TOTALLY GEODESIC MAPS INTO MANIFOLDS WITH NO FOCAL POINTS

BY JAMES DIBBLE

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Xiaochun Rong

and approved by

New Brunswick, New Jersey

October 2014

ABSTRACT OF THE DISSERTATION

Totally geodesic maps into manifolds with no focal points

BY JAMES DIBBLE

Dissertation Director:

Xiaochun Rong

The space of totally geodesic maps in each homotopy class $[F]$ from a compact Riemannian manifold M with non-negative Ricci curvature into a complete Riemannian manifold N with no focal points is path-connected. If $[F]$ contains a totally geodesic map, then each map in $[F]$ is energy-minimizing if and only if it is totally geodesic. When N is compact, each map from a product $W \times M$ into N is homotopic to a smooth map that's totally geodesic on the M-fibers. These results generalize the classical theorems of Eells–Sampson and Hartman about manifolds with non-positive sectional curvature and are proved using neither a geometric flow nor the Bochner identity. They can be used to extend to the case of no focal points a number of splitting theorems proved by Cao–Cheeger–Rong about manifolds with non-positive sectional curvature and, in turn, to generalize a theorem of Heintze–Margulis about collapsing.

The results actually require only an isometric splitting of the universal covering space of M and other topological properties that, by the Cheeger–Gromoll splitting theorem, hold when M has non-negative Ricci curvature. The flat torus theorem is combined with a theorem about the loop space of a manifold with no conjugate points to show that the space of totally geodesic maps in $[F]$ is path-connected. A center-of-mass method due to Cao–Cheeger–Rong is used to construct a homotopy to a totally geodesic map when M is compact. The asymptotic norm of a $\mathbb{Z}^m$-equivariant metric is used to show that the energies of C^1 maps in $[F]$ are bounded below by a constant involving the energy of an affine surjection from a flat Riemannian torus onto a flat semi-Finsler torus, with equality for a given map if and only if it is totally geodesic. This builds on work of Croke–Fathi.

It is also shown that the ratio of convexity radius to injectivity radius can be made arbitrarily small over the class of compact Riemannian manifolds of any fixed dimension at least two. This uses Gulliver's examples of manifolds with focal points but no conjugate points.

Acknowledgments

I'm happy to thank my adviser, Xiaochun Rong, for his patient guidance and insight, which have contributed to this work in ways large and small. I'm grateful to Hector Sussmann, Feng Luo, and Xiaojun Huang for arranging financial support over the years, to Feng Luo, Zheng-Chao Han, and Christopher Croke for serving on my dissertation committee, and to the Department of Mathematics at Rutgers University for its helpful environment. The Department of Mathematics at Capital Normal University in Beijing graciously hosted me at different points during this research.

I'd also like to thank Penny Smith for many interesting mathematical discussions and, moreover, for advice and encouragement that have been invaluable during difficult times in my graduate career. I'm greatly indebted to Jason DeBlois for many interesting and helpful conversations. I would like to thank Christopher Croke for informing me of Kleiner's counterexample to the flat torus theorem and for other useful suggestions. While working out an early proof of some of these results, I benefitted from conversations with Penny Smith and Armin Schikorra, whose comments, had I better understood them at the time, might have saved me a considerable amount of effort. A key step in that proof used the work of Karuwannapatana–Maneesawarng, which I was introduced to by Stephanie Alexander.

I'm certain that the people mentioned here, among many others, have given me far more than I've returned to them. I also know that any attempt to catalog all the ways in which I've been assisted through this process would prove inadequate. To those who have gone unmentioned, know that I'm grateful for your help.

Dedication

To Sofía, whose love and support have made this possible, and to our children.

Table of Contents

Chapter 1

Introduction

A Riemannian manifold N has **no conjugate points** if the exponential map at each point is non-singular and **no focal points** if the exponential map on the normal bundle of each geodesic is non-singular. When N is complete, these are equivalent to any two points in the universal Riemannian covering space $\overline{N}$ being joined by a unique geodesic and, respectively, every distance ball in $\overline{N}$ being strongly convex. Complete manifolds with no conjugate points or no focal points are known to share, to varying degrees, many of the geometric properties of those with non-positive sectional curvature. It almost goes without saying that manifolds with non-positive curvature have been extensively studied, with rigidity results dating back to the celebrated theorem of Gauss–Bonnet. It's perhaps less well known that the modern study of manifolds with no conjugate or no focal points dates back more than seventy years, with notable contributions by, for instance, Hedlund–Morse [HM] and Hopf [Ho]. The following characterizations, due to O'Sullivan [O'S1], show why these three types of spaces should be related.

Theorem 1.1. *(O'Sullivan) Let (N, g) be a complete Riemannian manifold. Then the following hold:*

(a) *N has non-positive sectional curvature if and only if, for every geodesic $\gamma : [0, \infty) \to N$ and every Jacobi field J along γ, $\frac{\mathrm{d}^2}{\mathrm{d}t^2}||J||^2 \geq 0$;*

(b) *N has no focal points if and only if, for every geodesic $\gamma : [0, \infty) \to N$ and every non-trivial Jacobi field J along γ that satisfies $J(0) = 0$, $\frac{\mathrm{d}}{\mathrm{d}t}||J||^2 > 0$; and*

(c) *N has no conjugate points if and only if, for every geodesic $\gamma : [0, \infty) \to N$, every non-trivial Jacobi field J along γ that satisfies $J(0) = 0$, and every positive time, $||J||^2 > 0$.*

It follows that complete Riemannian manifolds with non-positive sectional curvature have no focal points and that those with no focal points have no conjugate points. This refines the well-known theorem of Cartan–Hadamard. On the other hand, Gulliver [Gul] constructed examples of complete

Riemannian manifolds with positive sectional curvature but no focal points and examples with focal points but no conjugate points. The latter examples may be used to show that, for each $m \geq 2$, $\inf \frac{r(M)}{\mathrm{inj}(M)} = 0$ over the class of compact m-dimensional manifolds, where r and inj are the convexity and injectivity radiuses. This fills in a gap in the literature pointed out by Berger [Ber].

Even more so than in the case of no conjugate points, many of the major results about Riemannian manifolds with non-positive sectional curvature generalize to those with no focal points. These include the center theorem [O'S1], flat torus theorem [O'S2], and higher rank rigidity theorem [Wat], as well as the fact that, in dimensions up to three, all compact manifolds that admit metrics with no focal points also admit metrics with non-positive curvature [IK]. This is because many of the arguments used to prove results about non-positively curved manifolds actually depend only on the convexity of certain sets or functions that holds under the weaker assumption that there are no focal points. One such result is the flat strip theorem, which states that, in a complete and simply connected manifold N with no focal points, any two geodesic lines with finite Hausdorff distance bound a totally geodesic flat strip. This was proved by O'Sullivan [O'S2], using a result of Goto [Got], and independently by Eschenburg [Esc]. The flat strip theorem implies that the set of axes of an isometry is convex. Another is that N has no focal points if and only if, for each $y \in N$, $\mathrm{d}^2(\cdot, y)$ is a strictly convex function [Eb1]. By contrast, Burns showed that the flat strip theorem may fail for manifolds no conjugate points [Burn1], and an unpublished example of Kleiner shows the same for the flat torus theorem [Kle].

A celebrated tool in the study of manifolds of non-positive curvature, but which has seen little use in the study of manifolds with no focal points, is the harmonic map heat flow invented by Eells–Sampson. In their foundational paper [ES], they introduced the notion of a harmonic map between Riemannian manifolds as a smooth critical point of the energy functional $\mathrm{E}(u) := \int_M \|\mathrm{d}u\|^2 \,\mathrm{d}\mu_M$, which is a generalization of the Dirichlet energy of a real-valued function defined on Euclidean space. They also constructed a version of the heat equation as the negative gradient flow of this energy and showed that, for any C^1 map $u_0 : M \to N$ between compact manifolds, where N has non-positive sectional curvature, a unique solution $u : M \times [0, \infty) \to N$ to their heat equation

$$\begin{aligned} \frac{\partial u}{\partial t} &= \tau_u \\ u(\cdot, 0) &= u_0 \end{aligned} \tag{1.1}$$

exists for all time and uniformly subconverges as $t \to \infty$ to a harmonic map. That is to say, there exists a sequence of times $t_i \to \infty$ such that $u(\cdot, t_i)$ converges uniformly to a harmonic map. From this, they deduced that every homotopy class of maps contains an energy-minimizing representative. This portion of their argument uses the second variation of the energy under the flow, in which the sectional curvature of the target appears, to show that the energy is a convex function of time. This ensures long-term existence and uniform subconvergence. They also proved that, when the domain is a compact manifold with non-negative Ricci curvature, harmonic maps are totally geodesic.

Theorem 1.2. *(Eells–Sampson) Let M and N be Riemannian manifolds, where M is compact and has non-negative Ricci curvature and N is complete and has non-positive sectional curvature. If $f : M \to N$ is a harmonic map, then f is totally geodesic. If, in addition, M has positive Ricci curvature at a point, then f is constant.*

This portion of their argument uses an identity, ultimately inspired by the work of Bochner [Bo], that relates the second fundamental form of a harmonic map to the Ricci curvature of the domain and the sectional curvature of the target. Specifically, if f is harmonic and β_f its second fundamental form, then

$$\Delta e_f = \|\beta_f\|^2 + \langle f_*(\mathrm{Ric}_M(e_i)), f_*(e_i)\rangle - \langle R_N(f_*(e_i), f_*(e_j)) f_*(e_i), f_*(e_j)\rangle \tag{1.2}$$

Hartman [Har] improved upon the results of Eells–Sampson in the following ways.

Theorem 1.3. *(Hartman) Let M and N be Riemannian manifolds, where M is compact and N is complete and has non-positive sectional curvature. Let $f : M \to N$ be a C^1 map. Then the following hold:*

(a) *The solution $u_f : M \times [0, \infty) \to N$ to the heat equation with $u_f(\cdot, 0) = f(\cdot)$ exists for all time;*

(b) *The solution u_f converges uniformly to a harmonic map $u_\infty(\cdot) := \lim_{t\to\infty} u_f(\cdot, t)$ if and only if f is homotopic to a harmonic map;*

(c) *The set of harmonic maps homotopic to f is path-connected, and energy is constant on it;*

(d) *If f is homotopic to a harmonic map and $h : M \to N$ is another C^1 map such that $d(f, h) < \mathrm{inj}(N)$ uniformly, then $t \mapsto \|u_f(\cdot, t), u_h(\cdot, t)\|_p$ is non-increasing for each $1 \le p \le \infty$.*

As an application of a theorem of Li–Zhu [LZ], it's possible to recover the long-term existence and subconvergence shown by Eells–Sampson under the weaker assumption that the target has no focal points. However, the limit map under such a flow is only known to be harmonic. Since manifolds with no focal points can have positive curvature, the Bochner identity (1.2) cannot be gainfully applied, and it's not at the start clear that a harmonic map, or even an energy-minimizing map, into a manifold with no focal points is totally geodesic, nor that each homotopy class of maps contains a totally geodesic representative. The main point of this dissertation is to generalize the existence results of Eells–Sampson, and to a great extent Hartman's results, to energy-minimizing maps into compact manifolds with no focal points. This also yields a new proof of those results in the case of non-positive sectional curvature.

Theorem 1.4. *Let M be a compact Riemannian manifold with non-negative Ricci curvature, N a complete Riemannian manifold with no focal points, and $[F]$ a homotopy class of maps from M to N. Then the following hold:*

(a) *The set of totally geodesic maps in $[F]$ is path-connected;*

(b) *If $[F]$ contains a totally geodesic map, then each map in $[F]$ is energy-minimizing if and only if it is totally geodesic; and*

(c) *If N is compact, then $[F]$ contains a totally geodesic map.*

In particular, for compact N, a map in $[F]$ is energy-minimizing if and only if it is totally geodesic, and the set of energy-minimizing maps in $[F]$ is non-empty and path-connected. The proof of Theorem 1.4 is more geometric than previous approaches, in the sense that it uses neither the heat flow nor the Bochner identity and generally makes minimal use of partial differential equations. The key tools instead are a structure theorem about the loop space of a complete manifold with no conjugate points, the flat torus theorem [O'S2], the Cheeger–Gromoll splitting theorem [CG2], a center-of-mass construction due to Cao–Cheeger–Rong [CCR1], the asymptotic norm of a periodic metric on $\mathbb{Z}^m$ [BBI], and a characterization of totally geodesic maps that builds upon work of Croke [Cr1] and Croke–Fathi [CF] about energy and intersection. While the original proof of the Cheeger–Gromoll splitting theorem in [CG1] and [CG2] uses the theory of elliptic equations, it's worth pointing out, in the spirit of keeping the analysis as elementary as possible, that Eschenburg–Heintze [EH] later found a proof that uses the maximum principle instead. Furthermore, in the case of a flat domain, the splitting theorem is not needed. A qualitative corollary of Hartman's results also generalizes to manifolds with no focal points.

Theorem 1.5. *Let W and M be Riemannian manifolds, where M is compact and has non-negative Ricci curvature. Endow $W \times M$ with the product metric obtained from W and M. Let N be a compact Riemannian manifold with no focal points and $f : W \times M \to N$ a continuous function. Then f is homotopic to a smooth map that's totally geodesic on each M-fiber.*

To obtain Theorem 1.5 under the stronger assumption that N has non-positive sectional curvature, one could first apply a parameterized heat flow to $W \times M$. That is, one could simultaneously flow each individual M-fiber. The various parts of Theorem 1.3, along with Theorem 1.2, ensure that this flow exists for all time and uniformly converges to a map that's totally geodesic on each M-fiber.

Jost [J1] used the existence results of Eells–Sampson to give a new proof of the flat torus theorem in the case of non-positive sectional curvature. Theorem 1.4 is, loosely speaking, the converse, as it shows that the flat torus theorem can be used as one ingredient in a proof of the results of Eells–Sampson. In fact, the results here are more general than stated in Theorem 1.4 and Theorem 1.5, as they depend only on an isometric splitting of the universal covering space $\overline{M}$, the commutativity of a certain diagram, and a topological property of $[F]$ that, by the Cheeger–Gromoll splitting theorem, hold for compact manifolds with non-negative Ricci curvature. The main results also hold when N is a compact surface with no conjugate points. However, Kleiner's counterexample to the flat torus theorem [Kle], in conjunction with a result of Lemaire [Lem] and, independently, Sacks–Uhlenbeck [SaU], shows that they may fail for higher-dimensional manifolds with no conjugate points. Specifically, for each $n \geq 3$, there exist a compact n-dimensional manifold N with no conjugate points and an energy-minimizing map $T^2 \to N$ that is not totally geodesic.

In principle, it should be possible to use Theorem 1.4 and Theorem 1.5 to generalize to manifolds with no focal points those results for non-positively curved manifolds that depend only on energy-minimizing, rather than harmonic, maps being totally geodesic. As proof of concept, it's noted here, without detailed justification, that they extend to the case of no focal points a number of splitting theorems proved by Cao–Cheeger–Rong [CCR1] about manifolds with non-positive sectional curvature. These build on work in [CCR2]. One such generalization is the following.

Theorem 1.6. *Let M and N be compact Riemannian manifolds of the same dimension. Suppose M admits an F-structure. If there exists a continuous function $f : M \to N$ with non-zero degree, then every metric on N with no focal points admits a local splitting structure for which there is a* *consistency map homotopic to f.*

Precise definitions of the above terminology can be found in [CCR1]. Roughly speaking, a manifold admits an F-structure if it can be cut into pieces that, up to finite covers, admit effective torus actions that are compatible on overlaps. This generalizes the notion of a graph manifold. Theorem 1.6 implies that the universal cover of any such compact manifold with no focal points can be written as a union of isometric products $Z_i = D_i \times \mathbb{R}^{k_i}$, each of which is convex, whose Euclidean factors project to immersed submanifolds that, up to homotopy, contain the orbits of the torus actions. Using this, it's possible to extend a result of Heintze-Margulis [CCR1] to preclude collapsing with bounded sectional curvature for a large class of compact manifolds that admit metrics with no focal points.

Theorem 1.7. *For each $n \in \mathbb{N}$, there exists $\varepsilon = \varepsilon(n) > 0$ such that, if M is a compact n-dimensional manifold that admits a Riemannian metric with no focal points and negative Ricci curvature a point, then, for every metric on M with $|sec_M| \leq 1$, there is a point at which the injectivity radius is at least ε.*

The arguments of [CCR1] work, more or less verbatim, to prove Theorem 1.6 and Theorem 1.7, once Theorem 1.4 and Theorem 1.5 are invoked in the place of the results of Eells–Sampson and Hartman.

Chapter 2

Preliminaries

2.1 Algebraic topology

This section lays out some notation and terminology and proves two well-known algebraic results. It will be assumed that the reader is familiar with the essentials of algebraic topology. A **map** is a continuous function. If X is a topological space and $x, y \in X$, a map $\gamma : [a,b] \to X$ is a **path from** x **to** y if $\gamma(a) = x$ and $\gamma(b) = y$. A path $\gamma : [a,b] \to X$ is a **loop** if $\gamma(a) = \gamma(b)$. The set of natural numbers is $\mathbb{N} = \{1, 2, 3, \ldots\}$. If G is a group, its identity element will be denoted by e. Multiplicative notation for groups will be used throughout, so that, for each $g \in G$, its inverse is denoted g^{-1}. If $S \subseteq G$, then the **subgroup generated by** S is $<S> := \{a_1^{\varepsilon_1} \ldots a_k^{\varepsilon_k} \big| k \geq 1, a_i \in S, \varepsilon_i = \pm 1\} \cup \{e\}$. This is the intersection of all subgroups of G containing S. When $S = \{g_1, \ldots, g_n\}$ is finite, $<S>$ will be denoted $<g_1, \ldots, g_n>$. The **centralizer of** S is $\mathrm{Z}(S) := \{g \in G \big| gs = sg \text{ for all } s \in S\}$. This is always a subgroup of G. When S is finite, as above, the centralizer of S will be denoted $\mathrm{Z}(g_1, \ldots, g_n)$. Although this concept will see light use here, the **normalizer of** S is $\mathrm{N}(S) := \{g \in G \big| g^{-1} sg \in S \text{ for all } s \in S\}$. When H is a subgroup of G, $\mathrm{N}(H)$ is the union of all subgroups of G in which H is normal. The **commutator subgroup of** G is $[G,G] := \{a^{-1}b^{-1}ab \big| a, b \in G\}$. If N is a normal subgroup of G, then G/N is Abelian if and only if $[G,G] \subseteq N$. Generally speaking, the groups considered here will be contained in the fundamental group of a manifold at a point.

If M is a topological manifold and $p_1, p_2 \in M$, it is widely understood that $\pi_1(M, p_1)$ and $\pi_1(M, p_2)$ are isomorphic; an isomorphism between them may be constructed by conjugating loops in $\pi_1(M, p_1)$ by any fixed path connecting p_1 to p_2. In that way, one may speak of the abstract group $\pi_1(M)$, but this obscures the fact that there is no canonical isomorphism between the groups at different points. Since there is no single geometrically realized object $\pi_1(M)$, I will endeavor to avoid the notation $\pi_1(M)$ and explicitly note throughout the dependence on basepoints.

Let $\mathscr{P}(M)$ denote the set of homotopy classes of paths in M. For each $x_1, x_2 \in M$, denote by

$\mathscr{P}(x_1,x_2) \subseteq \mathscr{P}$ the set of homotopy classes of paths from x_1 to x_2. The set $\mathscr{P}(M)$ isn't quite a group under concatenation, since the concatenation of two paths only exists when the endpoints line up; but concatenation is associative whenever it's defined, the constant paths are like identity elements, and reversing the parameterization of a path is akin to inversion.[1] For each $[\alpha] \in \mathscr{P}(x_1,x_2)$, the function $A_{[\alpha]} : \pi_1(M,x_1) \to \pi_1(M,x_2)$ will denote conjugation by $[\alpha]$. More precisely, $A_{[\alpha]}([\gamma]) := [\alpha^{-1} \cdot \gamma \cdot \alpha]$. Each $A_{[\alpha]}$ is a group isomorphism, so the function $[\alpha] \mapsto A_{[\alpha]}$ resembles a group action of $\mathscr{P}(M)$ on $\coprod_{x\in M} \pi_1(M,x)$. Each $A_{[\alpha]}$ extends to an isomorphism $(\pi_1(M,x_1))^k \to (\pi_1(M,x_2))^k$ by setting $A_{[\alpha]}([\gamma_1],\ldots,[\gamma_k]) := (A_{[\alpha]}([\gamma_1]),\ldots,A_{[\alpha]}([\gamma_k]))$. This defines an equivalence relation on $\coprod_{x\in M}(\pi_1(M,x))^k$ by setting $([\sigma_1],\ldots,[\sigma_k]) \cong ([\gamma_1],\ldots,[\gamma_k])$ if and only if there exists $[\alpha] \in \mathscr{P}(M)$ such that $([\sigma_1],\ldots,[\sigma_k]) = A_{[\alpha]}([\gamma_1],\ldots,[\gamma_k])$. If $\mathscr{S}(M) := \{G \leq \pi_1(M,x) \big| x \in M\}$ denotes the set of subgroups of $\coprod_{x\in M} \pi_1(M,x)$ and $\mathscr{S}(x) := \{G \leq \pi_1(M,x)\}$ for each fixed $x \in M$, then something akin to an action of $\mathscr{P}(M)$ on $\mathscr{S}(M)$ is given by taking each $[\alpha] \in \mathscr{P}(x_1,x_2)$ to the function $G \mapsto A_{[\alpha]}(G)$. Overloading notation, this latter function will also be denoted by $A_{[\alpha]} : \mathscr{S}(x_1) \to \mathscr{S}(x_2)$, so that $A_{[\alpha]}(G) = \{A_{[\alpha]}([\gamma]) \big| [\gamma] \in G\}$. This may be used to define an equivalence relation on $\mathscr{S}(M)$.

The deck transformation group of a covering map $\psi : \tilde{M} \to M$ will be be denoted by some variant of the symbol Γ; the exact symbol will always be clear in context. In many situations, it will be necessary to discuss nested covering maps, that is, covering maps of spaces that themselves are covering spaces; a bar will typically be used to denote something associated with the higher covering space, while tilde will be reserved for the intermediate cover. In keeping with that convention, the universal covering map of M will typically be denoted by $\pi : \overline{M} \to M$; its deck transformation group might be denoted $\overline{\Gamma}$. Recall that $\pi : \overline{M} \to M$ is a **normal** covering map, which by definition is taken to mean that its deck group acts transitively on each fiber. Equivalently, one could define a covering map $\phi : \tilde{M} \to M$ to be normal if, for each $\tilde{p} \in \tilde{M}$, $\phi_*(\pi_1(\tilde{M},\tilde{p}))$ is a normal subgroup of $\pi_1(M,\phi(p))$. If $p \in M$ and $\tilde{p} \in \phi^{-1}(p)$, then the deck transformation group $\tilde{\Gamma}$ of a covering map ϕ is naturally identified with the quotient $\mathrm{N}\Big(\phi_*(\pi_1(\tilde{M},\tilde{p}))\Big)/\phi_*(\pi_1(\tilde{M},\tilde{p}))$, where $\mathrm{N}\Big(\phi_*(\pi_1(\tilde{M},\tilde{p}))\Big)$ is the normalizer of $\phi_*(\pi_1(\tilde{M},\tilde{p}))$ in $\pi_1(M,p)$ that is, the largest subgroup in which $\phi_*(\pi_1(\tilde{M},\tilde{p}))$ is normal. In particular, when ϕ is a normal covering map, $\tilde{\Gamma} \cong \pi_1(M,p)/\phi_*(\pi_1(\tilde{M},\tilde{p}))$. When $\pi : \overline{M} \to M$ is the universal covering map, M is simply connected, so $\pi_*(\pi_1(\overline{M},\overline{p})) = <e>$ for any $\overline{p} \in \overline{M}$, and one sees

[1]In the parlance of category theory, $\mathscr{P}(M)$ is often denoted $\pi_1(M)$ and called the fundamental groupoid of M. But such abstraction isn't important here.

that $\overline{\Gamma} \cong \pi_1(M, p)$. Under this identification, $[\sigma] \in \pi_1(M, p)$, represented by the loop $\sigma : [0,1] \to M$, corresponds to the deck transformation that takes $\overline{p}$ to $\overline{\sigma}(1)$, where $\overline{\sigma} : [0,1] \to \overline{M}$ is the lift of σ with $\overline{\sigma}(0) = \overline{p}$.

If G is a group and H a subgroup, the **index of H in G**, denoted $[G : H]$, is the number of either left or right cosets of H in G. The following well-known algebraic results have important consequences in the theory of covering spaces. They will be useful in Chapter 6.

Lemma 2.1.1. *Let H be a finite-index subgroup of a group G. Then H contains a subgroup N that is normal in G and has index satisfying* $[G : N] \leq [G : H]! < \infty$.

Proof. Note that, for each $g \in G$, $gaH = gbH$ if and only if $aH = bH$. One may therefore define an injective function $\phi_g : G/H \to G/H$ by setting $\phi_g(aH) := gaH$. If $aH \in G/H$, then $aH = \phi_g(g^{-1}aH)$, so ϕ_g is also surjective. Thus ϕ_g is a permutation of G/H; denote by $\Sigma(G/H)$ the group of such permutations under composition. Define $\Phi : G \to \Sigma(G/H)$ by $\Phi(g) := \phi_g$. For all $g_1, g_2 \in G$, $\phi_{g_1 g_2}(aH) = (g_1 g_2)aH = g_1(g_2 aH) = \phi_{g_1} \circ \phi_{g_2}(aH)$. That is to say, $\Phi(g_1 g_2) = \Phi(g_1) \circ \Phi(g_2)$, so Φ is a homomorphism. Thus $N := \ker(\Phi)$ is a normal subgroup of G and, by the first isomorphism theorem, $G/N \cong \Phi(G)$. Since $\Sigma(G/H)$ has cardinality $[G : H]!$, this implies that $[G : N] \leq [G : H]!$. As the kernel of Φ, N consists of those $g \in G$ such that $\phi_g(aH) = gaH = aH$ for all $a \in G$. It follows that $N = \cap_{a \in G} aHa^{-1}$ and, consequently, $N \subseteq eHe^{-1} = H$.

□

Lemma 2.1.2. *Let G_1 and G_2 be groups and H a normal subgroup of $G_1 \times G_2$ such that* $[G_1 \times G_2 : H] < \infty$. *Then there exist normal subgroups N_i of G_i, $i = 1, 2$, such that $N_1 \times N_2 \subseteq H$,* $[G_i : N_i] < [G_1 \times G_2 : H]$ *for each i, and* $[G_1 \times G_2 : N_1 \times N_2] \leq [G_1 \times G_2 : H]^2 < \infty$.

Proof. For each $i = 1, 2$, denote by $\frac{1}{2}\rho_i : G_1 \times G_2 \to G_i$ projection onto the i-th factor. Then each $\ker(\frac{1}{2}\rho_i)$ is a normal subgroup of $G_1 \times G_2$, so $H \cap \ker(\frac{1}{2}\rho_i)$ is normal as well. Since $\ker(\frac{1}{2}\rho_1) = \{e\} \times G_2$ and $\ker(\frac{1}{2}\rho_2) = G_1 \times \{e\}$, $\frac{1}{2}\rho_1|_{\ker(\frac{1}{2}\rho_2)} : \ker(\frac{1}{2}\rho_2) \to G_1$ and $\frac{1}{2}\rho_2|_{\ker(\frac{1}{2}\rho_1)} : \ker(\frac{1}{2}\rho_1) \to G_2$ are isomorphisms. Let $N_1 := \frac{1}{2}\rho_1(H \cap \ker(\frac{1}{2}\rho_2))$ and $N_2 := \frac{1}{2}\rho_2(H \cap \ker(\frac{1}{2}\rho_1))$. As the image of a normal subgroup under an isomorphism, each N_i is a normal subgroup of G_i. Note that $N_1 \times N_2$ equals the subset product $(H \cap \ker(\frac{1}{2}\rho_2))(H \cap \ker(\frac{1}{2}\rho_1))$, which is contained in H. By the second isomorphism theorem, $(\ker(\frac{1}{2}\rho_i)H)/H \cong \ker(\frac{1}{2}\rho_i)/(H \cap \ker(\frac{1}{2}\rho_i))$. Since $\ker(\frac{1}{2}\rho_2)/(H \cap \ker(\frac{1}{2}\rho_2)) \cong G_1/N_1$, it follows that $[G_1 : N_1] = [\ker(\frac{1}{2}\rho_i)H : H] \leq [G_1 \times G_2 : H]$. Similarly, $[G_2 : N_2] \leq [G_1 \times G_2 : H]$. Since $(G_1 \times G_2)/(N_1 \times N_2) \cong (G_1/N_1) \times (G_2/N_2)$, it follows that $[G_1 \times G_2 : N_1 \times N_2] = [G_1 : N_1][G_2 : N_2] \leq [G_1 \times G_2 : H]^2$.

□

2.2 Riemannian geometry

It will be assumed that the reader knows the basics of differential and Riemannian geometry. These include the notions of smooth manifolds, partitions of unity, Riemannian metrics, differential forms, vector bundles, connections, curvature, geodesics, Jacobi fields, and so forth. Excellent references for this background material include the books of Klingenberg [Kli1], Chavel [Chav], and Cheeger–Ebin [CE]. The textbooks by Lee on topological [Lee1], smooth [Lee2], and Riemannian [Lee3] manifolds were profoundly influential on my thinking during my first years as a graduate student, and an astute reader will no doubt hear their echoes here. In Chapter 7, it will also be assumed that the reader is at least somewhat familiar with length spaces and Finsler manifolds, although the prerequisites there are rather light. The treatment in the textbook by Burago–Burago–Ivanov [BBI] more than suffices. As far as Finsler manifolds are concerned, one need only have the intuition that they are endowed with a norm, rather than an inner product, on each tangent space.

Wherever possible, manifolds in this dissertation are assumed to be connected. The only exceptions are those that cannot be assumed so because they result from constructions that may produce disconnected spaces. For example, in Chapter 3, the inverse function theorem is invoked to construct a submanifold $\tilde{N}$ of a tensor bundle $\mathrm{T}^{(k,0)}N$, and this space will not in general be connected. Manifolds will also be assumed smooth. However, manifolds won't necessarily be assumed complete. When it's relevant that a manifold M may have boundary $\partial M \neq \emptyset$, this will be noted explicitly.

The tangent bundle of a smooth manifold M will be denoted by $\mathrm{T}M$ and the tangent space at each point $p \in M$ by T_pM. When M is Riemannian, its unit sphere bundle is $\mathrm{S}M$, and the normal bundle to a submanifold $S \subseteq M$ is $\mathrm{N}S$. One similarly has at each point the spaces S_pM and, for $p \in S$, N_pS. The sphere of radius $r \geq 0$ in T_pM will be denoted $\mathrm{S}_p(r)$. The set of (k,l)-tensors on the tangent space T_pM is denoted

$$\mathrm{T}_p^{(k,l)}M = \underbrace{\mathrm{T}_pM \otimes \cdots \otimes \mathrm{T}_pM}_{k \text{ times}} \otimes \underbrace{\mathrm{T}_p^*M \otimes \cdots \otimes \mathrm{T}_p^*M}_{l \text{ times}}$$

Here, T_p^*M is the dual space to T_pM. The vector bundle of (k,l)-tensors over M is $\mathrm{T}^{(k,l)}M = \coprod_{p \in M} \mathrm{T}_p^{(k,l)}M$. The symbol π will denote projection to the basepoint of a vector bundle. In presenting some of the background material, it will sometimes be used implicitly that a bilinear map $V \times W \to Z$, where V, W, and Z are vector spaces, may be canonically identified with a linear map $V \otimes W \to Z$. If V and W are vector spaces, then the space of linear maps from V to W will be denoted $\mathscr{L}(V, W)$.

To say that a map between manifolds C^k means that, in any local coordinates, all of its k-th order partial derivatives exist and are continuous. When $k = \infty$, this means that the map is C^k for all $k \in \mathbb{N}$. The space of C^k maps from M to N is denoted $\mathrm{C}^k(M, N)$. Where it exists, the push-forward, or differential, of a map $f : M \to N$ between manifolds is denoted f_*. When $f \in \mathrm{C}^1(M, N)$, its derivative $\mathrm{D}f$ exists at each point and is given by $\mathrm{D}f(v) = f_*(v)$. Formally speaking, this may be identified with a section of the vector bundle $\coprod_{p \in M} \mathscr{L}(\mathrm{T}_pM, \mathrm{T}_{f(p)}N)$ over M. That is, for each $p \in M$, $\mathrm{D}_pf : \mathrm{T}_pM \to \mathrm{T}_{f(p)}M$ is a linear map. Note that $f_*(v)$ may exist for all $v \in \mathrm{T}M$ without f being C^1.

The exponential map of a Riemannian manifold M will be denoted by $\exp : \mathrm{T}M \to M$ and its restriction to the tangent space at $p \in M$ by $\exp_p : \mathrm{T}_pM \to M$. For each $v \in \mathrm{T}M$, the geodesic determined by v will be written γ_v. That is, $\gamma_v(t) := \exp(tv)$ whenever that expression is defined. For complete M, the geodesic flow $\Psi : \mathbb{R} \times \mathrm{S}M \to \mathrm{SM}$ is the map defined by $\Psi(t, v) := \gamma_v'(t)$. For each fixed $T \in \mathbb{R}$, $\Psi^T : \mathrm{S}M \to \mathrm{S}M$ is defined by $\Psi^T(\cdot) := \Psi(T, \cdot)$.

The Levi-Civita connection will be the only one ever used. For any vector fields X and Y, their covariant derivative will be denoted $\nabla_X Y$. One has that $\nabla_X Y - \nabla_Y X = [X, Y]$, where $[X, Y]$ is the Lie bracket. The Levi-Civita connection induces a covariant derivative along each curve γ in M, denoted $\nabla_{\gamma'}$. It will be assumed that the reader is familiar with the basics of the calculus of variations, especially including the first and second variations of length and energy. A **geodesic loop** $\gamma : [a, b] \to M$ is a loop that's also a geodesic. Note that a geodesic loop may not be smoothly closed at the endpoints. That is, it's not necessarily the case that $\lim_{t \searrow a} \gamma'(t) = \lim_{t \nearrow b} \gamma'(t)$. If this does hold, then γ is a **closed geodesic**. A closed geodesic descends to a geodesic $S^1 \to M$ under the quotient map that identifies a and b. A geodesic $\gamma : \mathbb{R} \to M$ is a **closed geodesic** if it restricts to a closed geodesic on some finite interval or, equivalently, if it is periodic.

Unless otherwise stated, whenever a covering space of a Riemannian manifold is endowed with a Riemannian metric, that will be the pull-back metric from the covering map. When M is a Riemannian manifold, its isometry group will be denoted by $\mathscr{I}(M)$. If $\phi \in \mathscr{I}(M)$, an **axis of** ϕ is a geodesic $\gamma : \mathbb{R} \to M$ such that, for some $t_0 \in \mathbb{R}$, $\phi(\gamma(t)) = \gamma(t + t_0)$ for all $t \in \mathbb{R}$. The **displacement function of** ϕ is $x \mapsto \mathrm{d}(x, \phi(x))$, and the **minimum set of** ϕ, denoted $\min(\phi)$, is the set of points that minimize its displacement function. The deck transformation group $\overline{\Gamma}$ of $\pi : \overline{M} \to M$ acts by isometries, so one may speak of the axes of each deck transformation γ, which are naturally identified with the closed geodesics in M freely homotopic to any representative of $[\gamma]$.

Lemma 2.2.1. *Let M be a complete Riemannian manifold, $p \in M$, and γ a deck transformation of $\pi : \overline{M} \to M$. Then $\overline{\alpha} : \mathbb{R} \to M$ is an axis of γ if and only if $\pi \circ \overline{\alpha}$ is a closed geodesic in M whose restriction to some finite interval represents $A_{[\sigma]}([\gamma])$ for some path $\sigma : [a,b] \to M$.*

It's well-known that every non-trivial free homotopy class of loops in a compact manifold contains a closed geodesic. This may be proved using, say, the classical theorem of Arzelà–Ascoli. Therefore, when M is compact, the set of axes corresponding to any non-trivial deck transformation is non-empty.

It will be helpful to know that a local isometry from a complete Riemannian manifold into a simply connected manifold is a diffeomorphism. This was proved by Ambrose [Am].

Lemma 2.2.2. *(Ambrose) Let M and N be Riemannian manifolds and $f : M \to N$ a local isometry. If M is complete and N is simply connected, then f is a diffeomorphism.*

This is a special case of a more general result, a proof of which may be found in [CE].

Lemma 2.2.3. *Let M and N be Riemannian manifolds and $f : M \to N$ a local isometry. If M is complete, then f is a covering map.*

2.3 Integration

The basics of geometric measure theory will be taken for granted. The symbol μ will typically be used to denote a measure, with a subscript to indicate the space it measures. For example, the

Lebesgue measure induced by a Riemannian metric on a manifold M will be denoted μ_M. Most spaces under discussion will naturally be endowed with certain well-understood measures. The σ-algebra associated to a measure space will be suppressed in the notation, as it will usually be implied by the context.

The unit sphere bundle $\mathrm{S}M$ of a Riemannian manifold M is naturally endowed with the Liouville measure $\mu_{\mathrm{S}M}$. The key points about $\mu_{\mathrm{S}M}$, which are rigorously developed in [Chav], are that it's invariant under the geodesic flow Ψ and that it's locally the product measure $\mu_M \times \mu \mathrm{S}M$. The following is a classical fact about measure-preserving transformations, known as the Poincaré recurrence theorem.

Theorem 2.3.1. *(Poincaré) Let (X,μ) be a measure space with finite total measure and $f : X \to X$ a measure-preserving transformation. For any measurable $A \subseteq X$, almost all points of A return to A infinitely often under the iterates of f. Consequently, if $\mu(A) > 0$, then for any $K \in \mathbb{N}$ there exists $k \geq K$ such that $f^k(A) \cap A \neq \emptyset$.*

Applying this to the geodesic flow, one obtains the following result, also often referred to as the Poincaré recurrence theorem in the literature.

Corollary 2.3.2. *Let M be a complete Riemannian manifold with finite volume and $U \subseteq \mathrm{S}M$ a set with positive Liouville measure. For any $t > 0$, there exists $T \geq t$ such that $\Psi^T(U) \cap U \neq \emptyset$.*

The following result, known as Santaló's formula, will play an important role in Chapter 7. In the following, (M,g) is a compact Riemannian manifold with boundary $\partial M \neq \emptyset$. Let ν denote the inward-pointing unit normal vector field along ∂M, and define $\mathrm{S}(\partial M) := \{w \in \mathrm{S}M \,|\, \pi(w) \in \partial M\}$ and $\mathrm{S}^+\partial M := \{w \in \mathrm{S}(\partial M) \,|\, g(w,\nu) > 0\}$. The Liouville measure on $\mathrm{S}M$ restricts to a measure $\mu_{\mathrm{S}^+\partial M}$ on $\mathrm{S}^+\partial M$. For each $w \in \mathrm{S}^+\partial M$, $\ell(w) := \min\{t > 0 \,|\, \gamma_w(t) \in \partial M\}$ and $\varsigma_w : [0,\ell(w)] \to M$ is defined by $\varsigma_w(t) := f \circ \gamma_w|_{[0,\ell(w)]}$. That is, ς_w is the image under f of the geodesic γ_w, defined until γ_w hits ∂M again.

Theorem 2.3.3. *Let (M,g) be a compact Riemannian manifold with boundary $\partial M \neq \emptyset$. Let $f : M \to \mathbb{R}$ be a measurable function. Then*

$$\int_{\mathrm{S}M} f(v)\,\mathrm{d}\mu_{\mathrm{S}M} = \int_{\mathrm{S}^+\partial M} \Big[\int_0^{\ell(w)} f(\Psi^T(w))\,\mathrm{d}t \Big] g(w,\nu)\,\mathrm{d}\mu_{\mathrm{S}^+\partial M}$$

It will also be helpful to record the coarea formula, which allows one to convert integrals over the domain of a smooth surjection into integrals over the target. This result appears in many forms in the literature; the one presented here will be the most useful for the work to come.

Theorem 2.3.4. *Let M and N be Riemannian manifolds, $f : M \to N$ a smooth surjection whose push-forward f_* is surjective almost everywhere, and $\phi : M \to [0,\infty]$ a measurable function. Let $\mathscr{J}_f := |\det(f_*|_{\ker(f_*)^\perp})|$. That is, $\mathscr{J}_f$ is the Jacobian of the restriction of f_* to the orthogonal complement of its kernel. Then*

$$\int_M \phi(x)\,\mathrm{d}\mu_M = \int_N \Big[\int_{f^{-1}(y)} \frac{\phi(x)}{\mathscr{J}_f(x)}\,\mathrm{d}\mu_{f^{-1}(y)} \Big]\,\mathrm{d}\mu_N$$

One may check [Chav] for more information about Santaló's formula and the coarea formula. This section ends with a basic inequality that will be used in the proof of Theorem 7.4.4. For want of a better place, it's recorded here.

Lemma 2.3.5. *Let $f : [a,b] \to \mathbb{R}$ be a measurable function, and let $a = x_0 < x_1 < \cdots < x_n = b$ be a partition of $[a,b]$ for some $n \geq 1$. Then* $\displaystyle\sum_{i=0}^{n-1} \frac{[\int_{x_i}^{x_{i+1}} f(t)\,\mathrm{d}t]^2}{x_{i+1} - x_i} \geq \frac{[\int_a^b f(t)\,\mathrm{d}x]^2}{b-a}$.

Proof. The result is immediate in the case $n = 1$. Suppose that $n = 2$. It must be shown that

$$\frac{[\int_a^{x_1} f(t)\,\mathrm{d}t]^2}{x_1 - a} + \frac{[\int_{x_1}^{b} f(t)\,\mathrm{d}t]^2}{b - x_1} \geq \frac{[\int_a^b f(t)\,\mathrm{d}t]^2}{b-a}$$

Since $\int_a^b f(t)\,\mathrm{d}t = \int_a^{x_1} f(t)\,\mathrm{d}t + \int_{x_1}^b f(t)\,\mathrm{d}t$, elementary algebraic manipulations show the above to be equivalent to

$$[(b - x_1)\int_a^{x_1} f(t)\,\mathrm{d}t - (x_1 - a)\int_{x_1}^b f(t)\,\mathrm{d}t]^2 \geq 0$$

This proves the result in the case $n = 2$. The general case $n \geq 3$ follows by induction.

□

One may note that the inequality in Lemma 2.3.5 is equality whenever f is constant.

2.4 Convexity

A function $h : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, is **convex** if $h(st_1 + (1-s)t_2) \leq sh(t_1) + (1-s)h(t_2)$ for all $s \in [0,1]$ and $t_1, t_2 \in I$. It actually suffices to check this property at the midpoint of an arbitrary subinterval; that is, h is convex if and only if $h(\frac{t_1+t_2}{2}) \le \frac{h(t_1)+h(t_2)}{2}$ for all $t_1, t_2 \in I$. The function is **strictly convex** if the first inequality, or equivalently the second, is strict for all $s \in (0,1)$ and all $t_1, t_2 \in I$ with $t_1 \neq t_2$. If h is convex, then h is automatically continuous. If h is twice differentiable, then convexity is equivalent to the condition that $h'' \ge 0$. If $h'' > 0$, then h is strictly convex, but the converse does not hold. For example, the function $t \mapsto t^2$ is strictly convex, but its derivative vanishes at zero. However, a twice differentiable function h is strictly convex if and only if $h'' \ge 0$ and $h'' = 0$ at most once.

Let (M,g) be a Riemannian manifold. A function $h : M \to \mathbb{R}$ is **convex** if, for each geodesic $\gamma : [a,b] \to M$, the composition $g \circ \gamma : [a,b] \to \mathbb{R}$ is convex. It is **strictly convex** if each $h \circ \gamma$ is strictly convex. As before, convex functions on manifolds are automatically continuous. Note that the sum of convex functions is convex, and the sum of strictly convex functions is strictly convex. This generalizes to integrals of families of convex or strictly convex functions.

Lemma 2.4.1. *Let M be a complete Riemannian manifold and (A,μ) a measure space. Suppose that $\{h_\alpha : M \to \mathbb{R} \,\big|\, \alpha \in A\}$ is a family of convex functions. Then the function $F : M \to \mathbb{R}$ defined by $F(x) := \int_A h_\alpha(x)\,\mathrm{d}\mu$ is convex. If each h_α is strictly convex, then F is strictly convex.*

Proof. Let $\alpha : [a,b] \to M$ be a geodesic. Then

$$\begin{aligned} F\Big(\frac{t_1+t_2}{2}\Big) &= \int_A h_\alpha\Big(\frac{t_1+t_2}{2}\Big)\mathrm{d}\mu \\ &\le \int_A \frac{h_\alpha(t_1)+h_\alpha(t_2)}{2}\,\mathrm{d}\mu \\ &= \frac{F(t_1)+F(t_2)}{2} \end{aligned}$$

This shows that F is convex. If each h_α is strictly convex, the inequalities become strict, and F is strictly convex.

□

A key fact is that a C^2 function $h : M \to \mathbb{R}$ is strictly convex whenever $\nabla^2 h$ is everywhere positive-definite. Here, $\nabla^2 f$ denotes the Hessian of h, which is defined to equal the covariant derivative of the gradient $\operatorname{grad} h$. This is a symmetric $(0,2)$-tensor on M satisfying

$$\nabla^2 h(v,w) = g(\nabla_v \operatorname{grad} h, w)$$

for each $p \in M$ and $v, w \in \mathrm{T}_p M$. Furthermore, $\nabla^2 h$ is positive-definite if and only if, for every non-constant geodesic $\gamma : (-\varepsilon, \varepsilon) \to M$, $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\big|_{s=0}(h \circ \gamma)(s) > 0$.

There are many different notions of convexity that can be defined for subsets of a manifold M, three of which will be useful here. A subset $X \subseteq M$ is **convex** if, given any $p, q \in X$, there exists a minimal geodesic $\gamma : [a,b] \to M$ connecting p to q such that $\gamma([a,b]) \subseteq X$. Note that the minimal geodesics γ do not have to be unique for X to be convex. A subset $X \subseteq M$ is **strongly convex** if, given any $p, q \in X$, there is a unique minimal geodesic $\gamma : [a,b] \to M$ connecting p to q and, further, $\gamma([a,b]) \subseteq X$. A subset $X \subseteq M$ is **locally convex** if, for each $x \in \overline{X}$, there exists $0 < \varepsilon < r(x)$ such that $X \cap \mathrm{B}(x, \varepsilon)$ is strongly convex. Here, $r(x)$ is the injectivity radius of x, which is discussed in the next section. One may check that the closure of a convex set, when complete, is again convex and that a convex subset of a strongly convex set is strongly convex. Ozols [Oz] and, independently, Cheeger–Gromoll [CG2] proved the following structure theorem for locally convex sets.

Theorem 2.4.2. *(Ozols, Cheeger–Gromoll) Let M be a Riemannian manifold. If $X \subseteq M$ is a closed and locally convex set, then X is an embedded submanifold of M with smooth and totally geodesic interior and possibly non-smooth boundary.*

Some elementary facts about strictly convex functions defined on convex sets are listed in the following lemma.

Lemma 2.4.3. *Let M be a Riemannian manifold and $X \subseteq M$ a convex set. If $f : X \to \mathbb{R}$ is a strictly convex function, then the following hold:*

(a) *There exists at most one local minimum of h on X;*

(b) *If h has a local minimum at $p \in X$, then p is the unique global minimum of h on X; and*

(c) *If X is compact, then h has a unique local minimum on X, which is also its global minimum.*

Proof. **(a)** Assume that $p_1, p_2 \in X$ are distinct local minimums of X. Since X is convex, there exists a geodesic $\gamma : [a,b] \to M$ with $\gamma(a) = p_1$, $\gamma(b) = p_2$, and $\gamma([a,b]) \subseteq X$. The composition $h \circ \gamma : [a,b] \to \mathbb{R}$ is strictly convex with local minimums at a and b, which is a contradiction.

(b) Assume there exists $q \in X$ such that $q \neq p$ and $h(q) \leq h(p)$. As before, let $\gamma : [a,b] \to M$ be a geodesic connecting p to q that remains inside X. Then $h \circ \gamma$ is strictly convex. Since p is the unique local minimum in X, there must exist $a < t_0 < b$ such that $h \circ \gamma(t_0) > h(p)$, as otherwise one would have $h \circ \gamma(t) = h(a)$ for all t, contradicting uniqueness. It follows that $h \circ \gamma|_{[a,b]}$ must have a global maximum, which contradicts the fact that h is strictly convex.

(c) Since h is continuous, it must have a global minimum on X, which is also a local minimum. By part (a), this local minimum is unique.

□

If $X \subseteq M$, then the **convex hull of** X, denoted $\mathrm{conv}(X)$, is the smallest set containing X with the property that, for any $p, q \in \mathrm{conv}(X)$, the image of every minimal geodesic in M from p to q is contained in $\mathrm{conv}(X)$. This is equal to the intersection of all subsets of M that contain X and have this property. If X is contained in a strongly convex set, then $\mathrm{conv}(X)$ is the intersection of all strongly convex sets containing X.

2.5 Geometric radiuses

It's assumed throughout this section that (M, g) is a complete Riemannian manifold. It's well-known that there exist continuous functions $\mathrm{inj}, r : M \to (0, \infty]$ such that, for each $x \in M$,

$$\mathrm{inj}(x) = \max\{R > 0 \,\big|\, \exp_x|_{\mathrm{B}(0,s)} \text{ is injective for all } 0 < s < R\}$$

and

$$r(x) = \max\{R > 0 \,\big|\, \mathrm{B}(x, s) \text{ is strongly convex}\}$$

where $\mathrm{B}(0, s) \subset \mathrm{T}_x M$ denotes the Euclidean ball of radius s around the origin. The number $\mathrm{inj}(x)$ is called the **injectivity radius of** x, and $r(x)$ is called the **convexity radius of** x. If $S \subseteq M$, the **injectivity radius of** S is defined by $\mathrm{inj}(S) := \inf\{\mathrm{inj}(x) \,\big|\, x \in S\}$, and the **convexity radius of** S is defined by $r(S) := \inf\{r(x) \,\big|\, x \in S\}$. If S is compact, then those infimums are minimums and $\mathrm{inj}(S), r(S) > 0$. One may find proofs of these results in [Kli1] and [CE]. In the first, it's also shown that there exists a continuous function $\rho : M \to (0, \infty]$ with the property that

$$\rho(x) = \max\{R > 0 \,\big|\, \mathrm{B}(z, \rho(x)) \text{ is strongly convex for all } z \in \mathrm{B}(x, \rho(x))\}$$

This strengthens the defining property of r, and, as such, $\rho(x)$ is called the **strong convexity radius of** x. In fact, once it's known that r is positive and continuous, an elementary argument shows that, for any $\varepsilon > 0$,

$$\rho(x) \geq \min\left\{\varepsilon, r(z) \,\Big|\, \mathrm{d}(z, x) \leq \min\{\varepsilon, \mathrm{inj}(x)\}\right\} > 0$$

The continuity of ρ follows from the observation that $|\rho(z) - \rho(x)| \leq \mathrm{d}(z, x)$. By construction, one has that $\rho(x) = \infty$ for some $x \in M$ if and only if $r(M) = \infty$.

In a similar fashion, one may define the **conjugate radius of** x by

$$r_c(x) := \min\{T > 0 \,\big|\, \exists \text{ a non-trivial normal Jacobi field } J \text{ along a unit-speed geodesic } \gamma$$
$$\text{with } \gamma(0) = x, J(0) = 0 \text{ and } J(T) = 0\}$$

and the **focal radius of** x by

$$r_f(x) := \min\{T > 0 \,\big|\, \exists \text{ a non-trivial normal Jacobi field } J \text{ along a unit-speed geodesic } \gamma$$
$$\text{with } \gamma(0) = x, J(0) = 0 \text{ and } \|J\|'(T) = 0\}$$

If $\gamma_k : [0, T] \to M$ are unit-speed geodesics starting at x and J_k are normal Jacobi fields along γ satisfying $\|J\|'(0) = 1$, then, by passing a subsequence, one may without loss of generality suppose that γ_k and J_k converge uniformly, along with their derivatives, to a unit-speed geodesic γ starting at x and, respectively, a normal Jacobi field along γ satisfying $\|J\|'(0) = 1$. It follows by normalizing the Jacobi fields in their definitions that $r_c(x)$ and $r_f(x)$ are well-defined. It's well-known that $r_c, r_f > 0$, facts which underlie the standard proofs that $\mathrm{inj}, r > 0$. The essential point in proving this is that an upper bound K on the sectional curvatures in $U \subseteq M$ implies that, for any normal Jacobi field J along a unit-speed geodesic $\gamma : [0, T] \to U$ such that $J(a) = 0$, $\|J\|'(t) \geq \|\hat{J}\|'(t)$ for all $0 \leq t \leq r(S^2_K)$, where S^2_K is the model space of constant sectional curvature K and $\hat{J}$ is a normal Jacobi field along a unit-speed geodesic $\hat{\gamma} : [0, T] \to S^2_K$ satisfying $\hat{J}(0) = 0$ and $\|\hat{J}\|'(0) = \|J\|'(0)$. This follows from the standard proof of the Rauch comparison theorem,[2] which ultimately dates back to the work in [R]. Moreover, if J is a non-trivial Jacobi field with $J(0) = 0$ and $J(T) = 0$, then, since $\|J\|'(0) > 0$, there must exist $0 < t < T$ such that $\|J\|'(t) = 0$. It follows that $r_f < r_c$. If $S \subseteq M$, the **conjugate radius of** S is defined by $r_c(S) := \inf\{r_c(x) \,\big|\, x \in S\}$, and the **focal radius of** S is defined by $r_f(S) := \inf\{r_f(x) \,\big|\, x \in S\}$.

The geometric significance of the conjugate radius is that $\exp_x$ has invertible derivative at every vector in $\mathrm{B}(0, r_c(x)) \subseteq \mathrm{T}_x M$, and, consequently, is a local diffeomorphism. However, the global topology of M may cause $\exp_x$ to not be injective on $\mathrm{B}(0, r_c(x))$. For example, if T^n is a flat torus,

[2] The usual statement of the Rauch comparison theorem is more general than this, but under these assumptions implies that $\|J\|(t) \geq \|\hat{J}\|(t)$ for all $0 \leq t \leq \mathrm{inj}(S^2_K)$.

then $r_c(T^n) = \infty$, while $\mathrm{inj}(T^n) < \infty$. The geometric intuition behind the focal radius is given by the fact that, at least for $0 < t < \mathrm{inj}(x)$, $\|J\|'(t) = \mathrm{II}_t(J,J)$, where $\mathrm{II}_t(\cdot,\cdot)$ denotes the scalar second fundamental form of the distance sphere $\partial\mathrm{B}(x,t)$, measured with respect to the outward pointing normal $\gamma'(t)$. Therefore, for $0 < t < \min\{\mathrm{inj}(x), r_f(x)\}$, $\mathrm{II}_t(\cdot,\cdot)$ is positive-definite.

Let S be a submanifold of M, $z \in M$, and $\gamma : [a,b] \to M$ a geodesic. Then z **is focal to S along γ** if $\gamma(b) = z$ and there exists a variation $\Gamma : (-\varepsilon,\varepsilon)\times[a,b] \to M$ of γ through geodesics $\gamma_s(\cdot) = \Gamma(s,\cdot)$ such that $\gamma_s'(a) \in \mathrm{N}S$ for all $s \in (-\varepsilon,\varepsilon)$ and whose variation field J is non-trivial and vanishes at time b. For example, the north pole on S^2 is focal to the equator along any of the great circles connecting them. It turns out that the existence of focal points to S can be characterized in terms of the existence of certain types of Jacobi fields along geodesics normal to S. Recall that the second fundamental form II_S assigns to each $p \in S$ a symmetric bilinear map $\mathrm{T}_pS \times \mathrm{T}_pS \to \mathrm{T}_pS$. This is given by

$$\mathrm{II}_S(x,y) := \nabla^M_X Y - \nabla^S_X Y = (\nabla^M_X Y)^\perp$$

for any vector fields X and Y such that $X|_S, Y|_S \in \mathrm{T}S$, $X(p) = x$, and $Y(p) = y$. Formally speaking, up to the identification of bilinear maps $\mathrm{T}_pS \times \mathrm{T}_pS \to \mathrm{T}_pS$ with linear maps $\mathrm{T}_pS \otimes \mathrm{T}_pS \to \mathrm{T}_pS$, II_S is identified with a section of the vector bundle $\coprod_{p\in S} \mathscr{L}(\mathrm{T}_pS \odot \mathrm{T}_pS, \mathrm{T}_pS)$. In a similar way, one may define a section A_S of $\coprod_{p\in S} \mathscr{L}(\mathrm{N}_pS \otimes \mathrm{T}_pS, \mathrm{T}_pS)$ by setting

$$A_S(z,x) := (\nabla^M_X Z)^\perp - \nabla^M_X Z = -(\nabla^M_X Z)^\top$$

for any vector fields Z and X such that $Z|_S$ is a local section of $\mathrm{N}S$, $X|_S$ is a local section of $\mathrm{T}S$, $Z(p) = z$, and $X(p) = x$. The expressions on the right-hand side are bilinear over $\mathrm{C}^\infty(M,\mathbb{R})$, so A_S is well-defined and, at each point, bilinear. From the classical Weingarten equation

$$g(\nabla^M_X Z, Y) = -g(Z, \mathrm{II}_S(X,Y))$$

one sees that A_S and II_S are related by

$$g(A_S(z,x), y) = g(z, \mathrm{II}_S(x,y)) \tag{2.1}$$

It follows that, for each $p \in M$ and $z \in \mathrm{N}_pS$, $A_S(z,\cdot)$ is a symmetric operator on T_pS. A proof of the following may be found in [Her1].

Lemma 2.5.1. *Let M be a complete Riemannian manifold, $S \subseteq M$ a submanifold, $v \in \mathrm{T}_pS$, and $\gamma : [a,b] \to M$ a geodesic such that $\gamma'(a) \in \mathrm{N}_pS$. Then the following are equivalent:*

(i) *For any smooth curve $\sigma : (-\varepsilon,\varepsilon) \to N$ with $\sigma'(a) = v$, there exists a variation $\Gamma : (-\varepsilon,\varepsilon)\times[a,b] \to M$ of γ through geodesics $\gamma_s(\cdot) = \Gamma(s,\cdot)$ such that $\gamma_s'(a) \in \mathrm{N}S$ and $\Gamma(s,a) = \sigma(s)$ for all s; and*

(ii) *There exists a non-trivial normal Jacobi field J along γ such that $J(a) = v$, $J(b) = 0$, and $A_S(\gamma'(a), v) + \nabla_{\gamma'} J(a) \in \mathrm{N}_pS$.*

Consequently, $\gamma(b)$ is focal to S along γ if and only if there exists a non-trivial normal Jacobi field J along γ such that $J(a) \in \mathrm{T}_pS$, $J(b) = 0$, and $A_S(\gamma'(a), J(a)) + \nabla_{\gamma'} J(a) \in \mathrm{N}_pS$.

Focal points to a submanifold consisting of a single point are called conjugate points. That is, given $x \in M$, the point z **is conjugate to** x **along** γ if $\gamma(b) = z$ and there exists a variation $\Gamma : (-\varepsilon,\varepsilon) \to M$ of γ through geodesics $\gamma_s(\cdot) = \Gamma(s,\cdot)$ such that $\gamma_s(a) = x$ for all $s \in (-\varepsilon,\varepsilon)$ and whose variation field vanishes at time b. This is equivalent z being focal to the submanifold $\{x\}$ along γ. On S^k, each point is conjugate to its antipode along any of the great circles connecting them.

If $S \subseteq M$ is a submanifold, then the **cut locus of** S is the set

$$\mathrm{cut}(S) := \{v \in \mathrm{N}S \,\big|\, \gamma_v|_{[0,1]} \text{ minimizes the distance to } S \text{ while } \gamma_v|_{[0,T]} \text{ does not for all } T > 1\}$$

and the **focal locus of** S is the set

$$\mathrm{focal}(S) := \{v \in \mathrm{N}S \,\big|\, \exp|_{\mathrm{N}S} \text{ is singular at } v\}$$

For $x \in M$, the **cut locus of** x, denoted $\mathrm{cut}(x)$, is defined to be the cut locus of the set $\{x\}$ and takes the form

$$\mathrm{cut}(x) = \{v \in \mathrm{T}_xM \,\big|\, \gamma_v|_{[0,1]} \text{ is a minimal geodesic while } \gamma_v|_{[0,T]} \text{ is not minimal for all } T > 1\}$$

Similarly, the **conjugate locus of** x, denoted $\mathrm{conj}(x)$, is the focal locus of $\{x\}$. This is the set

$$\mathrm{conj}(x) = \{v_x \in \mathrm{T}_xM \,\big|\, \exp_x \text{ is singular at } v_x\}$$

It follows directly from the definitions that $\mathrm{inj}(x) < \infty$ if and only if $\mathrm{cut}(x) \neq \emptyset$, in which case $\mathrm{inj}(x) = \mathrm{d}_{\mathrm{T}_xM}(0, \mathrm{cut}(x))$ and, moreover, there exists $v \in \mathrm{cut}(x)$ such that $\|v\| = \mathrm{inj}(x)$. The names focal locus and conjugate locus are justified by the following alternative characterizations, which date back at least to the work of Morse [Mo].

