J. B.

Riemannian Geometry: An Almost Complete Introduction

May 8, 2024


T. 3. N°. 27.
1
SALIX Babylonica.
P. J. Redouté pinx.
SAULE pleureur. pag 113
Gabriel Sculp.
rawpixel

Preface

These are some Riemannian geometry notes written at Inria...

Contents

Chapter 1

A Rapid Introduction to Point-set Topology

Introduction

Manifolds are sets that locally look like a Euclidean space $\mathbb{R}^n$ while smooth manifolds are sets you can do calculus on (differentiate, integrate, etc.). Really simple examples of manifolds include squares or triangles or polyhedral surfaces like cubes, tetrahedra, or great cubicuboctohedra. Really simple examples of smooth manifolds include plane curves like circles or parabolas, or smooth surfaces like spheres, tori or ellipsoids. Higher dimensional examples include all points a constant distance from $\mathbf{0} \in \mathbb{R}^n$ (i.e. n - Spheres) or graphs of smooth functions between Euclidean spaces.

The simplest manifolds are topological manifolds which are topological spaces that encode what we mean by "locally looks like $\mathbb{R}^n$" and describe a notion of "closeness"; however, important applications of manifolds involve calculus. The application of manifolds to geometry involves volume and curvature; volume requires integration while curvature requires differentiation; the application of manifolds to classical mechanics involves solving systems of ordinary differential equations; the application of manifolds to general relativity involves solving partial differential equations. So we need to make sense of what it means to be "smooth" and what it means integrate and differentiate on a manifold. The problem of smoothness is NOT purely topological. A manifold has two layers: a topology and a smooth structure.

A Crash Course in Topology

What even is a topology...?

Let X be a set of stuff - it can be a continuous set of points or it can be a discrete collection of objects. A topology on X is a way of describing "closeness" or "nearby-ness" by describing neighbourhoods of elements of X. The entire set X is a neighbourhood of everything in X, so it should be part of a topology. If I join two neighbourhoods together (union) I should get another (larger) neighbourhood and if I look at the parts in common with two neighbourhoods I should get another (often smaller) neighbourhood. Two neighbourhoods may have nothing in common (class division) and so their intersection is the empty set - a neighbourhood noone lives in. It is convenient to include this in a Topology.

Let $X = \{1,2,3,4\}$. The set of all subsets of X describes all possible neighbourhoods of each element in X. This is called the discrete topology τ and (X,τ) is called a discrete space.

But why do we need a topology? We need the notion of topology in order to make sense of convergence - when and how are sequences of points getting close to one another or a limiting point; and we need the notion of topology in order to describe continuity - a continuous function.

Over the next few sections we'll cover the following:

- I will briefly say what a Topology is and what a topological space is with examples.

  I will tell you about two very basic ways to build a Topology on a space X and on subsets of X with examples.
- I will tell you what it means for a map between Topological spaces X and Y to be continuous and give examples
- I will tell you what it means for a map between Topological spaces to be a Homeomorphism and give examples.
- I will briefly mention what it means for a sequence to converge in a topology

Topological Spaces

Let's begin by formalizing the intuitive description of a topology given in the introduction.

Definition 1.1. *Let X be a set. A topology on X is a collection $\mathcal{T}$ of subsets of X, called open sets, satisfying*

a. *X and $\emptyset$ are open;*

b. *The union of any family of open subsets is open; and*

c. *The intersection of any finite family of open subsets is open*

A pair $(X,\mathcal{T})$ consisting of a set X together with a topology $\mathcal{T}$ is called a Topological space.

Let $X = \{1,2,3,4\}$ with τ the discrete topology - the set of all subsets of X. Let's verify that this is a topology:

a. X is a subset of X and the empty set is a subset of X - so these are both in τ;

b. Taking the union of any number of a subsets of X is again a subset of X;

c. Taking the intersection between subsets of X is again a subset of X.

So τ is indeed a topology on X

Topologies are not unique. It's important to note Definition 1 allows for many different topologies to be placed on a set (there are different ways to measure closeness).

Let $X = \{1, 2, 3, 4\}$. Then the following are alternative topologies

- $\tau = \{\emptyset, X\}$;
  $\tau = \{\emptyset, \{2\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}, X\}$;

It's also important to know that not every set of subsets is a topology.

Let $X = \{1, 2, 3, 4\}$. Then the following are *not* topologies on X

- $\tau = \{\emptyset, \{1, 2, 3\}, \{2, 3, 4\}, X\}$;
  $\tau = \{\emptyset, \{1, 2\}, \{2, 4\}, \{3\}, X\}$;

Basis for a Topology

In many cases it's not possible to specify all the open sets. Instead we can specify a basis of open sets and look at all the possible open sets generated by unions and intersections. This is completely analogous to a basis in a vector space. If X is a topological space then a collection $\mathcal{B}$ of subsets of X is called a basis for the topology of X if every open subset of X is the union of some collection of elements of $\mathcal{B}$. More specifically...

Definition 1.2. *If X is a set and $\mathcal{B}$ is a collection of subsets of X with*

a. *For each $x \in X$ there is at least one basis element B containing x ($X = \bigcup_{B \in \mathcal{B}} B$); and*

b. *If B_1, $B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$ then there exists a $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2$*

then we can define a topology on X by saying that a set U is open if for every $x \in U$ there exists a $B \in \mathcal{B}$ with $x \in B \subset U$. The topology is then the collection of all open sets.

Let $X := \mathbb{R}^n$ with the euclidean norm $\|\cdot\|_E$ and define an open ball of radius r, centered at $x_o \in \mathbb{R}^n$ by

$$\mathcal{B}\left(x_o, r\right) = \left\{x \in \mathbb{R}^n : \left\|x_o - x\right\|_E < r\right\}$$

The collection of all open balls at all points forms a basis for the metric topology on $\mathbb{R}^n$.

The Subspace Topology

The simplest way of making new topological spaces from old ones is by taking subsets.

Definition 1.3. *If X is a Topological space and $S \subseteq X$ an arbitrary subset of X then the subspace topology on S (sometimes called the relative topology) is defined by declaring:*

- *a subset $U \subseteq S$ to be open if there exists an open set $V \subseteq X$ such that $U = V \cap S$.*

A subset of S that is open in the subspace topology is said to be relatively open but it is NOT necessarily an open set of X.

The $S = [0, 1)$ is a subset of $\mathbb{R}$ and we give it the subspace topology, as above. The interval $U = [0, \frac{1}{2})$ is open in S with respect to the subspace topology since

$$U = (-1, \frac{1}{2}) \cap S$$

and $(-1, \frac{1}{2})$ is open in $\mathbb{R}$. But U is NOT open in $\mathbb{R}$ since there does not exist an open ball around 0 completely contained in U.

Continuous Maps

Definition 1.4. *A map $F : X \to Y$ between two topological spaces is continuous if for every open set $U \subset Y$ its pre-image $F^{-1}(U)$ is open in X.*

Let $X = \{R, G, B\}$ and $Y = \{1, 2, 3\}$ with topologies $\tau_X = \{\emptyset, \{R\}, \{B\}, \{R, G\}, \{R, B\}, X\}$ and $\tau_Y = \{\emptyset, \{1\}, \{1, 2\}, Y\}$. Then the map $f : X \to Y$ defined by $f(R) = 1$, $f(G) = 2$, and $f(B) = 3$ is continuous since

$$f^{-1}(\emptyset) = \emptyset,\ f^{-1}(\{1\}) = \{R\},\ f^{-1}(\{1, 2\}) = \{R, G\},\ f^{-1}(Y) = X$$

all of which are open in X.

The unit step function $u : \mathbb{R} \to \{0,1\}$ is given by

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Let $\mathbb{R}$ be equipped with the topology generated by all open intervals (balls) and let $\{0,1\}$ have the discrete topology. Then $u^{-1}(0) = (-\infty, 0)$ - which is open; but $u^{-1}(1) = [0, \infty)$ - which is not open. So the unit step function is discontinuous.

It's important to note that continuity is more a property of the topology than it is of the function

Consider the the unit step function deom Example 8. Instead of the discrete topology, give $\{0,1\}$ the trivial topology $\tau = \{\emptyset, \{0,1\}\}$. Then $u^{-1}(\emptyset) = \emptyset$ - which is open; and $u^{-1}(\{0,1\}) = \mathbb{R}$ - which is open! So the unit step function is now continuous! The problem is that there aren't enough open sets in the topology of the range to detect the discontinuity that our experience and intuition recognizes.

Suppose I modify the unit step function slightly, leave $\{0,1\}$ with the discrete topology and change the topology on the domain $\mathbb{R}$: $u : \mathbb{R} \to \{0,1\}$ is given by

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$

I give $\mathbb{R}$ the topology $\tau_{\mathbb{R}} = \{\emptyset, (-\infty, 0), (0, \infty), (-\infty, 0) \cup (0, \infty, 0), \mathbb{R}\}$. Then

$$u^{-1}(\emptyset) = \emptyset, \quad u^{-1}(\{0\}) = (-\infty, 0), \quad u^{-1}(\{1\}) = (0, \infty), \\ u^{-1}(\{0,1\}) = (-\infty, 0) \cup (0, \infty)$$

are all open! So this step function is now continuous! Similarly, the problem is that there aren't enough open sets in the topology of the domain to detect the "step" at 0 that our experience and intuition recognizes.

A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point x_0 if

$$\forall \epsilon > 0 \, \exists \delta > 0 \quad \text{such that} \quad |x_o - x| < \delta \implies |f(x_o) - f(x)| < \epsilon$$

where both the domain and range topologies are generated by open intervals. Show that this definition is equivalent to the open set definition above at x_o

Homeomorphisms

Definition 1.5. *A continuous bijective map $F : X \to Y$ between two topological spaces whose inverse F^{-1} is also continuous is called a homeomorphism. We say X and Y are homeomorphic.*

Let $X = \{R, G, B\}$ and $Y = \{1, 2, 3\}$ with topologies $\tau_X = \{\emptyset, \{R\}, \{B\}, \{R, G\}, \{R, B\}, X\}$ and $\tau_Y = \{\emptyset, \{1\}, \{1, 2\}, Y\}$. Then the map $f : X \to Y$ defined by $f(R) = 1$, $f(G) = 2$, and $f(B) = 3$ is bijective and continuous but is NOT a homeomorphism. Writing $g = f^{-1}$ which is a map from Y to X

$$g^{-1}(\{R, B\}) = \{1, 3\} \notin \tau_Y$$

and is therefore not open. That is, $g = f^{-1}$ is not continuous.

Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^3$. Then f is bijective; f is continuous (it's a polynomial but you can check it from the definition - Exercise); its inverse is given by $g(y) = y^{\frac{1}{3}}$ which is also continuous (it's a polynomial but check it). So f is a homeomorphism.

Definition 1.6. *A continuous map $F : X \to Y$ is a local homeomorphism if every point $p \in X$ has a neighbourhood (open set) $U \subset X$ such that $F(U)$ is open in Y and F is a homeomorphism between U and $F(U)$*

Let $f : \mathbb{R} \to S^1$ be given by $f(x) = (\cos 2\pi x, \sin 2\pi x)$ which maps $\mathbb{R}$ (with the standard topology) onto the unit circle in $\mathbb{R}^2$ (with the subspace topology). f maps intervals of the form $[n, n+1]$, for $n \in \mathbb{Z}$, onto S^1. It is not a global homeomorphism since $f(n) = f(n+1)$. However, f restricted to each of the open intervals $(n, n+1)$ and $(n - \frac{1}{2}, n + \frac{1}{2})$ is bijective and continuous with a continuous inverse so it is a local homeomorphism.

Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$ where $\mathbb{R}$ has the standard topology. Then f is not a local homeomorphism because there is no open interval around 0 on which f is bijective.

Convergence of Sequences

Definition 1.7. *Given a sequence $\{p_i\}_{i=1}^{\infty}$ of points in X and a point $p \in X$ we say the sequence converges to p if for every neighbourhood (open set) U of p there exists a natural number N such that for all $i \geq N$ $p_i \in U$. Write $p_i \to p$ or $\lim_{i\to\infty} p_i = p$.*

Let $p_i = \frac{1}{i} \in \mathbb{R}$ where $i \in \mathbb{N}$ and $\mathbb{R}$ has the standard topology. Then $p_i \to 0$ since for any interval $U = (-b, a) \subseteq \mathbb{R}$ $p_i \in U$ for all $i \geq \frac{1}{a}$.

Let $p_i = \sin\left(\frac{i\pi}{8}\right) \in \mathbb{R}$ where $i \in \mathbb{N}$ and $\mathbb{R}$ has the standard topology. Then p_i does not converge since it oscillates between 1 and -1 for all i.

Like continuity, convergence is more a property of the topology rather than of the sequence.

If we gave $\mathbb{R}$ the trivial topology $\tau = \{\emptyset, \mathbb{R}\}$ - i.e. the only open sets are the empty set and the entire space then both of the previous sequences converge AND they converge to every point in $\mathbb{R}$.

If we gave $\mathbb{R}$ the ray topology which is the topology generated by intervals of the form (a, ∞) (along with the empty set and all of $\mathbb{R}$) then the sequence $p_i = \frac{1}{i}$ converges to every negative number.

Summary

A topology gives us an idea or measure of "closeness" between points; a topology allows us to define continuous maps; a topology allows us to define convergence of sequences; however, we should be careful about what kind of topology we choose since some can generate pathological behaviour. In our subsequent study of manifolds we will put some limits on the kinds of topologies we will consider.

Chapter 2

Manifolds and Maps

Topological Manifolds

In the last chapter we saw what a topology was and how it is used to give a notion of

- closeness in terms of neighbourhoods;

  continuity;

- convergence.

which we used to define homeomorphisms and convergent sequences. However, we also saw that certain topologies come along with pathologies that contradict our general intuition and expectation. So the definition of a topological manifold should first place some conditions on the kinds of topologies we can consider so as to exclude weird behaviour and then use the concept of homeomorphism within an allowable topology to say a topological space looks like euclidean space. In particular,

Definition 2.1. *A topological manifold of dimension n is a topological space M in which*

a. *M is a Hausdorff space, meaning for every pair of distinct points p and q in M there are disjoint open sets U and V in M with $p \in U$ and $q \in V$;*

b. *M is second countable, meaning there exists a countable basis for the topology on M; and*

c. *each point of M has a neighbourhood (open set) U which is homeomorphic to an open subset $\hat{U}$ of $\mathbb{R}^n$. Specifically,*
   - *for each $p \in M$ there exists a neighbourhood U containing p;*

     a neighbourhood $\hat{U}$ of $\mathbb{R}^n$;

   - *and a homeomorphism φ mapping U onto $\hat{U}$.*

Remark 2.1. *M is a Hausdorff space - this condition ensures that the topology has enough open sets so that we don't get any weird continuity and convergence behaviour that we saw above - not every function is continuous and limits are unique.*

Remark 2.2. *M is second countable - countable means the open sets in the topology can be put into a bijective correspondence with the integers (infinite but not too infinite). This condition allows us to cover M with open sets in a way that is convenient for constructing functions and metrics.*

Remark 2.3. *The Eucliden space $\mathbb{R}^n$ is Hausdorff and second countable.*

The open set U and homeomorphism φ given in part 3 is called a coordinate chart; namely, a coordinate chart on M is a pair (U, φ) consisting of an open set U and a homeomorphism φ mapping U onto an open set $\hat{U}$ in $\mathbb{R}^n$ - just as in the definition. The set U is called a coordinate domain and φ the coordinate map. Given a point $p \in M$ we can write

$$\varphi(p) = \left(x^1(p), \ldots, x^n(p)\right) \in \mathbb{R}^n$$

where the component functions x^i are called the local coordinates .

Examples of Topological Manifolds

Here are some examples of topological manifolds

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^k$ be a continuous function, where U is open. The graph of f is a subset of $\mathbb{R}^n \times \mathbb{R}^k$ given by $G(f) := \{(x, f(x)) \in U \times \mathbb{R}^k\}$. The topology on $\mathbb{R}^n \times \mathbb{R}^k$ is the product topology where open sets have the form $V_1 \times V_2$ where V_1 is open in $\mathbb{R}^n$ and V_2 is open in $\mathbb{R}^k$. The topology on $G(f)$ will be the subspace topology. The set $G(f)$ is Hausdorff and second countable because $\mathbb{R}^n$ and $\mathbb{R}^k$ are.

Let $\pi_1 : \mathbb{R}^n \times \mathbb{R}^k$ denote projection onto the first factor. This is a continuous map: indeed, if $V \subset \mathbb{R}^n$ is open then $\pi_1^{-1}(V) = V \times \mathbb{R}^k$ which is open. Let $\varphi : G(f) \to U$ be the restriction of π_1 to $G(f)$ which is also continuous since the restriction of a continuous map is continuous. In fact, φ is a homeomorphism since it's inverse is given by $\varphi^{-1}(x) = (x, f(x))$ which is also continuous. Indeed if $W = W_1 \times W_2$ is open in $G(f)$ then $\left(\varphi^{-1}\right)^{-1}(W) = W_1$ - which is open. So $G(f)$ is homeomorphic to U - an open neighbourhood of $\mathbb{R}^n$ and is therefore a topological manifold of dimension n.

Let $\mathbb{S}^2$ be the 2-sphere which is a subset of $\mathbb{R}^3$. $\mathbb{S}^2$ is Hausdorff and second countable because it is a subspace of $\mathbb{R}^3$. For each index $i = 1, \ldots, 3$, let U_i^+ denote the subset of $\mathbb{R}^3$ where the i-th coordinate is positive:

$$U_i^+ = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_i > 0\}$$

Pick $i = 3$ for example $^*U_3^+$) and Let $f : B^2(0,1) \to \mathbb{R}$ be given by $f(u) = \sqrt{1-|u|^2}$, where $B^2(0,1)$ is the open ball of radius 1 in the x_1-x_2 plane. Now we can write

$$x_3 = 1 - x_1^2 - x_2^2 = f(x_1, x_2)$$

so that

$$U_3^+ = \{(x_1, x_2, f(x_1, x_2)) \in \mathbb{R}^2 \times \mathbb{R}\} = G(f)$$

- the graph of f. From our last example we saw that graphs of continuous functions are locally Euclidean of dimension n and U_3^+ is homeomorphic to $B^2(0,1)$! Exactly the same thing works for

$$U_3^- = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_i < 0\}$$

with f replaced by $-f$. Doing this for each combination of i and $\pm$ gives maps

$$\varphi_i^\pm : U_i^\pm \to B^2(0,1) \subset \mathbb{R}^2$$
$$\varphi_i^\pm(x^1, ..., x^3) = (x^1, ..., \hat{x^i}, ..., x^3)$$

which we call graph coordinates and since every point of $\mathbb{S}^2$ lies in some $U_i^\pm$ we see the $\mathbb{S}^2$ is a topological manifold of dimension 2.

$\mathbb{S}^n$ is a topological manifold of dimension n. The proof is identical to the proof that $\mathbb{S}^2$ is a topological manifold of dimension 2 but with more indices. The graph coordinates are

$$\varphi_i^\pm : U_i^\pm \to B^n(0,1) \subset \mathbb{R}^n$$
$$\varphi_i^\pm(x^1, ..., x^{n+1}) = (x^1, ..., \hat{x^i}, ..., x^{n+1}).$$

Fill in the details as an exercise!

NEED SOME COMMENTS ABOUT CHANGING COORDINATES

Smooth Manifolds

In section 2.1 we looked at topological manifolds and saw that their main defining characteristic is that they are locally homeomorphic to a Euclidean space. Part 3. of definition 8 gave us a notion of coordinates via coordinate charts. Let M be a topological manifold of dimension n and consider two coordinate charts (U, φ) and (V, ψ) with $U \cap V \neq \emptyset$. The composite map

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$$

is called the transition map from ψ to φ and is a essentially a change of coordinates map. Being the composition of two homeomorphisms, the transition map is also a homeomorphism. But if we are to do calculus on a manifold then this is not enough. The problem is that differentiability is not invariant under homeomorphism which really means that the smoothness properties of a function can depend on the coordinates we use - which is not what we want. Consider the following example

The map $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\varphi(x, y) = (x^{\frac{1}{3}}, y^{\frac{1}{3}})$$

is a homeomorphism and defines a coordinate chart for $\mathbb{R}^2$. On the other-hand the identity map $Id : \mathbb{R}^2 \to \mathbb{R}^2$ provides another coordinate chart for $\mathbb{R}^2$. In the identity chart the map $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x$ is smooth since $f \circ Id = x$. But in the φ chart $f \circ \varphi = x^{\frac{1}{3}}$ which is not differentiable at the origin. That is, changing from Id coordinates to φ coordinates via the transition $\varphi \circ Id^{-1}$ alters the smoothness properties of f.

So to make sense of calculus on a manifold we need to introduce some further structure in addition to its topology that will determine the differentiability of functions defined on it and ensure that coordinate changes are smooth. The model for this structure will be based on the familiar Euclidean spaces and maps between them. Let U and V be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively...

Definition 2.2. *A map $(f_1, \ldots, f_m) = F : U \to V$ is said to be smooth of class C^∞ (infinitely differentiable) if each of its component functions has continuous partial derivatives of every order. That is, writing*

$$F(\mathbf{x}) = ((f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})))$$

each f_i is a smooth function. If, in addition, F is bijective and has a smooth inverse then F is a diffeomorphism.

See Lee's Introduction to Smooth Manifolds, Appendix C for a review of smooth maps.

Definition 2.3.

Let M be a topological manifold of dimension n and consider two coordinate charts (U, φ) and (V, ψ). The two charts (U, φ) and (V, ψ) are smoothly compatible if either $U \cap V = \emptyset$ or the transition map

$$\varphi \circ \psi^{-1} : \psi (U \cap V) \subset \mathbb{R}^n \to \varphi (U \cap V) \subset \mathbb{R}^n$$

is a diffeomorphism in the sense of definition 9.

Definition 2.4. *An atlas on M is a collection of coordinate charts that cover M. A smooth atlas $\mathcal{A}$ on M is a collection of coordinate charts that cover M where any two such charts are smoothly compatible.*

To show that an atlas is smooth we usually just check that *every* transition map $\varphi \circ \psi^{-1}$ is smooth. It becomes automatic that this map is a diffeomorphism since its inverse $\psi \circ \varphi^{-1}$ is a transition map which was just checked to be smooth.

