]\Big\}$

$$c_0 - t \cdot \cos\sphericalangle(v, \Uparrow_{x_0}^p) - \epsilon_2 \cdot t \leq dist_p\Big(exp_{x_0}(tv)\Big).$$

Using α_2 and ϵ_2, on N_2 we have

$$\begin{aligned} dist_p\Big(exp_{x_0}(tv)\Big) &\geq c_0 - t \cdot \cos\alpha_2 - \epsilon_2 \cdot t \\ &> c_0 - t \cdot \cos\alpha_2 - t\Big(\frac{9}{20} - \cos\alpha_2\Big) \\ &= c_0 - \frac{9}{20}t. \end{aligned}$$

Define $R := \frac{1}{2}\min\{\rho_1, \rho_2\}$ and for $r \leq R$

$$N := \Big\{exp_{x_0}(tv) \;\Big|\; v \in S_{x_0} \setminus U_\delta,\ t \in (0, 2r]\Big\}.$$

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 $\Big(B(x_0, 2r) \setminus B(x_0, r)\Big)$ but are not in N, i.e. the points in

$$N_3 := \Big\{exp_{x_0}(tv) \;\Big|\; v \in \overline{U}_\delta,\ t \in [r, 2r)\Big\}.$$

Using contrapositive, we need to show that N_3 is not in $M^{c_0 - \frac{9}{10}r}$. Since $t < 2r \leq \rho_2$, we have $N_3 \subset N_2$. So on N_3,

$$dist_p\Big(exp_{x_0}(tv)\Big) > c_0 - \frac{9}{20}t > c_0 - \frac{9}{10}r.$$

Thus, the points in N_3 are not in $M^{c_0 - \frac{9}{10}R}$. $\square$

Lemma 30. *(Local Reduction Lemma) Suppose for $R > 0$, $\overline{B}(x_0, 2R)$ is contained in $int(M^{c_1} \setminus M^{c_{-1}})$. Then $M^{c_0 - \frac{9}{10}R} \cup \overline{B}(x_0, R)$ is a strong deformation retract of $int(M^{c_1})$.*

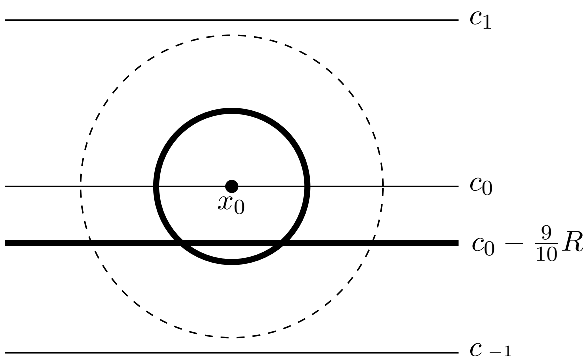

Figure 2.3: The sets involved in Lemma 30

Proof. Since $intM^{c_1} \setminus (M^{c\,-1} \cup \{x_0\})$ 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 $intM^{c_1} \setminus (M^{c\,-1} \cup \{x_0\})$ to end up in $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$. Define the function

$$\tau : intM^{c_1} \setminus (M^{c\,-1} \cup \{x_0\}) \longrightarrow \mathbb{R}$$

to be the minimum amount of time that it takes y to arrive in $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$ as it flows with ψ. Since each y in $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$ is already in the desired set, we have $\tau(y) = 0$.

Now we create a strong deformation retraction of the set $intM^{c_1}$ into the set $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$. Define $\phi : intM^{c_1} \times [0,1] \longrightarrow intM^{c_1}$ by

$$\phi(y,t) := \begin{cases} \psi\Big(y,\ \tau(y)\cdot t\Big) & \text{if } y \in intM^{c_1} \setminus (M^{c\,-1} \cup \{x_0\}) \\ y & \text{if } y \in M^{c\,-1} \cup \{x_0\} \end{cases}$$

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

$$y \in intM^{c_1} \setminus (M^{c\,-1} \cup \ \{x_0\}),$$

we have $\phi(y,0) = \psi(y,0) = y$. For $t = 1$, the points in $M^{c\,-1} \cup \{x_0\}$ remain fixed in $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$, and for the points y in $intM^{c_1} \setminus (M^{c\,-1} \cup \ \{x_0\})$, we have

$$\phi(y,1) = \psi(y,\tau(y)) \in M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R).$$

Further, for all y in $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$, we either have $\phi(y,t) = y$ when y is in $M^{c\,-1} \cup \{x_0\}$ or for $y \in intM^{c_1} \setminus (M^{c\,-1} \cup \{x_0\})$ we know $\tau(y) = 0$ and

$$\phi(y,t) = \psi\Big(y,\tau(y)\cdot t\Big) = \psi(y,0) = y.$$

Therefore, ϕ is a strong deformation retraction of $intM^{c_1}$ into $M^{c_0-\frac{9}{10}R} \cup \overline{B}(x_0,R)$. $\square$

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 $dist_p$ with $dist_p(x_0) = c_0$ and sub-index λ. Then the inclusion $M^{c_0-\epsilon} \hookrightarrow M^{c_0+\delta}$ 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 $i = 0, 1, \dots, (\lambda - 1)$.