Lemma 2.5.2. *Let M be a complete Riemannian manifold. Then the following hold:*

(a) *For any submanifold $S \subseteq M$,* $\mathrm{focal}(S) = \{v \in \mathrm{N}S \,|\, \exp(v)$ *is focal to S along* $\gamma_v|_{[0,1]}\}$.

(b) *For any $x \in M$,* $\mathrm{conj}(x) = \{v_x \in \mathrm{T}_x M \,|\, \exp_x(v_x)$ *is conjugate to x along* $\gamma_{v_x}|_{[0,1]}\}$.

Proof. **(a)** Let $v \in \mathrm{focal}(S)$. Then there exist $v \in \mathrm{N}S$ and $w_v \in \mathrm{T}_v(\mathrm{N}S)$ such that $(\exp|_{\mathrm{N}S})_*(w_v) = 0$. Note that $w_v \neq 0$. Let $\alpha : (-\varepsilon, \varepsilon) \to \mathrm{N}S$ be any curve with $\alpha'(0) = w_v$. Define a map $\Gamma : (-\varepsilon, \varepsilon) \times [0,1] \to N$ by $\Gamma(s,t) := \exp(t\alpha(s))$. Then Γ is a variation of $\gamma_v|_{[0,1]}$ through geodesics $\gamma_s(\cdot) := \Gamma(s,\cdot) = \gamma_{\alpha(s)}(\cdot)$. By construction, $\gamma_s'(0) \in \mathrm{N}S$ for all $s \in (-\varepsilon, \varepsilon)$. The variation field J of Γ satisfies $J(1) = \frac{\partial}{\partial s}\big|_{s=0}(\exp \circ \alpha)(s) = (\exp|_{\mathrm{N}S})_*(w_v) = 0$. It remains to show that J is non-trivial. If $J(0) \neq 0$, the result follows. If $J(0) = 0$, then $(\pi \circ \alpha)'(0) = 0$ and, with respect to the identification $\mathrm{T}_v(\mathrm{N}S) \cong \mathrm{T}_{\pi(v)}S \times \mathrm{N}_{\pi(v)}S$, w_v must lie entirely in the second component. If $v = 0$, then $(\exp|_{\mathrm{N}S})_*(w_v) = w_{\pi(v)} \neq 0$, which is a contradiction. Hence $v \neq 0$, and one may write $\nabla_{\gamma'} J(0) = \frac{w_{\pi(v)}}{\|v\|} \neq 0$. Thus J is non-trivial. It follows that $\exp(v)$ is focal to S along $\gamma_v|_{[0,1]}$.

The same argument works in reverse. Let $v \in \mathrm{N}S$ be arbitrary, and suppose that $\exp(v)$ is focal to S along $\gamma_v|_{[0,1]}$. Let $\Gamma : (-\varepsilon, \varepsilon) \times [0,1] \to M$ be a variation of $\gamma_v|_{[0,1]}$ through geodesics $\gamma_s(\cdot) = \Gamma(s,\cdot)$ such that $\gamma_s'(0) \in \mathrm{N}S$ for all $s \in (-\varepsilon, \varepsilon)$ and whose variation field J is non-trivial and satisfies $J(1) = 0$. Define a curve $\alpha : (-\varepsilon, \varepsilon) \to \mathrm{N}S$ by $\alpha(s) := \gamma_s'(0)$. Note that $(\exp|_{\mathrm{N}S})_*(\alpha'(0)) = \frac{\partial}{\partial s}\big|_{s=0}\Gamma(s,1) = J(1) = 0$. If $(\pi \circ \alpha)'(0) \neq 0$, then $\alpha'(0) \neq 0$, and $v \in \mathrm{focal}(S)$. If $(\pi \circ \alpha)'(0) = 0$, then $J(0) = 0$. If $v = 0$, then γ_v is a constant geodesic, and J must be an affine map into $\mathrm{N}_{\pi(v)}S$. Since $J(0) = J(1) = 0$, J must be trivial, which is a contradiction. Thus $v \neq 0$ and $\alpha'(0) = \|v\| \nabla_{\gamma'} J(0) \neq 0$. This shows that $v \in \mathrm{focal}(S)$.

(b) This follows from part (a) by setting $S = \{x\}$.

□

In later chapters, I'll be concerned only with points that are focal to totally geodesic submanifolds or, more precisely, with the absence of such points.

A general relationship between inj and r_c is described by the following well-known result of Klingenberg [Kli2]. Here, and throughout this section, $\ell(x)$ denotes the length of the shortest non-trivial geodesic loop based at x.

Theorem 2.5.3. *(Klingenberg) Let M be a complete Riemannian manifold and $x \in M$. If $v \in \mathrm{cut}(x)$ has length* $\mathrm{inj}(x)$*, then one of the following holds:*

(i) $v \in \mathrm{conj}(x)$*; or*

(ii) $\gamma_v|_{[0,2]}$ *is a geodesic loop.*

Consequently, $\mathrm{inj}(x) = \min\{r_c(x), \frac{1}{2}\ell(x)\}$.

Klingenberg used Theorem 2.5.3 to characterize $\mathrm{inj}(M)$. In the following, $\ell(M) := \inf\{\ell(x) \,|\, x \in M\}$ and, when M is compact, $\ell_c(M) > 0$ is the length of the shortest non-trivial closed geodesic in M.[3]

Corollary 2.5.4. *Let M be a complete Riemannian manifold. Then each of the following holds:*

(a) $\mathrm{inj}(M) = \min\{r_c(M), \frac{1}{2}\ell(M)\}$*; and*

(b) *If M is compact, then* $\mathrm{inj}(M) = \min\{r_c(M), \frac{1}{2}\ell_c(M)\}$.

Proof. Part **(a)** follows directly from Theorem 2.5.3 by taking infimums. To prove **(b)**, suppose that M is compact, and let $x \in M$ be a point at which inj realizes its minimum. It must be the case that $0 < \mathrm{inj}(x) < \infty$. Since $\mathrm{B}(0, 2\mathrm{inj}(x)) \subset \mathrm{T}_x M$ is compact, there must be a vector $v \in \mathrm{cut}(x)$ such that $\|v\| = \mathrm{inj}(x)$. By Theorem 2.5.3, $v \in \mathrm{conj}(x)$ or $\gamma_v|_{[0,2]}$ is a geodesic loop. In the first case, $\mathrm{inj}(M) \geq r_c(M)$, which implies that $\mathrm{inj}(M) = r_c(M) = \min\{r_c(M), \frac{1}{2}\ell(M)\}$. Since $\ell(M) \leq \ell_c(M)$, $\mathrm{inj}(M) = \min\{r_c(M), \frac{1}{2}\ell_c(M)\}$. In the second case, write $z := \gamma_v(1) = \exp_x(v)$. Then $\gamma_v'(1) \in \mathrm{cut}(z)$ has length $\mathrm{inj}(M)$, which must also equal to the distance from z to $\mathrm{cut}(z)$. Again applying Theorem 2.5.3, one has either that $\gamma_v'(1) \in \mathrm{conj}(z)$, which as before implies $\mathrm{inj}(M) = \min\{r_c(M), \frac{1}{2}\ell_c(M)\}$, or that $\gamma_{\gamma_v'(1)}|_{[0,2]}$ is a geodesic loop. In the latter case, it follows that γ_v is a closed geodesic, so $\mathrm{inj}(M) \geq \frac{1}{2}\ell_c(M) \geq \frac{1}{2}\ell(M)$ and, consequently, $\mathrm{inj}(M) = \frac{1}{2}\ell_c(M) = \min\{r_c(M), \frac{1}{2}\ell_c(M)\}$.

□

It will be instructive to outline the proof of Theorem 2.5.3, as described by Cheeger–Ebin [CE]. The argument that $\mathrm{inj}(x) \geq \min\{r_c(x), \frac{1}{2}\ell(x)\}$ is by contradiction. If $\mathrm{inj}(x) < \min\{r_c(x), \frac{1}{2}\ell(x)\}$, then one may, by passing to convergent subsequences of $v_i, w_i \in \mathrm{B}(0, \mathrm{inj}(x) + \varepsilon_i) \subset \mathrm{T}_x M$, where $v_i \neq w_i$ and $\varepsilon_i \to 0$, produce vectors $v, w \in T_x M$ such that $\|v\| = \|w\| = \mathrm{inj}(x)$ and $\exp_x(v) = \exp_x(w)$. Since

[3]A well-known theorem of Fet–Lyusternik [FL] asserts that every compact manifold contains a non-trivial closed geodesic.

$\mathrm{inj}(x) < r_c(x)$, $\exp_x$ is a local diffeomorphism at v and w, which shows that $v \neq w$. Since $\mathrm{inj}(x) < \frac{1}{2}\ell(x)$, the loop $\alpha : [0,2] \to M$ defined by

$$\alpha(t) := \begin{cases} \exp_x(tv) & \text{if} \quad 0 \le t \le 1 \\ \exp_x((2-t)w) & \text{if} \quad 1 \le t \le 2 \end{cases}$$

cannot be a geodesic, so $(\exp_x)_*(v_v) \neq -(\exp_x)_*(w_w)$, where, up to the identification $\mathrm{T}_v(\mathrm{T}_x M) \cong \mathrm{T}_x M$, $v_v \in \mathrm{T}_v(\mathrm{T}_x M)$ is being identified with itself, and similarly for $w_w \in \mathrm{T}_w(\mathrm{T}_p M)$. This means it's possible to perturb v and w slightly and produce shorter vectors that are mapped to each other by $\exp_x$, which contradicts the definition of $\mathrm{inj}(x)$. The fact that $\mathrm{inj}(x) \le \frac{1}{2}\ell(x)$ follows directly from the definition of $\mathrm{inj}(x)$. It only remains to show that $\mathrm{inj}(x) \le r_c(x)$. A standard argument for this fact is that, by the Morse index theorem [Mo], a geodesic cannot minimize past its first conjugate point.

This same reasoning can be used to show that $r(x) \le r_f(x)$. The key idea is that a unit-speed geodesic initially perpendicular to a submanifold S cannot minimize distance to S beyond its first focal point. For the sake of completeness, a somewhat narrower fact is proven here. This is modeled after the presentation in [Lee3]. Let $\gamma : [a,b] \to M$ be a unit-speed geodesic, and denote by $\mathscr{V}$ the vector space of piecewise C^2 vector fields along γ. Define the **index form** $\mathrm{I} : \mathscr{V} \times \mathscr{V} \to \mathbb{R}$ by

$$\mathrm{I}(V,W) := \int_a^b \left[g(\nabla_{\gamma'} V, \nabla_{\gamma'} W) - g(R(V,\gamma')\gamma', W) \right] \mathrm{d}t$$

This is a symmetric bilinear form on $\mathscr{V}$. Fix $V \in \mathscr{V}$. Denote by $a < a_1 < a_2 < \ldots < a_k < b$ the points where V is not C^2 and by $\Delta_i \nabla_{\gamma'} V$ the change in $\nabla_{\gamma'} V$ at a_i. That is,

$$\Delta_i \nabla_{\gamma'} V := \lim_{t \searrow a_i} \nabla_{\gamma'} V(t) - \lim_{t \nearrow a_i} \nabla_{\gamma'} V(t)$$

Integration by parts yields

$$\mathrm{I}(V,W) = -\int_a^b g(\nabla_{\gamma'} \nabla_{\gamma'} V + R(V,\gamma')\gamma', W)\, \mathrm{d}t - \sum_{i=1}^k g(\Delta_i \nabla_{\gamma'} V, W(a_i)) + g(\nabla_{\gamma'} V, W)\Big|_a^b \tag{2.2}$$

for each $W \in \mathscr{V}$.

Suppose V is the variation field of a variation $\Gamma : (-\varepsilon,\varepsilon) \times [a,b] \to M$ of γ. Then V extends to a vector field along Γ by setting $V := \Gamma_*(\frac{\partial}{\partial s})$. Write $V^\perp(\cdot) := V(0,\cdot) - g(V(0,\cdot),\gamma'(\cdot))\gamma'(\cdot)$ and $\gamma_s(\cdot) := \Gamma(s,\cdot)$. By taking the second variation of length, one finds that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \mathrm{L}(\gamma_s) = \mathrm{I}(V^\perp, V^\perp) + g(\nabla_V V(0,\cdot), \gamma'(\cdot))\Big|_a^b \tag{2.3}$$

In particular, if both $\Gamma(\cdot,a)$ and $\Gamma(\cdot,b)$ are known to be geodesics, then $\nabla_V V(0,a) = 0$, $\nabla_V V(0,b) = 0$, and, consequently, $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\big|_{s=0}\mathrm{L}(\gamma_s) = \mathrm{I}(V^\perp, V^\perp)$. In this sense, the index form is akin to the Hessian of the length functional on such variations.

In the special case that J is a Jacobi field along each $\gamma|_{[a_i,a_{i+1}]}$, one has that the Jacobi equation

$$\nabla_{\gamma'}\nabla_{\gamma'}J + R(J,\gamma')\gamma' = 0 \tag{2.4}$$

holds. Substituting (2.4) into (2.2) yields

$$\mathrm{I}(J,W) = -\sum_{i=1}^{k} g(\Delta_i \nabla_{\gamma'}J, W) + g(\nabla_{\gamma'}J, W)\Big|_a^b \tag{2.5}$$

for each $W \in \mathscr{V}$.

An important fact, which will be used implicitly throughout the remainder of this dissertation, is that $\mathrm{d}^2(\cdot,p)|_{\mathrm{B}(p,\mathrm{inj}(p))}$ is C^∞. This is because $\mathrm{d}^2(x,p) = \|\exp_p^{-1}(x)\|^2$ for all $x \in \mathrm{B}(p,\mathrm{inj}(p))$. In the same way, $\mathrm{d}(\cdot,p)$ is C^∞ on $\mathrm{B}(p,\mathrm{inj}(p)) \setminus \{p\}$. Within $r_f(p)$, more can be said. The following is a bit stronger than the corresponding result in [CE], although the argument is similar.

Lemma 2.5.5. *Let M be a complete Riemannian manifold and $p \in M$. Let $R := \min\{\mathrm{inj}(p), r_f(p)\}$. Then each of the following holds:*

(a) $\nabla^2\,\mathrm{d}(\cdot,p)$ *is positive-definite on* $\mathrm{B}(p,R)\setminus\{p\}$*; and*

(b) $\nabla^2\,\mathrm{d}^2(\cdot,p)$ *is positive-definite on* $\mathrm{B}(p,R)$.

Proof. **(a)** This is equivalent to the statement that $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\big|_{s=0}\mathrm{d}(\alpha(s),p) > 0$ for any non-constant geodesic $\alpha : (-\varepsilon,\varepsilon) \to \mathrm{B}(p,R)\setminus\{p\}$. Fix such an α. Write $L := \mathrm{d}(\alpha(0),p) > 0$. Since $R \le \mathrm{inj}(p)$, there exists a unit-speed minimal geodesic $\gamma : [0,L] \to \mathrm{B}(p,R)\setminus\{p\}$ connecting p to $\alpha(0)$, and one may construct a variation $\Gamma : (-\varepsilon,\varepsilon)\times[0,L] \to \mathrm{B}(p,R)\setminus\{p\}$ of γ by setting

$$\Gamma(s,t) := \exp_p\left(s \cdot \exp_p^{-1}(\alpha(s))\right)$$

Let $J := \Gamma_*(\frac{\partial}{\partial s})$. Since Γ is a variation through geodesics, each $J_s(\cdot) := J(s,\cdot)$ is a Jacobi field along $\gamma_s(\cdot) := \Gamma(s,\cdot)$. Since $J_0(0) = 0$, one may write $J_0(t) = J_0^\perp(t) + ct\gamma'(t)$, where $c := g(\nabla_{\gamma'}J(0), \gamma'(0))$ and $g(J_0^\perp, \gamma') = 0$. Since $J_0^\perp$ satisfies (2.4), it is also a Jacobi field. Note that $J_0^\perp(0) = 0$ and $J_0^\perp(L) = \alpha'(0) - g(\gamma'(L), \alpha'(0))\gamma'(L)$. If $J_0^\perp(L) = 0$, then α and γ are, up to affine reparameterization, the same geodesic, and the conclusion follows. If $J_0^\perp(L) \neq 0$, then $\nabla_{\gamma'} J_0^\perp(0) \neq 0$. Since $R \leq r_f(p)$, it follows from the definition of r_f that $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=L} \|J_0^\perp\| = \|J_0^\perp\|'(L) > 0$. One also has that

$$\begin{aligned}
\|J_0^\perp(L)\| \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=L} \|J_0^\perp\| &= g(\nabla_{\gamma'} J_0^\perp(L), J_0^\perp(L)) \\
&= \int_0^L \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=t'} g(\nabla_{\gamma'} J_0^\perp, J_0^\perp)\,\mathrm{d}t' \\
&= \int_0^L \Big[g(\nabla_{\gamma'} J_0^\perp(t'), \nabla_{\gamma'} J_0^\perp(t')) - g\Big(R(J_0^\perp(t'), \gamma'(t'))\gamma'(t'), J_0^\perp(t')\Big)\Big]\,\mathrm{d}t' \\
&= \mathrm{I}(J_0^\perp, J_0^\perp)
\end{aligned}$$

Thus $\mathrm{I}(J_0^\perp, J_0^\perp) > 0$. At the same time, for each $s \in (-\varepsilon, \varepsilon)$, $\mathrm{d}(\alpha(s), p) = \mathrm{L}(\gamma_s)$. Since $\Gamma(\cdot, 0) = p$ and $\Gamma(\cdot, L) = \alpha(\cdot)$ are geodesics, the discussion following (2.3) shows that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \mathrm{d}(\alpha(s), p) = \mathrm{I}(J_0^\perp, J_0^\perp)$$

Combining results shows that $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\big|_{s=0} \mathrm{d}(\alpha(s), p) > 0$.

(b) Let $\alpha : (-\varepsilon, \varepsilon) \to \mathrm{B}(p, R)$ be any non-constant geodesic. If $\alpha(0) = p$, then $\mathrm{d}^2(\alpha(s), p) = \|\alpha'\| s^2$ and, consequently, $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\big|_{s=0} \mathrm{d}^2(\alpha(s), p) = 2\|\alpha'\| > 0$. If $\alpha(0) \neq p$, then one computes

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \mathrm{d}^2(\alpha(s), p) = \Big[\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \mathrm{d}(\alpha(s), p)\Big]^2 + \mathrm{d}(\alpha(0), p) \frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} \mathrm{d}(\alpha(s), p) > 0$$

The final inequality there follows from part (a).

□

Lemma 2.5.5 may be used to show that $\mathrm{d}(\cdot, p)$ and $\mathrm{d}^2(\cdot, p)$ are strictly convex on $\mathrm{B}(p, R)$. Moreover, since $r(p) \leq \min\{\mathrm{inj}(p), r_f(p)\}$, one may replace R with $r(p)$.

Corollary 2.5.6. *Let M be a complete Riemannian manifold and $p \in M$. Then each of the following holds:*

(a) $\nabla^2 \mathrm{d}(\cdot, p)$ *is positive-definite on* $\mathrm{B}(p, r(p)) \setminus \{p\}$*; and*

(b) $\nabla^2 \mathrm{d}^2(\cdot, p)$ *is positive-definite on* $\mathrm{B}(p, r(p))$.

Hereafter, the function $\mathrm{d}^2(\cdot, p)$ will be preferred over $\mathrm{d}(\cdot, p)$ because of its regularity at p. Applying Corollary 2.5.6(b) to the strong convexity radius, one obtains the following.

Corollary 2.5.7. *Let M be a complete Riemannian manifold. Then the strong convexity radius $\rho : M \to (0, \infty]$ is a continuous function with the following properties:*

(a) *For each* $z \in \mathrm{B}(x,\rho(x))$*,* $\mathrm{B}(z,\rho(x))$ *is strongly convex;*

(b) *For each* $z \in \mathrm{B}(x,\rho(x))$*,* $\nabla^2\,\mathrm{d}^2(\cdot,z)$ *is positive-definite on* $\mathrm{B}(z,\rho(x))$*; and*

(c) $\rho(x) = \infty$ *for some* $x \in M$ *if and only if* $r(M) = \infty$*.*

Remark 2.5.8. Since $\mathrm{B}(x, \frac{1}{2}\rho(x)) \subseteq \mathrm{B}(z,\rho(x))$ whenever $\mathrm{d}(z,x) \leq \frac{1}{2}\rho(q)$, Corollary 2.5.7(b) implies that each $\mathrm{d}^2(\cdot,z)|_{\mathrm{B}\left(x,\frac{1}{2}\rho(x)\right)}$ is strictly convex. This is the key property that will be used in the construction of the center of mass in Chapter 6.

Remark 2.5.9. A version of Corollary 2.5.7 holds without the assumption that M is complete. The key point in such a generalization is that the function

$$x \mapsto \sup\left\{R > 0 \,\middle|\, \exp_x \text{ is defined on all of } \mathrm{B}(0,R) \subseteq \mathrm{T}_x M\right\} > 0$$

is continuous. Such a generalization would be convenient in, say, the proof of Lemma 2.6.2, but it's not strictly necessary for the results of this dissertation and will be omitted.

Another preliminary observation will help in the proof that $r(x) \leq r_f(x)$. Given any $V \in \mathscr{V}$, one may construct a variation $\Gamma : (-\varepsilon,\varepsilon)\times[a,b] \to M$ of γ with variation field V by setting $\Gamma(s,t) := \exp_{\gamma(t)}(sV(t))$. Since $\nabla_V V(\cdot,t) = 0$ for all t, the discussion following (2.3) shows that $\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0}\mathrm{L}(\gamma_s) = \mathrm{I}(V^\perp,V^\perp)$. Whenever $g(V(a),\gamma'(a)) = g(V(b),\gamma'(b)) = 0$, the first variation formula implies that $\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathrm{L}(\gamma_s) = 0$. If $\mathrm{I}(V^\perp,V^\perp) < 0$, it follows that $\mathrm{L}(\gamma_s) < \mathrm{L}(\gamma)$ whenever $|s| \neq 0$ is sufficiently small. In particular, in order to prove that $\mathrm{L}(\gamma)$ is a strict local maximum of the distance between geodesics γ_{v_0} and γ_{v_1}, where $v_0 \in \mathrm{T}_{\gamma(a)}M$ and $v_1 \in \mathrm{T}_{\gamma(b)}M$ are perpendicular to $\gamma'(a)$ and $\gamma'(b)$, respectively, one needs only to produce a piecewise C^2 normal vector field V along γ such that $V(a) = v_0$, $V(b) = v_1$, and $\mathrm{I}(V,V) < 0$.

Lemma 2.5.10. *Let* M *be a complete Riemannian manifold and* $x \in M$*. Then* $r(x) \leq r_f(x)$*.*

Proof. Assume that $r(x) > r_f(x)$. Let J be a non-trivial normal Jacobi field along a unit-speed geodesic γ with $\gamma(0) = x$, $J(0) = 0$, and $\|J\|'(T) = 0$, where $T := r_f(x)$. Set $z := \gamma(T)$ and $v_z := -\gamma'(T)$. Note that $J(T) \neq 0$, since otherwise there would exist $0 < t < T$ such that $\|J\|'(t) = 0$, in contradiction of the fact that $T = r_f(x) < r_c(x)$. Let S be the image of $\gamma_{J(T)}|_{(-\delta_0,\delta_0)}$ for $\delta_0 > 0$ small enough that $\gamma_{J(T)}|_{(-\delta_0,\delta_0)}$ is an embedding. Then S is a smooth and embedded submanifold of M, and $v_z \in \mathrm{N}S$ is a focal point of S. Since r is continuous, one may choose $\varepsilon > 0$ to be small enough that $r(x') > r_f(x) - \varepsilon$

whenever $\mathrm{d}(x,x') < \varepsilon$. The next step is to show that there exist arbitrarily small $0 < s_0 < \varepsilon$ and $0 < s_1, s_2 < \delta_0$ such that

$$\mathrm{d}(\gamma(-s_0), \gamma_{J(T)}(-s_1)),\ \mathrm{d}(\gamma(-s_0), \gamma_{J(T)}(s_2)) < T + s_0 \tag{2.6}$$

Suppose for the moment that this is possible. Choose such s_i to be small enough that $0 < s_0 < \varepsilon$ and $\gamma_{J(T)}(-s_1), \gamma_{J(T)}(s_2) \in \mathrm{B}\big(\gamma(-s_0), r(\gamma(-s_0))\big)$. Write $x' := \gamma(-s_0)$. Then $\mathrm{d}(x,x') < \varepsilon$, so $z \in \mathrm{B}(x', r(x'))$. This shows that $\max\{\mathrm{d}(x', \gamma_{J(T)}(-s_1)), \mathrm{d}(x', \gamma_{J(T)}(s_2))\} < r(x')$, so, for any fixed R satisfying

$$\max\{\mathrm{d}(x', \gamma_{J(T)}(-s_1)), \mathrm{d}(x', \gamma_{J(T)}(s_2))\} < R < r(x')$$

one has that $\mathrm{B}(x', R)$ is strongly convex. This means that $\gamma_{J(T)}|_{[-s_1,s_2]} \subset \mathrm{B}(x', R)$, so $\mathrm{d}(x', z) = \mathrm{d}(x', \gamma_{J(T)}(0)) < R$. Letting $R \searrow \max\{\mathrm{d}(x', \gamma_{J(T)}(-s_1)), \mathrm{d}(x', \gamma_{J(T)}(s_2))\}$, one has that

$$\mathrm{d}(x', z) \leq \max\{\mathrm{d}(x', \gamma_{J(T)}(-s_1)), \mathrm{d}(x', \gamma_{J(T)}(s_2))\} < T + s_0$$

At the same time, $\mathrm{d}(x', z) = \mathrm{L}(\gamma|_{[-s_0,T]}) = T + s_0$, since $\gamma|_{[-s_0,T]}$ maps into the strongly convex ball $\mathrm{B}(x', r(x'))$ and therefore must be the unique minimal geodesic connecting x' to z. This is a contradiction, which completes the proof, modulo the existence of s_i so that (2.6) holds.

By the earlier discussion about the index form, to prove the existence of such s_i, it suffices to let $0 < s_0 < \varepsilon$ be arbitrary and produce a piecewise C^2 normal vector field V along $\gamma|_{[-s_0,T]}$ such that $V(-s_0) = 0$, $V(T) = J(T)$, and $\mathrm{I}(V,V) < 0$. Define a vector field J_0 along $\gamma|_{[-s_0,T]}$ by

$$J_0(t) := \begin{cases} 0 & \text{if} \quad t \leq 0 \\ J(t) & \text{if} \quad t \geq 0 \end{cases}$$

Since $J(0) = 0$, J_0 is continuous, and it's apparent that J_0 is smooth on $[-s_0, 0]$ and $[0, T]$. Since J is normal, so is J_0. Note that

$$\Delta\nabla_{\gamma'} J_0 = \lim_{t \searrow 0} \nabla_{\gamma'} J_0(t) - \lim_{t \nearrow 0} \nabla_{\gamma'} J_0(t) = -\lim_{t \nearrow 0} \nabla_{\gamma'} J(t) \neq 0$$

since otherwise J would be identically zero. Let W be any smooth vector field along $\gamma|_{[-s_0,T]}$ with $W(0) = \Delta\nabla_{\gamma'} J_0$, $W(-s_0) = 0$, and $W(T) = 0$; such a vector field W can be constructed in local coordinates using a bump function. Set $V_\varepsilon := J_0 + \varepsilon W$. It follows from (2.5) that $\mathrm{I}(J_0, J_0) = 0$ and $\mathrm{I}(J_0, W) = -\|\Delta\nabla_{\gamma'} J_0\|^2 < 0$. Consequently,

$$\begin{aligned}\mathrm{I}(V_\varepsilon, V_\varepsilon) &= \mathrm{I}(J_0, J_0) + 2\varepsilon \mathrm{I}(J_0, W) + \varepsilon^2 \mathrm{I}(W, W) \\ &= -2\varepsilon \|\Delta\nabla_{\gamma'} J_0\|^2 + \varepsilon^2 \mathrm{I}(W, W)\end{aligned}$$

This is negative for all sufficiently small ε.

□

It's tempting to think that, akin to the case of the injectivity radius, $r(x) = \min\{r_f(x), \frac{1}{4}\ell(x)\}$ for each $x \in M$, but it's not clear that this holds in general. The best such pointwise bound that I've obtained is $r(x) \leq \min\{r_f(x), \frac{1}{2}\ell(x)\}$, which follows from Lemma 2.5.10 and the fact that $r(x) \leq \frac{1}{2}\ell(x)$. One may also note that $r_f(x) \leq r_c(x)$. In any event, one may still obtain global equalities akin to those in Corollary 2.5.4. These are presented in Theorem 2.5.12.

Lemma 2.5.11. *Let M be a complete Riemannian manifold. Then $r_f(M) \leq \frac{1}{2}r_c(M)$.*

Proof. Fix $\varepsilon > 0$, and let $x \in M$ be such that $r_c(x) < r_c(M) + \varepsilon$. Choose a unit-speed geodesic $\gamma : [0, r_c(x)] \to M$ with $\gamma(0) = x$ and a non-trivial normal Jacobi field J along γ with $J(0) = 0$ and $J(r_c(x)) = 0$. Write $z := \gamma(r_c(x))$. There must exist $0 < T < r_c(x)$ such that $\|J\|'(T) = 0$. If $T \leq \frac{1}{2}r_c(x)$, then $r_f(x) \leq \frac{1}{2}r_c(x) < \frac{1}{2}r_c(M) + \frac{1}{2}\varepsilon$. If $T \geq \frac{1}{2}r_c(x)$, then, since $t \mapsto \gamma(r_c(x) - t)$ is a unit-speed geodesic starting at z and $t \mapsto J(r_c(x) - t)$ is a non-trivial normal Jacobi field along it with $J(0) = 0$ and $\|J\|'(r_c(x) - T) = 0$, one has $r_f(z) \leq \frac{1}{2}r_c(x) - T < \frac{1}{2}r_c(M) + \frac{1}{2}\varepsilon$. Therefore, $r_f(M) < \frac{1}{2}r_c(M) + \frac{1}{2}\varepsilon$. Since the choice of $\varepsilon > 0$ was arbitrary, $r_f(M) \leq \frac{1}{2}r_c(M)$.

□

Theorem 2.5.12. *Let M be a complete Riemannian manifold. Then the following hold:*

(a) *$r(M) = \min\{r_f(M), \frac{1}{4}\ell(M)\}$; and*

(b) *If M is compact, then $r(M) = \min\{r_f(M), \frac{1}{4}\ell_c(M)\}$.*

Proof. **(a)** Lemma 2.5.10 implies that $r(M) \leq r_f(M)$. Assume that $r(M) > \frac{1}{4}\ell(M)$, and let $\varepsilon := \frac{4}{5}[r(M) - \frac{1}{4}\ell(M)] > 0$. Note that the case $\varepsilon = \infty$ is possible if $r(M) = \infty$. Let $\gamma : [0, 1] \to M$ be a non-trivial geodesic loop with $\mathrm{L}(\gamma) < \ell(M) + \varepsilon$. Then $\frac{1}{4}\mathrm{L}(\gamma) + \varepsilon < r(M)$, so $\mathrm{B}(\gamma(\frac{1}{4}), \frac{1}{4}\mathrm{L}(\gamma) + \varepsilon)$ and $\mathrm{B}(\gamma(\frac{3}{4}), \frac{1}{4}\mathrm{L}(\gamma) + \varepsilon)$ are strongly convex. However, both $\gamma(0)$ and $\gamma(\frac{1}{2})$ are in each of those balls; since $\gamma([0, \frac{1}{2}]) \subset \mathrm{B}(\gamma(\frac{1}{4}), \frac{1}{4}\mathrm{L}(\gamma) + \varepsilon)$ and $\gamma([\frac{1}{2}, 1]) \subseteq \mathrm{B}(\gamma(\frac{3}{4}), \frac{1}{4}\mathrm{L}(\gamma) + \varepsilon)$, it follows that each of $\gamma|_{[0, \frac{1}{2}]}$ and $-\gamma|_{[\frac{1}{2}, 1]}$ is the unique minimal geodesic connecting $\gamma(0)$ to $\gamma(\frac{1}{2})$. This is a contradiction, which shows that $r(M) \leq \frac{1}{4}\ell(M)$. Thus $r(M) \leq \min\{r_f(M), \frac{1}{4}\ell(M)\}$.

Assume that $r(M) < \min\{r_f(M), \frac{1}{4}\ell(M)\}$. Choose $p \in M$ such that $r(p) < \min\{r_f(M), \frac{1}{4}\ell(M)\}$. Let $\varepsilon_i > 0$ be a sequence with $\varepsilon_i \searrow 0$ and $r(p) + \varepsilon_1 < \min\{r_f(M), \frac{1}{4}\ell(M)\}$. By Lemma 2.5.11, $r_f(M) \leq$

$\frac{1}{2}r_c(M)$. By Theorem 2.5.4(a), $\text{inj}(M) = \{r_c(M), \frac{1}{2}\ell(M)\}$. Thus $r(p)+\varepsilon_1 < \frac{1}{2}\text{inj}(M)$. According to the definition of $r(p)$, one may, by passing to a subsequence of the ε_i, without loss of generality suppose that each $\mathrm{B}(p, r(p)+\varepsilon_i)$ is not strongly convex. Thus there exist $x_i, y_i \in \mathrm{B}(p, r(p)+\varepsilon_i)$ and minimal geodesics $\gamma_i : [0,1] \to M$ from x_i to y_i such that $\gamma_i([0,1]) \not\subset \mathrm{B}(p, r(p)+\varepsilon_i)$. Fix $\delta_i > 0$ such that $\max\{\mathrm{d}(p,x_i), \mathrm{d}(p,y_i)\} < r(p)+\delta_i < r(p)+\varepsilon_i$, and fix $t_i \in (0,1)$ such that $\mathrm{d}(p,\gamma_i(t_i)) \geq r(p)+\varepsilon_i$. Let (a_i,b_i) be the connected component of $\{t \in (0,1) \,|\, \mathrm{d}(p,\gamma_i(t)) > r(p)+\delta_i\}$ containing t_i. Without loss of generality, replace x_i and y_i with $\gamma_i(a_i)$ and $\gamma_i(b_i)$, respectively, so that $x_i, y_i \in \partial\mathrm{B}(p, r(p)+\delta_i)$. Also replace γ_i with $\gamma_i|_{[a_i,b_i]}$, reparameterizing the latter so that $\gamma_i(0) = x_i$, $\gamma_i(1) = y_i$, and $\mathrm{d}(p,\gamma_i(t)) > r(p)+\delta_i$ for all $t \in (0,1)$. Since $\overline{\mathrm{B}}(p, r(p)+\varepsilon_1)$ is compact and $\mathrm{L}(\gamma_i) \leq 2[r(p)+\varepsilon_1]$ for all i, one may, by passing to a subsequence, without loss of generality suppose that $x_i \to x \in \partial\mathrm{B}(p,r(p))$, $y_i \to y \in \partial\mathrm{B}(p,r(p))$, and γ_i uniformly converges to a minimal geodesic $\gamma : [0,1] \to M$ from x to y. Note that $\mathrm{d}(p,\gamma(t)) \geq r(p)$ for all $t \in [0,1]$.

The next step is to show that $x \neq y$. Assume that $x = y$, and choose $\delta > 0$ such that $r(p)+3\delta < \min\{r_f(M), \frac{1}{4}\ell(M)\}$. As above, one also has that $r(p)+3\delta < \frac{1}{2}\text{inj}(M)$. Let i be large enough that $x_i, y_i \in \mathrm{B}(x,\delta)$. Then $\mathrm{L}(\gamma_i) = \mathrm{d}(x_i,y_i) < 2\delta$, so

$$\gamma_i([0,1]) \subset \mathrm{B}(p, r(p)+3\delta) \subset \mathrm{B}(p, r_f(p)) \cap \mathrm{B}\Big(p, \frac{1}{2}\text{inj}(p)\Big)$$

By Lemma 2.5.5(b), $\mathrm{d}^2(p,\cdot)$ is strictly convex within $\mathrm{B}(p,r_f(p)) \cap \mathrm{B}(p,\text{inj}(p))$. Since

$$\mathrm{d}(p,\gamma_i(0)), \mathrm{d}(p,\gamma_i(1)) = r(p)+\delta_i < r(p)+3\delta$$

and, by construction, γ_i is not constant, this implies that $\mathrm{d}(p,\gamma_i(t)) < r(p)+\delta_i$ for all $t \in (0,1)$. This is a contradiction. So $x \neq y$, and γ is not constant.

Since $\mathrm{d}(x,y) \leq 2r(p) < \text{inj}(M)$, γ is the unique minimal geodesic connecting x to y. Since $x, y \in \partial\mathrm{B}(p,r(p))$, it's possible to choose sequences $w_i, z_i \in \mathrm{B}(p,r(p))$ such that $w_i \to x$ and $z_i \to y$. Since $\mathrm{B}(p,r(p))$ is strongly convex, there exist unique minimal geodesics $\sigma_i : [0,1] \to M$ from w_i to z_i with $\sigma_i([0,1]) \subset \mathrm{B}(p,r(p))$. By passing to a subsequence, one may, without loss of generality, suppose that σ_i converges uniformly to γ. This implies that

$$\gamma([0,1]) \subseteq \overline{\mathrm{B}}(p,r(p)) \subset \mathrm{B}(p,r_f(p)) \cap \mathrm{B}(p,\text{inj}(p))$$

Again using the strict convexity of $\mathrm{d}^2(p,\cdot)$, along with the fact that γ is not constant, one has that $\mathrm{d}(p,\gamma(t)) < r(p)$ for all $t \in (0,1)$. This is a contradiction. So $r(M) = \min\{r_f(M), \frac{1}{4}\ell(M)\}$.

(b) For compact M, one has that

$$r(M) = \min\{r_f(M), \frac{1}{4}\ell(M)\} \leq \min\{r_f(M), \frac{1}{4}\ell_c(M)\}$$

Since $\mathrm{inj}(M) = \min\{r_c(M), \frac{1}{2}\ell_c(M)\}$, the argument in the final three paragraphs of the proof of (a) shows, essentially without modification, that $r(M) \geq \min\{r_f(M), \frac{1}{4}\ell_c(M)\}$.

□

Remark 2.5.13. Corollary 2.5.4, Lemma 2.5.11, and Theorem 2.5.12 together imply the widely known inequality

$$r(M) \leq \frac{1}{2}\mathrm{inj}(M)$$

Berger [Ber] has noted that there are no examples in the literature of compact manifolds M for which this inequality is strict. Such examples may be found using Gulliver's method of constructing manifolds with focal points but no conjugate points [Gul]. The key idea in Gulliver's construction is to raise a blister on any compact hyperbolic manifold M with sufficiently large injectivity radius, introducing enough positive curvature in a ball of fixed radius, independent of M, to create focal points within that ball but not enough to create any conjugate points. The result of Mal'cev [Ma], also sometimes attributed to Selberg [Se], that finitely generated linear groups are residually finite implies that, in each dimension $m \geq 2$, there are hyperbolic manifolds of arbitrarily large injectivity radius. It follows from a short argument using Theorem 2.5.12(b) that $\inf_M \frac{r(M)}{\mathrm{inj}(M)} = 0$ over the class of compact manifolds of any fixed dimension $m \geq 2$.

2.6 Harmonic maps

Let (M, g) and (N, h) be Riemannian manifolds and $f : M \to N$ a map. When f is C^2, its tension field τ_f is a vector field along f whose vanishing characterizes when f is harmonic. Its second fundamental form β_f assigns to each point $p \in M$ a symmetric bilinear map $\mathrm{T}_p^2 M \to \mathrm{T}_{f(p)}N$, the vanishing of which characterizes when f is totally geodesic. These were introduced by Eells–Sampson [ES], who proved these characterizations. The first purpose of this section is to give intrinsic definitions of τ_f and β_f and sketch the standard proofs of these properties. This will not be done in full rigor; the interested reader may find the details in a number of textbooks that discuss harmonic maps, for example, [J2] or [X1]. The second purpose is to prove the regularity of arbitrary totally geodesic maps.

Without any regularity assumptions, f is called **totally geodesic** if it takes geodesics in M to geodesics in N. In other words, whenever $\gamma : [a,b] \to M$ is a geodesic, $f \circ \gamma$ is also a geodesic. When f is C^2, its **second fundamental form** is defined for $x, y \in \mathrm{T}_p M$ by

$$\beta_f(x,y) := (\nabla_X^{\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)} \mathrm{d}f)(Y, \cdot)$$

for any locally defined vector fields X and Y such that $X(p) = x$ and $Y(p) = y$. This expression needs to be unpacked. The expressions $f^{-1}(\mathrm{T}N)$ and $\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)$ are shorthand for the vector bundles $\coprod_{p \in M} \mathrm{T}_{f(p)} N$ and, respectively, $\coprod_{p \in M} \mathrm{T}_p^* M \otimes \mathrm{T}_{f(p)} N$ over M. Here, $\mathrm{T}_p^* M$ denotes the dual space to $\mathrm{T}_p M$. For each $p \in M$, the linear transformation $f_*|_{\mathrm{T}_p M} : \mathrm{T}_p M \to \mathrm{T}_{f(p)} N$ may be canonically identified with a unique $\mathrm{d}_p f \in \mathrm{T}_p^* M \otimes \mathrm{T}_{f(p)} N$ such that $f_*(v) = \mathrm{d}_p f(v, \cdot)$ for all $v \in \mathrm{T}_p M$. The map $p \mapsto \mathrm{d}_p f$ is a section of $\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)$. In coordinates $(x^1, \ldots, x^m)$ for M and $(y^1, \ldots, y^n)$ for N,

$$\mathrm{d}f = \frac{\partial f_\alpha}{\partial x^i} \mathrm{d}x^i \otimes \frac{\partial}{\partial y^\alpha}$$

Here and throughout, the Einstein summation notation is used. Note that $\mathrm{T}^* M$ is canonically endowed with the pull-back metric $< \cdot, \cdot >_{\mathrm{T}^* M} := g(\sharp(\cdot), \sharp(\cdot))$ under the musical isomorphism $\sharp := \flat^{-1} : \mathrm{T}^* M \to \mathrm{T}M$, where $\flat : \mathrm{T}M \to \mathrm{T}^* M$ takes $x \in \mathrm{T}_p M$ to the covector $x_\flat(\cdot) := g(x, \cdot) \in \mathrm{T}_p^* M$. The bundle $f^{-1}(\mathrm{T}N)$ inherits the metric $< \cdot, \cdot >_{f^{-1}(\mathrm{T}N)} := h_{f \circ \pi}(\cdot, \cdot)$ from $\mathrm{T}N$. These induce a metric $< \cdot, \cdot >_{\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)}$ on $\mathrm{T}^* M \otimes f^{-1}(TN)$, which is characterized by the property that, whenever $\{\epsilon_i\}$ and $\{v_j\}$ are orthonormal bases for $\mathrm{T}_p^* M$ and $\mathrm{T}_{f(p)} N$, respectively, $\{\epsilon_i \otimes v_j\}$ is an orthonormal basis for $\mathrm{T}_p^* M \otimes \mathrm{T}_{f(p)} N$. Corresponding to these bundle metrics are Levi-Civita connections $\nabla^{\mathrm{T}^* M}$, $\nabla^{f^{-1}(\mathrm{T}N)}$, and $\nabla^{\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)}$, the last of which obeys the product rule

$$\nabla_X^{\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)}(\omega \otimes Z) = (\nabla_X^{\mathrm{T}^* M} \omega) \otimes Z + \omega \otimes (\nabla_X^{f^{-1}(\mathrm{T}N)} Z)$$

For any locally defined vector fields X_i and Y_i and smooth functions g_i and h_i, where $i = 1, 2$,

$$(\nabla_{g_1 X_1 + g_2 X_2}^{\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)} \mathrm{d}f)(h_1 Y_1 + h_2 Y_2, \cdot) = \sum_{i,j=1}^{2} g_i h_j (\nabla_{X_i}^{\mathrm{T}^* M \otimes f^{-1}(\mathrm{T}N)} \mathrm{d}f)(Y_j, \cdot)$$

It follows that β_f is well-defined. In coordinates, one has

$$\nabla_{\frac{\partial}{\partial x^i}}^{\mathrm{T}^* M} \mathrm{d}x^k = -\Gamma_{ij}^k \mathrm{d}x^j$$

$$\nabla_{\frac{\partial}{\partial x^i}}^{f^{-1}(\mathrm{T}N)} \frac{\partial}{\partial y^\alpha} = \frac{\partial f_\beta}{\partial x^i} \hat{\Gamma}_{\alpha\beta}^\gamma \frac{\partial}{\partial y^\gamma}$$

where Γ^k_{ij} and $\hat{\Gamma}^\gamma_{\alpha\beta}$ are the Christoffel symbols of M and N, respectively. Using these, one computes

$$\beta_f\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\Big) = \Big(\frac{\partial^2 f_\alpha}{\partial x^i \partial x^j} + \hat{\Gamma}^\alpha_{\beta\gamma}\frac{\partial f_\beta}{\partial x^i}\frac{\partial f_\gamma}{\partial x^j} - \Gamma^k_{ij}\frac{\partial f_\alpha}{\partial x^k}\Big)\frac{\partial}{\partial y^\alpha}$$

It follows from the symmetry of the Christoffel symbols that β_f is symmetric. Formally, β_f is identified with a section of the vector bundle $\coprod_{p\in M}\mathscr{L}(\mathrm{T}_pM \odot \mathrm{T}_pM, \mathrm{T}_{f(p)}N)$ over M, where $\mathscr{L}(\cdot,\cdot)$ is the space of linear transformations and $\odot$ the symmetric product

Roughly speaking, the second fundamental form measures how far a C^2 map is from being totally geodesic. Let $\gamma : (-\varepsilon, \varepsilon) \to M$ be a geodesic. Then $f \circ \gamma$ is a curve in N, and a direct computation shows that

$$\nabla^N_{(f\circ\gamma)'}(f\circ\gamma)' = \beta_f(\gamma', \gamma')$$

Here, only the vector components on either side are being compared. It follows that a C^2 map f is totally geodesic if and only if $\beta_f = 0$.

The **tension field** of a C^2 map f is $\tau_f := \mathrm{trace}(\beta_f)$. This notion of trace will be explained briefly. If $\omega \in \mathrm{T}^*_pM \otimes \mathrm{T}^*_pM$ is a $(0,2)$-tensor, one may construct a $(1,1)$-tensor $\omega^\sharp$ by composing with $\sharp$ in, say, the first component. That is, $\omega^\sharp(\cdot,\cdot) := \omega(\sharp(\cdot),\cdot) \in \mathrm{T}_pM \otimes \mathrm{T}^*_pM$. This is called raising the first index and defines an isomorphism $\mathrm{T}^*_pM \otimes \mathrm{T}^*_pM \to \mathrm{T}_pM \otimes \mathrm{T}^*_pM$. The trace of a simple $(1,1)$-tensor is defined to be its contraction; that is, $\mathrm{trace}(z \otimes \zeta) := \zeta(z)$. This extends to a linear map $\mathrm{T}_pM \otimes \mathrm{T}^*_pM \to \mathbb{R}$. There is a canonical isomorphism $\mathrm{T}_pM \otimes \mathrm{T}^*_pM \cong \mathscr{L}(\mathrm{T}_pM, \mathrm{T}_pM)$, under which this agrees with the usual notion of trace.[4] One may now obtain a linear map $\mathrm{T}^*_pM \otimes \mathrm{T}^*_pM \to \mathbb{R}$ by setting $\mathrm{trace}(\omega) := \mathrm{trace}(\omega^\sharp)$. With respect to any orthonormal basis $\{e_1, \ldots, e_m\}$ for T_pM,

$$\mathrm{trace}(\omega) = \sum_{i=1}^m \omega(e_i, e_i) \tag{2.7}$$

In a similar way, one may define the trace, or contraction, of an arbitrary (k,l)-tensor, where $k + l \geq 2$, along any pair of indices, which lowers the rank of the tensor by two. This is done by either raising or lowering one of the two indices, if needed, then taking the trace of the $(1,1)$-tensor obtained by fixing the entries in the other slots.

The trace of a transformation in $\mathscr{L}(\mathrm{T}_pM \odot \mathrm{T}_pM, \mathrm{T}_{f(p)}N)$ may be defined by taking the trace of its components with respect to any orthonormal basis for $\mathrm{T}_{f(p)}N$. In the specific case of β_f, given

[4]This isomorphism identifies $T : \mathrm{T}_pM \to \mathrm{T}_pM$ with the $(1,1)$-form $(\zeta, z) \mapsto \zeta(T(z))$.

any such orthonormal basis $\{v_1,\dots,v_n\}$, there exist unique $\beta^i_f \in \mathrm{T}^*_pM\otimes\mathrm{T}^*_pM$ such that $\beta_f(p) = \beta^i_f v_i$, and the tension field at p takes the form

$$\tau_f(p) = \mathrm{trace}(\beta_f)(p) = \mathrm{trace}(\beta^i_f)v_i$$

This is independent of the choice of $\{v_1,\dots,v_n\}$. Since β_f is symmetric, this also doesn't depend on the choice of index raised in the construction. Applying (2.7), one finds that

$$\tau_f(p) = \sum_{i=1}^{m}\beta_f(e_i,e_i)$$

for any orthonormal basis $\{e_1,\dots,e_m\}$ for T_pM. In local coordinates,

$$\tau_f = g^{ij}\Big(\frac{\partial^2 f_\alpha}{\partial x^i\partial x^j} + \hat{\Gamma}^\alpha_{\beta\gamma}\frac{\partial f_\beta}{\partial x^i}\frac{\partial f_\gamma}{\partial x^j} - \Gamma^k_{ij}\frac{\partial f_\alpha}{\partial x^k}\Big)\frac{\partial}{\partial y^\alpha}$$

where $[g^{ij}] = [g_{ij}]^{-1}$. In particular, in exponential normal coordinates around p and $f(p)$,

$$\tau_f(p) = \sum_{i=1}^{m}\frac{\partial^2 f_\alpha}{\partial (x^i)^2}(p)\frac{\partial}{\partial y^\alpha}(f(p))$$

From this, one sees that τ_f generalizes the Laplacian Δ_f of a C^2 function $\mathbb{R}^k\to\mathbb{R}$.

Assuming only C^1 regularity of f, one may define its **energy density** $e_f : M\to[0,\infty)$ by $e_f := \frac{1}{2}\mathrm{trace}(<\cdot,\cdot>_{f^{-1}(\mathrm{T}N)})$. It follows from (2.7) that, whenever $p\in M$ and $\{e_1,\dots,e_m\}$ is an orthonormal basis for T_pM,

$$e_f(p) = \frac{1}{2}\sum_{i=1}^{m} <e_i,e_i>_{f^{-1}(\mathrm{T}N)} = \frac{1}{2}\sum_{i=1}^{m} h(f_*(e_i),f_*(e_i))$$

Thus e_f is, indeed, non-negative. One also has that $e_f = \frac{1}{2}\|\mathrm{d}f\|^2_{\mathrm{T}^*M\otimes f^{-1}(\mathrm{T}N)}$. When M is compact, possibly with boundary, the **energy of** f is

$$\mathrm{E}(f) := \int_M e_f\,\mathrm{d}\mu_M$$

In this case, f is **harmonic** if it is a smooth critical point of this energy functional $\mathrm{E}:\mathrm{C}^1(M,N)\to[0,\infty)$. The significance of τ is that it's the negative gradient of E. That is, for any C^1 variation $F:[0,\varepsilon)\times M\to N$ of f, taking the first variation of energy shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\mathrm{E}(F(t,\cdot)) = -\int_M g(\tau_f,V)\,\mathrm{d}\mu_M$$

where V is the variation field of F at time $t=0$. It follows that a C^2 map f is harmonic if and only if $\tau_f = 0$. Combining results, one obtains the characterizations of totally geodesic and, respectively, harmonic maps proved in [ES].

Theorem 2.6.1. *(Eells–Sampson) Let M and N be Riemannian manifolds and $f : M \to N$ a C^2 map. Then the following hold:*

(a) *f is totally geodesic if and only if $\beta_f = 0$; and*

(b) *f is harmonic if and only if $\tau_f = 0$.*

In particular, Theorem 2.6.1(b) implies that every C^2 map satisfying $\tau_f = 0$ is smooth. Since $\tau_f = \mathrm{trace}(\beta_f)$, every C^2 totally geodesic map must be harmonic. This puts Theorem 1.2 into perspective. It's also worth noting that Eells–Sampson proved that any C^1 local minimum of E is harmonic.

This section ends with the observation that totally geodesic maps are smooth, a small point which I haven't been able to locate elsewhere in the literature. The proof is elementary.

Lemma 2.6.2. *Let M and N be Riemannian manifolds. If $f : M \to N$ is continuous and totally geodesic, then f is smooth.*

Proof. It suffices to argue locally. Fix $p \in M$. Choose $0 < R < \infty$ to be small enough that $\exp_p$ is defined on all of $\overline{\mathrm{B}}(0, 2R) \subseteq \mathrm{T}_p M$. Since the sectional curvature on $\overline{B}(p, 2R)$ is bounded above, one may, if necessary, shrink R so that, for each $x \in \mathrm{B}(p, R)$, $\exp_x$ is injective on $\mathrm{B}(0, R) \subseteq \mathrm{T}_x M$. This follows from the Jacobi field comparison arguments discussed in the previous section. In the same way, one may choose $0 < R' < \infty$ such that $\exp_{f(p)}$ is defined on all of $\overline{\mathrm{B}}(0, 2R')$ and, for each $y \in \mathrm{B}(f(p), R')$, $\exp_y$ is injective on $\mathrm{B}(0, R') \subseteq \mathrm{T}_y N$.[5] Shrinking R once more, if necessary, one may suppose that $f(\mathrm{B}(p, R)) \subseteq \mathrm{B}(f(p), R')$. The essential step is to show that f is C^1 on $\mathrm{B}(p, R)$. Suppose for the moment that this is the case. Then the derivative $\mathrm{D}_p f : \mathrm{T}_p M \to \mathrm{T}_{f(p)} N$ exists and, in particular, is linear. Since f is totally geodesic, $f|_{\mathrm{B}(p,R)} = \exp_{f(p)} \circ \mathrm{D}_p f \circ \exp_p^{-1}|_{\mathrm{B}(p,R)}$, and it follows that $f|_{\mathrm{B}(p,R)}$ is smooth. This completes the proof, modulo $f|_{\mathrm{B}(p,R)}$ being C^1.

The fact that f is totally geodesic implies that, for each $w_z \in \mathrm{TB}(p, R)$, $f_*(w_z)$ exists. It also implies that $f \circ \gamma_{w_z} = \gamma_{f_*(w_z)}$ whenever these are defined. It will be shown that $f_*|_{\mathrm{SB}(p,R)}$ is continuous with respect to the usual topologies on $\mathrm{S}M$ and $\mathrm{S}N$. Assume not, and let $v_x, v^k_{x_k} \in \mathrm{SB}(p, R)$ be such that $v^k_{x_k} \to v_x$ but $f_*(v^k_{x_k}) \not\to f_*(v_x)$. One may suppose, without loss of generality, that no subsequence of $f_*(v^k_{x_k})$ converges to $f_*(v_x)$. Two special cases will be considered:

(i) $\|f_*(v^k_{x_k})\| \geq c$ for some $0 < c \leq 1$ and all $k \in \mathbb{N}$; and

[5]If M is complete, one may choose any $0 < R < \frac{1}{2}\rho_M(p)$ and $0 < R' < \frac{1}{2}\rho_N(f(p))$.

(ii) $\|f_*(v^k_{x_k})\| \to 0$.

Eliminating the possibility of cases (i) and (ii) shows that f_* is continuous, since, if (i) fails to hold, then one may pass to a subsequence for which (ii) holds.

Suppose that (i) holds. For all k large enough that $x_k \in \mathrm{B}(x,R)$ and, consequently, $f(x_k) \in \mathrm{B}(f(x),R')$, $f \circ \gamma_{v^k_{x_k}}$ is defined on at least the interval $[0, \min\left\{R, \frac{R'}{\|f_*(v^k_{x_k})\|}\right\}] \subseteq [0, \min\left\{R, \frac{R'}{c}\right\}]$. Let $R'' := c \cdot \min\{R, R'\} > 0$. Since f is totally geodesic, there exist $0 < t_k \leq \min\left\{R, \frac{R'}{c}\right\}$ such that

$$\mathrm{d}_N(f \circ \gamma_{v^k_{x_k}}(t_k), f(x_k)) = \mathrm{d}_N(\gamma_{f_*(v^k_{x_k})}(t_k), \gamma_{f_*(v^k_{x_k})}(0)) = R''$$

for all such k. By passing to a subsequence, one may without loss of generality suppose that $t_k \to T$. By the continuity of f and the exponential map,

$$\mathrm{d}_N(f \circ \gamma_{v^k_{x_k}}(t_k), f(x_k)) \to \mathrm{d}_N(f \circ \gamma_{v_x}(T), f(x))$$

It follows that $\mathrm{d}_N(f \circ \gamma_{v_x}(T), f(x)) = R''$. This implies that $T \neq 0$ and that $\|f_*(v^k_{x_k})\| \to \|f_*(v_x)\| = \frac{R''}{T}$. By passing to another subsequence, one may suppose without loss of generality that, for some $w \in \mathrm{T}_{f(x)}N$ such that $w \neq f_*(v_x)$, $f_*(v^k_{x_k}) \to w$. Since $0 < T\|w\| = T\|f_*(v_x)\| \leq R'$, $\exp_{f(p)}$ is injective on $\mathrm{B}(f(p), R')$, from which it follows that

$$\mathrm{d}_N(\gamma_w(T), \gamma_{f_*(v_x)}(T)) \neq 0$$

At the same time,

$$\begin{aligned}\mathrm{d}_N(\gamma_w(T), \gamma_{f_*(v_x)}(T)) &= \lim_{k\to\infty} \mathrm{d}_N(\gamma_{f_*(v^k_{x_k})}(T), \gamma_{f_*(v_x)}(T)) \\ &= \lim_{k\to\infty} \mathrm{d}_N(f \circ \gamma_{v^k_{x_k}}(T), f \circ \gamma_{v_x}(T)) \\ &= 0\end{aligned}$$

This is a contradiction.

Suppose that (ii) holds. Since $f_*(v^k_{x_k}) \nrightarrow f_*(v_x)$ and $f(x_k) \to f(x)$, one has that $f_*(v_x) \neq 0$. Let $T := \min\left\{R, \frac{R'}{\|f_*(v_x)\|}\right\}$. For all k large enough that $x_k \in \mathrm{B}(x,R)$ and $\|f_*(v^k_{x_k})\| < \frac{R'}{R}$, $f(x_k) \in \mathrm{B}(f(x), R')$ and $\gamma_{f_*(v^k_{x_k})}(T)$ is defined. Since $\|f_*(v^k_{x_k})\| \to 0$, $f \circ \gamma_{v^k_{x_k}}(T) = \gamma_{f_*(v^k_{x_k})}(T) \to f(x)$. At the same time, $f \circ \gamma_{v^k_{x_k}}(T) \to f \circ \gamma_{v_x}(T)$. Thus $f \circ \gamma_{v_x}(T) = f(x)$. However, since $T\|f_*(v_x)\| < R'$ and $f_*(v_x) \neq 0$, $f \circ \gamma_{v_x}(T) = \gamma_{f_*(v_x)}(T) \neq f(x)$. This is a contradiction. It follows that f_* is continuous and, consequently, f is C^1.

$\square$

Remark 2.6.3. It's reasonable to think that the assumption of continuity can be dropped from Lemma 2.6.2. However, the argument required would no doubt be more tedious than is worth attempting here.

It follows from Lemma 2.6.2 and Theorem 2.6.1 that a map f is totally geodesic if and only if it is smooth and satisfies $\beta_f = 0$.

2.7 Beta and gamma functions

A few standard results about the beta and gamma functions will be needed in the discussion of length and intersection in Chapter 7. These have been treated elegantly in many textbooks over the years; for example, the basic properties mentioned here may be found in [AAR]. In keeping with usual practice, these functions will be defined on domains in the complex plane, but in the application they will only be needed on the real numbers. Let $\mathbb{C}_+ := \{z \in \mathbb{C} \,\big|\, \mathrm{Re}(z) > 0\}$, where $\mathrm{Re}(z)$ denotes the real part of z. The **beta function** $B : \mathbb{C}_+ \times \mathbb{C}_+ \to \mathbb{C}$ is defined by

$$B(x, y) := \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t$$

The **gamma function** $\Gamma : \mathbb{C}_+ \to \mathbb{C}$ is defined by

$$\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t$$

The improper integral in the definition of Γ converges on $\mathbb{C}_+$, so Γ is well-defined. By analytic continuation, Γ extends to a meromorphic function on $\mathbb{C}$ with poles at the non-positive integers. These functions, which were introduced by Euler, have many beautiful properties. For example, $\Gamma(z+1) = z\Gamma(z)$, which since $\Gamma(1) = 1$ implies that $\Gamma(n) = (n-1)!$ for all $n \in \mathbb{N}$. That is, $z \mapsto \Gamma(z+1)$ is a continuous extension of the factorial. Another classical result is that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. A few other well-known facts are worth recording here.

Lemma 2.7.1. *Each of the following holds:*

(a) $B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ *for all* $x, y \in \mathbb{C}_+$*;*

(b) $c_{n-1} = \dfrac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$ *for all* $n \in \mathbb{N}$*; and*

(c) $B(\frac{k}{2}, \frac{l}{2}) = \frac{2c_{k+l-1}}{c_{k-1}c_{l-1}}$ *for all* $k, l \in \mathbb{N}$.