We can now say what it means for a function f to be smooth. The main idea is to call a function $f : M \to \mathbb{R}$ smooth if and only if

$$f \circ \varphi^{-1} : U \subset \mathbb{R}^n \to \mathbb{R}$$

is smooth as a map from $U \subset \mathbb{R}^n$ to $\mathbb{R}$ in every chart (U, φ) coming from an atlas $\mathcal{A}$. In this way we are able to write a map in coordinates so that we can differentiate it, and the fact that the transition maps are diffeomorphisms means that the smoothness doesn't depend on which chart (coordinate system) we use. However, there are many different smooth atlases we can put on a topological manifold that make the same maps smooth. A smooth atlas $\mathcal{A}$ on M is maximal if it is not properly contained in any larger smooth atlas. On the other hand there are many different smooth atlases which are NOT the same because the transition maps between them may not be smooth. A maximal atlas is a way of collecting all "equivalent" smooth structures. The maps given in example 24 are not smoothly compatible and hence possibly belong to different atlases. Collecting these ideas under a definition...

Definition 2.5. *A smooth manifold of dimension n is a pair $(M, \mathcal{A})$, where M is an n-dimensional topological manifold and $\mathcal{A}$ is a maximal smooth atlas.*

Remark 2.4. *One way to interpret the difference between topological and smooth manifolds is that topological manifolds are pieces of euclidean space that are sewn or glued together continuously via the transition maps while smooth manifolds are pieces of euclidean space that are sewn or glued together smoothly via the transition maps. One reason we require diffeomorphicity of the transition maps in the overlap is so that our notion of smoothness does not depend on the coordinates.*

The maximal atlas condition might also raise some questions/concerns. How would you even specify all of the charts in a maximal atlas. Fortunately, these concerns are settled by the following...

Proposition 2.1. *Every smooth atlas $\mathcal{A}$ on a topological manifold M is contained in a unique maximal smooth atlas. Two smooth atlases determine the same smooth structure if and only if their union is a smooth atlas.*

This is good. We only need to specify some smooth atlas to write down a smooth structure for a manifold.

Examples of Smooth Manifolds

Really simple examples of smooth structures are

n-dimensional Euclidean space with the smooth atlas consisting of a single chart $(\mathbb{R}^n, Id)$. Call this the standard smooth structure with standard coordinates.

Consider the homeomorphism $\psi : \mathbb{R} \to \mathbb{R}$

$$\psi(x) = x^3$$

the chart $(\mathbb{R}, \psi)$ defines a smooth atlas on $\mathbb{R}$ that is NOT compatible with the smooth structure given above since $Id \circ \psi^{-1} = x^{\frac{1}{3}}$ - which is not differentiable at zero. Therefore, the smooth structure defined by ψ is not the same as the standard one.

Recall the Sphere $\mathbb{S}^n$ - we showed that this is a topological manifold using graph coordinates

$$U_i^+ = \{(x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} : x_i > 0\}$$

$$\varphi_i^\pm : U_i^\pm \to B^n(0,1) \subset \mathbb{R}^n$$

$$\varphi_i^\pm(x^1, ..., x^{n+1}) = (x^1, ..., \hat{x^i}, ..., x^{n+1})$$

The inverse is given by

$$(\varphi_i^\pm)^{-1}(u^1, ...u^n) = (u^1, ..., \pm\sqrt{1 - |u|^2}, ..., u^n)$$

where the square root appears in the i-th position. So for two indices i and j (in this case $i > j$) the transition map is given by

$$\varphi_j^\pm \circ (\varphi_i^\pm)^{-1}(u^1, ...u^n) = (u^1, ..., \hat{u}^j, ... \pm \sqrt{1 - |u|^2}, ..., u^n)$$

which is smooth.So the sphere has the additional smooth structure and is therefore a smooth manifold.

Let V be a finite dimensional vector space over $\mathbb{R}$ with some norm determining open balls and hence a topology on V. Each basis $\{E_i\}$ of V determines an isomorphism $E : \mathbb{R}^n \to V$

$$E(\mathbf{x}) = \sum_{i=1}^n x^i E_i$$

This map is also a homeomorphism (NEED TO GIVE DETAILS HERE) so $\left(V, E^{-1}\right)$ is a chart. If $\{\tilde{E}_i\}$ is any other basis then it also determines a chart $\left(V, \tilde{E}_i^{-1}\right)$ in exactly the same way. The transition map $\tilde{E}_i^{-1} \circ E_i$ is nothing more than a change of basis matrix in $\mathbb{R}^n$ which is an invertible linear operator with constant coefficients and is therefore a diffeomorphism.

The set of $n \times m$ matrices is a vector space. So from the last example it is also a smooth manifold. What's the dimension?

Let M be a smooth manifold and $U \subseteq M$ any open subset. Define a smooth atlas on U by

$$\mathcal{A}_U = \{\text{smooth charts } (V, \varphi) \text{ for } M \text{ with } V \subseteq U\}$$

Every point $p \in U$ lies in some coordinate chart (W, ψ) so setting $V = W \cap U$ and restricting ψ to V gives a chart $(V, \psi|_V)$ in $\mathcal{A}_U$. This means the domains of charts in $\mathcal{A}_U$ cover U. Since $\mathcal{A}_U$ consists of smooth charts for M, which is a smooth manifold, the transition maps must be diffeomorphisms so that $\mathcal{A}_U$ is a smooth structure for U. We will call any open subset of M an open submanifold.

The General Linear Group $GL(n, \mathbb{R})$ is the set of $n \times n$ invertible matrices. The determinant map $\det : M(n, \mathbb{R}) \to \mathbb{R}$ is a surjective map which is polynomial in the entries of a matrix and is therefore continuous. Invertible matrices have a non-zero determinant. Since $\mathbb{R}/\{0\}$ is open $\det^{-1}(\mathbb{R}/\{0\}) = GL(n, \mathbb{R})$ is an open subset of $M(n, \mathbb{R})$ and is therefore a smooth manifold by the last example.

Manifolds With Boundary

Boundaries cause problems and it's important to be aware of this. But being an introductory text we will not spend so much time developing the technical details necessary for coping with these problems. We will give the definitions of a topological and smooth manifold with boundary and explain what it means for a function or map to be smooth on these objects but we will restrict our attention to manifolds without boundary in later chapters.

Sets can have boundaries. Examples include closed intervals in $\mathbb{R}$, closed balls in $\mathbb{R}^n$, and hemispheres in $\mathbb{S}^n$. Many applications involve boundaries so it's useful to extend our definition of topological and smooth manifold to include these sets.

Points in sets with boundaries will have neighbourhoods modelled on open sets in the closed upper half space

$$\mathbb{H}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n \geq 0\}.$$

Use the notation

$$\begin{aligned} \mathrm{Int}\mathbb{H}^n &= \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n > 0\} \\ \partial\mathbb{H}^n &= \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n = 0\} \end{aligned}$$

to denote the interior and the boundary; here the plane $x_n = 0$ models the boundary.

Definition 2.6. *An n-dimensional topological manifold with boundary is*

a. *A Hausdorff space, meaning for every pair of distinct points p and q in M there are disjoint open sets U and V in M with $p \in U$ and $q \in V$;*

b. *second countable, meaning there exists a countable basis for the topology on M; and*

c. *a space in which each point has a neighbourhood (open set) U which is homeomorphic to a (relatively) open subset $\hat{U}$ of $\mathbb{H}^n$.*

Just as in the case of manifolds without boundary a chart for M consists of

- An open set $U \subseteq M$; and

  a homeomorphism φ of U onto an open set $\hat{U} \in \mathbb{H}^n$

We say (U, φ) is an interior chart if $\varphi(U) \subset \mathrm{Int}\mathbb{H}^n$ and (U, φ) is a boundary chart if $\varphi(U) \subset \mathrm{Int}\mathbb{H}^n$ with $\varphi(U) \cap \partial\mathbb{H}^n \neq \emptyset$. A point $p \in M$ is an interior point if it is contained in an interior chart and a point $q \in M$ is a bounadry point if it is contained in a boundary chart and mapped into $\partial\mathbb{H}^n$.

It's not entirely obvious that a point which is interior in one chart can never be on the boundary in another chart. It's true that this cannot happen but the proof involves cohomological algebra which is more than we have right now, so we'll just state it and... run away.

Theorem 2.2. *(Topological Invariance of the Boundary) If M is a topological manifold with boundary then every point is either an interior point or a boundary point but never both. That is, the boundary of M, ∂M, and the interior of M, $\mathrm{Int}M$, are disjoint sets whose union is M.*

To understand what it means for a transition map to be smooth on a manifold with boundary we'll state what it means for a map to be smooth on a space like $\mathbb{H}^n$.

Definition 2.7. *If U is an open subset of $\mathbb{H}^n$, a map*

$$F : U \to \mathbb{R}^k$$

is smooth if for each $x \in U$ there exists an open subset $\tilde{U} \subset \mathbb{R}^n$ containing x and a smooth map

$$\tilde{F} : \tilde{U} \to \mathbb{R}^k$$

that agrees with F on $\tilde{U} \cap \mathbb{H}^n$.

If F is such a map then its restriction to $U \cap \mathrm{Int}\mathbb{H}^n$ is smooth in the usual sense (continuous partial derivatives of all orders) AND the continuity determines the values of the partial derivatives at $U \cap \partial\mathbb{H}^n$ so our definition is independent of the extension.

Let $B(0,1) \subset \mathbb{R}^2$ be the open unit disk, set $U = B(0,1) \cap \mathbb{H}^2$ and define $f : U \to \mathbb{R}$ by

$$f(x,y) = \sqrt{1 - x^2 - y^2}$$

This map f extends smoothly to all of $B(0,1)$ by the same formula so it's a smooth function on U.

The function $g : U \to \mathbb{R}$ given by

$$g(x,y) = \sqrt{y}$$

is continuous on U and smooth on $U \cap \mathrm{Int}\mathbb{H}^n$ but has no smooth extension across the origin since $\frac{\partial g}{\partial y} \to \infty$ as $y \to 0$. So g is not smooth on U.

A smooth structure for a manifold with boundary is completely analogous to a smooth structure for a manifold without...

Definition 2.8. *A smooth structure for a manifold with boundary M is defined to be a maximal Atlas $\mathcal{A}$ which consists of a collection of coordinate charts covering M whose transition maps are smooth in the sense described above.*

Let's see an example that illustrates the above ideas.

We'll show that the closed ball $\bar{B}(0,1) = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\} \subset \mathbb{R}^3$ is a smooth manifold with boundary. Let

$$U_z^+ = \{((x,y,z) \in \bar{B}(0,1) : z > 0\}$$

and consider the function $\varphi_z^+ : U_z^+ \to \mathbb{H}^3$ given by

$$\varphi_z^+(x,y,z) = (x, y, 1 - x^2 - y^2 - z^2)$$

The boundary portion $U_z^+ \cap \mathbb{S}^2$ gets mapped to the open disk in $\partial\mathbb{H}^3$ while the interior is mapped to an open dome in $\mathrm{Int}\mathbb{H}^3$. Since the component functions of φ_z^+ are continuous, φ_z^+ is continuous; its inverse is given by

$$(\varphi_z^+)^{-1}(u_1, u_2, u_3) = (u_1, u_2, \sqrt{1 - u_1^2 - u_2^2 - u_3})$$

which is also continuous so that φ_z^+ is a homeomorphism. Doing this for each possible coordinate chart $U_\cdot^\pm$ gives the closed ball the structure of a 3-dimensional topological manifold with boundary. To see that it's a smooth manifold consider

$$U_x^+ = \{((x, y, z) \in \bar{B}(0,1) : x > 0\}$$

with homeomorphism

$$\varphi_x^+(x, y, z) = (1 - x^2 - y^2 - z^2, y, z)$$

and look at the overlap $U_x^+ \cap U_z^+$. Using the expression for $(\varphi_z^+)^{-1}$ we get that the transition map is given by

$$\varphi_x^+ \circ (\varphi_z^+)^{-1}(u_1, u_2, u_3) = (u_3, u_2, \sqrt{1 - u_1^2 - u_2^2 - u_3})$$

which is smooth.

Smooth Maps

Definition 2.9. *Suppose M is a smooth n-manifold and*

$$f : M \to \mathbb{R}^k$$

is any function. We say that f is a **smooth function** *if for every $p \in M$ there exists a smooth chart (U, φ) for M whose domain contains p and such that the composite function*

$$f \circ \varphi^{-1}$$

is smooth between $\varphi(U) \subset \mathbb{R}^n$ and $\mathbb{R}^k$.

Remark 2.5. *If M has boundary then absolutely nothing changes except that $\varphi(U)$ lives in $\mathbb{H}^n$. If $\varphi(U) \cap \partial\mathbb{H}^n \neq \emptyset$ then smoothness of $f \circ \varphi^{-1}$ is interpreted to mean that each point of $\varphi(U)$ has a neighbourhood in $\mathbb{R}^n$ on which $f \circ \varphi^{-1}$ extends to a smooth function as above.*

Given a function $f : M \to \mathbb{R}^k$ and a coordinate chart (U, φ) the function

$$\hat{f} : \varphi(U) \to \mathbb{R}^k$$

is called the **coordinate representation of f**. Smooth functions have smooth coordinate representations in every chart (why?) The most important special case is a smooth real valued function

$$f : M \to \mathbb{R};$$

the set of all such functions are denoted $C^\infty(M)$. Many other functions are built out of C^∞ functions. For example, functions between $\mathbb{R}^n$ and $\mathbb{R}^k$ have real valued functions as their component functions. Diffeomorphisms of M, when written in coordinates have real valued functions as their component functions.

The above definition generalizes easily to smooth maps between manifolds M and N.

Definition 2.10. *Let M and N be any two m and n dimensional manifolds, respectively, and let*

$$F : M \to N;$$

be any map. We say that F is a **smooth map** *if for every $p \in M$ there exist smooth charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $F(U) \subset V$ and the composite map*

$$\psi \circ F \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^m \to \psi(V) \subseteq \mathbb{R}^n$$

is smooth as a map between Euclidean spaces.

Remark 2.6. *If M and/or N have boundary then the definition remains pretty much the same with the understanding that one or both of the smooth charts map to $\mathbb{H}^n$ and smoothness is interpreted as an extension to a smooth map in a neighbourhood of each point when the domain is a subset of $\mathbb{H}^n$ and/or as a smooth map into $\mathbb{R}^n$ when the codomain is a subset of $\mathbb{H}^n$. Our definition of smooth function is really just a special case of this definition when $N = \mathbb{R}^n$ and $\psi = Id$.*

Remark 2.7. *When we consider our definition of smooth map we realise that every chart in $\mathcal{A}$ actually consists of an open set U and a diffeomorphism φ. Why? If (U, φ) and (V, ψ) are any two charts for M then $\varphi \circ \psi^{-1} : V \subset \mathbb{R}^n \to U \subset \mathbb{R}^n$ is a smooth diffeomorphism. But this is simply the coordinate representation of φ and implies that φ is smooth as a map from M into $\mathbb{R}^n$. Similarly, for the inverse and therefore φ is a diffeomorphism. This provides another difference between topological manifolds and smooth manifolds: topological manifiolds are locally homeomorphic to $\mathbb{R}^n$ while smooth manifolds are locally diffeomorphic to $\mathbb{R}^n$.*

How do our definitions of smoothness interact with our previous topological ideas of continuity given in chapter 1?... Well, they interact well.

Proposition 2.3. *Every smooth map is continuous*

◦ *Proof.* Use the definition: A map F is smooth if for every $p \in M$ there exist smooth charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $F(U) \subseteq V$ and the composite map

$$\psi \circ F \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^m \to \psi(V) \subseteq \mathbb{R}^n$$

is smooth as a map between Euclidean spaces. The composite map is a composition of continuous maps and is therefore also continuous. Since both $\varphi : U \to \varphi(U)$ and $\psi : V \to \psi(V)$ are homeomorphisms

$$F|_U = \psi^{-1} \circ \left(\psi \circ F \circ \varphi^{-1}\right) \circ \varphi : U \to V$$

is also continuous so that F is continuous on a neighbourhood of each point and therefore continuous on M.

$\square$

Details Matter... To prove a map is smooth directly from the definition means for each $p \in M$ we prove the existence of coordinate charts (U, φ) containing p and (V, ψ) containing $F(p)$ with $F(U) \subset V$. This is included in the definition so as to automatically imply continuity. But if this condition is omitted then there exist smooth maps which are not continuous...

Chapter 3

The Tangent Space to a Smooth Manifold

A central idea in calculus is linear approximation. A function of a single variable can be approximated by its tangent line; a parameterized curve in $\mathbb{R}^n$ can be approximated by its velocity vector; or a surface in $\mathbb{R}^3$ by its tangent plane. The tangent space at a point of a smooth manifold M is a higher dimensional analogue of the above examples and can be thought of as a "linear model" for M near the point.

We'll go through two views of a tangent vector and the tangent space at a point for $\mathbb{R}^n$:

- Geometric tangent vectors, which can be visualised as arrows attached to points;

  Algebraic tangent vectors, which are linear maps acting on smooth functions (essentially as directional

We will see that the process of taking a directional derivative gives a one-to-one correspondence between geometric tangent vectors and linear maps from $C^\infty(\mathbb{R}^n)$ to $\mathbb{R}$ satisfying the product rule so that the two views are structurally equivalent.

Later, we'll also see how a smooth map between manifolds induces a linear map between tangent spaces, called the **differential** of the map. Using this we'll show that any smooth coordinate chart (U, φ) yields a basis for each tangent space. This will, in turn, allow us to see how the union of all tangent spaces at all points can be glued together to form a new manifold called **the tangent bundle** of the original manifold.

The Idea...

This introduction will be largely expository and mainly focus on the manifold $\mathbb{R}^n$ and a conceptualisation of the tangent space at one of its points.

What do we mean by a tangent vector at a point in $\mathbb{R}^n$? The answer to this question is related to how we think about elements of $\mathbb{R}^n$. On the one hand we think of them as points in space whose only property is location, in coordinates $(x_1, \ldots, x_n)$. On the other hand, we sometimes think of them as vectors $v = v^1 e_1 + \cdots + v^n e_n$ which are objects that have magnitude and direction but whose location is irrelevant. These are visualised as arrows with initial point being anywhere in $\mathbb{R}^n$.

What's really going on here with the vector point of view is that we are thinking about a separate copy of $\mathbb{R}^n$, existing at every (locational) point, in which our vector lives. So when we think about tangent vectors at a point $p \in \mathbb{R}^n$ we actually imagine them living in a separate copy of $\mathbb{R}^n$ with its origin translated to p. Here's a geometric definition of a tangent vector and tangent space for $\mathbb{R}^n$

Definition 3.1. *Given a point $p \in \mathbb{R}^n$, the tangent space to $\mathbb{R}^n$ at p is the set*

$$\mathbb{R}^n_p := \{p\} \times \mathbb{R}^n = \{(p, v) : v \in \mathbb{R}^n\}$$

A geometric tangent vector in $\mathbb{R}^n$ is an element of $\mathbb{R}^n_p$ for some $p \in \mathbb{R}^n$. We abbreviate (p, v) by v_p and think of it as the vector whose initial point is at p

The set $\mathbb{R}^n_p$ is a vector space in the usual sense

$$v_p + w_p = (v + w)_p \quad cv_p = (cv)_p$$

(just think about our experience of vector calculus). As a vector space $\mathbb{R}^n_p$ is exactly the same as $\mathbb{R}^n$; we have just collected all those vectors whose initial points are p. One consequence of this is that if p and q are two different points then $\mathbb{R}^n_p$ and $\mathbb{R}^n_q$ will be disjoint sets.

With this idea in mind we can define the tangent space to $\mathbb{S}^{n-1}$ at a point p as a subspace of $\mathbb{R}^n_p$; namely, the subspace consisting of all vectors orthogonal to the radial unit vector through p (where we are using the euclidean inner product on $\mathbb{R}^n$) This works just fine for any smooth manifold M that is a subset of $\mathbb{R}^n$; for example

- open sets of $\mathbb{R}^n$

  graphs of smooth functions

- tori, ellipoids, etc.

However, as far as our general definition of smooth manifold goes (which does not explicitly involve an embedding in Euclidean space) all we really have to work with are smooth functions, smooth maps and smooth coordinate charts. Let's look at $\mathbb{R}^n$ again and see if we can use the geometric tangent vector to come up with a different description of a tangent vector which makes use of functions and that will generalise to our setting of smooth manifold and smooth maps.

Let $\alpha : \mathbb{R} \to \mathbb{R}^n$ be a differentiable curve in $\mathbb{R}^n$, with $\alpha(0) = p$. Write

$$\alpha(t) = (x_1(t), \ldots, x_n(t))$$

so that

$$\alpha'(0) = (x_1'(0), \ldots, x_n'(0)) = v \in \mathbb{R}^n$$

Now let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. We can restrict f to the curve and express the directional derivative of f with respect to the vector v as

$$\frac{d\, f \circ \alpha}{dt}|_{t=0} = \sum_i \frac{df}{dx_i}|_{t=0} \cdot \frac{dx_i}{dt}|_{t=0} = \left(\sum_i x_i'(0) \frac{\partial}{\partial x_i} \right) f$$

This says that the directional derivative with respect to v is an operator on differentiable functions which uniquely depends on v. Write the directional derivative with respect to v at the point p as

$$\begin{aligned} D_v|_p &: C^\infty(\mathbb{R}^n) \to \mathbb{R} \\ D_v|_p &= \sum_i v_i \frac{\partial}{\partial x_i} \end{aligned}$$

where $v = (v_1, \ldots, v_n)$. It is linear and satisfies the product rule

$$\begin{aligned} D_v|_p\,(a \cdot f + b \cdot g) &= a \cdot D_v|_p f + b \cdot D_v|_p g \\ D_v|_p(f \cdot g) &= g \cdot D_v|_p f + f \cdot D_v|_p g \end{aligned}$$

for any two smooth functions f and g and any two constants a and b. With the above construction of directional derivative in mind let's make the following definition

Definition 3.2. *If p is a point in $\mathbb{R}^n$, a map*

$$w : C^\infty(\mathbb{R}^n) \to \mathbb{R}$$

is called a derivation at p if it is linear over $\mathbb{R}$ and satisfies the product rule

$$\begin{aligned} w\,(a \cdot f + b \cdot g) &= a \cdot w(f) + b \cdot w(g) \\ w(f \cdot g) &= g \cdot w(f) + f \cdot w(g) \end{aligned}$$

for any two smooth functions f and g and any two constants a and b.