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

$$\sphericalangle(v, \Uparrow_{x_0}^p) \geq \frac{\pi}{2}.$$

Since x_0 is a critical point, we know that $\sphericalangle(v, \Uparrow_{x_0}^p) \leq \frac{\pi}{2}$ for all tangent vectors v. So for all $v \in S_{x_0}$, we must have $\sphericalangle(v, \Uparrow_{x_0}^p) < \frac{\pi}{2}$. From this, the directional derivative tells us that $D_v(dist_p x_0) < 0$ 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 $1, 2, \dots,$ or $(n-1)$ within $int(M^{c_1} \setminus M^{c_{-1}})$ with boundary in $intM^{c_0}$ can be deformed into $intM^{c_0}$. Therefore, $\pi_i(M^{c_1}, M^{c_0}) = 0$ for $i = 0, 1, \dots, (n-1)$.

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_δ be the δ-neighborhood of $A(\Uparrow_{x_0}^p)$. Then by Lemma 29 there exists $R > 0$ such that $N \subset intM^{c_0}$ where

$$N := exp_{x_0}\Big(t(S_{x_0} \setminus U_\delta)\Big) \quad \text{for} \quad t \in (0, 2R].$$

Let E^k be a k-cell with $k = 0, 1, \ldots, (\lambda - 1)$ such that $E^k \subset int(M^{c_1} \setminus M^{c_{-1}})$ and $\partial E^k \subset intM^{c_0}$. Since $c_0 - dist_p$ is a continuous function on the compact set ∂E^k, there exists a minimum, say m. Then

$$dist_p(\partial E^k) \leq c_0 - m.$$

Choose $r < \min\left\{2R, \frac{1}{2}m\right\}$ so that $B(x_0, 2r)$ is contained in $int(M^{c_1} \setminus M^{c_{-1}})$ and $2r$ is smaller than the injectivity radius at x_0. Then $\overline{B}(x_0, r)$ is a ball around x_0 in $int(M^{c_1} \setminus M^{c_{-1}})$, and by the Local Reduction Lemma $M^{c_0 - \frac{9}{20}r} \cup \overline{B}\Big(x_0, \frac{1}{2}r\Big)$ is a strong deformation retract of $intM^{c_1}$. Since $r < \frac{1}{2}m$ we have

$$dist_p B(x_0, 2r) > c_0 - 2r > c_0 - m.$$

This means ∂E^k is outside the $2r$-ball. So the strong deformation retract moves E^k into $M^{c_0 - \frac{9}{20}r} \cup \overline{B}\Big(x_0, \frac{1}{2}r\Big)$ while keeping ∂E^k fixed. Since $r < 2R$, we know from Lemma 29 that

$$N_r := \left\{ exp_{x_0}\Big(t(S_{x_0} \setminus U_\delta)\Big) \;\Big|\; t \in (0, r] \right\} \subset intM^{c_0}.$$

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

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

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

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

$$\begin{aligned}
\dim E^k + \dim C_r A(\Uparrow_{x_0}^p) &\leq \Big(\lambda - 1\Big) + \Big(\dim A(\Uparrow_{x_0}^p) + 1\Big) \\
&= \lambda + \dim A(\Uparrow_{x_0}^p) \\
&= \Big(n - 1 - \dim A(\Uparrow_{x_0}^p)\Big) + \dim A(\Uparrow_{x_0}^p) \\
&= n - 1 \\
&< n.
\end{aligned}$$

This means by transversality we can apply a small homotopy so that

$$\left\{E^k \cap \overline{B}\Big(x_0, \frac{1}{2}r\Big)\right\} \bigcap C_r A(\Uparrow_{x_0}^p) = \emptyset$$

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

$$\left\{E^k \cap \overline{B}\Big(x_0, \frac{1}{2}r\Big)\right\} \subset B(x_0, r) \setminus C_r A(\Uparrow_{x_0}^p),$$

the following lemma shows that $E^k \cap \overline{B}\Big(x_0, \frac{1}{2}r\Big)$ can be moved into $\overline{N_r \Big\backslash B\Big(x_0, \frac{1}{2}r\Big)}$, a subset of $intM^{c_0}$, while keeping the boundary of the cell fixed. This will complete the proof.

