# THE HOPF-RINOW THEOREM

TANVI DESHPANDE

## 1. Introduction

A Riemannian manifold is a special type of manifold equipped with a positive-definite inner product on its tangent space. With just this condition, however, one can achieve some striking results. In particular, the Hopf-Rinow theorem makes a connection between completeness as a metric space, geodesic completeness, and compact sets on the manifold. This paper, which will prove a number of results building up to the Hopf-Rinow theorem, is based loosely off of [1] and [2]. The figures in this paper are taken from [1].

## 2. Riemannian manifolds

**Definition 2.1** (Riemannian metric)**.** A Riemannian metric $g$ is a smooth positive-definite function (technically, a covariant tensor field) on a manifold $M$ such that for all $p \in M$ and $v, w \in T_pM$:

(1) $g_p(v, w) = g_p(w, v)$.
(2) $g_p(v, v) > 0$ for $v \neq 0$.

A manifold coupled with a Riemannian metric is referred to as a Riemannian manifold, the set of which is denoted by $M$. Because $g$ is a tensor, it is bilinear on $T_pM$, and therefore it is an inner product that we will hereafter denote as $\langle v, w \rangle$, and correspondingly, the norm of some $v \in T_pM$ is $|v| = \sqrt{\langle v, v \rangle}$.

A coordinate chart, or a map $\phi : U \to V$ where $U$ is an open set of $M$ and $V$ is an open set of $\mathbb{R}^n$, is denoted by $(U, x^i)$. In any given coordinate chart, the Riemannian metric is denoted by $g = g_{ij} dx^i \otimes dy^i$. Because $g$ is smooth, its component functions are smooth (a universally true fact of tensor fields), and because $g$ is positive definite (Definition 2.1.2), $(g_{ij})$ is an invertible matrix which is also smooth.

*Example.* A simple example of a Riemmanian manifold is $\mathbb{R}^n$. Its corresponding metric is just $g = \delta_{ij} dx^i \otimes dy^i$ where

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases},$$

which works out to the typical distance metric in $\mathbb{R}^n$.

In addition, we can use these basic definitions to express the *induced covariant derivative* of a smooth curve $\alpha$, which can be written as

$$Z' = \left( \frac{dZ^k}{dt} + Z^i \Gamma^k_{ij} (\alpha')^j \right) \partial_k|_\alpha. \tag{2.1}$$

*Date*: June 2022.

The full derivation for $Z'$ can be found in [2]; it is not worth including. Generally, however, we will not have to work with technical objects such as tensor fields and coordinate charts in this paper; most concepts from differential geometry carry over to Riemannian manifolds and will suffice for our purposes. With preliminary definitions out of the way, in the rest of this section, we proceed to prove uniqueness and existence for geodesics using two brief lemmas.

**Lemma 2.2.** *For each $v \in T_pM$, there is an open interval $I$ about $0$ and a unique geodesic $\gamma : I \to M$ such that $\gamma'(0) = v$.*

*Proof.* Setting $Z = \gamma'$ in Equation 2.1, we have

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

Taking the case where $\gamma' = 0$, we obtain a system of second-order ODEs, from which the result follows based on the existence and uniqueness theorems for ODEs. □

Next, we must prove some facts about the uniqueness and congruence of geodesics.

**Lemma 2.3.** *Given two geodesics $\alpha, \beta : I \to M$, if $\alpha'(a) = \beta'(a)$ for $a \in I$, then we must have $a \equiv b$.*

*Proof.* Suppose there is some $t_0 \in I$ such that $\alpha(t_0) \neq \beta(t_0)$. WLOG, suppose that $t_0 > a$. Let $b$ denote the greatest lower bound of the set $\{t \in I : t > a, \alpha(t) \neq \beta(t)\}$. We want to show that $\alpha'(b) = \beta'(b)$; assume that $b > a$. Since $\alpha = \beta$ on $(a, b)$, so do their velocities; because velocity of a smooth curve is continuous, we have $\alpha'(b) = \beta'(b)$.

