THE HOPF-RINOW THEOREM

DANIEL SPIEGEL

Abstract. This paper is an introduction to Riemannian geometry, with an aim towards proving the Hopf-Rinow theorem on complete Riemannian manifolds. We assume knowledge of the basics of smooth manifolds, including the tangent and cotangent bundles and vector fields. After a brief introduction to tensors, we develop the foundations of Riemannian geometry: geodesics, the exponential map, and the Riemannian distance metric. We conclude with a proof of the Hopf-Rinow theorem, which states, among other things, that geodesics on a complete Riemannian manifold are defined for all time.

Contents

1. Introduction

Riemannian geometry is a geometry of manifolds that bears some resemblance to the Euclidean geometry of $\mathbb{R}^n$. In particular, every tangent space on a Riemannian manifold is equipped with an inner product. This leads to analogues of several concepts in $\mathbb{R}^n$, such as straight lines, convex sets, arc lengths, and distances, being defined on a Riemannian manifold. In this paper we develop these fundamental concepts of Riemannian geometry.

We take special interest in the case of geodesics, the analogues of straight lines. Euclid's first and second postulates are that any two points can be connected by a straight line segment and that any line segment can be extended indefinitely [4]. On a complete and connected Riemannian manifold, the latter of these statements translates into the Hopf-Rinow theorem, with the former following as a corollary.

Date: August 3, 2016.

We develop the basics of Riemannian geometry with an aim towards proving this theorem.

We assume basic knowledge of the theory of smooth manifolds and vector bundles. For the prerequisite knowledge, we refer the reader to chapters 1-3, 8, 10-11 of John M. Lee's *Introduction to Smooth Manifolds*. Much of our notation is the same as in Lee's book. Given a manifold M and a point $p \in M$ the tangent space at p is denoted by T_pM and the cotangent space by T_p^*M. We denote by $C^\infty(M)$ the ring of all smooth functions $M \to \mathbb{R}$. Local coordinate charts are denoted by any of (U, φ), (U, x^i), or (x^i) depending on what we want to emphasize. Given local coordinates (x^i) we let $\partial_i = \partial/\partial x^i$ and dx^i denote the induced coordinate vector and covector fields respectively. In addition, all manifolds are Hausdorff and second-countable, and "smooth" is synonymous with C^∞.

For ease of notation we use the Einstein summation convention. Explicitly, if an index is repeated as an upper and a lower index in a monomial term, then summation over that index is implied. An upper index in the denominator of a fraction or partial derivative is to be interpreted as a lower index for this purpose. For example,

$$a^i b_i + x^{jk} \frac{\partial y_j}{\partial z^k} = \sum_i a^i b_i + \sum_{j,k} x^{jk} \frac{\partial y_j}{\partial z^k}.$$

Before delving into Riemannian geometry, we briefly discuss tensors in Section 2. In Section 3 we develop the basic tools on Riemannian manifolds, including the Levi-Civita connection, geodesics, the exponential map, and normal neighborhoods. In Section 4 we define the arc length of an arbitrary curve and imbue connected Riemannian manifolds with a special metric, then finally we prove the Hopf-Rinow theorem.

2. Tensors

Tensors play an essential role in Riemannian geometry: they provide the inner products to a manifold's tangent spaces. In particular we will deal with tensor fields, which will allow us to smoothly string together inner products between tangent spaces. In addition, we will see that tensor fields have a number of useful properties, such as $C^\infty(M)$-linearity and pointwise action on vector fields. This section is adapted from Chapter 12 of [1]. Proofs are omitted from this section to prevent distraction from our goal of Riemannian geometry; proofs may be found in [1].

While more broad and abstract definitions of a tensor exist, it will be sufficient for us to define an ***(m,n)-tensor*** F to be a multilinear map $F : (V^*)^m \times V^n \to \mathbb{R}$, where V is a finite dimensional vector space. If $m = 0$, then F is said to be ***covariant n-tensor*** and if $n = 0$ then F is said to be ***contravariant m-tensor***. For example, any $\omega \in V^*$ is a covariant 1-tensor and any $v \in V$ can be thought of as a contravariant 1-tensor under the canonical isomorphism $V \cong V^{**}$. We will only be concerned with covariant tensors in this paper, though many of the statements can be generalized to arbitrary tensors.

Denote the set of covariant k-tensors by $T^k(V^*)$. Note that $T^k(V^*)$ is a vector space under the usual pointwise addition and scalar multiplication operations. Given $F \in T^j(V^*)$ and $W \in T^k(V^*)$, we define the ***tensor product*** of F and G to be the covariant $(j+k)$-tensor:

$$F \otimes G\,(v_1, \ldots, v_{j+k}) = F(v_1, \ldots, v_j)G(v_{j+1}, \ldots, v_{j+k})$$

It is easy to check that the tensor product operation is bilinear and associative, so we may drop any parentheses when taking multiple tensor products. Tensor products are important because they provide us with a basis for $T^k(V^*)$.

Proposition 2.1. *Let V be an n-dimensional vector space and let (e_i) be a basis for V. If (ε^i) is the dual basis for V^*, then for fixed k the set*

$$\mathscr{B} = \{\varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_k} : 1 \leq i_j \leq n \text{ for all } j\}$$

is a basis for $T^k(V^*)$. Thus, $\dim T^k(V^*) = n^k$.

The collection of all $T^k(T_p^*M)$ for each $p \in M$ is grouped together to form the ***bundle of covariant k-tensors on M***

$$T^kT^*M = \coprod_{p \in M} T^k(T_p^*M).$$

This tensor bundle has a natural structure as a smooth manifold and a vector bundle, though we shall neither need nor discuss the details of this structure.

A ***covariant k-tensor field*** on M is a map $A : M \to T^kT^*M$ such that $\pi \circ A = id_M$, where π is the natural projection on T^kT^*M. It assigns a tensor to each point in M, these tensors smoothly varying from point to point. We can add tensor fields, multiply them by functions $f : M \to \mathbb{R}$, and form tensor products of tensor fields, all in a pointwise fashion, like so: $(A \otimes B)_p = A_p \otimes B_p$. For example, in a coordinate chart (U, x^i) the covector fields dx^i form covariant 1-tensor fields and the $dx^{i_1} \otimes \cdots \otimes dx^{i_k}$ form covariant k-tensor fields. Since the covectors $dx^i|_p$ form a basis for T_p^*M for each $p \in U$, by Proposition 2.1 we can write any covariant k-tensor A in U as

$$A = A_{i_1 \ldots i_k} dx^{i_1} \otimes \cdots \otimes dx^{i_k},$$

where the ***component functions*** $A_{i_1 \ldots i_k}$ are functions on U. A smooth tensor field yields smooth component functions, and induces a few other smooth maps as well. In the following proposition, and the rest of the paper, we let $\mathfrak{X}(M)$ denote the set of all smooth vector fields on M.

Proposition 2.2. *Let A be a covariant k-tensor field on M. The following are equivalent:*

1. *A is a smooth map from M to T^kT^*M.*
2. *The component functions of A are smooth in every coordinate chart.*
3. *Each $p \in M$ is contained is a coordinate chart on which the component functions of A are smooth.*
4. *For any $X_1, \ldots, X_k \in \mathfrak{X}(M)$, the function $A(X_1, \ldots, X_k) : M \to \mathbb{R}$ defined as*

$$A(X_1, \ldots, X_k)(p) = A_p(X_1|_p, \ldots, X_k|_p)$$

is smooth.