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

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

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

Since $C_r A\left(\Uparrow_{x_0}^p\right)$ is a union of these radial geodesics, this restricts to an isotopy of $B(x_0, r) \setminus C_r A(\Uparrow_{x_0}^p)$ to

$$B(x_0, r) \Big\backslash \left\{B\Big(x_0, \frac{1}{2}r\Big) \bigcup C_r A(\Uparrow_{x_0}^p)\right\}.$$

Now we move the points into $\overline{N_r \Big\backslash B\Big(x_0, \frac{1}{2}r\Big)}$. For any $\delta \in (0, \frac{\pi}{2})$ the set $S_{x_0} \setminus U_\delta$ is a strong deformation retract of $S_{x_0} \setminus A(\Uparrow_{x_0}^p)$. Exponentiating this retract gives an isotopy of $B(x_0, r) \setminus C_r A(\Uparrow_{x_0}^p)$ that leaves the metric spheres around x_0 invariant. Thus, it carries $B(x_0, r) \Big\backslash \left\{B\Big(x_0, \frac{1}{2}r\Big) \bigcup C_r A(\Uparrow_{x_0}^p)\right\}$ into $\overline{N_r \Big\backslash B\Big(x_0, \frac{1}{2}r\Big)}$ without moving points in $\overline{N_r \Big\backslash B\Big(x_0, \frac{1}{2}r\Big)}$. 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 $M \setminus C_r A(\Uparrow_{x_0}^p)$ that fixes $M \setminus B(x_0, r)$. $\square$

Case 3: Suppose that both $A(\Uparrow_{x_0}^p)$ and its boundary are not empty. By definition, $A(\Uparrow_{x_0}^p)$ consists of unit tangent vectors at least $\frac{\pi}{2}$ away from $\Uparrow_{x_0}^p$. So, $A(\Uparrow_{x_0}^p)$ is a subset of a $\frac{\pi}{2}$-ball in the unit tangent sphere. This means we can choose a vector w_s in $A(\Uparrow_{x_0}^p)$ such that

$$A(\Uparrow_{x_0}^p) \subset \overline{B}\Big(w_s, \frac{\pi}{2}\Big). \tag{2.3}$$

Given that $A(\Uparrow_{x_0}^p)$ 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 $B\Big(x_0, \frac{3}{2}r\Big)$ 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 ψ. Then the desired homotopy $\Psi : M \times [0,1] \longrightarrow M$ is given by

$$\Psi(y,t) = \psi\Big(y, \tau(y) \cdot t\Big).$$

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 $M^{c_0 - \frac{9}{20}r} \cup \overline{B}\Big(x_0, \frac{1}{2}r\Big)$, it is enough to show that points in $\overline{B}\Big(x_0, \frac{1}{2}r\Big)$ 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 $y \in \overline{B}\Big(x_0, \frac{1}{2}r\Big)$ out of $B(x_0, 2r)$. By the triangle inequality, we have

$$\Big\| exp_{x_0}^{-1}(y) - rw_s \Big\| \leq \frac{7}{5}r < \frac{3}{2}r. \tag{2.4}$$

For $y \in \overline{B}\Big(x_0, \frac{1}{2}r\Big)$ we now claim that

$$\theta := \sphericalangle\Big(exp_{x_0}^{-1}(y) - rw_s,\ w_s\Big) > \frac{\pi}{2} + \delta.$$

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

$$\Psi(y, r) \subset \Big(N_r \cap B(x_0, 2r)\Big).$$

To prove the claim, first note that

$$\begin{aligned}
\cos\theta &= \frac{g\Big(exp_{x_0}^{-1}(y) - rw_s,\ w_s\Big)}{\|exp_{x_0}^{-1}(y) - rw_s\|} \\
&= \frac{g\Big(exp_{x_0}^{-1}(y),\ w_s\Big) - r}{\|exp_{x_0}^{-1}(y) - rw_s\|} \\
&\le \frac{\|exp_{x_0}^{-1}(y)\| - r}{\|exp_{x_0}^{-1}(y) - rw_s\|} \quad \text{by the Cauchy-Schwarz inequality} \\
&\le -\frac{r}{2\|exp_{x_0}^{-1}(y) - rw_s\|} \quad \text{since } y \in \overline{B}\Big(x_0, \frac{1}{2}r\Big) \\
&< -\frac{1}{3}. \qquad (2.5)
\end{aligned}$$

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

$$-\frac{1}{3} < -\cos\Big(\frac{\pi}{2} - \delta\Big) = -\sin\delta = \cos\Big(\frac{\pi}{2} + \delta\Big). \qquad (2.6)$$