Note that Lemma 2.7.1(c) follows immediately from parts (a) and (b). It will also help to record the following bound on the ratio of two gamma functions due to Gurland [Gur].

Theorem 2.7.2. *(Gurland) For any $n \in \mathbb{N}$, $\dfrac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} < \dfrac{n}{\sqrt{2n+1}}$.*

There is an extensive literature on bounding the ratio of two gamma functions. A fairly comprehensive survey article on the subject is by Qi [Q]. It's explained there how Theorem 2.7.2 both generalizes and improves upon the two-sided bound

$$\frac{1}{\sqrt{\pi(n+\frac{1}{2})}} < \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots(2n)} < \frac{1}{\sqrt{\pi n}}$$

due to Wallis in the seventeenth century and the later improvement

$$\frac{1}{\sqrt{\pi(n+\frac{1}{2})}} < \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots(2n)} < \frac{1}{\sqrt{\pi(n+\frac{1}{4})}}$$

by Kazarinoff [Ka]. The quantity being bounded in those is often called the Wallis ratio. Gurland's result, a relatively early one in the area, can be used to establish a bound involving the beta function.

Corollary 2.7.3. *Let $k, l \in \mathbb{N}$. Then $B(\frac{k+1}{2}, \frac{l}{2}) < \sqrt{\frac{k}{k+l}}B(\frac{k}{2}, \frac{l}{2})$.*

Proof. By Lemma 2.7.1(a), $B(\frac{k+1}{2}, \frac{l}{2}) = \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{l}{2})}{\Gamma(\frac{k+l+1}{2})}$ and $B(\frac{k}{2}, \frac{l}{2}) = \frac{\Gamma(\frac{k}{2})\Gamma(\frac{l}{2})}{\Gamma(\frac{k+l}{2})} = \frac{k+l}{2}\frac{\Gamma(\frac{k}{2})\Gamma(\frac{l}{2})}{\Gamma(\frac{k+l+2}{2})}$. Therefore, the desired inequality is equivalent to

$$\frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}\frac{\Gamma(\frac{k+l+2}{2})}{\Gamma(\frac{k+l+1}{2})} < \frac{1}{2}\sqrt{k(k+l)} \tag{2.8}$$

By Theorem 2.7.2,

$$\frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}\frac{\Gamma(\frac{k+l+2}{2})}{\Gamma(\frac{k+l+1}{2})} < \frac{k}{\sqrt{2k+1}}\frac{k+l+1}{\sqrt{2(k+l+1)+1}}$$

Hence (2.8) can be established by showing that

$$\frac{k}{\sqrt{2k+1}}\frac{k+l+1}{\sqrt{2(k+l+1)+1}} < \frac{1}{2}\sqrt{k(k+l)}$$

Elementary algebraic manipulations show this to be equivalent to

$$l^2 + \frac{3}{2}l + k(l - \frac{1}{2}) > 0$$

This holds since $k, l \geq 1$.

$\square$

Remark 2.7.4. The more general inequality $B(x+\frac{1}{2},y) < \sqrt{\frac{x}{x+y}}B(x,y)$ holds for all $x,y>0$. This is a small point that, to the best of my knowledge, may have gone unremarked upon in the literature. A well-known result of Bustoz–Ismail [BusI] is that the function

$$x \mapsto \sqrt{x}\frac{\Gamma(x)}{\Gamma(x+\frac{1}{2})}$$

is decreasing on $(0,\infty)$, a fact which by elementary manipulations is seen to be equivalent to the inequality for the beta function. It's interesting to note the similarity this bears to the well-known equality $B(x+1,y) = \frac{x}{x+y}B(x,y)$.

It turns out that Corollary 2.7.3 is equivalent to an inequality relating the volumes of spheres.

Corollary 2.7.5. *Suppose $k,n \in \mathbb{N}$ satisfy $n \geq k$. Then each of the following holds:*

(a) $nc_n^2c_{k-1}^2 \leq kc_{n-1}^2c_k^2$*; and*

(b) $nc_n^2c_{k-1}^2 = kc_{n-1}^2c_k^2$ *if and only if $n = k$.*

Proof. If $n = k$, then the equality in (b) is clear. The proof of (b) will be completed by showing that, whenever $n > k$,

$$nc_n^2c_{k-1}^2 < kc_{n-1}^2c_k^2$$

This will also prove (a). In that case, write $l = n-k > 0$. Then the desired inequality is equivalent to

$$\frac{4c_{k+l}^2}{c_k^2c_{l-1}^2} < \frac{k}{k+l}\frac{4c_{k+l-1}^2}{c_{k-1}^2c_{l-1}^2}$$

By Lemma 2.7.1(c), this is equivalent to

$$B^2\Big(\frac{k+1}{2},\frac{l}{2}\Big) < \frac{k}{k+l}B^2\Big(\frac{k}{2},\frac{l}{2}\Big)$$

This is equivalent to the inequality in Corollary 2.7.3.

□

Corollary 2.7.5 will be used in the proof of the final inequality in Theorem 7.4.4. It's worth pointing out that it, and not Theorem 2.7.2 itself, is all that's needed for this purpose. Since it will later be shown that Corollary 2.7.5, and consequently Corollary 2.7.3, may be derived from other results in Chapter 7, Gurland's result is not, strictly speaking, necessary for this dissertation. Presenting it here merely helps simplify the exposition and put Corollary 2.7.3 into context.

Chapter 3

Manifolds with no conjugate points

3.1 Topology and geometry

Recall that, by definition, a Riemannian manifold N has no conjugate points if the exponential map at each point is non-singular. This is equivalent to the universal covering space $\overline{N}$ having no conjugate points. When N is complete, this is also equivalent, by Theorem 1.1(c), to the condition that $r_c(N) = \infty$ and, similarly, that $r_c(\overline{N}) = \infty$. There are a few other widely known equivalences.

Theorem 3.1.1. *Let N be a complete and simply connected Riemannian manifold. Then the following are equivalent:*

(i) *N has no conjugate points;*

(ii) *For each $p \in N$, $\exp_p : \mathrm{T}_p N \to N$ is a diffeomorphism;*

(iii) $\mathrm{inj}(N) = \infty$*; and*

(iv) $r_c(N) = \infty$.

Proof. **(i)** $\Rightarrow$ **(ii)** Since N is complete and has no conjugate points, $\exp_p : \mathrm{T}_p N \to N$ is well-defined and a local diffeomorphism. When $\mathrm{T}_p N$ is endowed with the pull-back metric from $\exp_p$, $\exp_p$ is a local isometry. If $v^k \in \mathrm{T}_p N$ is a Cauchy sequence with respect to this metric, then there exists $R > 0$ such that $\mathrm{d}_{\mathrm{T}_p N}(0, v^k) \leq R$ for all k. By the definition of $\exp_p$, this implies that $v^k \in \overline{\mathrm{B}}(0, R)$ for all k. By compactness, v^k contains a convergent subsequence and, consequently, converges itself. Thus $\mathrm{T}_p N$ is complete. It follows from Lemma 2.2.2 that $\exp_p$ is a diffeomorphism.

(ii) $\Rightarrow$ **(iii)** This follows from the definition of $\mathrm{inj}(N)$.

(iii) $\Rightarrow$ **(iv)** This follows from Corollary 2.5.4.

(iv) $\Rightarrow$ **(i)** This follows from Lemma 2.5.2(b).

□

Thus a complete N has no conjugate points if and only if $\overline{N}$ satisfies any of the conditions (i)-(iv). Using the classical theorem of Hopf–Rinow [HR], (ii) may be reformulated as a synthetic condition, one which dates back to the first postulate in Euclid's *Elements*. Namely, N has no conjugate points if and only if, given any $\overline{p}, \overline{q} \in \overline{N}$, there exists a unique unit-speed geodesic from $\overline{p}$ to $\overline{q}$. This geodesic must necessarily be minimal. Even without the assumption that N is complete, this minimality requirement on $\overline{N}$ implies that N has no conjugate points.

Lemma 3.1.2. *Let N be a Riemannian manifold. If all geodesics in $\overline{N}$ are minimal, then N has no conjugate points.*

Proof. Let $\overline{p} \in \overline{N}$. Assume that $(\exp_{\overline{p}})_*(w_{v_{\overline{p}}}) = 0$ for some $v_{\overline{p}} \in \mathrm{T}_{\overline{p}}\overline{N}$ and $w_{v_{\overline{p}}} \in \mathrm{T}_{v_{\overline{p}}}(\mathrm{T}_{\overline{p}}\overline{N})$. As in the proof of Lemma 2.5.2, $v_{\overline{p}} \neq 0$ and, moreover, the Gauss lemma implies that $w_{v_{\overline{p}}} = w^{\perp}_{v_{\overline{p}}}$. Let $\alpha : (-\varepsilon, \varepsilon) \to N$ be the arc of the great circle on $\mathrm{S}_{\overline{p}}(\|v_{\overline{p}}\|)$ with $\alpha'(0) = w_{v_{\overline{p}}}$. The variation field J of the variation $\Gamma : (-\varepsilon, \varepsilon) \to \overline{N}$ defined by $\Gamma(s,t) := \exp(t\alpha(s))$ is a non-trivial normal Jacobi field along the unit-speed geodesic $\gamma(\cdot) := \Gamma(0, \cdot)$ satisfying $J(0) = 0$ and $J(\|v_{\overline{p}}\|) = 0$. By construction, $\Gamma(s, 0) = \overline{p}$ for all $s \in (-\varepsilon, \varepsilon)$. Thus $\gamma(0)$ and $\gamma(\|v_{\overline{p}}\|)$ are conjugate along γ. As discussed after Theorem 2.5.3, the Morse index theorem implies that a geodesic cannot minimize past its first conjugate point. This is a contradiction. It follows that $\overline{N}$ has no conjugate points, and, consequently, so does N.

□

From the synthetic characterization, one sees that a complete N has no conjugate points if and only if each basepoint-fixed homotopy class of paths $[\alpha] \in \mathscr{P}(p,q)$ contains a unique geodesic. This forces the set of closed geodesics in each free homotopy class to have many nice properties. The geometric structure of this set is related in fundamental ways to the algebraic structure of $\pi_1(N, y)$, and it is strongly affected by the properties of Busemann functions and the asymptotic behavior of geodesics in $\overline{N}$. The discussion here will mostly focus on how the topology of N directly affects its geometry, especially the way it affects the set of geodesic loops, and not on more analytical issues, even though the latter questions have a rich history and deep import. A key observation, which dates back to Busemann [Buse], is that any two freely homotopic closed geodesics must have the same length.

Lemma 3.1.3. *(Busemann) Let N be a complete Riemannian manifold with no conjugate points and*

$\gamma_1, \gamma_2 : [0,1] \to N$ *closed geodesics. If there exists a homotopy* $H : [a,b] \times [0,1] \to N$ *from* γ_1 *to* γ_2, *then* $\mathrm{L}(\gamma_1) = \mathrm{L}(\gamma_2)$.

Proof. Assume that $\mathrm{L}(\gamma_1) \neq \mathrm{L}(\gamma_2)$. Without loss of generality, let γ_1 be the longer of the two, so that there exists $\varepsilon > 0$ such that $\mathrm{L}(\gamma_1) = \mathrm{L}(\gamma_2) + \varepsilon$. Let $\alpha : [a,b] \to N$ be the curve defined by $\alpha(s) := H(s,0)$, and write $\ell := \mathrm{L}(\alpha)$. By slightly perturbing H, one may, without loss of generality, suppose that $\ell < \infty$. Let $n \in \mathbb{N}$ satisfy $n > \frac{2\ell}{\varepsilon}$. Let $H^n : [a,b] \times [0,n] \to N$ be defined by $H^n(s,t) := H(s, t - \lceil t \rceil)$, where $\lceil t \rceil$ denotes the integer part of t. In other words, for each $s \in [a,b]$, $H^n(s,\cdot)$ is periodic and iterates $H(s,\cdot)$ n times. This map H^n is continuous, and $H^n(a,\cdot)$ and $H^n(b,\cdot)$ are closed geodesics of length $n\mathrm{L}(\gamma_1)$ and $n\mathrm{L}(\gamma_2)$, respectively. If $\overline{H}^n : [a,b] \times [0,k] \to \overline{N}$ is any lift of H^n, then $\mathrm{L}(\overline{H}^n(a,\cdot)) = n\mathrm{L}(\gamma_1) > n\mathrm{L}(\gamma_2) + 2\ell$ and $\mathrm{L}(\overline{H}^n(b,\cdot)) = n\mathrm{L}(\gamma_2)$. Thus

$$\begin{aligned} \mathrm{d}(\overline{H}^n(a,0), \overline{H}^n(a,n)) &\leq \mathrm{d}(\overline{H}^n(a,0), \overline{H}^n(b,0)) + \mathrm{d}(\overline{H}^n(b,0), \overline{H}^n(b,n)) + \mathrm{d}(\overline{H}^n(b,n), \overline{H}^n(b,0)) \\ &\leq n\mathrm{L}(\gamma_2) + 2\ell \end{aligned}$$

This contradicts the fact that $\overline{H}^n(a,\cdot)$ is the minimal geodesic connecting $\overline{H}^n(a,0)$ and $\overline{H}^n(a,n)$. Therefore, $\mathrm{L}(\gamma_1) = \mathrm{L}(\gamma_2)$.

□

The proof of Lemma 3.1.3 demonstrates a general approach in this area, namely, to lift to $\overline{N}$ and use the synthetic characterization. Before continuing in that vein, it will help to establish a few results from general topology. If N is complete, then, by Theorem 3.1.1, $\exp_{\overline{p}} : \mathrm{T}_{\overline{p}}\overline{N} \to \overline{N}$ is a diffeomorphism and, consequently, $\overline{N} \cong \mathbb{R}^{\dim(N)}$. It follows that $\overline{N}$ is contractible. By the homotopy lifting property, this is equivalent to N being **aspherical**, which by definition means that $\pi_k(N,p) = 0$ for all $k \geq 2$ and $p \in N$. Thus N is an Eilenberg–Mac Lane space of the form $\mathrm{K}(\pi_1(N,q),1)$, determined up to homotopy equivalence by its fundamental group [EM]. It follows that any map $f : M \to N$ with $f_*(\pi_1(M,p)) = 0$ is homotopic to a constant map, a result which gives the first hint that maps into N are determined to a great extent by what they do at the level of fundamental group.

Proposition 3.1.4. *Let* M *and* N *be Riemannian manifolds, where* N *has no conjugate points, and let* $f : M \to N$ *be a continuous function. If* $f_*(\pi_1(M,p)) = \langle e \rangle$ *for any* $p \in M$, *then* f *is homotopic to a constant map.*

Proof. Let $\overline{q} \in \pi^{-1}(q) \subseteq \overline{N}$. Since $f_*(\pi_1(M,p)) = <e> \leq \pi_1(N,q)$, f lifts to a map $\overline{f} : M \to \overline{N}$ with $\overline{f}(p) = \overline{q}$. Since $\overline{N}$ is contractible, it follows that $\overline{f}$ is homotopic to a constant map, and any such homotopy descends to N.

□

This program will be further developed in this chapter, culminating in the result that maps $T^n \to N$ are, up to homotopy, determined by what they do at the level of fundamental group. This isn't a novel conclusion, since it holds for any $\mathrm{K}(\pi_1(N,q),1)$ space, but the non-conjugacy hypothesis enables a constructive method that will be useful in establishing Theorem 1.4(c). A well-known result is that the fundamental group of any aspherical manifold is torsion-free or, in other words, contains no elements of finite order. The original proof of this fact used a fixed-point theorem of Smith [Sm1], in an application that, as far as I can tell, was first recorded by Hurewicz [Hu].

Theorem 3.1.5. *(Hurewicz–Smith) Let N be an aspherical manifold and $y \in N$. Then $\pi_1(N,y)$ is torsion-free.*

Remark 3.1.6. Theorem 3.1.5 implies that, when N is a complete Riemannian manifold with no conjugate points, each $\pi_1(N,y)$ is torsion-free. Under the stronger assumption that N is compact, this follows immediately from the classical result that each free homotopy class of loops contains a closed geodesic.

More specifically, Smith's theorem in [Sm1] implies that, when N is an aspherical manifold, every periodic transformation $N \to N$ of prime order p has a fixed point. If $\pi_1(N,y)$ were to contain an element $\gamma \neq e$ of finite order, then there would exist $k \geq 1$ such that γ^k had prime order, and the corresponding deck transformation $\overline{N} \to \overline{N}$ would have no fixed points, a contradiction. This is essentially the argument given in [Hu], with the caveat that Smith's result only holds for transformations $K \to K$, where $K \subseteq \mathbb{R}^n$. Since Davis [D] constructed compact aspherical manifolds in any dimension $n \geq 4$ whose universal covers are not homeomorphic to $\mathbb{R}^n$, one must apply something like the well-known Whitney embedding theorem for this to work. A later generalization by Smith [Sm2] of his result to cell complexes makes this last step unnecessary. By contrast, the next proof uses the minimality of geodesics in $\overline{N}$ in an essential way.

Lemma 3.1.7. *Let M and N be Riemannian manifolds, where M is complete and has finite volume* *and all geodesics in $\overline{N}$ are minimal. If $f : M \to N$ is a totally geodesic map with the property that $f_*(\pi_1(M,x)) =< e >$ for any $x \in M$, then f is constant.*

Proof. The argument is by contradiction. By Lemma 2.6.2, f is smooth. Assume that, for some $v_x \in \mathrm{T}M$, $f_*(v_x) \neq 0$. Then there exists an open set $U \subset \mathrm{T}M$ containing v_x such that $\overline{U}$ is compact and $f_*(w_z) \neq 0$ for all $w_z \in \overline{U}$. Note that the latter condition implies that $\overline{U}$ is disjoint from the zero-section in $\mathrm{T}M$. Let $\varepsilon > 0$, $m := \max_{w_z \in \overline{U}} \|\mathrm{D}_z f\|$, $V := U \cap \mathrm{TB}(x, \frac{\varepsilon}{m})$, and $c := \min_{w_z \in \overline{V}} \|f_*(w_z)\| > 0$. By the Poincaré recurrence theorem, there exists $T > \frac{2\varepsilon}{c}$ such that $\Psi^T(V) \cap V \neq \emptyset$. Let $w_z \in V$ be such that $\Psi^T(w_z) = \gamma'_{w_z}(T) \in V$. Let $\alpha : [0,1] \to M$ be a minimal geodesic from x to z and $\beta : [0,1] \to M$ a minimal geodesic from x to $\gamma_{w_z}(T)$, so that $\mathrm{L}(\alpha), \mathrm{L}(\beta) \leq \frac{\varepsilon}{m}$. Write $y := f(x)$, and choose $\overline{y} \in \pi^{-1}(y)$. Since $f_*(\pi_1(M,x)) =< e >$, f lifts to a map $\overline{f} : M \to \overline{N}$ satisfying $\overline{f}(x) = \overline{y}$ and $f = \pi \circ \overline{f}$. Since the concatenation $\sigma := \alpha \cdot \gamma_{w_z}|_{[0,T]} \cdot \beta^{-1}$ is a loop based at x, $\overline{f} \circ \sigma = (\overline{f} \circ \alpha) \cdot (\overline{f} \circ \gamma_{w_z}|_{[0,T]}) \cdot (\overline{f} \circ \beta^{-1})$ is a loop based at $\overline{y}$. Since $\overline{f}$ is totally geodesic, each of $\overline{f} \circ \gamma_{w_z}$, $\overline{f} \circ \alpha$, and $\overline{f} \circ \beta^{-1}$ is a geodesic. By assumption, they are all minimal. But $\mathrm{L}(\overline{f} \circ \gamma_{w_z}|_{[0,T]}) = T|f_*(w_z)| > 2\varepsilon$ and, since m is a Lipschitz constant for $\overline{f}|_{\pi(V)}$, $\mathrm{L}(\overline{f} \circ \alpha), \mathrm{L}(\overline{f} \circ \beta) < \varepsilon$. This contradicts the triangle inequality. Therefore, $f_* = 0$ on $\mathrm{T}M$ and, consequently, f is constant.

□

Remark 3.1.8. The requirement that M have finite volume cannot be dropped from Lemma 3.1.7. For example, if N is any complete Riemannian manifold with no conjugate points, then the Riemannian universal covering map $\pi : \overline{N} \to N$ is totally geodesic and satisfies $\pi_*(\pi_1(\overline{N}, \overline{y})) =< e >$ for all $\overline{y} \in \overline{N}$.

Lemma 3.1.9. *Let M be a connected Riemannian manifold, N a complete Riemannian manifold with no conjugate points, and $H_1, H_2 : M \times [0,1] \to N$ continuous functions. Suppose*

$$H_1(x,0) = H_1(x,1) = H_2(x,0) = H_2(x,1)$$

for all $x \in M$ and that, for some $p \in M$, $[H_1(p,\cdot)] = [H_2(p,\cdot)] \in \pi_1(N, H_i(p,0))$. Then there exists a homotopy $F : [0,1] \times M \times [0,1] \to N$ from H_1 to H_2 such that

$$F(s,x,0) = F(s,x,1) = H_i(x,0) = H_i(x,1)$$

for each $x \in M$, $s \in [0,1]$, and $i = 1,2$.

Proof. Since M is connected, a standard argument shows that, for each $x \in M$, $[H_1(x,\cdot)] = [H_2(x,\cdot)] \in \pi_1(N, H_i(x,0))$. The map F will be characterized by the following property:

(*) If $\overline{H}_{i,x}$, $i = 1,2$, are lifts of the $H_i(x,\cdot)$ to $\overline{N}$ with $\overline{H}_{1,x}(0) = \overline{H}_{2,x}(0)$, then $F(s,x,t) = \pi \circ \overline{F}_x(s,t)$, where $\overline{F}_x : [0,1] \times [0,1] \to \overline{N}$ is defined by $\overline{F}_x(s,t) := \exp_{\overline{H}_{1,x}(t)}\left(s \cdot \exp^{-1}_{\overline{H}_{1,x}(t)}(\overline{H}_{2,x}(t))\right)$.

First note that, for any choice of lifts $\overline{H}_{i,x}$ as above, each $\overline{F}_x$ is well-defined since N has no conjugate points. Since $\overline{H}_{1,x}(1) = \overline{H}_{2,x}(1) \in \pi^{-1}(H_i(x,0))$, $\pi \circ \overline{F}_x(s,0) = \pi \circ \overline{F}_x(s,1) = H_i(x,0)$ for all $s \in [0,1]$ and $i = 1,2$. To see that the above yields a well-defined function F, it must be shown that the expression $\pi \circ \overline{F}_x(s,t)$ does not depend on the choice of lifts $\overline{H}_{i,x}$. This is equivalent to the statement that, for each deck transformation $[\alpha] \in \Gamma$,

$$[\alpha] \cdot \exp_{\overline{H}_{1,x}(t)}\left(s \cdot \exp^{-1}_{\overline{H}_{1,x}(t)}(\overline{H}_{2,x}(t))\right) = \exp_{[\alpha]\cdot\overline{H}_{1,x}(t)}\left(s \cdot \exp^{-1}_{[\alpha]\cdot\overline{H}_{1,x}(t)}([\alpha] \cdot \overline{H}_{2,x}(t))\right)$$

This holds since $[\alpha]$ is an isometry of $\overline{N}$ and takes geodesics to geodesics. Therefore, $F(s,x,t) := \pi \circ \overline{F}_x(s,t)$ is well-defined.

To see that F is continuous, note that there exist an evenly covered open set $U \subseteq N$ containing $H_i(x,0)$ and an open set $V \subseteq M$ containing x such that $H_i(V,0) \subseteq U$ for each i. The maps $H_i|_{V\times[0,1]}$ lift to maps $\overline{H}_i : V \times S^1 \to \pi^{-1}(U)$ with $\overline{H}_i(x,\cdot) = \overline{H}_{i,x}(\cdot)$. The function $\overline{F} : [0,1] \times U \times [0,1] \to \pi^{-1}(U)$ that's defined by $\overline{F}(s,y,t) := \exp_{\overline{H}_1(y,t)}\left(s \cdot \exp^{-1}_{\overline{H}_1(y,t)}(\overline{H}_2(y,t))\right)$ is continuous. Since $F|_{[0,1]\times U\times[0,1]} = \pi \circ \overline{F}$, it follows that F is continuous.

□

The literature on manifolds with no conjugate points is far too vast to be adequately surveyed in this space, but a few milestones are worth mentioning. Much of the interest in this subject dates to Hopf's result that any Riemannian metric on the torus T^2 without conjugate points must be flat [Ho] and his conjecture that the same should be true in any dimension. This was previously shown by Hedlund–Morse [HM] under the stronger assumption that the metric has no focal points. Hopf's proof used the theorem of Gauss–Bonnet, along with an analysis of the Riccati equation, and didn't generalize in an obvious way. In it, he introduced what would become known as the stable Jacobi tensor. A simple way of understanding the stable Jacobi tensor is that, for a fixed geodesic ray $\gamma : [0,\infty) \to M$, it assigns to each $v \in \mathrm{T}_{\gamma(0)}N$ the Jacobi field obtained as the limit as $t \to \infty$ of the Jacobi fields J_t along γ satisfying the initial conditions $J_t(0) = v$ and $J_t(t) = 0$. Conditions that ensure the existence and continuity of the stable Jacobi tensor were established by Eschenburg–O'Sullivan [EO'S], and a simplified proof of its existence was given by Innami [I].

Green [Gr1] showed that the total scalar curvature of a compact Riemannian manifold M with no conjugate points is non-positive and vanishes if and only if M is flat. Many others subsequently attacked Hopf's conjecture, mostly by focusing on the asymptotics of geodesics and their corresponding Busemann functions. Busemann functions and their level sets, called horospheres, were introduced in [Buse] and have since become a fundamental tool in Riemannian geometry. For example, they play an essential role in the original proof of the Cheeger–Gromoll splitting theorem in [CG1] and [CG2] and the simplification by Eschenburg–Heintze [EH]. Another notable contribution by Green [Gr2] was the claimed result that, when N is complete and has sectional curvature bounded below, geodesics from each point are uniformly divergent, although it must be noted that Eberlein identified [Eb2], and later filled [Eb3], a gap in Green's proof. Much effort has gone into studying the regularity of Busemann functions and horospheres. Eschenburg [Esc] proved their C^2 regularity when the stable Jacobi tensor is continuous, generalizing a result of Eberlein [Eb4] for the case of non-positive sectional curvature that was later recorded by Heintze–Im Hof[HI]. Eschenburg also introduced the class of manifolds with bounded asymptotes, which are manifolds with no conjugate points that have uniformly bounded stable Jacobi tensor, and proved the continuity of the stable Jacobi tensor for such spaces. However, the limitations of those methods were exposed when Ballmann–Brin–Burns [BBB] constructed a compact surface with no conjugate points and discontinuous stable Jacobi tensor, disproving a claim of Hopf.

Avez [Av] proved that any Riemannian torus with no focal points must be flat by examining the growth rate of the fundamental group. Croke [Cr2] showed that the volumes of balls in a simply connected Riemannian manifold N with no conjugate points are asymptotically at least as large as those in Euclidean space, with equality in the limit if and only if N is flat. Croke–Kleiner [CK1] proved that any Riemannian torus with no conjugate points and bounded asymptotes is flat, using a foliation of SN constructed by Heber [Heb] and a volume comparison argument. Around the same time, Burago–Ivanov [BurI] proved the Hopf conjecture, using an essentially different method than previous attempts. The key new tool in their proof was the asymptotic norm of a $\mathbb{Z}^m$-periodic metric, which was first described by Burago [Bura] and is discussed in Chapter 7.

Many other significant results have gone unmentioned here. Fundamental results on the geodesic flow of manifolds with no conjugate points or no focal points were proved by Eberlein in [Eb3] and [Eb1], building on the result of Klingenberg [Kli3] that any Riemannian manifold with Anosov geodesic flow has no conjugate points. Knieper [Kn] proved the ergodicity of the geodesic flow of a compact and non-flat surface with no conjugate points and continuous Jacobi tensors. Freire–Mañé [FM] proved that topological entropy and volume entropy are the same under a Hölder continuity assumption on the metric. A number of rigidity results for manifolds with non-positive sectional curvature have been shown by Croke and Croke–Kleiner to hold for manifolds with no conjugate points, for example the result in [CK2] that any complete Riemannian metric with no conjugate points on a non-compact flat manifold that agrees with a flat metric outside a compact set must itself be flat. It has also been shown by Bangert–Emmerich [BE] that any cylinder without conjugate points and whose loop length grows sublinearly in both directions is flat, which improves upon a result of Burns–Knieper [BK].

3.2 The set of geodesic loops

It will be shown that each covering space $\overline{N}/\mathrm{Z}([\gamma_1],\ldots,[\gamma_k])$ is realized concretely as a submanifold of the tensor bundle $\mathrm{T}^k N := \mathrm{T}^{(k,0)}N$. Specifically, the set

$$N_k := \{v_1 \otimes \cdots \otimes v_k \in \mathrm{T}^{(k,0)}_y N \,\big|\, y \in N, \exp(v_1) = \cdots = \exp(v_k) = y\}$$

is smooth submanifold of $\mathrm{T}^k N$ with respect to the inherited differentiable structure, and the space $\overline{N}/\mathrm{Z}([\gamma_1],\ldots,[\gamma_k])$ is realized as the connected component of N_k containing $w_1 \otimes \cdots \otimes w_k$, where each γ_{w_i} is the unique geodesic loop contained in $[\gamma_i]$. Roughly speaking, N_k is the submanifold of $\mathrm{T}^k N$ consisting entirely of the initial vectors of geodesic loops.

Throughout this section, the notation π is overloaded. In addition to referring to the universal covering map $\pi : \overline{N} \to N$, it also refers to the map $\pi : \mathrm{T}^k N \to N$ that takes $v_1 \otimes \cdots \otimes v_k$ to the common basepoint at which the v_i are located. The meaning will always be clear from context.

Lemma 3.2.1. *Let N be a complete n-dimensional Riemannian manifold with no conjugate points. For each $k \in \mathbb{N}$, N_k is a non-empty, smooth, and embedded submanifold of* $\mathrm{T}^k N$*, and $\pi|_{N_k} : N_k \to N$ is a local diffeomorphism.*[1]

[1] *I'm indebted to a student at Capital Normal University, whose name I unfortunately never learned, who suggested an elegant simplification of my original proof of this result.*

Proof. Since $0\otimes\cdots\otimes 0 \in N_k$ for each $x \in N$, $N_k \neq \emptyset$. Let $f : \mathrm{T}^k N \to N^{k+1}$ be the map defined by $f(v_1\otimes\cdots\otimes v_k) := (\pi(v_1\otimes\cdots\otimes v_k), \exp(v_1), \ldots, \exp(v_k))$. Since N has no conjugate points, for each $v_i \in \mathrm{T}N$, the derivative of $\exp_{\pi(v_i)}$ is non-singular at v_i; that is, $\mathrm{D}_{v_i}\exp_{\pi(v_i)} : \mathrm{T}_{v_i}(\mathrm{T}_{\pi(v_i)}N) \to \mathrm{T}_{\exp(v_i)}N$ is a linear isomorphism. The restriction $\mathrm{D}_{v_1\otimes\cdots\otimes v_k}\pi|_{\mathrm{T}_{\pi(v_1\otimes\cdots\otimes v_k)}N} : \mathrm{T}_{\pi(v_1\otimes\cdots\otimes v_k)}N \to \mathrm{T}_{\pi(v_1\otimes\cdots\otimes v_k)}N$ is also a linear isomorphism. These show that $\mathrm{rank}(f) = n(k+1)$ everywhere, since with respect to the splitting

$$\mathrm{T}_{v_1\otimes\cdots\otimes v_k}(\mathrm{T}^k N) \cong \mathrm{T}_{\pi(v_1\otimes\cdots\otimes v_k)}N \times \mathrm{T}_{v_1}(\mathrm{T}_{\pi(v_1)}N) \times\cdots\times \mathrm{T}_{v_k}(\mathrm{T}_{\pi(v_k)}N)$$

the derivative of the i-th component of f has rank n when restricted to the i-th component of $\mathrm{T}_{v_1\otimes\cdots\otimes v_k}(\mathrm{T}^k N)$ for each $1 \le i \le k+1$. Denote by $D := \{(y,\ldots,y) \in N^{k+1} \,\big|\, y \in N\}$ the diagonal in N^{k+1}. Since $\dim(\mathrm{T}^k N) = \dim(N^{k+1}) = n(k+1)$, the inverse function theorem implies that f is a local diffeomorphism, and $N_k = f^{-1}(D)$ is a smooth and embedded submanifold of $\mathrm{T}^k N$ of dimension n. Fix $v_1\otimes\cdots\otimes v_k \in N_k$. Assume that

$$\mathrm{T}_{v_1\otimes\cdots\otimes v_k}N_k \cap (\{0\}\times \mathrm{T}_{v_1}(\mathrm{T}_{\pi(v_1)}N)\times\cdots\times \mathrm{T}_{v_k}(\mathrm{T}_{\pi(v_k)}N)) \neq \{(0,\ldots,0)\},$$

and choose a non-zero $w = (0, w_1, \ldots, w_k) \in \mathrm{T}_{v_1\otimes\cdots\otimes v_k}N_k$. By the definition of N_k,

$$\mathrm{D}_{v_i}\exp_{\pi(v_1\otimes\cdots\otimes v_k)}(w_i) = \mathrm{D}_{v_1\otimes\cdots\otimes v_k}\pi|_{N_k}(w) = 0$$

for all $1 \le i \le k$. However, for some i, $w_i \neq 0$, so $\mathrm{D}_{v_i}\exp_{\pi(v_1\otimes\cdots\otimes v_k)}(w_i) \neq 0$. This is a contradiction. This implies that $\mathrm{rank}(\pi|_{N_k}) = n$, so $\pi|_{N_k}$ is a local diffeomorphism.

□

By Lemma 2.2.2, to show that $\pi|_{N_k}$ is a covering map, it suffices to show that N_k, when endowed with the pull-back metric $\pi|_{N_k}^*(g)$, is complete. This is done by showing that the lengths of vectors along a constant-speed path in N_k cannot blow up in finite time. To that end, define length functions $L_i : N_k \to [0,\infty)$ by setting $L_i(v_1\otimes\cdots\otimes v_k) := \|v_i\|$.

Lemma 3.2.2. *Let N be a complete Riemannian manifold with no conjugate points. With respect to the pull-back metric $\pi|_{N_k}^*(g)$ on N_k, each $L_i : N_k \to [0,\infty)$ is smooth and has gradient satisfying $\|\nabla L_i\| < 2$.*

Proof. On the zero-section in $\mathrm{T}^k N$, which is a connected component of N_k, L_i is identically zero and the result follows. On any connected component N_0 of N_k disjoint from the zero-section, L_i is smooth. Let $v = v_1 \otimes \cdots \otimes v_k \in N_0$, fix a unit vector $w = (w_0, \ldots, w_k) \in \mathrm{T}_v N_0$, and choose any curve $\tilde{\alpha} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_k) : (-\varepsilon, \varepsilon) \to N_0$ satisfying $\tilde{\alpha}'(0) = w$. Let $u := \pi_*(w) \in T_{\pi(v)}N$ and $\alpha := \pi \circ \tilde{\alpha}$, so that $\alpha'(0) = u$. Define a variation $V : (-\varepsilon, \varepsilon) \times [0,1] \to N$ of γ_{w_i} by setting $V(s,t) := \exp_{\alpha(t)}(t\tilde{\alpha}_i(s))$. Then V is a variation through geodesic loops, so its variation field J is a Jacobi field along γ_{w_i}. Since N_0 is endowed with the pull-back metric, $\|J(0)\| = \|\tilde{\alpha}'(0)\| = \|w\| = 1$. Set $\alpha_s(t) := V(s,t)$. By the first variation formula,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L(\alpha_s) &= \frac{g(\alpha_0'(1) - \alpha_0'(0), J(0))}{L(\alpha_0)} \\
&\leq \frac{\|\alpha_0'(1) - \alpha_0'(0)\| \, \|J(0)\|}{L(\alpha_0)} \\
&\leq \frac{[\|\alpha_0'(1)\| + \|\alpha_0'(0)\|] \, \|J(0)\|}{L(\alpha_0)} \\
&= 2\|J(0)\| \\
&= 2
\end{aligned}
$$

So $\|\nabla_w L_i\| = \|\frac{\partial}{\partial s}\big|_{s=0} L_i \circ \tilde{\alpha}(s)\| = \|\frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} L(\alpha_s)\| \leq 2$. Moreover, the first inequality is an equality if and only if $\alpha_0'(1) = -\alpha_0'(0)$, which is impossible since these are the initial and final vectors of a geodesic loop. So $\|\nabla_w L_i\| < 2$. It follows that $\|\nabla L_i\| = \max\{\|\nabla_w L_i\| \,\big|\, w \in \mathrm{T}_v N_0, \|w\| = 1\} < 2$.

□

Remark 3.2.3. Since $\alpha_0 = \gamma_{w_i}$, the above argument shows that ∇L_i points in the direction of $\gamma_{w_i}'(1) - \gamma_{w_i}'(0)$. This implies that w is a critical point of L_i if and only if w_i is the initial vector of a closed geodesic.

Lemma 3.2.4. *Let N be a complete Riemannian manifold with no conjugate points and N_0 a connected component of N_k. When endowed with the pull-back metric $\pi|_{N_0}^*(g)$, N_0 is complete.*

Proof. Let $v^n = v_1^n \otimes \cdots \otimes v_k^n \in N_0$ be a Cauchy sequence. Then there exists $C \geq 0$ such that $\mathrm{d}(v^n, v^1) < C$ for all n; in other words, $v^n \in \mathrm{B}(v^1, C)$. By Lemma 3.2.2, $\|\nabla L_i\| < 2$ for all i, which means that $\|L(v_i^n) - L(v_i^1)\| < 2C$ for all n. Therefore, each v^n lies in the set

$$
\bigcup_{y \in \overline{\mathrm{B}}(\pi(v^1), C)} \{w_1 \otimes \cdots \otimes w_k \in \mathrm{T}^k N \,\big|\, \|w_i\| \leq \|v_i^1\| + 2C, \pi(w_1 \otimes \cdots \otimes w_k) = y\}
$$

By compactness, v^n contains a convergent subsequence, which implies that v^n itself converges.

□

Using Lemma 3.2.4, one may prove the previously mentioned structure theorem for the loop space of a complete manifold with no conjugate points, which will play a central role in what follows. It will help to introduce some notation first. If $[\sigma] \in \pi_1(N, p)$, then, since N has no conjugate points, there exists a unique vector $v \in \mathrm{T}_p N$ such that γ_v is a geodesic loop in $[\sigma]$. Therefore, if $[\sigma_1], \ldots, [\sigma_k] \in \pi_1(N, p)$, there exists a unique $v = v_1 \otimes \cdots \otimes v_k \in N_k$ such that, for each i, $v_i \in [\sigma_i]$ and γ_{v_i} is a geodesic loop. The connected component of N_k containing v will be denoted $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$, and the restriction $\pi|_{\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}}$ will be denoted $\pi_{[\sigma_1],\ldots,[\sigma_k]}$.

Theorem 3.2.5. *Let N be a connected and complete Riemannian manifold with no conjugate points, $p \in N$, and $[\sigma_1], \ldots, [\sigma_k] \in \pi_1(N, p)$. Then each of the following holds:*

(a) *The set*

$$\tilde{N}_k := \{v_1 \otimes \cdots \otimes v_k \in \mathrm{T}^k N \,\big|\, \exp(v_1) = \cdots = \exp(v_k) = x\}$$

is a smooth submanifold of the tensor bundle $\mathrm{T}^k N$;

(b) *Each projection $\pi_{[\sigma_1],\ldots,[\sigma_k]} : \tilde{N}_{[\sigma_1],\ldots,[\sigma_k]} \to N$ is a smooth covering map; and*

(c) *The fundamental group of $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ satisfies*

$$(\pi_{[\sigma_1],\ldots,[\sigma_k]})_*(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}, v_1 \otimes \cdots \otimes v_k)) = \mathrm{Z}([\gamma_{v_1}], \ldots, [\gamma_{v_k}])$$

Consequently, $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ is diffeomorphic to $\overline{N}/\mathrm{Z}([\gamma_1], \ldots, [\gamma_k])$.

Proof. **(a)** This is shown in Lemma 3.2.1.

(b) This follows from Lemma 3.2.1, Lemma 3.2.4, and Lemma 2.2.2.

(c) Let $[\alpha] = (\pi_{[\sigma_1],\ldots,[\sigma_k]})_*([\tilde{\alpha}])$ for some $[\tilde{\alpha}] \in \pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}, v_1 \otimes \cdots \otimes v_k)$, and fix a representative $\tilde{\alpha} : [0,1] \to N$ of $[\tilde{\alpha}]$. Write $\tilde{\alpha} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_k)$. Then $\tilde{\alpha}_i(0) = \tilde{\alpha}_i(1) = v_i$ for each i, and $[\alpha] = [\pi_{[\sigma_1],\ldots,[\sigma_k]} \circ \tilde{\alpha}]$. Define a map $H_i : [0,1] \times [0,1] \to N$ by $H_i(s,t) := \gamma_{\tilde{\alpha}_i(s)}(t)$. This is a homotopy through geodesic loops. This map satisfies $H_i(0,\cdot) = H_i(1,\cdot) = \gamma_{v_i}(\cdot)$ and $H_i(\cdot,0) = H_i(\cdot,1) = \pi_{[\sigma_1],\ldots,[\sigma_k]} \circ \tilde{\alpha}(\cdot)$. It follows that $[\gamma_{v_i}] = [\pi_{[\sigma_1],\ldots,[\sigma_k]} \circ \tilde{\alpha}][\gamma_{v_i}][\pi_{[\sigma_1],\ldots,[\sigma_k]} \circ \tilde{\alpha}]^{-1} = [\alpha][\gamma_{v_i}][\alpha]^{-1}$, and, consequently, $[\alpha] \in \mathrm{Z}([\gamma_{v_1}], \ldots, [\gamma_{v_k}])$. So

$$(\pi_{[\sigma_1],\ldots,[\sigma_k]})_*(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}, v_1 \otimes \cdots \otimes v_k)) \leq \mathrm{Z}([\gamma_{v_1}], \ldots, [\gamma_{v_k}])$$

On the other hand, suppose $[\alpha] \in \mathrm{Z}([\gamma_{v_1}], \ldots, [\gamma_{v_k}])$, and fix a representative $\alpha : [0,1] \to N$ of $[\alpha]$. Lift α to a map $\tilde{\alpha} : [0,1] \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ with $\tilde{\alpha}(0) = v_1 \otimes \cdots \otimes v_k$. Write $\tilde{\alpha} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_i)$, and let

$H_i(s) := A_{[\alpha|_{[0,s]}]}([\gamma_{v_i}]) \in \pi_1(N, \alpha(s))$. By construction, $\tilde{\alpha}_i(s)$ is the initial vector of the unique geodesic loop in $H_i(s)$. Since $H_i(1) = A_{[\alpha]}([\gamma_{v_i}]) = [\alpha][\gamma_{v_i}][\alpha]^{-1} = [\gamma_{v_i}]$, it follows that $\tilde{\alpha}_i(1) = v_i$. So $\tilde{\alpha}(0) = \tilde{\alpha}(1) = v_1 \otimes \cdots \otimes v_k$, and $\tilde{\alpha} \in \pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}, v_1 \otimes \cdots \otimes v_k)$. This shows that

$$[\alpha] = (\pi_{[\sigma_1],\ldots,[\sigma_k]})_*([\tilde{\alpha}]) \in (\pi_{[\sigma_1],\ldots,[\sigma_k]})_*(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}, v_1 \otimes \cdots \otimes v_k)$$

and, consequently,

$$Z([\gamma_{v_1}], \ldots, [\gamma_{v_k}]) \leq (\pi_{[\sigma_1],\ldots,[\sigma_k]})_*(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}, v_1 \otimes \cdots \otimes v_k))$$

It follows by general theory that $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ is diffeomorphic to $\overline{N}/Z([\gamma_1], \ldots, [\gamma_k])$.

□

Remark 3.2.6. Each $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ corresponds to an equivalence class of the k-tuple $([\sigma_1], \ldots, [\sigma_k])$ under the conjugation map A, not just the equivalence classes of the individual $[\sigma_i]$. That is, $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]} = \tilde{N}_{[\gamma_1],\ldots,[\gamma_k]}$ if and only if $([\gamma_1], \ldots, [\gamma_k]) = A_{[\alpha]}([\sigma_1], \ldots, [\sigma_k])$ for some $[\alpha] \in \mathscr{P}$. If N is any compact hyperbolic surface, $y \in N$, and $[\beta], [\gamma] \in \pi_1(N, y)$ are independent in the sense that neither is a multiple of the other, then it follows from the well-known theorem of Preissmann [P] that $[\beta]^{-1}[\gamma][\beta] \neq [\gamma]$. Consequently, for $[\sigma] := [\beta]^{-1}[\gamma][\beta]$, $\tilde{N}_{[\beta],[\gamma]} \neq \tilde{N}_{[\beta],[\sigma]}$, even though $[\beta] \cong [\beta]$ and $[\gamma] \cong [\sigma]$.

A few properties of the spaces $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ are described in the next results.

Lemma 3.2.7. *Let N be a complete Riemannian manifold with no conjugate points, $p \in N$, and $[\sigma] \in \pi_1(N, p)$. If $v \in \tilde{N}_{[\sigma]}$, then $\tilde{\gamma}_v : [0, 1] \to \tilde{N}_{[\sigma]}$ defined by $\tilde{\gamma}_v(t) = \gamma'_v(t)$ is the unique closed geodesic based at v in its free homotopy class, $[\tilde{\gamma}_v]$ is in the stabilizer of the conjugation action of $\pi_1(\tilde{N}_{[\sigma]}, v)$ on itself, and $<[\tilde{\gamma}_v]>$ is a normal subgroup of $\pi_1(\tilde{N}_{[\sigma]}, v)$.*

Proof. Because $\tilde{N}_{[\sigma]}$ is a local isometry, $\tilde{\gamma}_{v_x}$ is a geodesic. Because $\tilde{\gamma}_{v_x}(0) = \gamma'_{v_x}(0) = \gamma'_{v_x}(1) = \tilde{\gamma}_{v_x}(1)$, $\tilde{\gamma}_{v_x}$ is closed. Since N has no conjugate points, $\tilde{\gamma}_v$ is the unique closed geodesic in $[\tilde{\gamma}_v]$. By Theorem 3.2.5(c), $(\pi_{[\sigma]})_*(\pi_1(\tilde{N}_{[\sigma]}, v)) = Z([\gamma_v])$; it follows that $[\tilde{\gamma}_v] = [\tilde{\alpha}]^{-1}[\tilde{\gamma}_v][\tilde{\alpha}]$ for any $[\tilde{\alpha}] \in \pi_1(\tilde{N}_{[\sigma]}, v)$. So $[\tilde{\gamma}_v]$ is in the stabilizer of the conjugation action, and $\tilde{\gamma}_v$ is the unique closed geodesic based at v in its free homotopy class. It follows immediately that $<[\tilde{\gamma}_v]>$ is normal.

□

Lemma 3.2.8. *Let T^n be a torus, $x \in T^n$, $\{[s_1], \ldots, [s_n]\}$ a minimal generating set for $\pi_1(T^n, x)$, N a complete Riemannian manifold with no conjugate points, $f : T^n \to N$ a continuous function, $y = f(x)$, and $[\sigma_i] = f_*([s_i])$ for each i. Then there exists a canonical lift $\tilde{F} : T^n \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ such that $f = \pi_{[\sigma_1],\ldots,[\sigma_n]} \circ \tilde{F}$.*

Proof. Since N has no conjugate points, each $[\sigma_i]$ contains a unique geodesic loop based at y. Let $v_i \in \mathrm{T}_x N$ be the initial vector of that loop. Then $v_1 \otimes \cdots \otimes v_n \in \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ and $\pi_{[\sigma_1],\ldots,[\sigma_n]}(v_1 \otimes \cdots \otimes v_n) = y$. Since $\pi_1(T^n, x) =< [s_1], \ldots, [s_n] >$ is Abelian, so is $f_*(\pi_1(T^n, x)) =< [\sigma_1], \ldots, [\sigma_n] >$. Note that

$$(\pi_{[\sigma_1],\ldots,[\sigma_n]})_*(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}, v_1 \otimes \cdots \otimes v_n)) = \mathrm{Z}([\sigma_1], \ldots, [\sigma_n]) =< [\sigma_1], \ldots, [\sigma_n] >$$

It follows that f lifts to a continuous function $\tilde{F} : T^m \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ satisfying $\tilde{F}(x) = v_1 \otimes \cdots \otimes v_n$ and $f = \pi_{[\sigma_1],\ldots,[\sigma_n]} \circ \tilde{F}$.

□

Loosely speaking, the covering map $\pi_{[\sigma]} : \tilde{N}_{[\sigma]} \to N$ unwraps the set of closed geodesics freely homotopic to any representative of $[\sigma]$ so that there's at most one through each point. This contrasts with the situation on N itself, as shown by the next example.

Example 3.2.9. Let $\alpha, \beta, \gamma : \mathbb{R}^3 \to \mathbb{R}^3$ be the isometries

$$\alpha(x, y, z) := (x+1, y, z)$$

$$\beta(x, y, z) := (x, y+1, z)$$

$$\gamma(x, y, z) := (y, -x+1, z-1)$$

Then $\gamma^{-1} \circ \alpha \circ \gamma = \beta$. The quotient $F := \mathbb{R}^3 / < \alpha, \beta, \gamma >$ is a compact and flat three-manifold. The axes of α and β descend to a foliation of F by two families of closed geodesics that are everywhere perpendicular. Since α and β are conjugate, those closed geodesics are all freely homotopic. This shows that, through each point in F, there exist two distinct freely homotopic closed geodesics.[2]

A bit more notation will be introduced. Recall that, by definition, each $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ is connected. It follows that, for each $p \in N$, $v \in \pi^{-1}_{[\sigma_1],\ldots,[\sigma_k]}(p)$, and $\overline{p} \in \pi^{-1}(p)$, there exists a covering map $\psi_{[\sigma_1],\ldots,[\sigma_k]} : \overline{N} \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ satisfying $\psi_{[\sigma_1],\ldots,[\sigma_k]}(\overline{p}) = v$ and $\pi = \pi_{[\sigma_1],\ldots,[\sigma_k]} \circ \psi_{[\sigma_1],\ldots,[\sigma_k]}$. Define

$$\tilde{C}_{[\sigma_1],\ldots,[\sigma_k]} := \{v_1 \otimes \cdots \otimes v_k \in \tilde{N}_{[\sigma_1],\ldots,[\sigma_k]} \,\big|\, \gamma_{v_i} \text{ is a closed geodesic for all } i\}$$

[2]This example was constructed during a conversation with Jason DeBlois.

In the case that $k = 1$, $\tilde{C}_{[\sigma]}$ should be thought of as the set of closed geodesics freely homotopic to any representative of $[\sigma]$. The following combines two results of Croke–Schroeder [CrS] about this set. It will be shown in Chapter 4 that, when N has no focal points, much more can be said.

Theorem 3.2.10. *(Croke–Schroeder) Let N be a compact Riemannian manifold with no conjugate points, $p \in N$, and $[\sigma] \in \pi_1(N, p)$. Then $\tilde{C}_{[\sigma]}$ is a compact and connected subset of $\tilde{N}_{[\sigma]}$.*

The proof that $\tilde{C}_{[\sigma]}$ is connected uses Morse theory and is more difficult than the proof that it's compact. It follows from this compactness that, for each $p \in N$, there are at most finitely many $v \in \mathrm{T}_p N$ such that γ_v is a closed geodesic freely homotopic to any representative of $[\sigma]$. Using this, one may show that each $\tilde{C}_{[\sigma_1],\ldots,[\sigma_k]}$ is compact. However, in the setting of no conjugate points, one may use Kleiner's counterexample to the flat torus theorem [Kle] to show that $\tilde{C}_{[\sigma_1],\ldots,[\sigma_k]}$ might not be connected for $k \geq 2$.

Remark 3.2.11. The question of whether each $\tilde{C}_{[\sigma]}$ is path-connected, or more specifically locally rectifiably path-connected in the sense of [CrS], is open. Ivanov–Kapovitch [IK] recently bypassed this issue to prove new results about the fundamental group of a compact manifold with no conjugate points.

3.3 The loop map

The principal application of Theorem 3.2.5 will be to construct an explicit homotopy between any two maps $f, g : T^n \to N$ satisfying $f_*(\pi_1(T^m, x)) \cong g_*(\pi_1(T^m, x))$. The idea is to first deform f and g into canonical forms corresponding to representatives in a suitable cover $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$, then to connect those canonical forms through a path in $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ that defines a homotopy in N. The key tool in doing so is a map $\Upsilon_{v_1 \otimes \cdots \otimes v_n} : T^n \to N$ that can be associated to $v_1 \otimes \cdots \otimes v_n \in \tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ whenever $< [\sigma_1], \ldots, [\sigma_k] >$ is Abelian.

It will help to introduce some notation. Let $\sim$ denote the quotient map on $[0, 1]$ that identifies the endpoints, so that $S^1 \cong [0, 1]/\sim$. Fix $\theta = (\theta_1, \ldots, \theta_n) \in T^n$, and, for each $i = 1, \ldots, n$, define a map $s_i : T^i \cong T^{i-1} \times ([0, 1]/\sim) \to T^n$ by setting $s_i(x, t) := (x, \theta_i + t, \theta_{i+1}, \ldots, \theta_n)$. When $i = 1$, this takes the form $s_1(t) := (\theta_1 + t, \theta_2, \ldots, \theta_n)$. When T^n has the standard product metric, each $s_i(x, \cdot)$ is a closed geodesic in T^n based at $(x, \theta_i, \ldots, \theta_n)$. Let $s_0 : T^0 \cong \{\mathrm{pt}\} \to M$ be defined by $s_0(T^0) = (\theta_1, \ldots, \theta_n)$.

Let $[\sigma_1],\ldots,[\sigma_n]$ generate an Abelian subgroup of $\pi_1(N,q)$. For each $v = v_1\otimes\cdots\otimes v_n \in \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$, the **loop map determined by** v, denoted $\Upsilon_v : T^n \cong ([0,1]/\sim)^n \to N$, is defined by the following inductive process: Write $p := \pi_{[\sigma_1],\ldots,[\sigma_n]}(v)$. Let $\Upsilon_1 : S^1 \cong ([0,1]/\sim) \to N$ be defined by

$$\Upsilon_1(t) := \gamma_v = \exp_p(tv_1)$$

Then Υ_1 is a geodesic loop based at $\Upsilon(\theta_1) = p$, and $(\Upsilon_1)_*(\pi_1(S^1,\theta_1)) =< [\sigma_1] >$. Suppose that, for some $1 \le i \le n-1$, a map $\Upsilon_i : T^i \to N$ has been defined that satisfies $\Upsilon_i(\theta_1,\ldots,\theta_i) = p$ and $(\Upsilon_i)_*\big(\pi_1(T^i,(\theta_1,\ldots,\theta_i))\big) =< [\sigma_1],\ldots,[\sigma_i] >$. Note that

$$(\pi_{[\sigma_1],\ldots,[\sigma_{i+1}]})_*(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_{i+1}]}, v_1\otimes\cdots\otimes v_{i+1})) = \mathrm{Z}([\sigma_1],\ldots,[\sigma_{i+1}]) =< [\sigma_1],\ldots,[\sigma_{i+1}] >$$

where the first equality holds by Theorem 3.2.5(c) and the second because $< [\sigma_1],\ldots,[\sigma_{i+1}] >$ is Abelian. Since $(\Upsilon_i)_*\big(\pi_1(T^i,(\theta_1,\ldots,\theta_i))\big) =< [\sigma_1],\ldots,[\sigma_i] >$, Υ_i lifts to a map $\tilde{\Upsilon}_i : T^i \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_i]}$ with $\tilde{\Upsilon}_i(\theta_1,\ldots,\theta_i) = (v_1,\ldots,v_{i+1})$. Write $\tilde{\Upsilon}_i = (V_1,\ldots,V_{i+1})$, where each V_i is a vector field along Υ_i consisting of the initial vectors of geodesic loops freely homotopic to σ_i. Let $\Upsilon_{i+1} : T^{i+1} \cong T^i\times([0,1]/\sim) \to N$ be defined by

$$\Upsilon_{i+1}(x,t) := \exp_{\Upsilon_i(x)}(tV_{i+1}(x))$$

As in the first case, each $\Upsilon_{i+1}(x,\cdot)$ is a geodesic loop based at $\Upsilon_i(x)$, $\Upsilon_{i+1}(\theta_1,\ldots,\theta_{i+1}) = p$, and

$$(\Upsilon_{i+1})_*\big(\pi_1(T^{i+1},(\theta_1,\ldots,\theta_{i+1}))\big) =< [\sigma_1],\ldots,[\sigma_{i+1}] >$$

Let $\Upsilon_v := \Upsilon_n$. By construction, for each $x \in T^{i-1}$, $\Upsilon_v\circ s_i(x,\cdot)$ is a geodesic loop based at $\Upsilon_v\circ s_i(x,0) = \Upsilon_v(x,\theta_i,\ldots,\theta_n)$.

Remark 3.3.1. If the loop map $\Upsilon_v : T^n \to N$ is totally geodesic, then $v \in \tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$. By the flat strip theorem, discussed in the next chapter, the converse is true when N has no focal points.

Lemma 3.3.2. *Let $\tilde{F} : M \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ be continuous. Then the composition $\Upsilon_{\tilde{F}} : M\times T^n \to N$ defined by $\Upsilon_{\tilde{F}}(x,t_1,\ldots,t_n) := \Upsilon_{\tilde{F}(x)}(t_1,\ldots,t_n)$ is continuous.*

Proof. This is seen inductively and follows from the continuity of $\exp : \mathrm{T}N \to N$.

$\square$

Theorem 3.3.3. *Let T^n be a flat Riemannian torus, $p \in T^n$, $\{[s_1], \ldots, [s_n]\}$ a minimal generating set for $\pi_1(T^n, p)$, and N a complete Riemannian manifold with no conjugate points. If $f, g : T^n \to N$ are continuous, then f and g are homotopic if and only if $f_*([s_1], \ldots, [s_n]) \cong g_*([s_1], \ldots, [s_n])$.*

Proof. If f and g are homotopic, then any homotopy $H : [a,b] \times T^n \to N$ from f to g yields a path $\alpha(\cdot) := H(\cdot, x)$ such that $A_{[\alpha]}(f_*([s_1], \ldots, [s_n])) = g_*([s_1], \ldots, [s_n])$. Conversely, suppose that $f_*([s_1], \ldots, [s_n]) \cong g_*([s_1], \ldots, [s_n])$. Without loss of generality, one may suppose that $p = (\theta_1, \ldots, \theta_n)$, where θ_i are as in the definition of the loop map Υ. Write $[\sigma_i] := f_*([s_i])$. Let $\tilde{F}$ be the canonical lift of f to $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ given by Lemma 3.2.8. Write $\tilde{F} = (V_1, \ldots, V_n)$ and $v_i := V_i(\theta_1, \ldots, \theta_n)$; the V_i are vector fields along f, each of which consists of the initial vectors of geodesic loops freely homotopic to $[\sigma_i]$. The idea is to first deform f and g into canonical forms using the loop map Υ, then to connect them by a path in $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$.