Now, suppose $b \in I$. Because $t \mapsto \alpha(b + t)$, $t \mapsto \beta(b + t)$ are geodesics (and have equal velocity at $t = 0$), 2.2 says that they agree on some interval about $b$; however, this contradicts the definition of $b$ as a greatest lower bound. In the case that $b$ is a right endpoint of $I$, we have $b = t_0$; therefore, $\alpha'(b) = \beta'(b)$, so we must have $\alpha(t_0) \neq \beta(t_0)$, a contradiction. □

From the previous two lemmas, we can prove the following theorem about constructing a *maximal geodesic* for each $v$ in the tangent plane of a point.

**Theorem 2.4.** *For each $v \in T_pM$ there is a unique geodesic $\gamma_v : I_v \to M$ such that $\gamma_v' = v$ and for any other geodesic $\eta : I \to M$ with $\eta'(0) = v$, we have $I \subset I_v$ and $\eta = \gamma_v|I$.*

*Proof.* Suppose $\mathcal{G}$ is the set of all geodesics with initial velocity $v$. By Lemma 2.3, $\eta_1, \eta_2$ agree on the intersection of their domains. If $I_v$ is the union of the domain of all geodesics in $\mathcal{G}$, there exists a well-defined curve $\gamma_v(t)$ such that $\gamma_v(t) = \eta_i(t)$ for all $\eta \in \mathcal{G}$. $\gamma_v$ is a geodesic with initial velocity $v$ because the differential and induced covariant derivative act locally. □

The maximal geodesic and its domain are denoted by $\gamma_v$ and $I_v$ respectively.

Next, we will introduce the exponential map, which relates the tangent plane of $p$ to the manifold $M$ itself.

## 2.1. **The Exponential Map.**

**Definition 2.5** (The Exponential Map). Let $\mathscr{D} = \{v \in TM : 1 \in I_v\}$ and $\mathscr{D}_p = \mathscr{D} \cap T_pM$ for $p \in M$. Then the exponential map $\exp : \mathscr{D} \to M$ is defined by $\exp(v) = \gamma_v(1)$.

The exponential map is well-defined by Theorem 2.4. $\mathscr{D}$ and $\mathscr{D}_p$ are the largest subsets of $TM$ on which exp and $\exp_p$ are defined.

**Lemma 2.6.** *Let $E : D \to M \times M$ be defined as $E(v) = (\pi(v), exp(v))$. Then, for each $p \in M$, there exists a neighborhood $U$ of $0_p \in TM$ such that $E$ is a diffeomorphism on $U$.*

We will not prove the lemma, but the proof can be found in [2].

**Remark 2.7.** Note that the exponential map locally carries straight lines on $T_pM$ on geodesics through $p$. See Figure 1 for a diagram demonstrating this.

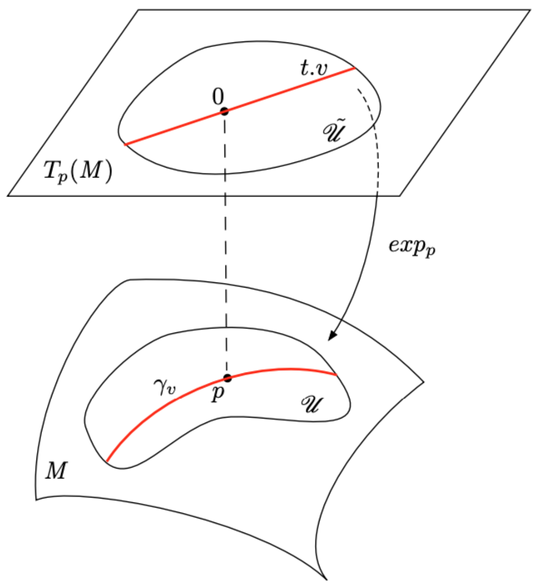


**Figure 1.** Exponential function

The next section introduce an important concept called the radial geodesic, which relates the exponential map to other notions of completeness found later in the paper; in particular, we will show that geodesics on manifolds are analogous to the notion of straight lines in vector spaces.

We will also need the concept of a normal neighborhood:

**Definition 2.8.** A neighborhood $\mathscr{U}$ of $p \in M$ is called normal if $\tilde{\mathscr{U}}$, a neighborhood of $0 \in T_pM$ such that $\exp_{p|\tilde{\mathscr{U}}}$ is a diffeomorphism, is starshaped.