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

$$\cos\theta < -\frac{1}{3} < \cos\Big(\frac{\pi}{2} + \delta\Big)$$

which means $\theta > \frac{\pi}{2} + \delta$ as claimed. $\square$

2.5 Lemmas Related to Conjugate Points

For the Relative π_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 $dist_p$ and $v \in \Uparrow_{x_0}^{p}$. Let $\gamma_v : [0,1] \longrightarrow M$ 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 γ_v with $J(0) = 0$ and $J(1) = 0$.*

Note that if p is conjugate to x_0 along γ_v, then $\ker\left(d\exp_{x_0}\right)_v$ 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 $\ker\left(d\exp_{x_0}\right)_v$. Then there is a unique Jacobi field J_w along γ_v so that*

$$J_w(0) = w \quad \textit{and} \quad J_w(1) = 0.$$

Proof. Let $\mathcal{N}$ be the familly of nonzero Jacobi fields N so that

$$N(0) = N(1) = 0.$$

Let $\mathcal{P}$ be the family of Jacobi fields P so that

$$P(1) = 0 \text{ and } P'(1) \perp N'(1) \text{ for all } N \in \mathcal{N}.$$

We have

$$\ker\left(d\exp_{x_0}\right)_v = \left\{N'(0) \,|\, N \in \mathcal{N}\right\}.$$

Since the Riccati operator on $\mathcal{N} \oplus \mathcal{P}$ is self adjoint we know that for all $P \in \mathcal{P}$ and all $N \in \mathcal{N}$

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

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

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

$$T_{2,H}^{x_0,w}(t) := dist_p x_0 - t \cdot \cos\sphericalangle(w, \Uparrow_{x_0}^p) + \frac{1}{2} H \cdot t^2.$$

Lemma 34. *Let H be:*

(1) $-g\Big(J_w'(0), J_w(0)\Big)$ if $w \in S_{x_0}$ is orthogonal to $ker(dexp_{x_0})_v$,

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

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

$$dist_p\Big(c_w(t)\Big) \leq T_{2,H}^{x_0,w}(t) + o(t^2).$$

Proof. Given $v \in \Uparrow_{x_0}^p$ and $w \in S_{x_0}$, it suffices to find a vector field V along γ_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 γ_v, there is a variation $\tilde{\gamma}$ whose variation field is V. By the first and second variation formulas, we know

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

So the Taylor polynomial gives us

$$Len(\tilde{\gamma}) \leq Len(\gamma_v) - t \cdot \cos\sphericalangle(w, v) + \frac{1}{2} t^2 \cdot I(V, V) + o(t^2).$$

Since distance is the minimum of length, there is an interval on which

$$dist_p(c_w(t)) \le dist_p x_0 - t \cdot \cos\sphericalangle(w, \Uparrow_{x_0}^p) + \frac{1}{2}t^2 \cdot I(V,V) + o(t^2).$$

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

Now suppose w is not orthogonal to $ker(dexp_{x_0})_v$. First, we consider the special case when w is in $ker(dexp_{x_0})_v$. By lemma 33, there exists a nonzero Jacobi field J along γ such that $J(0) = J(1) = 0$ and $J'(0) = w$. From this we create a vector field that does not vanish at both ends. Specifically, define a vector field V_ϵ by

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

where W_ϵ 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

$$\begin{aligned} I(V_\epsilon, V_\epsilon) &= g\left(\frac{J'(1)}{\|J(\epsilon)\|}, \frac{J(1)}{\|J(\epsilon)\|}\right) - g\left(\frac{J'(\epsilon)}{\|J(\epsilon)\|}, \frac{J(\epsilon)}{\|J(\epsilon)\|}\right) + g\Big(W_\epsilon'(\epsilon), W_\epsilon(\epsilon)\Big) - g\Big(W_\epsilon'(0), W_\epsilon(0)\Big) \\ &= -\frac{1}{\|J(\epsilon)\|^2} g\Big(J'(\epsilon), J(\epsilon)\Big) + g\left(W_\epsilon'(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\right) - g\Big(W_\epsilon'(0), J'(0)\Big). \end{aligned}$$

Note that the limit of the first term gives us

$$\begin{aligned} \lim_{\epsilon \to 0} \frac{-g\Big(J'(\epsilon), J(\epsilon)\Big)}{\|J(\epsilon)\|^2} &= \lim_{\epsilon \to 0} \frac{-g\Big(J'(\epsilon), J(\epsilon)\Big)}{g\Big(J(\epsilon), J(\epsilon)\Big)} \\ &= -\lim_{\epsilon \to 0} \frac{g\Big(J''(\epsilon), J(\epsilon)\Big) + g\Big(J'(\epsilon), J'(\epsilon)\Big)}{2g\Big(J'(\epsilon), J(\epsilon)\Big)} \\ &= -\infty \end{aligned}$$

since $J'(0) \neq 0$. So for an upper bound on $I(V_\epsilon, V_\epsilon)$, it suffices to bound

$$g\Big(W'_\epsilon(0), J'(0)\Big) \text{ and } g\bigg(W'_\epsilon(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\bigg)$$

independent of ϵ.