Let $T_p\mathbb{R}^n$ be the set of all derivations of $C^\infty(\mathbb{R}^n)$ at p. Then $T_p\mathbb{R}^n$ is a vector space:

$$(w_1 + w_2)f = w_1(f) + w_2(f), \quad w(c \cdot f) = c \cdot w(f)$$

I now want to show you that $T_p\mathbb{R}^n$ is actually finite dimensional and is isomorphic to $\mathbb{R}^n_p$ so that this algebraic notion of tangent space is structurally equivalent to our geometrically intuitive notion of tangent vector while allowing for a generalisation to arbitrary smooth manifolds without appealing to embeddings.

Lemma 3.1. *(Properties of Derivations) Suppose $p \in \mathbb{R}^n$, $w \in T_p\mathbb{R}^n$, and $f, g \in C^\infty(\mathbb{R}^n)$.*

a. *If f is a constant function then $w(f) = 0$*

b. *If $f(p) = g(p) = 0$ then $w(f \cdot g) = 0$*

Proof. We'll prove 1. for $f_1(x) = 1$; then $f(x) = c = c \cdot f_1(x)$ will follow from linearity. Notice that $f_1(x) = 1 = 1 \cdot 1 = f_1(x) \cdot f_1(x)$; the product rule then gives

$$w(f_1) = w(f_1 \cdot f_1) = f_1(a) \cdot w(f_1) + f_1(a)w(f_1) = 2w(f_1)$$

which implies $w(f_1) = 0$. Part 2. also follows from the product rule:

$$w(f \cdot g) =) = f(a) \cdot w(g) + g(a)w(f) = 0.$$

$\square$

This Theorem shows that there is a one-to-one correspondence between derivations and geometric tangent vectors.

Theorem 3.2. *Let $p \in \mathbb{R}^n$*

a. *For each geometric tangent vector $v_p \in \mathbb{R}^n_p$, the directional derivative $D_v|_p : C^\infty(\mathbb{R}) \to \mathbb{R}$ defined on slides 36 and 37 is a derivation*

b. *The map $v_p \mapsto D_v|_p$ is an isomorphism from $\mathbb{R}^n_p$ onto $T_p\mathbb{R}^n$*

Proof. For part 1. $D_v|_p$ is a directional derivative and therefore satisfies the product rule... so it's a derivation.

Part 2.: There are three things to check: (1) linearity; (2) injectivity; and (3) surjectivity. For linearity, take v_p and w_p; then

$$\begin{aligned} v_p + w_p \mapsto D_{v+w}|_p = \sum_{i=1}^n \left(v^i_p + w^i_p\right) \frac{\partial}{\partial x^i} &= \sum_{i=1}^n v^i_p \frac{\partial}{\partial x^i} + \sum_{i=1}^n w^i_p \frac{\partial}{\partial x^i} \\ &= D_v|_p + D_w|_p \end{aligned}$$

To check injectivity suppose that v_p induces the zero derivation $D_v|_p = 0$. Writing $v_p = v^i e_i|_p$ in terms of the standard basis, and taking f to be the j-th coordinate function $x^j : \mathbb{R}^n \to \mathbb{R}$ (which is a smooth function) we have

$$0 = D_v|_p(x^j) = v^i \frac{\partial}{\partial x^i}(x^j) = v^j.$$

Since this is true for each j, $v_p = 0$. To check surjectivity let $w \in T_p\mathbb{R}^n$ be any derivation. Define $v = v^i e_i$ via

$$w(x^i) = v^i.$$

We'll show that $w = D_v|_p$. Let f be any smooth real-valued function on $\mathbb{R}^n$. Taylor's Theorem allows us to write

$$\begin{aligned} f(x) = f(a) + &\sum_{i=1}^n \frac{\partial f}{\partial x^i}(x^i - a^i) + \\ &+ \sum_{i,j=1}^n (x^i - a^i)(x^j - a^j) \int_o^1 (1-t) \frac{\partial^2 f}{\partial x^i \partial x^j}(a + t(x-a))\, dt. \end{aligned}$$

The second sum is a product of two smooth functions which vanish at a. The Derivation evaluates to zero on this term by part 2 of Lemma 1. What remains is

$$
\begin{aligned}
w(f) &= w(f(a)) + \sum_{i=1}^{n} w\left(\frac{\partial f}{\partial x^i}(a)(x^i - a^i)\right) \\
&= 0 + \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(a)(w((x^i) - w(a^i)) \\
&= \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(a)v^i = D_v|_p(f).
\end{aligned}
$$

□

Corollary 3.3. *For any $p \in \mathbb{R}^n$, the n derivations*

$$\frac{\partial}{\partial x^1}|_p, \dots, \frac{\partial}{\partial x^n}|_p$$

defined by

$$\frac{\partial}{\partial x^i}|_p f = \frac{\partial f}{\partial x^n}(p)$$

form a basis for $T_p\mathbb{R}^n$, which therefore has dimension n.

Proof. Let $e_1, dots, e_n$ be a basis for $\mathbb{R}_p$. By the previous Theorem the map $v_p \mapsto D_v|_p$ is a linear isomorphism so that

$$D_{e_i}|_p = \frac{\partial}{\partial x^i}|_p$$

is a basis for $T_p\mathbb{R}^n$. □

We looked at two different views of a tangent vector and tangent space:

- The geometric tangent vector which consists of "arrows" at a point

  The algebraic tangent vector which consists of directional derivatives defined by geometric tangent vectors

We then showed that these two views are structurally equivalent where the latter has the advantage of not relying on an embedding and only considering the action of a derivation on smooth functions. It is good to have both views in mind: the geometric view is intuitive - this is useful; the algebraic view turns out to be more practical. In fact, almost everything we will see later - differentials, connections, covariant derivatives - are built out of derivations so it's a good idea to get used to them.

Algebraic Tangent Vectors

I'm going to define tangent vectors at a point p in a manifold as derivations at the point p and set the tangent space at p equal to the set of all derivations at p. Then we're going to see that this tangent space is isomorphic with the tangent space to $\mathbb{R}^n$ via a coordinate chart so that our algebraic definition of tangent space is structurally equivalent to our geometric intuition. To do this we'll need to introduce the differential of a smooth map, which we'll cover along the way. In the next section I'll give a second "geometric" definition of a tangent vector on a manifold. When talking about or thinking about tangent vectors you should visualize them as being arrows that are tangent to M at points of M. Proofs about tangent vectors will generally be done using the abstract definition of tangent vectors but intuition should be guided by the geometric picture.

For algebraic tangent vectors we're basically ready for general manifolds...

Definition 3.3. *For any point p in a smooth n-dimensional manifold M with or without boundary. A linear map $v : C^\infty(M) \to \mathbb{R}$ is called a derivation at p if it satisfies*

$$v(f \cdot g) = f(p) \cdot v(g) + g(p) \cdot v(f) \quad \forall\, f,\, g \in C^\infty(M).$$

The set of all derivations of $C^\infty(M)$ at p, also denoted T_pM , is a vector space called the tangent space to M at p and an element of T_pM is called a tangent vector.

We have the same properties of derivations on manifolds as we had for $\mathbb{R}^n$.

Lemma 3.4. *(Properties of Derivations) Suppose $p \in M$, $w \in T_pM$, and $f, g \in C^\infty(M)$.*

a. *If f is a constant function then $w(f) = 0$*

b. *If $f(p) = g(p) = 0$ then $w(f \cdot g) = 0$*

Proof. Exercise. Set up a coordinate chart and mimic the proof of Lemma 1. □

To relate the abstract definition of tangent spaces to manifolds to geometric tangent vectors in $\mathbb{R}^n$ we need to look at how smooth maps affect tangent vectors. In Euclidean space the total derivative of a map at a point (represented by its Jacobian matrix) is a linear map that represents the "best linear approximation" to the map near a point. In the manifold setting there is a similar map which is a map between tangent spaces.

Definition 3.4. *If M and N are smooth manifolds with or without boundary and $F : M \to N$ is a smooth map, for each $p \in M$ we define a map*

$$dF_p : T_pM \to T_{F(p)}N$$

called the differential of F at p as follows: given $v \in T_pM$, we let $dF_p(v)$ be the derivation at $F(p)$ that acts on $g \in C^\infty(N)$ by

$$dF_p(v) \cdot g = v(g \circ F)$$

Not that the definition makes sense since the composition of smooth functions is smooth (but check in a coordinate chart for practice). The operator $dF_p(v) : C^\infty(N) \to \mathbb{R}$ is linear because v is. It is also a derivation at $F(p)$ since for any $f, g \in C^\infty(N)$

$$\begin{aligned} dF_p(v)(f \cdot g) &= v\left([f \circ F] \cdot [g \circ F]\right) \\ &= f \circ F(p) \cdot v(g \circ F) + g \circ F(p) \cdot v(f \circ F) \\ &= f(F(p)) \cdot [dF_p(v) \cdot g] + g(F(p)) \cdot [dF_p(v) \cdot f] \end{aligned}$$

so the differential does indeed map derivations in T_pM to derivations in $T_{F(p)}N$ and definition 21 is working as it should. Let's record some useful properties of differentials.

Proposition 3.5. *Let M, N, and P be smooth manifolds with or without boundary, let $F : M \to N$ and $G : N \to P$ be smooth maps, and let $p \in M$.*

a. $dF_p : T_pM \to T_{F(p)}N$ *is linear;*

b. $d\left(G \circ F\right)_p = dG_{F(p)} \circ dF_p : T_pM \to T_{G \circ F(p)}P$;

c. $d(Id_M)_p = Id_{T_pM} : T_pM \to T_pM$; *and*

d. *If F is a diffeomorphism then $dF_p : T_pM \to T_{F(p)}N$ is an isomorphism and $d\left(F^{-1}\right)_{F(p)} = dF_p^{-1}$.*

Proof. I'll prove 2. and 4., you can prove 1. and 3. To prove these sorts of statements just check how the left hand side acts on a derivation and rewrite it in terms of the right hand side. For 1. let $v \in T_pM$ and pick any $h \in C^\infty(P)$. Then

$$\begin{aligned} d\left(G \circ F\right)_p(v) \cdot h &= v\left(h \circ G \circ F\right) \\ &= v\left([h \circ G] \circ F\right) \\ &= dF_p(v)(h \circ G) \\ &= dG_{F(p)}\left(dF_p(v)\right) \cdot h \\ &= dG_{F(p)} \circ dF_p(v) \cdot h \end{aligned}$$

Part 4. follows from parts 2. and 3:

$$\begin{aligned} Id_{T_pM} &= d\left(Id_M\right)_p \\ &= d\left(F^{-1} \circ F\right)_p \\ &= d(F^{-1})_{F(p)} \circ dF_p \end{aligned}$$

Similarly,

$$Id_{T_{F(p)}N} = dF_p \circ d(F^{-1})_{F(p)}$$

This shows that $d\left(F^{-1}\right)$ is a left and right inverse for dF_p which means

$$d\left(F^{-1}\right)_{F(p)} = dF_p^{-1}$$

so that dF_p is invertible between T_pM and $T_{F(p)}$ and hence is an isomorphism. $\square$ $\square$

We'll now use differentials of functions and coordinate charts to relate our abstract tangent space to the Euclidean tangent space. The Proposition I'll prove is the following:

Proposition 3.6. *If M is an n-dimensional smooth manifold then for each $p \in M$ the tangent space is an n-dimensional vector space with*

$$T_pM \cong T_{\varphi(p)}\mathbb{R}^n$$

for a chart (U, φ) around p.

What do we need to do to prove this? Recall that a smooth manifold M is a topological manifold with a smooth atlas $\mathcal{A}$ consisting of charts whose transition maps are diffeomorphisms (see Remark **??**). Given $p \in M$, let (U, φ) be a smooth coordinate chart containing p. Since φ is a diffeomorphism from U onto $\tilde{U} \subset \mathbb{R}^n$ it follows that $d\varphi_p$ is an isomorphism from T_pU onto $T_\varphi(p)\tilde{U}$ (by Proposition 3, 4.) and we have the following isomorphism/identification

$$T_pU \cong T_{\varphi(p)}\tilde{U}.$$

As far as the statement of Proposition 4 is concerned this isn't enough: there's a minor technical issue in that our definition of tangent space uses functions defined on all of M while coordinate charts are only defined on open sets of M. But it's ok... we can relate the tangent space to an open set U in M to the tangent space to M. Any open set $U \subset M$ is a manifold and has a tangent space T_pU to a point $p \in U$ consisting of derivations acting on smooth functions defined on U. The manifold M also has a tangent space at p, T_pM, which consists of derivations acting on smooth functions defined on M. In order to identify these two tangent spaces as the same we will first notice that derivations only rely on information about functions on neighbourhoods of points, and then we will show that derivations are independent of how we extend functions from neighbourhoods to all of M. In doing so we will establish $T_pM \cong T_pU$. The Euclidean space $\mathbb{R}^n$ is also a manifold and the diffeomorphic image of U under φ is an open set in $\mathbb{R}^n$ so a particular case of the isomorphism $T_pM \cong T_pU$ is $T_{\varphi(p)}\tilde{U} \cong T_{\varphi(p)}\mathbb{R}^n$. These, together with the isomorphism (**??**), complete the proof of Proposition 4.

Lemma 3.7. *Let M be a smooth manifold with or without boundary, $p \in M$, $v \in T_pM$. If f, $g \in C^\infty(M)$ agree on some neighbourhood of p then $v(f) = v(g)$.*

Proof. Take f, $g \in C^\infty(M)$ and suppose $f = g$ on a neighbourhood U of p. Set $h = f - g$ so that $h = 0$ on U and let $\psi \in C^\infty(M)$ with $\psi \equiv 1$ on M/U and 0 on U (ψ is called a bump function). Then since $\psi \equiv 1$ when h is non-zero we have $h = \psi \cdot h$ and since both $\psi(p) = h(p) = 0$ we have

$$v(h) = v(\psi \cdot h) = \psi(p) \cdot v(h) + h(p) \cdot v(\psi) = 0$$

using the properties of derivations so that $v(f) = v(g)$ by linearity. $\square$

Lemma 3 allows us to identify the tangent space to an open submanifold with the tangent space of the whole manifold. If A is a subset of B then the inclusion map ι is a map that sends each element x of A to x but treated as an element of B.

Lemma 3.8. *Let M be a smooth manifold with or without boundary, let $U \subseteq M$ be an open subset, and let $\iota : U \hookrightarrow M$ be the inclusion map. For every $p \in U$ the differential $d\iota_p : T_pU \to T_pM$ is an isomorphism.*

Proof. For injectivity suppose $v \in T_pU$ and $d\iota_p(v) = 0 \in T_pM$. Let B be a neighbourhood of p with $\bar{B} \subseteq U$. For any $f \in C^\infty(U)$ we can extend f to a smooth function $\tilde{f} \in C^\infty(M)$ in such a way that $f \equiv \tilde{f}$ on $\bar{B}$ (see the extension Lemma). Since $f \equiv \tilde{f}$ on a neighbourhood of p the previous proposition gives $v(f) = v(\tilde{f}|_U)$ so that

$$v(f) = v(\tilde{f}|_U) = v(\tilde{f} \circ \iota) = d\iota_p(v)f = 0.$$

As this holds for any f we must have $v = 0$.

For surjectivity suppose $w \in T_pM$. Define an operator $v : C^\infty(M) \to \mathbb{R}$ by setting $v(f) = w(\tilde{f})$ - where f and $\tilde{f}$ are as above. Lemma 3 implies that vf is independent of the choice of $\tilde{f}$ so that v is well defined and is a derivation because w is. Now pick a $g \in C^\infty(M)$:

$$d\iota_p(v)g = v(g \circ \iota) = v(g \tilde{\circ} \iota) = w(g).$$

Where $g \tilde{\circ} \iota$ is any extension of $g \circ \iota$ to M, and g, $g \circ \iota$, $g \tilde{\circ} \iota$ all agree on B. This shows surjectivity. $\square$

Geometric Tangent Vectors

What are the geometric tangent vectors for manifolds? For manifolds embedded in $\mathbb{R}^n$ You essentially take a curve $\gamma(t)$ which is constrained to lie in M and pass through a point $p \in M$ at $t = 0$ and look at its velocity vector $v = \frac{d}{dt}|_{t=0}\gamma(t)$ - the collection of all velocity vectors of all curves passing through p gives the tangent space. For general manifolds take a chart (U, φ) centered at p and let γ be a smooth map from $\mathbb{R}$ into M with $\gamma(0) = p$. Then we can define a tangent vector by $v = \frac{d}{dt}|_{t=0}\varphi \circ \gamma(t)$. One problem with each of these definitions is that different curves might produce the same v! So we define the following equivalence relation.

Definition 3.5. *For any point p in a smooth n-dimensional manifold M, two smooth curves $c_1 : \mathbb{R} \to M$ and $c_2 : \mathbb{R} \to M$ passing through p ($c_1(0) = c_2(0) = p$) are equivalent if and only if there is some chart (U, φ) at p such that*

$$(\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0)$$

This definition looks as though it depends on the choice of chart but this is not the case. The reason is that if (V, ψ) is any other chart then the transition map $\psi \circ \varphi^{-1}$ is a diffeomorphism on $U \cap V$ so that it's differential is an isomorphism by Lemma 3. Finally we can say

Definition 3.6. *For any point p in a smooth n-dimensional manifold M, a tangent vector to M at p is an equivalence class of smooth curves through p determined by the equivalence relation* (**??**)*. The set of all tangent vectors at p is denoted by T_pM.*

The vector space structure is: (1) the sum of two equivalence classes of curves $[c_1]$ and $[c_2]$ at a point p is the equivalence class of curves whose tangent vector in a chart at p is the sum of the tangent vectors corresponding to $[c_1]$ and $[c_2]$ written in the same chart; and (2) a scalar multiple $k \cdot [c]$ of an equivalence class $[c]$ at p is the equivalence class corresponding to the tangent vector of $[c]$ in a chart scaled by k. It should be clear why we favor the algebraic definition over the geometric one...

Tangent Vectors and Differentials in Coordinates

We've talked about tangent vectors adn we've talked about differentials of maps but how do you actually compute with these objects?

Last time we showed that if U is an open set in M and $p \in U$ then $T_pU \cong T_pM$ via Lemma 4. Instead of referencing the isomorphism we will just identify these two spaces as one and the same (because they are). So, if (U, φ) is a coordinate chart with φ a diffeomorphism onto $\hat{U} \subset \mathbb{R}^n$ the differential

$$d\varphi_p : T_pM \to T_{\varphi(p)}\mathbb{R}^n$$

is an isomorphism. Recall that coordinate derivations

$$\frac{\partial}{\partial x^1}|_{\varphi(p)}, \ldots, \frac{\partial}{\partial x^n}|_{\varphi(p)}$$

form a basis for $T_{\varphi(p)}\mathbb{R}^n$ (this was the content of Corollary 1). Therefore, the pre-images of these vectors under $d\varphi_p$ form a basis for T_pM. We continue to use the notation $\frac{\partial}{\partial x^i}|_p$ where it is implied that

$$\frac{\partial}{\partial x^i}|_p = d\varphi_p^{-1}\left(\frac{\partial}{\partial x^i}|_{\varphi(p)}\right).$$

Given a...

- $p \in M$;
  a chart (U, φ) around p; and
- a smooth function $f \in C^\infty(U)$

we can write the coordinate representations for p and f:

- $\hat{p} = \varphi(p) = (p_1, \ldots, p_n)$
  $\hat{f} = f \circ \varphi^{-1}$

and we see that the vector $\frac{\partial}{\partial x^i}|_p$ acts on f via

$$\frac{\partial}{\partial x^i}|_p \cdot f = d\varphi^{-1}_{\varphi(p)}\left(\frac{\partial}{\partial x^i}|_{\varphi(p)}\right) \cdot f = \frac{\partial}{\partial x^i}|_{\varphi(p)}\left(f \circ \varphi^{-1}\right) = \frac{\partial \hat{f}}{\partial x^i}(\hat{p})$$

where we have used the definition of the differential (definition 21) and the action of ordinary partial differentiation in Euclidean space. This shows that $\frac{\partial}{\partial x^i}|_p$ is the derivation that takes the i-th partial derivative of (the coordinate representation of) a function f at p. The $\frac{\partial}{\partial x^i}|_p$'s are called coordinate vectors at p associated with the given coordinate system. Now any tangent vector $v \in T_pM$ can be written as

$$v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}|_p.$$

The ordered basis $\{\frac{\partial}{\partial x^1}|_p, \ldots, \frac{\partial}{\partial x^n}|_p\}$ is called a coordinate basis for T_pM. The numbers $(v^1, \ldots, v^n)$ are called the components of v with respect to the coordinate basis.