Specifically, it will first be shown that f is homotopic to $\Upsilon_{v_1 \otimes \cdots \otimes v_m}$. Some preliminaries will help simplify the argument. A family of maps $\tilde{F}_i : T^{i-1} \to \tilde{N}_{[\sigma_i],\ldots,[\sigma_n]}$ can be defined by setting

$$\tilde{F}_i(\cdot) := (V_i \circ s_i(\cdot, 0), \ldots, V_n \circ s_i(\cdot, 0)) = (V_i(\cdot, \theta_i, \ldots, \theta_n), \ldots, V_m(\cdot, \theta_i, \ldots, \theta_n))$$

Since $\pi_{[\sigma_i],\ldots,[\sigma_n]} \circ \tilde{F}_i = f \circ s_{i-1}$, each $\tilde{F}_i$ is a lift of $f \circ s_{i-1}$ to $\tilde{N}_{[\sigma_i],\ldots,[\sigma_n]}$. The corresponding $\Upsilon_{\tilde{F}_i}$ are maps from T^n into N and satisfy $\Upsilon_{\tilde{F}_i} \circ s_{i-1} = f \circ s_{i-1}$. Note that $\tilde{F}_1$ has as its image $v_1 \otimes \cdots \otimes v_n$, so the goal is to show that f is homotopic to $\Upsilon_{\tilde{F}_1}$. This is done by showing inductively that f is homotopic to each $\Upsilon_{\tilde{F}_i}$, counting down from the base case $i = n$. For each $x \in T^{n-1}$, $f \circ s_n(x, 0) = \Upsilon_{\tilde{F}_n} \circ s_n(x, 0)$ and $[f \circ s_n(x, \cdot)] = [\Upsilon_{\tilde{F}_n} \circ s_n(x, \cdot)] \in \pi_1(N, f \circ s_n(x, 0))$. It follows from Lemma 3.1.9 that $f = f \circ s_n$ is homotopic to $\Upsilon_{\tilde{F}_n} = \Upsilon_{\tilde{F}_n} \circ s_n$. This completes the base case. The inductive step is similar, with one added wrinkle. Suppose that f is homotopic to $\Upsilon_{\tilde{F}_{i+1}}$. Since $\pi_{[\sigma_i],\ldots,[\sigma_n]} \circ \tilde{F}_{i+1} = f \circ s_i = \Upsilon_{\tilde{F}_{i+1}} \circ s_i$, $\tilde{F}_i$ is a lift of $\Upsilon_{\tilde{F}_{i+1}} \circ s_i$ to $\tilde{N}_{[\sigma_{i+1}],\ldots,[\sigma_n]}$. For each $x \in T^{i-1}$, $\Upsilon_{\tilde{F}_{i+1}} \circ s_i(x, 0) = \Upsilon_{\tilde{F}_i} \circ s_i(x, 0)$ and $[\Upsilon_{\tilde{F}_{i+1}} \circ s_i(x, \cdot)] = [\Upsilon_{\tilde{F}_i} \circ s_i(x, \cdot)] \in \pi_1(N, f \circ s_i(x, 0))$, so, as in the base case, Lemma 3.1.9 shows that $\Upsilon_{\tilde{F}_{i+1}} \circ s_i$ is homotopic to $\Upsilon_{\tilde{F}_i} \circ s_i$. Let $H : [a,b] \times T^i \to N$ be a homotopy from $\Upsilon_{\tilde{F}_{i+1}} \circ s_i$ to $\Upsilon_{\tilde{F}_i} \circ s_i$. By construction, $H(a, \theta_1, \ldots, \theta_i) = \Upsilon_{\tilde{F}_{i+1}} \circ s_i(\theta_1, \ldots, \theta_i) = f \circ s_i(x) = \pi_{[\sigma_{i+1}],\ldots,[\sigma_n]}(\theta_{i+1}, \ldots, \theta_n)$ and

$$\begin{aligned} H_*\Big(\pi_1([a,b] \times T^i, (a, \theta_1, \ldots, \theta_i))\Big) &= < [\sigma_1], \ldots, [\sigma_i] > \\ &\leq Z([\sigma_{i+1}], \ldots, [\sigma_n]) \\ &= (\pi_{[\sigma_{i+1}],\ldots,[\sigma_n]})_*\big(\tilde{N}_{[\sigma_{i+1}],\ldots,[\sigma_n]}, v_{i+1} \otimes \cdots \otimes v_n\big) \end{aligned}$$

Hence there exists a lift $\tilde{H} : [a,b] \times T^i \to \tilde{N}_{[\sigma_{i+1}],\ldots,[\sigma_n]}$ of H satisfying $\tilde{H}(a,\theta_1,\ldots,\theta_i) = (\theta_{i+1},\ldots,\theta_n)$ and $\pi_{[\sigma_{i+1}],\ldots,[\sigma_n]} \circ \tilde{H} = H$. The map $\Upsilon_{\tilde{H}}$ is a homotopy from $\Upsilon_{\tilde{F}_{i+1}}$ to $\Upsilon_{\tilde{F}_i}$. This completes the inductive step.

Applying the same argument to g, one has that g is homotopic to some $\Upsilon_{w_1 \otimes \cdots \otimes w_n}$. Since $f_*([s_1],\ldots,[s_n]) \cong g_*([s_1],\ldots,[s_n])$, $w_1 \otimes \cdots \otimes w_n \in \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$. By definition, $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ is connected, so there exists a path $\tilde{\gamma} : [a,b] \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ from $v_1 \otimes \cdots \otimes v_n$ to $w_1 \otimes \cdots \otimes w_n$. The map $\Upsilon_{\tilde{\gamma}}$ is a homotopy from $\Upsilon_{v_1 \otimes \cdots \otimes v_n}$ to $\Upsilon_{w_1 \otimes \cdots \otimes w_n}$.

$\square$

Remark 3.3.4. Since N is an Eilenberg–Mac Lane space, Theorem 3.3.3 may be proved using general topology, but that method is not constructive.

Remark 3.3.5. The conclusion of Theorem 3.3.3 may fail to hold if the condition $f_*([s_1],\ldots,[s_n]) \cong g_*([s_1],\ldots,[s_n])$ is weakened to $f_*(\pi_1(T^n,x)) \cong g_*(\pi_1(T^n,x))$. The difference is that the former condition keeps track of the order of terms, while the latter doesn't. For example, on a flat T^2, there exists a unique totally geodesic map $T : T^2 \to T^2$ that permutes $[s_1]$ and $[s_2]$. One has that $\mathrm{id}_*(\pi_1(T^2,x)) = \pi_1(T^2,x) = T_*(\pi_1(T^2,x))$, but T is not homotopic to the identity map.

Chapter 4

Manifolds with no focal points

4.1 The flat strip theorem and its consequences

Recall that, by definition, a Riemannian manifold N has no focal points if the exponential map on the normal bundle of each geodesic is non-singular. More specifically, this means that, for each interval $I \subseteq \mathbb{R}$ and each geodesic $\gamma : I \to \mathbb{R}$ such that $\gamma(I)$ is an embedded submanifold of N, $\mathrm{focal}(\gamma(I)) = \emptyset$. One sees that N has no focal points if and only if $\overline{N}$ has no focal points. Moreover, since each point $p \in N$ is an embedded submanifold equal to the image of a constant geodesic, N must also have no conjugate points. If, in addition, N is complete and simply connected, then all geodesics in N are minimal, and every non-constant geodesic extends to a maximal embedding.

Theorem 4.1.1. *Let N be a complete and simply connected Riemannian manifold. Then the following are equivalent:*

(i) *N has no focal points;*

(ii) *For each totally geodesic submanifold S of N, $\mathrm{focal}(S) = \emptyset$;*

(iii) *For each geodesic $\gamma : \mathbb{R} \to N$, $\exp|_{\mathrm{N}\gamma(\mathbb{R})} : \mathrm{N}\gamma(\mathbb{R}) \to N$ is a diffeomorphism;*

(iv) *$r(N) = \infty$;*

(v) *$r_f(N) = \infty$; and*

(vi) *For each $p \in M$, $\nabla^2 \mathrm{d}^2(\cdot, p) : N \to [0, \infty)$ is positive-definite.*

Proof. **(i) ⇔ (ii)** The implication (ii) ⇒ (i) is clear. Suppose $S \subseteq M$ is a totally geodesic submanifold and that $v \in \mathrm{focal}(S)$. By Lemma 2.5.2(a), $\exp(v)$ is focal to S along $\gamma := \gamma_v|_{[0,1]}$. By Lemma 2.5.1 and equation (2.1), there exists a non-trivial normal Jacobi field J along γ such that $J(0) \in \mathrm{T}_{\gamma(0)}S$, $J(1) = 0$ and $\nabla_{\gamma'} J(0) \in \mathrm{N}_{\gamma(0)}S$. Let $\alpha : (-\varepsilon, \varepsilon) \to N$ be the geodesic $\alpha := \gamma_{J(0)}$. Since $\alpha(-\varepsilon, \varepsilon)$ is totally geodesic, Lemma 2.5.1 implies that $\gamma(b) = \exp(v)$ is focal to $\alpha(-\varepsilon, \varepsilon)$ along γ, which by

Lemma 2.5.2(a) implies that $\operatorname{focal}(\alpha(-\varepsilon,\varepsilon)) \neq \emptyset$. It follows by contraposition that (i) $\Rightarrow$ (ii).

(i) $\Leftrightarrow$ (v) Suppose that $r_f(p) < \infty$ for some $p \in N$, and let J be a non-trivial normal Jacobi field along a unit-speed geodesic $\gamma : [0, r_f(p)] \to N$ such that $J(0) = 0$ and $\|J\|'(r_f(p)) = 0$. Since $r_f(p) < r_c(p)$, $J(r_f(p)) \neq 0$. Let $\alpha : (-\varepsilon, \varepsilon) \to N$ be the geodesic $\alpha := \gamma_{J(r_f(p))}$. Let $\hat{J}$ be the non-trivial normal Jacobi field $\hat{J}(t) := (r_f(p) - t)$ along the geodesic $\hat{\gamma}$ defined by $\hat{\gamma}(t) := \gamma(r_f(p) - t)$. Then $\hat{(}J)(r_f(p)) = 0$ and

$$g(\alpha'(0), \nabla_{\hat{\gamma}'} \hat{J}(0)) = g(\hat{J}(0), \nabla_{\hat{\gamma}'} \hat{J}(0)) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} g(\hat{J}(t), \hat{J}(t)) = \|J\|'(r_f(p))$$

Since $\alpha(-\varepsilon,\varepsilon)$ is totally geodesic, it follows from Lemma 2.5.1, equation (2.1), and Lemma 2.5.2(a) that $\operatorname{focal}(\alpha(-\varepsilon,\varepsilon)) \neq \emptyset$. Conversely, suppose that $\operatorname{focal}(\alpha(-\varepsilon,\varepsilon)) \neq \emptyset$ for some geodesic $\alpha : (-\varepsilon,\varepsilon) \to N$. Applying the previous reasoning in reverse, one may, without loss of generality, suppose that there exists a non-trivial normal geodesic $\gamma : [0, T] \to N$ such that $\gamma'(0) \in \mathrm{N}_{\alpha(0)}\alpha(-\varepsilon,\varepsilon)$, $J(T) = 0$, and $\|J\|'(0) = 0$. Note that the final equality is ensured by choosing T to be small enough that $\gamma(T)$ is not conjugate to $\gamma(0)$ along γ, in which case $\nabla_{\gamma'} J(0) = c\alpha'(0)$ for some $c \neq 0$. By reversing the parameterizations of γ and J, one finds that $r_f(\gamma(T)) \leq T$.

(v) $\Leftrightarrow$ (iv) Suppose $r_f(N) = \infty$. Since (i) holds, N has no conjugate points, which since N is complete and simply connected implies $\ell(N) = \infty$. It follows from Theorem 2.5.12(a) that $r(N) = \infty$. Conversely, if $r(N) = \infty$, Theorem 2.5.12(a) implies that $r_f(N) = \infty$.

(iv) $\Leftrightarrow$ (vi) If $r(N) = \infty$, then Corollary 2.5.6(b) states that each $\nabla^2 \mathrm{d}(\cdot, p)$ is positive-definite. Conversely, if each $\nabla^2 \mathrm{d}(\cdot, p)$ is positive-definite, then each $\mathrm{d}^2(\cdot, p)$ is strictly convex. This means that each distance ball in N is strongly convex, so $r(N) = \infty$.

(i) $\Leftrightarrow$ (iii) It's clear that (iii) $\Rightarrow$ (i). Suppose N has no focal points. The argument is similar to the proof that (i) $\Rightarrow$ (ii) in Theorem 3.1.1. It will use the following claim: If $p \in M$ and $\gamma : [0, T] \to N$ is a geodesic starting at p, then $\mathrm{d}_N(\exp(v), p) \geq T$ for all $v \in \mathrm{N}_{\gamma(T)}\gamma(\mathbb{R})$. To see this, first note that $\gamma_v : \mathbb{R} \to N$ satisfies $\gamma_v(0) = \gamma(T)$. It follows from the first variation formula that $t = 0$ is a critical point of the function $t \mapsto \mathrm{d}_N^2(\gamma_v(t), p)$. By (vi), $\mathrm{d}_N^2(\cdot, p)$ is strictly convex, so $\mathrm{d}_N^2(\gamma_v(0), p) = T^2$ is the strict global minimum of $\mathrm{d}_N^2(\cdot, p)$ along γ_v. The claim follows.

Since N is complete and $\operatorname{focal}(\gamma(\mathbb{R})) = \emptyset$, $\exp|_{\mathrm{N}\gamma(\mathbb{R})} : \mathrm{N}\gamma(\mathbb{R}) \to N$ is well-defined and a local diffeomorphism. When $\mathrm{N}\gamma(\mathbb{R})$ is endowed with the pull-back metric from $\exp|_{\mathrm{N}\gamma(\mathbb{R})}$, $\exp|_{\mathrm{N}\gamma(\mathbb{R})}$ is a local isometry. Write $p := \gamma(0)$. If $v^k \in \mathrm{N}\gamma(\mathbb{R})$ is a Cauchy sequence with respect to this metric, then there exists $R > 0$ such that $\mathrm{d}_{\mathrm{N}\gamma(\mathbb{R})}(v^k, 0_p) \leq R$ for all k. From the above claim, one sees that

$\pi(v^k) \in \gamma([-R,R])$ for all k. Since all geodesics in N are minimal, the triangle inequality implies that

$$v^k \in \{v \in \mathrm{N}\gamma(\mathbb{R}) \,\big|\, \pi(v) \in \gamma([-R,R]) \text{ and } \|v\| \leq 2R\}$$

By compactness, v^k contains a convergent subsequence and, consequently, converges. Thus $\mathrm{N}\gamma(\mathbb{R})$ is complete. The proof is completed by invoking Lemma 2.2.2.

□

Remark 4.1.2. Theorem 4.1.1 is widely known. Conditions (iii) and (vi) were shown by Eberlein [Eb1] and (iv), which is equivalent to Theorem 1.1(b), by O'Sullivan [O'S1]. It's worth pointing out that (iii) is a rather ancient synthetic condition: Given any point $p \in N$ and any geodesic γ not containing p, there exists a unique geodesic from p that meets γ perpendicularly.

Remark 4.1.3. The fact that (i) $\Rightarrow$ (iii) is a special case of a theorem of Hermann [Her2], who showed that, if N is a complete Riemannian manifold and $S \subseteq N$ a closed and connected submanifold with $\mathrm{focal}(S) = \emptyset$, then $\exp|_{\mathrm{N}S} : \mathrm{N}S \to N$ is a covering map.

By contrast, when N is complete and simply connected and has non-positive curvature, $\mathrm{d} : M \times M \to [0,\infty)$ is a strictly convex function. Equivalently, whenever $\alpha, \beta : [a,b] \to M$ are geodesics, $t \mapsto \mathrm{d}(\alpha(t), \beta(t))$ is strictly convex. It was pointed out in [O'S2] that this fails to hold for surfaces with any positive curvature. Still, it turns out that this distance function along asymptotic geodesics may be controlled. A key result that has allowed many theorems about manifolds with non-positive curvature to be generalized to those with no focal points is the flat strip theorem of O'Sullivan [O'S2] and, independently, Eschenburg [Esc].

Theorem 4.1.4. *(O'Sullivan, Eschenburg) Let N be a complete and simply connected Riemannian manifold with no focal points and $\alpha, \beta : \mathbb{R} \to N$ distinct unit-speed geodesics such that $s \mapsto \mathrm{d}_N(\alpha(s), \beta(s))$ is bounded. Then $\mathrm{d}_N(\alpha(s), \beta(s)) = c$ for some $c > 0$ and all $s \in \mathbb{R}$, and the map $V : \mathbb{R} \times [0,c] \to N$ defined by*

$$V(s,t) := \exp_{\alpha(s)}\left(t \cdot \exp^{-1}_{\alpha(s)}(\beta(s))\right)$$

is a totally geodesic embedding. Consequently, $\alpha(\mathbb{R})$ and $\beta(\mathbb{R})$ bound a totally geodesic, embedded, and flat strip of surface in N.

Remark 4.1.5. One may, without loss of generality, reparameterize β so that the Jacobi field $s \mapsto \exp^{-1}_{\alpha(s)}(\beta(s))$ is perpendicular and V is an isometry.

The proofs in [O'S2] and [Esc] use results of Eberlein [Eb3] that certain Jacobi fields in N must be parallel. O'Sullivan's original argument assumed that, for any non-trivial Jacobi field J along a geodesic $\gamma : [0,\infty) \to N$ that satisfies $J(0) = 0$, $\|J(t)\| \to \infty$ as $t \to \infty$. After his article was submitted, but before it was published, Goto [Got] showed that this is always the case. Eschenburg's argument required no such assumption. Versions of the flat strip theorem were proved earlier by Green [Gr3] for surfaces with no focal points and Eberlein–O'Neill [EO'N] for manifolds of any dimension with non-positive sectional curvature. Burns [Burn1] showed that the flat strip theorem may fail for surfaces with no conjugate points.

O'Sullivan [O'S2] used the flat strip theorem to prove a powerful result, known as the flat torus theorem.

Theorem 4.1.6. *(O'Sullivan) Let N be a complete Riemannian manifold with no focal points, $p \in N$, $\overline{p} \in \pi^{-1}(p) \subseteq \overline{N}$, and $G \leq \pi_1(N,p)$ an Abelian group of rank m. Then the following hold:*

(a) $\min(G) := \bigcap_{g \in G} \min(g) \subseteq \overline{N}$ *is strongly convex and isometric to $D \times \mathbb{R}^m$ for some closed and strongly convex $D \subseteq \overline{N}$,* $\min(G)$ *is invariant under the action of G on $\overline{N}$, and the action of G on* $\min(G) \cong D \times \mathbb{R}^m$ *is by translation on the $\mathbb{R}^m$ fibers, so that, for each $g \in G$ and $(p,x) \in D \times \mathbb{R}^m$, $g \cdot (p,x) = (p, g \cdot x)$;*

(b) *If $\overline{F}$ is any $\mathbb{R}^m$-fiber of* $\min(G) \cong D \times \mathbb{R}^m$*, then $\overline{F}/G$ is a flat torus T^m, and the restriction of π to $\overline{F}$ descends to an isometric and totally geodesic immersion $\iota : T^m \to N$ satisfying $\iota_*(\pi_1(T^m, x)) \cong G$ for any $x \in T^m$; and*

(c) *If N is compact, then $D \neq \emptyset$.*

Remark 4.1.7. O'Sullivan stated all of the above results for compact manifolds, but it's long been known that parts (a) and (b) hold for all complete with no focal points. His proof can effectively be split into two halves, one that shows (a) and (b) whenever $\min(G) \neq \emptyset$ and the other that shows $\min(G) \neq \emptyset$ whenever N is compact.

The flat torus theorem was first proved in the case of non-positive curvature by Gromoll–Wolf [GW] and, independently, Lawson–Yau [LY]. As a corollary, O'Sullivan showed that, when N is compact, every solvable subgroup of $\pi_1(N,p)$ is conjugate to the fundamental group of a compact flat manifold isometrically and totally geodesically immersed in N. This result is also often referred to as the flat torus theorem in the literature. It's worth pointing out that the arguments in this dissertation only use Theorem 4.1.6, not the statement about solvable subgroups. That said, the question about solvable subgroups has an interesting history of its own. It implies that every solvable subgroup of $\pi_1(N,p)$ is a Bieberbach group and, as a consequence of the following result of O'Sullivan [O'S1], that N is flat if and only if each $\pi_1(N,p)$ is solvable.

Theorem 4.1.8. *Let N be a compact Riemannian manifold with no focal points and $p \in N$. Then the following hold:*

(a) *For some $0 \le k \le \dim(N)$, $Z := \mathrm{Z}(\pi_1(N,p))$ is isomorphic to $\mathbb{Z}^k$;*

(b) *$\overline{N}$ is isometric to $\mathbb{R}^k \times N^*$, where N^* is a simply connected Riemannian manifold with no focal points, and the action of Z on $\overline{N} \cong \mathbb{R}^k \times N^*$ is by translation on each $\mathbb{R}^k$-fiber, so that $g \cdot (x,p) = (g \cdot x, p)$ for each $g \in Z$;*

(c) *M is foliated by totally geodesic and flat k-toruses, each of which is the image under π of an $\mathbb{R}^k$-fiber of $\overline{N} \cong \mathbb{R}^k \times N^*$; and*

(d) *There exist a flat torus T^k, a complete Riemannian manifold $\hat{N}$, and a Riemannian covering map $\psi : T^k \times \hat{N} \to N$ with Abelian deck transformation group $\hat{\Gamma}$ such that, for any $(\hat{x},\hat{p}) \in \psi^{-1}(p)$, $H := (\psi \circ \iota_{\hat{x}})_*(\pi_1(\hat{N},\hat{p}))$ is a normal subgroup of $\pi_1(N,p)$ containing the commutator subgroup $C := [\pi_1(N,p),\pi_1(N,p)]$ and each of the following sequences is exact:*

$$0 \to Z \times H \to \pi_1(N,p) \to \hat{\Gamma} \to 0$$

$$0 \to Z \times (H/C) \to H_1(N,\mathbb{Z}) \to \hat{\Gamma} \to 0$$

Theorem 4.1.8 is known as the center theorem. This was first proved by Wolf [Wo] in the case of non-positive sectional curvature. In the statement of part (d), $H_1(N,\mathbb{Z})$ denotes the first homology group with integer coefficients, and certain canonical identifications are being employed in the short exact sequences.

This history of this subject dates back at least to the theorem of Preissmann [P] that any Abelian subgroup of the fundamental group of a compact and negatively curved manifold is cyclic. Preissmann's result was generalized by Byers [By] to solvable subgroups. The flat torus theorem is most properly viewed as a continuation of this work. The theorem, as described by Jost [J2], developed over a long period. Al'ber proved it in the case of rank two Abelian subgroups [Al1], but his work may have been missed by Western mathematicians for some time. He also obtained results similar to those of Hartman in the case of negative curvature, and he made the first explicit connection between the existence results of Eells–Sampson and Preissmann [Al2]. Yau [Y] proved that any solvable subgroup of $\pi_1(N, y)$, where N is compact and has non-positive curvature, must be a Bieberbach group. Shortly thereafter, the general case of the flat torus theorem for solvable subgroups of $\pi_1(N, y)$, again for compact and non-positively curved N, was obtained independently by Gromoll–Wolf [GW] and Lawson–Yau [LY]. It was this work that O'Sullivan generalized to compact manifolds with no focal points. Gromoll–Wolf and, in the analytic case, Lawson–Yau also proved that, whenever the fundamental group of a compact manifold with non-negative sectional curvature has trivial center and splits as a product, then the manifold splits as an isometric product.

Although Kleiner [Kle] has showed that the flat torus theorem may fail to hold for compact manifolds with no conjugate points, it's interesting to note that Yau's result about solvable subgroups does generalize. After O'Sullivan's result, the progress on this latter question was incremental. Croke–Schroeder [CrS] proved it for analytic metrics; they also showed that, in the smooth case, every nilpotent subgroup is Abelian. The smooth case was ultimately resolved by Kleiner [Kle] and, independently, Lebedeva [Leb]. Ivanov–Kapovitch [IK] have since proved stronger results about the structure of the fundamental group of a compact manifold with no conjugate points.

As this dissertation links the flat torus theorem to the existence results of Eells–Sampson, it can be viewed as a furtherance of Al'ber's observation about Preissmann's theorem. The methods are different, however, as Al'ber's approach was based in the calculus of variations. It's worth noting that Hansen [Han], in something of an early spiritual predecessor to this work, pointed out that Al'ber's results could be obtained using only general topology and Preissmann's theorem. Jost [J1] further developed these ideas when he used the existence results of Eells–Sampson to give a new proof of the flat torus theorem for solvable subgroups in the case of non-positive curvature. The main new result in this dissertation, that the flat torus theorem may be used as part of a proof of the Eells–Sampson existence results, is essentially the converse of what Jost showed.

Lemma 4.1.9. *Let N be a complete Riemannian manifold with no focal points and $p \in N$. Suppose that $G = <[\sigma_1], \ldots, [\sigma_n]>$ is a maximal Abelian subgroup of $\pi_1(N, p)$. Let $D \subseteq \overline{N}$ be a non-empty,* *closed, and strongly convex set as in Theorem 4.1.6. Then* $\psi_{[\sigma_1],\dots,[\sigma_n]}|_D$ *is injective.*

Proof. Assume $\psi_{[\sigma_1],\dots,[\sigma_n]}|_D$ is not injective, and choose any $\overline{p},\overline{q} \in D$ such that $\psi_{[\sigma_1],\dots,[\sigma_n]}(\overline{p}) = \psi_{[\sigma_1],\dots,[\sigma_n]}(\overline{q})$. Since D is convex, and consequently path-connected, there exists a path $\overline{\alpha} : [a,b] \to D$ from $\overline{p}$ to $\overline{q}$. Under the identification $D \cong D \times \{0\} \subseteq D \times \mathbb{R}^m$, $\tilde{\alpha} := \psi_{[\sigma_1],\dots,[\sigma_n]} \circ \overline{\alpha}$ is a homotopically non-trivial loop in $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$. Since G is maximal, $\mathrm{Z}(G) = G$. Since $(\pi_{[\sigma_1],\dots,[\sigma_n]})_*([\tilde{\alpha}]) \in \mathrm{Z}(G)$, it must be that $[\tilde{\alpha}] \in G$. However, the action of G on $D \times \mathbb{R}^k$ is by translation in the $\mathbb{R}^k$-fibers, which since $[\tilde{\alpha}] \neq 0$ implies that $\overline{\alpha}(b) \notin D \times \{0\}$. This is a contradiction.

□

Theorem 4.1.10. *Let N be a compact Riemannian manifold with no focal points and* $p \in N$*. Suppose that* $[\sigma_1],\dots,[\sigma_n] \in \pi_1(N,p)$ *generate an Abelian subgroup* $G = <[\sigma_1],\dots,[\sigma_n]>$ *of rank m. Then* $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$ *is a non-empty, compact, convex, and locally convex subset of* $\tilde{N}_{[\sigma_1],\dots,[\sigma_n]}$*. If G is a maximal Abelian subgroup, then* $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$ *is isometric to* $C \times T^m$*, where* $C \subseteq \tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$ *is strongly convex.*

Proof. Let $D \subseteq \overline{N}$ be a non-empty, closed, and strongly convex set as in Theorem 4.1.6, so that, in particular, $\min(G)$ is isometric to $D \times \mathbb{R}^m$. If $(\overline{p},\overline{x}) \in D \times \mathbb{R}^m$, then $[\sigma_i] \cdot (\overline{p},\overline{x}) = (\overline{p},[\sigma_i] \cdot \overline{x})$, and since $(\overline{p},\overline{x}) \in \min(G)$ the vector $\overline{w}_i := (0,[\sigma_i] \cdot \overline{x} - \overline{x}) \in \mathrm{T}_{(\overline{p},\overline{x})}(D \times \mathbb{R}^m) \cong \mathrm{T}_{\overline{p}}D \times \mathrm{T}_{\overline{x}}\mathbb{R}^m$ descends via π_* to the initial vector of a closed geodesic in $[\sigma_i]$. Therefore, $\psi_{[\sigma_1],\dots,[\sigma_n]}(\tilde{p},\overline{x}) = (\pi_*(\overline{w}_1),\dots,\pi_*(\overline{w}_n)) \in \tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$. On the other hand, if $(w_1,\dots,w_n) \in \tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$, then each w_i is the initial vector of a closed geodesic in $[\sigma_i]$, so $\psi^{-1}_{[\sigma_1],\dots,[\sigma_n]}(w_1,\dots,w_n) \in \min(G)$. Therefore,

$$\tilde{C}_{[\sigma_1],\dots,[\sigma_n]} = \psi_{[\sigma_1],\dots,[\sigma_n]}(D \times \mathbb{R}^m) \neq \emptyset$$

Since there exists a covering map $\tilde{N}_{[\sigma_1]} \to \tilde{N}_{[\sigma_1],\dots,[\sigma_n]}$ that maps the set $\tilde{C}_{[\sigma_1]}$ onto a set containing $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$, Theorem 3.2.10 implies that $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$ is compact. If $\tilde{v},\tilde{w} \in \tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$, one may choose $\overline{v},\overline{w} \in D \times \mathbb{R}^m$ such that $\psi_{[\sigma_1],\dots,[\sigma_n]}(\overline{v}) = \tilde{v}$ and $\psi_{[\sigma_1],\dots,[\sigma_n]}(\overline{w}) = \tilde{w}$. Since $D \times \mathbb{R}^m$ is strongly convex, there exists a unique geodesic path $\overline{\alpha}$ from $\overline{v}$ to $\overline{w}$, and $\tilde{\alpha} := \overline{\alpha}$ is a geodesic path in $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$ connecting $\tilde{v}$ to $\tilde{w}$. Hence $\tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$ is convex.

Let $v \in \tilde{C}_{[\sigma_1],\dots,[\sigma_n]}$, and choose $\overline{v} \in \overline{N}$, an open set $U \subseteq \overline{N}$ containing $\overline{v}$, and an open set $V \subseteq \tilde{N}_{[\sigma_1],\dots,[\sigma_n]}$ containing v such that $\psi_{[\sigma_1],\dots,[\sigma_n]}|_U : U \to V$ is an isometry. It will be shown that, for any deck transformation $\overline{\gamma}$ of $\psi_{[\sigma_1],\dots,[\sigma_n]}$, $\overline{\gamma}(U \cap (D \times \mathbb{R}^k)) = \overline{\gamma}(U) \cap (D \times \mathbb{R}^k)$. Choose any point

$\overline{q} \in U \cap (D \times \mathbb{R}^k)$. Then $\psi_{[\sigma_1],\ldots,[\sigma_n]}(\overline{q}) \in \tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$, so $\psi_{[\sigma_1],\ldots,[\sigma_n]}(\overline{q}) = (u_1, \ldots, u_n)$, where each γ_{u_i} is a closed geodesic in $[\sigma_i]$. Write $q := \psi_{[\sigma_1],\ldots,[\sigma_n]}(\overline{q})$. By Theorem 3.2.5(iii), $\overline{\gamma}$ is identified with an element of $(\psi_{[\sigma_1],\ldots,[\sigma_n]})_*\Big(\pi_1(\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}, (u_1, \ldots, u_n))\Big) = \mathrm{Z}([\gamma_{u_1}], \ldots, [\gamma_{u_n}])$, and any path connecting $\overline{q}$ to $\overline{\gamma} \cdot \overline{q}$ descends via π to a loop $\gamma \in \mathrm{Z}([\gamma_{u_1}, \ldots, [\gamma_{u_n}])$. Therefore, $[\gamma]^{-1}[\gamma_{u_i}][\gamma] = [\gamma_{u_i}]$, which since N has no conjugate points implies that $\psi_{[\gamma_1],\ldots,[\gamma_n]} \circ \overline{\gamma}(\overline{q}) = (u_1, \ldots, u_n) = \psi_{[\gamma_1],\ldots,[\gamma_n]}(\overline{q})$. So $\overline{\gamma}(\overline{q}) \in \overline{\gamma}(U) \cap (D \times \mathbb{R}^k)$, and it's been shown that $\overline{\gamma}(U \cap (D \times \mathbb{R}^k)) \subseteq \overline{\gamma}(U) \cap (D \times \mathbb{R}^k)$. Applying the same argument to $\overline{\gamma}^{-1}$ shows that $\overline{\gamma}(U) \cap (D \times \mathbb{R}^k) \subseteq \overline{\gamma}(U \cap (D \times \mathbb{R}^k))$. So $\overline{\gamma}(U \cap (D \times \mathbb{R}^k)) = \overline{\gamma}(U) \cap (D \times \mathbb{R}^k)$. By shrinking U and V, one may now suppose that $U = \mathrm{B}(\overline{v}, \varepsilon)$ for some $0 < \varepsilon \leq r(\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}(v))$ and, moreover, that $U \cap (D \times \mathbb{R}^k)$ is strongly convex. Then $V \cap \tilde{C}_{[\sigma_1],\ldots,[\sigma_n]} = \psi_{[\sigma_1],\ldots,[\sigma_n]}(U \cap (D \times \mathbb{R}^k))$ is strongly convex. So $\tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$ is locally convex.

In the case that G is a maximal Abelian subgroup of $\pi_1(N, p)$, $\mathrm{Z}(G) = G$. Since $\mathrm{Z}(G)$ is, up to the canonical identification, the deck transformation group of $\psi_{[\sigma_1],\ldots,[\sigma_n]}$, which in this case acts on $D \times \mathbb{R}^m$ by translation in the $\mathbb{R}^m$-factor, $\psi_{[\sigma_1],\ldots,[\sigma_n]}(D \times \mathbb{R}^k)$ is isometric to $D \times \mathbb{R}^m / G \cong D \times T^m$. At the same time, by Lemma 4.1.9, $\psi_{[\sigma_1],\ldots,[\sigma_n]}$ is injective on D, so its image $C := \psi_{[\sigma_1],\ldots,[\sigma_m]}(D) \neq \emptyset$ is a convex subset of $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ contained in $\tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$, and $\psi_{[\sigma_1],\ldots,[\sigma_n]}|_D : D \to C$ is an isometry with respect to the intrinsic metric on C. It will be shown that C is strongly convex. For any $\tilde{p}, \tilde{q} \in C$, there is exactly one geodesic $\tilde{\alpha} : [a, b] \to C$ connecting $\tilde{p}$ to $\tilde{q}$. It must be shown that $\tilde{\alpha}$ is minimal in $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$. Let $\tilde{\beta} : [a, b] \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ be any minimal geodesic from $\tilde{p}$ to $\tilde{q}$. Write $\overline{p} := \psi_{[\sigma_1],\ldots,[\sigma_n]}|_D^{-1}(\tilde{p})$, and lift $\tilde{\alpha}$ and $\tilde{\beta}$ via $\psi_{[\sigma_1],\ldots,[\sigma_n]}$ to geodesics $\overline{\alpha}$ and $\overline{\beta}$, respectively, based at $\overline{p}$. Then $\overline{\beta}$ is a minimal geodesic. By construction, $\overline{\alpha}([a, b]) \subseteq D$ and $\overline{\beta}(b) \in \psi^{-1}_{[\sigma_1],\ldots,[\sigma_n]}(\tilde{q}) \subseteq D \times \mathbb{R}^k$. Since $D \times \mathbb{R}^m$ is strongly convex, one must have that $\overline{\beta}([a, b]) \subseteq D \times \mathbb{R}^k$. Since $D \times \mathbb{R}^k$ has a product metric, $\overline{\beta}([a, b]) \subseteq D$. Since D is strongly convex, $\overline{\beta} = \overline{\alpha}$ and, consequently, $\tilde{\beta} = \tilde{\alpha}$. So C is strongly convex. It follows that $\tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$ is isometric to $C \times T^m$.

□

Remark 4.1.11. When the Abelian subgroup G in Theorem 4.1.10 is not maximal, one still has that, in the covering space $\overline{N}/G$, which lies above $\tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$, the image of $D \times \mathbb{R}^k$ under the quotient map $\overline{N} \to \overline{N}/G$ is a closed and strongly convex set that splits isometrically as $D \times T^m$.

Example 4.1.12. When N is complete but non-compact, the set $\tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$ may be empty. For instance, the cylinder $S^1 \times \mathbb{R}$ can be endowed with a complete and negatively curved metric, as a surface of rotation, that has no closed geodesics.

Corollary 4.1.13. *Let T^n be a flat Riemannian torus and N a compact Riemannian manifold with no focal points. If $f : T^n \to N$ is continuous, then f is homotopic to a totally geodesic map.*

Proof. This almost follows from Theorem 4.1.6 and Theorem 3.3.3, but it's not in general the case that the totally geodesic immersion ι guaranteed by the flat torus theorem is defined on a torus of the same dimension as T^n. Instead, it follows from Theorem 4.1.10 and Theorem 3.3.3.

□

Corollary 4.1.14. *Let T^n be a flat Riemannian torus, N a complete Riemannian manifold with no focal points, and $[F]$ a homotopy class of maps from T^n to N. Then the set of totally geodesic maps in $[F]$ is path-connected.*

Proof. Let $f_1, f_2 \in [F]$ be totally geodesic. Fix $x \in T^n$, and choose a minimal generating set $\{[\varsigma_1], \ldots, [\varsigma_n]\}$ for $\pi_1(T^n, x)$. Write $[\sigma_i] := (f_1)_*([\varsigma_i])$. Since f_1 and f_2 are homotopic, one has that $g_*([\varsigma_1], \ldots, [\varsigma_n]) \cong ([\sigma_1], \ldots, [\sigma_n])$. By Lemma 3.2.8, each f_i lifts canonically to a map $\tilde{F}_i : T^n \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$. Since each f_i is totally geodesic, $\tilde{F}_i(T^n) \subseteq \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$. By Theorem 4.1.10, $\tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$ is convex. For any geodesic $\tilde{\alpha} : [a,b] \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_n]}$ from $\tilde{F}_1(x)$ to $\tilde{F}_2(x)$ such that $\tilde{\alpha}([a,b]) \subseteq \tilde{C}_{[\sigma_1],\ldots,[\sigma_n]}$, $\Upsilon_{\tilde{\alpha}}$ is a homotopy from f_1 to f_2 through totally geodesic maps.

□

A theorem of Lemaire [Lem] and, independently, Sacks–Uhlenbeck [SaU] states that, if f is a map from a compact orientable surface into a compact manifold N, $\dim(N) \geq 3$, $\pi_2(N, y) = 0$, and f_* is injective, then f is homotopic to an energy-minimizing map. In particular, this holds for any map from a two-dimensional flat torus into a compact manifold with no conjugate points whose induced homomorphism is injective. If N is any compact manifold with no conjugate points, $\dim(N) \geq 3$, and $< [\sigma_1], [\sigma_2] > \leq \pi_1(N, p)$ is an Abelian subgroup of rank two, then, for the unique initial vectors v_i of closed geodesics $\gamma_{v_i} \in [\sigma_i]$, the loop map $\Upsilon_{v_1 \otimes v_2}$ exists and has injective induced homomorphism $(\Upsilon_{v_1 \otimes v_2})_*$. As $\pi_k(N, y) = 0$ for all $k \geq 2$, it follows that $\Upsilon_{(v_1, v_2)}$ is homotopic to an energy-minimizing map. Since Kleiner [Kle] has shown that the flat torus theorem may fail for compact manifolds with no conjugate points in any dimension greater than two, one obtains the following surprising result.

Theorem 4.1.15. *For each $n \geq 3$, there exist a flat torus T^2, a compact n-dimensional Riemannian manifold N with no conjugate points, and an energy-minimizing map $T^2 \to N$ that is not totally geodesic.*

This contrasts with the situation for compact manifolds with no focal points, where, as shown in Theorem 1.4(b)-(c), every energy-minimizing map from a flat torus is totally geodesic. It is apparently, and rather surprisingly, an open question whether every homotopy class of maps $T^n \to N$, where $n > 3$ and N is a compact manifold with no conjugate points, contains an energy-minimizing representative.

4.2 Heat flow methods

It will be instructive to detour from the proof of Theorem 1.4 and Theorem 1.5 and explore what can be said about the existence, uniqueness, and convergence of solutions to the Eells–Sampson heat equation (1.1) when the target is a compact Riemannian manifold with no focal points. Specifically, it will be shown that, for compact domains, solutions to (1.1) exist for all time, are unique, and uniformly subconverge to harmonic maps. The results of this section won't be used elsewhere in the dissertation, but their limitations help illustrate the need for the work to come in later chapters.

Recall that Eells–Sampson [ES] proved the short-term existence and uniqueness of solutions $u : M \times [0,\infty) \to N$ to (1.1) when M and (N,h) are arbitrary compact Riemannian manifolds. When N has non-positive sectional curvature, they proved long-term existence and uniqueness, as well as uniform subconvergence to harmonic maps. To say that u **subconverges** means that there exists a sequence of times $t_i \to \infty$ such that $u(\cdot,t_i)$ converges.[1] A key point in their argument is that the energy of the map under the flow is a non-increasing and convex function of time, the convexity being a consequence of the curvature assumption on N. When N is only assumed to have no focal points, the examples of Gulliver [Gul] show that the sectional curvatures of N may be unbounded, in which case this convexity may not hold and the argument falls apart.

A great deal of effort over the years has gone into determining whether and how solutions to (1.1) may blow up, which is to say, when they may fail to exist for all time or to converge in the limit. The history of this problem will be sketched briefly, with the caveat that this discussion is far

[1] Actually, their methods show that a stronger statement holds, namely, for each sequence of times $t_i \to \infty$, there exists a subsequence t_{i_j} such that $u(\cdot,t_{i_j})$ converges to a smooth harmonic map.

from comprehensive and omits mention of great work done by many authors. In the case of $M = S^1$, Ottarsson [Ot] proved long-term existence and uniform subconvergence to closed geodesics for all compact targets N. The work of rigorously defining the heat flow on domains with boundary, including proving short-term existence of solutions to the Dirichlet problem, was done by Hamilton [Ham], who also generalized the results of [ES] to this case. The great breakthrough in this area was provided by Struwe [St1], who showed that, when M is a compact surface, a weak solution $u : M \times [0, \infty) \to N$ to (1.1) exists and is smooth except possibly at finitely many points, around each of which a bubbling phenomenon similar to that described in [SaU] occurs. Roughly, at a singular time $(x, t) \in M \times (0, \infty)$, one may find a sequence of times $t_k \nearrow t$ and radiuses $r_k \searrow 0$ such that, after rescaling the metrics on $\mathrm{B}(x, r_k)$ appropriately, the maps $u(\cdot, t_k)|_{\mathrm{B}(x,r_k)}$ converge to a map $u_\infty : \mathbb{R}^2 \to N$ that extends by stereographic projection to a non-constant harmonic map $\overline{u}_\infty : S^2 \to N$. A harmonic map $S^m \to N$ is called a **harmonic sphere**. When all harmonic spheres from S^2 into N are constant, the long-term existence, uniqueness, and uniform subconvergence to harmonic maps of solutions to (1.1) follows. Struwe's results were generalized to domains with boundary by Chang [Chan].

The first examples of blow-up were provided by Coron-Ghidaglia [CoG] in dimensions greater than two and, later, by Chang–Ding-Ye [CDY] in dimension two. Many subsequent examples have been constructed, including by Chen-Ding [CD] and Topping. Topping's examples are notable, as they show that the flow may fail to converge in the limit without any bubbling ever occurring [Topp1], develop multiple bubbles at the same point at infinite time [Topp2], and converge to a discontinuous function at the first finite-time singularity [Topp3]. Chen–Struwe [ChS] generalized the existence of a weak solution $u : M \times [0, \infty) \to N$ to the case of $n = \dim(M) \geq 3$ and proved its regularity away from a singular set of locally finite n-dimensional Hausdorff measure. In this case, Lin–Wang [LW] showed that singularities may occur only if N admits a harmonic sphere from S^m for some $2 \leq m \leq n - 1$ or a non-constant harmonic map from $(\mathbb{R}^m, e^{\frac{-|x|^2}{2(m-2)}} \mathrm{d}s^2)$ with finite energy, where $3 \leq m \leq n$ and $\mathrm{d}s^2$ is the Euclidean metric on $\mathbb{R}^m$. A harmonic map from $(\mathbb{R}^m, e^{\frac{-|x|^2}{2(m-2)}} \mathrm{d}s^2)$ with finite energy is called a **quasi-harmonic sphere**. A result in this direction was also proved by Struwe [St2]. Using their result, Lin–Wang showed that, when all harmonic and quasi-harmonic spheres into N are constant, classical solutions to (1.1) exist for all time.

Following Gordon [Gor], a Riemannian manifold is called **convex-supporting** if every compact subset has an open neighborhood that admits a function with positive-definite Hessian. Gordon showed that any harmonic map from a compact Riemannian manifold M with finite fundamental group into a Riemannian manifold N that has a convex-supporting covering space must be constant. In particular, if each $\pi_1(M,x)$ is finite and $\overline{N}$ admits such a function, any harmonic map $M \to N$ must be constant. This precludes the existence of non-constant harmonic spheres. To prove the long-term existence and regularity of the heat flow, it remains to show that N doesn't admit non-constant quasi-harmonic spheres. This was done by Li–Zhu [LZ], who also applied techniques of Li–Tian [LT] and Lin [Li] to prove that globally defined heat flows subconverge in $\mathrm{C}^2(M,N)$ to harmonic maps.[2]

Theorem 4.2.1. *(Li–Zhu) Let N be a compact Riemannian manifold. Suppose that $\overline{N}$ admits a* C^2 *function $\rho : \overline{N} \to [0,\infty)$ such that $\nabla^2\rho$ is everywhere positive-definite and $\rho(\overline{y}) \leq C(1+\mathrm{d}(\overline{y},\overline{y}_0))^k$ for some $\overline{y}_0 \in \overline{N}$, $C \geq 0$, and $k \geq 1$ and all $\overline{y} \in \overline{N}$. Then the following hold:*

(a) *Every quasi-harmonic sphere into N is constant; and*

(b) *If M is a compact Riemannian manifold, then, given any* C^1 *map $u_0 : M \to N$, a unique solution $u : M \times [0,\infty) \to N$ to the heat equation* (1.1) *exists, is smooth on $M \times (0,\infty)$, and subconverges in* $\mathrm{C}^2(M,N)$ *to a smooth harmonic map as $t \to \infty$.*

Theorem 4.2.1(a) improves upon a result of Ding–Lin [DL], which required a stronger assumption on $\nabla^2\rho$. Part (b) implies, in particular, uniform subconvergence to harmonic maps. Combining part (b) with Theorem 4.1.1, one obtains the corresponding result when N has no focal points.

Corollary 4.2.2. *Let M and N be compact Riemannian manifolds, where N has no focal points. Given any* C^1 *map $u_0 : M \to N$, a unique solution $u : M \times [0,\infty) \to N$ to the heat equation* (1.1) *exists, is smooth on $M \times (0,\infty)$, and subconverges in* $\mathrm{C}^2(M,N)$ *to a smooth harmonic map as $t \to \infty$.*

Note that the limit map of the flow in Corollary 4.2.2 is not known to minimize energy. Using the direct method, rather than a geometric flow, Xin [X2] showed that, when N has no focal points, every homotopy class of maps $M \to N$ contains an energy-minimizing harmonic representative. The argument essentially combines Theorem 4.1.1, Gordon's result, and a regularity theorem of Schoen–Uhlenbeck [ScU]. A much more general existence result of this form, for homotopy classes of maps

[2] Li–Zhu make no mention of uniqueness in the statement of their results, but this may be taken for granted, as in [ES] it's shown that classical solutions are unique for as long as they exist. Similarly, though Li–Zhu only state their result for smooth initial data, the short-term existence for C^1-initial data shown in [ES], when combined with the result of Li–Zhu, implies long-term existence.

between metric measure spaces, is given in [J3], where references to other work in this direction may be found. Applying a theorem of Alexander–Bishop [AB], Burns [Burn2] showed that any complete and simply connected surface with no conjugate points is convex-supporting. Along with a result of Eells [Ee], which similarly combined the work of Gordon and Schoen–Uhlenbeck, this shows that every homotopy class of maps $M \to N$ into a compact surface with no conjugate points contains an energy-minimizing harmonic representative.

The most significant limitation of Corollary 4.2.2, and of the many related existence results, is that the map in the limit is only known to be harmonic or, at best, energy-minimizing. Recall that Eells–Sampson also proved Theorem 1.2, which in part states that, when M has non-negative Ricci curvature and N has non-positive sectional curvature, every C^2 harmonic map $M \to N$ is totally geodesic. The proof of that fact will be sketched here. For details, one may check the original paper [ES] or the textbooks [J2] or [X1].

The proof of Theorem 1.2 depends heavily upon the Bochner identity (1.2). The notation in (1.2) is rather terse. In it, Δ denotes the Laplace-Beltrami operator, which is the divergence of the gradient. The norm $\|\beta_f\|$ at each point $p \in M$ is the operator norm on the space $\mathscr{L}(\mathrm{T}_pM \odot \mathrm{T}_pM, T_{f(p)}N)$ that may be computed as the sum of the Hilbert-Schmidt norms of the $(0,2)$-tensors, or more precisely the linear operators $\mathrm{T}_pM \to \mathrm{T}_pM$ with which they're canonically identified, that appear as the components of β_f with respect to any orthonormal basis for $\mathrm{T}_{f(p)}N$. The inner products are $<\cdot,\cdot>=<\cdot,\cdot>_{f^{-1}(\mathrm{T}N)}$. The Ricci tensor Ric_M is being identified with a $(1,1)$-tensor, and repeated subscripts in a term indicate contraction. After unpacking everything, one finds that any C^2 harmonic map $f : M \to N$ satisfies

$$\Delta e_f - \|\beta_f\|^2 = Q_f$$

where $Q_f : M \to \mathbb{R}$ at each point $x \in M$ takes the form

$$Q_f(x) = \sum_{i,j=1}^{m} h\Big(R_N(f_*(e_i), f_*(e_j))f_*(e_j), f_*(e_i)\Big) + \sum_{i=1}^{m} \mathrm{Ric}_M(e_i, e_i)\|f_*(e_i)\|^2$$

for any orthonormal basis $\{e_1, \ldots, e_m\}$ of T_xM consisting of principal directions of the Ricci tensor on M. By the curvature assumptions, Q_f is non-negative. By the divergence theorem, $\int_M \Delta e_f \,\mathrm{d}\mu_M = 0$. It follows that $\int_M \|\beta_f\|^2 \,\mathrm{d}\mu_M \leq 0$ and, consequently, $\beta_f = 0$. By Theorem 2.6.1, this is equivalent to f being totally geodesic.

When N is only assumed to have no focal points, the Bochner identity is not useful, and it's not clear at the start that energy-minimizing maps $M \to N$ are totally geodesic, nor that every homotopy class of maps contains a totally geodesic representative. This is the main technical difficulty overcome in this dissertation, by completely different methods. The more general question about whether every harmonic map $M \to N$ is totally geodesic remains open.

Recall also that Hartman [Har] improved upon the results of Eells–Sampson in Theorem 1.3. Hartman's method is to use the bound on sectional curvature, by the parabolic maximum principle as well as direct computation, to show that certain functions associated with the heat flow are non-increasing in time. It suffices to say that this method also fails when N is only assumed to have no focal points. An early proof of Theorem 1.5 that bypassed this issue used parameterized heat flows on loops. The argument was first to prove Theorem 1.4(c) by successively applying a parameterized heat flow to each of the families of loops in the expression $T^k = S^1 \times \cdots \times S^1$ and, after finite time, projecting the family being flowed onto closed geodesics. This used the long-term existence of solutions to the heat equation (1.1) shown by Ottarsson [Ot] and a new result that the heat flow from a compact manifold with boundary doesn't leave an arbitrary compact and locally convex set before the image of the boundary does. This latter result generalizes a theorem of Hamilton [Ham] and may be proved by combining the maximum principle in [Ham] with ideas of Evans [Ev] about viscosity solutions. Looking closely at the equations relating the energy of a loop with its total geodesic curvature under the heat flow, one may show, using the work of Lenbury–Maneesawarng [LM] and Karuwannapatana–Maneesawarng [KM] about length and total curvature in singular spaces with curvature bounded above, that, after finite time, the loops in each family being flowed are uniformly C^0-close to closed geodesics. After the flow is lifted to an appropriate covering space $\tilde{N}_{[\sigma_1],\ldots,[\sigma_m]}$, a theorem of Walter [Wal] about the nearest-point projection onto compact and locally convex sets ensures that it may be stopped at this time and the loops simultaneously projected onto closed geodesics in the set $\tilde{C}_{[\sigma_1],\ldots,[\sigma_m]}$. The argument proceeds inductively, using the loop map Υ and the flat strip theorem while flowing the $(i+1)$-th family to extend to a totally geodesic map along each of the already-completed T^i-fibers. Once Theorem 1.4(c) is established, Theorem 1.5 may be shown using the center of mass and a partition of unity as in the proof of Theorem 6.3.2.

Chapter 5

Commutative diagrams

5.1 Manifolds with non-negative Ricci curvature

Here and throughout, M_i and N_i, where $i = 1,2$, will be connected topological manifolds and $M_i \to M_i \times N_i \xrightarrow{\pi_i} N_i$ trivial fiber bundles, where π_i in each case is projection onto the second component and $\rho_i : M_i \times N_i \to M_i$ onto the first. The functions $\chi : M_0 \to M_1$, $\psi : M_0 \times N_0 \to M_1 \times N_1$, and $\phi : N_0 \to N_1$ will be covering maps. The spaces under study here will be those for which the diagram

$$\begin{tikzcd} M_0 \times N_0 \arrow[r, "\pi_0"] \arrow[d, "\psi"] & N_0 \arrow[d, "\phi"] \\ M_1 \times N_1 \arrow[r, "\pi_1"] & N_1 \end{tikzcd} \tag{5.1}$$

commutes. The diagram (5.1) will be said to commute **isometrically** when M_0 and the N_i and $M_i \times N_i$ are Riemannian manifolds, $M_0 \times N_0$ has the product metric obtained from $M_0 \times N_0$, and ψ and ϕ are local isometries. It's worth emphasizing, pointedly, that in this case $M_1 \times N_1$ is not assumed to have a product metric and that π_1 is not necessarily a Riemannian submersion. Intuitively, the space $M_1 \times N_1$ is obtained from $M_0 \times N_0$ by a quotient in which twisting only occurs along the M_0-fibers. When this diagram commutes, the spaces involved will be shown to possess certain nice topological properties. One such property is that there exist a covering map $\chi : M_0 \to M_1$ and a homeomorphism $\varphi : M_0 \times N_0 \to M_0 \times N_0$ such that the diagram

$$\begin{tikzcd} M_0 \times N_0 \arrow[r, "\varphi"] \arrow[d, "\chi\times\phi"] & M_0 \times N_0 \arrow[r, "\pi_0"] \arrow[d, "\psi"] & N_0 \arrow[d, "\phi"] \\ M_1 \times N_1 \arrow[r, "\mathrm{id}"] & M_1 \times N_1 \arrow[r, "\pi_1"] & N_1 \end{tikzcd} \tag{5.2}$$

commutes.

Diagrams of the form (5.1) arise naturally in relation to compact manifolds with non-negative Ricci curvature, as shown by Cheeger–Gromoll in [CG1] and [CG2]. This is an application of their famous splitting theorem.

Theorem 5.1. *(Cheeger–Gromoll) Let M be a complete Riemannian manifold with non-negative Ricci curvature. Then M is isometric to $\hat{M}\times\mathbb{R}^k$, where $\hat{M}$ admits no minimal geodesics $\gamma:\mathbb{R}\to\hat{M}$ and $\mathbb{R}^k$ has its standard flat metric.*

This generalizes a result of Toponogov [Topo] for manifolds with non-negative sectional curvature. Using Theorem 5.1, Cheeger–Gromoll showed that every compact manifold with non-negative Ricci curvature is finitely covered by a space $M_1\times T^k$ in a diagram of the form (5.1) that commutes isometrically.

Corollary 5.2. *(Cheeger–Gromoll) Let M be a compact Riemannian manifold with non-negative Ricci curvature. Then there exist Riemannian manifolds M_0 and M_1, where M_0 is compact and simply connected, a constant metric on $\mathbb{R}^k$, a flat torus T^k, and Riemannian covering maps $\psi: M_0\times\mathbb{R}^k\to M_1\times T^k$ and $\phi:\mathbb{R}^k\to T^k$ such that the diagram* (5.1) *commutes isometrically.*

Proof. This is almost exactly the statement of Theorem 9.2 of [CG2], but it's not immediately clear from their diagram that the diffeomorphism $\hat{M}\to M_1\times T^k$ makes the diagram (5.1) commute. However, following their proof, one finds that the diffeomorphism they construct is canonical and does indeed take the image of each M_0-fiber to the correct M_1-fiber.

□

Theorem 1.4 will actually be shown to hold for all M that are finitely covered by the space $M_1\times T^k$ in a diagram (5.1) that commutes isometrically, provided the homotopy class of maps $[F]$ satisfies a certain topological property which always holds when each $\pi_1(M_1,\tilde{p})$ is finite. This latter condition holds whenever M_0 is simply connected. Moreover, it was shown by Cheeger–Gromoll [CG1] that there are manifolds with non-negative Ricci curvature for which the covering map ψ_0 in the corresponding diagram (5.1) cannot be finite. Their example is to let $\mathbb{Z}$ act on $S^2\times\mathbb{R}$ by an irrational rotation on the first component and translation on the second. Then the quotient under this action is diffeomorphic to $S^2\times S^1$ but is never finitely covered by $S^2\times\mathbb{R}$. Modifying this so that the initial metric on S^1 is no longer round but still admits an isometric S^1 action, one obtains a space that satisfies diagram (5.1) but which may have negative Ricci curvature.

5.2 Geometric structure

This section is dedicated to rigorously developing the geometric structure of spaces in a diagram (5.1) that commutes. Most of the results are technical lemmas, with elementary proofs, that will be used in the next two chapters. The notation ι is used throughout to signify inclusion into either the first or second component of a product, while ρ_i indicates projection onto the i-th component. For example, if $\overline{x} \in N_0$, then $\iota_{\overline{x}} : M_0 \to M_0 \times N_0$ is the map $\iota_{\overline{x}}(\cdot) := (\cdot, \overline{x})$. Similarly, if $\overline{p} \in M_0$, then $\iota_{\overline{p}} : N_0 \to M_0 \times N_0$ is $\iota_{\overline{p}}(\cdot) := (\overline{p}, \cdot)$. The maps $\rho_1 : M_1 \times N_1 \to M_1$ and $\rho_2 : M_1 \times N_1 \to N_1$ are projection. The meaning will always be clear in context.

Lemma 5.3. *Suppose the diagram* (5.1) *commutes. Let $\overline{x} \in N_0$, and write $\tilde{x} := \phi(\overline{x})$. Then the restriction $\psi|_{M_0 \times \{\overline{x}\}} : M_0 \times \{\overline{x}\} \to M_1 \times \{\tilde{x}\}$ is a covering map.*

Proof. Fix any $\overline{x} \in N_0$, and write $\tilde{x} := \phi(\overline{x})$. Since $\phi \circ \pi_0 = \pi_1 \circ \psi$, $\psi(M_0 \times \{\overline{x}\}) \subseteq M_1 \times \{\tilde{x}\}$. To see that $\psi|_{M_0 \times \{\overline{x}\}} : M_0 \times \{\overline{x}\} \to M_1 \times \{\tilde{x}\}$ is surjective, fix $\tilde{p} \in M_1$. M_1. Choose any $(\tilde{q}, \tilde{x}) \in M_1 \times \{\tilde{x}\}$. Since M_1 is connected, there exists a path $\sigma : [a,b] \to M_1$ from $\tilde{p}$ to $\tilde{q}$. Let $\tilde{\sigma} := \iota_{\tilde{x}} \circ \sigma$, and denote by $\overline{\sigma} : [a,b] \to M_0 \times \{\overline{x}\}$ the lift of $\tilde{\sigma}$ under ψ with $\overline{\sigma}(a) = (\overline{p}, \overline{x})$. Since $\phi \circ \pi_0 \circ \overline{\sigma}(t) = \pi_1 \circ \psi \circ \overline{\sigma}(t) = \pi_1 \circ \tilde{\sigma}(t) = \tilde{z}$ for all t, $\pi_0 \circ \overline{\sigma}(t) \in \phi^{-1}(\tilde{x})$ for all t. Since $\pi_0 \circ \overline{\sigma}(a) = \overline{z}$ and $\phi^{-1}(\tilde{x})$ is a discrete subset of N_0, it follows from continuity that $\pi_0 \circ \overline{\sigma}(b) = \overline{x}$. Therefore, $\psi(\overline{\sigma}(b)) = \tilde{\sigma}(b) = (\tilde{q}, \tilde{z})$. So $\psi \circ \iota_{\overline{x}}$ is surjective. Since ψ is a covering map and $M_0 \times \{\overline{x}\}$ and $M_1 \times \{\tilde{x}\}$ are embedded submanifolds of $M_0 \times N_0$ and $M_1 \times N_1$, evenly covered neighborhoods of ψ in $M_1 \times N_1$ restrict to evenly covered neighborhoods of $\psi \circ \iota_{\overline{x}}$ in $M_1 \times \{\tilde{x}\}$. Thus $\psi \circ \iota_{\overline{x}}$ is a covering map.

□

Proposition 5.4. *Suppose the diagram* (5.1) *commutes. For any $\overline{x} \in N_0$, write $\chi_{\overline{x}} := \rho_1 \circ \psi \circ \iota_{\overline{x}}$. Then there exists a homeomorphism $\varphi : M_0 \times N_0 \to M_0 \times N_0$ such that the diagram* (5.2) *commutes when $\chi = \chi_{\overline{x}}$.*

Proof. By Lemma 5.3, each $\psi|_{M_0 \times \{\overline{z}\}} : M_0 \times \{\overline{z}\} \to M_1 \times \{\tilde{z}\}$ is a covering map. For the given $\overline{x} \in N_0$, write $\tilde{x} := \phi(\overline{x})$. Since $\rho_1|_{M_1 \times \{\tilde{x}\}} : M_1 \times \{\tilde{x}\} \to M_1$ and $\iota_{\overline{x}} : M_0 \to M_0 \times \{\overline{x}\}$ are homeomorphisms, $\chi = \chi_{\overline{x}}$ is indeed a covering map. It remains to produce the homeomorphism φ. Any such φ would have to be a lift of $\chi \times \phi$ along ψ. Let $\overline{p} \in M_0$, and write $\tilde{p} := \rho_1 \circ \phi(\overline{p}, \overline{x})$, so that $\phi(\overline{p}, \overline{x}) = (\tilde{p}, \tilde{x}) = \chi \times \phi(\overline{p}, \overline{x})$.