Let $\{E_i\}_{i=1}^{n-1}$ be an orthonormal parallel frame for the normal space of γ_v with $E_1(0) = J'(0)$. 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 $E_1(0) = J'(0)$, we know $f'_1(0) = 1$ and $f'_i(0) = 0$ for all $i = 2, ..., n-1$. Since J is a Jacobi field with $J(0) = 0$,

$$J''(0) = \sum_{i=1}^{n-1} f''_i(0)E_i(0) = -R\Big(J(0), \gamma'(0)\Big)\gamma'(0) = 0.$$

So $f''_i(0) = 0$ for all i. Using Taylor's Theorem, there exists an interval on which

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

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

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

Taking the square root, we have

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

Combining gives us

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

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

and for $i \geq 2$

$$\frac{f_i(t)}{t + \mathcal{O}(t^3)} = \frac{\mathcal{O}(t^3)}{t + \mathcal{O}(t^3)} = \mathcal{O}(t^2).$$

In order to approximate W_ϵ', 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 ϵ. Since the space of Jacobi fields with bounded endpoints is compact, there is a bound B on $[0,\epsilon]$, independent of ϵ, so that

$$\|g_{\epsilon,i}''\| = \|R(W_\epsilon, \gamma', \gamma', E_i)\| \leq B. \tag{2.8}$$

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 ϵ, on which

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

and

$$g_{\epsilon,i}(t) = g_{\epsilon,i}'(0)t + \mathcal{O}(t^2) \text{ for } i = 2, \ldots, n-1.$$

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

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

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

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

$$\left\| g\Big(W_\epsilon'(0), J'(0)\Big) \right\| = \left\| g\Big(W_\epsilon'(0), w\Big) \right\| \leq \|\mathcal{O}(\epsilon)\|. \tag{2.9}$$

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

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

This means by the Cauchy-Schwarz inequality

$$\left\| g\Big(W_\epsilon'(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\Big) \right\| \leq \|W_\epsilon'(\epsilon)\| \leq \|\mathcal{O}(\epsilon)\|. \tag{2.10}$$

Thus from (2.9) and (2.10) we have

$$\begin{aligned} I(V_\epsilon, V_\epsilon) &= -\frac{1}{\|J(\epsilon)\|^2} g\Big(J'(\epsilon), J(\epsilon)\Big) + g\Big(W_\epsilon'(\epsilon), \frac{J(\epsilon)}{\|J(\epsilon)\|}\Big) - g\Big(W_\epsilon'(0), J'(0)\Big) \\ &\leq -\frac{1}{\|J(\epsilon)\|^2} g\Big(J'(\epsilon), J(\epsilon)\Big) + \|\mathcal{O}(\epsilon)\| \longrightarrow -\infty \ \text{ as } \epsilon \to 0. \end{aligned}$$

So choose any number for H. Then choose $0 < \epsilon < m$ such that

$$-\frac{1}{\|J(\epsilon)\|^2} g\Big(J'(\epsilon), J(\epsilon)\Big) + \|\mathcal{O}(\epsilon)\| \leq H.$$

Now suppose w is not orthogonal to $ker(dexp_{x_0})_v$ and w is not in $ker(dexp_{x_0})_v$. Write $w = w_{tang} + w_\perp$ with respect to $ker(dexp_{x_0})_v$. Then there exists a Jacobi field U_w along γ_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 $J'(0) = w_{tang}$. Define the vector field V_ϵ as in the proof of the special case previously discussed, and let $V_{\epsilon,1} := \frac{J}{\|J(\epsilon)\|}$. Then

$$I(U_w, U_w) = -g\Big(U_w'(0), U_w(0)\Big) \tag{2.11}$$

and

$$I(V_\epsilon, V_\epsilon) \leq -\frac{1}{\|J(\epsilon)\|^2} g\Big(J'(\epsilon), J(\epsilon)\Big) + \|\mathcal{O}(\epsilon)\| \longrightarrow -\infty \ \text{ as } \epsilon \to 0. \tag{2.12}$$

So $U_w + V_\epsilon$ is a vector field along γ_v with $(U_w + V_\epsilon)(0) = w$ and $(U_w + V_\epsilon)(1) = 0$. Now consider

$$I(U_w + V_\epsilon, U_w + V_\epsilon) = I(U_w, U_w) + 2I(U_w, V_\epsilon) + I(V_\epsilon, V_\epsilon). \tag{2.13}$$