Given a tangent vector v how can we calculate its components in a basis? If a tangent vector v is known we can calculate its components in the basis by using the above formulas for the coordinate functions $f = x^j$:

$$v(x^j) = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}|_p(x^j) = \sum_{i=1}^{n} v^i \frac{\partial x^j}{\partial x^i}|_p = v^j$$

What about differentials? Start with the simplest case of a map $F : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^m$. We'll compute the matrix of the differential of F at a point p, $dF_p : T_p\mathbb{R}^n \to \mathbb{R}^m$, with respect to the standard coordinate basis. Let $(x^1, \ldots, x^n)$ be coordinates on $\mathbb{R}^n$ and $(y^1, \ldots, y^m)$ on $\mathbb{R}^m$. Then for any $f \in C^\infty(V)$

$$dF_p\left(\frac{\partial}{\partial x^i}\right) \cdot f = \frac{\partial}{\partial x^i}(f \circ F) = \frac{\partial f}{\partial y^j}(F(p)) \cdot \frac{\partial F^j}{\partial x^i}(p) = \left(\frac{\partial F^j}{\partial x^i}(p)\frac{\partial}{\partial y^j}|_{F(p)}\right) \cdot f$$

from which we read off

$$dF_p\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial F^j}{\partial x^i}(p)\frac{\partial}{\partial y^j}|_{F(p)}.$$

The corresponding matrix expression is simply the Jacobian matrix of F

$$dF_p = \begin{bmatrix} \frac{\partial F^1}{\partial x^1}(p) & \dots & \frac{\partial F^1}{\partial x^n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial F^m}{\partial x^1}(p) & \dots & \frac{\partial F^m}{\partial x^n}(p) \end{bmatrix}$$

Now consider a map $F : M \to N$ and pick coordinate charts (U, φ) on M containing p and (V, ψ) on N containing $F(p)$. Let $\hat{p} = \varphi(p)$ be the coordinate representation of p; the coordinate representation of F is given by

$$\hat{F} = \psi \circ F \circ \varphi{-1} : \varphi(U) \to \psi(V).$$

From the above calculation $d\hat{F}_{\hat{p}}$ is the Jacobian matrix of $\hat{F}$ at $\hat{p}$. But how exactly does this relate to dF_p? To relate $d\hat{F}_{\hat{p}}$ to dF_p we use $F \circ \varphi^{-1} = \psi^{-1} \circ \hat{F}$ which implies

$$dF_p \circ d\varphi_{\hat{p}}^{-1} = d\psi_{\hat{F}(\hat{p})}^{-1} \circ d\hat{F}_{\hat{p}}$$

and the coordinate basis of T_pM and $T_{F(p)}N$ as the pre-images of the coordinate bases in $\mathbb{R}^*$:

$$\frac{\partial}{\partial x^i}|_p = d\varphi_{\hat{p}}^{-1} \frac{\partial}{\partial x^i}|_{\hat{p}}, \quad \frac{\partial}{\partial y^i}|_{F(p)} = d\psi_{\hat{F}(\hat{p})}^{-1} \frac{\partial}{\partial y^i}|_{\hat{F}(\hat{p})}$$

Now

$$dF_p \left(\frac{\partial}{\partial x^i}|_p \right) = dF_p \left(d\varphi_{\hat{p}}^{-1} \frac{\partial}{\partial x^i}|_{\hat{p}} \right) = d\psi_{\hat{F}(\hat{p})}^{-1} \left(d\hat{F}_{\hat{p}} \left(\frac{\partial}{\partial x^i}|_{\hat{p}} \right) \right)$$

and recalling that $d\hat{F}_{\hat{p}} \left(\frac{\partial}{\partial x^i}|_{\hat{p}} \right) = \frac{\partial \hat{F}^j}{\partial x^i}(\hat{p}) \frac{\partial}{\partial y^j}|_{\hat{F}(\hat{p})}$ we continue as

$$\begin{aligned} dF_p \left(\frac{\partial}{\partial x^i}|_p \right) = dF_p \left(d\varphi_{\hat{p}}^{-1} \frac{\partial}{\partial x^i}|_{\hat{p}} \right) &= d\psi_{\hat{F}(\hat{p})}^{-1} \left(d\hat{F}_{\hat{p}} \left(\frac{\partial}{\partial x^i}|_{\hat{p}} \right) \right) \\ &= d\psi_{\hat{F}(\hat{p})}^{-1} \left(\frac{\partial \hat{F}^j}{\partial x^i}(\hat{p}) \frac{\partial}{\partial y^j}|_{\hat{F}(\hat{p})} \right) \\ &= \frac{\partial \hat{F}^j}{\partial x^i}(\hat{p}) \cdot d\psi_{\hat{F}(\hat{p})}^{-1} \frac{\partial}{\partial y^j}|_{\hat{F}(\hat{p})} \\ &= \frac{\partial \hat{F}^j}{\partial x^i}(\hat{p}) \frac{\partial}{\partial y^j}|_{F(p)} \end{aligned}$$

That is, dF_p is represented by the Jacobian matrix of the coordinate representative of F. The definition we started with, involving derivations, is just a coordinate free definition of the Jacobian matrix.

In geometry the differential of a map is sometimes called the **tangent map**, the **derivative**, or the **push-forward** - because it pushes tangent vectors forward. Alternative notation includes

$$F'(p) \quad DF \quad DF(p) \quad F_* \quad TF \quad T_pF$$

How do we change coordinates? Suppose (U, φ) and (V, ψ) are two charts on M and $p \in U \cap V$. Let x^i be the coordinates of φ and $\tilde{x}^i$ those of ψ. Any tangent vector can be expressed in either basis: $\frac{\partial}{\partial x^i}|_p$ or $\frac{\partial}{\partial \tilde{x}^i}|_p$. But how are the two expressions related? Well, if the transition map $\psi \circ \varphi^{-1}$ is the change of coordinate map between $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^n$ then it's differential $d\psi \circ \varphi_p^{-1}$ maps basis vectors of $T_{\varphi(p)}\mathbb{R}^n$ to basis vectors $T_{\psi(p)}\mathbb{R}^n$. This will give the change of basis map on T_pM.

Write the transition map $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ as

$$\psi \circ \varphi^{-1}(x) = (\tilde{x}^1(x), \ldots, \tilde{x}^n(x))$$

Then the differential is given by the Jacobian matrix (which can be computed from the definition of the differential and the chain rule, as above):

$$d(\psi \circ \varphi^{-1})_{\varphi(p)}(\frac{\partial}{\partial x^i}|_{\varphi(p)}) = \frac{\partial \tilde{x}^j}{\partial x^i}(\varphi(p))\frac{\partial}{\partial \tilde{x}^j}|_{\psi(p)}.$$

Now using the definition of the coordinate vectors and property 2. of differentials

$$\begin{aligned}
\frac{\partial}{\partial x^i}|_p = d\varphi^{-1}_{\varphi(p)}\left(\frac{\partial}{\partial x^i}|_{\varphi(p)}\right) &= d\psi^{-1}_{\psi(p)} d(\psi \circ \varphi)^{-1}_{\varphi(p)}\left(\frac{\partial}{\partial x^i}|_{\varphi(p)}\right) \\
&= d\psi^{-1}_{\psi(p)}\left(\frac{\partial \tilde{x}^j}{\partial x^i}(\varphi(p))\frac{\partial}{\partial \tilde{x}^j}|_{\psi(p)}\right) \\
&= \frac{\partial \tilde{x}^j}{\partial x^i}(\varphi(p)) \cdot d\psi^{-1}_{\psi(p)}\left(\frac{\partial}{\partial \tilde{x}^j}|_{\psi(p)}\right) \\
&= \frac{\partial \tilde{x}^j}{\partial x^i}(\varphi(p)) \cdot \frac{\partial}{\partial \tilde{x}^j}|_p
\end{aligned}$$

Applying the above formula to a vector

$$v = v^i \frac{\partial}{\partial x^i} = \tilde{v}^j \frac{\partial}{\partial \tilde{x}^j}$$

we see that the components are related by

$$\tilde{v}^j = \frac{\partial \tilde{x}^j}{\partial x^i} v^i$$

The transition map between standard and polar coordinates on $\mathbb{R}^2$ is given by

$$(x, y) = (r\cos\theta, r\sin\theta)$$

Let p bet the point in $\mathbb{R}^2$ whose polar coordinate representation is $(r, \theta) = (2, \frac{\pi}{2})$ and $v \in T_p\mathbb{R}^2$ whose polar coordinate representation at p is

$$v = 3\frac{\partial}{\partial r}|_p - \frac{\partial}{\partial \theta}|_p$$

Find the cartesian representation of v at p. We worked out the change of coordinate formula as

$$\frac{\partial}{\partial x^i}|_p = \frac{\partial \tilde{x}^j}{\partial x^i}(\varphi(p)) \cdot \frac{\partial}{\partial \tilde{x}^i}|_p$$

Applying this formula to our situation we have

$$\frac{\partial}{\partial r}|_p = \frac{\partial x}{\partial r}(2, \frac{\pi}{2})\frac{\partial}{\partial x}|_p + \frac{\partial y}{\partial r}(2, \frac{\pi}{2})\frac{\partial}{\partial y}|_p = \frac{\partial}{\partial y}|_p$$

$$\frac{\partial}{\partial \theta}|_p = \frac{\partial x}{\partial \theta}(2, \frac{\pi}{2})\frac{\partial}{\partial x}|_p + \frac{\partial y}{\partial \theta}(2, \frac{\pi}{2})\frac{\partial}{\partial y}|_p = -2\frac{\partial}{\partial x}|_p$$

So $v = 3\frac{\partial}{\partial y}|_p + 2\frac{\partial}{\partial x}|_p$.

Alternatively we can view the last set of formulas as a linear coordinate change map C which maps the polar basis vectors to the cartesian basis vectors:

$$C(\frac{\partial}{\partial r}|_p) = \frac{\partial x}{\partial r}(2, \frac{\pi}{2})\frac{\partial}{\partial x}|_p + \frac{\partial y}{\partial r}(2, \frac{\pi}{2})\frac{\partial}{\partial y}|_p = \frac{\partial}{\partial y}|_p$$

$$C(\frac{\partial}{\partial \theta}|_p) = \frac{\partial x}{\partial \theta}(2, \frac{\pi}{2})\frac{\partial}{\partial x}|_p + \frac{\partial y}{\partial \theta}(2, \frac{\pi}{2})\frac{\partial}{\partial y}|_p = -2\frac{\partial}{\partial x}|_p$$

The image of each polar basis vector is a column of a matrix corresponding to C.

$$C = \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}$$

And applying this to $v = (3, -1)^T$ gives the same result.

The Tangent Bundle

It's useful to consider the set of all tangent vectors at all points of a manifold.

Definition 3.7. *Given a smooth manifold, with or without boundary, the* **tangent bundle** *of M, denoted by TM, is the disjoint union of the tangent spaces at all points of M*

$$TM = \coprod_{p \in M} T_pM$$

We usually write an element of this disjoint union as an ordered pair (p, v), with $p \in M$ and $V \in T_pM$. We have a (fairly natural) projection

$$\pi : TM \to M$$

which sends each element in TM to the point p at which the corresponding tangent vector is based

$$\pi(p, v) = p$$

For $M = \mathbb{R}^n$ we saw that derivations at p, T_pM, were isomorphic to geometric tangent vectors $\mathbb{R}^n_p = \{(p,v) \in \{p\} \times \mathbb{R}^n\}$ so that the tangent bundle of $\mathbb{R}^n$ can be canonically identified with the union of its geometric tangent spaces, which is just the cartesian product of $\mathbb{R}^n$ with itself

$$TM = \coprod_{p\in\mathbb{R}^n} T_p\mathbb{R}^n \cong \coprod_{p\in\mathbb{R}^n} \mathbb{R}^n_p = \coprod_{p\in\mathbb{R}^n} \{p\} \times \mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$$

CAUTION: In general a tangent bundle cannot be identified with a cartesian product because there is usually no way to identify tangent spaces at different points - the above is particular to $\mathbb{R}^n$.

It turns out that the tangent space is much more than a disjoint union of vector spaces: it is, itself, a smooth manifold.

Theorem 3.9. *For any smooth n-manifold M the tangent bundle TM has a natural topology and smooth structure that make it into a smooth $2n$-dimensional manifold. With respect to this structure the projection $\pi : TM \to M$ is smooth.*

○ *Proof.* Pick a chart (U, φ) for M and note that $\pi^{-1}(U) \subseteq TM$ is the set of all tangent vectors to M at all points of U. Let $(x^1, \ldots, x^n)$ denote the coordinates of φ and define a map $\tilde{\varphi} : \pi^{-1}(U) \to \mathbb{R}^{2n}$ by

$$\tilde{\varphi}\left(v^i \frac{\partial}{\partial x^i}|_p\right) = (x^1(p), \ldots, x^n(p), v^1, \ldots, v^n)$$

The image of $\tilde{\varphi}$ is $\varphi(U) \times \mathbb{R}^n$ which is an open set in $\mathbb{R}^{2n}$. It's a bijection with inverse

$$\tilde{\varphi}^{-1}(x^1(p), \ldots, x^n(p), v^1, \ldots, v^n) = v^i \frac{\partial}{\partial x^i}|_{\varphi^{-1}(x)}$$

Now take two charts (U, φ) and (V, ψ) for M and let $(\pi^{-1}(U), \tilde{\varphi})$ $(\pi^{-1}(V), \tilde{\psi})$ be the corresponding charts on TM. The sets

$$\tilde{\varphi}(\pi^{-1}(U) \cap \pi^{-1}(V)) = \varphi(U \cap V) \times \mathbb{R}^n$$
$$\tilde{\psi}(\pi^{-1}(U) \cap \pi^{-1}(V)) = \psi(U \cap V) \times \mathbb{R}^n$$

are open in $\mathbb{R}^{2n}$ and the transition map $\tilde{\psi} \circ \tilde{\varphi}^{-1} : \varphi(U \cap V) \times \mathbb{R}^n \to \psi(U \cap V) \times \mathbb{R}^n$ can be written using the change of coordinates formula for vectors as

$$\tilde{\psi} \circ \tilde{\varphi}^{-1}(x^1(p), \ldots, x^n(p), v^1, \ldots, v^n) = (\tilde{x}^1(x), \ldots, \tilde{x}^n(x), \frac{\partial \tilde{x}^1}{\partial x^j} v^j, \ldots, \frac{\partial \tilde{x}^n}{\partial x^j} v^j)$$

which is smooth. Choosing a countable cover $\{U_i\}$ of charts of M we obtain a countable cover $\{\pi^{-1}(U_i)\}$ of charts of TM. If (p, v) and (q, w) are two tangent vectors in TM belonging to different tangent spaces then there exist disjoint neighbourhoods U of p and V of q so that $\pi^{-1}(U)$ and $\pi^{-1}(V)$ are disjoint neighbourhoods containing (p, v) and (q, w), respectively. This is the Hausdorff condition. So TM is a smooth manifold. To see that π is smooth observe that its coordinate representation with respect to the charts (U, φ) $(\pi^{-1}(U), \tilde{\varphi})$ is $\pi(x, v) = x$. □

The above proof showed that the tangent bundle TM is locally diffeomorphic to an open set in $\mathbb{R}^{2n}$ of the form $U \times \mathbb{R}^n$. However, it is very rare for the tangent bundle to be even globally homeomorphic to $M \times \mathbb{R}^n$.

By putting together all the differentials of a smooth map F on M, we obtain a globally defined map between tangent bundles, called the **global differential** and denoted by

$$dF : TM \to TN$$

This is the map whose restriction to each tangent space $T_pM \subset TM$ is dF_p. The smooth structure on TM makes dF into a smooth map

Proposition 3.10. *If $F : M \to N$ is a smooth map then its global differential $dF : TM \to TN$ is a smooth map*

Proof. Using the coordinate representation for dF_p the coordinate representation for dF in coordinates for TM and TN is given by

$$dF(x^1, \ldots, x^n, v^1, \ldots, v^n) = (F^1(x), \ldots, F^n(x), \frac{\partial F^1}{\partial x^j} v^j, \ldots, \frac{\partial F^n}{\partial x^j} v^j)$$

which is smooth because F is. □ □

We also have exactly the same properties for global differentials as we do for differentials.

Proposition 3.11. *Let M, N, and P be smooth manifolds with or without boundary, let $F : M \to N$ and $G : N \to P$ be smooth maps..*

a. *$dF : TM \to TN$ is linear;*

b. *$d(G \circ F) = dG \circ dF : TM \to TP$;*

c. *$d(Id_M) = Id_{TM} : TM \to TM$; and*

d. *If F is a diffeomorphism then $dF : TM \to TN$ is also a diffeomorphism and $d\left(F^{-1}\right) = dF^{-1}$.*

Velocity Vectors of Curves in Manifolds

A curve in a smooth manifold M is a continuous Map

$$\gamma : J \subseteq \mathbb{R} \to M$$

The velocity vector of γ at t_o is the vector

$$\gamma'(t_o) = d\gamma\left(\frac{d}{dt}|_{t_o}\right) \in T_{\gamma(t_o)}M.$$

This tangent vector acts on functions by

$$\gamma'(t_o)\cdot f = d\gamma\left(\frac{d}{dt}|_{t_o}\right)\cdot f = \frac{d}{dt}|_{t_o}(f\circ\gamma) = (f\circ\gamma)'(t_0).$$

That is, $\gamma'(t_o)$ is the derivation at $\gamma(t_o)$ obtained by taking the derivative of a function along it. If (U,φ) is a smooth chart with coordinates (x^i) we can write the coordinate representation for γ as

$$\gamma(t) = (\gamma^1(t),\ldots,\gamma^n(t))$$

and for its differential

$$\gamma'(t_o) = \frac{d\gamma^i}{dt}(t_o)\frac{\partial}{\partial x^i}|_{\gamma(t_o)}$$

So $\gamma'(t_o)$ is essentially the same formula as it would be in Euclidean space. Recall that our alternative definition of tangent vector was that it is an equivalence class of curves. This next proposition shows that every tangent vector, as a derivation, is the velocity vector of some smooth curve.

Proposition 3.12. *Suppose M is a smooth manifold and $p\in M$. Every $v\in T_pM$ is the velocity vector of some smooth curve in M.*

Proof. Let (U,φ) be a chart centered at p ($\varphi(p)=0$) and write $v = v^i\frac{\partial}{\partial x^i}|_p$ Let $\gamma : (-\epsilon,\epsilon)\to M$ be the curve whose coordinate representation is given by

$$\gamma(t) = (tv^1, ...tv^n)$$

This is a smooth curve passing through p at $t=0$ with $\gamma'(0) = v^i\frac{\partial}{\partial x^i}|_p = v$. □

Velocity vectors also behave well under composition with smooth maps

Proposition 3.13. *Let $F : M\to N$ be a smooth map and let $\gamma : I\to M$ be a smooth curve. For any $t_o\in I$ the velocity at $t=t_o$ of the composite curve $F\circ\gamma : I\to N$ is given by*

$$(F\circ\gamma)'(t_o) = dF(\gamma'(t_o))$$

Proof.

$$(F\circ\gamma)'(t_o) = d(F\circ\gamma)\left(\frac{d}{dt}|_{t_o}\right) = dF\circ d\gamma\left(\frac{d}{dt}|_{t_o}\right) = dF(\gamma'(t_o))$$

□

The last two Propositions actually give an alternative way of computing the differential of a smooth map. We can compute $dF_p(v)$ for any $v\in T_pM$ by choosing a smooth curve $\gamma(t)$ whose velocity vector is v and then compute $(F\circ\gamma)'(0)$.

Proposition 3.14. *Let $F : M\to N$ be a smooth map, $p\in M$ and $v\in T_pM$; then*

$$dF(v) = (F\circ\gamma)'(0)$$

where $\gamma : I\to M$ is any smooth curve with $\gamma(0)=p$ and $\gamma'(0)=v$.

Vector Fields on Manifolds

A vector field on a smooth manifold M is a smooth map

$$X : M \to TM$$

which assigns to each point $p \in M$ a point $(p, v) \in TM$ and satisfies

$$\pi \circ X = Id_M$$

We usually write the value of a vector field X at p by $X_p \in T_pM$. Vector fields on manifolds are visualized in exactly the same way that we visualize vector fields on Euclidean space: as arrows at every point which are tangent to M. Using a coordinate chart (U, φ), any vector field at a point p can be written in terms of component functions $X^i : U \to \mathbb{R}$ and the coordinate basis vectors:

$$X_p = X^i(p) \frac{\partial}{\partial x^i}|_p$$

Alternatively, recall that coordinates (x^i) via a chart (U, φ) on a smooth manifold M induce coordinates on its tangent bundle TM so that we can write a vector field in coordinates as

$$\begin{aligned} \Phi : \pi^{-1}(U) &\to \mathbb{R}^n \times \mathbb{R}^n \\ (p, v) &\mapsto (x^1(p), \ldots, x^n(p), X^1(p), \ldots, X^n(p)) \end{aligned}$$

For each i we get a smooth coordinate vector field $\frac{\partial}{\partial x^i}$ whose value at a point p is $\frac{\partial}{\partial x^i}|_p$ with coordinate representation

$$(x^1(p), \ldots, x^n(p), 0, \ldots, 1, 0, \ldots, 0) \in \varphi(U) \times \mathbb{R}^n$$

Observe that the coordinate vector fields $\frac{\partial}{\partial x^1}, \ldots \frac{\partial}{\partial x^n}$ form a basis for $\mathbb{R}^n$ at every point $\varphi(p) \in \varphi(U) \subset \mathbb{R}^n$. We say that the coordinate vector fields form a local frame for M on the open set U, called a **coordinate frame**. More generally, any collection of vector fields $X_1, \ldots X_n$ defined on an open set U whose coordinate representations in a chart for TM are linearly independent at every base-point are said to form a **local frame** for M. If $U = M$ then the frame is a **global frame**.

An essential property of vector fields is that they define derivations on the space the of smooth functions over M. If $X \in \mathcal{X}(M)$ and f is a smooth real-valued function on M we obtain a new function $Xf : M \to \mathbb{R}$ defined by

$$(Xf)(p) = X_p f$$

Indeed, writing the expression Xf in coordinates and observing that the component functions of X are smooth verifies that Xf is a smooth function:

$$Xf(x) = \left(X^i(x) \frac{\partial}{\partial x^i}|_x \right) f = X^i(x) \frac{\partial f}{\partial x^i}(x)$$

In particular, a smooth vector field defines a map

$$X : C^\infty(M) \to C^\infty(M)$$
$$f \mapsto Xf$$

The map is linear over $\mathbb{R}$ and the product rule for tangent vectors translates into the product rule for vector fields:

$$X(f \cdot g) = f \cdot Xg + g \cdot Xf$$

In general a map $X : C^\infty(M) \to C^\infty(M)$ is called a derivation if it is linear and satisfies the product rule. As in the case of tangent vectors, a map D is a derivation if and only if it is of the form $Df = Xf$ for some smooth vector field X.