Since $\chi\times\phi$ and ψ are covering maps, such a lift φ will exist and be a homeomorphism exactly when $(\chi\times\phi)_*\big(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))\big)=\phi_*\big(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))\big)$. Recall that

$$(\chi\times\phi)_*\big(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))\big)=(\iota_{\tilde{x}}\circ\chi)_*(\pi_1(M_0,\overline{p}))(\iota_{\tilde{p}}\circ\phi)_*(\pi_1(N_0,\overline{x}))$$

and

$$\psi_*(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))=(\psi\circ\iota_{\overline{x}})_*(\pi_1(M_0,\overline{p}))(\psi\circ\iota_{\overline{p}})_*(\pi_1(N_0,\overline{x}))$$

It therefore suffices to show that $(\iota_{\tilde{x}}\circ\chi)_*(\pi_1(M_0,\overline{p}))=(\psi\circ\iota_{\overline{x}})_*(\pi_1(M_0,\overline{p}))$ and $(\iota_{\tilde{p}}\circ\phi)_*(\pi_1(N_0,\overline{x}))=(\psi\circ\iota_{\overline{p}})_*(\pi_1(N_0,\overline{x}))$. The first is a matter of as

$$\begin{aligned}(\iota_{\tilde{x}}\circ\chi)_*(\pi_1(M_0,\overline{p}))&=(\iota_{\tilde{x}}\circ\rho_1\circ\psi\circ\iota_{\overline{x}})_*(\pi_1(M_0,\overline{p}))\\&=(\psi\circ\iota_{\overline{x}})_*(\pi_1(M_0,\overline{p}))\end{aligned}$$

As for the second,

$$\begin{aligned}(\iota_{\tilde{p}}\circ\phi)_*(\pi_1(N_0,\overline{x}))&=(\iota_{\tilde{p}}\circ\phi\circ\pi_0)_*\big(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))\big)\\&=(\iota_{\tilde{p}}\circ\pi_1\circ\psi)_*\big(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))\big)\\&=\psi_*\big(\pi_1(M_0\times N_0,(\overline{p},\overline{x}))\big)\\&=(\iota_{\overline{p}}\circ\psi)_*(\pi_1(M_0,\overline{p}))\end{aligned}$$

Therefore, the desired homeomorphism φ exists.

$\square$

In light of Proposition 5.4, whenever the diagram (5.1) commutes, there is no loss of generality in assuming that the diagram (5.2) commutes as well.

Lemma 5.5. *Suppose the diagram* (5.1) *commutes isometrically. Then each M_1-fiber of $M_1\times N_1$ is totally geodesic.*

Proof. Let $(\tilde{p}_1,\tilde{z})\in M_1\times N_1$, and choose $(\overline{p}_1,\overline{z})\in\psi^{-1}(\tilde{p}_1,\tilde{z})$. Let $U\subseteq M_0\times N_0$ and $V\subseteq M_1\times N_1$ be open sets around $(\overline{p}_1,\overline{z})$ and $(\tilde{p}_1,\tilde{z})$, respectively, such that $\psi|_U:U\to V$ is an isometry. Shrinking U and V, if necessary, one may suppose that $\phi\circ\pi_0$ is injective on U and that V is strongly convex. Note that V intersects $M_1\times\{\tilde{z}\}$ in an open subset of $M_1\times\{\tilde{z}\}$ containing $(\tilde{p}_1,\tilde{z})$. Let $(\tilde{p}_2,\tilde{z})\in V$, and fix a minimal geodesic $\tilde{\alpha}:[a,b]\to V$ from $(\tilde{p}_1,\tilde{z})$ to $(\tilde{p}_2,\tilde{z})$. Let $\overline{\alpha}:[a,b]\to U$ be the lift of $\tilde{\alpha}$ based at $(\overline{p}_1, \overline{z})$. Then $\overline{\alpha}$ is a minimal geodesic and, since $\phi \circ \pi_0 = \pi_1 \circ \psi$, $\phi \circ \pi_0 \circ \overline{\alpha}(0) = \tilde{z} = \phi \circ \pi_0 \circ \overline{\alpha}(1)$. Since $\phi \circ \pi_0$ is injective on U, $\overline{\alpha}(1) = \overline{\alpha}(0) = \overline{z}$. Since $M_0 \times N_0$ has a product metric, $\pi_0 \circ \overline{\alpha}(t) = \overline{z}$ for all t. Thus $\pi_1 \circ \tilde{\alpha}(t) = \pi_1 \circ \psi \circ \overline{\alpha}(t) = \phi \circ \pi_0 \circ \overline{\alpha}(t) = \tilde{z}$ for all t. This shows that $M_1 \times \{\tilde{z}\}$ is totally geodesic.

□

Lemma 5.6. *Suppose the diagram* (5.1) *commutes, where M_0, N_0, and the $M_i \times N_i$ are Riemannian manifolds, $M_0 \times N_0$ has the product metric obtained from M_0 and N_0, and ψ is a local isometry. Let $\overline{\Gamma}$ denote the deck transformation group of $M_0 \times N_0$. Then $\overline{\Gamma} \subseteq \mathscr{I}(M_0) \times \mathscr{I}(N_0)$.*

Proof. It must be shown that each $\gamma \in \overline{\Gamma}$ is of the form $\gamma = \alpha \times \beta$ for some $\alpha \in \mathscr{I}(M_0)$ and $\beta \in \mathscr{I}(N_0)$. This is done by first showing that γ takes M_0-fibers to M_0-fibers. Fix $\overline{x} \in N_0$, and let $\overline{\sigma} : [a,b] \to M_0 \times \{\overline{x}\}$ be continuous. Then $\pi_0 \circ \overline{\sigma} = \overline{x}$ for all t. Let $\tilde{\sigma} := \psi \circ \overline{\sigma}$. Then $\pi_1 \circ \tilde{\sigma}(t) = \pi_1 \circ \circ \psi \circ \overline{\sigma}(t) = \phi \circ \pi_0 \circ \overline{\sigma}(t) = \phi(\overline{x})$ for all t. Let $\overline{\sigma}_0 := \gamma \circ \overline{\sigma}$, and $\overline{x}_0 := \overline{\sigma}_0(a)$. Then $\psi \circ \overline{\sigma}_0 = \tilde{\sigma}$, so as before one has that $\tilde{x} = \pi_1 \circ \tilde{\sigma}(t) = \pi_1 \circ \psi \circ \overline{\sigma}_0(t) = \phi \circ \pi_0 \circ \overline{\sigma}_0(t)$ for all t. So $\pi_0 \circ \overline{\sigma}_0 \in \phi^{-1}(\tilde{x})$ for all t. Since $\phi^{-1}(\tilde{x})$ is discrete, this implies that $\pi_0 \circ \overline{\sigma}_0(t) = \overline{x}_0$ for all t. Since M_0 is connected, it follows that γ takes each each M_0 fiber into a single M_0-fiber. Since γ is an isometry, it must preserve the normal distribution to the M_0-fibers; since $M_0 \times N_0$ has a product metric, that is exactly the tangent distribution to the N_0-fibers. Since N_0 is connected, γ also takes each N_0-fiber into an N_0-fiber. Because $\pi_0 \circ \gamma$ is constant along each M_0-fiber and $\rho_0 \circ \gamma$ is constant along each N_0-fiber, there exist $\alpha : M_0 \to M_0$ and $\beta : N_0 \to N_0$ such that $\alpha \times \beta$. Since γ is an isometry, α and β are isometries as well.

□

As an application of Lemma 5.6, one may show that, when ϕ is normal, the metric on N_1 need not be given in advance, but is instead uniquely determined by the rest of the diagram (5.1).

Proposition 5.7. *Suppose the diagram* (5.1) *commutes, where M_0 and N_0 are Riemannian manifolds, $M_0 \times N_0$ has the product metric obtained from M_0 and N_0, ϕ is normal, and ψ is a local isometry. Then there exists a unique Riemannian metric on N_1 with respect to which ϕ is a local isometry.*

Proof. Denote by h the metric on N_0, g the metric on M_0, and Γ the deck transformation group of ϕ. Since ϕ is a normal covering map, h will descend to a metric on N_1 if and only if h is Γ-equivariant.

Fix any $\overline{x} \in N_0$ and $v_1, v_2 \in \mathrm{T}_{\overline{x}}N_0$, write $\tilde{x} := \phi(\overline{x})$, and let $\gamma \in \Gamma$. Note that $\phi \circ \gamma(\overline{x}) = \tilde{x}$. Fix any $\overline{p}_1 \in M_0$. Since $\psi(\overline{p}_1, \overline{x}) \in M_1 \times \{\tilde{x}\}$, Lemma 5.3 states that $\psi|_{M_0 \times \{\gamma(\overline{x})\}} : M_0 \times \{\gamma(\overline{x})\} \to M_1 \times \{\tilde{x}\}$ is surjective, and there exists $\overline{p}_2 \in M_0$ such that $\psi(\overline{p}_2, \gamma(\overline{x})) = \psi(\overline{p}_1, \overline{x})$. Since ψ is a normal covering map, there exists a deck transformation σ of ψ that takes $(\overline{p}_1, \overline{x})$ to $(\overline{p}_2, \gamma(\overline{x}))$. By Lemma 5.6, $\sigma = \alpha \times \beta$ for $\alpha \in \mathscr{I}(M_0)$ and $\beta \in \mathscr{I}(N_0)$. One has that

$$\begin{aligned} \phi \circ \beta &= \phi \circ \pi_1 \circ \sigma \circ \iota_{\overline{x}} \\ &= \pi_0 \circ \psi \circ \sigma \circ \iota_{\overline{x}} \\ &= \pi_0 \circ \psi \circ \iota_{\overline{x}} \\ &= \phi \circ \pi_0 \circ \iota_{\overline{x}} \\ &= \phi \end{aligned}$$

Thus β is a deck transformation of ϕ. Since $\beta(\overline{x}) = \gamma(\overline{x})$, it follows that $\beta = \gamma$, and one has

$$\begin{aligned} h(v_1, v_2) &= g \times h((0, v_1)_{(\overline{p}_1, \overline{x})}, (0, v_2)_{(\overline{p}_2, \overline{x})}) \\ &= g \times h(\sigma_*(0, v_1)_{(\overline{p}_1, \overline{x})}, \sigma_*(0, v_1)_{(\overline{p}_1, \overline{x})}) \\ &= g \times h\Big((\alpha_*(0), \beta_*(v_1))_{(\overline{p}_2, \gamma(\overline{x}))}, (\alpha_*(0), \beta_*(v_2))_{(\overline{p}_2, \gamma(\overline{x}))}\Big) \\ &= g \times h\Big((0, \beta_*(v_1))_{(\overline{p}_2, \gamma(\overline{x}))}, (0, \beta_*(v_2))_{(\overline{p}_2, \gamma(\overline{x}))}\Big) \\ &= h(\beta_*(v_1), \beta_*(v_2)) \\ &= h(\gamma_*(v_1), \gamma_*(v_2)) \end{aligned}$$

Thus h descends to a unique metric on N_0 with respect to which ϕ is a local isometry.

□

In the following, a **fundamental domain** for a covering map $\varphi : Y \to Z$ is defined to be a set $X \subseteq Y$ such that $\psi|_X : X \to Z$ is bijective. In practice, fundamental domains should be nice sets, at least measurable with boundary measure zero, but that is not taken as part of the definition here.

Lemma 5.8. *Suppose the diagram* (5.2) *commutes. If $A \subseteq M_0$ and $B \subseteq N_0$ are fundamental domains of χ and ϕ, respectively, then $\varphi(A \times B)$ is a fundamental domain of ψ.*

Proof. It must be shown that $\psi \circ \varphi(A \times B) = M_1 \times N_1$ and that $\psi|_{\varphi(A \times B)}$ is injective. To see the first, let $(\tilde{p}, \tilde{x}) \in M_1 \times N_1$. Since $\chi(A) = M_0$ and $\phi(B) = N_0$, there exist $\overline{p} \in A$ and $\overline{x} \in B$ such that $\phi(\overline{x}) = \tilde{x}$.

Therefore, $\psi \circ \varphi(\overline{p}, \overline{x}) = \chi \times \phi(\overline{p}, \overline{x}) = (\tilde{p}, \tilde{x})$. To see the second, suppose $(\overline{p}_i, \overline{x}_i) \in \varphi(A \times B)$, $i = 1, 2$, satisfy $\psi(\overline{p}_1, \overline{x}_1) = \psi(\overline{p}_2, \overline{x}_2)$. Then there exist $(\overline{q}_i, \overline{z}_i) \in A \times B$ such that $(\overline{p}_i, \overline{x}_i) = \varphi(\overline{q}_i, \overline{z}_i)$. Thus

$$\begin{aligned}(\chi(\overline{q}_i), \phi(\overline{z}_i)) &= \chi \times \phi(\overline{q}_i, \overline{z}_i) \\ &= \psi \circ \varphi(\overline{q}_i, \overline{z}_i) \\ &= \psi(\overline{p}_i, \overline{z}_i)\end{aligned}$$

for each i. Therefore, $\chi(\overline{q}_1) = \chi(\overline{q}_2)$ and $\phi(\overline{z}_1) = \phi(\overline{z}_2)$. Since $\chi|_A$ and $\phi|_B$ are injective, $(\overline{q}_1, \overline{z}_1) = (\overline{q}_2, \overline{z}_2)$, and consequently $(\overline{p}_1, \overline{z}_1) = \varphi(\overline{q}_1, \overline{z}_1) = \varphi(\overline{q}_2, \overline{z}_2) = (\overline{p}_2, \overline{z}_2)$, and $\psi|_{\varphi(A\times B)}$ is injective.

□

Lemma 5.9. *Suppose the diagram* (5.2) *commutes. If χ is finite with $\#(\chi)$ sheets and $B \subseteq N_0$ is a fundamental domain of ϕ, then $M_0 \times B$ is the union of $\#(\chi)$ fundamental domains of ψ.*

Proof. Since χ has $\#(\chi)$ sheets, M_0 is the union of $\#(\chi)$ fundamental domains of χ. By Lemma 5.8, whenever $B_0 \subseteq N_0$ is a fundamental domain of ϕ, $\varphi(M_0 \times B_0)$ is the union of $\#(\chi)$ fundamental domains of ψ; so the result will follow if it's shown that, for some such B_0, $M_0 \times B = \varphi(M_0 \times B_0)$. Fix $\overline{p} \in M_0$. Since $\phi \circ \pi_0 \circ \varphi \circ \iota_{\overline{p}} = \pi_1 \circ (\chi \times \phi) \circ \iota_{\overline{p}} = \phi$, $\pi_0 \circ \varphi \circ \iota_{\overline{p}}$ is a deck transformation of ϕ. Therefore, $B_0 := (\pi_0 \circ \varphi \circ \iota_{\overline{p}})^{-1}(B)$ is a fundamental domain of ϕ. Note that, for each $\overline{z} \in N_0$, $\phi \circ \pi_0 \circ \varphi \circ \iota_{\overline{z}}(\cdot) = \pi_1 \circ (\chi \times \phi) \circ \iota_{\overline{z}}(\cdot) = \phi(\overline{z})$. As in the proof of Lemma 5.3, it follows by continuity that $\pi_0 \circ \varphi \circ \iota_{\overline{z}}$ is constant. One may therefore show that $\rho_0 \circ \varphi \circ \iota_{\overline{z}} : M_0 \to M_0$ is a homeomorphism. If $(\overline{q}, \overline{z}) \in M_0 \times B_0$, then

$$\begin{aligned}\pi_0 \circ \varphi \circ \iota_{\overline{q}}(\overline{z}) &= \pi_0 \circ \varphi \circ \iota_{\overline{z}}(\overline{q}) \\ &= \pi_0 \circ \varphi \circ \iota_{\overline{z}}(\overline{p}) \\ &= \pi_0 \circ \varphi \circ \iota_{\overline{p}}(\overline{z})\end{aligned}$$

Since $\pi_0 \circ \varphi \circ \iota_{\overline{p}}(\overline{z}) \in B$, it follows that $\varphi(\overline{q}, \overline{z}) = \varphi \circ \iota_{\overline{q}}(\overline{z}) \in M_0 \times B$. So $\varphi(M_0 \times B_0) \subseteq M_0 \times B$. On the other hand, if $(\overline{q}, \overline{x}) \in M_0 \times B$, let $\overline{z} := (\pi_0 \circ \varphi \circ \iota_{\overline{p}})^{-1}(\overline{x})$ and $\overline{q}_0 := (\rho_0 \circ \varphi \circ \iota_{\overline{z}})^{-1}(\overline{q})$. As above, $\pi_0 \circ \varphi(\overline{q}_0, \overline{z}) = \pi_0 \circ \varphi \circ \iota_{\overline{p}}(\overline{z}) = \overline{x}$. At the same time, $\rho_0 \circ \varphi(\overline{q}_0, \overline{z}) = \overline{q}$. So $(\overline{q}, \overline{x}) = \varphi(\overline{q}_0, \overline{z}) \in \varphi(M_0 \times B_0)$. So $M_0 \times B \subseteq \varphi(M_0 \times B_0)$.

□

Lemma 5.10. *Suppose the diagram* (5.2) *commutes isometrically. Let N be a Riemannian manifold and $\tilde{f} : M_1 \times N_1 \to N$ a continuous function. Then the following hold:*

(a) *If $\tilde{f}$ is constant along each M_1-fiber and, for each $\tilde{p} \in M_1$, $\tilde{f} \circ \iota_{\tilde{p}}$ is totally geodesic, then $\tilde{f}$ is totally geodesic; and*

(b) *If $M_1 \times N_1$ has finite volume, each geodesic in $\overline{N}$ is minimal, $\tilde{f}$ is totally geodesic, and, for any $(\tilde{p}, \tilde{x}) \in M_1 \times N_1$, $(\tilde{f} \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p})) =< e >$, then $\tilde{f}$ is constant along each M_1-fiber and, for each $\tilde{p} \in M_1$, $\tilde{f} \circ \iota_{\tilde{p}}$ is totally geodesic.*

Proof. **(a)** Suppose $\tilde{f}$ is constant along each M_1-fiber and that each $\tilde{f} \circ \iota_{\tilde{p}} : N_1 \to N$ is totally geodesic. Write $\overline{f} := \tilde{f} \circ \psi$. Since $\pi_1 \circ \psi = \phi \circ \pi_0$, $\overline{f}$ is constant along each M_0-fiber. Since $M_0 \times N_0$ has a product metric, to prove that $\overline{f}$ is totally geodesic, it suffices to show that, for each $\overline{p} \in M_0$, $\overline{f} \circ \iota_{\overline{p}} : N_0 \to N$ is totally geodesic. Let $(\overline{p}, \overline{x}) \in M_0 \times N_0$, and write $(\tilde{p}, \tilde{x}) := \psi(\overline{p}, \overline{x})$. Let $\overline{\gamma} : [a,b] \to N_0$ be a geodesic geodesic with $\overline{\gamma}(a) = \overline{x}$, and write $\tilde{\gamma} := \phi \circ \overline{\gamma}$. Since ϕ is a local isometry, $\tilde{\gamma}$ is a geodesic. Note

$$
\begin{aligned}
\pi_1 \circ \psi \circ \iota_{\overline{p}} &= \phi \circ \pi_0 \circ \iota_{\overline{p}} \\
&= \phi \\
&= \pi_1 \circ \iota_{\tilde{p}} \circ \phi
\end{aligned}
$$

Thus $\pi_1 \circ \iota_{\tilde{p}} \circ \tilde{\gamma} = \pi_1 \circ \iota_{\tilde{p}} \circ \phi \circ \overline{\gamma} = \pi_1 \circ \psi \circ \iota_{\overline{p}} \circ \overline{\gamma}$. Since $\tilde{f}$ is constant along each M_1-fiber, it follows that $\tilde{f} \circ \iota_{\tilde{p}} \circ \tilde{\gamma} = \tilde{f} \circ \psi \circ \iota_{\overline{p}} \circ \overline{\gamma} = \overline{f} \circ \iota_{\overline{p}} \circ \overline{\gamma}$. Since the former is a geodesic, so is the latter. Since $(\overline{p}, \overline{x})$ and $\overline{\gamma}$ were arbitrary, $\overline{f}$ is totally geodesic.

(b) By Lemma 2.6.2, $\tilde{f}$ is smooth. By Lemma 5.5, each M_1-fiber of $M_1 \times N_1$ is totally geodesic. Since $M_1 \times N_1$ has finite volume and its M_1-fibers are totally geodesic, the coarea formula may be used to show that almost all M_1-fibers have finite volume. By Lemma 3.1.7 and a continuity argument, $\tilde{f}$ must be constant along every M_1-fiber. Let $(\tilde{p}, \tilde{x}) \in M_1 \times N_1$ and $(\overline{p}, \overline{x}) \in \psi^{-1}(\tilde{p}, \tilde{x})$, and let $\tilde{\gamma} : [a,b] \to N_1$ be a geodesic with $\tilde{\gamma}(a) = \tilde{x}$. Since $\phi \circ \pi_0 = \pi_1 \circ \psi$, $\overline{x} \in \phi^{-1}(\tilde{x})$. map $\overline{f} := \tilde{f} \circ \psi$ is totally geodesic. Let $\overline{\gamma} : [a,b] \to N_0$ be the lift of $\tilde{\gamma}$ satisfying $\overline{\gamma}(a) = \overline{x}$ and $\tilde{\gamma} = \phi \circ \overline{\gamma}$. Then $\iota_{\overline{p}} \circ \overline{\gamma}$ is a geodesic, which implies that $\overline{f} \circ \iota_{\overline{p}} \circ \overline{\gamma}$ is also a geodesic. As shown above, $\tilde{f} \circ \iota_{\tilde{p}} \circ \tilde{\gamma} = \overline{f} \circ \iota_{\overline{p}} \circ \overline{\gamma}$. Thus $\tilde{f} \circ \iota_{\tilde{p}} \circ \tilde{\gamma}$ is a geodesic. Since $\tilde{x}$ and $\tilde{\gamma}$ were arbitrary, $\tilde{f} \circ \iota_{\tilde{p}}$ is totally geodesic.

□

Lemma 5.11. *Suppose the diagram (5.1) commutes. Suppose also that the N_i are Riemannian manifolds such that $\phi : N_0 \to N_1$ is a local isometry, that N_0 is simply connected, and that M_1 and N_1 are compact. Let N be a Riemannian manifold, $\pi : \overline{N} \to N$ its Riemannian universal covering map, and $\tilde{f} : M_1 \times N_1 \to N$ a continuous function such that $(\tilde{f} \circ \iota_{\tilde{z}})_*(\pi_1(M_1, \tilde{q})) = <e>$ for any $(\tilde{q}, \tilde{z}) \in M_1 \times N_1$. Write $\overline{f} := \tilde{f} \circ \psi$. For each $R \geq 0$, there exists $C = C(R, \tilde{f}) \geq 0$ such that the following hold:*

(a) *Given any $(\overline{p}, \overline{x}) \in M_0 \times N_0$ and $\overline{y} \in \pi^{-1}(\overline{f}(\overline{p}, \overline{x}))$, $\overline{f}$ lifts to a map $\overline{F} : M_0 \times N_0 \to \overline{N}$ such that $\overline{F}(\overline{p}, \overline{x}) = \overline{y}$ and $\pi \circ \overline{F} = \overline{f}$; and*

(b) *Every such lift $\overline{F}$ satisfies $\mathrm{d}_{\overline{N}}(\overline{F}(\overline{q}_0, \overline{z}_0), \overline{F}(\overline{q}_1, \overline{z}_1)) \leq C$ whenever $\mathrm{d}_{N_0}(\overline{z}_0, \overline{z}_1) \leq R$.*

Proof. **(a)** Write $(\tilde{p}, \tilde{x}) := \psi(\overline{p}, \overline{x})$. Since $(\tilde{f} \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p})) = <e>$, the homotopy lifting property implies that $(\overline{f} \circ \iota_{\overline{x}})_*(\pi_1(M_0, \overline{p})) = <e>$. This and the fact that N_0 is simply connected that

$$\overline{f}_*\Big(\pi_1(M_0 \times N_0, (\overline{p}, \overline{x}))\Big) = (\overline{f} \circ \iota_{\overline{x}})_*(\pi_1(M_0, \overline{p}))(\overline{f} \circ \iota_{\overline{p}})_*(\pi_1(N_0, \overline{x})) = <e>$$

Therefore, $\overline{f}$ lifts to a map $\overline{F} : M_0 \times N_0 \to \overline{N}$ such that $\overline{F}(\overline{p}, \overline{x}) = \overline{y}$ and the diagram

$$\begin{tikzcd}
M_0 \times N_0 \arrow[r, "\overline{F}"] \arrow[d, "\psi"] \arrow[dr, "\overline{f}"] & \overline{N} \arrow[d, "\pi"] \\
M_1 \times N_1 \arrow[r, "\tilde{f}"] & N
\end{tikzcd}$$

commutes.

(b) Fix $(\overline{p}, \overline{x}) \in M_0 \times N_0$, $(\tilde{p}, \tilde{x}) := (\overline{p}, \overline{x})$, $\overline{y} \in \pi^{-1}(\overline{f}(\overline{p}, \overline{x}))$, and $\overline{F} : M_0 \times N_0 \to \overline{N}$ as in part (a). Let $\chi = \chi_{\overline{x}} : M_0 \to M_1$ be the covering map and $\varphi : M_0 \times N_0 \to M_0 \times N_0$ the homeomorphism from Proposition 5.4, so that the diagram (5.2) commutes. It will be shown that, for each $(\tilde{q}, \overline{z}) \in M_1 \times N_0$, $\overline{F} \circ \varphi$ is constant on $(\chi \times \mathrm{id})^{-1}(\tilde{q}, \overline{z})$. Write $\tilde{z} := \phi(\overline{z})$. Let $\overline{q}_0, \overline{q}_1 \in \chi^{-1}(\tilde{q})$, and fix any path $\overline{\sigma} : [a, b] \to M_0$ from $\overline{q}_0$ to $\overline{q}_1$. Then $\tilde{\sigma} := \chi \circ \overline{\sigma}$ is a loop in M_1 based at $\tilde{q}$. Note that, for each $i = 0, 1$,

$$\begin{aligned}
\pi \circ \overline{F} \circ \varphi(\overline{q}_i, \overline{z}) &= \tilde{f} \circ \psi \circ \varphi(\overline{q}_i, \overline{z}) \\
&= \tilde{f} \circ \chi \times \phi(\overline{q}_i, \overline{z}) \\
&= \tilde{f}(\tilde{q}, \tilde{z})
\end{aligned}$$

So $\pi \circ \overline{F} \circ \varphi(\overline{q}_0, \overline{z}) = \pi \circ \overline{F} \circ \varphi(\overline{q}_1, \overline{z})$; in other words, $\overline{F} \circ \varphi(\overline{q}_0, \overline{z})$ and $\overline{F} \circ \varphi(\overline{q}_1, \overline{z})$ lie in the same fiber of π. Note that $\overline{F} \circ \varphi \circ \iota_{\bar{z}} \circ \overline{\sigma}$ is a path connecting $\overline{F} \circ \varphi(\overline{q}_0, \bar{z})$ to $\overline{F} \circ \varphi(\overline{q}_1, \bar{z})$. By assumption,

$$\begin{aligned} 0 &= (\tilde{f} \circ \iota_{\tilde{z}})_*([\tilde{\sigma}]) \\ &= [\tilde{f} \circ \iota_{\tilde{z}} \circ \chi \circ \overline{\sigma}] \\ &= [\tilde{f} \circ \chi \times \phi \circ \iota_{\bar{z}} \circ \overline{\sigma}] \\ &= [\tilde{f} \circ \psi \circ \varphi \circ \iota_{\bar{z}} \circ \overline{\sigma}] \\ &= [\pi \circ \overline{F} \circ \varphi \circ \iota_{\bar{z}} \circ \overline{\sigma}] \end{aligned}$$

It follows from the homotopy lifting property that $\overline{F} \circ \varphi(\overline{q}_0, \bar{z}) = \overline{F} \circ \varphi(\overline{q}_1, \bar{z})$. Thus $\overline{F} \circ \varphi$ is constant on each fiber of $\chi \times \mathrm{id}$. This implies that $\overline{F} \circ \varphi$ descends along $\chi \times \mathrm{id}$ to a map $\tilde{F} : M_1 \times N_0 \to \overline{N}$ such that the diagram

$$\begin{tikzcd} M_0 \times N_0 \arrow[r, "\varphi"] \arrow[d, "\chi\times\mathrm{id}"] & M_0 \times N_0 \arrow[r, "\overline{F}"] \arrow[d, "\psi"] \arrow[dr, "\overline{f}"] & \overline{N} \arrow[d, "\pi"] \\ M_1 \times N_0 \arrow[r, "\mathrm{id}\times\phi"] \arrow[urr, bend left=15, "\tilde{F}"] & M_1 \times N_1 \arrow[r, "\tilde{f}"] & N \end{tikzcd}$$

commutes.

The next step is to define a function $D : N_1 \to [0, \infty)$ by $D(\tilde{z}) := \mathrm{diam}\Big(\tilde{F}(M_1 \times \overline{B}(\bar{z}, R))\Big)$ for any $\bar{z} \in \phi^{-1}(\tilde{z})$. To see that D is well-defined, fix $\tilde{z} \in N_1$ and $\bar{z}_0, \bar{z}_1 \in \phi^{-1}(\tilde{z})$. Denote by $\overline{\alpha} : N_0 \to N_0$ the deck transformation of ϕ that takes $\bar{z}_0$ to $\bar{z}_1$. Since ϕ is a local isometry, $\overline{\alpha}$ is an isometry, and $\overline{\alpha}(\overline{B}(\bar{z}_0, R)) = \overline{B}(\bar{z}_1, R)$. One has that

$$\begin{aligned} \pi \circ \tilde{F} \circ \mathrm{id} \times \overline{\alpha} &= \tilde{f} \circ \mathrm{id} \times \phi \circ \mathrm{id} \times \overline{\alpha} \\ &= \tilde{f} \circ \mathrm{id} \times (\phi \circ \overline{\alpha}) \\ &= \tilde{f} \circ \mathrm{id} \times \phi \\ &= \pi \circ \tilde{F} \end{aligned}$$

In other words, $\tilde{F}$ and $\tilde{F} \circ \mathrm{id} \times \overline{\alpha}$ are lifts along π of the same map. Since π is normal, it follows from general theory that there exists a deck transformation $\overline{\beta} : \overline{N} \to \overline{N}$ of π such that $\overline{\beta} \circ \tilde{F} = \tilde{F} \circ \mathrm{id} \times \overline{\alpha}$. Since π is a local isometry, $\overline{\beta}$ is an isometry. Therefore,

$$\begin{aligned} \mathrm{diam}\Big(\tilde{F}(M_1 \times \overline{B}(\overline{z_0}, R))\Big) &= \mathrm{diam}\Big(\overline{\beta} \circ \tilde{F}(M_1 \times \overline{B}(\overline{z_0}, R))\Big) \\ &= \mathrm{diam}\Big(\tilde{F} \circ \mathrm{id} \times \overline{\alpha}(M_1 \times \overline{B}(\overline{z_0}, R))\Big) \\ &= \mathrm{diam}\Big(\tilde{F}(M_1 \times \overline{B}(\overline{z_1}, R))\Big) \end{aligned}$$

This shows that D is well-defined. Since M_1 is compact, $D(\tilde{z}) < \infty$ for each $\tilde{z} \in N_1$. Since D is continuous and N_1 is compact, it follows that D is bounded above by some $C \geq 0$. If $(\overline{q}_i, \overline{z}_i) \in M_0 \times N_0$, $i = 0, 1$, satisfy $\mathrm{d}_{N_0}(\overline{z}_0, \overline{z}_1) \leq R$, then $\mathrm{d}_{\overline{N}}(\overline{F}(\overline{q}_0, \overline{z}_0), \overline{F}(\overline{q}_1, \overline{z}_1)) \leq \mathrm{diam}\Big(\tilde{F}(M_1 \times \overline{B}(\overline{z}_0, R))\Big) = D(\phi(\overline{z}_0)) \leq C$. It only remains to show that this inequality holds for an arbitrary lift $\overline{G} : M_0 \times N_0 \to \overline{N}$ of $\overline{f}$. This follows from the fact that there exists an isometry $\overline{\gamma} : \overline{N} \to \overline{N}$ such that $\overline{\gamma} \circ \overline{F} = \overline{G}$.

□

Remark 5.12. In the case of $R = 0$, Lemma 5.11(b) is equivalent to the statement that every such lift $\overline{F}$ satisfies $\mathrm{diam}(\overline{F} \circ \iota_{\overline{z}}(M_0)) \leq C$ for each $\overline{z} \in N_0$. For this to hold, the diagram (5.1) need only commute, not necessarily isometrically. One may define a pseudo-metric $\mathrm{d}_{\overline{F}}$ on $M_0 \times N_0$ by setting

$$\mathrm{d}_{\overline{F}}((\overline{q}_0, \overline{z}_0), (\overline{q}_1, \overline{z}_1)) := \mathrm{d}_{\overline{N}}(\overline{F}(\overline{q}_0, \overline{z}_0), \overline{F}(\overline{q}_1, \overline{z}_1))$$

for any such lift $\overline{F}$. One may also define a pseudo-metric d_H on N_0 by setting $\mathrm{d}_H(\overline{z}_0, \overline{z}_1)$ to be the Hausdorff distance between the compact sets $\overline{F}(\pi_0^{-1}(z_0))$ and $\overline{F}(\pi_0^{-1}(z_1))$. The result for $R = 0$ is then equivalent to the statement that $(M_0 \times N_0, \mathrm{d}_{\overline{F}})$ is within finite Gromov-Hausdorff distance of (N_0, d_H).

Lemma 5.13. *Suppose the diagram* (5.1) *commutes isometrically, where M_0 has finite volume. Then there exists a number $\#(\psi) \in \mathbb{N}$ such that each $\psi|_{\overline{z}}$ has $\#(\psi)$ sheets and each M_1-fiber of $M_1 \times N_1$ has volume equal to* $\frac{1}{\#(\psi)}\mathrm{vol}(M_0)$.

Proof. By Lemma 5.5, each M_1-fiber of $M_1 \times N_1$ is totally geodesic. Since the M_0-fibers of $M_0 \times N_0$ are also totally geodesic and ψ is a local isometry, Lemma 5.3 implies that each $\psi|_{M_0 \times \{\overline{z}\}} : M_0 \times \{\overline{z}\} \to M_1 \times \{\phi(\overline{z})\}$ is a Riemannian covering map. Since M_0 has finite volume, each $\psi|_{M_0 \times \{\overline{z}\}}$ must therefore be a finite covering map, and each M_1-fiber must have finite volume. Let $\chi = \chi_{\overline{x}} := \rho_1 \circ \psi \circ \iota_{\overline{x}}$ and $\varphi : M_0 \times N_0 \to M_0 \times N_0$ be as in Proposition 5.4, so that the diagram (5.2) commutes. As shown in the proof of Lemma 5.9, each $\pi_0 \circ \varphi \circ \iota_{\overline{z}}$ is constant, and, consequently, each $\rho_0 \circ \varphi \circ \iota_{\overline{z}} : M_0 \to M_0$ is a homeomorphism. Therefore, $\chi = (\rho_1 \circ \psi \circ \iota_{\pi_0 \circ \varphi \circ \overline{z}}) \circ (\rho_0 \circ \varphi \circ \iota_{\overline{z}})$ for each $\overline{z}$. That is to say, $\rho_0 \circ \varphi \circ \iota_{\overline{z}}$ maps each fiber of χ to a fiber of $\rho_1 \circ \psi \circ \iota_{\pi_0 \circ \varphi \circ \overline{z}}$; so their fibers must have the same cardinality. Since $\pi_0 \circ \varphi$ is surjective, this shows that each $\varphi|_{M_0 \times \{\overline{z}\}}$ has the same number of sheets as χ. Let $\#(\psi) \in \mathbb{N}$ denote that common number of sheets. Then each M_1-fiber must have volume equal to $\frac{1}{\#(\psi)}\mathrm{vol}(M_0)$.

□

Remark 5.14. By Lemma 5.6, whenever (5.1) commutes isometrically, the deck transformation group $\overline{\Gamma}$ of ψ satisfies $\overline{\Gamma} \subseteq \mathscr{I}(M_0) \times \mathscr{I}(N_0)$. In the abstract, M_1 is realized as the quotient of M_0 by the kernel of the projection onto the first component of that splitting. This provides another way of seeing that each $\psi|_{\overline{z}}$ has the same number of sheets.

As a consequence of Lemma 2.1.1 and Lemma 2.1.2, when a diagram of the form (5.1) commutes isometrically and $\psi_1 : M_1 \times N_1 \to M$ is a finite Riemannian covering map, one may, by passing to covers of the spaces in the diagram, replace ψ_1 with a normal covering map. That is the content of the next lemma, which is rather technical but will play an important role in the work to come. A few more algebraic preliminaries will help clarify the argument. A key fact, which will be used implicitly and repeatedly, is that, when $M \times N$ is a product manifold and $(p,q) \in M \times N$, the map $\pi_1(M,p) \times \pi_1(N,q) \to \pi_1(M \times N,(p,q))$ defined by

$$([\alpha],[\beta]) \mapsto [\iota_{q,1} \circ \alpha][\iota_{p,2} \circ \beta]$$

is an isomorphism, where $\iota_{q,1} : M \to M \times N$ and $\iota_{p,2} : N \to M \times N$ are the standard inclusion maps $\iota_{q,1}(\cdot) := (\cdot, q)$ and $\iota_{p,2}(\cdot) := (p, \cdot)$. Suppose $A \leq \pi_1(M,p)$ and $B \leq \pi_1(N,q)$. Since each of $\pi_1(M,p) \cong \pi_1(M,q(\iota_{q,1})_*(\pi_1(M,p))$ and $\pi_1(N,q) \cong (\iota_{p,2})_*(\pi_1(N,q))$, viewed as subgroups of $\pi_1(M,p) \times \pi_1(N,q) \cong \pi_1(M \times N,(p,q))$, is contained in the centralizer of the other, the subgroup product AB is a subgroup of $\pi_1(M \times N,(p,q))$. If $f : M \times N \to X$ is continuous and $x := f(p,q)$, then, since $f_* : \pi_1(M \times N,(p,q)) \to \pi_1(X,x)$ is a homomorphism, $f_*(A \times B) = f_*(AB) = f_*(A)f_*(B)$, the latter also being the subgroup product. When $X = Y \times Z$ and f is a product map of the form $f = g \times h$ for $g : M \to Y$ and $h : N \to Z$, where $y = g(p)$ and $z = h(q)$, these equalities take the form $(g \times h)_*(A \times B) = g_*(A) \times h_*(B) = (\iota_{z,1} \circ g)_*(A)(\iota_{x,2} \circ h)_*(B)$.

Lemma 5.2.1. *Suppose the diagram* (5.1) *commutes. Let M be a manifold and $\psi_1 : M_1 \times N_1 \to M$ a finite covering map. Then there exist manifolds $\hat{M}_i$ and $\hat{N}_i$, $i = 1,2$, and covering maps $\hat{\psi}_0 : \hat{M}_0 \times \hat{N}_0 \to \hat{M}_1 \times \hat{N}_1$ and $\hat{\phi} : \hat{N}_0 \to \hat{N}_1$ such that the diagram*

$$\begin{tikzcd} \hat{M}_0 \times \hat{N}_0 \arrow[r, "\hat{\pi}_0"] \arrow[d, "\hat{\psi}_0"] & \hat{N}_0 \arrow[d, "\hat{\phi}"] \\ \hat{M}_1 \times \hat{N}_1 \arrow[r, "\hat{\pi}_1"] & \hat{N}_1 \end{tikzcd} \tag{5.3}$$

commutes, where each $\hat{\pi}_i$ denotes projection onto the second component. There also exist finite and normal covering maps $\hat{\psi}_1 : \hat{M}_1 \times \hat{N}_1 \to M$, $\zeta_1 : \hat{N}_1 \to N_1$, $\xi_1 \times \zeta_1 : \hat{M}_1 \times \hat{N}_1 \to M_1 \times N_1$ and, for each

$i = 1,2$, $\xi_i : \hat{M}_i \to M_i$. If N_0 is simply connected, then $\hat{N}_0 = N_0$ and $\xi_0 \times \mathrm{id} : \hat{M}_0 \times \hat{N}_0 \to M_0 \times N_0$ is a finite and normal covering map. Moreover, if the M_i, N_i, and $M_i \times N_i$ are Riemannian manifolds, where $M_0 \times N_0$ has the product metric obtained from M_0 and N_0, and ψ_0, ψ_1, and ϕ are local isometries, then the $\hat{M}_i$, $\hat{N}_i$, and $\hat{M}_i \times \hat{N}_i$ may be endowed with Riemannian metrics, where the metric on $\hat{M}_0 \times \hat{N}_0$ is the product metric obtained from $\hat{M}_0$ and $\hat{N}_0$, that make $\hat{\psi}_0$, $\hat{\psi}_1$, and $\hat{\phi}$ local isometries.

Proof. Let $(\tilde{p},\tilde{x}) \in M_1 \times N_1$ and $(\overline{p},\overline{x}) \in \psi_0^{-1}(\tilde{p},\tilde{x})$, where, since $\phi \circ \pi_0 = \pi_1 \circ \psi_0$, $\phi(\overline{x}) = \tilde{x}$. Write $x := \psi_1(\tilde{p},\tilde{x})$. Since ψ_1 has finitely many sheets, $G := (\psi_1)_*\big(\pi_1(M_1 \times N_1, (\tilde{p},\tilde{x}))\big)$ is a finite-index subgroup of $\pi_1(M,x)$. Note that $\pi_1(M_1 \times N_1, (\tilde{p},\tilde{x})) \cong \pi_1(M_1,\tilde{p}) \times \pi_1(N_1,\tilde{x})$. Write $G_1 := (\psi_1 \circ \iota_{\tilde{x}})_*(\pi_1(M_1,\tilde{p}))$ and $G_2 := (\psi_1 \circ \iota_{\tilde{p}})_*(\pi_1(N_1,\tilde{x}))$. Then $G \cong G_1 \times G_2$. By Lemma 2.1.1, there exists a normal subgroup H of $\pi_1(M,x)$ such that $H \subseteq G$ and $[G : H] \leq [\pi_1(M,x) : H] < \infty$. By Lemma 2.1.2, there exist normal subgroups H_i of G_i, $i = 1,2$, such that $H_1 \times H_2 \subseteq H$, $[G_i : H_i] \leq [G : H]$ for each i, and $[G : H_1 \times H_2] < \infty$. Since $(\psi_1 \circ \iota_{\tilde{x}})_* : \pi_1(M_1,\tilde{p}) \to G_1$ and $(\psi_1 \circ \iota_{\tilde{p}})_* : \pi_1(N_1,\tilde{x}) \to G_2$ are isomorphisms, $\hat{H}_1 := (\psi_1 \circ \iota_{\tilde{x}})_*^{-1}(H_1)$ and $\hat{H}_2 := (\psi_1 \circ \iota_{\tilde{p}})_*^{-1}(H_2)$ are finite-index and normal subgroups of $\pi_1(M_1,\tilde{p})$ and $\pi_1(N_1,\tilde{x})$, respectively. By the general theory of covering spaces, there exist a manifold $\hat{M}_1$, a point $\hat{p}_1 \in \hat{M}_1$, and a finite and normal covering map $\xi_1 : \hat{M}_1 \to M_1$ such that $\xi_1(\hat{p}_1) = \tilde{p}$ and $(\xi_1)_*(\pi_1(\hat{M}_1,\hat{p}_1)) = \hat{H}_1$. Similarly, there exist a manifold $\hat{N}_1$, $\hat{x}_1 \in \hat{N}_1$, and a finite and normal covering map $\zeta_1 : \hat{N}_1 \to N_1$ such that $\zeta_1(\hat{x}_1) = \tilde{x}$ and $(\zeta_1)_*(\pi_1(\hat{N}_1,\hat{x}_1)) = \hat{H}_2$. Since ξ_1 and ζ_1 are finite finite and normal, so is $\xi_1 \times \zeta_1$. By construction, $\hat{\psi}_1 := \psi_1 \circ (\xi_1 \times \zeta_1)$ is a covering map that satisfies $\hat{\psi}_1(\hat{p}_1,\hat{x}_1) = x$ and $(\hat{\psi}_1)_*\big(\pi_1(\hat{M}_1 \times \hat{N}_1, (\hat{p}_1,\hat{x}_1))\big) = H_1 \times H_2$. It follows that $\hat{\psi}_1$ is finite and normal.

By Lemma 5.3, $\psi_0|_{M_0 \times \{\overline{x}\}} : M_0 \times \{\overline{x}\} \to M_1 \times \{\tilde{x}\}$ is a covering map. So $\chi := \frac{1}{2}\rho_1 \circ \psi_0 \circ \iota_{\overline{x}} : M_0 \to M_1$ is a covering map satisfying $\chi(\overline{p}) = \tilde{p}$. Let $I := \hat{H}_1 \cap \chi_*(\pi_1(M_0,\overline{p})) \leq \pi_1(M_1,\tilde{p})$. Then there exist a manifold $\hat{M}_0$, $\hat{p}_0 \in \hat{M}_0$, and covering maps $\hat{\chi} : \hat{M}_0 \to \hat{M}_1$ and $\xi_0 : \hat{M}_0 \to M_0$ such that $\hat{\chi}(\hat{p}_0) = \hat{p}_1$, $\xi_0(\hat{p}_0) = \overline{p}$, $\xi_1 \circ \hat{\chi} = \chi \circ \xi_0$, and $(\xi_1 \circ \hat{\chi})_*(\pi_1(\hat{M}_0,\hat{p}_0)) = I = (\chi \circ \xi_0)_*(\pi_1(\hat{M}_0,\hat{p}_0))$. Note that I is a normal subgroup of $\chi_*(\pi_1(M_0,\overline{p}))$. By the second isomorphism theorem,

$$[\chi_*(\pi_1(M_0,\overline{p})) : I] = [\chi_*(\pi_1(M_0,\overline{p}))\hat{H}_1 : \hat{H}_1] \leq [\pi_1(M_1,\tilde{q}) : \hat{H}_1] < \infty$$

So ξ_0 is finite and normal. The constructions involving $\hat{N}_0$ are different and somewhat simpler: Let $\zeta_0 : \hat{N}_0 \to N_0$ be the universal covering map of N_0 and $\hat{x}_0 \in \zeta_0^{-1}(\overline{x})$. Then there exists a covering map $\hat{\phi} : \hat{N}_0 \to \hat{N}_1$ such that $\hat{\phi}(\hat{x}_0) = \hat{x}_1$ $\zeta_1 \circ \hat{\phi} = \phi \circ \zeta_0$.

The next step is to define a covering map $\hat{\psi}_0 : \hat{M}_0 \times \hat{N}_0 \to \hat{M}_1 \times \hat{N}_1$ so that the diagram

$$\begin{tikzcd} & \hat{M}_0 \times \hat{N}_0 \arrow[dl, "\hat{\psi}_0"'] \arrow[dr, "\xi_0 \times \zeta_0"] & \\ \hat{M}_1 \times \hat{N}_1 \arrow[dr, "\xi_1 \times \zeta_1"'] & & M_0 \times N_0 \arrow[dl, "\psi_0"] \\ & M_1 \times N_1 & \end{tikzcd}$$

commutes. This is done by lifting along $\xi_1 \times \zeta_1$. Note that

$$\psi_0 \circ (\xi_0 \times \zeta_0)(\hat{p}_0, \hat{x}_0) = \psi_0(\overline{p}, \overline{x}) = (\tilde{p}, \tilde{x}) = \xi_1 \times \zeta_1(\hat{p}_1, \hat{x}_1)$$

Since $\xi_1 \times \zeta_1$ is a covering map, a map $\hat{\psi}_0$ satisfying the above diagram and $\hat{\psi}_0(\hat{p}_0, \hat{x}_0) = (\hat{p}_1, \hat{x}_1)$ exists exactly when

$$(\psi_0 \circ (\xi_0 \times \zeta_0))_*\Big(\pi_1(\hat{M}_0 \times \hat{N}_0, (\hat{p}_0, \hat{x}_0))\Big) \leq (\xi_1 \times \zeta_1)_*\Big(\pi_1(\hat{M}_1 \times \hat{N}_1, (\hat{p}_1, \hat{x}_1))\Big)$$

Note that the diagram

$$\begin{tikzcd} \hat{M}_0 \arrow[r, "\xi_0"] \arrow[d, "\hat{\chi}"] & M_0 \arrow[r, "\iota_{\overline{x}}"] \arrow[d, "\chi"] & M_0 \times N_0 \arrow[d, "\psi_0"] \\ \hat{M}_1 \arrow[r, "\xi_1"] & M_1 \arrow[r, "\iota_{\tilde{x}}"] & M_1 \times N_1 \end{tikzcd}$$

commutes. Therefore,

$$\begin{aligned} (\psi_0 \circ \iota_{\overline{x}} \circ \xi_0)_*(\pi_1(\hat{M}_0, \hat{p}_0)) &= (\iota_{\tilde{x}} \circ \xi_1 \circ \hat{\chi})_*(\pi_1(\hat{M}_0, \hat{p}_0)) \\ &= (\iota_{\tilde{x}})_*(I) \\ &\leq (\iota_{\tilde{x}})_*(\hat{H}_1) \\ &= (\iota_{\tilde{x}} \circ \xi_1)_*(\pi_1(\hat{M}_1, \hat{p}_1)) \end{aligned}$$

It follows that

$$\begin{aligned} (\psi_0 \circ (\xi_0 \times \zeta_0))_*\Big(\pi_1(\hat{M}_0 \times \hat{N}_0, (\hat{p}_0, \hat{x}_0))\Big) &= (\psi_0)_*\Big((\iota_{\overline{x}} \circ \xi_0)_*(\pi_1(\hat{M}_0, \hat{p}_0))(\iota_{\overline{p}} \circ \zeta_0)_*(\pi_1(\hat{N}_0, \hat{x}_0))\Big) \\ &= (\psi_0 \circ \iota_{\overline{x}} \circ \xi_0)_*(\pi_1(\hat{M}_0, \hat{p}_0))(\psi_0 \circ \iota_{\overline{p}} \circ \zeta_0)_*(\pi_1(\hat{N}_0, \hat{x}_0)) \\ &= (\psi_0 \circ \iota_{\overline{x}} \circ \xi_0)_*(\pi_1(\hat{M}_0, \hat{p}_0)) \\ &\leq (\iota_{\tilde{x}} \circ \xi_1)_*(\pi_1(\hat{M}_1, \hat{p}_1)) \\ &\leq (\iota_{\tilde{x}} \circ \xi_1)_*(\pi_1(\hat{M}_1, \hat{p}_1))(\iota_{\tilde{p}} \circ \zeta_1)_*(\pi_1(\hat{N}_1, \hat{x}_1)) \\ &= (\xi_1 \times \zeta_1)_*\Big(\pi_1(\hat{M}_1 \times \hat{N}_1, (\hat{p}_1, \hat{x}_1))\Big) \end{aligned}$$

where the equality in the fourth line follows from the fact that $\hat{N}_0$ is simply connected. Thus $\hat{\psi}_0$ exists. Since $\psi_0 \circ (\xi_0 \times \zeta_0)$ and $\xi_1 \times \zeta_1$ are covering maps between manifolds, $\hat{\psi}_0$ must also be a covering map.

It will next be shown that $\hat{\pi}_1 \circ \hat{\psi}_0 = \hat{\phi} \circ \hat{\pi}_0$. Write $\hat{\psi}_0 = (\hat{\alpha}, \hat{\beta})$ for maps $\hat{\alpha} : \hat{M}_0 \times \hat{N}_0 \to \hat{M}_1$ and $\hat{\beta} : \hat{M}_0 \times \hat{N}_0 \to \hat{N}_1$. Then $\pi_1 \circ (\xi_1 \times \zeta_1) \circ \hat{\psi}_0 = \pi_1(\xi_1 \circ \hat{\alpha}, \zeta_1 \circ \hat{\beta}) = \zeta_1 \circ \hat{\beta}$. At the same time, by the definition of $\hat{\psi}_0$, $\pi_1 \circ (\xi_1 \times \zeta_1) \circ \hat{\psi}_0 = \pi_1 \circ \psi_0 \circ (\xi_0 \times \zeta_0) = \phi \circ \pi_0 \circ (\xi_0 \times \zeta_0) = \phi \circ \zeta_0 \circ \hat{\pi}_0 = \zeta_1 \circ \hat{\phi} \circ \hat{\pi}_0$. Thus $\zeta_1 \circ \hat{\beta} = \zeta_1 \circ \hat{\phi} \circ \hat{\pi}_0$. Put differently, $\hat{\beta}$ and $\hat{\phi} \circ \hat{\pi}_0$ are lifts via ζ_1 of the same function. Since $\hat{\beta}(\hat{p}_0, \hat{x}_0) = \hat{\pi}_1 \circ \hat{\psi}_0(\hat{p}_0, \hat{x}_0) = \hat{\pi}_1(\hat{p}_1, \hat{x}_1) = \hat{x}_1 = \hat{\phi}(\hat{x}_0) = \hat{\phi} \circ \hat{\pi}_0(\hat{p}_0, \hat{x}_0)$, it follows that $\hat{\pi}_1 \circ \hat{\psi}_0 = \hat{\beta} = \hat{\phi} \circ \hat{\pi}_0$.

In the case that N_0 is simply connected, $\hat{N}_0 = N_0$ and $\zeta_0 = \mathrm{id}$, and $\xi_0 \times \zeta_0 = \xi_0 \times \mathrm{id}$ is a finite covering map. In the case that the M_i, N_i, and $M_i \times N_i$ are Riemannian manifolds, where $M_0 \times N_0$ has the product metric obtained from M_0 and N_0, and ψ_0, ψ_1, and ϕ are local isometries, each of the manifolds $\hat{M}_i$, $\hat{N}_i$, and $\hat{M}_i \times \hat{N}_i$ may be endowed with the pull-back metrics from the corresponding ξ_i, ζ_i, and $\xi_i \times \zeta_i$, respectively. Since $M_0 \times N_0$ has a product metric, the pull-back metric on $\hat{M}_0 \times \hat{N}_0$ from $\xi_0 \times \zeta_0$ is the product metric obtained from $\hat{M}_0$ and $\hat{N}_0$. Since $\xi_1 \circ \zeta_1 \circ \hat{\psi}_0 = \psi_0 \circ \xi_0 \times \zeta_0$, $\hat{\psi}_1 = \psi_1 \circ (\xi_1 \times \zeta_1)$, and $\zeta_1 \circ \hat{\phi} = \phi \circ \zeta_0$, the maps $\hat{\psi}_0$, $\hat{\psi}_1$, and $\hat{\phi}$ are local isometries.

$\square$

Chapter 6

Deformation into totally geodesic maps

6.1 Riemannian center of mass

A key tool that will be used is the center of mass. According to Berger [Ber], this topic has a long history. It was first used by Cartan to study manifolds of non-positive curvature and has since been used productively by a number of prominent geometers. Its modern renaissance seems to date to the work of Grove–Karcher [GK]. It turns out that the construction works equally well on the universal cover of a complete manifold with no focal points, since the essential property it requires is that the distance squared to each point is strictly convex. This chapter begins with a proof that the center of mass exists. Key properties that will be used are from Corollary 2.5.7(a)-(b). Namely, if N is a complete Riemannian manifold and $q \in N$, then $\mathrm{B}(q, \frac{1}{2}\rho(q))$ is strongly convex and, for each $z \in \mathrm{B}(q, \frac{1}{2}\rho(q))$, the function $\mathrm{d}^2(\cdot, z)|_{\mathrm{B}(q, \frac{1}{2}\rho(q))}$ is strictly convex.

Lemma 6.1.1. *Let N be a Riemannian manifold, $q \in N$, and $Y \subset \mathrm{B}(q, \frac{1}{2}\rho(q))$ a compact and convex set. If $p \in \mathrm{B}(q, \frac{1}{2}\rho(q)) \setminus Y$, let $\gamma : [0,1] \to N$ be any geodesic with $\gamma(0) = p$, $\gamma(1) \in Y$, and $\mathrm{L}(\gamma) = \mathrm{d}(p, Y)$. Then, for each $y \in Y$, the function $t \mapsto \mathrm{d}(\gamma(t), y)$ is strictly decreasing.*

Proof. It suffices to show that $t \mapsto \mathrm{d}^2(\gamma(t), y)$ is strictly decreasing, since $\sqrt{\cdot} : (0,\infty) \to (0,\infty)$ is strictly increasing. This is equivalent to $s \mapsto \mathrm{d}^2(\gamma(1-s), y)$ being strictly increasing. Let $\alpha : [0,1] \to \mathrm{B}(q, \frac{1}{2}\rho(q))$ be the unique minimal geodesic connecting y to $\gamma(1)$. Since Y is convex, $\alpha([0,1]) \subseteq Y$. Since $y \in \mathrm{B}(q, \frac{1}{2}\rho(q))$, the function $\mathrm{d}^2(\cdot, y)$ is strictly convex on $\mathrm{B}(q, \frac{1}{2}\rho(q))$. Since γ realizes the distance from p to Y, the first variation formula implies that $g(\alpha'(1), \gamma'(1)) \le 0$. The result follows by constructing a variation

$$A : [0,\varepsilon) \times [0,1] \to \mathrm{B}\Big(q, \frac{1}{2}\rho(q)\Big)$$

of α by setting $A(s,t) := \exp_y\Big(t \exp_y^{-1}(\gamma(1-s))\Big)$ and applying the first variation formula; this shows that $s \mapsto \mathrm{d}^2(\gamma(1-s), y)$ has non-negative derivative at $s = 0$. By strict convexity, it must be strictly increasing on $[0,1]$.

$\square$

A **mass distribution** is a measurable function $m : Z \to M$, where (Z,ν) is a measure space of total measure one. Whenever m maps into a ball $\mathrm{B}(q,R)$, where $0 < R < \frac{1}{2}\rho(q)$, the function $x \mapsto \int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\nu$, defined on $\mathrm{B}(q,\frac{1}{2}\rho(q))$, is strictly convex and, consequently, attains a unique minimum on $\overline{\mathrm{B}}(q,R)$. The above lemma implies that this minimum must occur inside $\overline{\mathrm{conv}}(m(Z))$ and be a global minimum on $\mathrm{B}(q,\frac{1}{2}\rho(q))$.

Proposition 6.1.2. *Let M be a complete Riemannian manifold, $q \in M$, and $m : Z \to \mathrm{B}(q,\frac{1}{2}\rho(q))$ a mass distribution. Then the function $x \mapsto \int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\nu$ is strictly convex on $\mathrm{B}(q,\frac{1}{2}\rho(q))$.*

Proof. By the construction of $\frac{1}{2}\rho(q)$, for each $z \in Z$, the function $\mathrm{d}^2(\cdot,m(z))$ is strictly convex on $\mathrm{B}(q,\frac{1}{2}\rho(q))$. The result now follows from Lemma 2.4.1.

$\square$

Proposition 6.1.3. *Let M be a complete Riemannian manifold, $q \in M$, and $m : Z \to M$ a mass distribution. Suppose $m(Z) \subseteq \mathrm{B}(q,R)$ for some $0 < R < \frac{1}{2}\rho(q)$. Then the function $x \mapsto \int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\nu$ attains a unique minimum Φ_m on $\mathrm{B}(q,\frac{1}{2}\rho(q))$, which must lie inside $\overline{\mathrm{conv}}(m(Z))$. If $R < \frac{\rho(q)}{6}$, then Φ_m is the unique minimum on N.*

Proof. By the construction of $\frac{1}{2}\rho(q)$, the ball $\mathrm{B}(q,R)$ is strongly convex, so $\overline{\mathrm{B}}(q,R)$ is as well. Since $m(Z) \subseteq \mathrm{B}(q,R)$, $\overline{\mathrm{conv}}(m(Z)) \subseteq \overline{\mathrm{B}}(q,R)$. Since $x \mapsto \int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\nu$ is strictly convex, it attains a unique minimum Φ_m on $\overline{\mathrm{conv}}(m(Z))$. Assume there exists $x \in \mathrm{B}(q,\frac{1}{2}\rho(q)) \setminus \overline{\mathrm{conv}}(m(Z))$ such that $\int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\nu \leq \int_Z \mathrm{d}^2(\Phi_m,m(z))\,\mathrm{d}\nu$. Let $\gamma : [0,1] \to M$ be any geodesic with $\gamma(0) = x$, $\gamma(1) \in \overline{\mathrm{conv}}(m(Z))$, and $\mathrm{L}(\gamma) = \mathrm{d}\big(x, \overline{\mathrm{conv}}(m(Z))\big)$. By Lemma 6.1.1, $t \mapsto \mathrm{d}(\gamma(t),m(z))$ is strictly decreasing for each $z \in Z$, so $\mathrm{d}^2(\gamma(1),m(z)) < \mathrm{d}^2(x,m(z))$, and, consequently, $\int_Z \mathrm{d}^2(\Phi_x,m(z))\,\mathrm{d}\mu \leq \int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\mu < \int_Z \mathrm{d}^2(\gamma(1),m(z))\,\mathrm{d}\mu$. This is a contradiction, which shows that Φ_m is the unique minimum on $\mathrm{B}(q,\frac{1}{2}\rho(q))$. Suppose $R < \frac{\rho(q)}{6}$. Then there cannot exist $x \in N \setminus \overline{\mathrm{B}}(q,\frac{1}{2}\rho(q))$ with $\int_Z \mathrm{d}^2(x,m(z))\,\mathrm{d}\nu \leq \int_Z \mathrm{d}^2(\Phi_m,m(z))\,\mathrm{d}\nu$, since then $\mathrm{d}(x,m(z)) \geq \frac{\rho(q)}{3}$ and $\mathrm{d}(\Phi_m,m(z)) < \frac{\rho(q)}{3}$ for all $z \in Z$. So Φ_m is the unique minimum on N.

$\square$

The point $\Phi_m \in \mathrm{B}(q, \frac{1}{2}\rho(q))$ is called the **center of mass of** m. In Grove–Karcher [GK], its existence is shown using a somewhat different tack, although ultimately the strict convexity of the square of the distance function is the key point in any approach. They prove that, when the mass distribution $m(Z)$ maps into a sufficiently small ball, whose size is given explicitly in terms of upper and lower bounds for the sectional curvature and the injectivity radius, the function $x \mapsto \int_Z \exp_x^{-1}(m(z))\,\mathrm{d}\nu$ has a unique zero. This zero is Φ_m.