Based on (2.11) and (2.12), it remains to show that we have a bound on

$$\begin{aligned} I(U_w, V_\epsilon) &= g\Big(V_{\epsilon,1}'(1), U_w(1)\Big) - g\Big(V_{\epsilon,1}'(\epsilon), U_w(\epsilon)\Big) + g\Big(W_\epsilon'(\epsilon), U_w(\epsilon)\Big) - g\Big(W_\epsilon'(0), U_w(0)\Big) \\ &= -g\Big(V_{\epsilon,1}'(\epsilon), U_w(\epsilon)\Big) + g\Big(W_\epsilon'(\epsilon), U_w(\epsilon)\Big) - g\Big(W_\epsilon'(0), U_w(0)\Big). \end{aligned}$$

From the proof of the special case when w is in $ker(dexp_{x_0})_v$, we know that $\|W'_\epsilon(\epsilon)\| \leq \|\mathcal{O}(\epsilon)\|$ and $\|W'_\epsilon(0)\| \leq \|\mathcal{O}(\epsilon)\|$. So

$$\left\| g\Big(W'_\epsilon(\epsilon), U_w(\epsilon)\Big) \right\| \leq \|W'_\epsilon(\epsilon)\| \cdot \|U_w(\epsilon)\| \leq \|\mathcal{O}(\epsilon)\| \cdot \|U_w(\epsilon)\|$$

and

$$\left\| g\Big(W'_\epsilon(0), w_\perp\Big) \right\| \leq \|W'_\epsilon(0)\| \cdot \|w_\perp\| \leq \|\mathcal{O}(\epsilon)\|.$$

To estimate $g\Big(V'_{\epsilon,1}(\epsilon), U_w(\epsilon)\Big)$, 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

$$f_1(t) = \|w_{tang}\| \cdot t + \mathcal{O}(t^3) \;\text{ and }\; f_i(t) = \mathcal{O}(t^3) \text{ for } i \geq 2.$$

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, $h_1(t) = \mathcal{O}(t)$ and $h_i(t) = h_i(0) + \mathcal{O}(t)$ for $i \geq 2$ on a uniform interval. Then

$$\begin{aligned}
\left\| g\Big(V'_{\epsilon,1}(\epsilon), U_w(\epsilon)\Big) \right\| &= \frac{1}{\|J(\epsilon)\|} \sum_{i=1}^{n-1} f'_i(\epsilon) h_i(\epsilon) \\
&= \frac{1}{\|J(\epsilon)\|} \Big[\Big(\|w_{tang}\| + \mathcal{O}(\epsilon^2)\Big)\mathcal{O}(\epsilon) + \sum_{i=2}^{n-1} \mathcal{O}(\epsilon^2)\Big(h_i(0) + \mathcal{O}(\epsilon)\Big) \Big] \\
&= \frac{\mathcal{O}(\epsilon)}{\|J(\epsilon)\|}
\end{aligned}$$

Therefore,

$$\begin{aligned}
I(U_w, V_\epsilon) &= -g\Big(\frac{J'(\epsilon)}{\|J(\epsilon)\|}, U_w(\epsilon)\Big) + \Big(W'_\epsilon(\epsilon), U_w(\epsilon)\Big) - g\Big(W'_\epsilon(0), w_\perp\Big) \\
&\leq -\frac{\mathcal{O}(\epsilon)}{\|J(\epsilon)\|} + \|\mathcal{O}(\epsilon)\| \cdot \|U_w(\epsilon)\| - \|\mathcal{O}(\epsilon)\| \qquad (2.14)
\end{aligned}$$

which is smaller than a constant K that is independent of ϵ.

Thus using (2.11), (2.12), and (2.14), equation (2.13) becomes

$$I(U_w + V_\epsilon, U_w + V_\epsilon) \leq -g\Big(U'_w(0), U_w(0)\Big) + 2K - \frac{1}{\|J(\epsilon)\|^2} g\Big(J'(\epsilon), J(\epsilon)\Big) + \|\mathcal{O}(\epsilon)\|.$$

So for a given H, choose ϵ such that

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

$\square$

The following lemma provides an explicit formula for $dist_p$ when p and x_0 are not conjugate along any γ_v for $v \in \Uparrow_{x_0}^p$. In this case, $ker(dexp_{x_0})_v$ is zero for all $v \in \Uparrow_{x_0}^p$. 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 γ_v for $v \in \Uparrow_{x_0}^p$. Given $w \in S_{x_0}$ we set*

$$w(\Uparrow_{x_0}^p) := \Big\{ v \in \Uparrow_{x_0}^p \ \Big| \ \sphericalangle(w, v) = \sphericalangle(w, \Uparrow_{x_0}^p) \Big\}.$$

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

$$dist_p\Big(c_w(t)\Big) = \min_{v \in w(\Uparrow_{x_0}^p)} \Big\{ dist_p x_0 - t \cdot \cos \sphericalangle(w, v) + \frac{1}{2} H \cdot t^2 \Big\} + o(t^2)$$

where $H = -g\Big(J_w'(0), J_w(0)\Big)$.