A vector field V along a curve $\gamma : I \subseteq \mathbb{R} \to M$ is a smooth map that associates to each point on $\gamma(t)$ a tangent vector $V(t) \in T_{\gamma(t)}M$. We say V is induced by a vector field $X \in \mathcal{X}(M)$ if $V(t) = X(\gamma(t))$.

The Lie Bracket

Let X and Y be smooth vector fields on a manifold M. Given a smooth function $f : M \to \mathbb{R}$ we can apply X to f to obtain another smooth function Xf. Similarly we can Y to this function and obtain yet another smooth function $YXf = Y(Xf)$. However, the operation $f \mapsto YXf$ generally does not satisfy the product rule so that YX is generally not a derivation and hence not a vector field. For example, consider $X = \frac{\partial}{\partial x}$ and $Y = x\frac{\partial}{\partial y}$ on $\mathbb{R}^2$ and $f(x, y) = x$, $g(x, y) = y$. But we can apply the same vector fields X and Y in the opposite order and obtain a (usually) different function XYf. Applying both of these operators to f and subtracting defines an operator

$$[X, Y] : C^\infty(M) \to C^\infty(M)$$

called the Lie bracket of X and Y, defined by

$$[X, Y]f = XYf - YXf$$

The key point is that this operator IS a vector field because it satisfies the product rule (check it!). Given coordinates (x^i) on M, the Lie bracket $[X, Y]$ has the coordinate expression

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}$$

... prove it. This expression immediately implies that the coordinate vector fields $\frac{\partial}{\partial x^i}$ satisfy

$$[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$$

for all i and j. The Lie bracket satisfies the following identities for all X, Y, and $Z \in \mathcal{X}(M)$

a. Bilinearity: for a and $b \in \mathbb{R}$

$$[aX + by, Z] = a[X, Z] + b[Y, Z]$$
$$[Z, aX + bY] = a[Z, X] + b[Z, Y]$$

b. Antisymmetry: [X,Y]=-[Y,X]

c. The Jacobi Identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

d. For any f and $g \in C^{\infty}(M)$

$$[fX, gY] = fg[X, Y] + (fXg)Y - (gYf)X$$

Finally, suppose $F : M \to N$ is a smooth map between manifolds and suppose that $X_1, X_2 \in \mathcal{X}(M)$ and $Y_1, Y_2 \in \mathcal{X}(N)$ with

$$dF_p \cdot X_1 = Y_1(F(p))$$
$$dF_p \cdot X_2 = Y_2(F(p));$$

then

$$dF_p[X_1, X_2] = [Y_1, Y_2](F(p))$$

- this follows directly from the definition of the differential and the Lie bracket.

Chapter 4

Riemannian Metrics

Inner Products

Basically everything we know about $\mathbb{R}^n$ is, or can be, derived from the dot product:

$$v \cdot w = \sum_{i=1}^{n} v^i w^i$$

for vectors $v = (v^1, ..., v^n)$ and $w = (w^1, ..., w^n)$. The dot product admits a natural generalization to any vector space V, called an inner product which is a map

$$\begin{aligned} \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} \\ (v, w) \mapsto \langle v, w \rangle \end{aligned}$$

and satisfies the following properties for all $u,\ v,\ w \in V$ and $a,\ b \in \mathbb{R}$:

a. SYMMETRY: $\langle v, w \rangle = \langle w, v \rangle$

b. LINEARITY: $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$

c. POSITIVE DEFINITENESS: $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$.

Inner products allow us to measure lengths and angles. The length or norm of a vector $v \in V$ is given by

$$\|v\| = \sqrt{(v, v)}$$

An inner product is completely determined by knowledge of the lengths of all vectors:

Lemma 4.1. *Suppose $\langle \cdot, \cdot \rangle$ is an inner product on a vector space V. Then, for all $v,\ w \in V$*

$$\langle v, w \rangle = \frac{1}{4} \left(\langle v + w, v + w \rangle - \langle v - w, v - w \rangle \right)$$

The angle between two non-zero vectors v and $w \in V$ is is defined by the unique $\theta \in [0, \pi]$ satisfying

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}$$

Two vectors v and $w \in V$ are orthogonal if $\langle v, w \rangle = 0$ which means either their angle is $\frac{\pi}{2}$ or one of the vectors is zero. If $S \subset V$ is a linear subspace then

$S^\perp \subset V$ is the set of all vectors orthogonal to every vector in S called the orthogonal complement of S. Vectors $v_1, \ldots, v_k$ are called orthonormal if they have length 1 and are all mutually orthogonal: $\langle v_i, v_j \rangle = \delta_{i,j}$. Let $\{v_1, \ldots, v_n\}$ be a basis for a vector space V. If $\langle \cdot, \cdot \rangle$ is an inner product on V then there is an associated symmetric, positive definite matrix representing $\langle \cdot, \cdot \rangle$ whose entries are given by

$$A_{ij} = \langle v_i, v_j \rangle$$

In fact, we have the principal axis Theorem which says that there exists an ordered basis for V in which A is diagonal (see Knapp, for example).

Riemannian Metrics

Definition 4.1. *A Riemannian metric g on a smooth manifold M is a correspondence which associates to each $p \in M$ an inner product*

$$g_p(\cdot, \cdot)$$

on the tangent space T_pM which varies smoothly with p.

What does this mean? In local coordinates (U, φ), a metric is represented by a symmetric, positive definite matrix

$$(g_{ij}(x))_{i,j=1,\ldots,n} =: G(x)$$

whose coefficients depend smoothly on x. The product of two tangent vectors v and w in T_pM with coordinate representations $v = v^i \frac{\partial}{\partial x^i}|_{\varphi(p)}$, $w = w^i \frac{\partial}{\partial x^i}|_{\varphi(p)}$ is given by

$$g_p(v, w) = g_{ij}(\varphi(p))v^i w^i = \mathbf{v}^T G(\varphi(p))\mathbf{w}$$

In particular, the components of the metric are given by

$$g_{ij}(\varphi(p)) = g_p(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$$

Theorem 4.2. *Every smooth manifold can be equipped with a Riemannian metric.*

Proof. The proof of this Theorem requires the use of a partition of unity which is a way of blending together local smooth objects to create global smooth objects. Let M be a smooth manifold with atlas $\mathcal{A} = \{(U_\alpha, \varphi_\alpha\}_{\alpha \in A}$ indexed by a set A. A partition of unity subordinate to $\mathcal{A}$ is an indexed family of smooth functions $\psi_\alpha : M \to \mathbb{R}$ with the following properties:

a. $0 \leq \psi_\alpha(x) \leq 1$ for all $\alpha \in A$ and all $x \in M$;

b. $\mathrm{supp}\psi_\alpha \subseteq U_\alpha$ for all $\alpha \in A$;

c. Every point in M has a neighbourhood which intersects $\mathrm{supp}\psi_\alpha$ for only finitely many values of α

d. $\sum_{\alpha \in A} \psi_\alpha(x) = 1$ for all $x \in M$

The finiteness condition in 3. ensures that the sum in 4. only has finitely many terms so there are no convergence issues. That partitions of unity exist is the contents of Lemma 2.23 in Lee - we will assume this for now. So suppose we have an atlas $\mathcal{A}$ with a partition of unity subordinate to $\mathcal{A}$. For v and w in T_pM and $\alpha \in A$ with $p \in U_\alpha$ let the coordinate representations be $(v^1_\alpha, \ldots, v^n_\alpha)$ and $(w^1_\alpha, \ldots, w^n_\alpha)$. Then set

$$g_p(v, w) = \sum_{\alpha \in A \text{ with } p \in U_\alpha} \psi_\alpha(p) v^i_\alpha w^j_\alpha;$$

this defines a Riemannian metric by piecing together the Euclidean metrics of the coordinate images together which are smoothly interpolated by the partition of unity. $\square$

Riemannian metrics allow us to measure lengths, angles, and orthogonality of tangent vectors just as in Euclidean space. The length of a tangent vector will be denoted by

$$\|v\|_g = \sqrt{g(v, v)_p}$$

Definition 4.2. *A Riemannian manifold is a pair (M, g), where M is a smooth manifold and g is a specific choice of Riemannian metric.*

Isometries

Suppose (M, g) and (N, h) are Riemannian manifolds.

Definition 4.3. *An isometry from (M, g) to (N, h) is a a diffeomorphism $\eta : M \to N$ such that for each $p \in M$ and all v and $w \in T_pM$*

$$g(v, w)_p = h(d\eta_p \cdot v, d\eta_p \cdot w)_{\eta}(p).$$

We say (M, g) and (N, h) are isometric if there exists and isometry between them

The isometry condition is often written

$$\eta^* h = g$$

and reads: "the pullback of h by η to M is equal to g". Pulling back h by a diffeomorphism is equivalent to pushing forward tangent vectors by the diffeomorphism and then evaluating h on them. Compositions of isometries and inverses of isometries are also isometries so being isometric is an equivalence relation on Riemannian manifolds and you can think of an isometry as a change of coordinates between Riemannian manifolds. As with almost all branches of mathematics, a primary concern of Riemannian geometry is the study of properties of Riemannian manifolds that are preserved by isometries; that is, invariants of Riemannian manifolds.

If (M, g) and (N, h) are Riemannian manifolds then a map $\eta : M \to N$ is a local isometry if every point $p \in M$ has a neighbourhood U such that $\eta|_U$ is an isometry onto an open subset of N. A Riemannian n-manifold is said to be **flat** if it is locally isometric to a Euclidean space; that is, if every point $p \in M$ has a neighbourhood that is locally isometric to an open subset of $\mathbb{R}^n$ with its Euclidean metric.

Let w^1 and w^2 be linearly independent in $\mathbb{R}^2$. Two points z_1 and z_2 in $\mathbb{R}^2$ are equivalent if there exist integers m_1 and m_2 such that

$$z_1 - z_2 = m_1 w_1 + m_2 w_2$$

Let π be the projection mapping to its equivalence class. The 2-Torus $\mathbb{T}^2 := \pi(\mathbb{R}^2)$ is a smooth manifold with charts defined as follows: Let Δ be an open set in $\mathbb{R}^2$ not containing any equivalent points and define $U = \pi(\Delta)$ with coordinate map $\varphi = \pi^{-1}|_U$ as our set of charts. We can define a metric on $\mathbb{T}^2$ in such a way that $\mathbb{T}^2$ is locally isometric to Euclidean space: On each chart of the form $(U, \pi^{-1}|_U)$ we use the Euclidean metric on the components of $\pi^{-1}(U)$ with each component being diffeomorphic to U. For v and w in $T_q\mathbb{T}^2$ with $q \in U$ and $p \in \pi^{-1}(q)$, define a metric by

$$g^{\mathbb{T}^2}(v, w)_q = g^{\mathbb{R}^2}(d\pi_p^{-1} \cdot v, d\pi_p^{-1} \cdot w)$$

The image of $d(\pi^{-1})_p$ is $\coprod_{m_1, m_2 \in \mathbb{Z}} T_{p+m_1 w_1 + m_2 w_2}\mathbb{R}^2$ - all those vectors whose base points differ by $m_1 w_1 + m_2 w_2$ for some integers m_1 and m_2. So we just need to check that the metric we've defined is well defined (i.e. does not depend on the component of $\pi^{-1}(U)$). Since the translations $z \mapsto z + m_1 w_1 + m_2 w_2$ are Euclidean isometries the Euclidean metrics on the different disjoint components of $\pi^{-1}(U)$ (which are obtained from each other via such translations) yield the same metric on U. Indeed, suppose T is a translation by $m = (m_1, m_2)$ and $p' = p + m$. Then $T(p) = p'$ and $\pi \circ T = \pi$ so that

$$d\pi_p = d(\pi \circ T)_p = d\pi_{p'} \circ dT_p \implies d\pi_p^{-1} = dT_p^{-1} \circ d\pi_{p'}^{-1}$$

and

$$\begin{aligned} g^{\mathbb{R}^2}(d\pi_p^{-1} \cdot v, d\pi_p^{-1} \cdot w) &= g^{\mathbb{R}^2}(dT_p^{-1} \circ d\pi_{p'}^{-1} \cdot v, dT_p^{-1} \circ d\pi_{p'}^{-1} \cdot w) \\ &= g^{\mathbb{R}^2}(d\pi_{p'}^{-1} \cdot v, d\pi_{p'}^{-1} \cdot w). \end{aligned}$$

It is not true that every manifold is locally isometric to Euclidean space by virtue of a chart. The above example used very particular properties of the construction of the torus and they way these properties relate to both the chart structure used and to Euclidean isometries themselves. For example, there is no open subset of the sphere which is isometric to an open subset of $\mathbb{R}^n$ for any n - let's see why...

Although we haven't covered this we will assume that the sphere has a metric under which the distance between two points is given by the length of the great arc segment connecting them (it's the restriction of the Euclidean metric). Briefly, we'll see that in a sufficiently small neighbourhood U of a point there exists a unique length minimizing geodesic γ joining it to every other point in U. If γ is not a great arc segment then we can reflect it across the great arc segment connecting its endpoints and obtain a second geodesic of the same length which contradicts uniqueness.

It's pretty clear the a two point metric space embeds isometrically in $\mathbb{R}$. It's not as clear but any 3 point metric space $\{a, b, c\}$ embeds isometrically in $\mathbb{R}^2$:

a. Say $d(a, b) = x_0$, $d(a, c) = x_1$ and $d(b, c) = x_2$. Set $f(a) = (0, 0)$ and rotate space if necessary so that $f(b) = (0, x_o)$. Now $d(f(a), f(b)) = d(a, b)$.

b. Now we want $d(f(a), f(c)) = x_1$ so let's draw a circle of radius x_1 around $f(a)$; but we also want $d(f(b), f(c) = x_2$ so let's draw a circle of radius x_2 around $f(b)$.

c. The intersection point of these two circles should be where we set $f(c)$. But do they intersect?

d. Yes! The triangle inequality says $x_1 + x_2 \geq x_0$ so the sum of the radii is greater than the distance between the two centres.

What's surprising is that there are 4-point metric spaces that don't embed in any Euclidean space of any dimension. Here's one: start with a 3-point metric space $Y = \{a, b, c\}$ with $d(a, b) = d(a, c) = d(b, c) = 2L > 0$. The image of Y under and isometric embedding is an equilateral triangle with side lengths $2L$

a. Add to Y a point x to get a new metric space X and set $d(b, x) = d(c, x) = L$ but don't specify $d(a, x)$ yet!

b. Any isometric embedding f from X to a euclidean space necessarily maps x to the midpoint M of the line joining $f(b)$ and $f(c)$.

c. Since the Euclidean distance between $f(a)$ and M is $\sqrt{3}L$ the isometric character of f forces $d(a, x) = \sqrt{3}L$.

d. So if we specify $d(a,x)$ as something different in accordance with the triangle inequality then we have constructed a proof by contradiction for the non-existence of an isometric embedding between X and any Euclidean space. Take, for example, $d(a,x) = 2L$.

HOWEVER! With the above choice and $L = \frac{\pi}{4}$, X does embed isometrically in $S^2 \subset \mathbb{R}^3$.

a. Send a to the North pole $N = (0,0,1)$, b to $B = (1,0,0)$, c to $C = (0,1,0)$ and x to $X = \frac{1}{\sqrt{2}}(1,1,0)$.

b. This shows that no portion of the sphere that contains the points N, B, C, and X can be isometrically embedded in any Euclidean space.

We can adapt the above argument to show that no Neighbourhood of the sphere is isometric to an open set of any Euclidean space. Let N be the north pole and choose nearby points P and Q which form an equilateral triangle of side length 2θ so that the angle made by the position vectors with one another at $\mathcal{O}$ is 2θ. Let M be the mid-point of the (short) great circular arc joining P and Q and call the length of the arc joining M with N ϕ so that the angle made by the position vector of N with that of M is ϕ. We've now specified a 4-point metric space whose distances are proportionally equivalent to the example we constructed earlier and so if there is an isometric embedding of these points in a Euclidean space then $\phi = \sqrt{3}\theta$. We'll now show that this can't happen... First, the position vectors of P and Q satisfy

$$\|\mathbf{p}+\mathbf{q}\|^2 = \|\mathbf{p}\|^2 + \mathbf{p}\cdot\mathbf{q} + \|\mathbf{q}\|^2 = 2(1+\cos 2\theta) = 4\cos^2\theta$$

and therefore the angle ϕ satisfies

$$\cos\phi = \mathbf{m}\cdot\mathbf{n} = \frac{\mathbf{p}+\mathbf{q}}{\|\mathbf{p}+\mathbf{q}\|}\cdot\mathbf{n} = \frac{\cos(2\theta)}{\cos\theta}.$$

Lemma 4.3. *If the real numbers a and b satisfy $a^2 + b^2 = 1$ and $g : \mathbb{R} \to \mathbb{R}$ is given by*

$$g(s) = \cos(s) - \cos(as)\cos(bs)$$

then $g'''(0) = g''(0) = g'(0) = g(0) = 0$ and $g''''(0) = -4a^2b^2$.

Proof. It's a calculation so you do it. □

It follows that there exists a $\delta > 0$ such that

$$g(2\theta) < g(0) = 0$$

for $0 < \theta < \delta$ - look at the Taylor series (which is valid). Picking $a = \frac{1}{2}$ and $b = \frac{\sqrt{3}}{2}$ we get

$$\cos(2\theta) - \cos(\theta)\cos(\sqrt{3}\theta) = g(2\theta) < 0$$

so that

$$\cos(\phi) = \frac{\cos(2\theta)}{\cos(\theta)} < \cos(\sqrt{3}\theta)$$

which implies $\phi > \sqrt{3}\theta$. That is, there is no isometric embedding of the above four points in any Euclidean space.

Riemannian Submanifolds and Induced Metrics

Every submanifold of a Riemannian manifold automatically inherits a Riemannian metric. Before stating the key fact let's recall what an immersion is:

Definition 4.4. *A smooth map $F : M \to N$ is an immersion if its differential $dF_p : T_pM \to T_{F(p)}N$ is injective at every point $p \in M$.*

Any diffeomorphism is an immersion.

The smooth map $X : \mathbb{R}^2 \to \mathbb{R}^3$

$$X(x, y) = ((2 + \cos 2\pi x)\cos 2\pi y, (2 + \cos 2\pi x)\sin 2\pi y, \sin 2\pi x)$$

is a smooth immersion of $\mathbb{R}^2$ into $\mathbb{R}^3$ whose image is a donut.

$$DX = \begin{bmatrix} -2\pi\sin(2\pi x)\cos(2\pi y) & -2\pi(2 + \cos(2\pi x))\sin(2\pi y) \\ -2\pi\sin(2\pi x)\sin(2\pi y) & 2\pi(2 + \cos(2\pi x))\cos(2\pi y) \\ 2\pi\cos(2\pi x) & 0 \end{bmatrix}$$

is injective since the first two rows are linearly independent vectors in $\mathbb{R}^2$.

Lemma 4.4. *Suppose (N, h) is a Riemannian manifold, M is a smooth manifold, and $F : M \to N$ is a smooth map between them. Then $g := F^*h = h(dF\cdot, dF\cdot)$ is a Riemannian metric on M if and only if F is an immersion.*

Proof. If F is a smooth immersion and $g := F^*h = h(dF\cdot, dF\cdot)$ then g varies smoothly with its basepoint because dF and h do; g is bi-linear because dF and h are; g is symmetric because h is; g is positive definite because h is and is non-degenerate (i.e. $g(v, v) = 0 \Leftrightarrow v = 0$) because dF is injective. That is, g is a Riemannian metric.

Since h is a Riemannian metric $h(dF_p \cdot v, dF_p \cdot v) = 0 \Leftrightarrow dFp \cdot v = 0$ and since g is a metric $g(v, v) = 0 \Leftrightarrow v = 0$. Consequently, $dFp \cdot v = 0 \Leftrightarrow v = 0$; that is, dF_p is injective so that F is an immersion. □ □

Suppose we have a Riemannian manifold (N, h) and a smooth immersion $F : M \to N$. The metric

$$g := F^* h = h(dF\cdot, dF\cdot)$$

is called the metric induced by F. On the other hand if M already has a metric g, an immersion $F : M \to N$ satisfying $g = F^* h$ is called an isometric immersion. It should be remarked that this is not the same as an isometry which required F to be a diffeomorphism. The most important examples of induced metrics are those which occur on submanifolds via the inclusion map. Just as a reminder, the inclusion map ι from a submanifold S into the ambient manifold M $\iota : S \hookrightarrow M$ maps a point $x \in S$ to x but now considered as an element of M. The induced metric on S is the metric

$$\tilde{g} := \iota^* g = g(d\iota\cdot, d\iota\cdot)$$

and with this metric S is a Riemannian submanifold of M. In order to better understand

$$\tilde{g} := \iota^* g = g(d\iota\cdot, d\iota\cdot)$$

it's useful to view $d\iota_p$ as a linear map which identifies T_pS with its image in T_pM as inclusion at the level of tangent spaces (ι is inclusion at the point set level, $d\iota$ is inclusion at the level of linear tangent space). In other words, the induced metric $\tilde{g}$ is just the restriction of g to tangent vectors to S. The clearest example of this is the example of the sphere $\mathbb{S}^n$ which we talked about last time: the round metric $\tilde{g}$ on the sphere is the restriction of the ambient Euclidean metric to $T\mathbb{S}^n$.

Computations on any submanifold $S \subset M$ are most conveniently carried out in terms of a smooth local parametrization. This is a smooth map on an open set U

$$X : U \subset \mathbb{R}^n \to S$$

which is a diffeomorphism onto its image. Putting $V = X(U)$ and $\varphi = X^{-1}$ we see that (V, φ) is a chart for S. Since X is a diffeomorphism it's differential is an isomorphism and we can describe the tangent space to S at every point in $X(U)$. We saw earlier that every tangent vector is the velocity of some curve. Fixing all coordinates except x^i, a tangent vector in the x^i direction is given by

$$E_i = \frac{\partial X}{\partial x^i}$$

and the collection $\{E_j\}$ gives a local frame for S in coordinates on U. Since the induced metric is just the restriction of the ambient metric to tangent vectors on the submanifold, the components of the induced metric in U coordinates are given by $\tilde{G}_{ij} = g(E_i, E_j)$. Let's use this to compute the induced metric on the sphere $\mathbb{S}^2$ as a submanifold of $\mathbb{R}^3$.