Lemma 6.1.4. *Let M be a complete Riemannian manifold, $q \in M$, and $m : Z \to M$ a mass distribution. Suppose $m(Z) \subseteq \mathrm{B}(q,R)$ for some $0 < R < \frac{1}{2}\rho(q)$. Then Φ_m equals the unique minimum of the function $x \mapsto \int_Z \exp_x^{-1}(m(z))\,\mathrm{d}\nu$ on $\mathrm{B}(q,R)$.*

Proof. Let $v \in \mathrm{T}_x\mathrm{B}(x,R)$. For each $z \in Z$, construct a variation $V : (-\varepsilon,\varepsilon) \times [0,1] \to \mathrm{B}(x,R)$ by $V(s,t) := \exp_{m(z)}\big(t \cdot \exp_{m(z)}^{-1}(\gamma_v(s))\big)$. As shown in the proof of Lemma 2.5.5,

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \mathrm{d}^2(\gamma_v(s), m(z)) = -2g\big(v, \exp_q^{-1}(m(z))\big)$$

Therefore, the gradient $\nabla_v \int_Z \mathrm{d}^2(\cdot, m(z))\,\mathrm{d}\nu$ satisfies

$$\begin{aligned}
\nabla_v \int_Z \mathrm{d}^2(\cdot, m(z))\,\mathrm{d}\nu &= \frac{\mathrm{d}}{\mathrm{d}s}|_{s=0} \int_Z \mathrm{d}^2(\gamma(s), m(z))\,\mathrm{d}\nu \\
&= \int_Z \frac{\mathrm{d}}{\mathrm{d}s}|_{s=0}\, \mathrm{d}^2(\gamma(s), m(z))\,\mathrm{d}\nu \\
&= -2\int_Z g\big(v, \exp_x^{-1}(m(z))\big)\,\mathrm{d}\nu \\
&= -2g\big(v, \int_Z \exp_x^{-1}(m(z))\big)\,\mathrm{d}\nu
\end{aligned}$$

Since Φ_m minimizes $x \mapsto \int_Z \mathrm{d}^2(x, m(z))\,\mathrm{d}\nu$, it follows that $g\big(v, \int_Z \exp_{\Phi_m}^{-1}(m(z))\,\mathrm{d}\nu\big) = 0$ for all vectors $v \in \mathrm{T}_{\Phi_m}\mathrm{B}(x,R)$, and consequently $\int_Z \exp_{\Phi_m}^{-1}(m(z))\,\mathrm{d}\nu = 0$. Since Φ_m is the unique minimum on $\mathrm{B}(x,R)$, the function $x \mapsto \int_Z \exp_x^{-1}(m(z))\,\mathrm{d}\nu$ can have no other zeros on $\mathrm{B}(x,R)$.

□

Another useful property is that, roughly speaking, the center of mass in a product space is the product of the centers of mass.

Lemma 6.1.5. *Let M be a complete Riemannian manifold, $q \in M$, $R < \frac{1}{2}\rho(q)$, $m : Z \to \mathrm{B}(q,R)$ a mass distribution, and S a closed and convex subset of* $\mathrm{B}(q,R)$. *Suppose that S is isometric to the product*

$S_1 \times \cdots \times S_n$*, where the* S_i *are convex subsets of* S*. For each* $i = 1,\ldots,n$*, denote by* $\pi_i : S \to S_i$ *the standard projection. If* $m(Z) \subseteq S$*, then* $\Phi_m = (\Phi_{\pi_1 \circ m}, \ldots, \Phi_{\pi_n \circ m})$*.*

Proof. Note that S and the S_i are strongly convex, since $\mathrm{B}(q,R)$ is strongly convex. Note that $\overline{\mathrm{conv}}(\pi_i \circ m(Z)) \subseteq S_i$. Thus $\Phi_{\pi_i \circ m} \in S_i$. It follows that $(\Phi_{\pi_1 \circ m}, \ldots, \Phi_{\pi_n \circ m})$ is identified with a point in S. By the same reasoning, $\Phi_m \in \overline{\mathrm{conv}}(m(Z)) \subseteq S$, which, since S is isometric to $S_1 \times \cdots \times S_n$, means that there exist $s_i \in S_i$ such that $\Phi_m = (s_1, \ldots, s_n)$. Since S and S_i are convex, $\mathrm{d}_S = \mathrm{d}$ on $S \times S$, $\mathrm{d}_{S_i} = \mathrm{d}$ on $S_i \times S_i$, and $\mathrm{d}^2((x_1,\ldots,x_n),(y_1,\ldots,y_n)) = \sum_{i=1}^n \mathrm{d}^2(x_i,y_i)$ for $x_i, y_i \in S_i$. Therefore,

$$\begin{aligned}
\int_Z \mathrm{d}^2(\Phi_m, m(z))\,\mathrm{d}\nu &= \int_Z \mathrm{d}^2\Big((s_1,\ldots,s_n),(\pi_1 \circ m(z),\ldots,\pi_n \circ m(z))\Big)\,\mathrm{d}\nu \\
&= \sum_{i=1}^n \int_Z \mathrm{d}^2(s_i, \pi_i \circ m(z))\,\mathrm{d}\nu \\
&\geq \sum_{i=1}^n \int_Z \mathrm{d}^2(\Phi_{\pi_i \circ m}, \pi_i \circ m(z))\,\mathrm{d}\nu
\end{aligned}$$

with equality if and only if $s_i = \Phi_{\pi_i \circ m}$ for all i. Since Φ_m is the unique minimum of the function $x \mapsto \int_Z \mathrm{d}^2(\cdot, m(z))\,\mathrm{d}z$, it follows that $\Phi_m = (\Phi_{\pi_1 \circ m}, \ldots, \Phi_{\pi_n \circ m})$.

□

It's also worth noting that, since the center of mass is defined only using the distance function, it commutes with isometries.

Lemma 6.1.6. *Let* M *be a complete Riemannian manifold,* $q \in M$*, and* $m : Z \to M$ *a mass distribution. Suppose* $m(Z) \subseteq \mathrm{B}(q,R)$ *for some* $0 < R < \frac{1}{2}\rho(q)$*. Let* $\gamma \in \mathscr{I}(M)$*. Then* $\gamma(\Phi_m) = \Phi_{\gamma \circ m}$*.*

The following will also be useful; it will be applied to show that the center of mass can be used to smooth out a map without leaving a locally convex set.

Lemma 6.1.7. *Let* M *be a complete Riemannian manifold,* $q \in M$*,* $0 < R < \frac{1}{2}\rho(q)$*,* S *a compact and locally convex subset of* $\mathrm{B}(q,R)$*, and* $m : Z \to S$ *a mass distribution. If* $\nu(m^{-1}(S^\circ)) > 0$*, where* S° *is defined with respect to the structure* S *has as a topological manifold, possibly with boundary, then* $\Phi_m \in S^\circ$*.*

Proof. Since S is convex, $\overline{\mathrm{conv}}(m(Z)) \subseteq S$, so $\Phi_m \in S$. Assume $\Phi_m \in \partial S$. Let P be a supporting hyperplane to S at Φ_m. Since $\nu(m^{-1}(S^\circ)) > 0$, $m^{-1}(S^\circ) \neq \emptyset$. There are two unit normal vectors to P at Φ_m; let ν be the one with the property that $g\Big(\exp_{\Phi_m}^{-1}(m(z)), \nu\Big) > 0$ for all $z \in m^{-1}(S^\circ)$. Note that

$g\Big(\exp^{-1}_{\Phi_m}(m(z)), v\Big) \geq 0$ for all $z \in Z$. It follows that $g\Big(\int_Z \exp^{-1}_{\Phi_m}(m(z))\,\mathrm{d}\nu, v\Big) = \int_Z g\Big(\exp^{-1}_{\Phi_m}(m(z)), v\Big)\mathrm{d}\nu \geq \int_{m^{-1}(S^\circ)} g\Big(\exp^{-1}_{\Phi_m}(m(z)), v\Big)\mathrm{d}\nu > 0$. This contradicts the fact that $\int_Z \exp^{-1}_{\Phi_m}(m(z))\,\mathrm{d}\nu = 0$. Therefore, $\Phi_m \in S^\circ$.

□

I will need to use the center of mass mostly when Y is a discrete set of point masses of the form $\{y_i \,\big|\, 1 \leq i \leq k\} \subset \mathrm{B}(q, \frac{1}{2}\rho(q))$, where each y_i has mass $\lambda_i \in [0,1]$ and $\sum_{i=1}^k \lambda_i = 1$. To bring this in line with the definition given above, Z is taken to be a finite set $\{z_i \,\big|\, 1 \leq i \leq k\}$, ν is the discrete measure $\nu(z_i) = \lambda_i$, and $m : Z \to N$ is defined by $m(z_i) = y_i$. In this special case, much of the abstract formalism can be done away with. For any $Y = (y_1, \ldots, y_k) \in \overline{N}^k$ and $\Lambda = (\lambda_1, \ldots, \lambda_k) \in [0,1]^k$, where $\sum_{i=1}^k \lambda_i = 1$, the trio (Z, ν, m) will be implicitly identified with the pair (Y, Λ), and Φ_m will be denoted $\Phi_\Lambda(Y)$. Since $y_i \in \mathrm{B}(q, \frac{1}{2}\rho(q))$, there exists $0 < R < \frac{1}{2}\rho(q)$ such that $y_i \in \mathrm{B}(q, R)$, so $\Phi_\Lambda(Y)$ exists, uniquely minimizes $x \mapsto \sum_{i=1}^k \lambda_i \mathrm{d}^2(x, y_i)$ on N, and lies inside $\overline{\mathrm{conv}}\{y_1, \ldots, y_n\}$. Since the function $\sum_{i=1}^k \lambda_i \mathrm{d}^2(\cdot, y_i)$ varies continuously with y_i and λ_i, this defines a continuous function $\Phi : \overline{N}^k \times \mathscr{Z} \to \mathrm{B}(q, R)$, where $\mathscr{Z}$ denotes the zero set of $(\lambda_1, \ldots, \lambda_k) \mapsto (\sum_{i=1}^k \lambda_i) - 1$, viewed as a function on $[0,1]^k$. The previous lemmas, in this special case, take the following forms.

Lemma 6.1.8. *Let M be a complete Riemannian manifold, $q \in M$, $R < \frac{1}{2}\rho(q)$, and S a closed and convex subset of* $\mathrm{B}(q, R)$. *Suppose that S is isometric to the product $S_1 \times \cdots \times S_n$, where the S_i are convex subsets of S. For each $i = 1, \ldots, n$, denote by $\pi_i : S \to S_i$ the standard projection. If $Y \in S^k$ and $\Lambda \in [0,1]^k$, where $\sum_{i=1}^k \lambda_i = 1$, then $\Phi_\Lambda(Y) = \Big(\Phi_\Lambda(\pi_1(Y)), \ldots, \Phi_\Lambda(\pi_n(Y))\Big)$, where $\pi_i(Y) = \pi_i(y_1, \ldots, y_k) = (\pi_i(y_1), \ldots, \pi_i(y_k))$.*

Lemma 6.1.9. *Let M be a complete Riemannian manifold, $q \in M$, $0 < R < \frac{1}{2}\rho(q)$, S a compact and convex subset of* $\mathrm{B}(q, R)$, *$Y \in S^k$, and $\Lambda \in [0,1]^k$, where $\sum_{i=1}^k \lambda_i = 1$. If there exists $y_i \in S^\circ$ with $\lambda_i > 0$, where S° is defined with respect to the structure S has as a topological manifold, possibly with boundary, then $\Phi_\Lambda(Y) \in S^\circ$.*

Lemma 6.1.10. *Let M be a complete Riemannian manifold, $q \in M$, and $\Lambda \in [0,1]^k$, where $\sum_{i=1}^k \lambda_i =$* 1. *Suppose $Y \in \mathrm{B}(q, \frac{1}{2}\rho(q))^k$. Let $\phi : M \to M$ be an isometry. Then $\phi(\Phi_\Lambda(Y)) = \Phi_\Lambda(\phi(Y))$.*

It follows immediately from the definition of Φ that, for any $\Lambda = (\lambda_1, \ldots, \lambda_k) \in \mathscr{Z}^k$ and $x \in M$, $\Phi_\Lambda(x, \ldots, x) = x$. Moreover, the center of mass is invariant under permutations of indices. That is, if $\sigma : \{1,\ldots,k\} \to \{0,\ldots,k\}$ is any permutation, $\sigma(\Lambda) := (\lambda_{\sigma(1)},\ldots,\lambda_{\sigma(k)})$, and $\sigma(x_1,\ldots,x_k) := (x_{\sigma(1)},\ldots,x_{\sigma(k)})$, then $\Phi_{\sigma(\Lambda)}(\sigma(x)) = \Phi_\Lambda(x)$.

6.2 Construction of a homotopy

A key insight of Cao–Cheeger–Rong is that the center of mass function can be used to glue together homotopies in a way that preserves the property of being totally geodesic [CCR1]. They were working in the context of Riemannian manifolds N with non-positive sectional curvature, but the essential property required is only that, for each $\overline{p} \in \overline{N}$, $\mathrm{d}^2(\cdot,\overline{p})$ is strictly convex. By Theorem 4.1.1, this holds when N has no focal points. In this case, by Corollary 2.5.7, $\rho(\overline{N}) = \infty$ and the center of mass is defined globally on $\overline{N}$.

Theorem 6.2.1. *Let M be a manifold, N a complete Riemannian manifold with no focal points, and $h_1,\ldots,h_k : [a,b]\times M \to N$ continuous functions such that $h_i(a,\cdot) = h_j(a,\cdot)$ for all i, j. Let $\Lambda = (\lambda_1,\ldots,\lambda_k) \in \mathscr{Z}_k$. Then there exists a continuous function $h : [a,b]\times M \to N$ characterized by the following property:*

(*) If $x \in M$, then, for any lifts $\overline{h}_i(\cdot) : [a,b] \to \overline{N}$ of the curves $h_i(\cdot,x)$ satisfying $\overline{h}_i(a) = \overline{h}_j(a)$ for all i, j, $h(\cdot,x) = \pi\circ\Phi_\Lambda(\overline{h}_1(\cdot),\ldots,\overline{h}_k(\cdot))$.

Consequently, this map satisfies $h(a,\cdot) = h_i(a,\cdot)$ for all i. If M is a flat Riemannian torus and each $h_i(b,\cdot)$ is totally geodesic, then $h(b,\cdot)$ is totally geodesic.

Proof. The map h will be defined by property (*), but first it must be shown that $\pi\circ\Phi_\Lambda(\overline{h}_1,\ldots,\overline{h}_k)$ is independent of the choice of basepoint $\overline{h}_i(a) \in \overline{N}$. This will follow from the fact that Φ_Λ commutes with the deck transformations of $\pi : \overline{N} \to N$. For any $x \in M$, let $\overline{y}_1,\overline{y}_2 \in \pi^{-1}(h_i(a,x)) \subseteq \overline{N}$. For each $l = 1,2$, denote by $\overline{h}_{i,l} : [a,b] \to \overline{N}$ the lifts of the paths $h_i(\cdot,x)$ with $\overline{h}_{i,l}(a) = \overline{y}_l$. Denote by $\gamma : \overline{N} \to \overline{N}$ the deck transformation satisfying $\gamma(\overline{y}_1) = \overline{y}_2$. Note that $\gamma\circ h_{i,1} = h_{i,2}$. Since $\gamma\circ\Phi_\Lambda(\overline{h}_{1,1},\ldots,\overline{h}_{k,1}) = \Phi_\Lambda(\gamma\circ\overline{h}_{1,1},\ldots,\gamma\circ\overline{h}_{k,1}) = \Phi_\Lambda(\overline{h}_{1,2},\ldots,\overline{h}_{k,2})$, it follows that $\pi\circ\Phi_\Lambda(\overline{h}_{1,1},\ldots,\overline{h}_{k,1}) = \pi\circ\gamma\circ\Phi_\Lambda(\overline{h}_{1,1},\ldots,\overline{h}_{k,1}) = \pi\circ\Phi_\Lambda(\overline{h}_{1,2},\ldots,\overline{h}_{k,2})$. Thus h is well-defined.

For any $x \in M$ and $\overline{y} \in \pi^{-1}(h_i(a,x))$, let $\overline{h}_i(\cdot) : [a,b] \to \overline{N}$ be lifts of $h_i(\cdot,x)$ satisfying $\overline{h}_i(a) = \overline{y}$ for all i, j. Then $h(a,x) = \pi\circ\Phi_\Lambda(\overline{h}_1(a),\ldots,\overline{h}_k(a)) = \pi\circ\Phi_\Lambda(\overline{y},\ldots,\overline{y}) = \pi(\overline{y}) = h_i(a,x)$. Thus $h(a,\cdot) = h_i(a,\cdot)$. To see that h is continuous, first note that, for each $x \in M$, $h(a,x)$ is contained in an open set V

evenly covered by π. Since $h(a,\cdot) = h_i(a,\cdot)$ is continuous, $h^{-1}(a,V)$ is open. Therefore, the functions $h_i(a,\cdot)$ have lifts on $h^{-1}(a,V)$ that extend to continuous functions $\overline{h}_i : [a,b]\times h^{-1}(a,V) \to \overline{N}$. As the composition of continuous functions, $h = \pi \circ \Phi_\Lambda(\overline{h}_1,\ldots,\overline{h}_n)$ is continuous.

Suppose that M is a flat torus T^n and that each $h_i(b,\cdot)$ is totally geodesic. The goal is to show that $h(b,\cdot)$ is totally geodesic. Since each geodesic $\alpha_0 : [a,b] \to T^n$ may be uniformly closely approximated by some portion of a closed geodesic, it suffices by taking limits to prove that the composition of $h(b,\cdot)$ with any closed geodesic $\alpha : S^1 \cong ([0,1]/\sim) \to T^n$ is totally geodesic. Let $q := h(a,\alpha(0))$ and $[\sigma] := [h(a,\alpha(\cdot))] \in \pi_1(N,q)$, and let $w \in \mathrm{T}_qN$ be the initial vector of the unique geodesic loop in $[\sigma]$. Note that the image of the map induced on $\pi_1([a,b]\times S^1,(a,0))$ by each $h_i(\cdot,\alpha(\cdot)) : [a,b]\times S^1 \to N$ is the subgroup generated by $[\sigma]$. At the same time, $(\pi_{[\sigma]})_*(\pi_1(\tilde{N}_{[\sigma]},w)) = \mathrm{Z}([\sigma])$. It follows that the $h_i(\cdot,\alpha(\cdot))$ lift to maps $\tilde{H}_i : [a,b]\times S^1 \cong [a,b]\times([0,1]/\sim) \to \tilde{N}_{[\sigma]}$ such that $\tilde{H}_i(a,0) = \tilde{H}_j(a,0) = w$ and $\tilde{H}_i(a,\cdot) = \tilde{H}_j(a,\cdot)$ for all i, j. Since each $h_i(b,\cdot)$ is totally geodesic, $\tilde{H}_i(b,\cdot)$ is a closed geodesic, which means that $\tilde{H}_i(b,S^1) \subseteq \tilde{C}_{[\sigma]}$. Let $\overline{H}_i : [a,b]\times[0,1) \to \overline{N}$ be lifts of $\tilde{H}_i|_{[a,b]\times[0,1)}$ to $\overline{N}$ with $\overline{H}_i(a,\cdot) = \overline{H}_j(a,\cdot)$ for all i, j. Then all $\overline{H}_i(b,[0,1))$ lie in the same component C of $\psi^{-1}_{[\sigma]}(\tilde{C}_{[\sigma]})$. By Theorem 4.1.6, C is strongly convex and splits isometrically as $C_0\times\mathbb{R}^{\mathrm{rank}\{[\sigma]\}}$, where $\mathrm{rank}\{[\sigma]\} \in \{0,1\}$ and each $\overline{H}_i(b,\cdot)$ maps into an $\mathbb{R}^{\mathrm{rank}\{[\sigma]\}}$-fiber. It follows that each $\overline{H}_i(b,\cdot)$ takes the form $\overline{H}_i(b,t) = (c_i, d_i + \nu t)$, where $c_i \in C_0$, $\nu, d_i \in \mathbb{R}$, and, by Lemma 3.1.3, all $\overline{H}_i(b,\cdot)$ have the same speed $|\nu|$. This implies, by Lemma 6.1.8, that $\Phi_\Lambda(\overline{H}_1(b,t),\ldots,\overline{H}_k(b,t)) = (\Phi_\Lambda(c_1,\ldots,c_k),[\sum_{i=1}^k \lambda_i d_i]\nu t)$. Therefore, $h(b,\alpha(\cdot)) = \pi\circ\Phi_\Lambda(\overline{H}_1(b,\cdot),\ldots,\overline{H}_k(b,\cdot))$ is a geodesic. It follows that $h(b,\cdot)$ is totally geodesic.

□

Lemma 6.2.2. *Let M be a manifold, N a complete Riemannian manifold with no focal points, $h_1,\ldots,h_k : [a,b]\times M \to N$ continuous functions such that $h_i(a,\cdot) = h_j(a,\cdot)$ for all i, j, $\Lambda = (\lambda_1,\ldots,\lambda_k) \in \mathscr{L}_k$, and $X \subseteq M$ a path-connected set. Let $h : [a,b]\times M \to N$ be the continuous function characterized by property (*) that's guaranteed by Theorem 6.2.1. If each $h_i(b,\cdot)$ is constant on X, then $h(b,\cdot)$ is constant on X.*

Proof. Let $p_k \in X$ for $k = 1,2$, and write $y_k := h(a,p_k)$. Then $y_k = h_i(a,p_k)$ for all i. Choose $\overline{y}_1 \in \pi^{-1}(z_1)$. For each i, let $\overline{h}_{i,1} : [a,b] \to \overline{N}$ denote the lift of $h_i(\cdot,p_1)$ with $\overline{h}_{i,1}(a) = \overline{y}_1$. Since X is path-connected, there exists a path $\sigma : [0,1] \to M$ from p_1 to p_2 such that $\sigma([0,1]) \subseteq X$. Then $h(a,\sigma(\cdot))$

is a path from y_1 to y_2. Since each $h_i(b,\cdot)$ is constant on X, each $h_i(b,\sigma(\cdot))$ is a constant path. Let $\overline{\sigma}:[0,1]\to\overline{N}$ be the lift of $h(a,\sigma(\cdot))$ with $\overline{\sigma}(0)=\overline{y}_1$, and write $\overline{y}_2:=\overline{\sigma}(1)$. Then $\overline{y}_2\in\pi^{-1}(y_2)$. Let $\overline{h}_{i,2}:[a,b]\to\overline{N}$ denote the lift of $h_i(\cdot,p_2)$ with $\overline{h}_{i,2}(a)=\overline{y}_2$. Define a map $H_i:[0,1]\times[a,b]\to N$ by $H_i(s,t):=h_i(t,\sigma(s))$. Since $H_i(0,a)=h_i(a,p_1)=y_1$, H_i lifts to a map $\overline{H}_i:[0,1]\times[a,b]\to\overline{N}$ satisfying $\overline{H}_i(0,a)=\overline{z}_1$. By construction, $\overline{H}_i(0,\cdot)=\overline{h}_{i,1}(\cdot)$; since $\overline{H}_i(a,\cdot)=\overline{\sigma}$, one has $\overline{H}_i(a,1)=\overline{\sigma}(1)=\overline{y}_2$, which means $\overline{H}_i(1,\cdot)=\overline{h}_{i,2}(\cdot)$. Since $H_i(\cdot,b)$ is constant, so is $\overline{H}_i(\cdot,b)$. It follows that $\overline{h}_{i,1}(b)=\overline{H}_i(0,b)=\overline{H}_i(1,b)=\overline{h}_{i,2}(b)$. Therefore, by property (*),

$$\begin{aligned}
h(b,p_1)&=\pi\circ\Phi_\Lambda(\overline{h}_{1,1}(b),\ldots,\overline{h}_{n,1}(b))\\
&=\pi\circ\Phi_\Lambda(\overline{h}_{1,2}(b),\ldots,\overline{h}_{n,2}(b))\\
&=h(b,p_2)
\end{aligned}$$

This shows that $h(b,\cdot)$ is constant on X.

□

The following theorems develop further machinery based upon observations of Cao–Cheeger–Rong [CCR1] about the center of mass. By a well-known theorem of Bieberbach, every compact flat manifold F admits a finite and normal covering by a torus. By averaging over the deck transformation group of such a covering $\pi_0:T^n\to F$, one may prove that, whenever $f:F\to N$ is continuous and $\tilde{f}:=f\circ\psi:T^n\to N$ is homotopic to a totally geodesic map, f is also homotopic to a totally geodesic map. It's worth emphasizing, however, that the given homotopy on T^n may not descend to F. A somewhat more general form of this principle, to finite covering maps that aren't necessarily normal, is stated in Lemma 6.2.3, with the normal case discussed in Remark 6.2.4.

Theorem 6.2.3. *Let M and $\tilde{M}$ be Riemannian manifolds, $\psi:\tilde{M}\to M$ a finite covering map with n sheets, $\Lambda\in\mathscr{Z}_n$, N a complete Riemannian manifold with no focal points, $f:M\to N$ a continuous function, and $\tilde{f}:=f\circ\psi$ the lift of f to $\tilde{M}$. Suppose $\tilde{H}:[a,b]\times\tilde{M}\to N$ is a continuous function with $\tilde{H}(a,\cdot)=\tilde{f}(\cdot)$. Then there exists a continuous function $h:[a,b]\times M\to N$ characterized by the following property:*

(**) If $x\in M$ and $\psi^{-1}(x)=\{x_1,\ldots,x_n\}$, then, for any lifts $\overline{h}_i(\cdot):[a,b]\to\overline{N}$ of the curves $\tilde{H}(\cdot,x_i)$ satisfying $\overline{h}_i(a)=\overline{h}_j(a)$ for all i,j, $h(\cdot,x)=\pi\circ\Phi_\Lambda(\overline{h}_1(\cdot),\ldots,\overline{h}_n(\cdot))$.

Consequently, this map satisfies $h(a,\cdot)=f(\cdot)$.

Proof. For each $x \in M$, let $U_x \subseteq M$ be an open set containing x that's evenly covered by ψ, and, for each $i = 1, \dots, n$, let $U_{x,i} \subseteq \tilde{M}$ be the component of $\psi^{-1}(U_x)$ containing x_i. Let $h_{x,i} : [a,b] \times U_x \to N$ be defined by $h_{x,i}(t,y) := \tilde{H}(t, \psi|_{U_{x,i}}^{-1}(y))$. Then $h_{x,i}(a,\cdot) = f|_{U_x}(\cdot)$ for all i. Let $h_x : [a,b] \times U_x \to N$ be the map guaranteed by Theorem 6.2.1 with respect to the maps $h_{x,i}$. Note that, for each $y \in U_x$, $h_x(\cdot, y)$ is characterized by property (**), which is independent of x and U_x. Therefore, setting $h(t,y) := h_x(t,y)$ for any $x \in N$ such that $y \in U_x$ yields a well-defined map $h : [a,b] \times M \to N$ characterized by property (**). By construction, $h(a,\cdot) = f(\cdot)$.

□

Remark 6.2.4. Suppose the covering map ψ in Theorem 6.2.3 is normal. Denote by $\tilde{\Gamma} = \{\tilde{\gamma}_1, \dots, \tilde{\gamma}_n\}$ the deck transformation group of ψ. The maps $\tilde{H}_i : [a,b] \times \tilde{M} \to N$ defined by $\tilde{H}_i(t,x) := \tilde{H}(t, \tilde{\gamma}_i(x))$ satisfy $\tilde{H}_i(a,\cdot) = \tilde{H}(\cdot)$, and the map $\tilde{h} : [a,b] \times \tilde{M} \to N$ guaranteed by Theorem 6.2.1 with respect to the maps $\tilde{H}_i$ is $\tilde{\Gamma}$-equivariant. Since $\tilde{\Gamma}$ acts transitively on each fiber $\psi^{-1}(x)$, it follows that $\tilde{h} = h \circ \psi$, where $h : [a,b] \times M \to N$ is the map from Theorem 6.2.3. Roughly speaking, h is the average of $\tilde{H}$ over $\tilde{\Gamma}$.

The next result will play an important technical role in the proof of the main theorem. The key idea is that one may glue together locally defined homotopies using a partition of unity and the center of mass.

Theorem 6.2.5. *Let T^n be a flat Riemannian torus and W and M manifolds, where M is connected, and endow $W \times M \times T^n$ with the product metric obtained from any Riemannian metrics on W and M and the given flat metric on T^n. Let N be a complete Riemannian manifold with no focal points and $f : W \times M \times T^n \to N$ a continuous function. Suppose that, around each $w \in W$, there exist an open set V_w and a continuous function $h_w : [a,b] \times V_w \times M \times T^n \to N$ such that $h_w(a,\cdot) = f|_{V_w \times M \times T^n}(\cdot)$, $h_w(b,\cdot)$ is constant along each M-fiber, and $h_w(b,\cdot)$ is totally geodesic along each T^n-fiber. Then there exists a continuous function $H : [a,b] \times W \times M \times T^n \to N$ such that $H(a,\cdot) = f(\cdot)$, $H(b,\cdot)$ is constant along each W-fiber, and $H(b,\cdot)$ is totally geodesic along each T^n-fiber.*

Proof. For each $w \in W$, let $U_w \subseteq W$ be an open set containing w such that $U_w \subset \overline{U}_w \subset V_w$, and let $\{\lambda_w\}$ be a partition of unity subordinate to the open cover $\{U_w\}$ of W. The goal is to define a homotopy $H : [a,b] \times W \times M \to N$ by gluing together the homotopies h_w with weights λ_w. For each $p \in W$, there exists a neighborhood W_p around p such that at most finitely many λ_w are non-zero on W_p. Denote these by $\lambda_{w_1}, \ldots, \lambda_{w_k}$. If $p \notin \mathrm{supp}(\lambda_{w_i})$ for some i, then one may replace W_p with $W_p \cap \mathrm{supp}(\lambda_{w_i})^{\complement}$ and discard λ_{w_i}. In that way, one may, without loss of generality, suppose that $p \in \mathrm{supp}(\lambda_{w_i})$ for each i and, consequently, that $p \in V_{w_i}$. Further shrinking W_p, namely by replacing it with $W_p \cap (\cap_{i=1}^k V_{w_i})$, one may, without loss of generality, suppose that each H_{w_i} is defined on $[0,1] \times W_p \times M \times T^n$. Then $\Lambda_p := (\lambda_{w_1}, \ldots, \lambda_{w_k})$ is a continuous function from W_p into $\mathscr{Z}_k \subset [0,1]^k$.

For each $w \in W_p$, let $H_{p,w} : [a,b] \times \{w\} \times M \times T^n \to N$ be the map characterized by property (*) guaranteed by Theorem 6.2.1, where $\Lambda = \Lambda(w)$ and $h_i = h_{w_i}|_{W_p \times M \times T^n}$. Then $H_{p,w}(a, \cdot) = f|_{\{w\} \times M \times T^n}(a, \cdot)$. Note that Λ is constant along each M-fiber and each T^n-fiber. Because each $H_{p,w}$ is characterized by property (*) along each T^n-fiber, Theorem 6.2.1 states that $H_{p,w}(b, \cdot)$ is totally geodesic along each T^n-fiber. Similarly, by Lemma 6.2.2, $H_{p,w}$ is constant along each M-fiber. Define $H_p : [a,b] \times W_p \times M \times T^n$ by $H_p(t,w,x,y) := H_{p,w}(t,x,y)$. Since the center of mass $\Phi : \overline{N}^k \times \mathscr{Z}_k \to \overline{N}$ is continuous, the proof that H_p is continuous on $[a,b] \times W_p \times M \times T^n$ proceeds along exactly the same lines as in the proof of Theorem 6.2.1.

The desired homotopy $H : [a,b] \times W \times M \times T^n \to N$ will be defined on each $[a,b] \times W_p \times M \times T^n$ to equal H_p. To see that H is well-defined, suppose $W_{p_1} \cap W_{p_2} \neq \emptyset$. Then p_1 and p_2 are in the support of every λ_w that's non-zero on W_{p_1} and, similarly, every λ_w that's non-zero on W_{p_2}. Therefore, p_1 and p_2 determine the same w_i and, perhaps up to a permutation of their elements, $\Lambda_{p_1} = \Lambda_{p_2}$ on $W_{p_1} \cap W_{p_2}$. So $H_{p_1} = H_{p_2}$ on $[a,b] \times (W_{p_1} \cap W_{p_2}) \times M \times T^n$. Since each H_p is continuous, so is H. By construction, $H(a, \cdot) = f(\cdot)$, $H(b, \cdot)$ is constant along each M-fiber, and $H(b, \cdot)$ is totally geodesic along each T^n-fiber.

$\square$

Remark 6.2.6. One may take M to be a single point in Theorem 6.2.5, in which case it becomes a statement about maps $f : W \times T^n \to N$. It will be interpreted this way in the proof of Theorem 6.2.8.

Remark 6.2.7. Theorem 6.2.5, along with most of the results in this section, generalizes to fiber bundles $T^n \to M \xrightarrow{\pi} B$ endowed with Riemannian bundle metrics with flat T^n-fibers. Roughly speaking, a Riemannian bundle metric is a metric defined only along the fibers of the bundle; equivalently, it is the restriction to each fiber of a Riemannian metric on M. In that way, bundle metrics describe the intrinsic geometry of the fibers. Since the proof of the Theorem 6.2.5 nowhere requires the T^n-fibers to be totally geodesic, the result naturally generalizes. However, such a generalization is not necessary for later developments.

From this point on, it will be assumed that M_0, $\mathbb{R}^k$, T^k, and $M_1 \times T^k$ have Riemannian metrics, that the metric h on $\mathbb{R}^k$ is constant and flat, and that the diagram

$$\begin{tikzcd} M_0 \times \mathbb{R}^k \arrow[r, "\pi_0"] \arrow[d, "\psi_0"] & \mathbb{R}^k \arrow[d, "\phi"] \\ M_1 \times T^k \arrow[r, "\pi_1"] & T^k \end{tikzcd} \tag{6.1}$$

commutes isometrically. By definition, this means that $M_0 \times \mathbb{R}^k$ has the product metric obtained from M_0 and $\mathbb{R}^k$ and that ψ_0 and ϕ are Riemannian covering maps, but not necessarily that $M_1 \times T^k$ has a product metric nor that π_1 is a Riemannian submersion. Since ϕ is a Riemannian covering map, the metric on T^k must be flat.

Theorem 6.2.8. *Suppose the diagram* (6.1) *commutes isometrically. Let M be a Riemannian manifold, $\psi_1 : M_1 \times T^k \to M$ a finite Riemannian covering map, N a compact Riemannian manifold with no focal points, $f : M \to N$ a continuous function, and $\tilde{f} := f \circ \psi_1$. Suppose $(\tilde{f} \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p})) = <e>$ for any $(\tilde{p}, \tilde{x}) \in M_1 \times T^k$. Then f is homotopic to a totally geodesic map whose lift to $M_1 \times T^k$ is constant along each M_1-fiber.*

Proof. By Lemma 5.2.1, one may replace M_0 and M_1 with the spaces $\hat{M}_0$ and $\hat{M}_1$, respectively, in diagram (5.3), replace T^k with a possibly larger flat torus, and replace ψ_0, ψ_1, and ϕ with the corresponding covering maps, so that ψ_1 is normal and the diagram

$$\begin{tikzcd} M_0 \times N_0 \arrow[r, "\pi_0"] \arrow[d, "\psi_0"] & N_0 \arrow[d, "\phi"] \\ M_1 \times N_1 \arrow[r, "\pi_1"] & N_1 \end{tikzcd}$$

commutes isometrically. It's worth emphasizing that $M_0 \times \mathbb{R}^k$ still possesses the product metric obtained from M_0 and $\mathbb{R}^k$. In the first stage of the proof, $M_1 \times T^k$ will be taken to have the product metric obtained from M_1 and T^k, which is not necessarily the Riemannian covering metric with respect to ψ_1. For each $\tilde{x} \in M_1$, let $U_{\tilde{x}}$ be a contractible neighborhood of $\tilde{x}$. It will be shown that $\tilde{f}|_{U_{\tilde{x}} \times T^k}$ is homotopic to a map $\tilde{g}_{\tilde{x}} : U_{\tilde{x}} \times T^k \to N$ that's totally geodesic along each T^k-fiber with respect to the product metric. Since $U_{\tilde{x}}$ is contractible, $\tilde{f}|_{U_{\tilde{x}} \times T^k}$ is homotopic to a map $\tilde{f}_{1,\tilde{x}}$ that's constant along each $U_{\tilde{x}}$-fiber. The loop map Υ may be used to construct a homotopy from $\tilde{f}_{1,\tilde{x}}$ to a map $\tilde{f}_{2,\tilde{x}}$ that equals the totally geodesic map guaranteed by the flat torus theorem on each each T^k-fiber. Applying Theorem 6.2.5 to these sets $U_{\tilde{x}}$ and the homotopies from $\tilde{f}|_{U_{\tilde{x}}\times T^k}$ to $\tilde{f}_{2,\tilde{x}}$, one obtains a homotopy from $\tilde{f}$ to a map $\tilde{g}$ that's totally geodesic along each T^k-fiber with respect to the product metric. This is equivalent to $\tilde{g}\circ\iota_{\tilde{p}} : T^k \to N$ being totally geodesic for each $\tilde{p}\in M_1$.

The map $\tilde{g}$ lifts canonically to a map $\tilde{G} : M_1\times T^k \to \tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$. Since $\tilde{g}$ is totally geodesic along each T^k-fiber, $\tilde{G}(M_1\times T^k)\subseteq \tilde{C}_{[\sigma_1],\ldots,[\sigma_k]}$. Fix $(\tilde{p},\tilde{x})\in M_1\times T^k$. Since $\tilde{G}_*\Big(\pi_1(M_1\times\{\tilde{x}\},(\tilde{p},\tilde{x}))\Big) = <e>$, $\tilde{G}|_{M_1\times\{\tilde{x}\}}$ lifts to a map $\overline{G} : M_1\times\{\tilde{x}\}\to\overline{N}$ such that

$$\overline{G}(M_1\times\{\tilde{x}\})\subseteq \min([\sigma_1],\ldots,[\sigma_k]) = \psi^{-1}_{[\sigma_1],\ldots,[\sigma_k]}(\tilde{C}_{[\sigma_1],\ldots,[\sigma_k]})$$

Since $\min([\sigma_1],\ldots,[\sigma_k])$ is strongly convex, $\overline{G}$ is homotopic to a constant map via a homotopy that remains inside that minimal set. Composing the loop map Υ with the map $\psi_{[\sigma_1],\ldots,[\sigma_k]}\circ\overline{G}$, one obtains a homotopy from $\tilde{g}$ to a map $\tilde{g}_0$ that's totally geodesic along each T^k-fiber and constant along each M_0-fiber. By Lemma 5.10(a), $\tilde{g}_0$ is totally geodesic with respect to the pull-back metric from ψ_1.

According to Remark 6.2.4, applying Lemma 6.2.3 to the homotopy from $\tilde{f}$ to $\tilde{g}_0$ yields a continuous map $h : [a,b]\times M\to N$ such that its lift $\tilde{h} : [a,b]\times\tilde{M}\to N$ is characterized by property (*) with respect to the maps $\tilde{H}_i(t,\tilde{p},\tilde{x}) := \tilde{H}(t,\tilde{\gamma}_i(\tilde{p},\tilde{x}))$, where $\tilde{\Gamma} = \{\tilde{\gamma}_1,\ldots,\tilde{\gamma}_n\}$ is the deck transformation group of ψ_1. It will be shown that $\tilde{h}(b,\cdot)$ is totally geodesic along each T^k-fiber with respect to the product metric on $M_1\times T^k$. By the general theory of covering spaces, corresponding to each $\tilde{\gamma}_i$ there is a deck transformation $\overline{\gamma}_i$ of $\psi_1\circ\psi_0$ such that the diagram

$$\begin{array}{ccc} M_0\times\mathbb{R}^k & \xrightarrow{\overline{\gamma}_i} & M_0\times\mathbb{R}^k \\ \big\downarrow{\scriptstyle\psi_0} & & \big\downarrow{\scriptstyle\psi_0} \\ M_1\times T^k & \xrightarrow{\tilde{\gamma}_i} & M_1\times T^k \end{array}$$

commutes. By Lemma 5.6, each $\overline{\gamma}_i$ is of the form $\overline{\gamma}_i = \overline{\alpha}_i\times\overline{\beta}_i$ for some $\overline{\alpha}_i\in\mathscr{I}(M_0)$ and $\overline{\beta}_i\in\mathscr{I}(\mathbb{R}^k)$. For each i, one has that

$$\begin{aligned} \pi_1\circ\tilde{\gamma}_i\circ\psi_0 &= \pi_1\circ\psi_0\circ\overline{\gamma}_i \\ &= \phi\circ\pi_0\circ\overline{\gamma}_i \\ &= \phi\circ\overline{\beta}_i\circ\pi_0 \end{aligned}$$

Let $\tilde{p}_1, \tilde{p}_2 \in M_1$, $\tilde{x} \in T^k$, and $\overline{x} \in \phi^{-1}(\tilde{x})$. By Lemma 5.3, $\psi_0|_{M_0\times\{\overline{x}\}} : M_0 M_0 \times \{\overline{x}\} \to M_1 \times \{\tilde{x}\}$ is surjective. Thus there exist $\overline{p}_j \in M_0$, $j = 1,2$, such that $\psi_0(\overline{p}_j, \overline{x}) = (\tilde{p}_j, \tilde{x})$. It follows that

$$\begin{aligned}\pi_1 \circ \tilde{\gamma}_i(\tilde{p}_j, \tilde{x}) &= \pi_1 \circ \tilde{\gamma}_i \circ \psi_0(\overline{p}_j, \overline{x}) \\ &= \phi \circ \overline{\beta}_i \circ \pi_0(\overline{p}_j, \overline{x}) \\ &= \phi \circ \overline{\beta}_i(\overline{x})\end{aligned}$$

This shows that $\tilde{\gamma}_i(\tilde{p}_1, \tilde{x})$ and $\tilde{\gamma}_i(\tilde{p}_2, \tilde{x})$ lie in the same M_1-fiber. Since $\tilde{g}_0$ is constant along each M_1-fiber, one has

$$\begin{aligned}\tilde{H}_i(b, \tilde{p}_1, \tilde{x}) &= \tilde{H}(b, \tilde{\gamma}_i(\tilde{p}_1, \tilde{x})) \\ &= \tilde{g}_0 \circ \tilde{\gamma}_i(\tilde{p}_1, \tilde{x}) \\ &= \tilde{g}_0 \circ \tilde{\gamma}_i(\tilde{p}_2, \tilde{x}) \\ &= \tilde{H}(b, \tilde{\gamma}_i(\tilde{p}_2, \tilde{x})) \\ &= \tilde{H}_i(b, \tilde{p}_2, \tilde{x})\end{aligned}$$

That is to say, each $\tilde{H}_i(b,\cdot)$ is constant along each M_1-fiber. Let $\sigma : [a,b] \to T^k$ be a geodesic starting at $\tilde{x} \in T^k$, and let $\tilde{\sigma} := \iota_{\tilde{p}} \circ \sigma$ for $\tilde{p} \in M_1$. Choose $(\overline{p}, \overline{x}) \in \psi^{-1}(\tilde{p}, \tilde{x})$, and let $\overline{\sigma} : [a,b] \to M_0 \times \mathbb{R}^k$ be the lift of $\tilde{\sigma}$ satisfying $\overline{\sigma}(a) = (\overline{p}, \overline{x})$ and $\psi \circ \overline{\sigma} = \tilde{\sigma}$. Since $\tilde{H}_i(b,\cdot)$ is constant along each M_1-fiber and $\pi_1 \circ \tilde{\sigma} = \pi_1 \circ \psi \circ \overline{\sigma} = \phi \circ \pi_0 \circ \overline{\sigma} = \phi \circ \pi_0 \circ \iota_{\overline{p}} \circ \pi_0 \circ \overline{\sigma} = \pi_1 \circ \psi \circ \iota_{\overline{p}} \circ \pi_0 \circ \overline{\sigma}$,

$$\tilde{H}_i(b, \tilde{\sigma}(\cdot)) = \tilde{H}_i(b, \psi \circ \iota_{\overline{p}} \circ \pi_0 \circ \overline{\sigma}(\cdot))$$

Arguing as before, one has that

$$\begin{aligned}\tilde{H}_i(b, \psi \circ \iota_{\overline{p}} \circ \pi_0 \circ \overline{\sigma}(\cdot)) &= \tilde{H}(b, \psi \circ \overline{\gamma}_i \circ \iota_{\overline{p}} \circ \pi_0 \circ \overline{\sigma}(\cdot)) \\ &= \tilde{H}(b, \psi \circ (\overline{\alpha}_i \times \overline{\beta}_i) \circ \iota_{\overline{p}} \circ \pi_0 \circ \overline{\sigma}(\cdot)) \\ &= \tilde{H}\big(b, \psi(\overline{\alpha}_i(\overline{p}), \overline{\beta}_i \circ \pi_0 \circ \overline{\sigma}(\cdot))\big)\end{aligned}$$

Since $\phi \circ \pi_0 \circ \overline{\sigma} = \pi_1 \circ \psi \circ \overline{\sigma} = \pi_1 \circ \iota_{\tilde{p}} \circ \tilde{\sigma} = \tilde{\sigma}$ and ϕ is a local isometry, $\pi_0 \circ \overline{\sigma}$ is a geodesic, so $\overline{\beta}_i \circ \pi_0 \circ \overline{\sigma}$ is a geodesic in $\mathbb{R}^k$. Since $M_0 \times \mathbb{R}^k$ has a product metric, $(\overline{\alpha}_i(\overline{p}), \overline{\beta}_i \circ \pi_0 \circ \overline{\sigma}(\cdot))$ is also a geodesic, which since ψ is a local isometry means that $\psi(\overline{\alpha}_i(\overline{p}), \overline{\beta}_i \circ \pi_0 \circ \overline{\sigma}(\cdot))$ is as well. Since $\tilde{H}(b,\cdot) = \tilde{g}_0(\cdot)$ is totally geodesic, $\tilde{H}_i(b, \tilde{\sigma}(\cdot))$ is a geodesic. So each $\tilde{H}_i(b, \iota_{\tilde{p}}(\cdot))$ is totally geodesic.

It follows from Lemma 5.10(a) that each $\tilde{H}_i(b, \iota_{\tilde{p}}(\cdot))$ is totally geodesic. Theorem 6.2.1 now implies that each $\tilde{h}(b, \iota_{\tilde{p}}(\cdot))$ is totally geodesic. Moreover, by Lemma 6.2.2, $\tilde{h}(b, \cdot)$ is constant along each M_1-fiber. Another application of Lemma 5.10(a) shows that $\tilde{h}(b, \cdot)$ is totally geodesic with respect to the pull-back metric, which implies that $h(b, \cdot)$ is totally geodesic.

□

Remark 6.2.9. When M_0 is compact and simply connected, $\pi_1(M_1, \tilde{p})$ is finite. By Theorem 3.1.5, $\pi_1(N, y)$ is torsion-free. It follows that $(\tilde{f} \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p})) = <e>$.

Theorem 6.2.10. *Suppose the diagram* (6.1) *commutes isometrically. Let M be a Riemannian manifold, $\psi_1 : M_1 \times T^k \to M$ a finite Riemannian covering map, N a complete Riemannian manifold with no focal points, and* $[F]$ *a homotopy class of maps from M to N such that, for any $f \in [F]$ and $\tilde{f} := f \circ \psi_1$, $(\tilde{f} \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p})) = <e>$ for any $(\tilde{p}, \tilde{x}) \in M_1 \times T^k$. Then the set of totally geodesic maps in* $[F]$ *is path-connected.*

Proof. By Lemma 6.2.4, one may take ψ_1 to be normal. Suppose $f, g \in [F]$ are are totally geodesic. It follows from Lemma 5.10(b) that $\tilde{f}$ is constant along each M_1-fiber and totally geodesic along each T^k-fiber. Thus $\tilde{f}$ descends to a totally geodesic map $\hat{f} : T^k \to N$. It's similarly true that $\tilde{g} := g \circ \psi_1$ descends to a totally geodesic map $\hat{g} : T^k \to N$. Theorem 4.1.14 states that $\hat{f}$ and $\hat{g}$ are homotopic via a homotopy through totally geodesic maps. This homotopy can be extended to a homotopy through totally geodesic maps on $M_1 \times T^k$ that are constant along each M_1-fiber. Using Theorem 6.2.5, as discussed in Remark 6.2.4, this homotopy descends to one on M. It follows exactly as in the proof of Theorem 6.2.8 that this homotopy is through totally geodesic maps.

□

By Corollary 5.2, when M is a compact manifold with non-negative Ricci curvature, one may suppose that the diagram (6.1) commutes isometrically for a compact and simply connected M_0 and that there exists a finite Riemannian covering map $\psi_1 : M_1 \times T^k \to M$. As in Remark 6.2.9, $\tilde{f} := f \circ \psi_1$ satisfies $(\tilde{f} \circ \iota_{\tilde{z}})_*(\pi_1(M_1, \tilde{p})) = <e>$ for all $(\tilde{p}, \tilde{z}) \in M_1 \times T^k$. Theorem 1.4(c) is therefore a corollary of Theorem 6.2.8, and Theorem 1.4(a) is a corollary of Theorem 6.2.10.

6.3 Generalization to product domains

This section contains the proof of Theorem 1.5. Much of the proof of the next result resembles that of Theorem 6.2.5.

Lemma 6.3.1. *Let M and N be Riemannian manifolds, and let $S \subseteq N$ be a closed and convex set. If $f : M \to Y$ is continuous, then f is homotopic to a smooth map $u : M \to N$ whose image lies in S° via a homotopy $H : [0,1] \times M \to S$.*

Proof. If $\partial S = \emptyset$, there is little to prove, so one may suppose $\partial S \neq \emptyset$. The key idea is to construct a continuous function $R : [0,1] \times S \to S$ such that $R(0,\cdot) = \mathrm{id}(\cdot)$ and $R(1,S) \subseteq S^\circ$. Assume this is possible for the moment. Define $H_0(s,x) := R(s, f(x))$. Then $H_0(1,M) \subset S^\circ$. Since S is closed and convex, S is a submanifold of N with totally geodesic interior and possibly non-smooth boundary. In particular, S° is a smooth manifold. Therefore, $H_0(1,\cdot)$ can be approximated by a smooth function $u : M \to S^\circ$ such that $\mathrm{d}(H_0(1,x), u(x)) < \mathrm{inj}(S)$. This means that $x \overset{V}{\mapsto} \exp^{-1}_{H_0(1,x)}(u(x))$ is a well-defined vector field along $H_0(1,M)$. Note that $V(x) \in \mathrm{T}_{H_0(1,x)}S \subseteq \mathrm{T}_{H_0(1,x)}N$. The map $H_1 : [0,1] \times M \to S$ defined by $H_1(s,x) := \exp_{H_0(1,x)}(sV(x))$ is a homotopy from $H_0(1,\cdot)$ to $u(\cdot)$. So $H := H_1 \cdot H_0$ is the desired homotopy, which completes the proof, modulo the existence of R.

Since S is a topological manifold with boundary, each point $x \in S$ is contained in an open ball $\mathrm{B}(x,\varepsilon(x))$ such that $\mathrm{B}(x,\varepsilon(x)) \cap S$ is homeomorphic to an open subset of the closed upper half-space in $\mathbb{R}^n$. Without loss of generality, one may suppose that $0 < \varepsilon(x) \leq \rho(x)$. Using these homeomorphisms, one may construct a family of deformation retractions $R_x : [0,1] \times \mathrm{B}(x,\varepsilon(x)) \cap S \to \mathrm{B}(x,\varepsilon(x))$ onto subsets of $\mathrm{B}(x,\varepsilon(x)) \cap S^\circ$. Let $\{\lambda_x\}$ be a partition of unity subordinate to the open cover $\{\mathrm{B}(x, \frac{1}{2}\varepsilon(x)) \cap S \mid x \in S\}$ of S. The map R will be defined by gluing together the maps R_x using the center of mass with weights λ_x. Around each $p \in S$, there exists an open set V_p such that at most finitely many λ_x are non-zero on V_p. Denote these by $\lambda_{x_1}, \ldots, \lambda_{x_k}$. If $p \notin \mathrm{supp}(\lambda_{x_i})$ for some i, then one may replace V_p with $V_p \cap \mathrm{supp}(\lambda_{x_i})^\complement$ and discard λ_{x_i}. In that way, one may, without loss of generality, suppose that $p \in \mathrm{supp}(\lambda(x_i))$ for each i and, consequently, that $p \in \mathrm{B}(x_i, \frac{1}{2}\varepsilon(x_i))$. Further shrinking V_p, namely by replacing it with $V_p \cap \left(\cap_{i=1}^k \mathrm{B}(x_i, \frac{1}{2}\varepsilon(x_i)) \right)$, one may, without loss of generality, suppose that each R_{x_i} is defined on $[0,1] \times V_p \cap S$. Then $\Lambda_p := (\lambda_{x_1}, \ldots, \lambda_{x_k})$ is a continuous function from V_p into $\mathscr{X}_k$.

For each $x \in S$ and $k \in \mathbb{N}$, the function Φ_Λ is well-defined and continuous on $\mathrm{B}(x,\rho(x))^k \times \mathscr{X}^k$.

The desired map R will be defined on each $[0,1]\times V_p \cap S$ to equal $\Phi_{\Lambda_p}(R_{x_1},\ldots,R_{x_k})$. To see that R is well-defined, suppose $V_{p_1}\cap V_{p_2}\neq\emptyset$. Then p_1 and p_2 are in the support of every λ_x that's non-zero on V_{p_1} and, similarly, every λ_x that's non-zero on V_{p_2}. Therefore, p_1 and p_2 determine the same x_i and, perhaps up to a permutation of their elements, $\Lambda_{p_1}=\Lambda_{p_2}$ on $V_{p_1}\cap V_{p_2}$. So $\Phi_{\Lambda_{p_1}}(R_{x_1},\ldots,R_{x_k}) = \Phi_{\Lambda_{p_2}}(R_{x_1},\ldots,R_{x_k})$ on $[0,1]\times(V_{p_1}\cap V_{p_2}\cap S)$. Since each R_p is continuous, so is R. By construction, $R(0,\cdot)=\mathrm{id}(\cdot)$. By Lemma 6.1.7, $R(1,S)\subseteq S^\circ$.

□

Theorem 6.3.2. *Suppose the diagram* (6.1) *commutes isometrically. Let M be a Riemannian manifold, $\psi_1 : M_1\times T^k\to M$ a finite Riemannian covering map, N a compact Riemannian manifold with no focal points, W a manifold, and $f : W\times M\to N$ a continuous function. Write $\tilde{f} := f\circ(\mathrm{id}\times\psi_1) : W\times M_1\times T^k\to N$. Endow $W\times M$ with the product metric obtained from any Riemannian metric on W and the given metric on M. Suppose $\tilde{f}_*(\pi_1(\{w\}\times M_1\times\{\tilde{x}\},(w,\tilde{p},\tilde{x})) = <e>$ for any $(w,\tilde{p},\tilde{x})\in W\times M_1\times T^k$. Then f is homotopic to a map that's totally geodesic on each M-fiber.*

Proof. The argument proceeds much like the proof of Theorem 6.2.8. By Lemma 6.2.4, one may, without loss of generality, take ψ_1 to be a normal covering map. Around each $w\in W$, there exists a contractible neighborhood U_w, and each $f|_{U_w\times M}$ is homotopic to a map f_w that's constant on each U_w-fiber. By Theorem 6.2.8, each f_w is homotopic to a map g_w that's constant on each U_w-fiber, totally geodesic on each M-fiber, and whose lift $\tilde{g}_w$ to $M_1\times T^k$ is constant on each M_1-fiber. Let $h_w : [a,b]\times U_w\times M\to N$ be a homotopy from $f|_{U_w\times M}$ to g_w, and define $\tilde{h}_w : [a,b]\times U_w\times M_1\times T^k\to N$ by $\tilde{h}_w(t,v,\tilde{p},\tilde{x}) := h_w(t,v,\psi_1(\tilde{p},\tilde{x}))$. Since $\tilde{h}_w(b,\cdot)$ is constant along each U_w-fiber and, because ψ_1 is a local isometry, totally geodesic along each $M_1\times T^k$-fiber, one may now show, by the argument in the proof of Lemma 5.10(b), that $\tilde{h}_w(b,\cdot)$ is totally geodesic along each T^k-fiber with respect to the product metric on $U_w\times M\times T^k$. Theorem 6.2.5 states that there exists a continuous function $\tilde{H} : [a,b]\times W\times M\times T^n\to N$ such that $\tilde{H}(a,\cdot)=f(\cdot)$, $\tilde{H}(b,\cdot)$ is constant along each W-fiber, and $\tilde{H}(b,\cdot)$ is totally geodesic along each T^n-fiber with respect to the product metric. Since ψ_1 is normal, Theorem 6.2.3 guarantees the existence of a continuous map $h : [a,b]\times W\times M\to N$ such that its lift $\tilde{h} : [a,b]\times W\times M\to N$ is characterized by property (*) with respect to the maps $\tilde{H}_i(t,w,\tilde{p},\tilde{x}) := \tilde{H}(t,w,\tilde{\gamma}_i(\tilde{p},\tilde{x}))$, where $\tilde{\Gamma}=\{\tilde{\gamma}_1,\ldots,\tilde{\gamma}_n\}$ is the deck transformation group of ψ_1. Arguing exactly as in the proof of Theorem 6.2.8, one may show that each $\tilde{H}_i(b,\cdot)$ is constant along each M_1-fiber and totally geodesic along each T^k-fiber with respect to the product metric and, consequently, that $\tilde{h}(b,\cdot)$ has those same properties. Thus $\tilde{h}(b,\cdot)$ is totally geodesic with respect to the pull-back metric, and h is a homotopy from $h(a,\cdot) = f(\cdot)$ to the totally geodesic map $h(b,\cdot)$.

□

Since the map $\tilde{h}(b,\cdot)$ is constant along each M_1-fiber, it descends to a map $\hat{h} : W \times T^k \to N$ that's totally geodesic along each T^k-fiber. By Lemma 6.3.1, one may perturb the canonical lift of $\hat{h}$ to $\tilde{N}_{[\sigma_1],\ldots,[\sigma_k]}$ through a homotopy that remains inside $\tilde{C}_{[\sigma_1],\ldots,[\sigma_k]}$ for all time. Using the loop map Υ, one may use this to define a homotopy from $\hat{h}$ to a smooth map through maps that are totally geodesic on the T^k-fibers at each time. This may be used to define a homotopy from $h(b,\cdot)$ to a smooth map through maps that are totally geodesic on the M-fibers at each time. One obtains the following.

Corollary 6.3.3. *Suppose the diagram* (6.1) *commutes isometrically. Let M be a Riemannian manifold, $\psi_1 : M_1 \times T^k \to M$ a finite Riemannian covering map, N a compact Riemannian manifold with no focal points, W a manifold, and $f : W \times M \to N$ a continuous function. Write $\tilde{f} := f \circ (\mathrm{id} \times \psi_1) : W \times M_1 \times T^k \to N$. Endow $W \times M$ with the product metric obtained from any Riemannian metric on W and the given metric on M. Suppose $\tilde{f}_*(\pi_1(\{w\} \times M_1 \times \{\tilde{x}\}, (w, \tilde{p}, \tilde{x})) = <e>$ for any $(w, \tilde{p}, \tilde{x}) \in W \times M_1 \times T^k$. Then f is homotopic to a smooth map that's totally geodesic on each M-fiber.*

Theorem 1.5 follows from Corollary 6.3.3 and Corollary 5.2. In fact, it generalizes with little additional effort to fiber bundles $M \to W \xrightarrow{\pi} B$ that are locally isometrically trivial and whose M-fibers have non-negative Ricci curvature.

Chapter 7

Energy-minimizing versus totally geodesic

7.1 Energy, length, and intersection

The principal goal of this chapter is to describe when a map f from a compact Riemannian manifold M with non-negative Ricci curvature into a Riemannian manifold N with no conjugate points is totally geodesic. This is done in Theorem 7.4.4 and Theorem 7.4.5, which together state that f is totally geodesic if and only if $\mathrm{E}(f) = \mathrm{vol}(M)e_T$, where e_T is the energy density of an affine surjection $T : T^k \to T^m$, T^k is a flat Riemannian torus determined by M, and $(T^m, \|\cdot\|)$ is a flat semi-Finsler torus that depends on the homotopy class of f. Furthermore, $\mathrm{vol}(M)e_T$ is a lower bound for the energy of maps homotopic to f, one which, by Kleiner's counterexample to the flat torus theorem [Kle], may not be realized within that homotopy class, even for compact N. It turns out that this builds upon work of Croke [Cr1] and Croke–Fathi [CF] about energy and intersection, and some effort, beyond what's truly necessary for the larger goal of generalizing the results of Eells–Sampson and Hartman to manifolds with no focal points, is made to explore that relationship. This occasions mention of the beta and gamma functions, which appear in results having to do with the volume of spheres.

A key observation of Croke [Cr1] is that, because the energy density e_f of a C^1 map between Riemannian manifolds is a trace, it may be computed at each $x \in M$ as an average over the unit sphere S_xM. Specifically, $e_f(x) = \frac{1}{2}\mathrm{trace}_x(<\cdot,\cdot>_{f^{-1}(\mathrm{T}N)})$ is equal to the average of the length of the push-forward of vectors in S_xM.[1]

Lemma 7.1.1. *(Croke) Let M and N be Riemannian manifolds with $n = \mathrm{dim}(M) > 0$ and $f : M \to N$ a C^1 map. For each $x \in M$, $e_f(x) = \frac{n}{2c_{n-1}} \int_{S_xM} \|f_*(w)\|^2 \,\mathrm{d}\mu_{S_xM}$, where c_{n-1} denotes the volume of S^{n-1}.*

[1] The result in [Cr1] differs from the one here by a multiplicative factor of n. The presentation here agrees with that in [CF]. The results in the former were also stated for smooth maps, but the arguments only require C^1 regularity.