Furthermore, if A is smooth then the function $v \mapsto A_p(v, \ldots, v)$ is a smooth map on TM.

Item (4) of the above proposition shows that a smooth k-tensor A induces a map $\mathfrak{X}(M)^k \to C^\infty(M)$. It follows from multilinearity of A that this map is not only multilinear but is ***$C^\infty(M)$-multilinear***, i.e. for $f, f' \in C^\infty(M)$ we have

$$A(X_1, \ldots, fX_i + f'X_i', \ldots, X_k) = fA(X_1, \ldots, X_i, \ldots, X_k) + f'A(X_1, \ldots, X_i', \ldots, X_k).$$

In fact this property characterizes smooth covariant tensors.

Proposition 2.3. *A map $A : \mathfrak{X}(M)^k \to C^\infty(M)$ is induced by a covariant k-tensor as above if and only if it is $C^\infty(M)$-multilinear.*

The proof uses $C^\infty(M)$-linearity of A to show that for any $X_1, \ldots, X_k \in \mathfrak{X}(M)$, the value $A(X_1, \ldots, X_k)(p)$ depends only on the values of the vector fields at p. This allows us to obtain a well-defined tensor field $A_p(v_1, \ldots, v_k) = A(V_1, \ldots, V_k)(p)$, where $V_1, \ldots, V_k$ are arbitrary extensions of $v_1, \ldots, v_k$ to smooth vector fields.

In the following section, we will make use of tensor fields to define the Riemannian metric and $C^\infty(M)$-linearity to define the Levi-Civita connection.

3. Riemannian Geometry

3.1. **Basic Constructs.** In this section we imbue our manifold with a special kind of tensor field, called a Riemannian metric. This will make every tangent space to the manifold an inner product space, hence the name Riemannian geometry. In particular, we shall be concerned with geodesics, the analogue of straight lines on manifolds. This section is adapted from Chapter 3 of [3].

Definition 3.1. A ***Riemannian metric*** g is a smooth covariant 2-tensor field on a manifold M such that for all $p \in M$ and $v, w \in T_pM$:

(1) $g_p(v, w) = g_p(w, v)$, i.e. g is *symmetric*.
(2) $g_p(v, v) > 0$ for $v \neq 0$, i.e. g is *positive-definite*.

A pair (M, g) of a manifold and Riemannian metric is called a ***Riemannian manifold***. We will refer to Riemannian manifolds by the underlying set M.

Throughout the rest of the paper M will be a Riemannian manifold. By definition of a tensor g_p is bilinear for each $p \in M$, thus g_p is an inner product on T_pM. We use the notation $g_p(v, w) = \langle v, w\rangle$ to reflect this. As usual, we denote the induced norm on T_pM by $|v| = \sqrt{\langle v, v\rangle}$. It is a continuous function $TM \to \mathbb{R}$, smooth away from any zeros in TM.

In any coordinate chart (U, x^i) on M the Riemannian metric can be written as $g = g_{ij}\, dx^i \otimes dx^j$. Since g is smooth the component functions g_{ij} are smooth on U by Proposition 2.2. Furthermore, since g is positive-definite it follows that the matrix (g_{ij}) is invertible at each $p \in U$. We denote the elements of the inverse matrix by g^{ij}. Since the g^{ij} are rational functions of the g_{ij}, they are also smooth.

Example 3.2. Euclidean space $\mathbb{R}^n$ is the most basic example of a Riemannian manifold. With the covector fields dx^i defined with respect to the standard basis, $\mathbb{R}^n$ is equipped with the Riemannian metric $g = \delta_{ij} dx^i \otimes dx^j$ where δ_{ij} is the Kronecker delta, which is just the usual inner product on $\mathbb{R}^n$ through the canonical isomorphism $\mathbb{R}^n \cong T_p\mathbb{R}^n$.

As diffeomorphisms provide a notion of equivalence between smooth manifolds, we have the following notion of equivalence between Riemannian manifolds.

Definition 3.3. An ***isometry*** between Riemannian manifolds (M, g_M) and (N, g_N) is a diffeomorphism of smooth manifolds $\varphi : M \to N$ that preserves the metrics i.e. for any $p \in M$ and $v, w \in T_pM$,

$$\langle d\varphi_p(v), d\varphi_p(w)\rangle = \langle v, w\rangle .$$

We will not encounter any true isometries between Riemannian manifolds in this paper, but we include the definition for the sake of completeness.

Our final construct (which we shall have use for) will be the Levi-Civita connection. Given vector fields $V, W \in \mathfrak{X}(M)$, the Levi-Civita connection $D : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ gives us a third vector field, notated $D(V, W) = D_V W$, which describes the change of W at $p \in M$ in the V_p direction. On $\mathbb{R}^n$ this intuitive notion translates easily into the vector field:

$$D_V W := V(W^i)\partial_i.$$

However, this coordinate dependent definition does not lend itself to abstraction to an arbitrary Riemannian manifold. But notice the following coordinate independent properties of the above vector field:

(i) $D_V W$ is $C^\infty(M)$-linear in V,
(ii) $D_V W$ is $\mathbb{R}$-linear in W,
(iii) $D_V(fW) = (Vf)D_V W + fD_V W$ for all $f \in C^\infty(M)$,
(iv) $[V, W] = D_V W - D_W V$,
(v) $X\langle V, W\rangle = \langle D_X V, W\rangle + \langle V, D_X W\rangle$ for all $X \in \mathfrak{X}(M)$.

These are the properties that allow us to define the Levi-Civita connection.

Theorem 3.4. *There exists a unique map $D : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, called the* ***Levi-Civita connection****, satisfying properties* (i) *through* (v) *above. Furthermore, D is characterized by the* ***Koszul formula****:*

$$\begin{aligned} 2\langle D_V W, X\rangle = V\langle W, X\rangle + W\langle V, X\rangle - X\langle V, W\rangle \\ - \langle V, [W, X]\rangle + \langle W, [X, V]\rangle + \langle X, [V, W]\rangle. \end{aligned}$$

The proof is lengthy, but mostly straightforward. We omit the proof to avoid the detour. See Proposition 3.10 and Theorem 3.11 in [3] for the proof.

We will want to be able to deal with the Levi-Civita connection in local coordinates. If (U, x^i) are coordinates on M then we define the ***Christoffel symbols*** $\Gamma^k_{ij} : U \to \mathbb{R}$ to be the smooth component functions:

$$D_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k.$$

It follows from property (iii) that:

$$D_{\partial_i} W = \left\{ \frac{\partial W^k}{\partial x^i} + W^j \Gamma^k_{ij} \right\} \partial_k.$$

This helps us to compute $D_V W$ for any $V, W \in \mathfrak{X}(M)$. The following coordinate representation of the Christoffel symbols will also be useful:

Proposition 3.5. *With respect to a coordinate system (x^i), the Christoffel symbols are given by*

$$\Gamma^k_{ij} = \frac{1}{2} g^{km} \left\{ \frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m} \right\}.$$

Proof. In the Koszul formula we set $V = \partial_i$, $W = \partial_j$, and $X = \partial_m$. The Lie brackets vanish, leaving:

$$\begin{aligned} 2\langle D_{\partial_i}\partial_j, \partial_m\rangle &= \partial_i\langle \partial_j, \partial_m\rangle + \partial_j\langle \partial_i, \partial_m\rangle - \partial_m\langle \partial_i, \partial_j\rangle \\ 2\Gamma^\ell_{ij} g_{\ell m} &= \frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m}. \end{aligned}$$

Multiplying by g^{km} and summing over m yields the desired result. ∎

3.2. **Geodesics.** We now construct geodesics on M—the analogues of straight lines in $\mathbb{R}^n$. We shall see that it is the property of vanishing acceleration that is abstracted to define the geodesics. Therefore before we define geodesics we need to make sense of the acceleration of a curve.