Recall that a Riemannian metric can be represented in coordinates as a symmetric positive definite matrix $G(x)$ whose components vary smoothly with x. The components of the induced metric $\tilde{G}$ can be calculated from a local frame via

$$\tilde{G}_{ij} = g(E_i, E_j)$$

So to construct the induced metric on the sphere $\mathbb{S}^2$ we need to construct a suitable parameterization of it and calculate the Euclidean dot products of the associated local frame at every point p of $\mathbb{S}^2$. We can parameterize the sphere via two angles: the inclination θ and the azimuth ϕ:

$$\begin{aligned} x &= \sin\theta\cos\phi \\ y &= \sin\theta\sin\phi \\ z &= \cos\theta \end{aligned}$$

with domain $V = \{(\theta, \phi) : 0 < \phi < 2\pi,\ 0 < \theta < \pi\}$. To compute the Riemannian metric we need to compute a basis $(E_1(\theta,\phi), E_2(\theta,\phi))$ for the tangent space $T_p\mathbb{S}^2$ at $p = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Recall that each tangent vector at p is the velocity vector of some curve passing through p. For a fixed ϕ, $p(\theta,\phi)$ traces out a longitudinal line, so let's take

$$E_1(\theta,\phi) = \frac{\partial p}{\partial\theta} = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)$$

For a fixed θ, $p(\theta,\phi)$ traces out a meridional ring, so let's take

$$E_2(\theta,\phi) = \frac{\partial p}{\partial\phi} = (-\sin\theta\sin\phi, \sin\theta\cos\phi, 0)$$

Then

$$\tilde{g}_{11}(\theta,\phi) = \langle E_1, E_1\rangle = \cos^2\theta\cos^2\phi + \cos^2\theta\sin^2\phi + \sin^2\theta = 1$$

$$\tilde{g}_{12}(\theta,\phi) = \langle E_1, E_2\rangle = -\cos\theta\cos\phi\sin\theta\sin\phi + \cos\theta\sin\phi\sin\theta\cos\phi = 0$$

$$\tilde{g}_{22}(\theta,\phi) = \langle E_2, E_2\rangle = \sin^2\theta\sin^2\phi + \sin^2\theta\cos^2\phi = \sin^2\theta$$

so that

$$\tilde{G}(\theta,\phi) = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}$$

If $v = aE_1 + bE_2 \in T_p\mathbb{S}^2$ and $w = cE_1 + dE_2 \in T_p\mathbb{S}^2$ then

$$\tilde{g}(v,w) = ac + bd\sin^2\theta$$

Let H be a half plane given by $\{(r, z) : r > 0\}$ and suppose $C \subset H$ is some open curve in H. The surface of revolution determined by C is the subset $S_C \subset \mathbb{R}^3$ given by

$$S_C = \{(x, y, z) : (\sqrt{x^2 + y^2}, z) \in C\}$$

Every smooth local parametrization of C, $\gamma(t) = (a(t), b(t))$, gives a smooth parametrization of S_C of the form

$$X(t, \theta) = (a(t) \cos\theta, a(t) \sin\theta, b(t))$$

provided that t and θ are restricted to a sufficiently small neighbourhood of the plane. Now take as basis vectors for the tangent space at a point $X(t, \theta)$:

$$E_1 = \frac{\partial X}{\partial t} = (a'(t) \cos\theta, a'(t) \sin\theta, b'(t))$$
$$E_2 = \frac{\partial X}{\partial \theta} = (-a(t) \sin\theta, a(t) \cos\theta, 0).$$

The components of the induced metric are

$$\tilde{g}_{11} = \langle E_1, E_1 \rangle = (a'(t))^2 + (b'(t))^2$$
$$\tilde{g}_{12} = \langle E_1, E_2 \rangle = 0$$
$$\tilde{g}_{22} = \langle E_2, E_2 \rangle = (a(t))^2$$

so that

$$\tilde{G}(\theta, \phi) = \begin{bmatrix} (a'(t))^2 + (b'(t))^2 & 0 \\ 0 & (a(t))^2\theta. \end{bmatrix}$$

If $v = xE_1 + yE_2 \in T_{X(t,\theta)}S_C$ and $w = cE_1 + dE_2 \in T_{X(t,\theta)}S_C$ then

$$\tilde{g}(v, w) = xc\left((a'(t))^2 + (b'(t))^2\right) + yd(a(t))^2.$$

Chapter 5

Connections

Previously we looked at curves $\gamma : I \subseteq \mathbb{R} \to M$ in a manifold and saw that the velocity vector of a curve at a point $\gamma(t_o)$ is given by

$$\gamma'(t_o) = d\gamma\left(\frac{d}{dt}|_{t_o}\right) \in T_{\gamma(t_o)}M$$

whose representation in coordinates is given by

$$\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t)) \quad \gamma'(t) = (\gamma_1'(t), \ldots, \gamma_n'(t)).$$

This expression is well-defined and there are really no obstacles to calculating the velocity of a curve in a manifold M. What about acceleration? Can't we just differentiate $\gamma'(t)$ with respect to t? The problem is that to define the acceleration $\gamma''(t)$ by differentiating $\gamma'(t)$ with respect to t, we have to take a limit of difference quotients involving the vectors $\gamma'(t+h)$ and $\gamma'(t)$ but these live in completely different vector spaces:

$$\gamma'(t+h) \in T_{\gamma(t+h)}M \quad \text{and} \quad \gamma(t) \in T_{\gamma(t)}M$$

and so there is no natural way to compare them, let alone subtract them. We will see in the next example that ignoring the above problem and naively defining the acceleration as simply the second derivative can lead to answers that depend on the coordinate system chosen, which is highly undesirable when trying to make sense of shortest paths or "geodesics".

Consider the circle S^1 as a submanifold of $\mathbb{R}^2$ parameterized in cartesian coordinates by

$$\gamma(t) = (x(t), y(t)) = (\cos t, \sin t)$$

The acceleration vector written with respect to the ambient standard basis of $\mathbb{R}^2$ is given by

$$\begin{aligned}\gamma''(t) &= x''(t)\frac{\partial}{\partial x}|_{\gamma(t)} + y''(t)\frac{\partial}{\partial y}|_{\gamma(t)} \\ &= -\cos t\frac{\partial}{\partial x}|_{\gamma(t)} + -\sin t\frac{\partial}{\partial y}|_{\gamma(t)}\end{aligned}$$

However, in polar coordinates the same circle is described by $(r(t), \theta(t)) = (1, t)$ and if we try to naively compute the acceleration in the same way using the same standard basis we get

$$\begin{aligned}\gamma''(t) &= r''(t)\frac{\partial}{\partial x}|_{\gamma(t)} + \theta''(t)\frac{\partial}{\partial y}|_{\gamma(t)} \\ &= 0\end{aligned}$$

So the two expressions for acceleration are inequivalent, which begs the question: which one is the correct one?

The above discussion equally applies to vector fields on a manifold, the above being a particular case of a vector field along a curve in a manifold. At first glance it is not obvious what it means to calculate the rate of change of a vector field since vectors at different points live in spaces that are not comparable. So what we need to do is come up with a way to identify or "**connect**" tangent spaces at different points of a manifold M in order to calculate acceleration or, more generally, differentiate vector fields.

Affine Connections

We need to come up with a way of differentiating vector fields. This involves defining an operation which takes in two vector fields as arguments and outputs a new vector field. We'll start by writing the properties we would like this map to have and get a feel for what this map should look like. Once we've done this we'll prove that such a map does actually exist and interacts well with a given Riemannian metric.

Definition 5.1. *An affine connection ∇ on a smooth manifold M is a mapping*

$$\begin{aligned}\nabla : \mathcal{X}(M) \times \mathcal{X}(M) &\to \mathcal{X}(M) \\ (X, Y) &\mapsto \nabla_X Y\end{aligned}$$

which satisfies the following properties:

a. $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$

b. $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$

c. $\nabla_X(fY) = f\nabla_X Y + X(f)\cdot Y$

where X, Y, $Z \in \mathcal{X}(M)$ and f, $g \in C^\infty(M)$

$\nabla_X Y$ is called the covariant derivative of Y in the direction X.

How should a connection appear in a local frame? (See section 5.7 for a reminder of a local frame). If $\{E_i\}$ is a smooth local frame on an open subset $U \subset M$ then for every choice of indices i and j we can expand $\nabla_{E_i}E_j$ in the frame:

$$\nabla_{E_i}E_j = \Gamma^k_{ij}E_k$$

As i, j, and k range from 1 to $\dim M = n$, this relation determines n^3 smooth functions $\Gamma_{ij}^k : U \to M$ called the connection coefficients of ∇ with respect to the frame $\{E_i\}$. In fact, the connection is completely determined in U by its connection coefficients...

Proposition 5.1. *Let M be a smooth manifold and let ∇ be a connection on M. Suppose $\{E_i\}$ is a smooth local frame over an open subset $U \subset M$ defining connection coefficients $\{\Gamma_{ij}^k\}$ for ∇. For any two vector fields X, $Y \in \mathcal{X}(M)$ written in terms of this frame over U as $X = X^i E_i$, $Y = Y^i E_i$ on has*

$$\nabla_X Y = \left(X(Y^k) + X^i Y^j \Gamma_{ij}^k \right) E_k$$

Proof. Just compute using the defining properties of a connection:

$$\begin{aligned} \nabla_X Y = \nabla_X \left(Y^i E_i \right) &= X(Y^i) E_i + Y^i \nabla_X E_i \\ &= X(Y^i) E_i + Y^i \nabla_{X^j E_j} E_i \\ &= X(Y^j) E_j + Y^j X^i \nabla_{E_i} E_j \\ &= X(Y^j) E_j + Y^j X^i \Gamma_{ij}^k E_k \\ &= \left(X(Y^k) + X^i Y^j \Gamma_{ij}^k \right) E_k \end{aligned}$$

where we have relabelled the indices in the first term of the penultimate equality.
□ □

Proposition 10 shows slightly more

Proposition 5.2. *The value of $\nabla_X Y$ at $p \in M$ depends only the values of Y in an arbitrarily small neighbourhood $U \subset M$ of p and the value of X at p.*

Proof. Consider a chart around p where the local frame $\{E_i\}$ is given by the coordinate vector fields $\frac{\partial}{\partial x^i}$. From Proposition 10

$$\nabla_X Y|_p = \left(X(Y^k)|_p + X^i(p) Y^j(p) \Gamma_{ij}^k \right) E_k$$

which depends only on the value of X at p and, in view of the derivational terms $X(Y^k)$, the values of Y in a neighbourhood of p. □

Once the connection coefficients have been computed in a local frame, they can be determined in any other local frame on the same open set via the transformation law:

Proposition 5.3. *Let M be a smooth manifold with connection ∇ and suppose we are given two local frames $\{E_i\}$ and $\{\tilde{E}_i\}$ on an open subset $U \subset M$ related by $\tilde{E}_j = A_j^i E_i$ for some matrix of functions $\{A_j^i\}$. Let Γ_{ij}^k and $\tilde{\Gamma}_{ij}^k$ denote the connection coefficients of ∇ with respect to each of these frames. Then*

$$\tilde{\Gamma}_{ij}^k = (A^{-1})_p^k A_i^q A_j^r \Gamma_{qr}^p + (A^{-1})_p^k A_i^q E_q(A_j^p)$$

Proof. Write $\tilde{E}_i = A_i^q E_q$ and $\tilde{E}_j = A_j^r E_r$ and compute as before using the defining properties of the connection

$$\begin{aligned}\tilde{\Gamma}_{ij}^k \tilde{E}_k = \nabla_{\tilde{E}_i} \tilde{E}_j = \nabla_{\tilde{E}_i} A_j^r E_r &= \tilde{E}_i(A_j^r)E_r + A_j^r \nabla_{\tilde{E}_i} E_r \\ &= A_i^q E_q(A_j^r)E_r + A_j^r \nabla_{A_i^q E_q} E_r \\ &= A_i^q E_q(A_j^r)E_r + A_j^r A_i^q \nabla_{E_q} E_r \\ &= A_i^q E_q(A_j^r)E_r + A_j^r A_i^q \Gamma_{qr}^p E_p \\ &= \left(A_i^q E_q(A_j^p) + A_j^r A_i^q \Gamma_{qr}^p\right) E_p.\end{aligned}$$

Then writing $\tilde{\Gamma}_{ij}^k \tilde{E}_k = \tilde{\Gamma}_{ij}^k A_k^p E_p$, equating the coefficients and solving for $\tilde{\Gamma}_{ij}^k$ gives the result. □

So we've defined what a connection should be; we've looked at how a connection is determined by a local frame and the corresponding connection coefficients; we've seen how the connection transforms under a change of frame but we haven't produced a single example of a connection! Let's see some...

Given two vector fields X and $Y \in \mathcal{X}(\mathbb{R}^n)$ define

$$\nabla_X Y = X(Y^1)\frac{\partial}{\partial x^1} + \cdots + X(Y^n)\frac{\partial}{\partial x^n}$$

which is the standard directional derivative. It is straight forward to verify that this definition satisfies

a. $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$

b. $\nabla_X Y + Z = \nabla_X Y + \nabla_X Z$

c. $\nabla_X(fY) = f\nabla_X Y + X(f)\cdot Y$

and is therefore a connection for $\mathbb{R}^n$.

Let $M \subseteq \mathbb{R}^n$ be a submanifold. Define a connection $\nabla^\top$ on M, called the **tangential connection**, by setting

$$\nabla_X^\top Y = \pi^\top\left(\nabla_{\tilde{X}} \tilde{Y}|_M\right)$$

where $\pi^\top$ is the orthogonal projection onto TM, ∇ the Euclidean connection above, and $\tilde{X}$ and $\tilde{Y}$ are smooth extensions of X and Y to an open subset of $\mathbb{R}^n$. Let's check that this is well-defined and actually a connection...

The Euclidean connection was given by

$$\nabla_{\tilde{X}} \tilde{Y}|_p = \tilde{X}(\tilde{Y}^1)(p)\frac{\partial}{\partial x^1}|_p + \cdots + \tilde{X}(\tilde{Y}^n)(p)\frac{\partial}{\partial x^n}|_p.$$

$\nabla_{\tilde{X}}\tilde{Y}|_p$ is independent of the extension $\tilde{X}$ of X since it only depends on $\tilde{X}_p = X_p$. To calculate the directional derivatives $\tilde{X}(\tilde{Y}^i)$ we need only take a curve $\gamma(t) \in M$, $t \in (-\epsilon, \epsilon)$ with $\gamma'(0) = X_p$ so that $\tilde{X}(\tilde{Y}^i) = \frac{d}{dt}\left(\tilde{Y}^i \circ \gamma(t)\right) = \frac{d}{dt}\left(Y^i \circ \gamma(t)\right)$ depends only on Y. So $\nabla^\top$ is well-defined. Conditions 1. and 2. of a connection are clear so we just need to check 3.

Let $f \in C^\infty(M)$ and let $\tilde{f}$ be a smooth extension of f to a neighbourhood of M ins $\mathbb{R}^n$. Then $\tilde{f} \cdot \tilde{Y}$ is a smooth extension of $f \cdot Y$ to a neighbourhood of M; so

$$\begin{aligned}
\nabla_X^\top f \cdot Y &= \pi^\top\left(\nabla_{\tilde{X}}\tilde{f} \cdot \tilde{Y}|_M\right) \\
&= \pi^\top\left(\tilde{X}(\tilde{f})\tilde{Y}|_M\right) + \pi^\top\left(\tilde{f}\nabla_{\tilde{X}}\tilde{Y}|_M\right) \\
&= X(f)Y + f \cdot \nabla_X^\top \cdot Y
\end{aligned}$$

There are actually loads of connections on any given manifold M. Take a chart $(U_\alpha, \varphi_\alpha)$ for M with local coordinate frame $\{\frac{\partial}{\partial x^i}\}$. Define a connection ∇^α locally through the choice of n^3 smooth functions Γ_{ij}^k on U by

$$\nabla^\alpha_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}.$$

Doing this over each chart in an atlas for M and blending the local connections together smoothly via a partition of unity subordinate to the atlas gives a smooth connection on M:

$$\nabla = \sum_\alpha \psi_\alpha \nabla^\alpha$$

So there are many connections... Many.

Covariant Derivatives

Last time we saw how a connection can be used to take directional derivatives of vector fields on manifolds. One goal was to make sense of 'acceleration" of a curve in a manifold M. We can refine the idea of a connection slightly to take derivatives of vector fields along curves; in particular this will allow us to calculate the directional derivative of the velocity of a curve. Recall that a vector field V along a curve $\gamma : I \subseteq \mathbb{R} \to M$ is a smooth map that associates to each point on $\gamma(t)$ a tangent vector $V(t) \in T_{\gamma(t)}M$. We say V is induced by a vector field $X \in \mathcal{X}(M)$ if $V(t) = X(\gamma(t))$.

Theorem 5.4. *Let M be a smooth manifold and let ∇ be a connection on TM. For each smooth curve $\gamma : I \subseteq \mathbb{R} \to M$ the connection ∇ determines a unique operator $D_t : \mathcal{X}(\gamma) \to \mathcal{X}(\gamma)$ called the covariant derivative along γ, satisfying the following properties:*

a. *Linearity over $\mathbb{R}$*

$$D_t(aV + bW) = a \cdot D_tV + v \cdot D_tW$$

b. *Product Rule:*

$$D_t\,(f\cdot V) = f'\cdot V + fD_tV$$

c. *If V is induced by a vector field X then*

$$D_tV = \nabla_{\gamma'(t)}X$$

Proof. We'll start by assuming the existence of D_t and prove its uniqueness... Suppose D_t is an operator on vector fields along a curve $\gamma(t)$ that satisfies properties 1. - 3. given in the Theorem. Choose smooth coordinates (x^i) for M in a neighbourhood U of $\gamma(t_o)$ and consider the local coordinate vector frame $\{\frac{\partial}{\partial x^i}\} = \{E_i\}$. Write V in these coordinates and frame as

$$V(t) = v^i(t)E_i|_{\gamma(t)}$$

for t near t_o. Since the vector fields $E_i|_{\gamma(t)}$ along γ are induced by the local coordinate frame E_i, property 3 applies. In particular, using properties 2. and 3. we get

$$\begin{aligned} D_tV(t) = D_t\left(v^i(t)E_i|_{\gamma(t)}\right) &= \dot{v}^i(t)E_i|_{\gamma(t)} + v^i(t)D_tE|_{\gamma(t)} \\ &= \dot{v}^i(t)E_i|_{\gamma(t)} + v^i(t)\left(\nabla_{\dot{\gamma}(t)}E_i\right)|_{\gamma(t)} \\ &= \dot{v}^i(t)E_i|_{\gamma(t)} + v^i(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))E_k|_{\gamma(t)} \\ &= \left(\dot{v}^k(t) + v^i(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))\right)E_k|_{\gamma(t)} \end{aligned}$$

Since this expression depends only on the connection coefficients this shows that the operator is uniquely determined by ∇. To prove existence look at a coordinate chart around a point $\gamma(t_o)$ of γ and define D_tV by

$$D_tV(t) = \left(\dot{v}^k(t) + v^i(t)\dot{\gamma}^j(t)\Gamma^k_{ij}(\gamma(t))\right)E_k|_{\gamma(t)}.$$

This expression is valid in the case that V is induced by another vector field X and since differentiation is $\mathbb{R}$ linear and satisfies the product rule it's clear that this expression satisfies properties 1.-3. Uniqueness guarantees that the coordinate expressions in different charts agree on the overlaps. $\square$

Geodesics

This is great, the last section gave us a way of differentiating the velocity vector of a curve γ in the direction of γ; in particular,

Definition 5.2. *Given a smooth manifold M with connection ∇ on TM, we define the acceleration of a smooth curve $\gamma : I\subseteq\mathbb{R}\to M$ to be the vector field $D_t\gamma'$ along $\gamma(t)$.*

Definition 5.3. *Given a smooth manifold M with connection ∇ on TM, a smooth curve $\gamma : I\subseteq\mathbb{R}\to M$ is called a geodesic with respect to ∇ if it has zero acceleration: $D_t\gamma' = 0$.*

In coordinates (x^i) on $U \subset M$ if we write a curve γ as $\gamma(t) = (x^1(t), \ldots, x^n(t))$ then it follows from the general formula for D_t:

$$D_t V(t) = \left(\dot{v}^k(t) + v^i(t)\dot{\gamma}^j(t)\Gamma_{ij}^k(\gamma(t))\right) E_k|_{\gamma(t)}$$

that γ is a geodesic if and only if its component functions satisfy the geodesic equation

$$\ddot{x}^k + \dot{x}^i \dot{x}^j \Gamma_{ij}^k(x(t)) = 0.$$

A very natural question is do such curves exist?