By Lemma 7.1.1,

$$\mathrm{E}(f) = \frac{n}{2c_{n-1}} \int_M \int_{S_xM} \|f_*(w)\|^2 \, \mathrm{d}\mu_{S_xM} \, \mathrm{d}\mu_M = \frac{n}{2c_{n-1}} \int_{\mathrm{S}M} \|f_*(w)\|^2 \, \mathrm{d}\mu_{\mathrm{S}M}$$

where $\mu_{\mathrm{S}M}$ denotes the Liouville measure on the unit sphere bundle $\mathrm{S}M \subset \mathrm{T}M$. The last equality was used in [CF] and follows from the fact that each S_qM, when endowed with the induced metric it inherits from the canonical flat metric on T_qM, is isometric to the standard round sphere $S^{n-1} \subset \mathbb{R}^n$, and with respect to these identifications the Liouville measure is locally the product measure $\mu_{\mathrm{S}M} = \mu_M \times \mu_{S^{n-1}}$.

Note that the expression in Lemma 7.1.1 depends only on the induced Finsler norm on N, and therefore may be taken as the definition of the energy density of a map into a Finsler manifold, or even a semi-Finsler manifold. Here, the definition of a semi-Finsler manifold parallels that of a Finsler manifold, except it is endowed with a semi-norm, rather than a norm, on each tangent space. If M is an n-dimensional Riemannian manifold with $n = \dim(M) > 0$, N is a semi-Finsler manifold, $f : M \to N$ is C^1, and $x \in M$, then the **energy density of** f **at** x is defined to be

$$e_f(x) := \frac{n}{2c_{n-1}} \int_{S_xM} \|f_*(w)\|^2 \, \mathrm{d}\mu_{S_xM}$$

and the **energy of** f to be $\mathrm{E}(f) := \int_M e_f(x) \, \mathrm{d}\mu_M$. In [J3], Jost generalized the notions of energy and energy density to maps from measure spaces into metric spaces. Centore [Cen] showed that the above definitions agree with Jost's in the Finsler setting and yield a sensible notion of energy, in that its minimizers have vanishing Laplacian.[2]

In parallel with energy density, the **length density of** f **at** x is defined to be

$$\ell_f(x) := \sqrt{\frac{n}{2c_{n-1}}} \int_{S_xM} \|f_*(w)\| \, \mathrm{d}\mu_{S_xM}$$

and the **length of** f to be $\mathrm{L}(f) := \int_M \ell_f(x) \, \mathrm{d}\mu_M$. One may check that, for $M = S^1$ and with respect to the convention $c_0 = \mathrm{vol}(\{-1, 1\}) = 2$, this agrees with the usual length of a loop in a Finsler manifold. In general, by the same reasoning as above,

$$\mathrm{L}(f) = \sqrt{\frac{n}{2c_{n-1}}} \int_M \int_{S_xM} \|f_*(w)\| \, \mathrm{d}\mu_{S_xM} \, \mathrm{d}\mu_M = \sqrt{\frac{n}{2c_{n-1}}} \int_{\mathrm{S}M} \|f_*(w)\| \, \mathrm{d}\mu_{\mathrm{S}M}$$

[2]The definitions given here are actually special cases of those in [Cen], where M as well as N is allowed to be a Finsler manifold, but a proper treatment of Centore's work, including what is meant by the Laplacian of a map between Finsler manifolds, is beyond the scope of this dissertation.

By the Cauchy-Schwarz inequality, one has $e_f \geq \frac{1}{c_{n-1}} \ell_f^2$ and $\mathrm{E}(f) \geq \frac{1}{\mathrm{vol}(SM)} \mathrm{L}^2(f) = \frac{1}{c_{n-1}\mathrm{vol}(M)} \mathrm{L}^2(f)$.

As an application of Lemma 7.1.1 and Santaló's formula, Croke [Cr1] characterizes the energy of a map $f: M \to N$ between Riemannian manifolds, where M is compact and has boundary $\partial M \neq \emptyset$, as an integral over the inward-pointing portion of the sphere bundle along ∂M. The notation in the following is the same as in Theorem 2.3.3.

Theorem 7.1.2. *(Croke) Let (M, g) be a compact Riemannian manifold with $n = \dim(M) > 0$ and boundary $\partial M \neq \emptyset$. Let N be a Riemannian manifold and $f: M \to N$ a C^1 map. Then* $\mathrm{E}(f) = \frac{n}{2c_{n-1}} \int_{\mathrm{S}^+\partial M} \mathrm{E}(\varsigma_w) g(w, \nu) \, \mathrm{d}\mu_{\mathrm{S}^+\partial M}$.

The proof of Theorem 7.1.2 doesn't use the Riemannian structure of N in any essential way, so the conclusions hold when N is only a semi-Finsler manifold. The results should also generalize to any reasonable notion of a compact manifold M with corners. In the application to come, they will only be needed when M is a **parallelotope** in flat Euclidean space, that is, when M is the set $P_V := \{\sum_{i=1}^k t_i v_i \,\big|\, 0 \leq t_i \leq 1\}$ generated by a set of linearly independent vectors $V = \{v_1, \ldots, v_k\} \subset \mathbb{R}^n$. In that case, by approximating M from within by smooth manifolds with boundary that slightly round off its corners, one obtains the following.

Corollary 7.1.3. *Let $P_V \subset \mathbb{R}^n$ be the parallelotope corresponding to a linearly independent set of vectors $V = \{v_1, \ldots, v_n\} \subset \mathbb{R}^n$ for $n > 0$. Let N be a semi-Finsler manifold and $f: P_V \to N$ a C^1 map. Then* $\mathrm{E}(f) = \frac{n}{2c_{n-1}} \int_{\mathrm{S}^+\partial P_V} \mathrm{E}(\varsigma_w) g(w, \nu) \, \mathrm{d}\mu_{\mathrm{S}^+\partial P_V}$.

To be clear, in the statement of Corollary 7.1.3, ν denotes the inward-pointing normal vector field along each $(n-1)$-face of ∂P_V, and f being C^1 means that f_* exists and is continuous on $\mathrm{T}^+ P_V := \{w \in \mathrm{T}(P_V) \,\big|\, \gamma_w(t) \in P_V \text{ for some } t > 0\}$. The set $\mathrm{S}^+\partial P_V$ and curves ς_w have their original definitions with respect to ν. One may remark that versions of Theorem 7.1.2 and Corollary 7.1.3 hold for length as well as energy.

Lemma 7.1.1 was used by Croke–Fathi [CF] to give a sufficient condition for when a map $f: M \to N$ is a **homothety**, that is, when the image of every geodesic in M is homotopy-minimizing in N and there exists $c \geq 0$ such that $f^*(h) = cg$. Note that every homothety is totally geodesic and that a homothety can be non-constant only when M has no conjugate points. To do this, they first associate to f a quantity called its intersection. For any $w \in SM$ and $t \geq 0$, denote by $\phi_t(w)$ the minimum length of all paths in N basepoint-fixed homotopic to $f \circ \gamma_w|_{[0,t]}$. Then the **intersection**

of f is the non-negative real number $I(f) := \inf_{t>0} \frac{1}{t} \int_{SM} \phi_t(w)\, d\mu_{SM}$. Croke–Fathi show that $I(f) = \lim_{t\to\infty} \frac{1}{t} \int_{SM} \phi_t(w)\, d\mu_{SM}$. They also show that $I(f)$ depends only on the homotopy class of f. If $[F]$ is a homotopy class of maps from M to N, then $I([F])$ will denote the intersection of any map in $[F]$.

Theorem 7.1.4. *(Croke–Fathi) Let M and N be Riemannian manifolds, where M is compact and has dimension $n > 0$. Let $[F]$ be a homotopy class of maps from M to N and $f \in [F]$ a C^1 map. Then the following hold:*

(a) $\mathrm{E}(f) \geq \frac{n}{2c_{n-1}^2 \mathrm{vol}(M)} I^2([F])$*; and*

(b) *If* $\mathrm{E}(f) = \frac{n}{2c_{n-1}^2 \mathrm{vol}(M)} I^2([F])$*, then f is a homothety.*

The next sections will build upon Theorem 7.1.4 in the case that M satisfies a diagram of the form (5.1), which, for instance, is the case when M is compact and has non-negative Ricci curvature.

7.2 Asymptotic norm of a periodic metric on $\mathbb{Z}^m$

A key tool used by Burago–Ivanov [BurI] in their proof of the Hopf conjecture was a certain norm on $\mathbb{R}^m$ corresponding to a $\mathbb{Z}^m$-equivariant Riemannian metric. This norm, which was introduced by Burago [Bura], captures the large-scale geometry of the metric and, in fact, approximates its distance function up to an additive constant. In that sense, viewed from far enough away, $\mathbb{R}^m$ with such a metric is nearly indistinguishable from a normed space. When a metric on $\mathbb{R}^m$ is not necessarily Riemannian, but is at least induced by a length structure, this norm still exists with the same properties. In fact, a semi-norm on $\mathbb{R}^m$ exists corresponding to any $\mathbb{Z}^m$-equivariant metric on $\mathbb{Z}^m$. The results in this section are all taken from [BBI].

Theorem 7.2.1. *Let d be a metric on $\mathbb{Z}^m$ that's equivariant under the natural action of $\mathbb{Z}^m$ on itself by addition. Then there exists a unique semi-norm $\|\cdot\|$ on $\mathbb{R}^m$ such that $\|v\| = \lim_{n\to\infty} \frac{\mathrm{d}(0,nv)}{n}$ for all $v \in \mathbb{Z}^m$. This semi-norm has the property that $\frac{\mathrm{d}(0,v)}{\|v\|} \to 1$ uniformly as $\|v\| \to \infty$.*

The semi-norm guaranteed by the above theorem is called the **asymptotic semi-norm** of d and, when it's a norm, the **asymptotic norm**. This is a norm, for example, whenever d is the orbit metric of a free and properly discontinuous action, with compact quotient, of $\mathbb{Z}^m$ on a length space. The principal application of the asymptotic semi-norm here will be to the orbit metric of the action of an Abelian subgroup of the fundamental group of a manifold on its universal cover. Let N be a Riemannian manifold, $y \in N$, and G a finitely generated, free, and Abelian subgroup of $\pi_1(N,y)$. Then G may be isomorphically, although not canonically, embedded as the integer lattice in $\mathbb{R}^m$. If $\pi : \overline{N} \to N$ is the universal covering map of N, then G acts canonically on $\overline{N}$, and one may define the **orbit metric of** G, denoted $\mathrm{d}_G : G \times G \to [0,\infty)$, by

$$\mathrm{d}_G([\gamma_1],[\gamma_2]) := \mathrm{d}_{\overline{N}}([\gamma_1]\cdot\overline{y},[\gamma_2]\cdot\overline{y})$$

for any choice of $\overline{y} \in \pi^{-1}(y)$. Since the action of G on $\overline{N}$ is by isometries, d_G is both G-invariant and, indeed, independent of the choice of $\overline{y}$. With respect to a fixed isomorphism $\iota : G \to \mathbb{Z}^m \subset \mathbb{R}^m$, d_G induces a $\mathbb{Z}^m$-equivariant orbit metric $\mathrm{d}_{\mathbb{Z}^m,\iota} : \mathbb{Z}^m \times \mathbb{Z}^m \to [0,\infty)$, defined by $\mathrm{d}_{\mathbb{Z}^m,\iota}(\iota([\gamma_1]),\iota([\gamma_2])) = \mathrm{d}_G([\gamma_1],[\gamma_2])$. By Theorem 7.2.1, the corresponding asymptotic semi-norm $\|\cdot\|_{G,\iota}$ exists. Moreover, for any other isomorphism $\jmath : G \to \mathbb{Z}^m$, $(\mathbb{R}^m,\|\cdot\|_{G,\iota})$ and $(\mathbb{R}^m,\|\cdot\|_{G,\jmath})$ are isomorphic. When the group G and isomorphism ι are understood, they will sometimes be suppressed in the notation. That is, $\mathrm{d}_{\mathbb{Z}^m,\iota}$ and $\|\cdot\|_{G,\iota}$ will sometimes be written $\mathrm{d}_{\mathbb{Z}^m}$ and $\|\cdot\|$, respectively.

Lemma 7.2.2. *Let N be a Riemannian manifold, $y_1, y_2 \in N$, $\alpha : [a,b] \to N$ a path from y_1 to y_2, and $[\gamma] \in \pi_1(N,y_1)$. Then*

$$|\,\mathrm{d}(\overline{y}_2, A_{[\alpha]}([\gamma])\cdot\overline{y}_2) - \mathrm{d}(\overline{y}_1,[\gamma]\cdot\overline{y}_1)| \leq 2\mathrm{L}(\alpha)$$

for all $\overline{y}_1 \in \pi^{-1}(y_1)$ and $\overline{y}_2 \in \pi^{-1}(y_2)$.

Proof. Since the deck transformation group of $\pi : \overline{N} \to N$ acts by isometries, it suffices to prove the result for the given $\overline{y}_1 \in \pi^{-1}(y_1)$ and any choice of $\overline{y}_2 \in \pi^{-1}(y_2)$. Let $\overline{\alpha}_1 : [a,b] \to \overline{N}$ be the lift of α with $\overline{\alpha}_1(a) = \overline{y}_1$. Since $\overline{\alpha}_1(b) \in \pi^{-1}(y_2)$, one may without loss of generality take $\overline{y}_2 = \overline{\alpha}_1(b)$. Write $\beta := \alpha^{-1}\cdot\gamma\cdot\alpha$, and let $\overline{\beta} : [0,1] \to \overline{N}$ be, up to reparameterization, the lift of β with $\overline{\beta}(0) = \overline{y}_2$. By definition,

$$\begin{aligned} A_{[\alpha]}([\gamma])\cdot\overline{y}_2 &= [\alpha^{-1}\cdot\gamma\cdot\alpha]\cdot\overline{y}_2 \\ &= [\beta]\cdot\overline{y}_2 \\ &= \overline{\beta}(1) \end{aligned}$$

Let $\overline{\gamma} : [0,1] \to \overline{N}$ be, up to reparameterization, the lift of γ with $\overline{\gamma}(0) = \overline{y}_1$. Let $\overline{\alpha}_2 : [a,b] \to \overline{N}$ be the lift of α with $\overline{\alpha}_2(0) = [\gamma]\cdot\overline{y}_1 = \overline{\gamma}(1)$. By construction, $\overline{\alpha}_2(b) = \overline{\beta}(1)$. Therefore,

$$\begin{aligned}
\mathrm{d}(\overline{y}_2, A_{[\alpha]}([\gamma])\cdot\overline{y}_2) &\le \mathrm{d}(\overline{y}_2,\overline{y}_1) + \mathrm{d}(\overline{y}_1,[\gamma]\cdot\overline{y}_1) + \mathrm{d}([\gamma]\cdot\overline{y}_1, A_{[\alpha]}([\gamma])\cdot\overline{y}_2)\\
&= \mathrm{d}(\overline{\alpha}_1(b),\overline{\alpha}_1(a)) + \mathrm{d}(\overline{y}_1,[\gamma]\cdot\overline{y}_1) + \mathrm{d}(\overline{\alpha}_2(a),\overline{\alpha}_2(b))\\
&\le \mathrm{d}(\overline{y}_1,[\gamma]\cdot\overline{y}_1) + 2\mathrm{L}(\alpha)
\end{aligned}$$

Since $[\gamma] = A_{[\alpha^{-1}]}(A_{[\alpha]}([\gamma]))$, the same argument shows that

$$\begin{aligned}
\mathrm{d}(\overline{y}_1,[\gamma]\cdot\overline{y}_1) &= \mathrm{d}(\overline{y}_1, A_{[\alpha^{-1}]}(A_{[\alpha]}([\gamma]))\cdot\overline{y}_1)\\
&\le \mathrm{d}(\overline{y}_2, A_{[\alpha]}([\gamma])\cdot\overline{y}_2) + 2\mathrm{L}(\alpha^{-1})\\
&= \mathrm{d}(\overline{y}_2, A_{[\alpha]}([\gamma])\cdot\overline{y}_2) + 2\mathrm{L}(\alpha)
\end{aligned}$$

These combine to yield the desired inequality.

□

Lemma 7.2.3. *Let N be a Riemannian manifold, $y_1, y_2 \in N$, $\alpha : [a,b] \to N$ a path from y_1 to y_2, and G a finitely generated, free, and Abelian subgroup of $\pi_1(N, y_1)$. Let $\imath : G \to \mathbb{Z}^m$ be an isomorphism and $\jmath := \imath \circ A_{[\alpha]}^{-1}$. Then $\|\jmath\circ A_{[\alpha]}([\gamma])\|_{A_{[\alpha]}(G),\jmath} = \|\imath([\gamma])\|_{G,\imath}$ for all $[\gamma] \in G$. Consequently, $\|\cdot\|_{A_{[\alpha]}(G),\jmath} = \|\cdot\|_{G,\imath}$ on $\mathbb{R}^m$.*

Proof. Since $A_{[\alpha]} : \pi_1(N,y_1) \to \pi_1(N,y_2)$ is a group isomorphism, $A_{[\alpha]}(G)$ is finitely generated, free, and Abelian. Since $\jmath$ is a group isomorphism, the semi-norm $\|\cdot\|_{A_{[\alpha]}(G),\jmath}$ is well-defined. Let $\overline{y}_i \in \pi^{-1}(y_i)$ for each $i = 1,2$. For each $n \in \mathbb{N}$ and $[\gamma] \in G$, Lemma 7.2.2 states that

$$|\,\mathrm{d}_{\overline{N}}(\overline{y}_2, A_{[\alpha]}(n[\gamma])\cdot\overline{y}_2) - \mathrm{d}_{\overline{N}}(\overline{y}_1, n[\gamma]\cdot\overline{y}_1)| \le 2\mathrm{L}(\alpha)$$

Therefore,

$$\begin{aligned}
\Big|\|\jmath\circ A_{[\alpha]}([\gamma])\|_{A_{[\alpha]}(G),\jmath} - \|\imath([\gamma])\|_{G,\imath}\Big| &= \Big|\lim_{n\to\infty}\frac{\mathrm{d}_{\mathbb{Z}^m,\jmath}(0, n\jmath\circ A_{[\alpha]}([\gamma]))}{n} - \lim_{n\to\infty}\frac{\mathrm{d}_{\mathbb{Z}^m,\imath}(0, n\imath([\gamma]))}{n}\Big|\\
&= \lim_{n\to\infty}\frac{1}{n}\Big|\mathrm{d}_{A_{[\alpha]}(G)}([e], A_{[\alpha]}(n[\gamma])) - \mathrm{d}_G([e], n[\gamma])\Big|\\
&= \lim_{n\to\infty}\frac{1}{n}\Big|\mathrm{d}_{\overline{N}}(\overline{y}_2, A_{[\alpha]}(n[\gamma])\cdot\overline{y}_2) - \mathrm{d}_{\overline{N}}(\overline{y}_1, n[\gamma]\cdot\overline{y}_1)\Big|\\
&\le \lim_{n\to\infty}\frac{2\mathrm{L}(\alpha)}{n}\\
&= 0
\end{aligned}$$

It follows that $\|\jmath \circ A_{[\alpha]}([\gamma])\|_{A_{[\alpha]}(G),\jmath} = \|\iota([\gamma])\|_{G,\iota}$. Note that $\iota \circ A_{[\gamma]}^{-1} \circ \jmath^{-1}$ is the identity map on $\mathbb{Z}^m$. It follows that $\|\cdot\|_{A_{[\alpha]}(G),\jmath} = \|\cdot\|_{G,\iota}$ on $\mathbb{Z}^m$, and, consequently, on all of $\mathbb{R}^m$.

□

In other words, if G a finitely generated, free, and Abelian subgroup of $\pi_1(N,y)$, then, with respect to an isomorphism $G \to \mathbb{Z}^m$ and up to the identification above, all subgroups equivalent to G define the same asymptotic semi-norm on $\mathbb{R}^m$.

Lemma 7.2.4. *Let N be a Riemannian manifold, $y \in N$, G a finitely generated, free, and Abelian subgroup of $\pi_1(N,y)$, and $\iota : G \to \mathbb{Z}^m$ an isomorphism. Then $\mathrm{d}_{\mathbb{Z}^m,\iota}(0,v) \geq \|v\|_{G,\iota}$ for all $v \in \mathbb{Z}^m$.*

Proof. This follows from the triangle inequality. Since $\mathrm{d}_{\mathbb{Z}^m,\iota}$ is a metric, $\mathrm{d}_{\mathbb{Z}^m,\iota}(0,nv) \leq n\,\mathrm{d}_{\mathbb{Z}^m,\iota}(0,v)$ for all $v \in \mathbb{Z}^m$. Therefore, $\|v\|_{G,\iota} = \lim_{n\to\infty} \frac{\mathrm{d}_{\mathbb{Z}^m,\iota}(0,nv)}{n} \leq \lim_{n\to\infty} \frac{n\,\mathrm{d}_{\mathbb{Z}^m,\iota}(0,v)}{n} = \mathrm{d}_{\mathbb{Z}^m,\iota}(0,v)$.

□

7.3 The Finsler torus and affine surjection associated to $[F]$

As in Chapter 6, it will hereafter be assumed that M_0, $\mathbb{R}^k$, T^k, and $M_1 \times T^k$ have Riemannian metrics, that the metric h on $\mathbb{R}^k$ is constant and flat, and that the diagram (6.1) commutes isometrically. Recall that, by definition, this means that $M_0 \times \mathbb{R}^k$ has the product metric obtained from M_0 and $\mathbb{R}^k$ and that ψ_0 and ϕ are Riemannian covering maps, but not necessarily that $M_1 \times T^k$ has a product metric nor that π_1 is a Riemannian submersion. In particular, since ϕ is a Riemannian covering map, the metric on T^k is flat. It will be assumed that M is a Riemannian manifold of dimension $n > 0$ and that $\psi_1 : M_1 \times T^k \to M$ is a finite Riemannian covering map; the number of sheets of ψ_1 will be denoted by $\#(\psi_1) < \infty$. The manifolds M_i will be taken to have dimension $l \geq 0$, so that $n = k + l$. It will also be assumed that M_0 is compact, which by Lemma 5.3 forces M_1 and M to be compact as well. According to Corollary 5.2, these assumptions are satisfied when M is compact and has non-negative Ricci curvature.

Let M be as described above, N a Riemannian manifold, and $[F]$ a homotopy class of maps from M to N. For each $f \in [F]$, let $\tilde{f} : M_1 \times T^k \to N$ and $\overline{f} : M_0 \times \mathbb{R}^k \to N$ be the compositions $\tilde{f} := f \circ \psi_1$ and $\overline{f} := \tilde{f} \circ \psi_0$. The homotopy class $[F]$ under consideration will always be supposed to have the property that $(\tilde{f} \circ \iota_{\tilde{x}})_*(\pi_1(M_1,\tilde{p})) = <e>$ for any $f \in [F]$ and $(\tilde{p},\tilde{x}) \in M_1 \times T^k$. When M has non-negative Ricci curvature, the manifold M_0 may be assumed to be simply connected and M_1 to have finite fundamental group, in which case this condition is satisfied whenever the groups $\pi_1(N, y)$ are torsion-free. By Theorem 3.1.5, this is the case when N is aspherical and, in particular, when N has no conjugate points. Let $0 \le m \le k$ equal the rank of $\tilde{f}_*\Big(\pi_1(M_1 \times T^k, (\tilde{p}, \tilde{x}))\Big)$ for all $f \in [F]$ and $(\tilde{p}, \tilde{x}) \in M_1 \times T^k$.

As an application of the asymptotic semi-norm of a $\mathbb{Z}^m$-equivariant metric, a semi-Finsler torus $(T^m, \|\cdot\|)$ and an affine surjection $T : T^k \to T^m$ will be associated to $[F]$. This construction is not canonical, but depends up to affine isometry on a choice of point in $M_1 \times T^k$, representative of $[F]$, and group isomorphism with $\mathbb{Z}^m$. Fix $(\tilde{p}, \tilde{x}) \in M_1 \times T^k$, $\overline{x} \in \phi^{-1}(\tilde{x})$, and $f_0 \in [F]$. Let $p := \psi_1(\tilde{p}, \tilde{x})$. Since h is constant, one may, without loss of generality, suppose that ϕ is the quotient map induced by a lattice Γ_1 equal to the span, with integer coefficients, of a set of vectors $V = \{v_1, \ldots, v_k\}$. Each geodesic $t \mapsto \overline{x} + tv_i$, $t \in [0, 1]$, descends via ϕ to a closed geodesic s_i based at $\tilde{x}$, and $\{[s_1], \ldots, [s_k]\}$ is a minimal generating set for $\pi_1(T^k, \tilde{x})$. Write $y := f_0(p) = \tilde{f}_0(\tilde{p}, \tilde{x})$, $G := (\tilde{f}_0)_*\Big(\pi_1(M_1 \times T^k, (\tilde{p}, \tilde{x}))\Big) \subseteq \pi_1(N, y)$, and $\sigma_i := \tilde{f}_0 \circ \iota_{\tilde{p}} \circ s_i$. Since $(\tilde{f}_0 \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p})) = <e>$,

$$G = (\tilde{f}_0 \circ \iota_{\tilde{x}})_*(\pi_1(M_1, \tilde{p}))(\tilde{f}_0 \circ \iota_{\tilde{p}})_*(\pi_1(T^k, \tilde{x})) = < [\sigma_1], \ldots, [\sigma_m] >$$

Thus G is an Abelian subgroup of $\pi_1(N, y)$ of rank m. Let $H := \mathrm{span}_{\mathbb{Z}}\{[s_1], \ldots, [s_m]\}$. For reasons that will become clear toward the end of the section, I will suppose that V has the following two properties:

(A) $(\tilde{f}_0 \circ \iota_{\tilde{p}})_*|_H : H \to G$ is an isomorphism; and

(B) $[\sigma_i] = e$ for each $m + 1 \le i \le k$.

These may be assumed without loss of generality. To attain (A), one need only reorder the v_i. Once (A) holds, one may achieve (B) by replacing each v_i, $m + 1 \le i \le k$, with $v_i - \sum_{j=1}^m d_{ij} v_j$, where the $d_{ij} \in \mathbb{Z}$ are the unique coefficients such that $[\sigma_i] = \sum_{j=1}^m d_{ij} [\sigma_j]$. One may check that, at the end of this process, it's still true that $\Gamma_1 = \mathrm{span}_{\mathbb{Z}} V$. Note that (A) and (B) are together equivalent to the condition that $[\sigma_i] = \sum_{j=1}^m \delta_{ij} [\sigma_j]$ for each $1 \le i \le k$, where δ_{ij} is the Kronecker delta.

Let d_G denote the orbit metric on G obtained from the canonical action of G on $\overline{N}$. For a fixed isomorphism $\iota : G \to \mathbb{Z}^m \subset \mathbb{R}^m$, let $\mathrm{d}_{\mathbb{Z}^m}$ denote the orbit metric on $\mathbb{Z}^m$ corresponding to ι and $\|\cdot\|_{[F],\iota}$ the corresponding asymptotic semi-norm on $\mathbb{R}^m$. Let $T^m := \mathbb{R}^m / \mathbb{Z}^m$ be endowed with the constant semi-Finsler metric $\|\cdot\|_{[F],\iota}$. By (A), $\{[\sigma_1], \ldots, [\sigma_m]\}$ is a minimal generating set for

G; for each $1 \le i \le m$, let $w_i := \iota([\sigma_i]) \in \mathbb{Z}^m$. Let $\overline{T}_{[F],\iota} : \mathbb{R}^k \to \mathbb{R}^m$ be the affine map satisfying $\overline{T}_{[F],\iota}(\overline{x} + v_i) = \sum_{j=1}^m \delta_{ij} w_j$. Then $\overline{T}_{[F],\iota}$ descends to an affine surjection $T_{[F],\iota} : T^k \to T^m$. One may check that $\|\cdot\|_{[F],\iota}$ and $T_{[F],\iota}$ are well-defined, in the sense that they are independent of the choice of $\overline{x} \in \phi^{-1}(\tilde{x})$. Moreover, up to equivalence by isometries in the following sense, they are determined by the homotopy class $[F]$.

Lemma 7.3.1. *Fix $(\tilde{q}, \tilde{z}) \in M_1 \times T^k$, $g_0 \in [F]$, and an isomorphism $\jmath : (\tilde{g}_0)_*\Big(\pi_1(M_1 \times T^k, (\tilde{q}, \tilde{z}))\Big) \to \mathbb{Z}^m$. Denote by $\|\cdot\|_{[F],\jmath}$ and $T_{[F],\jmath}$ the corresponding semi-Finsler metric and affine surjection. Then there exist isometries $\Psi : T^k \to T^k$ and $\Phi : (T^m, \|\cdot\|_{[F],\iota}) \to (T^m, \|\cdot\|_{[F],\jmath})$ such that $T_{[F],\jmath} = \Phi \circ T_{[F],\iota} \circ \Psi^{-1}$.*

Proof. Let $\overline{\Psi} : \mathbb{R}^k \to \mathbb{R}^k$ be the translation that takes $(\tilde{p}, \tilde{x})$ to $(\tilde{q}, \tilde{z})$. Then $\overline{\Psi}$ descends to the desired isometry Ψ. Fix any path $\tilde{\alpha} : [a, b] \to M_1 \times T^k$ from $(\tilde{p}, \tilde{x})$ to $(\tilde{q}, \tilde{z})$ and any homotopy $H : M \times [0, 1] \to N$ from f_0 to g_0. Denote by $\tilde{H} : M_1 \times T^k \times [0, 1] \to N$ the lift $\tilde{H}(\cdot, t) := H(\psi_1(\cdot), t)$. Then $\tilde{H}$ is a homotopy from $\tilde{f}_0$ to $\tilde{g}_0$. Define $\beta : [0, 1] \to N$ by $\beta(t) := \tilde{H}(\tilde{\alpha}(t), t)$ and $\iota_0 : (\tilde{g}_0)_*\Big(\pi_1(M_1 \times T^k, (\tilde{q}, \tilde{z}))\Big) \to \mathbb{Z}^m$ by $\iota_0 := \iota \circ A_{[\beta]}^{-1}$. Denote by $\|\cdot\|_{[F],\iota_0}$ and $T_{[F],\iota_0}$ the semi-Finsler metric and affine surjection corresponding to $(\tilde{q}, \tilde{z})$, g_0, and ι_0. The composition $\overline{\Phi}_0 := \iota_0 \circ A_{[\beta]} \circ \iota^{-1} : \mathbb{Z}^m \to \mathbb{Z}^m$ extends to a linear isomorphism $\overline{\Phi}_0 : (\mathbb{R}^m, \|\cdot\|_{[F],\iota}) \to (\mathbb{R}^m, \|\cdot\|_{[F],\iota_0})$. By Lemma 7.2.3, $\overline{\Phi}_0$ is an isometry, so $\overline{\Phi}_0$ descends to an isometry $\Phi_0 : (T^m, \|\cdot\|_{[F],\iota}) \to (T^m, \|\cdot\|_{[F],\iota_0})$. One may check that $T_{[F],\iota_0} = \Phi_0 \circ T_{[F],\iota} \circ \Psi^{-1}$. The composition $\overline{\Phi}_1 := \jmath \circ \iota_0^{-1}$ also extends to a linear isometry $(\mathbb{R}^m, \|\cdot\|_{[F],\iota_0}) \to (\mathbb{R}^m, \|\cdot\|_{[F],\jmath})$, which in turn descends to an isometry $\Phi_1 : (T^m, \|\cdot\|_{[F],\iota_0}) \to (T^m, \|\cdot\|_{[F],\jmath})$ such that $T_{[F],\jmath} = \Phi_1 \circ T_{[F],\iota_0}$. The proof is completed by setting $\Phi := \Phi_1 \circ \Phi_0$.

□

As the point $(\tilde{p}, \tilde{x}) \in M_1 \times T^k$, map $f_0 \in [F]$, and isomorphism $\iota : (\tilde{f}_0)_*\Big(\pi_1(M_1 \times T^k, (\tilde{p}, \tilde{x}))\Big) \to \mathbb{Z}^m$ will remain fixed, they will typically be suppressed in the notation. By Lemma 7.3.1, there is little danger in any confusion, as important geometric quantities associated to $T_{[F]}$, namely $e_{T_{[F]}}$ and $\ell_{T_{[F]}}$, will not change if any of those are varied. When the homotopy class $[F]$ is understood, it will also be suppressed. That is, it should be understood that $\|\cdot\| = \|\cdot\|_{[F]} = \|\cdot\|_{[F],\iota}$ and $T = T_{[F]} = T_{[F],\iota}$.

Lemma 7.3.2. *Let M, N, and $[F]$ be as above, and denote by $(T^m, \|\cdot\|)$ and $T : T^k \to T^m$ the corresponding semi-Finsler torus and affine surjection. Let $\overline{F}_0 : M_0 \times \mathbb{R}^k \to \overline{N}$ be a lift of $\overline{f}_0$ of the form guaranteed by Lemma 5.11. Then there exists $D \ge 0$ such that, for any $\tilde{z} \in T^k$, $\tilde{v} \in \mathrm{T}_{\tilde{z}} T^k$, and $\overline{\gamma} : [0, \infty) \to M_0 \times \mathbb{R}^k$ satisfying $\phi \circ \pi_0 \circ \overline{\gamma} = \gamma_{\tilde{v}}$, the following hold:*

(a) $\|T_*(t\tilde{v})\| \leq \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)) + D$ *for all* $t \geq 0$*; and*

(b) $\|T_*(\tilde{v})\| = \lim\limits_{t\to\infty} \dfrac{\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0))}{t}$.

Proof. **(a)** By Lemma 5.11, there exists $D_0 \geq 0$ such that

$$\mathrm{d}_{\overline{N}}(\overline{F}_0(\overline{q}_0, \overline{z}_0), \overline{F}_0(\overline{q}_1, \overline{z}_1)) \leq D_0 \tag{7.1}$$

whenever $\mathrm{d}_{\mathbb{R}^k}(\overline{z}_0, \overline{z}_1) \leq \mathrm{diam}(T^k)$. Fix $(\overline{p}, \overline{x}) \in \psi_0^{-1}(\tilde{p}, \tilde{x})$. Since Γ_1 is $\mathrm{diam}(T^k)$-dense in $\mathbb{R}^k$, so is the lattice $\overline{x} + \Gamma_1 := \{\overline{x} + u \,|\, u \in \Gamma_1\}$. Thus there exist $u_0, u_1 \in \Gamma_1$ such that

$$\mathrm{d}_{\mathbb{R}^k}(\pi_0 \circ \overline{\gamma}(it), \overline{x} + u_i) \leq \mathrm{diam}(T^k) \tag{7.2}$$

for each $i = 0, 1$. Let $\tilde{\alpha}_i : [0,1] \to M_1 \times T^k$ be defined by $\tilde{\alpha}_i(s) := (\tilde{p}, \phi(\overline{x} + su_i))$. Since $u_i \in \Gamma_1$ and $\phi \circ \pi_0 = \pi_1 \circ \psi_0$, each $\tilde{\alpha}_i$ is a loop based at $(\tilde{p}, \tilde{x})$. Let $\overline{\alpha}_i : [0,1] \to M_0 \times \mathbb{R}^k$ be the lift of $\tilde{\alpha}_i$ along ψ_0 satisfying $\overline{\alpha}_i(0) = (\overline{p}, \overline{x})$. Since $\pi_0 \circ \overline{\alpha}_i(t) = \overline{x} u_i$, combining (7.1) and (7.2) yields

$$\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(it), \overline{F}_0 \circ \overline{\alpha}_i(1)) \leq D_0$$

Therefore,

$$\left| \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\alpha}_1(1), \overline{F}_0 \circ \overline{\alpha}_0(1)) - \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)) \right| \leq 2D_0 \tag{7.3}$$

It remains to bound $\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\alpha}_0(1), \overline{F}_0 \circ \overline{\alpha}_1(1))$ from below in terms of $\|T_*(t\tilde{v})\|$.

Since $\overline{T}$ is affine, it is Lipschitz continuous, with a Lipschitz constant of $\|\overline{T}\| := \max_{|\overline{w}|=1} \|\overline{T}(\overline{x} + \overline{w})\|$. Set $D_1 := \|\overline{T}\| \cdot \mathrm{diam}(T^k)$. Since $\pi_1 \circ \tilde{\alpha}_i \in \sum_{j=1}^k d_{ij}[s_j]$, where $d_{ij} \in \mathbb{Z}$ are the unique coefficients such that $u_i = \sum_{j=1}^k d_{ij} v_j$, one has that $\tilde{f}_0 \circ \tilde{\alpha}_i \in \sum_{j=1}^k d_{ij}[\sigma_j] = \sum_{j=1}^m d_{ij}[\sigma_j]$. By the definition of d_G,

$$\begin{aligned}
\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\alpha}_1(1), \overline{F}_0 \circ \overline{\alpha}_0(1)) &= \mathrm{d}_G([\tilde{f}_0 \circ \tilde{\alpha}_1], [\tilde{f}_0 \circ \tilde{\alpha}_0]) \\
&= \mathrm{d}_G([e], [\tilde{f}_0 \circ \tilde{\alpha}_1] - [\tilde{f}_0 \circ \tilde{\alpha}_0]) \\
&= \mathrm{d}_G([e], \sum_{j=1}^m (d_{1j} - d_{0j})[\sigma_j]) \\
&= \mathrm{d}_{\mathbb{Z}^m}(0, \sum_{j=1}^m (d_{1j} - d_{0j}) w_j) \\
&= \mathrm{d}_{\mathbb{Z}^m}(0, \overline{T}(\overline{x} + u_1 - u_0)) \\
&\geq \|\overline{T}(\overline{x} + u_1 - u_0)\|
\end{aligned} \tag{7.4}$$

where the equality in the second line is due to the fact that d_G is G-equivariant and the final inequality is from Lemma 7.2.4. Since $\overline{T}$ is affine, $\overline{T}(\overline{x}+u_1-u_0) = \overline{T}(\overline{x}+u_1) - \overline{T}(\overline{x}+u_0)$. Therefore,

$$\Big| \|\overline{T}(\overline{x}+u_1-u_0)\| - \|\overline{T}(\pi_0\circ\overline{\gamma}(t)) - \overline{T}(\pi_0\circ\overline{\gamma}(0))\| \Big| \le \sum_{i=0}^{1} \|\overline{T}(\pi_0\circ\overline{\gamma}(it)) - \overline{T}(\overline{x}+u_i)\|$$

At the same time, $\overline{T}(\overline{x}+\pi_0\circ\overline{\gamma}(t)-\pi_0\circ\overline{\gamma}(0)) = \overline{T}(\pi_0\circ\overline{\gamma}(t)) - \overline{T}(\pi_0\circ\overline{\gamma}(0))$. Inserting this into the left-hand side of the previous inequality and applying (7.2) and the Lipschitz continuity of $\overline{T}$ to the right-hand side yield

$$\Big| \|\overline{T}(\overline{x}+u_1-u_0)\| - \|\overline{T}(\overline{x}+\pi_0\circ\overline{\gamma}(t)-\pi_0\circ\overline{\gamma}(0))\| \Big| \le 2D_1$$

Since the Riemannian metric on T^k is constant and $T : T^k \to T^m$ is affine, $T_*(t\tilde{v}) = \overline{T}(\overline{x}+t\overline{v}) \in \mathrm{T}_{T(\tilde{z})}T^m \cong \mathbb{R}^m$, where $\overline{v} \in \mathbb{R}^k$ is identified with $\tilde{v}$ under the canonical identification of $\mathrm{T}_{\tilde{z}}T^k$ with $\mathbb{R}^k$. The assumption that $\phi\circ\pi_0\circ\overline{\gamma} = \gamma_{\tilde{v}}$ implies that $\pi_0\circ\overline{\gamma}(t) - \pi_0\circ\overline{\gamma}(0) = t\overline{v}$. As $\|\cdot\|$ is a constant Finsler metric on T^m, $\|T_*(t\tilde{v})\| = \|\overline{T}(\overline{x}+t\overline{v})\|$. Combining these with the previous inequality yields

$$\Big| \|\overline{T}(\overline{x}+u_1-u_0)\| - \|T_*(t\tilde{v})\| \Big| \le 2D_1 \tag{7.5}$$

The proof is completed by setting $D := 2D_0 + 2D_1$ and combining (7.3), (7.4), and (7.5).

(b) As in the proof of (a), one has that $\|T_*(\tilde{v})\| = \|\overline{T}(\overline{x}+\overline{v})\|$ for $(\overline{p},\overline{x}) \in \psi_0^{-1}(\tilde{p},\tilde{x})$ and $\overline{v} \in \mathbb{R}^k$. If $\overline{v} = 0$, then $\pi_0\circ\overline{\gamma}$ must be a constant map, and the result follows immediately from Lemma 5.11(b). The remaining arguments will assume that $\overline{v} \neq 0$. It will be helpful to use conditions (A) and (B). Let $V_0 := \mathrm{span}_{\mathbb{R}}\{v_1,\ldots,v_m\}$ and $V_1 := \mathrm{span}_{\mathbb{R}}\{v_{m+1},\ldots,v_k\}$. The result will first be proved when $\overline{v} \in V_0$ or $\overline{v} \in V_1$.

Suppose $\overline{v} \in V_0$. Fix $\varepsilon > 0$. By Theorem 7.2.1, there exists $K \ge 0$ such that

$$\Big| \mathrm{d}_{\mathbb{Z}^m}(0,w) - \|w\| \Big| \le \varepsilon\|w\| \tag{7.6}$$

whenever $w \in \mathbb{Z}^m$ satisfies $\|w\| \ge K$. It follows from (A) that the map $v \mapsto \overline{T}(\overline{x}+v)$ restricts to an isomorphism from V_0 onto $\mathbb{R}^m$ and, consequently, $v \mapsto \|\overline{T}(\overline{x}+v)\|$ defines a norm on V_0. Since $\overline{v} \neq 0$, $\|\overline{T}(\overline{x}+\overline{v})\| \neq 0$. Therefore, there exists $t_0 \ge 0$ such that $\|\overline{T}(\overline{x}+t\overline{v})\| \ge K + 2D_1$ for all $t \ge t_0$. Fix any such t. As in part (a), there exist $u_0, u_1 \in \Gamma_1$ such that, with respect to the corresponding curves $\overline{\alpha}_i$, (7.2)-(7.5) hold. By (7.5), $\|\overline{T}(\overline{x}+u_1-u_0)\| \ge K$. Substituting into (7.6) yields

$$\Big| \mathrm{d}_{\mathbb{Z}^m}(0,\overline{T}(\overline{x}+u_1-u_0)) - \|\overline{T}(\overline{x}+u_1-u_0)\| \Big| \le \varepsilon\|\overline{T}(\overline{x}+u_1-u_0)\| \tag{7.7}$$

As in (7.4), $\mathrm{d}_{\mathbb{Z}^m}(0, \overline{T}(\overline{x}+u_1-u_0)) = \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\alpha}_1(1), \overline{F}_0 \circ \overline{\alpha}_0(1))$. Applying this in conjunction with (7.3), (7.5), and (7.7) yields

$$\left|\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)) - \|\overline{T}(\overline{x}+t\overline{v})\|\right| \leq \varepsilon\|\overline{T}(\overline{x}+t\overline{v})\| + 2D_0 + 2\varepsilon D_1 + 2D_1$$

Therefore,

$$\left|\frac{\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0))}{t} - \|\overline{T}(\overline{x}+\overline{v})\|\right| \leq \varepsilon\|\overline{T}(\overline{x}+\overline{v})\| + \frac{2D_0 + 2\varepsilon D_1 + 2D_1}{t}$$

Since the choice of $\varepsilon > 0$ was arbitrary, it follows that $\|\overline{T}(\overline{x}+\overline{v})\| = \lim_{t\to\infty} \frac{\mathrm{d}_{\overline{N}}\left(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)\right)}{t}$.

Suppose $\overline{v} \in V_1$. Fix $t \geq 0$. As in the previous case, there exist $u_0, u_1 \in \Gamma_1$ such that (7.2)-(7.5) hold with respect to the corresponding curves $\overline{\alpha}_i$. One may, without loss of generality, suppose that $u_1 - u_0 = \sum_{i=1}^{k-m} d_i v_{m+i}$ for some $d_i \in \mathbb{Z}$. It follows from (B) that $\overline{F}_0 \circ \overline{\alpha}_0(1) = \overline{F}_0 \circ \overline{\alpha}_1(1)$. It follows from (7.3) that

$$\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)) \leq 2D_0 \tag{7.8}$$

At the same time, (7.4) and (7.5) imply that

$$\|\overline{T}(\overline{x}+\overline{v})\| \leq \lim_{t\to\infty} \frac{2D_1}{t} = 0$$

This and (7.8) together imply that $\|\overline{T}(\overline{x}+\overline{v})\| = 0 = \lim_{t\to\infty} \frac{\mathrm{d}_{\overline{N}}\left(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)\right)}{t}$.

Finally, suppose that $\overline{v} \in \mathbb{R}^k$ is arbitrary. As $\mathbb{R}^k = V_0 \oplus V_1$, there exist $a_i \in \mathbb{R}$ and $\overline{v}_i \in V_i$ such that $\overline{v} = a_0\overline{v}_0 + a_1\overline{v}_1$. Since $\left|\|\overline{T}(\overline{x}+\overline{v})\| - \|\overline{T}(\overline{x}+a_0\overline{v}_0)\|\right| \leq \|\overline{T}(\overline{x}+a_1\overline{v}_1)\| = 0$, one has that $\|\overline{T}(\overline{x}+\overline{v})\| = \|\overline{T}(\overline{x}+a_0\overline{v}_0)\|$. Let $\overline{\gamma}_0 : [0,\infty) \to M_0 \times \mathbb{R}^k$ be defined by $\overline{\gamma}_0(t) := (\overline{p}, \overline{x}+ta_0\overline{v}_0)$. Applying the first special case to $a_0\overline{v}_0$ and $\overline{\gamma}_0$ shows that

$$\|\overline{T}(\overline{x}+a_0\overline{v}_0)\| = \lim_{t\to\infty} \frac{\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}_0(t), \overline{F}_0 \circ \overline{\gamma}_0(0))}{t} \tag{7.9}$$

Fix $t \geq 0$, and define $\overline{\gamma}_1 : [0,t] \to M_0 \times \mathbb{R}^k$ by $\overline{\gamma}_1(s) := (\overline{p}, \overline{x}+ta_0\overline{v}_0 + sa_1\overline{v}_1)$. By Lemma 5.11(b),

$$\left|\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)) - \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}_0(t), \overline{F}_0 \circ \overline{\gamma}_0(0))\right| \leq \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}_1(t), \overline{F}_0 \circ \overline{\gamma}_1(0)) + 2D_0$$

Applying (7.8) to $a_1\overline{v}_1$ and $\overline{\gamma}_1$ shows that

$$\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}_1(t), \overline{F}_0 \circ \overline{\gamma}_1(0)) \leq 2D_0$$

Therefore,

$$\left|\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}(t), \overline{F}_0 \circ \overline{\gamma}(0)) - \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \overline{\gamma}_0(t), \overline{F}_0 \circ \overline{\gamma}_0(0))\right| \leq 2D_0 + 2D_1 \tag{7.10}$$

The proof is completed by combining the equality $\|T_*(\tilde{v})\| = \|\overline{T}(\overline{x}+a_0\overline{v}_0)\|$ with (7.9) and (7.10).

$\square$

7.4 Main inequalities

This section continues to use the notation and constructions from the previous section. It begins by relating the intersection $I([F])$ to the length density ℓ_T of the affine surjection $T : T^k \to T^m$. Note that the length density $\ell_T : T^k \to [0,\infty)$ is constant. Recall that $l = \dim(M_0)$ and $n = \dim(M) = k+l$, where, by assumption, $n > 0$. This latter assumption ensures that $I([F])$ is well-defined. In any case, were n to be zero, many of the results to come could be given sensible formulations that would hold trivially.

In order to relate $I([F])$ to ℓ_T, define constants

$$d_{k,l} := \begin{cases} 0 & \text{if} \quad k = 0 \\ \frac{c_{k+l}}{c_k}\sqrt{\frac{2c_{k-1}}{k}} & \text{if} \quad k > 0 \end{cases}$$

For the purposes of the following theorem, it turns out that the values of $d_{0,l}$ are irrelevant, because when $k = 0$ both $\ell_T = 0$ and $I([F]) = 0$.

Theorem 7.4.1. *Let M, N, and $[F]$ be as in the previous section, and denote by $(T^m, \|\cdot\|)$ and $T : T^k \to T^m$ the corresponding semi-Finsler torus and affine surjection. Then $I([F]) = d_{k,l}\mathrm{vol}(M)\ell_T$.*

Proof. It will help to first establish some general equalities. By Lemma 7.3.1, the length density ℓ_T is independent of the choice of map $f \in [F]$ used in the construction of $\|\cdot\|$ and T. Thus one may, without loss of generality, suppose that f_0 is C^1. Since M is compact, there exists $C \geq 0$ such that $|(f_0)_*(w)| \leq C$ for all $w \in \mathrm{S}M$. Let $\overline{F}_0 : M_0 \times \mathbb{R}^k \to \overline{N}$ be a lift of $\overline{f}_0$ of the form guaranteed by Lemma 5.11. For each $w \in \mathrm{S}M$, one has, directly from the definition of ϕ_t, that

$$\phi_t(w) = \mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \gamma_{\overline{w}}(t), \overline{F}_0 \circ \gamma_{\overline{w}}(0)) \tag{7.11}$$

for any $\overline{w} \in \mathrm{S}(M_0 \times \mathbb{R}^k)$ such that $(\psi_1 \circ \psi_0)_*(\overline{w}) = w$. It follows that $\frac{\phi_t(w)}{t} \leq C$ for all $t > 0$ and $w \in \mathrm{S}M$. Thus

$$\begin{aligned} I([F]) &= \int_{\mathrm{S}M} \lim_{t\to\infty} \frac{\phi_t(w)}{t}\,\mathrm{d}\mu_{\mathrm{S}M} \\ &= \int_M \int_{\mathrm{S}_q M} \lim_{t\to\infty} \frac{\phi_t(w)}{t}\,\mathrm{d}\mu_{\mathrm{S}_q M}\,\mathrm{d}\mu_M \end{aligned} \tag{7.12}$$

where the first equality follows from the bounded convergence theorem and the second from the fact that, up to the isometric identification of each S_qM with S^{n-1}, $\mu_{\mathrm{S}M}$ is locally the product measure $\mu_{\mathrm{S}M} = \mu_M \times \mu_{S^{n-1}}$. For each $\overline{w} \in \mathrm{T}(M_0 \times \mathbb{R}^k)$, Lemma 7.3.2(b) implies that

$$\|(T \circ \phi \circ \pi_0)_*(\overline{w})\| = \lim_{t\to\infty} \frac{\mathrm{d}_{\overline{N}}(\overline{F}_0 \circ \gamma_{\overline{w}}(t), \overline{F}_0 \circ \gamma_{\overline{w}}(0))}{t}$$

Combining this with (7.11) and the fact that $\psi_1 \circ \psi_0$ is a local isometry yields

$$\int_{\mathrm{S}_qM} \lim_{t\to\infty} \frac{\phi_t(w)}{t} \,\mathrm{d}\mu_{\mathrm{S}_qM} = \int_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} \|(T \circ \phi \circ \pi_0)_*(\overline{w})\| \,\mathrm{d}\mu_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} \tag{7.13}$$

for each $q \in M$ and $(\overline{q},\overline{z}) \in M_0 \times \mathbb{R}^k$ such that $\psi_1 \circ \psi_0(\overline{q},\overline{z}) = q$.

The argument now splits into three cases. If $k = 0$, then the result holds trivially, as $\ell_T = 0$, $\|\cdot\| = 0$, and, by (7.12) and (7.13), $I([F]) = 0$. Suppose for the remainder of the proof that $k > 0$. Since ϕ is a local isometry,

$$\int_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} \|(T \circ \phi)_*(\overline{v})\| \,\mathrm{d}\mu_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} = \int_{\mathrm{S}_{\phi(\overline{z})}T^k} \|T_*(\tilde{v})\| \,\mathrm{d}\mu_{\mathrm{S}_{\phi(\overline{z})}T^k} = \sqrt{\frac{2c_{k-1}}{k}}\,\ell_T \tag{7.14}$$

for each $\overline{z} \in \mathbb{R}^k$. If $l = 0$, then $M_0 \times T^k \cong T^k$ and

$$\int_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} \|(T \circ \phi \circ \pi_0)_*(\overline{w})\| \,\mathrm{d}\mu_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} = \int_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} \|(T \circ \phi)_*(\overline{w})\| \,\mathrm{d}\mu_{\mathrm{S}_{\overline{z}}\mathbb{R}^k}$$

for the only $\overline{q} \in M_0$ and each $\overline{z} \in \mathbb{R}^k$. Since $d_{k,l} = \sqrt{\frac{2c_{k-1}}{k}}$, the result follows from (7.12)-(7.14). Suppose $l > 0$. Fix $(\overline{q},\overline{z}) \in M_0 \times \mathbb{R}^k$. Since $k > 0$, $\mathrm{S}_{(\overline{q},\overline{z})}(M_0 \times \mathbb{R}^k) \setminus \ker(\pi_0)_* \neq \emptyset$. Define a smooth surjection $H : \mathrm{S}_{(\overline{q},\overline{z})}(M_0 \times \mathbb{R}^k) \setminus \ker(\pi_0)_* \to \mathrm{S}_{\overline{z}}\mathbb{R}^k$ by $H(\overline{w}) := \frac{(\pi_0)_*(\overline{w})}{|(\pi_0)_*(\overline{w})|}$. Applying the coarea formula to H yields

$$\begin{aligned}
\int_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} \|(T \circ \phi \circ \pi_0)_*(\overline{w})\| \,\mathrm{d}\mu_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} &= \int_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} \int_{H^{-1}(\overline{v})} |(\pi_0)_*(\overline{w})|^{k-1} \|(T \circ \phi \circ \pi_0)_*(\overline{w})\| \,\mathrm{d}\mu_{H^{-1}(\overline{v})} \,\mathrm{d}\mu_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} \\
&= \int_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} \|(T \circ \phi)_*(\overline{v})\| \int_{H^{-1}(\overline{v})} |(\pi_0)_*(\overline{w})|^k \,\mathrm{d}\mu_{H^{-1}(\overline{v})} \,\mathrm{d}\mu_{\mathrm{S}_{\overline{z}}\mathbb{R}^k}
\end{aligned}$$

For each fixed $\overline{v} \in \mathrm{S}_{\overline{z}}\mathbb{R}^k$, another application of the coarea formula shows that

$$\begin{aligned}
\int_{H^{-1}(\overline{v})} |(\pi_0)_*(\overline{w})|^k \,\mathrm{d}\mu_{H^{-1}(\overline{v})} &= \int_0^1 \int_{\mathrm{S}_{\overline{q}}(\sqrt{1-r^2})} \frac{r^k}{\sqrt{1-r^2}} \,\mathrm{d}\mu_{\mathrm{S}_{\overline{q}}(\sqrt{1-r^2})} \,\mathrm{d}r \\
&= c_{l-1} \int_0^1 r^k (1-r^2)^{\frac{l}{2}-1} \,\mathrm{d}r \\
&= \frac{1}{2} c_{l-1} \int_0^1 t^{\frac{k-1}{2}} (1-t)^{\frac{l}{2}-1} \,\mathrm{d}t \\
&= \frac{1}{2} c_{l-1} B\Big(\frac{k+1}{2}, \frac{l}{2}\Big)
\end{aligned}$$

where the sets $S_{\overline{q}}(\sqrt{1-r^2})$ for $0 \leq r < 1$ are non-empty since $l > 0$. Combining the last two results yields

$$\int_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} \|(T\circ\phi\circ\pi_0)_*(\overline{w})\| \,\mathrm{d}\mu_{\mathrm{S}_{(\overline{q},\overline{z})}(M_0\times\mathbb{R}^k)} = \frac{1}{2}c_{l-1}B\Big(\frac{k+1}{2},\frac{l}{2}\Big)\int_{\mathrm{S}_{\overline{z}}\mathbb{R}^k} \|(T\circ\phi)_*(\overline{v})\| \,\mathrm{d}\mu_{\mathrm{S}_{\overline{z}}\mathbb{R}^k}$$

Combining this with (7.12)-(7.14) shows that

$$\begin{aligned} I([F]) &= c_{l-1}\sqrt{\frac{c_{k-1}}{2k}}B\Big(\frac{k+1}{2},\frac{l}{2}\Big)\mathrm{vol}(M)\ell_T \\ &= d_{k,l}\mathrm{vol}(M)\ell_T \end{aligned}$$

where the latter equality follows from Lemma 2.7.1(c).

□

Two special cases of Theorem 7.4.1 are worth recording explicitly.

Remark 7.4.2. When $k > 0$, the conclusion of Theorem 7.4.1 takes the form

$$I([F]) = \frac{c_n}{c_k}\mathrm{vol}(M)\int_{S_{\tilde{z}}T^k} \|T_*(\tilde{v})\| \,\mathrm{d}\mu_{S_{\tilde{z}}T^k}$$

for any $\tilde{z} \in T^k$.

Remark 7.4.3. In the case that $l = 0$, or equivalently that M_1 is a point, Theorem 7.4.1 implies that

$$I([F]) = \sqrt{\frac{2c_{n-1}}{n}}\frac{\mathrm{L}(T)}{\#(\psi_1)}$$

In particular, if $M = T^k$, then

$$I([F]) = \sqrt{\frac{2c_{n-1}}{n}}\mathrm{L}(T) = \mathrm{vol}(T^k)\int_{S_{\tilde{z}}T^k} \|T_*(\tilde{v})\| \,\mathrm{d}\mu_{S_{\tilde{z}}T^k}$$

for any $\tilde{z} \in T^k$.

It's interesting to note that the sequence $d_{k,l}$ is not, in general, monotone in l. For instance, $d_{1,l} = \frac{c_{l+1}}{\pi}$ is strictly increasing until its maximum at $l = 5$ and strictly decreasing thereafter [BH]. This means that the intersection of a homotopically non-trivial map $T : S^1 \to S^1$ will, for small values of l, increase when it's extended to a map $T^{l+1} = T^l \times S^1 \to S^1$ that's constant along the T^l-fibers.

It's now possible to prove this chapter's main theorem. Before stating the result, it will help to dispose of the situation where $k = 0$. In that case, $I([F]) = e_T = \ell_T = 0$ and $\mathrm{E}(f) \geq 0$, with equality if and only if f is constant. This is the trivial case of the following phenomenon.