If $\alpha : I \to M$ is a smooth curve, a vector field on α is a smooth map $Z : I \to TM$ such that $\pi \circ Z = \alpha$. We denote the set of all such vector fields by $\mathfrak{X}(\alpha)$. The most important examples of vector fields on α are the velocity field α' and vector fields of the form $V|_\alpha = V \circ \alpha$ for $V \in \mathfrak{X}(M)$. Since we want to talk about the acceleration of a curve, we look to taking a time derivative of $Z \in \mathfrak{X}(\alpha)$.

Theorem 3.6. *Let $\alpha : I \to M$ be a smooth curve, with I an open interval. There exists a unique map $\mathfrak{X}(\alpha) \to \mathfrak{X}(\alpha)$, denoted $Z \mapsto Z'$, such that:*

(i) $(a_1Z_1 + a_2Z_2)' = a_1Z_1' + a_2Z_2'$ *for all* $a, b \in \mathbb{R}$,
(ii) $(hZ)' = h'Z + hZ'$ *for all* $h \in C^\infty(I)$,
(iii) $(V|_\alpha)'(t) = D_{\alpha'(t)}V(\alpha(t))$ *for all* $V \in \mathfrak{X}(M)$ *and* $t \in I$.

Furthermore,

$$\frac{d}{dt}\langle Z_1, Z_2\rangle = \langle Z_1', Z_2\rangle + \langle Z_1, Z_2'\rangle .$$

The map $Z \mapsto Z'$ is called the ***induced covariant derivative****.*

Again, the full proof is a little too lengthy to be included. See Proposition 3.18 in [3] for the proof.

In writing property (iii) we have taken advantage of the fact that D_VW is tensorial in V, as it follows from Propositions 2.2 and 2.3 that D_vW is well-defined for any $v \in TM$. Note also that if $Z \in \mathfrak{X}(\alpha)$ and (U, x^i) is a coordinate chart, then on $\alpha^{-1}(U)$ we have the coordinate representation $Z = Z^i\partial_i|_\alpha$, where the Z^i are smooth real valued functions. Applying all three properties we can compute a coordinate representation for the induced covariant derivative

$$Z' = \frac{dZ^i}{dt}(\partial_i \circ \alpha) + Z^i(D_{\alpha'}\partial_i \circ \alpha) = \left(\frac{dZ^k}{dt} + Z^i\Gamma^k_{ij}(\alpha')^j\right)\partial_k|_\alpha. \tag{1}$$

We can now define a ***geodesic*** to be a curve $\gamma : I \to M$ whose velocity γ' has vanishing induced covariant derivative on $\operatorname{int} I$. In other words, a geodesic is a curve of acceleration zero $\gamma'' \equiv 0$. Acceleration zero implies a constant speed in the sense that

$$\frac{d}{dt}\langle\gamma', \gamma'\rangle = 2\langle\gamma', \gamma''\rangle = 0.$$

In particular, if $\gamma'(t_0) \neq 0$ for some t_0, then γ' is never zero. Geodesics depend not only on their image in M but also on their parameterization.

Proposition 3.7. *Let $\gamma : I \to M$ be a nonconstant geodesic and $h : J \to I$ a smooth function between intervals. Then $\gamma \circ h$ is a geodesic if and only if $h'' \equiv 0$.*

Proof. The velocity field of the reparameterized curve is $(\gamma \circ h)' = h' \cdot (\gamma' \circ h)$. Applying the product rule and Eq. (1) to a coordinate representation around any point yields:

$$(\gamma \circ h)'' = h'' \cdot (\gamma' \circ h) + (h')^2 \cdot (\gamma'' \circ h).$$

That γ is a nonconstant geodesic implies that $\gamma'' = 0$ and γ' is never zero. Thus we see that $(\gamma \circ h)'' = 0$ is equivalent to $h'' = 0$. □

Our goal for the remainder of this section will be to prove an existence and uniqueness theorem for geodesics. We require two short lemmas.

Lemma 3.8. *Given $v \in T_pM$ there exists an open interval I about 0 and a unique geodesic $\gamma : I \to M$ such that $\gamma'(0) = v$.*

Proof. In local coordinates we rewrite Eq. (1) for $Z = \gamma'$:

$$\gamma'' = \left\{ \frac{d^2\gamma^k}{dt^2} + \frac{d\gamma^i}{dt}\frac{d\gamma^j}{dt}\Gamma^k_{ij} \right\} \partial_k. \tag{2}$$

Setting this to zero yields a system of second order ordinary differential equations. The result then follows by the existence and uniqueness theorem for ODEs. □

Lemma 3.9. *If $\alpha, \beta : I \to M$ are geodesics such that $\alpha'(a) = \beta'(a)$ for some $a \in I$, then $\alpha \equiv \beta$.*

Proof. Suppose there exists $t_0 \in I$ such that $\alpha(t_0) \neq \beta(t_0)$. Let $t_0 > a$; the proof for $t_0 < a$ is similar. Since the set $\{t \in I : t > a \text{ and } \alpha(t) \neq \beta(t)\}$ is nonempty and bounded below, it has a greatest lower bound b. We want to show that $\alpha'(b) = \beta'(b)$. Since this is trivial for $a = b$, assume $b > a$. Then α and β agree on (a, b), so their velocities agree on (a, b) as well. By continuity of the velocity of a smooth curve, we have $\alpha'(b) = \beta'(b)$.

Suppose $b \in \operatorname{int} I$. Since $t \mapsto \alpha(b+t)$ and $t \mapsto \beta(b+t)$ are geodesics by Proposition 3.7, and since they share a common velocity at $t = 0$, Lemma 3.8 implies that they agree on some interval J about b, but this contradicts the definition of b. Note that if b is a right endpoint of I then $b = t_0$ so $\alpha'(b) = \beta'(b)$ contradicts $\alpha(t_0) \neq \beta(t_0)$. □

Lemma 3.8 gives us a geodesic whose domain is an unspecified interval I about 0. With the help Lemma 3.9, we can make the interval as large as possible.

Theorem 3.10. *Given $v \in T_pM$ there exists a unique geodesic $\gamma_v : I_v \to M$ such that $\gamma_v'(0) = v$ and for any other geodesic $\eta : I \to M$ with $\eta'(0) = v$ we have $I \subset I_v$ and $\eta = \gamma_v|_I$, i.e. the domain of I_v is maximal.*

Proof. Let $\mathcal{G}$ be the collection of all geodesics with initial velocity v. Any $\eta_1, \eta_2 \in \mathcal{G}$ agree on the intersection of their domains by Lemma 3.9. Thus, if I_v is the union of all the domains of elements of $\mathcal{G}$, then there exists a well defined curve $\gamma_v : I_v \to M$ whose value $\gamma(t)$ is equal to the common value $\eta(t)$ for all $\eta \in \mathcal{G}$ with t in their domain. That γ_v is a geodesic with initial velocity v follows immediately from the fact that the differential and induced covariant derivative act locally. □

The notation γ_v and I_v will be used frequently to denote the maximal geodesic and its domain. To conclude this section, consider again Eq. (2). The solutions to an ODE are smooth not only in the parameter but also in the initial data, a fact that immediately gives us the following lemma, which we shall use shortly.