## 2.2. **Radial Geodesics.**

**Proposition 2.9.** *Given $p \in M$ and $U$ a normal neighborhood of $p$, there exists a unique geodesic $\sigma : [0,1] \to U$ connecting $p$ with each $q \in U$, where $\sigma'(0) = exp_p^{-1}(q)$.*

*Proof.* Define $v = \exp_p^{-1}(q) \in \tilde{U} \subset t_p M$ and $\rho(t) = tv$ for $t \in [0,1]$ and $v \in U$. Then, $\sigma = \exp_p \circ \rho$ is a geodesic from $p$ to $q$ in $U$; this proves existence. Also, because $\exp_p \circ \rho(t) = \exp_p(tv) = \gamma_v(t)$, $\sigma'(0) = v$, as claimed.

For uniqueness, let $\tau$ be some other geodesic connecting $p$ and $q$ in $U$, 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, and we must show that $w = v$, which will imply that $\tau = \sigma$.

First, we will show that $w \in \tilde{U}$ (if it is not, then $t \mapsto tw$ will leave $\tilde{U}$). So, there is some $t_0 \in [0,1]$ for which $\exp_p(t_0 w) \in U \setminus \tau([0,1])$; however, this contradicts the definition of $\tau$.

Since $w \in \tilde{U}$, we have

$$\exp_p(w) = \tau(1) = q = \exp_p(v),$$

and because the exponential map is injective, $w = v$, completing the proof. □

**Remark 2.10.** A subset of a Riemannian manifold $M$ is convex if it is a normal neighborhood of each of its points. For any points $p, q$ in a convex set, there exists a unique geodesic $\eta$ defined on $[0,1]$ joining them.

*Example.* A simple example of a convex set is $\mathbb{R}^n$; $\mathbb{R}^n \setminus 0$, on the other hand, is not convex: there is no geodesic connecting $(1, 0, \ldots, 0)$ to $(0, \ldots, 0, -1)$.

*Example.* $\mathbb{S}^2$ is not convex: there are infinitely many great circles (which are geodesics) joining any two points on the sphere.

## 3. Arc Length, Distance, and Completeness

This section relates arc length and distance to radial geodesics and introduces the radius function.

### 3.1. Arc Length.

**Definition 3.1.** Given a smooth curve $\alpha : [a, b] \to M$, its length is

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

which should not come as a surprise; this integral is always defined because $\alpha'$ and the norm are both continuous on $TM$.

Similarly to what has been discussed in class, length is also independent of monotone reparametrization, and unit-speed reparametrizations exist for curves on $M$.

### 3.2. Radius Function.

In particular, the radial geodesics have special properties relating to length in a normal neighborhood $U$ of $p \in M$.

**Definition 3.2.** The radius function on $U$ is defined as $r(q) = |\exp_p^{-1}(q)|$ for $q \in U$.

See Figure 2 for a visual representation of what the radius function looks like; it is a sort of sphere on the manifold, and this diagram represents $r = \delta$ on the sphere. This definition of the radius function also leads us to the following proposition.

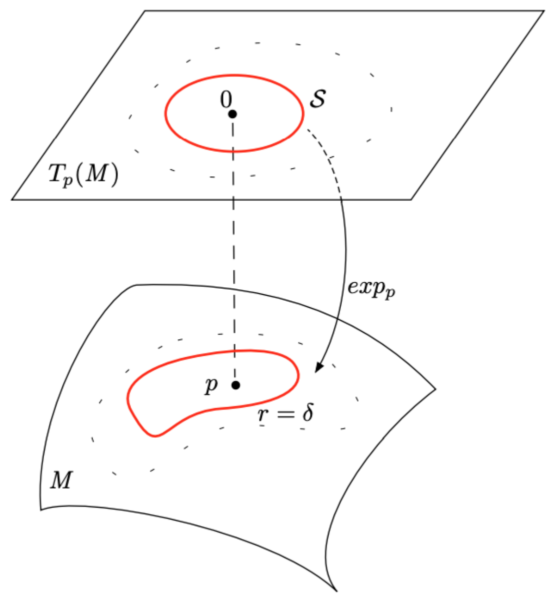


**Figure 2.** Radius function

**Proposition 3.3.** *Suppose $U$ is a normal neighborhood of $p \in M$. If $q \in U$ and $\sigma : [0,1] \to U$ is the radial geodesic from $p$ to $q$, $r(q) = L(\sigma)$.*

*Proof.* Because $\sigma$ is a geodesic, its speed is constant. Furthermore, $\sigma'(0) = \exp_p^{-1}(q)$, so

$$L(\alpha) = \int_0^1 |\sigma'(t)|dt = \int_0^1 |\sigma'(0)|dt = \int_0^1 |\exp_p^{-1}(q)|dt = r(q).$$

□

We will also state, but not prove, the following necessary proposition:

**Proposition 3.4.** *For a normal neighborhood $U$ of $p \in M$, the radial geodesic $\sigma : [0,1] \to U$ from $p$ to $q$ is distance-minimizing and unique up to monotone reparametrization.*

Next, we will define Riemannian distance, which establishes the Riemannian manifold as a metric space.