Theorem 7.4.4. *Let M, N, and $[F]$ be as in the previous section, and denote by $(T^m, \|\cdot\|)$ and $T : T^k \to T^m$ the corresponding semi-Finsler torus and affine surjection. Suppose that $k > 0$. Let $f \in [F]$ be any* C^1 *map. Then*

$$\mathrm{E}(f) \geq \mathrm{vol}(M)e_T \geq \frac{1}{c_{k-1}}\mathrm{vol}(M)\ell_T^2 \geq \frac{nc_n^2c_{k-1}}{kc_{n-1}^2c_k^2}\mathrm{vol}(M)\ell_T^2$$

Moreover, each of the following holds:

(a) *If* $\mathrm{E}(f) = \mathrm{vol}(M)e_T$, *then f is totally geodesic;*

(b) $\mathrm{E}(f) = \frac{1}{c_{k-1}}\mathrm{vol}(M)\ell_T^2$ *if and only if $\overline{f}$ is constant along each M_0-fiber and a homothety along each $\mathbb{R}^k$-fiber;*

(c) $\mathrm{E}(f) = \frac{nc_n^2c_{k-1}}{kc_{n-1}^2c_k^2}\mathrm{vol}(M)\ell_T^2$ *if and only if f is a homothety and either M_1 is a point or f is constant.*

Proof. By Lemma 7.3.1, the quantities ℓ_T and e_T are independent of the choice of representative $f_0 \in [F]$ used in the construction of $\|\cdot\|$ and T, so one may, without loss of generality, suppose that $f_0 = f$. The inequality $\mathrm{E}(f) \geq \mathrm{vol}(M)e_T$ will be proved first. Note that $\mathrm{E}(\tilde{f}) = \#(\psi_1)\mathrm{E}(f)$. By Lemma 5.3, $\chi = \chi_{\overline{x}} := \rho_1 \circ \psi_0 \circ \iota_{\overline{x}} : M_0 \to M_1$ is a covering map; denote its number of sheets by $\#(\chi) < \infty$. Fix $\overline{p} \in \chi^{-1}(\tilde{p})$. Since $\phi \circ \pi_0 = \pi_1 \circ \psi_0$, $\psi_0(\overline{p}, \overline{x}) = (\tilde{p}, \tilde{x})$, and consequently $\overline{f}(\overline{p}, \overline{x}) = y$. Fix $\overline{y} \in \overline{N}$, and let $\overline{F} : M_0 \times \mathbb{R}^k \to \overline{N}$ be the lift of $\overline{f}$ guaranteed by Lemma 5.11 that satisfies $\overline{F}(\overline{p}, \overline{x}) = \overline{y}$. For each $r \in \mathbb{N}$, define a parallelotope $P_r := \{\sum_{i=1}^k t_i v_i \,\big|\, 0 \leq t_i \leq r\}$. Then P_r differs from a union of r^k fundamental domains of ϕ by a set of measure zero. It follows from Lemma 5.9 that $M_0 \times P_r$ differs from a union of $\#(\chi)r^k$ fundamental domains of ψ_0 by a set of measure zero. Thus

$$\mathrm{E}(f) = \frac{1}{\#(\psi_1)\#(\chi)r^k}\mathrm{E}(\overline{F}|_{M_0 \times P_r}) \tag{7.15}$$

Let $\{e_1, \ldots, e_k\}$ be an orthonormal set of vector fields on $M_0 \times \mathbb{R}^k$ that everywhere spans $\mathrm{T}\mathbb{R}^k \subset \mathrm{T}(M_0 \times \mathbb{R}^k)$. If $U \subseteq M_0$ is a connected open set and $\{e_{k+1}, \ldots, e_{k+l}\}$ is an orthonormal set of vector fields on $U \times \mathbb{R}^k$ that at each point spans $\mathrm{T}U \subset \mathrm{T}(U \times \mathbb{R}^k)$, so that $\{e_1, \ldots, e_{k+l}\}$ is an orthonormal frame for $U \times \mathbb{R}^k$, then on $U \times \mathbb{R}^k$ the energy density $e_{\overline{F}}$ satisfies

$$\begin{aligned} e_{\overline{F}} &= \sum_{i=1}^{k+l} \|\overline{F}_*(e_i)\|^2 \\ &\geq \sum_{i=1}^{k} \|\overline{F}_*(e_i)\|^2 \\ &= e_{\overline{F} \circ \iota_{\overline{q}}} \circ \pi_0 \end{aligned}$$

with equality on all of $U \times \mathbb{R}^k$ if and only if $\overline{F}$ is constant on each U-fiber. Therefore,

$$\begin{aligned} E(\overline{F}|_{M_0\times P_r}) &= \int_{M_0\times P_r} e_{\overline{F}}\, \mathrm{d}\mu_{M_0\times\mathbb{R}^k} \\ &\geq \int_{M_0\times P_r} e_{\overline{F}\circ\iota_{\overline{q}}}\circ\pi_0\, \mathrm{d}\mu_{M_0\times\mathbb{R}^k} \\ &= \int_{M_0}\int_{P_r} e_{\overline{F}\circ\iota_{\overline{q}}}\circ\pi_0\, \mathrm{d}\mu_{\mathbb{R}^k}\, \mathrm{d}\mu_{M_0} \\ &= \int_{M_0} \mathrm{E}(\overline{F}\circ\iota_{\overline{q}}|_{P_r})\, \mathrm{d}\mu_{M_0} \end{aligned} \tag{7.16}$$

with equality if and only if $\overline{F}$ is constant along each M_0-fiber. For each $\overline{q}\in M_0$ and $w\in \mathrm{S}^+\partial P_r$, let $\varsigma_{\overline{q},w} : [0,\ell(w)] \to N$ be defined by $\varsigma_{\overline{q},w} := \overline{F}\circ\iota_{\overline{q}}\circ\gamma_w$. Applying Corollary 7.1.3 to $\mathrm{E}(\overline{F}\circ\iota_{\overline{q}}|_{P_r})$ shows that

$$\mathrm{E}(\overline{F}\circ\iota_{\overline{q}}|_{P_r}) = \frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r} \mathrm{E}(\varsigma_{\overline{q},w})h(w,\nu)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \tag{7.17}$$

Combining (7.15)-(7.17) yields

$$\mathrm{E}(f) \geq \frac{1}{\#(\psi_0)\#(\chi)r^k}\int_{M_0}\frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r} \mathrm{E}(\varsigma_{\overline{q},w})h(w,\nu)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r}\, \mathrm{d}\mu_{M_0} \tag{7.18}$$

with equality if and only if $\overline{F}$ is constant along each M_0-fiber. At the same time, the Cauchy-Schwarz inequality states that

$$\mathrm{E}(\varsigma_{\overline{q},w}) \geq \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} \tag{7.19}$$

with equality if and only if $\varsigma_{\overline{q},w}$ has constant speed. Note that

$$\mathrm{L}(\varsigma_{\overline{q},w}) \geq \mathrm{d}_{\overline{N}}\Big(\varsigma_{\overline{q},w}(\ell(w)), \varsigma_{\overline{q},w}(0)\Big)$$

By Theorem 7.3.2(a), there exists $D \geq 0$, independent of $\overline{q}$ and w, such that

$$\mathrm{d}_{\overline{N}}\Big(\varsigma_{\overline{q},w}(\ell(w)), \varsigma_{\overline{q},w}(0)\Big) \geq \ell(w)\|\overline{T}_*(w)\| - D \tag{7.20}$$

Regardless of the sign of $\ell(w)\|\overline{T}_*(w)\| - D$, one may conclude that

$$\frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} \geq \ell(w)\|\overline{T}_*(w)\|^2 - 2D\|\overline{T}_*(w)\|$$

Combining results yields

$$\begin{aligned} \mathrm{E}(f) &\geq \frac{1}{\#(\psi_0)\#(\chi)r^k}\int_{M_0}\frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r}[\ell(w)\|\overline{T}_*(w)\|^2 - 2D\|\overline{T}_*(w)\|]h(w,\nu)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r}\, \mathrm{d}\mu_{M_0} \\ &\geq \frac{\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)r^k}\frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r}\ell(w)\|\overline{T}_*(w)\|^2 h(w,\nu)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} - \frac{D\|\overline{T}\|\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)}\frac{k\mathrm{vol}(\mathrm{S}^+\partial P_r)}{c_{k-1}r^k} \end{aligned} \tag{7.21}$$

The two summands will be handled separately. First note that

$$\operatorname{vol}(M) = \frac{\operatorname{vol}(M_0 \times P_r)}{\#(\psi_0)\#(\chi)r^k} = \frac{\operatorname{vol}(M_0)\operatorname{vol}(P_r)}{\#(\psi_0)\#(\chi)r^k} \tag{7.22}$$

Since T is affine, $\mathrm{E}(\overline{T}|_{P_r}) = \operatorname{vol}(P_r)e_T$. At the same time, applying Corollary 7.1.3 to $\mathrm{E}(\overline{T}|_{P_r})$ yields

$$\begin{aligned}\mathrm{E}(\overline{T}|_{P_r}) &= \frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r} \mathrm{E}(\overline{T}\circ\gamma_w|_{[0,\ell(w)]})h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_r}\\ &= \frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r}\Big[\int_0^{\ell(w)}\|\overline{T}_*(w)\|^2\,\mathrm{d}t\Big]h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_r}\\ &= \frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r}\ell(w)\|\overline{T}_*(w)\|^2 h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_r}\end{aligned}$$

It follows that

$$e_T = \frac{k}{2c_{k-1}\operatorname{vol}(P_r)}\int_{\mathrm{S}^+\partial P_r}\ell(w)\|\overline{T}_*(w)\|^2 h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \tag{7.23}$$

Taken together, (7.22) and (7.23) imply that

$$\operatorname{vol}(M)e_T = \frac{\operatorname{vol}(M_0)}{\#(\psi_0)\#(\chi)r^k}\frac{k}{2c_{k-1}}\int_{\mathrm{S}^+\partial P_r}\ell(w)\|\overline{T}_*(w)\|^2 h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \tag{7.24}$$

For each $1 \le i \le k$, let Λ_i denote the parallelotope in $\mathbb{R}^k$ determined by $\{v_1,\dots,v_k\}\setminus\{v_i\}$, and let λ_i equal the $(k-1)$-dimensional volume of Λ_i. Since ∂P_r consists, up to translations, of two copies of each of $r\Lambda_i := \{rv \,\big|\, v \in \Lambda_i\}$, one has that

$$\begin{aligned}\operatorname{vol}(\mathrm{S}^+\partial P_r) &= \frac{c_{k-1}}{2}\operatorname{vol}(\partial P_r)\\ &= c_{k-1}\sum_{i=1}^{k}\operatorname{vol}(r\Lambda_i)\\ &= c_{k-1}r^{k-1}\sum_{i=1}^{k}\lambda_i\end{aligned}$$

It follows that

$$\frac{D\|\overline{T}\|\operatorname{vol}(M_0)}{\#(\psi_0)\#(\chi)}\frac{k\operatorname{vol}(\mathrm{S}^+\partial P_r)}{c_{k-1}r^k} = \frac{D\|\overline{T}\|\operatorname{vol}(M_0)}{\#(\psi_0)\#(\chi)}\frac{k\sum_{i=1}^k\lambda_i}{r} \to 0 \tag{7.25}$$

as $r \to \infty$. The inequality $\mathrm{E}(f) \ge \operatorname{vol}(M)e_T$ follows from (7.21), (7.24), and (7.25) by letting $r \to \infty$.

The remaining inequalities are much simpler. The Cauchy-Schwarz inequality implies that $\operatorname{vol}(M)e_T \ge \frac{1}{c_{k-1}}\operatorname{vol}(M)\ell_T^2$. The inequality $\frac{1}{c_{k-1}}\operatorname{vol}(M)\ell_T^2 \ge \frac{nc_n^2c_{k-1}}{kc_{n-1}^2c_k^2}\operatorname{vol}(M)\ell_T^2$ follows from Corollary 2.7.5(a).

(a) Suppose that $\mathrm{E}(f) = \operatorname{vol}(M)e_T$. Fix $r \in \mathbb{N}$. For each $R = (r_1,\dots,r_k) \in \{0,\dots,r-1\}^k$, let $P_R := P_1 + \sum_{i=1}^k r_iv_i$. Note that any two parallelotopes in $\{P_R \,\big|\, R \in \{0,\dots,r-1\}^k\}$, if they intersect at all, do so only along boundary faces, and that P_r is the union of all such P_R. By Corollary 7.1.3, one has that

$$\begin{aligned}\int_{\mathrm{S}^+\partial P_r} \mathrm{E}(\varsigma_{\overline{q},w}) h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} &= \mathrm{E}(\overline{F} \circ \iota_{\overline{q}}|_{P_r}) \\ &= \sum_{R\in\{0,\ldots,r-1\}^k} \mathrm{E}(\overline{F} \circ \iota_{\overline{q}}|_{P_R})\end{aligned}$$

for each $\overline{q} \in M_0$. Since $\phi \circ \pi_0 = \pi_1 \circ \psi_0$, there exist isometric deck transformations $\gamma_R : M_0 \times \mathbb{R}^k \to M_0 \times \mathbb{R}^k$ of ψ_0 such that $\gamma_R(M_0 \times P_1) = M_0 \times P_R$. This implies that

$$\int_{M_0} \mathrm{E}(\varsigma_{\overline{q},w}|_{M_0\times P_1})\, \mathrm{d}\mu_{M_0} = \int_{M_0} \mathrm{E}(\varsigma_{\overline{q},w}|_{M_0\times P_R})\, \mathrm{d}\mu_{M_0}$$

Combining results, one has that

$$\int_{M_0}\int_{\mathrm{S}^+\partial P_r} \mathrm{E}(\varsigma_{\overline{q},w}) h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} = r^k \int_{M_0}\int_{\mathrm{S}^+\partial P_1} \mathrm{E}(\varsigma_{\overline{q},w}) h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_1}\, \mathrm{d}\mu_{M_0} \tag{7.26}$$

for each $r \in \mathbb{N}$. Taken together, (7.15)-(7.26) show that

$$\begin{aligned}\mathrm{vol}(M) e_T &= \mathrm{E}(f) \\ &= \frac{1}{\#(\psi_1)\#(\chi)} \mathrm{E}(\overline{F}|_{M_0\times P_1}) \\ &\geq \frac{1}{\#(\psi_0)\#(\chi)} \int_{M_0} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_1} \mathrm{E}(\varsigma_{\overline{q},w}) h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_1}\, \mathrm{d}\mu_{M_0} \\ &= \liminf_{r\to\infty} \frac{1}{\#(\psi_0)\#(\chi) r^k} \int_{M_0} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_r} \mathrm{E}(\varsigma_{\overline{q},w}) h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r}\, \mathrm{d}\mu_{M_0} \\ &\geq \mathrm{vol}(M) e_T\end{aligned}$$

It follows that the above inequalities are all equalities. By the condition for equality in (7.18), $\overline{F}$ must be constant along each M_0-fiber.

It will next be shown that each $\varsigma_{\overline{q},w}$ has constant speed. This will be done using the condition for equality in (7.19), in an argument similar to the one just employed, but which requires another inequality be inserted into those above. It will be shown that

$$\int_{\mathrm{S}^+\partial P_r} \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \geq \frac{1}{2^k} \int_{\mathrm{S}^+\partial P_{2r}} \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} h(w,v)\, \mathrm{d}\mu_{\mathrm{S}^+\partial P_{2r}} \tag{7.27}$$

for each $\overline{q} \in M_0$ and $r \in \mathbb{N}$. Fix such $\overline{q}$ and r. For each $S = (s_1,\ldots,s_k) \in \{0,1\}^k$, let $Q_S := P_r + r\sum_{i=0}^k s_i v_i$. At the risk of overloading notation, let $\chi_S : \mathrm{S}\mathbb{R}^k \to \{0,1\}$ denote the indicator function of

$\mathrm{S}Q_S$. That is,

$$\chi_S(w) := \begin{cases} 1 & \text{if} \quad w \in \mathrm{S}Q_S \\ 0 & \text{if} \quad w \notin \mathrm{S}Q_S \end{cases}$$

Define a continuous function $\ell_S : \mathrm{S}Q_S \to [0,\infty)$ by setting $\ell_S(w)$ equal to the length of the line segment $Q_S \cap \gamma_w(\mathbb{R})$, and let $\mathrm{L}_S : \mathrm{S}Q_S \to [0,\infty)$ be defined by $\mathrm{L}_S(w) := \int_{-\infty}^{\infty} (\chi_S \circ \gamma'_w) \|(\overline{F} \circ \iota_{\overline{q}} \circ \gamma_w)'\| \, \mathrm{d}t$. In other words, $\mathrm{L}_S(w)$ is the length of the curve $\overline{F} \circ \iota_{\overline{q}}|_{Q_S \cap \gamma_w(\mathbb{R})}$. Note that $Q_0 = P_r$, where $0 = (0,\ldots,0) \in \{0,1\}^k$, and that $\ell_0 \circ \gamma'_w(t) = \ell(w)$ and $\mathrm{L}_0 \circ \gamma'_w(t) = \mathrm{L}(\varsigma_{\overline{q},w})$ for all $w \in \mathrm{S}^+\partial P_r$ and $t \in [0,\ell(w)]$. Since $\overline{F}$ is constant along each M_0-fiber, $\overline{F}$ is $\mathbb{Z}^k$-equivariant under the action of $\mathbb{Z}^k$ on $M_0 \times \mathbb{R}^k$ that translates the second component by elements of Γ_0. It follows from this symmetry that

$$\int_{\mathrm{S}P_r} \frac{\mathrm{L}_0^2(w)}{\ell_0^2(w)} \, \mathrm{d}\mu_{\mathrm{S}P_r} = \int_{\mathrm{S}P_{2r}} \chi_S(w) \frac{\mathrm{L}_S^2(w)}{\ell_S^2(w)} \, \mathrm{d}\mu_{\mathrm{S}P_{2r}}$$

for each $S \in \{0,1\}^k$, where the right-hand integrand is well-defined away from a set of measure zero under the convention that it equals zero whenever χ_S does. Thus

$$\begin{aligned} 2^k \int_{\mathrm{S}^+\partial P_r} \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} h(w,v) \, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} &= 2^k \int_{\mathrm{S}^+\partial P_r} \Big[\int_0^{\ell(w)} \frac{\mathrm{L}_0^2 \circ \gamma'_w}{\ell_0^2 \circ \gamma'_w} \, \mathrm{d}t \Big] h(w,v) \, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \\ &= 2^k \int_{\mathrm{S}P_r} \frac{\mathrm{L}_0^2(w)}{\ell_0^2(w)} \, \mathrm{d}\mu_{\mathrm{S}P_r} \\ &= \int_{\mathrm{S}P_{2r}} \Big[\sum_{S \in \{0,1\}^k} \chi_S(w) \frac{\mathrm{L}_S^2(w)}{\ell_S^2(w)} \Big] \mathrm{d}\mu_{\mathrm{S}P_{2r}} \\ &= \int_{\mathrm{S}^+\partial P_{2r}} \Big[\int_0^{\ell(w)} \sum_{S \in \{0,1\}^k} (\chi_S \circ \gamma'_w) \frac{\mathrm{L}_S^2 \circ \gamma'_w}{\ell_S^2 \circ \gamma'_w} \, \mathrm{d}t \Big] h(w,v) \, \mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \\ &= \int_{\mathrm{S}^+\partial P_{2r}} \Big[\sum_{\ell_S(w) \neq 0} \frac{\mathrm{L}_S^2(w)}{\ell_S(w)} \Big] h(w,v) \, \mathrm{d}\mu_{\mathrm{S}^+\partial P_{2r}} \end{aligned} \tag{7.28}$$

where the second and fourth equalities are by Santaló's formula and the last is obtained by dividing $[0,\ell(w)]$ into subintervals $I_S := \{t \in [0,\ell(w)] \,\big|\, \gamma_w(t) \in Q_S\}$, each of length $\ell_S(w) \neq 0$. Since $\sum_{\ell_S(w) \neq 0} \ell_S(w) = \ell(w)$ and $\mathrm{L}(w) = \sum_{\ell_S(w) \neq 0} \mathrm{L}_S(w)$ for each $w \in \mathrm{S}^+\partial P_{2r}$, Lemma 2.3.5 implies that

$$\sum_{\ell_S(w) \neq 0} \frac{\mathrm{L}_S^2(w)}{\ell_S(w)} \geq \frac{\mathrm{L}^2(w)}{\ell(w)} \tag{7.29}$$

The inequality (7.27) follows immediately from (7.28) and (7.29). Combining (7.27) with (7.15)-(7.25), and taking into account the fact that $\overline{F}$ is constant along each M_0-fiber, one has for each

$\overline{q} \in M_0$ that

$$\begin{aligned}
\mathrm{vol}(M)e_T &= \mathrm{E}(f) \\
&= \frac{\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_1} \mathrm{E}(\varsigma_{\overline{q},w}) h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_1}\,\mathrm{d}\mu_{M_0} \\
&\geq \frac{\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_1} \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_1}\,\mathrm{d}\mu_{M_0} \\
&\geq \liminf_{r\to\infty} \frac{\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)2^{kr}} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_{2^r}} \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_{2^r}}\,\mathrm{d}\mu_{M_0} \\
&\geq \mathrm{vol}(M)e_T
\end{aligned} \tag{7.30}$$

All of these inequalities must therefore be equalities, and by the condition for equality in (7.19) each $\varsigma_{\overline{q},w}$ must have constant speed.

It will next be shown that each $\varsigma_{\overline{q},w}$ has speed $\|\overline{T}_*(w)\|$ and then, as a consequence, that each is a geodesic. It follows from (7.20) that $\|\varsigma'_{\overline{q},w}\| \geq \|\overline{T}_*(w)\|$ for each $r \geq 1$ and $w \in \mathrm{S}^+\partial P_r$, since, if one had $\|\varsigma'_{\overline{q},w}\| = \|\overline{T}_*(w)\| - \varepsilon$ for some $\varepsilon > 0$, then by increasing r and applying a translation in Γ_0 one could without loss of generality suppose that $\ell(w) > \frac{D}{\varepsilon}$. This would imply that

$$\begin{aligned}
\mathrm{d}_{\overline{N}}\Big(\varsigma_{\overline{q},w}(\ell(w)), \varsigma_{\overline{q},w}(0)\Big) &\leq \mathrm{L}(\varsigma_{\overline{q},w}) \\
&= \ell(w)\|\varsigma'_{\overline{q},w}\| \\
&= \ell(w)\|\overline{T}_*(w)\| - \ell(w)\varepsilon \\
&< \ell(w)\|\overline{T}_*(w)\| - D
\end{aligned}$$

The above contradicts (7.20). On the other hand, assume that $\|\varsigma'_{\overline{q},w}\| > \|\overline{T}_*(w)\|$ for some $r \geq 1$ and $w \in \mathrm{S}^+\partial P_r$. Then

$$\begin{aligned}
\mathrm{vol}(M)e_T &= \frac{\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_1} \frac{\mathrm{L}^2(\varsigma_{\overline{q},w})}{\ell(w)} h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_1}\,\mathrm{d}\mu_{M_0} \\
&> \frac{\mathrm{vol}(M_0)}{\#(\psi_0)\#(\chi)r^k} \frac{k}{2c_{k-1}} \int_{\mathrm{S}^+\partial P_r} \ell(w)\|\overline{T}_*(w)\|^2 h(w,v)\,\mathrm{d}\mu_{\mathrm{S}^+\partial P_r} \\
&= \mathrm{vol}(M)e_T
\end{aligned}$$

where the final equality is (7.24). This is a contradiction. Thus each $\varsigma_{\overline{q},w}$ has speed $\|\overline{T}_*(w)\|$. To see that each $\varsigma_{\overline{q},w}$ is a geodesic, it suffices by the density of rational vectors to prove it for all rational $w \in \mathrm{S}^+\partial P_r$, that is, those w such that, for some $s > 0$, $\gamma_w(s) - \gamma_w(0) \in \Gamma_0$. Assume that $\varsigma_{\overline{q},w}$ is not a geodesic for some such w. As before, one may without loss of generality suppose that $\ell(w) > s$.

Since $\overline{F}$ is constant along each M_0-fiber, $\pi \circ \varsigma_{\overline{q},w}$ is a loop in N. Following the lines of (7.4), in particular invoking Lemma 7.2.4, one has

$$\begin{aligned} s\|\overline{T}_*(w)\| &= \mathrm{L}(\varsigma_{\overline{q},w}|_{[0,s]}) \\ &> \mathrm{d}(\varsigma_{\overline{q},w}(s), \varsigma_{\overline{q},w}(0)) \\ &\geq s\|\overline{T}_*(w)\| \end{aligned}$$

This is a contradiction. It follows that $\overline{F} \circ \iota_{\overline{q}}$ is totally geodesic for each $\overline{q} \in M_0$. An application of Lemma 5.10(a) now shows that $\overline{F}$ is totally geodesic. Consequently, so is f.

(b) If $\mathrm{E}(f) = \frac{1}{c_{k-1}}\mathrm{vol}(M)\ell_T^2$, then, by the main inequalities, one also has that $\mathrm{E}(f) = \mathrm{vol}(M)e_T$. As shown in the proof of (a), $\overline{F}$ must be constant along each M_0-fiber, and therefore $\overline{f}$ is as well. By the condition for equality in (7.16), one has for each $\overline{q} \in M_0$ that

$$\begin{aligned} \mathrm{E}(\tilde{f}) &= \frac{1}{\#(\chi)}\mathrm{E}(\overline{F}|_{M_0 \times P_1}) \\ &= \frac{\mathrm{vol}(M_0)}{\#(\chi)}\mathrm{E}(\overline{F} \circ \iota_{\overline{q}}|_{P_1}) \\ &= \frac{\mathrm{vol}(M_0)}{\#(\chi)}\mathrm{E}(\tilde{f} \circ \iota_{\chi(\overline{q})}) \end{aligned}$$

where the domain of the map $\tilde{f} \circ \iota_{\chi(\overline{q})}$ is T^k with its given flat metric. At the same time, by assumption,

$$\frac{1}{c_{k-1}}\mathrm{vol}(M)\ell_T^2 = \mathrm{E}(f) = \frac{1}{\#(\psi_1)}\mathrm{E}(\tilde{f})$$

Combining results shows that $\mathrm{E}(\tilde{f} \circ \iota_{\chi(\overline{q})}) = \frac{\#(\psi_1)\#(\chi)}{c_{k-1}} \frac{\mathrm{vol}(M)}{\mathrm{vol}(M_0)}\ell_T^2$. Note that $\mathrm{vol}(M_1 \times T^k) = \#(\psi_1)\mathrm{vol}(M)$ and, at the same time, $\mathrm{vol}(M_1 \times T^k) = \frac{1}{\#(\chi)}\mathrm{vol}(M_0 \times P_1) = \frac{1}{\#(\chi)}\mathrm{vol}(M_0)\mathrm{vol}(T^k)$. It follows that $\mathrm{vol}(T^k) = \#(\psi_1)\#(\chi)\frac{\mathrm{vol}(M)}{\mathrm{vol}(M_0)}$. So $\mathrm{E}(\tilde{f} \circ \iota_{\chi(\overline{q})}) = \frac{\mathrm{vol}(T^k)}{c_{k-1}}\ell_T^2$. By Theorem 7.4.1, the intersection of the homotopy class $[\tilde{f} \circ \iota_{\chi(\overline{q})}]$ is $I([\tilde{f} \circ \iota_{\chi(\overline{q})}]) = \frac{2c_{k-1}}{k}\mathrm{vol}(T^k)\ell_T^2$. An application of Theorem 7.1.4(b) now shows that $\tilde{f} \circ \iota_{\chi(\overline{q})}$ is a homothety, which is equivalent to $\overline{f} \circ \iota_{\overline{q}}$ being a homothety.

On the other hand, if $\overline{f}$ is constant along each M_0-fiber and each $\overline{f} \circ \iota_{\overline{q}}$ is a homothety, then as before one has that $\mathrm{E}(f) = \frac{\mathrm{vol}(M_0)}{\#(\psi_1)\#(\chi)}\mathrm{E}(\tilde{f} \circ \iota_{\tilde{q}}) = \frac{\mathrm{vol}(M)}{\mathrm{vol}(T^k)}\mathrm{E}(\tilde{f} \circ \iota_{\tilde{q}})$ for any fixed $\tilde{q} \in M_1$. Since $\tilde{f} \circ \iota_{\tilde{q}}$ is a homothety, there exists $a \geq 0$ such that $\|(\tilde{f} \circ \iota_{\tilde{q}})_*(\tilde{v})\| = a|\tilde{v}|$ for any $\tilde{z} \in T^k$ and $\tilde{v} \in \mathrm{T}_{\tilde{z}}T^k$. Directly computing the energy integral shows that $\mathrm{E}(\tilde{f} \circ \iota_{\tilde{q}}) = \frac{a^2k}{2}\mathrm{vol}(T^k)$, so $\mathrm{E}(f) = \frac{a^2k}{2}\mathrm{vol}(M)$. At the same time, since the image of each geodesic in T^k under $\tilde{f} \circ \iota_{\tilde{q}}$ is homotopy-minimizing in N, Lemma 7.3.2(b) implies that $\|T_*(\tilde{v})\| = a|\tilde{v}|$ for each $\tilde{q} \in T^k$ and $\tilde{v} \in \mathrm{T}_{\tilde{z}}T^k$, and another direct computation shows that $\ell_T = a\sqrt{\frac{kc_{k-1}}{2}}$. Thus $\frac{1}{c_k}\mathrm{vol}(M)\ell_T^2 = \frac{a^2k}{2}\mathrm{vol}(M)$. This shows that $\mathrm{E}(f) = \frac{1}{c_k}\mathrm{vol}(M)\ell_T^2$.

(c) If $\mathrm{E}(f) = \frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2}\mathrm{vol}(M)\ell_T^2$, then, by the main inequalities, it's also the case that $\mathrm{E}(f) = \frac{1}{c_{k-1}}\mathrm{vol}(M)\ell_T^2$. By part (b), $\tilde{f}$ is a homothety along each T^k-fiber. If $\ell_T = 0$, then $\mathrm{E}(f) = 0$, so f is constant and, trivially, a homothety. If $\ell_T > 0$, then $\frac{nc_n^2 c_{k-1}^2}{kc_{n-1}^2 c_k^2} = 1$, and it follows from Corollary 2.7.5(b) that $k = n$. This implies that M_1 is a point and, consequently, that $\tilde{f}$ is a homothety. This is equivalent to f being a homothety.

Conversely, suppose f is a homothety. If f is constant, then $\mathrm{E}(f) = 0$ and $\ell_T = 0$, so one has that $\mathrm{E}(f) = \frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2}\mathrm{vol}(M)\ell_T^2$. If M_1 is a point, then $k = n$, so $\frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2} = \frac{1}{c_{k-1}}$. It follows from part (b) that $\mathrm{E}(f) = \frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2}\mathrm{vol}(M)\ell_T^2$.

□

When every geodesic in $\overline{N}$ is minimal, the converse to Theorem 7.4.4(a) also holds, and the equivalence in Theorem 7.4.4(c) can be simplified.

Theorem 7.4.5. *Let M, N, and $[F]$ be as in the previous section, and denote by $(T^m, \|\cdot\|)$ and $T : T^k \to T^m$ the corresponding semi-Finsler torus and affine surjection. Suppose that $k > 0$ and that every geodesic in $\overline{N}$ is minimal. Let $f \in [F]$ be any C^1 map. Then the following hold:*

***(a)** If f is totally geodesic, then $\mathrm{E}(f) = \mathrm{vol}(M)e_T$; and*

***(b)** If f is a homothety, then $\mathrm{E}(f) = \frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2}\mathrm{vol}(M)\ell_T^2$.*

Proof. **(a)** By Lemma 5.10(b), $\tilde{f}$ is constant along each M_1-fiber, and consequently $\overline{F}$ is constant along each M_0-fiber. Moreover, for each $\tilde{q} \in M_0$, $\tilde{f} \circ \iota_{\tilde{q}}$ is totally geodesic. Since N has no conjugate points, for each $w \in \mathrm{S}T^k$, $\varsigma_{\overline{q},w} : \mathbb{R} \to N$ must be a minimal geodesic; an application of Lemma 7.4.1(b) thus shows that $\varsigma_{\overline{q},w}$ has speed $\|\overline{T}_*(w)\|$. A direct computation, which may be simplified by invoking Lemma 7.4.6[3], now shows that $\mathrm{E}(f) = \mathrm{vol}(M)e_T$.

(b) Since every homothety is totally geodesic, Lemma 5.10(b) implies that $\overline{F}$ is constant along each M_0-fiber. At the same time, there exists $a \geq 0$ such that $\|\overline{F}_*(\overline{w})\| = a|\overline{w}|$ for any $\overline{w} \in \mathrm{T}(M_0 \times \mathbb{R}^k)$. When f is non-constant, this is possible only if each M_i is a point. Thus $n = k$ and, by Corollary 2.7.5(b), $\frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2} = \frac{1}{c_{k-1}}$. The result now follows from Theorem 7.4.4(b). When f is constant, $\mathrm{E}(f) = 0$ and $\ell_T = 0$, which completes the proof.

□

[3] Should one choose not to use Lemma 7.4.6, one may still compute $\#(\psi_1)\#(\chi)\mathrm{E}(f)$ by integrating over $\mathrm{S}(M_0 \times P_1)$ and using the coarea formula, as in the proof of Theorem 7.4.1, to reduce it to an integral over $\mathrm{S}P_1$.

Lemma 7.4.6. *Let M, N, and $[F]$ be as in the previous section, and denote by $(T^m, \|\cdot\|)$ and $T : T^k \to T^m$ the corresponding semi-Finsler torus and affine surjection. Suppose that $k > 0$ and that every geodesic in $\overline{N}$ is minimal. If $[F]$ contains a totally geodesic map, then $\|\cdot\|$ is a semi-Riemannian norm.*

Proof. Let $f \in [F]$ be totally geodesic. By Lemma 7.3.1, one may suppose without loss of generality that $f_0 = f$. A consequence of Lemma 5.10(b) is that, for each $\overline{q} \in M_0$, $\overline{F} \circ \iota_{\overline{q}}$ is totally geodesic, which implies that $P := \overline{F} \circ \iota_{\overline{q}}(\mathbb{R}^k)$ is a totally geodesic and flat m-dimensional submanifold of $\overline{N}$. Since $\overline{F}$ is also constant along each M_0-fiber, both P and the lattice $\Gamma_2 := \overline{F} \circ \iota_{\overline{q}}(\Gamma_1) \subseteq P$ are independent of the choice of $\overline{q}$, and $\overline{f}$ is $\mathbb{Z}^k$-equivariant under the action that translates the $\mathbb{R}^k$-fibers by elements of Γ_1. Recall that $\overline{y} = \overline{F}(\overline{p}, \overline{x})$. As in the construction of the loops s_i in the previous section, for each $u = \overline{F}(\overline{p}, v) \in \Gamma_2$, where $v \in \Gamma_1$, the map $t \mapsto \overline{F}(\overline{p}, \overline{x} + tv) = \overline{y} + tu$, $t \in [0,1]$, descends via π to a closed geodesic in G. This identification defines a group isomorphism $\wp : \Gamma_2 \to G$, which in turn defines an isomorphism $\jmath := \imath \circ \wp : \Gamma_2 \to \mathbb{Z}^m$. Since each $\varsigma_{\overline{p},w}$ extends for all time to a geodesic in $\overline{N}$, which by assumption is minimal, an application of Lemma 7.3.2(b) shows that $\|\cdot\|_\jmath := \|\jmath(\cdot)\|$ agrees on Γ_2 with the Euclidean semi-norm possessed by P. It is now an elementary exercise to show that $\|\cdot\|$ is induced by a semi-definite bilinear form or, in other words, that $\|\cdot\|$ is a semi-Riemannian norm.

□

Remark 7.4.7. Suppose the induced homomorphism of the homotopy class $[F]$ in Theorem 7.4.4 is non-trivial, so that $m > 0$. If N is compact, then by modifying the proof of Theorem 8.3.19 in [BBI] one may show that d_G is bi-Lipschitz equivalent to a word metric and, as a consequence, that $(T^m, \|\cdot\|)$ is a Finsler torus. In particular, for compact N, $(T^m, \|\cdot\|)$ is a Riemannian torus whenever $[F]$ contains a totally geodesic map.

Remark 7.4.8. If the induced homomorphism of $[F]$ is non-trivial and N is compact, then $e_T > 0$ and, consequently, every map in $[F]$ has energy bounded below by a positive constant. There are apparently no known examples of bubbles forming in the heat flow between compact manifolds when the initial map lies in a homotopy class with that property. When N has no conjugate points, or more generally is aspherical, such an example of blow-up would be fantastically interesting. On the other hand, it would also be interesting, albeit somewhat less so, if one could show that the heat flow exists for all time and uniformly subconverges when the target is compact and has no conjugate points. Whether this is true is apparently, and rather surprisingly, an open question.

Remark 7.4.9. The energy of a totally geodesic map $T : T^m \to T^k$ between flat Riemannian toruses can be computed using the Gram-Schmidt procedure. The answer is a bit messy and involves a number of determinants of matrices containing the coefficients of the metrics on T^m and T^k.

Remark 7.4.10. As discussed in Chapter 2, Corollary 2.7.3 may be deduced from other results in this chapter. Suppose that $n > k > 0$, so that $l > 0$. Let T^n have any flat metric, and let $\rho : T^n = T^k \times T^l \to T^k$ denote projection. Then ρ is constant along each T^l-fiber and a homothety along each T^k-fiber, so by Theorem 7.4.4(b), $\mathrm{E}(\rho) = \frac{1}{c_{k-1}}\mathrm{vol}(T^n)\ell_T^2$, where $T : T^k \to T^k$ in this case is the identity map. Since ρ is not a homothety, Theorem 7.1.4 and Theorem 7.4.1 imply that $\mathrm{E}(\rho) > \frac{nc_n^2 c_{k-1}}{kc_{n-1}^2 c_k^2}\mathrm{vol}(T^n)\ell_T^2$. Thus $nc_n^2 c_{k-1}^2 < kc_{n-1}^2 c_k^2$, which, as shown in the proof of Corollary 2.7.5, is equivalent to the inequality in Corollary 2.7.3. Note that, since Corollary 2.7.3 was only used to prove Corollary 2.7.5, which in turn was used in the proof of Theorem 7.4.4 only to obtain part (c) and the final inequality, the potential for circular reasoning has been avoided.

Corollary 7.4.11. *Let M and N be Riemannian manifolds, where M is compact with non-negative Ricci curvature and every geodesic in $\overline{N}$ is minimal. Let $[F]$ be a homotopy class of maps from M to N. Then the following hold:*

(a) *There exists $C \geq 0$ such that, for each C^1 map $f \in [F]$, $\mathrm{E}(f) \geq C$, with equality if and only if f is totally geodesic; and*

(b) *If $[F]$ contains a totally geodesic map, then each C^1 map in $[F]$ is energy-minimizing if and only if it is totally geodesic.*

Proof. **(a)** This follows immediately from the first inequality in Theorem 7.4.4, Theorem 7.4.4(a), and Theorem 7.4.5(a) by setting $C := \mathrm{vol}(M)e_T$.

(b) Let $f \in [F]$ be totally geodesic. By Lemma 2.6.2, f is smooth. Part (a) now implies that $\mathrm{E}(f) = C$. Since any other C^1 map $g \in [F]$ satisfies $\mathrm{E}(g) \geq C$, it follows that g is energy-minimizing if and only if g is totally geodesic.

□

Corollary 7.4.11 immediately implies Theorem 1.4(b). It's important to keep in mind that the constant in Corollary 7.4.11 is merely a lower bound for the energy of maps in $[F]$, not necessarily the infimum. Moreover, not every homotopy class of maps between compact manifolds contains an energy-minimizing representative. For example, as shown by White [Wh], any compact and connected manifold M with $\dim(M) \geq 2$ and trivial first and second fundamental groups has the property that $\inf_{f \in [\mathrm{id}]} \mathrm{E}(f) = 0$, where $\mathrm{id} : M \to M$ is the identity map. In that case, [id] cannot contain an energy-minimizing representative. On the other hand, Croke–Fathi [CF] showed that the identity map is energy-minimizing for any compact manifold with no conjugate points and that any energy-minimizing map in [id] is an isometry. When N has no conjugate points, Corollary 7.4.11 shows that this lower bound C is realized exactly when $[F]$ contains a totally geodesic representative. Theorem 4.1.15 shows that $[F]$ may not have that property. However, by Theorem 1.4(c), it does whenever N has no focal points.

Chapter 8

Further questions

There are a number of interesting questions related to this work that I've been unable to resolve. Note that Theorem 1.4 is a result about energy-minimizing maps $M \to N$, rather than all harmonic maps. It's unknown whether every such harmonic map must be totally geodesic. By contrast, when N has non-positive curvature, all harmonic maps are energy-minimizing and totally geodesic by Theorem 1.2 and Theorem 1.3(c). Moreover, by Theorem 4.1.15, there exist compact manifolds N with no conjugate points that admit energy-minimizing maps $T^2 \to N$ which aren't totally geodesic.

Question 8.1. Is every harmonic map from a compact Riemannian manifold with non-negative Ricci curvature into a compact Riemannian manifold with no focal points energy-minimizing? By Theorem 1.4(b)-(c), this is equivalent to asking whether every such map is totally geodesic. The essential case is when the domain is a flat torus T^k for $k \geq 2$.

Question 8.2. Does there exist an energy-minimizing map from a compact Riemannian manifold with non-negative Ricci curvature into a complete but non-compact Riemannian manifold with no focal points that's not totally geodesic? Again, the important case is when the domain is a flat torus. By the flat strip theorem, this is equivalent to asking whether, whenever $\min(g) = \emptyset$ for some $g \in f_*(\pi_1(T^k, p))$, it's possible for $[f]$ to contain an energy-minimizing representative.

It's surprising to note that the existence of energy-minimizing, or more generally harmonic, maps into compact manifolds with no conjugate points is an open question, except in the case of surfaces, where it follows by combining the work of Eells [Ee], Gordon [Gor], and Schoen–Uhlenbeck [ScU] with the result of Burns [Burn2] that the universal cover of every complete surface with no conjugate points is convex-supporting.

Question 8.3. Does every homotopy class of maps from a compact Riemannian manifold with non-negative Ricci curvature into a compact Riemannian manifold with no conjugate points contain an energy-minimizing representative?

It's also not known whether the heat flow from compact domains exists for all time and uniformly subconverges into compact manifolds with no conjugate points, except in the case of maps between surfaces, where the affirmative answer follows from the work of Struwe [St1], Gordon [Gor], and Burns [Burn2]. As discussed in Remark 7.4.8, a negative answer would be fabulously interesting. By Theorem 4.2.2, it does when the target has no focal points.

Question 8.4. Does the heat flow from a compact Riemannian manifold exist for all time and uniformly subconverge when the target is a compact Riemannian manifold with no conjugate points?

In light of Lemma 7.4.6, one might ask whether, for compact N, the Finsler torus in the statement of Theorem 7.4.4 must always be Riemannian. By Theorem 1.4(c), it is whenever N has no focal points. It needn't be in general, as Bangert [Ba] showed that there exist Riemannian metrics on T^2 whose asymptotic norms in the sense of [Bura] are not Riemannian. If g is such a metric and h is a flat metric on T^2, then $\mathrm{id} : (T^2, h) \to (T^2, g)$ is a map whose corresponding Finsler torus is not Riemannian. For each $k < \infty$, examples of C^k metrics on high-dimensional toruses T^n whose asymptotic norms are non-differentiable at almost every irrational vector in $S^{n-1} \subseteq \mathbb{R}^n$ were constructed by Burago–Ivanov–Kleiner [BIK]. However, if one restricts to target manifolds with no conjugate points, the solution to the Hopf conjecture by Burago–Ivanov [BurI] precludes counterexamples on toruses. It would be surprising if the Finsler torus in the case of no conjugate points were always Riemannian, but to my knowledge there are no known counterexamples.

Question 8.5. If N is compact and has no conjugate points, must the Finsler torus in Theorem 7.4.4 be Riemannian?

Finally, one might note that the results of Chapter 6 nowhere assume that the manifold M_0 in the diagram (6.1) is compact, unlike the results of Chapter 7.

Question 8.6. To what extent do the conclusions of Theorem 7.4.4 and Theorem 7.4.5 hold when M_0 is only assumed to be complete and have finite volume?

References

[AAR] Andrews, George, Richard Askey, and Ranjan Roy, *Special functions*, Encyclopedia of Mathematics and its Applications, 71. Cambridge University Press, Cambridge (1999).

[AB] Alexander, Stephanie, and Richard Bishop, *Convex-supporting domains on spheres*, Illinois J. Math. 18 (1974), 37–47.

[Al1] Al'ber, Solomon, *The topology of functional manifolds and the calculus of variations in the large*, Usphei Mat. Nauk 25 (1970) 57–122.

[Al2] Al'ber, Solomon, *Spaces of mappings into a manifold of negative curvature*, Dokl. Akad. Nauk SSSR 178 (1968) 13–16.

[Am] Ambrose, Warren, *Parallel translation of Riemannian curvature*, Ann. of Math. 64 (1956) 337–363.

[Av] Avez, André, *Variétés riemanniennes sans points focaux*, C. R. Acad. Sci. Paris Sér. I Math. 270 (1970), 191–199.

[Ba] Bangert, Victor, *Geodesic rays, Busemann functions and monotone twist maps*, Calc. Var. Partial Differential Equations 2 (1994), 49–63.

[BBB] Ballmann, Werner, Michael Brin, and Keith Burns, *On surfaces with no conjugate points*, J. Differential Geom. 25 (1987), 249–273.

[BBI] Burago, Dmitri, Yuri Burago, and Sergei Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, 33. American Mathematical Society, Providence (2001).

[BE] Bangert, Victor, and Patrick Emmerich, *Area growth and rigidity of surfaces without conjugate points*, J. Differential Geom. 94 (2013), 367–385.

[Ber] Berger, Marcel, *A panoramic view of Riemannian geometry. Volume 2.*, Springer-Verlag: Berlin, 2003.

[BH] Böhm, Johannes, and Eike Hertel, *Polyedergeometrie in n-dimensionalen Räumen konstanter Krümmung*, Lehrbücher und Monographien aus dem Gebiete der Exakten Wissenschaften. Mathematische Reihe, 70. Birkhäuser Verlag, Basel (1981).

[BIK] Burago, Dmitri, Sergei Ivanov, and Bruce Kleiner, *On the structure of the stable norm of periodic metrics*, Math. Res. Lett. 4 (1997), 791–808.

[BK] Burns, Keith and Gerhard Knieper, *Rigidity of surfaces with no conjugate points*, J. Differential Geom. 34 (1991), 623–650.

[Bo] Bochner, Salomon, *Harmonic surfaces in Riemannian metric*, Trans. Amer. Math. Soc. 47 (1940), 146–154.

[Bura] Burago, Dmitri, *Periodic metrics*, Advances in Soviet Math. 9 (1992), 205–210.

[BurI] Burago, Dmitri, and Sergei Ivanov, *Riemannian tori without conjugate points are flat*, Geom. Funct. Anal. 4 (1994), 259–269.

[Burn1] Burns, Keith, *The flat strip theorem fails for surfaces with no conjugate points*, Proc. Amer. Math. Soc. 115 (1992), 199–206.

[Burn2] Burns, Keith, *Convex supporting domains on surfaces*, Bull. London Math. Soc. 17 (1985), 271–274.

[Buse] Busemann, Herbert, *The geometry of geodesics*, Academic Press Inc., New York (1955).

[BusI] Bustoz, Joaquin, and Mourad Ismail, *On gamma function inequalities*, Math. Comp. 47 (1986), 659–667.

[By] Byers, William, *On a theorem of Preissmann*, Proc. Amer. Math. Soc. 24 (1970), 50–51.

[CCR1] Cao, Jianguo, Jeff Cheeger, and Xiaochun Rong, *Local splitting structures on nonpositively curved manifolds and semirigidity in dimension 3*, Comm. Anal. Geom. 12 (2004), 389–415.

[CCR2] Cao, Jianguo, Jeff Cheeger, and Xiaochun Rong, *Splittings and Cr-structures for manifolds with nonpositive sectional curvature*, Invent. Math. 144 (2001), 139–167.

[CD] Chang, Kung-Ching, and Weiyue Ding, *A result on the global existence for heat flows of harmonic maps from $D2$ into $S2$*, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 332 (1991), 37–47.

[CDY] Chang, Kung-Ching, Weiyue Ding, and Rugang Ye, *Finite-time blow-up of the heat flow of harmonic maps from surfaces*, J. Differential Geom. 36 (1992), 507–515.

[CE] Cheeger, Jeff, and David Ebin, *Comparison theorems in Riemannian geometry*, North-Holland Mathematical Library, Vol. 9. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York (1975).

[Cen] Centore, Paul, *Finsler Laplacians and minimal-energy maps*, Internat. J. Math. 11 (2000), 1–13.

[CF] Croke, Christopher, and Albert Fathi, *An inequality between energy and intersection*, Bull. London Math. Soc. 22 (1990), 489–494.

[CG1] Cheeger, Jeff, and Detlef Gromoll, *On the structure of complete manifolds of nonnegative curvature*, Ann. of Math. 96 (1972), 413–443.

[CG2] Cheeger, Jeff, and Detlef Gromoll, *The splitting theorem for manifolds of nonnegative Ricci curvature*, J. Differential Geom. 6 (1971), 119–128.

[Chan] Chang, Kung-Ching, *Heat flow and boundary value problem for harmonic maps*, Ann. Inst. H. Poincarée Anal. Non Linéaire 6 (1989), 363–395.

[Chav] Chavel, Isaac, *Riemannian geometry: a modern introduction. Second edition.*, Cambridge Studies in Advanced Mathematics, 98. Cambridge University Press, Cambridge (2006).

[ChS] Chen, Yunmei, and Michael Struwe, *Existence and partial regularity results for the heat flow for harmonic maps*, Math. Z. 201 (1989), 83–103.

[CK1] Croke, Christopher, and Bruce Kleiner, *On tori without conjugate points*, Invent. Math. 120 (1995), 241–257.

[CK2] Croke, Christopher, and Bruce Kleiner, *A rigidity theorem for manifolds without conjugate points*, Ergodic Theory Dynam. Systems 18 (1998), 813–829.

[CoG] Coron, Jean-Michel, and Jean-Michel Ghidaglia *Explosion en temps fini pour le flot des applications harmoniques*, C. R. Acad. Sci. Paris Sér. I Math. 308 (1989), 339–344.

[Cr1] Croke, Christopher, *Lower bounds on the energy of maps*, Duke Math. J. 55 (1987), 901–908.

[Cr2] Croke, Christopher, *Volumes of balls in manifolds without conjugate points*, Internat. J. Math. 3 (1992), 455–467.

[CrS] Croke, Christopher, and Viktor Schroeder, *The fundamental group of compact manifolds without conjugate points*, Comment. Math. Helv. 61 (1986), 161–175.

[D] Davis, Michael Walter, *Groups generated by reflections and aspherical manifolds not covered by Euclidean space*, Ann. of Math. (2) 117 (1983), 292–324.

[DL] Ding, Weiyue, and Fanghua Lin, *A generalization of Eells–Sampson's theorem*, J. Partial Differential Equations 5 (1992), 13–22.

[Eb1] Eberlein, Patrick, *When is a geodesic flow of Anosov type? II*, J. Differential Geom. 8 (1973), 565–577.

[Eb2] Eberlein, Patrick, *Geodesic flow in certain manifolds without conjugate points*, Trans. Amer. Math. Soc. 167 (1972), 151–170.

[Eb3] Eberlein, Patrick, *When is a geodesic flow of Anosov type? I*, J. Differential Geom. 8 (1973), 437–463.

[Eb4] Eberlein, Patrick. unpublished.

[Ee] Eells, James, Jr., *Regularity of certain harmonic maps*, Global Riemannian Geometry, 139–147. Ellis Horwood Ser. Math. Appl., Horwood (1984).

[EH] Eschenburg, Jost-Hinrich, and Ernst Heintze, *An elementary proof of the Cheeger–Gromoll splitting theorem*, Ann. Global Anal. Geom. 2 (1984), 141–151.

[EM] Eilenberg, Samuel, and Saunders Mac Lane, *Relations between homology and homotopy groups of spaces*, Ann. of Math. 46 (1945), 480–509.

[EO'N] Eberlein, Patrick, and Barrett O'Neill, *Visibility manifolds*, Pacific J. Math. 46 (1973), 45–109.

[EO'S] Eschenburg, Jost-Hinrich, and John O'Sullivan, *Growth of Jacobi fields and divergence of geodesics*, Math. Z. 150 (1976), 221–237.

[ES] Eells, James, Jr., and Joseph Sampson, *Harmonic mappings of Riemannian manifolds*, Amer. J. Math. 86 (1964), 109–160.

[Esc] Eschenburg, Jost-Hinrich, *Horospheres and the stable part of the geodesic flow*, Math. Z. 153 (1977), 237–251.

[Ev] Evans, Lawrence Christopher, *A strong maximum principle for parabolic systems in a convex set with arbitrary boundary*, Proc. Amer. Math. Soc. 138 (2010), 3179–3185.

[FL] Fet, Abram, and Lazar Lyusternik, *Variational problems on closed manifolds*, Doklady Akad. Nauk SSSR (N.S.) 81 (1951), 17–18.

[FM] Freire, Alexandre, and Ricardo Mañé, *On the entropy of the geodesic flow in manifolds without conjugate points*, Invent. Math. 69 (1982), 375–392.

[GK] Grove, Karsten, and Hermann Karcher, *How to conjugate* C^1*-close group actions*, Math. Z. 132 (1973), 11–20.

[Gor] Gordon, William, *Convex functions and harmonic maps*, Proc. Amer. Math. Soc. 33 (1972), 433–437.

[Got] Goto, Midori, *Manifolds without focal points*, J. Differential Geom. 13 (1978), 341–359.

[Gr1] Green, Leon, *A theorem of E. Hopf*, Michigan Math. J. 5 (1958), 31–34.

[Gr2] Green, Leon, *Geodesic instability*, Proc. Amer. Math. Soc. 7 (1956), 438–448.

[Gr3] Green, Leon, *Surfaces without conjugate points*, Trans. Amer. Math. Soc. 76 (1954), 529–546.

[Gul] Gulliver, Robert, *On the variety of manifolds without conjugate points*, Trans. Amer. Math. Soc. 210 (1975), 185–201.

[Gur] Gurland, John, *On Wallis' formula*, Amer. Math. Monthly 63 (1956), 643–645.

[GW] Gromoll, Detlef, and Joseph A. Wolf, *Some relations between the metric structure and the algebraic structure of the fundamental group in manifolds of nonpositive curvature*, Bull. Amer. Math. Soc. 77 (1971), 545–552.

[Ham] Hamilton, Richard, *Harmonic maps of manifolds with boundary*, Lecture Notes in Mathematics, Vol. 471. Springer-Verlag, Berlin-New York (1975).

[Han] Hansen, Vagn Lundsgaard, *On a theorem of Al'ber on spaces of maps*, J. Differential Geom. 12 (1977), 565–566.

[Har] Hartman, Philip, *On homotopic harmonic maps*, Canad. J. Math. 19 (1967), 673–687.

[Heb] Heber, Jens, *On the geodesic flow of tori without conjugate points*, Math. Z. 216 (1994), 209–216.

[Her1] Hermann, Robert, *Differential geometry and the calculus of variations*, Second edition. Interdisciplinary Mathematics, XVII. Math Sci Press, Brookline (1977).

[Her2] Hermann, Robert, *Focal points of closed submanifolds of Riemannian spaces*, Nederl. Akad. Wetensch. Proc. Ser. A 66 (1963), 613–628.

[HI] Heintze, Ernst and Hans-Christoph Im Hof, *Geometry of horospheres*, J. Differential Geom. 12 (1977), 481–491.

[HM] Hedlund, Gustav, and Marston Morse, *Manifolds without conjugate points*, Trans. Amer. Math. Soc. 51 (1942), 362–386.

[Ho] Hopf, Eberhard, *Closed surfaces without conjugate points*, Proc. Nat. Acad. Sci. U.S.A. 34 (1948), 47–51.

[HR] Hopf, Heinz, and Willi Rinow, *Ueber den Begriff der vollständigen differentialgeometrischen Fläche*, Comment. Math. Helv. 3 (1931), 209–225.

[Hu] Hurewicz, W., *Anmerkung bei der Korrektur*, Monatsch. Math. Phys. 41 (1934), 423.

[I] Innami, Nobuhiro, *A proof of existence of the stable Jacobi tensor*, Proc. Japan Acad. Ser. A Math. Sci. 63 (1987), 39–41.

[IK] Ivanov, Sergei, and Vitali Kapovitch, *Manifolds without conjugate points and their fundamental groups*, J. Differential Geom. 96 (2014), 223–240.

[J1] Jost, Jürgen, *Nonlinear methods in Riemannian and Kählerian geometry*, Second edition. DMV Seminar, 10. Birkhäuser Verlag, Basel (1991).

[J2] Jost, Jürgen, *Riemannian geometry and geometric analysis*, Sixth edition. Universitext. Springer, Heidelberg (2011).

[J3] Jost, Jürgen, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics : ETH Zürich. Birkhäuser Verlag, Basel (1997).

[Ka] Kazarinoff, Donat, *On Wallis' formula*, Edinburgh Math. Notes 40 (1956), 19–21.

[Kle] Kleiner, Bruce. unpublished.

[Kli1] Klingenberg, Wilhelm, *Riemannian geometry*, Second edition. de Gruyter Studies in Mathematics, 1. Walter de Gruyter & Co., Berlin (1995).

[Kli2] Klingenberg, Wilhelm, *Contributions to Riemannian geometry in the large*, Ann. of Math. 69 (1959), 654–666.

[Kli3] Klingenberg, Wilhelm, *Riemannian manifolds with geodesic flow of Anosov type*, Ann. of Math. 99 (1974), 1–13.

[KM] Karuwannapatana, Wichitra, and Chaiwat Maneesawarng, *The lower semi-continuity of total curvature in spaces of curvature bounded above*, East-West J. Math. 9 (2007), 1–9.

[Kn] Knieper, Gerhard, *Mannigfaltigkeiten ohne konjugierte Punkte*, Dissertation, Rheinische Friedrich-Wilhelms-Universität, Bonn (1985).

[Leb] Lebedeva, Nina, *On the fundamental group of a compact space without conjugate points*, PDMI preprint, http://www.pdmi.ras.ru/preprint/2002/02-05.html, 2002.

[Lee1] Lee, John M. *Introduction to topological manifolds*, Second edition. Springer-Verlag: New York, 2011.

[Lee2] Lee, John M. *Introduction to smooth manifolds*, Second edition. Springer: New York, 2013.

[Lee3] Lee, John M. *Riemannian manifolds: an introduction to curvature*, Springer-Verlag: New York, 1997.

[Lem] Lemaire, Luc, *Applications harmoniques de surfaces riemanniennes*, J. Differential Geom. 13 (1978), 51–78.

[Li] Lin, Fanghua, *Gradient estimates and blow-up analysis for stationary harmonic maps*, Ann. of Math. 149 (1999), 785–829.

[LM] Lenbury, Yongwimon, and Chaiwat Maneesawarng, *Total curvature and length estimates for curves in CAT(K) spaces*, Differential Geom. Appl. 19 (2003), 211–222.

[LT] Li, Jiayu, and Gang Tian, *A blow-up formula for stationary harmonic maps*, Internat. Math. Res. Notices 14 (1998), 735–755.

[LW] Lin, Fanghua, and Changyou Wang, *Harmonic and quasi-harmonic spheres*, Comm. Anal. Geom. 7 (1999), 397–429.

[LY] Lawson, H. Blaine, Jr., and Shing-tung Yau, *Compact manifolds of nonpositive curvature*, J. Differential Geom. 7 (1972), 211–228.

[LZ] Li, Jiayu, and Xiangrong Zhu, *Non existence of quasi-harmonic spheres*, Calc. Var. Partial Differential Equations 37 (2010), 441–460.

[Ma] Mal'cev, Anatoliĭ, *On isomorphic matrix representations of infinite groups of matrices*, Rec. Math. [Mat. Sbornik] N.S. 8 (1940), 405–422.

[Mo] Morse, Marston, *The calculus of variations in the large*, Reprint. American Mathematical Society Colloquium Publications, 18. American Mathematical Society, Providence (1996).

[O'S1] O'Sullivan, John, *Manifolds without conjugate points*, Math. Ann. 210 (1974), 295–311.

[O'S2] O'Sullivan, John, *Riemmanian manifolds without focal points*, J. Differential Geom. 11 (1976), 321–333.

[Ot] Ottarsson, S. K., *Closed geodesics on Riemannian manifolds via the heat flow*, J. Geom. Phys. 2 (1985), 49–72.

[Oz] Ozols, Vilnis, *Critical points of the displacement function of an isometry*, J. Differential Geometry 3 (1969), 411–432.

[P] Preissmann, Alexandre *Quelques propriétés globales des espaces de Riemann*, Comment. Math. Helv. 15 (1943), 175–216.

[Q] Qi, Feng, *Bounds for the ratio of two gamma functions*, J. Inequal. Appl. (2010), Art. ID 493058, 84pp.

[R] Rauch, Harry, *A contribution to differential geometry in the large*, Ann. of Math. 54 (1951), 38–55.

[Se] Selberg, Atle, *On discontinuous groups in higher-dimensional symmetric spaces*, Contributions to Function Theory, Tata Institute of Fundamental Research, Bombay (1960), 147–164.

[SaU] Sacks, Jonathan, and Karen Uhlenbeck, *Minimal immersions of closed Riemann surfaces*, Trans. Amer. Math. Soc. 271 (1982), 639–652.

[ScU] Schoen, Richard, and Karen Uhlenbeck, *A regularity theory for harmonic maps*, J. Differential Geom. 17 (1982), 307–335.

[Sm1] Smith, Paul, *A theorem on fixed points for periodic transformations*, Ann. of Math. 35 (1934), 572–578.

[Sm2] Smith, Paul, *Fixed-point theorems for periodic transformations*, Amer. J. Math. 63 (1941), 1–8.

[St1] Struwe, Michael, *On the eovlution of harmonic mappings of Riemannian surfaces*, Comment. Math. Helv. 60 (1985) 558–581.

[St2] Struwe, Michael, *On the evolution of harmonic maps in higher dimensions*, J. Differential Geom. 28 (1988), 485–502.

[Topo] Toponogov, Viktor, *Riemannian spaces containing straight lines*, Dokl. Akad. Nauk SSSR 127 (1959), 977–979.

[Topp1] Topping, Peter, *Rigidity in the harmonic map heat flow*, J. Differential Geom. 45 (1997), 593–610.

[Topp2] Topping, Peter, *An example of a nontrivial bubble tree in the harmonic map heat flow*, Harmonic morphisms, harmonic maps, and related topics (Brest, 1997), 185–191. Chapman & Hall/CRC Res. Notes Math., 413. Chapman & Hall/CRC, Boca Raton (2000).

[Topp3] Topping, Peter, *Winding behaviour of finite-time singularities of the harmonic map heat flow*, Math. Z. 247 (2004), 279–302.

[Wal] Walter, Rolf, *On the metric projection onto convex sets in Riemannian spaces*, Arch. Math. 25 (1974), 91–98.

[Wat] Watkins, Jordan, *The higher rank rigidity theorem for manifolds with no focal points*, Geom. Dedicata 164 (2013), 319–349.

[Wh] White, Brian, *Infima of energy functionals in homotopy classes of mappings*, J. Differential Geom. 23 (1986), 127–142.

[Wo] Wolf, Joseph A., *Homogeneity and bounded isometries in manifolds of negative curvature*, Illinois J. Math. 8 (1964), 14–18.

[X1] Xin, Yuanlong, *Geometry of harmonic maps*, Progress in Nonlinear Differential Equations and their Applications, 23. Birkhäuser Boston, Inc., Boston (1996).

[X2] Xin, Yuanlong, *Nonexistence and existence for harmonic maps in Riemannian manifolds*, Proc. DD2, 1981. Science Press (1984), 529–538.

[Y] Yau, Shing-tung, *On the fundamental group of compact manifolds of non-positive curvature*, Ann. of Math. 93 (1971), 579–585.