Lemma 3.11. *If $v \in TM$, then there exists a neighborhood $U \subset TM$ of v and an interval I about 0 such that $(w, t) \mapsto \gamma_w(t)$ is a smooth map from $U \times I$ to M.*

3.3. **The Exponential Map.**

In this section we define and examine the exponential map. As we shall see, it provides us with a means of locally relating M and T_pM, and from it we can construct especially nice coordinate charts on M.

Definition 3.12. Let $\mathscr{D} = \{v \in TM : 1 \in I_v\}$ and $\mathscr{D}_p = \mathscr{D} \cap T_pM$ for each $p \in M$. We define the ***exponential map*** $\exp : \mathscr{D} \to M$ by

$$\exp(v) = \gamma_v(1).$$

for all $v \in \mathscr{D}$. For each $p \in M$ we denote the restriction of exp to $\mathscr{D}_p \subset T_pM$ by $\exp_p : \mathscr{D}_p \to M$.

Note that exp is well-defined by Theorem 3.10 and that $\mathscr{D}$ (resp. $\mathscr{D}_p$) is the largest subset of TM on which exp (resp. $\exp_p$) is defined. In fact, $\mathscr{D}$ and $\mathscr{D}_p$ are open sets of TM and T_pM, although we will not need this fact; for a proof, see [3]. Given $v \in T_pM$ and $t \in \mathbb{R}$, Proposition 3.7 tells us that $s \mapsto \gamma_v(st)$ is a geodesic, and by the chain rule it has initial velocity $t\gamma_v'(0) = tv$. By uniqueness of geodesics $\gamma_{tv}(s) = \gamma_v(st)$ for all s such that either side is defined. If $v \in \mathscr{D}_p$ then this implies

$$\exp(tv) = \gamma_{tv}(1) = \gamma_v(t),$$

so the exponential map takes lines in T_pM to geodesics in M. Since each geodesic γ_v is defined on an interval I_v, this also implies that $\mathscr{D}_p$ is starshaped for each $p \in M$.

We want to show that exp is smooth on a neighborhood of $0_p \in TM$ for each $p \in M$. By Lemma 3.11 there exists a neighborhood $U \subset TM$ of 0_p and an interval $(-\varepsilon, \varepsilon)$ such that $(v, t) \mapsto \gamma_v(t)$ is smooth on $U \times (-\varepsilon, \varepsilon)$. If $1 < \varepsilon$, then restricting this map to $t = 1$ yields the exponential map, so the result follows automatically. Suppose $\varepsilon \leq 1$. It follows from the definition of the topology on TM that we can find an open set V containing 0_p such that $V \subset U$ and $2v/\varepsilon \in U$ for all $v \in V$. Thus for all $v \in V$ we know:

$$\gamma_v(1) = \gamma_{2v/\varepsilon}(\varepsilon/2)$$

is defined. Therefore the composite map $(v, t) \mapsto (2v/\varepsilon, \varepsilon t/2) \mapsto \gamma_{2v/\varepsilon}(\varepsilon t/2)$ is defined and smooth on $V \times (-2, 2)$. Composing once more with $v \mapsto (v, 1)$ yields the exponential map, which is smooth on V as it is a composition of smooth functions. Since the inclusion $i_p : T_pM \to TM$ is smooth and $\exp_p = \exp \circ i_p$, this shows that $\exp_p$ is smooth on a neighborhood of 0 as well. The next proposition shows that we can do even better.

Proposition 3.13. *For each $p \in M$ there exists a neighborhood $\widetilde{U}$ of 0 in T_pM such that $\exp_p$ is a diffeomorphism from $\widetilde{U}$ onto $U = \exp_p(\widetilde{U})$.*

Proof. Consider the differential $d(\exp_p)_0 : T_0(T_pM) \to T_pM$. We claim that $d(\exp_p)_0$ is the canonical isomorphism $v_0 \mapsto v$ between $T_0(T_pM)$ and T_pM. Given $v \in T_pM$, define $\rho(t) = tv$, so $\rho'(0) = v_0 \in T_0(T_pM)$. Composing with $\exp_p$ yields:

$$d(\exp_p)_0(v_0) = d(\exp_p)_0(\rho'(0)) = d(\exp_p \circ \rho)_0(0) = \gamma_v'(0) = v.$$

Thus $d(\exp_p)_0$ is an isomorphism, so the result follows by the inverse function theorem on manifolds. □

With the aid of the exponential map, we now begin constructing a collection of useful neighborhoods and coordinates. We call $U \subset M$ a ***normal neighborhood*** of p if it is the image $\exp_p(\widetilde{U})$ of a starshaped neighborhood $\widetilde{U}$ of $0 \in T_pM$ on which $\exp_p$ is a diffeomorphism. When working on a normal neighborhood, $\exp_p^{-1}$ will denote the inverse function of $\exp_p|_{\widetilde{U}}$. Since any open neighborhood of 0 contains a starshaped neighborhood of 0, it follows from Proposition 3.13 that we can find a normal neighborhood of any point in M. Normal neighborhoods are also starshaped in the following sense.

Proposition 3.14. *If $p \in M$ and U is a normal neighborhood of p, then for any $q \in U$ there exists a unique geodesic $\sigma : [0,1] \to U$ from p to q, called the **radial geodesic** from p to q. Furthermore, $\sigma'(0) = \exp_p^{-1}(q)$*

Proof. Fix $q \in U$ and let $\widetilde{U} \subset T_pM$ be a starshaped neighborhood of 0 such that $\exp_p|_{\widetilde{U}}$ is a diffeomorphism onto U. Let $v = \exp_p^{-1}(q)$ and let $\rho(t) = tv$ for $t \in [0,1]$. Since $\widetilde{U}$ is starshaped, ρ maps into $\widetilde{U}$, so $\sigma = \exp_p \circ \rho$ is a geodesic lying in U from p to q. As in Proposition 3.13, its initial velocity is:

$$\sigma'(0) = d(\exp_p \circ \rho)_0(0) = \gamma_v'(0) = v.$$

We show that σ is unique. Suppose $\tau : [0,1] \to U$ is a geodesic from p to q and let $w = \tau'(0)$. Then $w \in \mathscr{D}_p$ and $\tau(t) = \exp_p(tw)$ since these geodesics have the same initial velocities. We want to show that $v = w$, so that $\tau = \sigma$ by uniqueness of geodesics. If $tw \in \widetilde{U}$ for all $t \in [0,1]$, then

$$\exp_p(w) = \tau(1) = q = \sigma(1) = \exp_p(v).$$

Since $\exp_p$ is injective on $\widetilde{U}$ this implies that $v = w$ as desired.