### 3.3. **Riemannian Distance.**

**Definition 3.5.** Suppose $M$ is a connected Riemannian manifold. Then, the *Riemannian distance* between two points $p, q$ on $M$ is a function $d : M \times M \to \mathbb{R}$, defined as

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

where $\Omega(p,q)$ is the set of all piecewise smooth curves connecting $p$ and $q$. Hereafter, we will assume that $M$ is connected; Riemannian distance may not be defined for two given points on a disconnected manifold.

We can also define an open ball around $p$:

$$B_\varepsilon(p) = \{q \in M : d(p,q) < \varepsilon\}.$$

(a) Front View. (b) Back View.

**Figure 3.** The largest $\varepsilon$-neighborhood of a point on the cylinder

**Proposition 3.6.** *This open ball is normal for sufficiently small $\varepsilon$; in this case, we refer to it as an $\varepsilon$-neighborhood.*

*Proof.* Suppose that $U$ is a neighborhood of $p \in M$ and $\tilde{U} \subset T_pM$, so that $\exp_p : \tilde{U} \to U$ is a diffeomorphism. Then, for small enough $\varepsilon$, the open ball of radius $\varepsilon$ of $T_pM$ is starshaped and contained in $\tilde{U}$. Therefore, $N = \exp_p(B_e)$ is a normal neighborhood: for each $x \ inN$, there exists $v \in T_pM$ such that $|v| < \varepsilon$ and $x = \exp_p(v)$. Then, $r(x) < \varepsilon$ so $x \in \{x \in M | r(x) < \varepsilon\}$.

Then, by Proposition 3.4, $r(x) = d(p, x)$ for any $x \in N$, meaning $N_\varepsilon$ is a normal neighborhood. □

As a corollary, each point has a normal neighborhood, given by the $\varepsilon$-neighborhood with suitably small $\varepsilon$.

*Example.* The largest $\varepsilon$-neighborhood of a point $p$ on a cylinder is shown in Figure 3; it can be thought of as a circle centered at $p$ wrapped around the cylinder.

We will also consider *minimizing curves*, which, intuitively, are curves that realize the lower bound of the distance.

**Proposition 3.7.** *Given an $\varepsilon$-neighborhood $B$ of a point $p \in M$, the radial geodesic from $p$ to a point $q \in B$ is the unique shortest path in $M$ connecting the two:*

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

The significance of this proposition is that $\varepsilon$-neighborhoods are stronger than normal neighborhoods, because they admit unique minimizing curves between two points, and because no other minimal curves lie anywhere else in the manifold.

*Proof.* From Proposition 3.4, we have that $\sigma_{pq}$, the radial geodesic connecting $p$ and $q$, is the unique shortest curve connecting them in $B$. Also, $L(\sigma_{pq}) = r(q) < \varepsilon$ (because $q$ is in the $\varepsilon$-neighborhood of $p$).

Next, consider a curve $\alpha$ connecting $p$ and $q$ that leaves $B$. If $\alpha$ leaves $B$, then for all $\rho < \varepsilon$, there is some $t_\rho$ such that

$$r(\alpha(t_\rho)) = \rho.$$

Because $\alpha(t_\rho) \in B$, we have

$$L(\alpha) > L(\alpha_\rho) = r(\alpha(t_\rho)) = \rho,$$

where $\alpha_\rho$ is the shortest path connecting $p$ and $\alpha(t_\rho)$. As this is true for all $\rho < \varepsilon$, $L(\alpha) \geq \varepsilon$, so if $\alpha$ leaves $B$, it must not be minimal. □

The next task is to show that Riemannian distance together with a manifold $M$ forms a metric space.