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 $\Big\{(\sigma_{s_i}^{-1})'(1)\Big\}$, the sequence of tangent vectors at p, has a subsequence that converges to a vector u. Let σ 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

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

for all $t \in U$. Using $c_w(t)$, we can produce a variation of σ by geodesics, called α, 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,

$$\left.\frac{dLen(\alpha)}{dt}\right|_{t=0} = -\cos\sphericalangle\Big(w, \sigma'(0)\Big) \quad \text{and} \quad \left.\frac{d^2Len(\alpha)}{dt^2}\right|_{t=0} = -g\Big(J_w'(0), J_w(0)\Big).$$

So,

$$\begin{aligned} dist_p\Big(c_w(t)\Big) &= Len(\sigma_{s_i}) \\ &= dist_p x_0 - t\cdot\cos\sphericalangle\Big(w, \sigma'(0)\Big) - t^2\cdot g\Big(J_w'(0), J_w(0)\Big) + o(t^2). \end{aligned}$$

On the other hand, $dist_p$ is directionally differentiable, and each σ_{s_i} is a segment from x_0 to p. So, $\sigma'(0) \in \Uparrow_{x_0}^p$ which means

$$\sphericalangle\Big(w, \sigma'(0)\Big) \geq \sphericalangle(w, \Uparrow_{x_0}^p) = \min_{v\in w(\Uparrow_{x_0}^p)} \sphericalangle(w, v).$$

Thus, for $H = -g\Big(J_w'(0), J_w(0)\Big)$

$$dist_p\Big(c_w(t)\Big) \geq \min_{v\in w(\Uparrow_{x_0}^p)} \Big\{dist_p x_0 - t\cdot\cos\sphericalangle(w,v) + \frac{1}{2}H\cdot t^2\Big\} + o(t^2). \tag{2.15}$$

By Lemma 34, there exists an interval on which

$$\begin{aligned} dist_p\Big(c_w(t)\Big) &\leq T_{2,H}^{x_0,w}(t) + o(t^2) \\ &= \min_{v\in w(\Uparrow_{x_0}^p)} \Big\{dist_p x_0 - t\cdot\cos\sphericalangle(w,v) + \frac{1}{2}H\cdot t^2\Big\} + o(t^2) \qquad (2.16) \end{aligned}$$

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

$$dist_p\Big(c_w(t)\Big) = \min_{v\in w(\Uparrow_{x_0}^p)} \Big\{dist_p x_0 - t\cdot\cos\sphericalangle(w,v) + \frac{1}{2}H\cdot t^2\Big\} + o(t^2)$$

on an interval $[0, m]$ with $H = -g\Big(J_w'(0), J_w(0)\Big)$. $\square$

2.6 Proof of the Relative π_1 Theorem

We now restate Theorem 27 and present its proof.

Theorem. *Let x_0 be an isolated critical point for $dist_p$ with $dist_p(x_0) = c_0$. 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 $\Uparrow^p_{x_0}$ is a pair of antipodal points, i.e. there are only two segments from p to x_0 and they make angle π 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 $A(\Uparrow^p_{x_0})$ is not empty but its boundary is. This means $A(\Uparrow^p_{x_0})$ must be a sub-sphere of S_{x_0}. Also,

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

which means $\dim A(\Uparrow^p_{x_0}) = n - 2$. So, $\Uparrow^p_{x_0}$ 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, $\Uparrow^p_{x_0}$ consists of two antipodal points, say v and $-v$.

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

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

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

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

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

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 $dist_p\Big(c_w(t)\Big)$ 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 $intM^{c_0}$ while keeping the boundary fixed. This contradicts the fact that $\pi_1(M^{c_1}, M^{c_0}) \neq 0$. $\square$

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*

$$diamM \leq D, \quad volM \geq v \quad and \quad \sec M \geq K,$$

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*

$$diamM \leq D, \;\; volM \geq v, \quad and \;\; |secM| \leq K$$

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*

$$diamM \leq D, \;\; volM \geq v, \quad and \;\; secM \geq -K^2.$$

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 $dist_p q \leq \delta$, then either p is α-regular for q or q is α-regular for p.