We now show that $tw \in \widetilde{U}$ for all $t \in [0,1]$. Let $t_0 = \sup\{t \in [0,1] : tw \in \widetilde{U}\}$. Since $\widetilde{U}$ is starshaped $tw \in \widetilde{U}$ for all $t < t_0$. If $t_0 < 1$, then because $\widetilde{U}$ is open we know $t_0 w \notin \widetilde{U}$, otherwise we could extend the ray tw slightly past $t_0 w$. Let (t_n) be a sequence such that $t_n \to t_0$ and $0 \le t_n < t_0$ for each n. Since τ is continuous, $\tau(t_n) \to \tau(t_0)$. Since $\exp_p^{-1} : U \to \widetilde{U}$ is continuous and $\tau(t_n), \tau(t_0) \in U$, we know

$$t_n w = \exp_p^{-1}(\tau(t_n)) \to \exp_p^{-1}(\tau(t_0)) \in \widetilde{U}.$$

But $t_n w \to t_0 w$, so $t_0 w = \exp_p^{-1}(\tau(t_0)) \in \widetilde{U}$. Since $t_0 w \in \widetilde{U}$ we know t_0 is not less than 1, so $t_0 = 1$. This concludes the proof. ☐

The exponential map can be used to construct a special coordinate system on a normal neighborhood U of p. Fix an orthonormal basis (e_i) for T_pM (orthonormal with respect to g_p). Let $E : \mathbb{R}^n \to T_pM$ be the isomorphism $E(x^1, \ldots, x^n) = x^i e_i$. Then $\varphi = E^{-1} \circ \exp_p^{-1}$ is a diffeomorphism on U and therefore (U, φ) is a coordinate chart on M. Such charts are called ***normal coordinates***. The components x^i of φ map a point $q \in U$ to the coordinates of the vector $v = \exp_p^{-1}(q)$ in the (e_i) basis. In other words, if (f^i) is the dual basis on T_p^*M, then $x^i = f^i \circ \exp_p^{-1}$. The next proposition shows that normal coordinates make the Riemannian metric and Christoffel symbols simple at p.

Proposition 3.15. *Let (x^i) be a normal coordinate system at p. Then*

$$g_{ij}(p) = \delta_{ij} \qquad \text{and} \qquad \Gamma_{ij}^k(p) = 0.$$

Proof. Let $v \in T_pM$ and write $v = v^i e_i$. Then:

$$x^i(\gamma_v(t)) = f^i(\exp_p^{-1}(\gamma_v(t))) = f^i(tv) = tv^i.$$

Then for the initial velocity we have:

$$v = \gamma_v'(0) = \frac{d(x^i \circ \gamma_v)}{dt}(0)\partial_i|_p = v^i \partial_i|_p.$$

For fixed index i, setting $v^j = \delta_{ij}$ proves that $e_i = \partial_i|_p$ for all i. Therefore:

$$g_{ij}(p) = \langle \partial_i|_p, \partial_j|_p \rangle = \langle e_i, e_j \rangle = \delta_{ij}.$$

Notice that this is why we chose (e^i) to be an orthonormal basis.

To see that $\Gamma^k_{ij}(p) = 0$, note that Eq. (2) for γ_v at $t = 0$ reduces to $\Gamma^k_{ij}(p)v^i v^j = 0$ for each k. Since we can choose v arbitrarily, this implies that the quadratic form with matrix representation $(\Gamma^k_{ij}(p))$ is zero. The associated bilinear form is then also identically zero, and since $(\Gamma^k_{ij}(p))$ is symmetric, the associated bilinear form has the same matrix. Therefore $\Gamma^k_{ij}(p) = 0$ for all i, j, k. □

3.4. **Convex Sets.** In this section we construct another useful type of neighborhood in M. A ***convex*** set is an open set that is a normal neighborhood of each of its points. Our goal is to prove that there is a convex neighborhood of every $p \in M$. We require the following lemma.

Lemma 3.16. *Define $E : \mathscr{D} \to M \times M$ by $E(v) = (\pi(v), \exp(v))$. For each $p \in M$ there exists a neighborhood U of $0_p \in TM$ such that E is a diffeomorphism on U.*

Proof. By our discussion in the previous section we know we can choose coordinates (V, φ) at p such that exp, and therefore E, are smooth on the corresponding coordinates $(\widetilde{V}, \widetilde{\varphi})$. We show that dE is nonsingular at 0_p. Suppose $v \in T_{0_p}(TM)$ such that $dE(v) = 0$; we want to show that $v = 0$. Let $\pi : TM \to M$ be the usual projection and $\pi_1 : M \times M \to M$ be the projection onto the first component. Then $\pi = \pi_1 \circ E$, so $d\pi(v) = d\pi_1(dE(v)) = 0$.

Let us consider what this means in terms of our chosen coordinates. In terms of these coordinates π has the coordinates expression:

$$\varphi \circ \pi \circ \widetilde{\varphi}^{-1}(x^1, \ldots, x^n, v^1, \ldots, v^n) = (x^1, \ldots, x^n),$$

where n is the dimension of the manifold M. Therefore the matrix of the differential takes the form:

$$d\pi = \left(\begin{array}{c|c} I & 0 \end{array} \right)$$

where I and 0 represent the $n \times n$ identity and zero matrix respectively. Thus $d\pi(v) = 0$ implies v is of the form $v = (0, \ldots, 0, v^1, \ldots, v^n)$ in local coordinates.

Now let $\widehat{\varphi}$ be the induced coordinate chart on T_pM and consider the inclusion map $i_p : T_pM \to TM$. Its coordinate representation is

$$\widetilde{\varphi} \circ i_p \circ \widehat{\varphi}^{-1}(w^1, \ldots, w^n) = (\varphi(p), w^1, \ldots, w^n).$$

Thus di_p has the form:

$$di_p = \left(\frac{0}{I} \right).$$

We see that v is in the image of di_p, so let $w \in T_0(T_pM)$ such that $di_p(w) = v$.

Finally, with $\pi_2 : M \times M \to M$ as the second component projection, we have:

$$d(\pi_2 \circ E \circ i_p)(w) = d\pi_2(dE(di_p(w))) = d\pi_2(dE(v)) = 0.$$

But $\pi_2 \circ E \circ i_p = \exp_p$, and we know $d\exp_p$ is nonsingular at $0 \in T_pM$. This implies that $w = 0$, so $v = 0$ by linearity of di_p. Therefore dE is nonsingular at 0_p by Proposition 3.13, so the result follows by the inverse function theorem on manifolds. □

Proposition 3.17. *Every $p \in M$ is contained in a convex neighborhood.*

Proof. Let V be a normal neighborhood of p and let $\varphi = (x^i)$ be the corresponding normal coordinates. Consider the symmetric covariant 2-tensor field on V:

$$B = \left(\delta_{ij} - \sum_k \Gamma_{ij}^k x^k\right) dx^i \otimes dx^j.$$

Since $\Gamma_{ij}^k(p) = 0$, $B_p = g_p$ on T_pM by Proposition 3.15. In particular, B_p is positive definite, so by continuity B is positive definite on a neighborhood of p contained in V. Relabel V to be this neighborhood, so that B is positive definite on V.

By Lemma 3.16 there exists a neighborhood W' of $0_p \in TM$ such that E maps W' diffeomorphically onto a subset of $V \times V$. We may choose W' to be a coordinate ball about 0_p by restricting E if necessary. Then for each $q \in M$ we know $W' \cap T_qM$ is a coordinate ball in T_qM, and is therefore starshaped.