**Proposition 3.8.** *A manifold $M$, coupled with the Riemannian distance, is a metric space $(M, d)$.*

*Proof.* We must show that for $p, q, r \in M$:

(1) $d(p, q) \geq 0$
(2) $d(p, q) = 0 \iff p = q$
(3) $d(p, q) = d(q, p)$
(4) $d(p, r) \leq d(p, q) + d(q, r)$

The first 2 conditions are obvious from the definition of the Riemannian metric. The third condition, symmetry, is also obvious since any piecewise smooth curve $\alpha \in \Omega(p, q)$ can be reversed to obtain an identical curve in $\Omega(q, p)$.

Lastly, we must show that the triangle inequality is upheld. By the definition of distance, for $\varepsilon > 0$, we have $\alpha \in \Omega(p, q)$ and $\beta \in \Omega(q, r)$ such that

$$L(\alpha) \leq d(p, q) + \varepsilon$$

and

$$L(\beta) \leq d(q, r) + \varepsilon.$$

Joining these two curves to form a curve $\gamma \in \Omega(p, r)$ yields

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

which satisfies the triangle inequality as $\varepsilon \to 0$. □

**Proposition 3.9.** *Every point admits a convex neighborhood.*

*Proof.* The proof of this fact is long, technical, and out of the scope of this paper, so we omit it. □

**Definition 3.10.** A geodesic $\gamma : [a, b) \to M$ is continuously extendible if it has a continuous extension to a curve on $[a, b]$; it is geodesically extendible if it has an extension to a geodesic defined on $[a, c)$ with $c > b$.

The following corollary to Proposition 3.9 unites the two notions.

**Corollary 3.11.** *A geodesic $\gamma : [0, b) \to M$ is geodesically extendible iff it is continuously extendible.*

*Proof.* If $\gamma$ is geodesically extendible, it is obviously continuously extendible by definition. For the other direction, suppose that $\gamma$ is continuously extendible to a function $\tilde{\gamma} : [0, b] \to M$. Let $U$ be a convex neighborhood (according to Proposition 3.9, each point admits one). Because $\gamma$ is continuous, there exists $0 \leq a < b$ such that $\tilde{\gamma}([a, b]) \subset U$.

Now, let $p = \gamma(a)$ and $v = \exp_p^{-1}(\tilde{\gamma}(b))$ (this exists because $U$ is a normal neighborhood of $p$). Because $\exp_p^{-1}(U)$ is open, there exists $t_0 > 1$ such that $t_0 v \in \exp_p^{-1}(U)$. Since the radial geodesic is unique, $\gamma|[a, b)$ is equal to the geodesic $\eta$ mapping $t$ to $\exp_p((t-a)v(b-a))$, restricting

$t$ to $[a, b)$. However, this geodesic is defined on the interval $[a, c)$ where $c = t_0(b-a)+a > b$; therefore, adjoining it to $\gamma|_{[0,a]}$ is a geodesic extension of $\gamma$. □

## 4. The Hopf-Rinow Theorem

Riemannian manifolds have two types of completeness. On one hand, they can be complete in the sense of metric spaces.

**Definition 4.1.** A metric space $M$ is complete if every Cauchy sequence of points in $M$ is convergent; in the case of Riemmanian manifolds, we use the Riemannian distance metric.

Alternatively, a Riemannian manifold can be *geodesically complete.*

**Definition 4.2.** A Riemannian manifold is geodesically complete if every maximal geodesic is defined on all of $\mathbb{R}$.

The Hopf-Rinow theorem, which is the main result of this paper, unites these two notions of completeness.

**Theorem 4.3.** *If $M$ is a connected Riemannian manifold, then the following statements are equivalent:*

1. *$M$ is a complete metric space.*
2. *$M$ is geodesically complete.*
3. *There exists a point $p \in M$ such that $\mathscr{D}_p = T_pM$.*
4. *Closed and bounded subsets of $M$ are compact.*

We present two lemmas required in the proof of the theorem, and then prove the theorem.