The geodesic equation is a system of second order ODE's M which can be recast as a system of first order ODE's on TM by introducing $\dot{x}^k = v^k$:

$$\begin{aligned} \dot{x}^k &= v^k \\ \dot{v}^k &= -\dot{v}^i \dot{v}^j \Gamma_{ij}^k(x) \end{aligned}$$

where we are using natural coordinates $(x^1, \ldots, x^n, v^1, \ldots, v^n)$ in a chart $U \times \mathbb{R}^n$ for TM. In these coordinates the right hand side of the above system defines a vector field on TM written in the chart $U \times \mathbb{R}^n$ and given by

$$F(x, v) = v^k \frac{\partial}{\partial x^k}|_{(x,v)} - \dot{v}^i \dot{v}^j \Gamma_{ij}^k(x) \frac{\partial}{\partial v^k}|_{(x,v)}$$

Observe that this vector field

$$F(x, v) = v^k \frac{\partial}{\partial x^k}|_{(x,v)} - \dot{v}^i \dot{v}^j \Gamma_{ij}^k(x) \frac{\partial}{\partial v^k}|_{(x,v)}$$

is smooth! Why? The coordinate vector fields are smooth, the x^i's are smooth, the v^i's are smooth, the connection coefficients Γ_{ij}^k are smooth, and multiplication is smooth. By the fundamental theorem of existence and uniqueness of ODE's we conclude that for each $(p, w) \in U \times \mathbb{R}^n$ and $t_o \in \mathbb{R}$ there exists an open interval I around t_o and a unique smooth solution $\xi : I \to U \times \mathbb{R}^n$ to this system satisfying $\xi(t_o) = (p, w)$. Writing $\xi(t) = (x^1(t), \ldots, x^n(t), v^1(t), \ldots, v^n(t))$ in coordinates we see the existence of a curve $\gamma(t) = (x^1(t), \ldots, x^n(t))$ in M which solves the geodesic equation. To see uniqueness suppose that γ and $\tilde{\gamma}$ are both geodesics on M defined on some open interval I with $\gamma(t_o) = \tilde{\gamma}(t_o)$ and $\gamma'(t_o) = \tilde{\gamma}'(t_o)$:

- Looking at a coordinate chart U around $\gamma(t_o)$ we can define curves $\xi, \tilde{\xi} : (t_o - \epsilon, t_o + \epsilon) \to U \times \mathbb{R}^n$ which solve the first order system above.
- These two curves satisfy the same initial value problem and so by the uniqueness of solutions the curves ξ and $\tilde{\xi}$ agree on $(t_o - \epsilon, t_o + \epsilon)$.
- Since all components $x^i(t)$ and $\tilde{x}^i(t)$ agree on $(t_o - \epsilon, t_o + \epsilon)$ the curves $\gamma(t)$ and $\tilde{\gamma}(t)$ and their derivatives agree on $(t_o - \epsilon, t_o + \epsilon)$.
- Covering the curve with overlapping charts and repeating the above argument shows that γ and $\tilde{\gamma}$ agree on I; for example pick $b \in (t_o, t_o + \epsilon)$ and repeat the above for a chart around $\gamma(b)$.

We have proved the following Theorem

Theorem 5.5. *Let M be a smooth manifold and ∇ a connection on TM. For every $p \in M$ and every $v \in T_pM$, and $t_o \in \mathbb{R}$, there exists an open interval $I \subseteq \mathbb{R}$ around t_o and a geodesic $\gamma : I \to M$ satisfying $\gamma(t_o) = p$ and $\gamma'(t_o) = v$. Any two such geodesics agree on their common domain.*

Definition 5.4. *A geodesic $\gamma : I \to M$ is said to be maximal if it cannot be extended to a larger interval; that is, if there does not exist a geodesic $\tilde{\gamma} : \tilde{I} \to M$ defined on an interval $\tilde{I}$ and satisfying $\tilde{\gamma}_I = \gamma$.*

Theorem 5.6. *Let M be a smooth manifold and ∇ a connection on TM. For every $p \in M$ and every $v \in T_pM$ there is a unique maximal geodesic $\gamma : I \to M$ satisfying $\gamma(t_o) = p$ and $\gamma'(t_o) = v$ defined on some open interval containing 0.*

Proof. Let I be the union of all open intervals containing 0 on which there is a geodesic with the given initial conditions. The proof of uniqueness in the last theorem showed that all such geodesics agree when they overlap and so they define a geodesic $\gamma : I \to M$ which is the unique maximal geodesic with the given initial conditions. □

Parallel Transport

Let M be a smooth manifold without boundary and ∇ a connection on M.

Definition 5.5. *A smooth vector field V along a smooth curve γ in M is parallel along γ with respect to ∇ if $D_tV = 0$*

A geodesic can be a characterised a curve whose velocity vector is parallel. In coordinates the equation of parallelism takes the form

$$\dot{V}^k = -V^j \dot{\gamma}^i \Gamma^k_{ij}(\gamma)$$

This is a first order linear differential equation and so solutions exist, are unique, and are defined for as long as the curve γ is defined. In particular...

Theorem 5.7. *Suppose M is a smooth manifold and ∇ is a connection on TM. Given a smooth curve $\gamma : I \to M$ with $t_o \in I$ and a vector $T_{\gamma(t_o)}M$ there exists a unique parallel vector field V along γ such that $V(t_o) = v$.*

We'll prove the Theorem by first proving existence and uniqueness of first order linear differential equations on $\mathbb{R}^n$ and then on a manifold by writing the problem in a chart, applying the Existence and Uniqueness result for $\mathbb{R}^n$, and then smoothly welding together the solutions across chart overlaps. Here we go...

Lemma 5.8. *Let $I \subseteq \mathbb{R}$ be an open interval and for $1 \leq j, k \leq n$, let $A_j^k : I \to \mathbb{R}$ be smooth functions. For all $t_o \in I$ and every initial vector $(c_1, \ldots, c_n) \in \mathbb{R}^n$ the linear initial value problem*

$$\begin{aligned} \dot{V}^k(t) &= A_j^k(t) V^j(t) \\ V^k(t_o) &= c^k \end{aligned}$$

has a unique smooth solution on all of I and the solution depends smoothly on $(t, c) \in I \times \mathbb{R}^n$.

Proof. The fundamental theorem of existence and uniqueness of ODE's guarantees the existence of a unique solution on some interval $(t_o - \epsilon, t_o + \epsilon)$. We now show that this interval of existence can be extended to I. Suppose $A_j^k(t)$ are defined on an interval $(a, b) \subseteq \mathbb{R}$ containing t_o but $V(t)$ is only defined on (a_o, b_o) with $b_o < b$. On the compact interval $[0, b_o]$ each $A_j^k(t)$ is bounded from above so that the matrix $A(t)$ whose entries are given by $A_j^k(t)$ satisfies $|A(t)| \leq M$ (for the Frobenius norm). Using the differential equation we can compute in the Euclidean norm

$$\begin{aligned} \frac{d}{dt} \|V(t)\|^2 = 2\langle \dot{V}, V\rangle = 2\langle A(t)V(t), V(t)\rangle &= 2\,|A(t)|\,\|V(t)\|^2 \\ &= 2M\,\|V(t)\|^2 \end{aligned}$$

which implies an estimate of the form

$$\|V(t)\| \leq e^{Mt}\,\|V(0)\|$$

The above estimate implies the solution V lies in a compact set $B_R(0) \subseteq \mathbb{R}^n$, where $R = e^{Mb_o}\,\|V(0)\|$. If $\{t_i\}$ is any sequence of times approaching b_o from below then $\{V(t_i)\}$ is a sequence in K which has a convergent subsequence with limit $q \in K$. The domain of the $A_j^k(t)$ is open and so applying the fundamental theorem of existence and uniqueness to the ODE with initial condition $V(t_i)$ for i sufficiently large so that it's infinitesimally close to q, and noting that the two solutions agree by uniqueness we have extended the original solution past the time b_o. Repeating this we see that $b_o = b$ and the solution $V(t)$ is defined on all of I. $\square$

Now to prove Theorem 8 let $A_j^k(t) := \dot{\gamma}^i(t)\Gamma_{ij}^k(\gamma(t))$ and write the parallel transport equation as a matrix equation where $A_j^k(t)$ denotes the entries of a matrix $A(t)$ - these entries are defined for as long as γ is. If the image of γ lies in a single chart then apply the previous lemma and we're done. If not, then cover γ with overlapping charts and observe that the lemma can be used to continue the solution to the parallel transport equation uniquely across the overlaps.

The vector field $V(t)$ given in Theorem 8 is called the parallel transport of v along γ. For each t_o, $t_1 \in I$ define a map

$$P^{\gamma}_{t_o,t_1} : T_{\gamma(t_o)}M \to T_{\gamma(t_1)}M,$$

called the parallel transport map, by setting $P^{\gamma}_{t_o,t_1}(v) = V(t_1)$ for each $v \in T_{\gamma(t_o}M$. This map is:

a. linear because the equations of parallelism are linear; and

b. is an isomorphism because the uniqueness part of the above theorem implies that $P^{\gamma}_{t_1,t_0}$ is an inverse for $P^{\gamma}_{t_o,t_1}$.

Parallel transport is exceptionally useful: since parallel transport is an isomorphism take a basis $\{b_i\}$ of $T_{\gamma(t_o)}M$ and transport them along γ to obtain vector fields $E_i(t) = P^{\gamma}_{t_o,t}b_i$ which form a **parallel frame** along γ. Now every vector field $V(t)$ along γ can be expressed in terms of this frame: $V(t) = v^i(t)E_i(t)$. Since the E_i's are parallel we have

$$D_tV(t) = \dot{v}^i(t)E_i(t) + v^i(t)D_tE_i(t) = \dot{v}^i(t)E_i(t)$$

and we find that a vector field along γ is parallel if and only if its component functions are constant with respect to the parallel frame $\{E_i(t)\}$.

So far we've seen that a connection ∇ on TM determines covariant differentiation along a curve D_t which in turn determines parallel transport $P^{\gamma}_{t_o,t_1}$. This order of determination can be reversed; in particular, parallel transport determines covariant differentiation along a curve, which in turn determines the connection itself...

Theorem 5.9. *Let M be a smooth manifold and ∇ a connection on TM. Suppose $\gamma : I \to M$ is a smooth curve and V is a smooth vector field along γ. For each $t_o \in I$*

$$D_tV(t_o) = \lim_{t_1 \to t_o} \frac{P^{\gamma}_{t_1,t_o}V(t_1) - V(t_o)}{t_1 - t_o}$$

Proof. Let $\{E_i\}$ be a parallel frame along γ and write $V(t) = v^i(t)E_i(t)$. On the one hand, the formula above says $D_tV(t) = \dot{v}^i(t)E_i(t)$. On the other hand, for each $t_1 \in I$ the parallel transport of the vector $V(t_1)$ along γ is the constant coefficient vector field $W(t) = v^i(t_1)E_i(t_o)$ so that $P^{\gamma}_{t_1,t_o}V(t_1) = v^i(t_1)E_i(t_o)$ and

$$\begin{aligned}
\lim_{t_1 \to t_o} \frac{P^{\gamma}_{t_1,t_o}V(t_1) - V(t_o)}{t_1 - t_o} &= \lim_{t_1 \to t_o} \frac{v^i(t_1)E_i(t_o) - v^i(t_o)E_i(t_o)}{t_1 - t_o} \\
&= \dot{v}^i(t_o)E_i(t_o) \\
&= D_tV(t_o)
\end{aligned}$$

□

Theorem 5.10. *Let M be a smooth manifold and ∇ a connection on TM. Suppose X and Y are smooth vector fields on M. For every p in M*

$$\nabla_X Y|_p = \lim_{h \to 0} \frac{P^{\gamma}_{h,0}Y_{\gamma(h)} - Y_p}{h}$$

where γ is a smooth curve in M with $\gamma(0) = p$ and $\gamma'(0) = X_p$

Proof. Let γ be a curve in M with $\gamma(0) = p$ and $\gamma'(0) = X_p$ and let $V(t)$ be the vector field along γ induced by Y; in this case $D_t V(0) = \nabla_X Y|_p$ and so the result follows from the previous theorem. $\square$

Pull-back Connections

If ∇ is a connection on N then there is a connection $F^*\nabla$ on M obtained by pulling back ∇ to M via F. The connection $F^*\nabla$ is defined by the relation

$$(F^*\nabla)_X Y = dF^{-1} \cdot (\nabla_{dF \cdot X} dF \cdot Y)$$

In words this says: first push forward X and Y to N with F and compute the covariant derivative using ∇, then push the result back to M via F^{-1}. That this expression is indeed a connection (i.e. is $\mathbb{R}$-linear in Y, $C^\infty(M)$ linear in X, and satisfies the product rule) essentially follows from the linearity of the pushforward and the fact that ∇ is a connection. But check it as an exercise.

Proposition 5.11. *Suppose M and N are smooth manifolds and $F : M \to N$ is a diffeomorphism. Let ∇^N be a connection on TN and $\nabla^M = F^*\nabla^N$ the pullback connection on TM. Suppose $\gamma : I \subseteq \mathbb{R} \to M$ is a smooth curve.*

a. *F takes covariant derivatives along curves to covariant derivatives along curves: if V is a smooth vector field along γ then*

$$dF \cdot D_t^M V = D_t^N (dF \cdot V)$$

b. *F takes geodesics to geodesics: if γ is a ∇^M geodesic in M then $F \circ \gamma$ is a ∇^N geodesic in N*

c. *F takes parallel transport to parallel transport: for every t_o and $t_1 \in I$*

$$dF_{\gamma(t_1)} \circ P^{\gamma}_{t_o,t_1} = P^{F\circ\gamma}_{t_o,t_1} \circ dF_{\gamma(t_1)}$$

Proof. *a.* Compute in a chart just as we did when deriving a coordinate expression for D_t. Express V on M using a coordinate vector frame: $V(t) = v^i(t) E_i|_{\gamma(t)}$; then

$$\begin{aligned}
dF \cdot D_t^M V &= dF \cdot D_t^M \left(v^i(t) E_i|_{\gamma(t)}\right) \\
&= dF \left(\dot{v}^i E_i|_\gamma + v^i D_t^M E_i|_\gamma\right) \\
&= dF \left(\dot{v}^i E_i|_\gamma + v^i \left(\nabla^M_{\dot\gamma} E_i\right)|_\gamma\right) \\
&= dF \left(\dot{v}^i E_i|_\gamma + v^i dF^{-1} \left(\nabla^N_{dF\cdot\dot\gamma}(dF \cdot E_i)\right)|_{F\circ\gamma}\right) \\
&= \dot{v}^i (dF \cdot E_i)|_{F\circ\gamma} + v^i \left(\nabla^N_{dF\cdot\dot\gamma}(dF \cdot E_i)\right)|_{F\circ\gamma} \\
&= D_t^N (dF \cdot V)
\end{aligned}$$

Parts 2. and 3. follow directly from 1. and properties of differentials:

b. Follows directly from 1. and properties of differentials: If γ is a geodesic in M with respect to ∇^M then $D_t^M \gamma' = 0$ and 1. implies that $0 = D_t^N (dF \cdot \gamma') = D_t^N (F \circ \gamma)'$ so that $F \circ \gamma$ is a geodesic with respect to ∇^N.

c. Follows directly from 1. and properties of differentials: similarly, $dF \cdot V$ is parallel with respect to ∇^N if and only if V is parallel with respect to ∇^M which implies the relation stated in the proposition.

$\square$

Chapter 6

The Levi-Civita Connection

Using the example of the tangential connection on a submanifold of $\mathbb{R}^n$ given above we'll see two key properties that are sufficient to determine a unique connection on any given Riemannian manifold. These are...

a. Compatibility with the metric; and

b. Symmetry

Compatibility

Recall that the Euclidean connection on $\mathbb{R}^n$ was given by the ordinary directional derivative:

$$\nabla_X Y = X(Y^1)\frac{\partial}{\partial x^1} + \cdots + X(Y^n)\frac{\partial}{\partial x^n}$$

for any two vector fields X and Y, and this induces covariant differentiation along a curve γ given by D_t. Recall that the general formula for a connection on a manifold M given in a chart U was $\nabla_X Y = \left(X(Y^k) + X^i Y^j \Gamma_{ij}^k\right) E_k$ and so the connection coefficients Γ_{ij}^k of the Euclidean connection are all zero. The Euclidean connection ∇ and D_t satisfy one very nice property with respect to the Euclidean metric $\langle \cdot, \cdot \rangle$... Let $V(t)$ and $W(t)$ be any two vector fields along a smooth curve $\gamma : I \subseteq \mathbb{R} \to \mathbb{R}^n$; then

$$\frac{d}{dt}\langle V(t), W(t)\rangle = \langle D_t V, W\rangle + \langle V, D_t W\rangle$$

Starting with the right hand side of the above equation and noting that in the Euclidean case the connection coefficients Γ_{ij}^k are all zero we have

$$D_t V = \dot{v}^k(t) E_k|_{\gamma(t)}$$
$$D_t W = \dot{w}^k(t) E_k|_{\gamma(t)}$$

so that

$$\langle D_t V, W\rangle + \langle V, D_t W\rangle = \dot{v}^k(t) w^k(t) + v^k(t)\dot{w}^k = \frac{d}{dt}\langle V, W\rangle.$$

As a consequence, the connection satisfies

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$$

for any vector fields X, Y, and $Z \in \mathcal{X}(\mathbb{R}^n)$. Indeed for a point $p \in M$ let $\gamma(t)$ be a curve through p at t_o with $\gamma'(t_o) = X_p$. Then

$$\begin{aligned} X_p\langle Y, Z\rangle &= \frac{d}{dt}\left(\langle Y, Z\rangle \circ \gamma(t)\right)|_{t=t_o} \\ &= \langle D_t\left(Y \circ \gamma(t)\right)|_{t=t_o}, Z\rangle + \langle Y, D_t\left(Z \circ \gamma(t)\right)|_{t=t_o}\rangle \\ &= \langle \nabla_{X_p} Y, Z\rangle + \langle Y, \nabla_{X_p} Z\rangle \end{aligned}$$

where the last line follows from the definition of a directional derivative, or more generally from property 3 of the operator D_t.

Recall that the Tangential connection on a smooth submanifold $M \subseteq \mathbb{R}^n$ was given by orthogonal projection of the Euclidean connection onto TM

$$\nabla^{\top}_X Y = \pi^{\top}\left(\nabla_{\tilde{X}}\tilde{Y}\right)$$

for any two vector fields X and Y and smooth extensions $\tilde{X}$ and $\tilde{Y}$ to a neighbourhood of M in $\mathbb{R}^n$. This induces covariant differentiation along a curve γ in M given by $D^{\top}_t$. Giving M the metric g induced by the Euclidean metric it is immediate that

$$Xg(Y, Z) = g(\nabla^{\top}_X Y, Z) + g(Y, \nabla^{\top}_X Z).$$

...Why?... Let $\tilde{X}$, $\tilde{Y}$ and $\tilde{Z}$ be smooth extensions of vector fields X, Y, and Z on M to a neighbourhood of M in $\mathbb{R}^n$. Starting with the right-hand side of the equation again and recalling that the induced metric is the restriction of the Euclidean metric to vectors tangent to M we have at points p of M

$$\begin{aligned} g(\nabla^{\top}_X Y, Z) + g(Y, \nabla^{\top}_X Z) &= g(\pi^{\top}\left(\nabla_{\tilde{X}}\tilde{Y}\right), \tilde{Z}) + g(\tilde{Y}, \pi^{\top}\left(\nabla_{\tilde{X}}\tilde{Z}\right)) \\ &= \langle \pi^{\top}\left(\nabla_{\tilde{X}}\tilde{Y}\right), \tilde{Z}\rangle + \langle \tilde{Y}, \pi^{\top}\left(\nabla_{\tilde{X}}\tilde{Z}\right)\rangle \\ &= \langle\left(\nabla_{\tilde{X}}\tilde{Y}\right), \tilde{Z}\rangle + \langle \tilde{Y}, \left(\nabla_{\tilde{X}}\tilde{Z}\right)\rangle \\ &= \tilde{X}\langle \tilde{Y}, \tilde{Z}\rangle \\ &= X\langle Y, Z\rangle \\ &= X \cdot g(Y, Z) \end{aligned}$$

The above property of both the Euclidean and tangential connections makes sense on a general Riemannian manifold (M, g):

Definition 6.1. *Given a smooth Riemannian manifold (M, g), a connection ∇ on TM is said to be compatible with g and called a metric connection if for all X, Y, and $Z \in \mathcal{X}(M)$:*

$$\nabla_X g(Y, Z) = X \cdot g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

It turns out that compatibility is not enough to uniquely pin down a natural connection for a given metric. We need to turn to another key property of both the Euclidean and tangential connections...

Symmetry

The Euclidean connection satisfies

$$[X, Y] = \nabla_X Y - \nabla_Y X$$

for any smooth vector fields X and Y on $\mathbb{R}^n$. This follows directly from the definition of the Euclidean connection as the directional derivative and the definition of the Lie bracket (section 5.8).

Given a smooth submanifold M of $\mathbb{R}^n$ the tangential connection $\nabla^\top$ also satifies

$$[X, Y] = \nabla^\top_X Y - \nabla^\top_Y X$$

for any smooth vector fields X and Y in $\mathcal{X}(M)$. Indeed, if $\tilde{X}$ and $\tilde{Y}$ are smooth extensions of X and Y to a neighbourhood of M and $\iota : M \hookrightarrow \mathbb{R}^n$ is the natural inclusion map then

$$\begin{aligned} d\iota_p X &= \tilde{X}(\iota(p)) \\ d\iota_p Y &= \tilde{Y}(\iota(p)) \end{aligned}$$

so that $d\iota_p[X, Y] = [\tilde{X}, \tilde{Y}](\iota(p))$. In particular $[\tilde{X}, \tilde{Y}]$ is tangent to M and coincides with $[X, Y]$ at points of M so that

$$\begin{aligned} \nabla^\top_X Y - \nabla^\top_Y X &= \pi^\top \left(\nabla_{\tilde{X}} \tilde{Y}|_M - \nabla_{\tilde{Y}} \tilde{X}|_M\right) \\ &= \pi^\top \left([\tilde{X}, \tilde{Y}]|_M\right) \\ &= [\tilde{X}, \tilde{Y}]|_M \\ &= [X, Y] \end{aligned}$$

What the above shows is that if we wish to single out a connection on each Riemannian manifold in such a way that it matches the tangential connection when the manifold is presented as an embedded submanifold of $\mathbb{R}^n$ with the induced metric then we must require that the connection be at least compatible with the metric and symmetric. It is a very convenient fact that compatibility and symmetry are enough to determine a unique connection.