Let $N = \sum_i (x^i)^2$ and let $V(\delta) = \{q \in V : N(q) < \delta\}$ for $\delta > 0$. The image of $V(\delta)$ under φ is the set $B_\delta(0) \cap \varphi(V)$, so for δ sufficiently small $V(\delta)$ is diffeomorphic to $B_\delta(0)$. By shrinking δ further we achieve $V(\delta) \times V(\delta) \subset E(W')$, so E becomes a diffeomorphism of a set $W \subset W'$ onto $V(\delta) \times V(\delta)$. We claim this set $U = V(\delta)$ is the desired convex set.

Fix $q \in U$; we show that U is a normal neighborhood of q. Let $W_q = W \cap T_qM$. Since E is a diffeomorphism of W onto $U \times U$, we know $\exp_q$ is a diffeomorphism of W_q onto U. We must show that W_q is starshaped about 0. Pick $v \in W_q$ and let $x = \exp_q(v) \in U$. Then $\sigma(t) = \exp_q(tv)$ defined for $t \in [0,1]$ is a geodesic from q to x. If $\sigma([0,1]) \in U$ then $\exp_q^{-1}(\sigma(t)) = tv \in W_q$ for all $t \in [0,1]$, so W_q would be starshaped.

Suppose σ leaves U. Since $W \subset W'$, we know $v \in W' \cap T_qM$, which is starshaped. Since $\exp_q$ maps $W' \cap T_qM$ into V, we know σ never leaves V. Now $N \circ \sigma$ is a continuous real valued function on a compact set, so it attains a maximum value at some t_0. Since σ leaves $U = V(\delta)$, we know $N(\sigma(t_0)) \geq \delta$. Since $q, x \in U$ we know $N(q), N(x) < \delta$, so $0 < t_0 < 1$. We compute:

$$\frac{d^2(N \circ \sigma)}{dt^2} = 2 \sum_i \left\{ \left(\frac{d\sigma^i}{dt}\right)^2 + \sigma^i \frac{d^2\sigma^i}{dt^2} \right\}$$

Since σ is a geodesic, we can apply Eq. (2) to get rid of the second derivative:

$$\begin{aligned} \frac{d^2(N \circ \sigma)}{dt^2} &= 2 \sum_k \left\{ \left(\frac{d\sigma^k}{dt}\right)^2 - \sigma^k \frac{d\sigma^i}{dt}\frac{d\sigma^j}{dt}\Gamma_{ij}^k \right\} \\ &= 2\left\{ \delta_{ij} - \sum_k \Gamma_{ij}^k \sigma^k \right\} \frac{d\sigma^i}{dt}\frac{d\sigma^j}{dt}. \end{aligned}$$

Substituting in B yields:

$$\frac{d^2(N \circ \sigma)}{dt^2}(t_0) = 2B(\sigma'(t_0), \sigma'(t_0)).$$

Since $\sigma(t_0) \in V$ and B is positive definite on V, this implies that $(N \circ \sigma)''(t_0) > 0$, contradicting that $N(\sigma(t_0))$ is a maximum. □

We say a geodesic $\gamma : [a,b) \to M$ ***continuously extendible*** if it has a continuous extension to a curve defined on $[a,b]$ and say it is ***geodesically extendible*** if it has an extension to a geodesic defined on $[a,c)$ where $c > b$. The following corollary to Proposition 3.17 unites these two notions.

Corollary 3.18. *A geodesic* $\gamma : [0, b) \to M$ *is geodesically extendible if and only if it is continuously extendible.*

Proof. If γ is geodesically extendible, then it is obviously continuously extendible. Suppose γ is continuously extendible to $\tilde{\gamma} : [0, b] \to M$. Let U be a convex neighborhood of $\tilde{\gamma}(b)$. By continuity of γ, there exists $0 \leq a < b$ such that $\tilde{\gamma}([a, b]) \subset U$. Denote $p = \gamma(a)$ and $v = \exp_p^{-1}(\tilde{\gamma}(b))$, which exists since U is a normal neighborhood of p. Since $\exp_p^{-1}(U)$ is open, there exists $t_0 > 1$ such that $t_0 v \in \exp_p^{-1}(U)$. By uniqueness in Proposition 3.14, $\gamma|_{[a,b)}$ is equal to the geodesic $t \mapsto \exp_p((t - a)v/(b - a))$ with t restricted to $[a, b)$. But this latter geodesic is defined at least on $[a, c)$ where $c = t_0(b - a) + a > b$, so adjoining it to $\gamma|_{[0,a]}$ yields a geodesic extension of γ, as desired. □

4. Length, Distance, and Completeness

In this section we continue to develop the Riemannian analogues to properties of lines in Euclidean space. Having developed the notions of a geodesic and the acceleration of a curve, we now consider the length of a curve and the distance between two points. This distance will turn out to be a metric on M. In the final section we'll prove the Hopf-Rinow Theorem to unite two notions of completeness on M. This will have the interesting corollary that any two points in a connected and complete Riemannian manifold can be joined by a length minimizing geodesic.

4.1. **Arc Length.**

Unlike the acceleration of a curve, the definition of the length of a curve is thankfully easy to state.

Definition 4.1. Let $\alpha : [a, b] \to M$ be a smooth curve. We define the ***length*** of α to be the integral:

$$L(\alpha) = \int_a^b |\alpha'(t)| \, dt,$$

where $|\alpha'(t)| = \langle \alpha'(t), \alpha'(t) \rangle^{1/2}$. This integral is always defined by continuity of the velocity field α' and norm $|\cdot|$ on TM.

The length of a curve typically does not depend on its parameterization. If $\alpha : I \to M$ is a smooth curve, a reparameterization of α is a curve $\beta = \alpha \circ h$ where $h : J \to I$ is a smooth function between intervals. If h' does not change sign, then we say the reparameterization is ***monotone***.

Proposition 4.2. *Let* $\alpha : [a, b] \to M$ *be a smooth curve. Monotone reparameterizations of* α *do not change its length. Furthermore, if* $|\alpha'| > 0$ *then there exists a strictly increasing reparameterization* β *of* α *such that* $|\beta'| \equiv 1$.

We omit the proof; it is the same as the proof for the analogous statement for curves in $\mathbb{R}^n$.

In a normal neighborhood U of $p \in M$, the lengths of radial geodesics (from Proposition 3.14) have a few special properties. We define the ***radius function*** on U to be $r(q) = \left|\exp_p^{-1}(q)\right|$ for $q \in U$. Then we have the following proposition:

Proposition 4.3. *Let* U *be a normal neighborhood of* $p \in M$. *If* $q \in U$ *and* $\sigma : [0, 1] \to U$ *is the radial geodesic from* p *to* q, *then* $r(q) = L(\sigma)$.

Proof. Since σ is a geodesic, $|\sigma'|$ is constant along its trajectory. By Proposition 3.14 we know $\sigma'(0) = \exp_p^{-1}(q)$. Therefore:

$$L(\sigma) = \int_0^1 |\sigma'(t)|\, dt = \int_0^1 |\sigma'(0)|\, dt = \left|\exp_p^{-1}(q)\right| = r(q) \qquad \square$$

With the notion of arc length Proposition 3.14 can be strengthened as follows.

Proposition 4.4. *If U is a normal neighborhood of $p \in M$ and $q \in U$, then the radial geodesic $\sigma : [0,1] \to U$ from p to q is the shortest curve in U from p to q, and is unique up to monotone reparameterization.*

The proof of this fact requires more development than we have room for. See Lemma 5.1, Corollary 5.3, and Lemma 5.14 in [3] for the necessary development and a proof. We will modify this proposition one more time on our way to the Hopf-Rinow theorem, but first we need the Riemannian distance.