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

$$\sphericalangle(v, \Uparrow_x^p) > \pi - \alpha.$$

A regular point for $dist_p$ 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*

$$diamM \leq D, \;\; volM \geq v, \quad and \;\; secM \geq -K^2$$

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 that 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*

$$diam\ M_i \leq D \quad and \quad \sec M_i \geq k.$$

Suppose

$$M_i \overset{G-H}{\longrightarrow} X$$

with $\dim X = n - m$, 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 $dist_{q_i}$ and q_i is critical for $dist_{p_i}$. Define

$$w(p_i, \theta_i) := exp_{p_i}\Big(D \cdot B\Big(A(\Uparrow_{p_i}^{q_i}), \theta_i\Big)\Big) \quad and \quad w(q_i, \theta_i) := exp_{q_i}\Big(D \cdot B\Big(A(\Uparrow_{q_i}^{p_i}), \theta_i\Big)\Big).$$

If

$$dist_{p_i} q_i \longrightarrow 0,$$

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

Note that $\dim w(p_i, 0) = \dim A(\Uparrow_{x_0}^{p}) + 1$.

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

$$M_i = B(p_i, r_i) \cup B(q_i, r_i) \cup w(p_i, \theta_i) \cup w(q_i, \theta_i). \tag{2.17}$$

Without loss of generality say $dist_{p_i} q_i = \frac{1}{i}$. Set

$$r_i := \Big(\frac{1}{i}\Big)^{\frac{1}{4}}.$$

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 $a_i := dist_{q_i} m_i$, $b_i := dist_{p_i} m_i$, $c_i := dist_{p_i} q_i = \frac{1}{i}$, and α_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 $dist_{q_i}$, 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, $\alpha_i \geq \overline{\alpha}_i$ where $\overline{\alpha}_i$ is the angle at p_i in the space form S^n_k. By the Law of Cosines with $k = 0$, we have

$$a_i^2 = b_i^2 + \frac{1}{i^2} - \frac{2}{i} b_i \cos \overline{\alpha}_i. \tag{2.18}$$

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

$$\begin{aligned} 0 &\geq b_i^2 - a_i^2 \\ &= \frac{2}{i} b_i \cos \overline{\alpha}_i - \frac{1}{i^2} \quad \text{by equation (2.18)} \\ &\geq \frac{2}{i}\left(\frac{1}{i}\right)^{\frac{1}{4}} \cos \overline{\alpha}_i - \frac{1}{i^2} \\ &= \frac{2}{i^{\frac{5}{4}}} \cos \overline{\alpha}_i - \frac{1}{i^2}. \end{aligned}$$

Since $\frac{\pi}{2} \geq \alpha_i \geq \overline{\alpha}_i$,

$$0 \leq \cos \alpha_i \leq \cos \overline{\alpha}_i \leq \frac{1}{2}\left(\frac{1}{i}\right)^{\frac{3}{4}}.$$

Thus, as i approaches infinity

$$\left(\frac{1}{i}\right)^{\frac{3}{4}} \longrightarrow 0$$

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

$$j := \max\Big\{ \dim w(p_i, 0),\ \dim w(q_i, 0) \Big\}.$$

Then $j < n - m$. Let $\beta_X(\epsilon)$ be the maximal number of ϵ-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 ϵ-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 ϵ-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 ϵ-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

$$1 + c_1 \left(\frac{\epsilon_2}{2}\right)^{-j} < c_0 \epsilon_2^{-(n-m)}. \tag{2.19}$$

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). $\square$

Bibliography

[1] Yu. Burago, M. Gromov, and G. Perel′man. A. D. Aleksandrov spaces with curvatures bounded below. *Uspekhi Mat. Nauk*, 47(2(284)):3–51, 222, 1992.

[2] Jeff Cheeger and David G. Ebin. *Comparison theorems in Riemannian geometry.* AMS Chelsea Publishing, Providence, RI, 2008. Revised reprint of the 1975 original.

[3] Jeff Cheeger and Detlef Gromoll. On the structure of complete manifolds of nonnegative curvature. *Ann. of Math. (2)*, 96:413–443, 1972.

[4] Manfredo Perdigão do Carmo. *Riemannian geometry.* Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[5] Fuquan Fang and Xiaochun Rong. Curvature, diameter, homotopy groups, and cohomology rings. *Duke Math. J.*, 107(1):135–158, 2001.

[6] Michael Gromov. Curvature, diameter and Betti numbers. *Comment. Math. Helv.*, 56(2):179–195, 1981.

[7] Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 1999. 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.

[8] Karsten Grove and Peter Petersen, V. Bounding homotopy types by geometry. *Ann. of Math. (2)*, 128(1):195–206, 1988.

[9] Karsten Grove and Katsuhiro Shiohama. A generalized sphere theorem. *Ann. Math. (2)*, 106(2):201–211, 1977.

[10] Allen Hatcher. *Algebraic topology.* Cambridge University Press, Cambridge, 2002.

[11] J. Milnor. *Morse theory.* Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

[12] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics.* Springer, New York, second edition, 2006.

[13] Conrad Plaut. Metric spaces of curvature $\geq k$. In *Handbook of geometric topology*, pages 819–898. North-Holland, Amsterdam, 2002.