**Lemma 4.4.** *Suppose $\gamma_1 : [a, b] \to M$ is a geodesic from $p$ to $q$ and $\gamma_2 : [b, c] \to M$ is a geodesic from $q$ to $r$, where both $\gamma_1$ and $\gamma_2$ have the same speed. If the curve $\gamma : [a, c] \to M$ obtained by adjoining $\gamma_1$ and $\gamma_2$ has length $d(p, r)$, then $\gamma$ is a geodesic.*

*Proof.* Suppose $U$ is a convex neighborhood of $q$. Then, there is some $d \in [a, b)$ and some $e \in (b, c]$ such that $\gamma_1|_{[d,b]}$ and $\gamma_2|_{[b,e]}$ are both in $U$. Now, adjoin these two curves together to get a curve $\gamma$ from $d$ to $e$. $\gamma|_{[d,e]}$ must be length-minimizing, or else $p$ and $r$ could be connected by a shorter curve, a contradiction.

Because $U$ is a normal neighborhood of $\gamma(d)$, Proposition 3.4 means that $\gamma|_{[d,e]}$ is a monotone reparametrization of a radial geodesic. $\gamma_1$ and $\gamma_2$ are both geodesics and have the same speed, so this reparametrization must be linear and $\gamma|_{[d,e]}$ is a geodesic. We also know that $\gamma'' = 0$ except for possibly at $\gamma(b)$; however, since $b \in [d, e]$, and $\gamma|_{[d,e]}$ is a geodesic, $\gamma$ is a geodesic. □

**Lemma 4.5.** *If there exists a $p \in M$ such that $\mathscr{D}_p = T_pM$, then for any $q \in M$ there is a minimal geodesic segment connecting $p$ and $q$.*

*Proof.* The proof uses the previous lemma (which is why it was included), but it is long and technical, so we leave it to [2]. □

*Proof.* (1) → (2). We may show this by proving that a unit-speed geodesic $\gamma : [0, b) \to M$ is geodesically extendable, from Corollary 3.11. Suppose $(t_n)$ is a sequence in $[0, b)$ converging to $b$. Then, $\gamma(t_n)$ is a Cauchy sequence converging to some point $p$: we have $d(\gamma(t_n), \gamma(t_m)) \leq |t_n - t_m|$.

Also, for any other sequence $(s_n)$ in the same interval approaching $b$, $\gamma(s_n)$ also converges to $p$ since $d(\gamma(t_n), \gamma(s_n)) \leq |t_n - s_n|$. Therefore, setting $\gamma(b) = p$ yields a continuous extension of $\gamma$; by Corollary 3.10, $\gamma$ is also geodesically extendable.

(2) $\rightarrow$ (3). By the definition of geodesic completeness, $\gamma_v$ is defined on $\mathbb{R}$ for each $v \in T_pM$; in particular, it is defined on 1, so $v \in \mathscr{D}_p$.

(3) $\rightarrow$ (4). $M$ is a metric space, so any compact set is immediately closed and bounded. For the other direction, letting $A \subset M$ be closed and bounded, we have by Lemma 4.5 that for each $q \in A$, there is a minimizing geodesic segment $\sigma_q : [0,1] \rightarrow M$ connecting $p$ and $q$.

Because $A$ is bounded, $|\sigma_q'(0)| = L(\sigma_q) = d(p,q)$ are bounded above by the triangle inequality, say by some value $R$ depending on $q$. So, each $\sigma_q'(0)$ lies in the compact ball $B_R = \{v \in T_pM : |v| \leq R\}$. If $q \in A$, we have $\exp_p(\sigma_q'(0)) = q$, so $A \in \exp_p(B_R)$, which is compact; since $A$ is also closed, $A$ is compact.

(4) $\rightarrow$ (1). Suppose $(x_n)$ is a Cauchy sequence, or a sequence such that for any given small positive $\delta$, all but finitely many terms of the sequence have a distance less than $\delta$ from the next term. The set of terms $\{x_n\}$ is bounded; therefore, its closure is compact and $(x_n)$ has a convergent subsequence. It is also Cauchy, so it must converge to the limit of the subsequence.

This completes the proof in all directions. $\square$

## References


[1] Thomas Franzinetti. *Geodesics in Riemannian Manifolds: Completeness and Hopf-Rinow Theorem.* July 2016. URL: `http://perso.eleves.ens-rennes.fr/people/Gwendal.Soisnard/Rapports_2016/Thomas_Franzinetti_38996.pdf`.

[2] Daniel Spiegel. *The Hopf-Rinow Theorem.* Aug. 2016. URL: `http://math.uchicago.edu/~may/REU2016/REUPapers/Spiegel.pdf`.