4.2. **Riemannian Distance.** In this section we give our Riemannian manifold M a metric analogous to the standard metric on $\mathbb{R}^n$. Once distance is defined, we refine Proposition 4.4 to the form that is used in proving the Hopf-Rinow theorem.

Definition 4.5. Let M be connected. The ***Riemannian distance*** $d : M \times M \to \mathbb{R}$ between two points $p, q \in M$ is the infimum of all lengths of piecewise smooth curves connecting p and q, i.e.

$$d(p,q) = \inf\{L(\alpha) : \alpha \in \Omega(p,q)\}.$$

where $\Omega(p,q)$ is the set of all piecewise smooth curves connecting p and q. We denote the ε-ball about p as $B_\varepsilon(p) = \{q \in M : d(p,q) < \varepsilon\}$.

Since a disconnected Riemannian manifold need not admit any curves between two given points, the Riemannian distance may fail to be defined. Therefore for the remainder of the paper we assume M to be connected.

Proposition 4.6. *The Riemannian distance is a metric on M that induces its standard topology.*

Proof. Symmetry and nonnegativity are obvious. To prove the triangle inequality, fix $\varepsilon > 0$, let α be a curve from p to q, and let β be a curve from q to r such that

$$L(\alpha) < d(p,q) + \varepsilon \qquad\qquad L(\beta) < d(q,r) + \varepsilon.$$

Adjoining β to the end of α yields a curve γ from p to r with length

$$d(p,r) \le L(\gamma) = L(\alpha) + L(\beta) < d(p,q) + d(q,r) + 2\varepsilon.$$

Since this holds for all $\varepsilon > 0$, the triangle inequality follows.

To prove that $d(p,q) > 0$ for $p \neq q$ we must first show that the ε-balls $B_\varepsilon(p)$ are compatible with the open sets of M. We first show that if U is open and $p \in U$, then there exists $\varepsilon > 0$ such that $p \in B_\varepsilon(p) \subset U$. Let $V \subset U$ be a normal neighborhood of p. If ε is sufficiently small then the closure of $\widetilde{B} = \{v \in T_pM : |v| < \varepsilon\}$ is contained in $\exp_p^{-1}(V)$, so the closure of $B = \exp_p(\widetilde{B})$ is contained in V. Note that B is open and normal as it is the image of an open, convex set under $\exp_p|_V$. We claim $B = B_\varepsilon(p)$. If $q \in B$ and σ is the radial geodesic from p to q, then $L(\sigma) = \left|\exp_p^{-1}(q)\right| < \varepsilon$, so $B \subset B_\varepsilon(p)$.

To show the reverse inclusion, we suppose $q \in B_\varepsilon(p) \setminus B$ and derive a contradiction. Let $\alpha : [0,1] \to M$ be a piecewise smooth curve from p to q. Let

$t_0 = \inf\{t \in [0,1] : \alpha(t) \notin B\}$. Since V contains the closure of B, we know $r \circ \alpha$ is defined on t_0, and since B is open we know $\alpha(t_0) \notin B$. Therefore $r(\alpha(t_0)) \geq \varepsilon$. Since $r(\alpha(t_0)) \leq L(\alpha|_{[0,t_0]})$, this implies that $L(\alpha) \geq \varepsilon$. Since α was arbitrary, this implies that $d(p,q) \geq \varepsilon$, contradicting that $q \in B_\varepsilon(p)$.

Next we show that each ε-ball is open. It suffices to show that for each $p \in M$ the ball $B_\varepsilon(p)$ is open for sufficiently small ε, for then the triangle inequality allows us to write any $B_\eta(p)$ as a union of open ε-balls. But as we saw above, for sufficiently small ε the ball $B_\varepsilon(p)$ is the image of the set $\widetilde{B}$ under $\exp_p$ on a normal neighborhood, implying that $B_\varepsilon(p)$ is open. Thus, d induces the topology on M.

Finally, we show that $d(p,q) > 0$ for $p \neq q$. Let U and V be disjoint open sets containing p and q respectively. By the above arguments there exists $\varepsilon > 0$ such that $p \in B_\varepsilon(p) \subset U$. Since $q \notin B_\varepsilon(p)$, this implies that $d(p,q) > \varepsilon$, as desired. □

The above proof yields the following refinement of Proposition 4.4.

Proposition 4.7. *If $p \in M$, then for sufficiently small $\varepsilon > 0$ the neighborhood $B_\varepsilon(p)$ is normal. In this case, the radial geodesic σ from p to $q \in B_\varepsilon(p)$ is the unique (up to monotone reparameterization) shortest curve in M from p to q, i.e.*

$$L(\sigma) = r(q) = d(p,q).$$

Proof. The proof of Proposition 4.6 shows that $B_\varepsilon(p)$ is normal for sufficiently small ε. Let $B_\varepsilon(p)$ be a normal neighborhood of p and let α be a piece-wise smooth curve from p to q. If α lies in $B_\varepsilon(p)$, then Proposition 4.4 implies that $L(\alpha) \geq L(\sigma)$. If α leaves $B_\varepsilon(p)$, then the third paragraph of the proof of Proposition 4.6 shows that $L(\alpha) \geq \varepsilon > L(\sigma)$, so the result follows. □

We are now ready to discuss completeness and prove the Hopf-Rinow Theorem.

4.3. **The Hopf-Rinow Theorem.** On a Riemannian manifold there are two notions of completeness. The first is completeness in the sense of metric spaces, using the Riemannian distance metric. On the other hand, a Riemannian manifold is called ***geodesically complete*** if every maximal geodesic is defined on the entire real line. If the Riemannian manifold is connected then these notions, and a few others, are equivalent.

Theorem 4.8 (Hopf-Rinow)**.** *Let M be a connected Riemannian manifold. The following are equivalent:*

- *I. M is complete as a metric space.*
- *II. M is geodesically complete.*
- *III. There exists a point $p \in M$ such that $\mathscr{D}_p = T_pM$.*
- *IV. A subset of M is compact if and only if it is closed and bounded.*

The proof requires two lemmas.

Lemma 4.9. *Let $\gamma_1 : [a,b] \to M$ be a geodesic from p to q, $\gamma_2 : [b,c] \to M$ a geodesic from q to r, and suppose γ_1 and γ_2 have the same speed. If the curve $\gamma : [a,c] \to M$ obtained by adjoining γ_1 and γ_2 has length $L(\gamma) = d(p,r)$, then γ is a geodesic.*

Proof. Let U be a convex neighborhood of q. Then there exists $d \in [a,b)$ and $e \in (b,c]$ such that $\gamma_1|_{[d,b]}$ and $\gamma_2|_{[b,e]}$ lie within U. The combination $\gamma|_{[d,e]}$ of these two curves must be length-minimizing, otherwise we would be able to connect p and

r by a shorter curve than γ. Since U is a normal neighborhood of $\gamma(d)$, Proposition 4.4 implies that $\gamma|_{[d,e]}$ is a monotone reparameterization of a radial geodesic. Since γ_1 and γ_2 are geodesics, the reparameterization must be piecewise linear. Since γ_1 and γ_2 have the same speed, the reparameterization must be entirely linear, so $\gamma|_{[d,e]}$ is a geodesic by Lemma 3.7. Since the only point where we could have had $\gamma'' \neq 0$ was at $\gamma(b)$ and $b \in [d, e]$, this implies that γ is a geodesic as well. □

Lemma 4.10. *If there exists $p \in M$ such that $\mathscr{D}_p = T_pM$, then for any $q \in M$ there is a minimizing geodesic segment from p to q.*

Proof. Let $B = B_\varepsilon(p)$ be a normal ε-ball about p and let $\widetilde{B}$ be the corresponding neighborhood of $0 \in T_pM$. If $q \in B$, then the result is trivial by Proposition 4.7, so suppose $q \notin B$. For $\delta < \varepsilon$, the set $S = \{x \in B : r(x) = \delta\}$ is the image of a sphere in $\widetilde{B}$ under $\exp_p$, and is therefore compact. The map $s \mapsto d(s, q)$ defined on S is continuous, and therefore attains a minimum at some $m \in S$.

Let $\gamma : [0, \infty) \to M$ be the unit speed reparameterization of the radial geodesic from p to m, extended to the interval $[0, \infty)$. Notate $d = d(p, q)$ and let T be the set of all $t \in [0, d]$ such that

$$t + d(\gamma(t), q) = d. \tag{3}$$

We want to show that $d \in T$, for the above equation then implies $d(\gamma(d), q) = 0$, so $\gamma(d) = q$. Since γ has unit speed $L(\gamma|_{[0,d]}) = d = d(p, q)$, which is the desired result.

The set T is nonempty since $0 \in T$, and is closed since $t \mapsto d(\gamma(t), q)$ is continuous. Therefore T contains a maximum $t_0 \leq d$. We suppose $t_0 < d$ and derive a contradiction. As before, let $B' = B_{\varepsilon'}(\gamma(t_0))$ be a normal neighborhood about $\gamma(t_0)$ with $\varepsilon' < d(\gamma(t_0), q)$ (so $q \notin B'$) and let $S' = \{x \in B' : \tilde{r}(x) = \delta'\}$ where $\delta' < \varepsilon'$ and $\tilde{r}$ is the radius function on B'. Let $m' \in S'$ be a minimum of the function $s \mapsto d(s, q)$ defined on S'. Note that Proposition 4.7 implies that $d(\gamma(t_0), m') = \tilde{r}(m') = \delta'$.

We claim

$$d(\gamma(t_0), q) = \delta' + d(m', q). \tag{4}$$

Let $\alpha : [0, b] \to M$ be a piecewise smooth curve from $\gamma(t_0)$ to q. Since α leaves B', we know there exists $a \in [0, b]$ such that $\alpha(a) \in S' \subset B'$. Then Proposition 4.7 implies

$$L(\alpha) = L(\alpha|_{[0,a]}) + L(\alpha|_{[a,b]}) \geq \delta' + d(m', q).$$

Therefore $d(\gamma(t_0), q) \geq \delta' + d(m', q)$. But $d(\gamma(t_0), q) \leq \delta' + d(m', q)$ by the triangle inequality, so our claim is proven. Since $t_0 \in T$, substituting (3) into (4) yields

$$t_0 + \delta' + d(m', q) = d. \tag{5}$$

Since $d = d(p, q)$, by the triangle inequality we have:

$$t_0 + \delta' + d(m', q) \leq d(p, m') + d(m', q),$$

so $t_0 + \delta' \leq d(p, m')$.

Let $\sigma : [t_0, t_0 + \delta'] \to B'$ be a unit speed geodesic from $\gamma(t_0)$ to m'. If η is the curve obtained by adjoining $\gamma|_{[0,t_0]}$ and σ, then η is a piecewise smooth curve from p to m' and $L(\eta) = t_0 + \delta' \leq d(p, m')$. Therefore η is length-minimizing, so by Lemma 4.9 η is a geodesic. This implies that γ and σ have the same velocity at t_0 and are therefore equal on their shared domain. Thus, Eq. (5) yields

$$t_0 + \delta' + d(\gamma(t_0 + \delta'), q) = t_0 + \delta' + d(m', q) = d,$$

contradicting that t_0 is the maximum of T. □

We can now prove the Hopf-Rinow theorem.

Proof of Theorem. I $\Rightarrow$ II. It suffices to show that a unit speed geodesic $\gamma : [0, b) \to M$ is geodesically extendible. Let (t_n) be a sequence in $[0, b)$ such that $t_n \to b$. Then $\gamma(t_n)$ is a Cauchy sequence since $d(\gamma(t_n), \gamma(t_m)) \leq |t_n - t_m|$, so it converges to a point p. If (s_n) is any other sequence in $[0, b)$ approaching b then $\gamma(s_n)$ converges to p since $d(\gamma(t_n), \gamma(s_n)) \leq |t_n - s_n|$. Therefore defining $\gamma(b) = p$ yields a continuous extension of γ, so γ is geodesically extendible by Corollary 3.18.

II $\Rightarrow$ III. For any $v \in T_pM$, the maximal geodesic γ_v is defined on $\mathbb{R}$. In particular, it is defined on 1, so $v \in \mathscr{D}_p$.

III $\Rightarrow$ IV. Since M is a metric space, any compact set is automatically closed and bounded. Conversely, let $A \subset M$ be closed and bounded. By the previous lemma, for each $q \in A$ there exists a minimizing geodesic segment $\sigma_q : [0, 1] \to M$ from p to q. Since A is bounded, the values $\left|\sigma_q'(0)\right| = L(\sigma_q) = d(p, q)$ are bounded by the triangle inequality, say $R \geq d(p, q)$ for each $q \in A$. Then each $\sigma_q'(0)$ is contained in the compact ball $B_R = \{v \in T_pM : |v| \leq R\}$. If $q \in A$, then $\exp_p(\sigma_q'(0)) = q$, so $A \subset \exp_p(B_R)$. Since $\exp_p(B_R)$ is compact and A is closed, this implies that A is compact.

IV $\Rightarrow$ I. Let (x_n) be a Cauchy sequence in M. The set $\{x_n\}$ is bounded, so its closure is compact. Therefore (x_n) has a convergent subsequence and since (x_n) is Cauchy, it must converge to the limit of the subsequence. □

To conclude, the Hopf-Rinow theorem has the following satisfying corollary, which follows immediately from geodesic completeness and Lemma 4.10.

Corollary 4.11. *If M is a connected and complete Riemannian manifold, then any two points in M may be connected by a length minimizing geodesic.*

Acknowledgments. I would like to thank my mentor, Reid Harris, for his assistance on this project. His reading assignments and problem sets were enormously helpful in guiding me through the material in an effective and timely manner. Thanks to him I've developed a much stronger foundation in differential geometry, gotten a glimpse into the rewarding subject of Riemannian geometry, and feel prepared to pursue new topics in differential geometry as well.

References

[1] Lee, John M. *Introduction to Smooth Manifolds*. 2nd ed. New York: Springer, 2013. Print.
[2] Lee, John M. *Riemannian Manifolds: An Introduction to Curvature*. New York: Springer, 1997. PDF e-book.
[3] O'Neill, Barrett. *Semi-Riemannian Geometry: With Applications to Relativity*. New York: Academic, 1983. Print.
[4] Weisstein, Eric W. "Euclid's Postulates." From *MathWorld*–A Wolfram Web Resource. http://mathworld.wolfram.com/EuclidsPostulates.html