The Fundamental Theorem of Riemannian Geometry

Theorem 6.1. *Let (M, g) be a Riemannian manifold. There exists a unique connection ∇ on TM that is compatible with the metric g and symmetric.*

Proof. As usual we will assume existence and prove uniqueness. Uniqueness will then dictate how to define it and hence prove existence. So suppose ∇ is a symmetric connection compatible with g and let X, Y, $Z \in \mathcal{X}(M)$. Writing the compatibility condition three times with X, Y, and Z cyclically permuted we get:

$$\begin{aligned} X \cdot g(Y,Z) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ Y \cdot g(Z,X) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \\ Z \cdot g(X,Y) &= g(\nabla_Z X, Y) + g(X, \nabla_Z Y) \end{aligned}$$

Next we incorporate compatibility:

$$\begin{aligned} X \cdot g(Y,Z) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z - \nabla_Z X + -\nabla_Z X) \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_Z X) + g(Y, [X,Z]) \end{aligned}$$

Doing this for the other two equations above gives

$$\begin{aligned} Y \cdot g(Z,X) &= g(\nabla_Y Z, X) + g(Z, \nabla_X Y) + g(Z, [Y,X]) \\ Z \cdot g(X,Y) &= g(\nabla_Z X, Y) + g(X, \nabla_Y Z) + g(X, [Z,Y]) \end{aligned}$$

Adding the first two of the above equations and subtracting the third results in

$$\begin{aligned} X \cdot g(Y,Z) + &Y \cdot g(Z,X) - Z \cdot g(X,Y) = \\ &2g(\nabla_X Y, Z) + g(Y,[X,Z]) + g(Z,[Y,X]) - g(X,[Z,Y]). \end{aligned}$$

Solving for $2g(\nabla_X Y, Z)$ gives

$$\begin{aligned} 2g(\nabla_X Y, Z) = X \cdot g(Y,Z) + &Y \cdot g(Z,X) - Z \cdot g(X,Y) \\ &- g(Y,[X,Z]) - g(Z,[Y,X]) + g(X,[Z,Y]). \end{aligned}$$

Now let ∇^1 and ∇^2 be any two compatible and symmetric connections. Since the right hand side of the formula derived above does not depend on any connection we must have

$$g(\nabla^1_X Y - \nabla^2_X Y, Z) = g(\nabla^1_X Y, Z) - g(\nabla^2_X Y, Z) = 0$$

Since this equation holds for any vector field Z, including $Z = \nabla^1_X Y - \nabla^2_X Y$, we must have $\nabla^1_X Y - \nabla^2_X Y = 0$. And since this holds for all vector fields X and Y we must have $\nabla^1 = \nabla^2$; this proves uniqueness.

To prove existence we define the connection ∇ through the formula we derived above. I'll do this more explicitly in a moment by giving the exact expressions for the connection coefficients but before doing this observe that any connection defined through the above formula is necessarily compatible and symmetric; Indeed, using the formula with Y and Z interchanged we have

$$\begin{aligned} 2g(\nabla_X Y, Z) + 2g(Y, \nabla_X Z) &= X \cdot g(Y,Z) + Y \cdot g(Z,X) - Z \cdot g(X,Y) \\ &- g(Y,[X,Z]) - g(Z,[Y,X]) + g(X,[Z,Y]) \\ &+ X \cdot g(Y,Z) + Z \cdot g(Y,X) - Y \cdot g(X,Z) \\ &- g(Z,[X,Y]) - g(Y,[Z,X]) + g(X,[Y,Z]) \\ &= 2X \cdot g(Y,Z) \end{aligned}$$

Symmetry follows in an analogous way by subtracting $2g(\nabla_Y X, Z)$ from $2g(\nabla_X Y, Z)$ and applying the above formula to each of these terms:

$$2g(\nabla_X Y - \nabla_Y X, Z) = 2g([X, Y], Z).$$

Since this expression holds for any vector field Z we have

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

Whatever form this connection takes we now know it is unique by the first part of the proof so if we compute in a chart the expression will automatically agree on overlaps.

Take a chart U with coordinates (x^i) and apply the above formula to the coordinate vector fields $\frac{\partial}{\partial x^i} = E_i$ while remembering that $[E_i, E_j] = 0$ for all i and j:

$$2g(\nabla_{E_i} E_j, E_l) = E_i \cdot g(E_j, E_l) + E_j \cdot g(E_k, E_i) - E_l \cdot g(E_i, E_j).$$

Recall the definitions of the metric and connections coefficients and the action of the coordinate vector fields on functions:

$$g_{ij} = g(E_i, E_j) \quad \nabla_{E_i} E_j = \Gamma^m_{ij} E_m \quad E_i f = \frac{\partial f}{\partial x^i}.$$

With these definitions

$$2g(\nabla_{E_i} E_j, E_l) = 2g(\Gamma^m_{ij} E_m, E_l) = 2\Gamma^m_{ij} g_{ml}$$

and

$$2\Gamma^m_{ij} g_{ml} = \frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l}$$

Denoting the components of the inverse metric by g^{kl} and noting that $g_{ml} g^{kl} = \delta^k_m$ we get

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right)$$

which defines the connection coefficients and hence the connection in each chart. □ □

In the proof of the Fundamental Theorem we picked up two formulas for the Levi-Civita connection. We'll record them independently:

a. (Koszul's Formula): If X, Y, and Z are smooth vector fields on M then

$$\begin{aligned} 2g(\nabla_X Y, Z) = X \cdot g(Y, Z) &+ Y \cdot g(Z, X) - Z \cdot g(X, Y) \\ &- g(Y, [X, Z]) - g(Z, [Y, X]) + g(X, [Z, Y]). \end{aligned}$$

b. In any smooth coordinate chart on M the Levi-Civita connection coefficients are given by

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right)$$

We saw that the Euclidean and Tangential connections are compatible and symmetric but the fundamental theorem says there's only one connection with these properties and that's the Levi-Civita connection.

Proposition 6.2. *Suppose (M, g) and (N, h) are Riemannian manifolds and let ∇^g denote the Levi-Civita connection of g and ∇^h that of h. If $\phi : M \to N$ is an isometry then $\nabla^g = \phi^*\nabla^h$.*

Proof. Since the Levi-Civita connection is unique it's enough to show that the pull-back connection $\phi^*\nabla^h$ is symmetric and compatible with g.Since ϕ is an isometry we have

$$g(X, Y) = h(d\phi \cdot X, d\phi \cdot Y) \circ \phi$$

for all vector fields X and Y on M. Now using the definition of the differential and the pull-back connection we compute...

$$\begin{aligned}
&X \cdot g(Y, Z) \\
&= X\left(h(d\phi \cdot Y, d\phi \cdot Z) \circ \phi\right) \\
&= ((d\phi \cdot X) \cdot h(d\phi \cdot Y, d\phi \cdot Z)) \circ \phi \\
&= \big(h(\nabla^h_{d\phi\cdot X} d\phi \cdot Y, d\phi \cdot Z) + h(d\phi \cdot Y, \nabla^h_{d\phi\cdot X} d\phi \cdot Z)\big) \circ \phi \\
&= \big(h(d\phi \circ d\phi^{-1}\nabla^h_{d\phi\cdot X} d\phi \cdot Y, d\phi \cdot Z) + h(d\phi \cdot Y, d\phi \circ d\phi^{-1}\nabla^h_{d\phi\cdot X} d\phi \cdot Z)\big) \circ \phi \\
&= \big(h(d\phi \cdot (\phi^*\nabla^h)_X Y, d\phi \cdot Z) + h(d\phi \cdot Y, d\phi \cdot (\phi^*\nabla^h)_X Z)\big) \circ \phi \\
&= g((\phi^*\nabla^h)_X Y, Z) + g(Y, (\phi^*\nabla^h)_X Z)
\end{aligned}$$

which proves compatibility. To prove symmetry use the definition of the pull-back connection properties of the Lie bracket under differentials

$$\begin{aligned}
(\phi^*\nabla^h)_X Y - (\phi^*\nabla^h)_Y X &= d\phi^{-1}\nabla^h_{d\phi\cdot X} d\phi \cdot Y - d\phi^{-1}\nabla^h_{d\phi\cdot Y} d\phi \cdot X \\
&= d\phi^{-1}[d\phi \cdot X, d\phi \cdot Y]^N \\
&= d\phi^{-1} d\phi [X, Y]^M \\
&= [X, Y]^M
\end{aligned}$$

□ □

As a corollary we get

Corollary 6.3. *Suppose (M, g) and (N, h) are Riemannian manifolds and $\phi : M \to N$ is a local isometry. If γ is a geodesic of the Levi-Civita connection in M then $\phi \circ \gamma$ is a geodesic of the Levi-Civita connection in N.*

Proof. This follows from the general properties of pull-back connections (see the previous section), the previous theorem, and the fact that being a geodesic is a local property. □

The Exponential Map

From now on we let (M,g) denote a smooth Riemannian manifold without boundary endowed with its Levi-Civita connection ∇. We've seen that each point $(p,v) \in TM$ determines a a unique maximal geodesic γ_{v_p}. To deepen our understanding of geodesics we need to study their *collective* behaviour and , in particular, address the following question: "How do geodesics change if we vary the initial position p or the initial velocity v?". The dependence of geodesics on the initial data is encoded in a map from tangent bundle to the manifold, called the *exponential* and *map*.

Geodesics with proportional velocities are equal with a proportional scaling of time.

Lemma 6.4. *For every $(p,v) \in TM$ and every $c,\, t \in \mathbb{R}$ we have*

$$\gamma_{cv}(t) = \gamma_v(ct)$$

Proof. At $t = 0$ both sides of the equation equal p. The initial velocity of γ_{cv} is cv while the initial velocity of γ_v is v so that

$$\frac{d}{dt}\gamma_v(ct)|_{t=0} = c \cdot \gamma_v'(0) = cv.$$

Since the two curves $\gamma_{cv}(t)$ and $\gamma_v(ct)$ have the same initial conditions the fundamental theorem on existence and uniqueness of ODE's implies the two curves are the same. □ □

If you're not convinced that $\gamma_v(ct)$ is a geodesic for $c \neq 1$ then take its first and second derivatives and show that they satisfy the geodesic equation - it's good practice anyway.

The assignment

$$(p,v) \mapsto \gamma_{v_p}$$

in conjunction with the above Lemma allows us to define a map from TM to M which sends each line in T_pM through the origin to a geodesic. Define a subset $\mathcal{E} \subseteq TM$ - the domain of the exponential map - by

$$\mathcal{E} = \{(p,v) \in TM :\ \gamma_{v_p} \text{is defined on an interval containing } [0,1]\}$$

and then define the exponential map

$$\begin{aligned} \exp : \mathcal{E} &\to M \\ \exp(v_p) &= \gamma_{v_p}(1) \end{aligned}$$

For each $p \in M$ the restricted exponential map at p, denoted by $\exp_p$, is the restriction of exp to the set $\mathcal{E}_p = \mathcal{E} \cap T_pM$.

Remark 6.1. *The exponential map of a Riemannian manifold should NOT be confused with the exponential map of a Lie group. When a Lie group has a bi-invariant metric the two are related but in general they are not.*

A subset S of a vector space V is said to be star-shaped with respect to a point $x \in S$ if for every $y \in S$ the line segment from x to y is contained in S.

Proposition 6.5. *Let* $\exp : \mathcal{E} \to M$ *be the exponential map of a Riemannian manifold* (M, g).

a. *For each* $(p, v) \in TM$ *the geodesic* γ_{v_p} *is given by*

$$\gamma_{v_p}(t) = \exp(tv_p)$$

b. $\mathcal{E}$ *is an open subset of* TM *containing the image of the zero section and each set* $\mathcal{E}_p$ *is star shaped with respect to* 0.

c. *The exponential map is smooth*

d. *At each point* $p \in M$ *the differential of* $\exp_p$ *at* 0 , $d\exp_p(0) : T_0(T_pM) \cong T_pM \to T_pM$, *is the identity map of* T_pM.

Proof. Part 1. follows from the definition of the exponential map and rescaling Lemma above with $t = 1$: $\exp(cv) = \gamma_{cv}(1) = \gamma(c)$. If $v \in \mathcal{E}$ then γ_v is defined on at least $[0, 1]$. The rescaling lemma above gives

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

and since the right hand side is defined for at least $0 \le t \le 1$, the line segment tv lies in $\mathcal{E}_p$ which shows that $\mathcal{E}_p$ is star-shaped with respect to 0. To show that $\mathcal{E}$ is open we return to the fundamental theorem on existence and uniqueness of solutions to ODE's...

Recall the we recast the second order geodesic equation on M as a first order equation on TM whose solutions were given by integral curves of the smooth vector field

$$F(x, v) = v^k \frac{\partial}{\partial x^k}|_{(x,v)} - \dot{v}^i \dot{v}^j \Gamma^k_{ij}(x) \frac{\partial}{\partial v^k}|_{(x,v)}$$

written in a chart using natural coordinates $(x^1, \ldots, x^n, v^1, \ldots, v^n)$ on TM. The fundamental theorem says given a smooth vector field on TM there exists an open set $\mathcal{D} \subset \mathbb{R} \times TM$ containing $\{0\} \times TM$ and a smooth map $\theta : \mathcal{D} \to TM$ such that each curve $\theta^{(p,v)}(t) = \theta(t, p, v)$ is the unique maximal integral curve of F starting at (p, v), defined on an open interval around 0. Now suppose $(p, v) \in \mathcal{E}$. This means the geodesic γ_{v_p} is defined on at least $[0, 1]$ and therefore so is the integral curve of F starting at $(p, v) \in TM$. Since $(1, (p, v)) \in \mathcal{D}$ there is a neighbourhood of $(1, (p, v))$ in $\mathbb{R} \times TM$ on which the flow of F is defined. In particular, this means there is a neighbourhood of (p, v) on which the flow exists for $t \in [0, 1]$ and therefore on which the exponential map is defined. That is, $\mathcal{E}$ is open.

To show that $\exp_p$ is smooth we can use the above to write the exponential map as

$$\exp_p(v) = \gamma_{v_p}(1) = \pi \circ \theta(1, (p, v))$$

which is a composition of smooth maps and is therefore smooth. To show that $d\exp_p(0)$ is the identity map let $c(t) = tw \in T_pM$ so that $c(0) = 0$ and $c'(0) = w$ and compute

$$d\exp_p(0) \cdot w = \frac{d}{dt}|_{t=0}(\exp_p \circ c)(t) = \frac{d}{dt}|_{t=0} \exp_p(tv) = \frac{d}{dt}|_{t=0}\gamma_{v_p}(t)$$
$$= v$$

since this holds for all $w \in T_pM$ $d\exp_p(0) = I$. □ □

Our previous results on isometries between Riemannian manifolds, their Levi-Civita connections, and geodesics translates into the following Theorem for the exponential map

Proposition 6.6. *Suppose (M, g) and (N, h) are Riemannian manifolds and $\phi : M \to N$ is a local isometry. Then for every $p \in M$*

$$\phi \circ \exp_p^g = \exp_{\phi(p)}^h \circ d\phi_p$$

A consequence of the last Proposition is that local isometries on connected manifolds are completely determined by their values and differentials at a single point...

Proposition 6.7. *Suppose (M, g) and (N, h) are Riemannian manifolds with M connected and ϕ, $\psi : M \to N$ are local isometries such that for some point $p \in M$ we have $\phi(p) = \psi(p)$ and $d\phi_p = d\psi_p$. Then $\phi = \psi$.*

Proof. From the previous Proposition we have

$$\phi \circ \exp_p^g = \exp_{\phi(p)}^h \circ d\phi_p = \exp_{\psi(p)}^h \circ d\psi_p = \psi \circ \exp_p^g$$

In particular, ϕ and ψ must take the same values on their common domains of definition in M. □ □

The restricted exponential map $\exp_p$ maps $\mathcal{E}_p \subset T_pM$ smoothly into M. Since $d\exp_p(0)$ is invertible the inverse function Theorem guarantees that there exists a neighbourhood V of 0 in T_pM and a neighbourhood U of p in M such that

$$\exp_p : V \to U$$

is a diffeomorphism.

Definition 6.2. *A neighbourhood U of p in M that is the diffeomorphic image of a star-shaped neighbourhood of 0 in T_pM is called a normal neighbourhood of p.*

Every orthonormal basis (b_i) for T_pM determines a basis isomorphism $B : \mathbb{R}^n \to T_pM$ by

$$B(x^1, \dots, x^n) = x^i b_i$$

If U is a normal neighbourhood of p we can combine this isomorphism with the exponential map.

Definition 6.3. *The coordinates defined by the coordinate chart map*

$$\phi = B^{-1} \circ \left(\exp_p|_V\right)^{-1} : U \to \mathbb{R}^n$$

are called normal coordinates centered at p.

Proposition 6.8. *Let (M, g) be a Riemannian manifold, p a point of M, and U a normal neighbourhood of p. For every normal coordinate chart on U, centered at p, the coordinate basis is orthonormal at p; and for every orthonormal basis (b_i) of T_pM there is a unique normal coordinate chart (x^i) on U such that $\partial_i|_p = b_i$. Any two normal coordinate charts (x^i) adn $(\tilde{x}^i)$ are related by*

$$\tilde{x}^j = A_i^j x^i$$

for some constant matrix $A_i^j \in O(n)$.

Proof. Let ϕ be a normal coordinate chart on U centered at p with coordinate functions (x^i). This means that $\phi = B^{-1} \circ \left(\exp_p|_V\right)^{-1}$ where $B : \mathbb{R}^n \to T_pM$ is the basis isomorphism determined by some orthonormal basis (b_i) for T_pM. Note that $d\phi_p^{-1} = d\exp_p(0) \circ dB(0) = B$ because $d\exp_p(0) = I$ and B is linear. Therefore $\partial_i|_p = d\phi^{-1}(\partial_i|_0) = B(\partial_i|_0) = b_i$, which shows that the coordinate basis at p is orthonormal. Conversely, every orthonormal basis (b_i) determines a basis isomorphism B and thus a normal coordinate chart $\phi = B^{-1} \circ \left(\exp_p|_V\right)^{-1}$ which satisfies $\partial_i|_p = b_i$ by the above computation. If $\tilde{\phi}^{-1} = \tilde{B}^{-1} \circ \left(\exp_p|_V\right)^{-1}$ is another normal coordinate chart then the change of coordinate map is

$$\tilde{\phi}^{-1} \circ \phi^{-1} = \tilde{B}^{-1} \circ \left(\exp_p|_V\right)^{-1} \circ \left(\exp_p|_V\right) \circ B = \tilde{B}^{-1} \circ B$$

which is a linear isometry of $\mathbb{R}^n$ and therefore has the form given in the statement of the Theorem with respect to standard coordinates. Normal coordinates $(\tilde{x}^j)$ and (x^i) are equal if and only if $(A_j^i) = A = \tilde{B}^{-1} \circ B$ is the identity matrix which shows that normal coordinates associated with a given orthonormal basis are unique. □ □

Proposition 6.9. *Let (M, g) be a Riemannian manifold and let $(U, (x^i))$ be any normal coordinate chart centered at $p \in M$. Then*

a. *The coordinates of p are $(0, \ldots, 0)$.*

b. *The components of the metric at p are $g_{ij} = \delta_{ij}$.*

c. *For every $v = v^i \partial_i|_p \in T_pM$, the geodesic γ_v starting at p with initial velocity v is represented in normal coordinates by the line*

$$\gamma_v(t) = (tv^1, \ldots, tv^n)$$

as long as t is in some interval I containing 0 such that $\gamma_v(I) \subset U$.

d. *The Levi-Civita connection coefficients (Christoffel symbols) vanish at p*

Proof. Part 1. follows from the definition of Normal coordinates. The normal coordinate vectors are orthonormal so

$$g_{ij}(p) = g(0)(\partial_i, \partial_j) = \delta_{ij}$$

This is part 2. From the properties of the exponential map a geodesic γ_v is given by $\gamma_v(t) = \exp_p(tv)$. Writing this in a normal coordinate chart:

$$\phi(\gamma_v(t)) = B^{-1} \circ \left(\exp_p\right)^{-1}\right) \circ \exp_p(tv) = B^{-1}(tv) = (tv^1, \ldots, tv^n)$$

This is part 3. To prove part 4. let $v = v^i \partial_i|_p \in T_pM$ be arbitrary. Using part 3. the geodesic equation for $\gamma_v(t) = (tv^1, \ldots, tv^n)$ simplifies to

$$\Gamma^k_{ij}(tv)v^i v^j = 0$$

Evaluating at $t = 0$ this shows that $\Gamma^k_{ij}(0)v^i v^j = 0$ for every index k and every index k and every vector v. In particular, set $v = \partial_a$ for some fixed a - this shows that $\Gamma^k_{aa} = 0$ (no summation) for each a and k. Now set $v = \partial_a + \partial_b$ so that

$$0 = \Gamma^k_{ij}(0)v^i v^j = 2\Gamma^k_{ab}$$

This is part 4. □

The Last thing we need for later chapters is the idea of a uniformly normal neighbourhood.

Definition 6.4. *A subset $W \subseteq M$ is called uniformly normal if there exists a $\delta > 0$ such that W is contained in a geodesic ball of radius δ around each of its points.*

Proposition 6.10. *Given any $p \in M$ and any neighbourhood $U \subseteq M$ of p there exists a uniformly normal neighbourhood of p contained in U.*

Proof. Choose a normal coordinate chart $(U_o, (x^i))$ centered at p and contained in U and identify U_o with an open subset of $\mathbb{R}^n$. Let (x^i, v^i) be the natural coordinates on $\pi^{-1}(U_o) \subseteq TM$ so that $\pi^{-1}(U_o) \cong U_o \times \mathbb{R}$. Consider the map $E : U_o \times \mathbb{R} \to U_o \times U_o$ defined by

$$E(x, v) = (x, \exp_x v).$$

The differential at $(p, 0)$ is given by

$$dE(p,0) = \begin{bmatrix} \frac{\partial x^i}{\partial x^j} & \frac{\partial x^i}{\partial v^j} \\ \frac{\partial \exp^i}{\partial x^j} & \frac{\partial \exp^i}{\partial v^j} \end{bmatrix} |_{(p,0)} = \begin{bmatrix} Id & 0 \\ \star & Id \end{bmatrix}$$

which is invertible. By the inverse function theorem E is a diffeomorphism in a neighbourhood of $(p, 0)$; that is, there exists a neighbourhood $V \in TM$ of $(p, 0)$

which maps diffeomorphically onto a neighbourhood $W' \subset U_o \times U_o$ of (p, p). Choose a $\delta > 0$ so that

$$\{(q, v) \in V : \|v\| < \delta\} = V' \subseteq V$$

and choose any neighbourhood W of p such that $W \times W \subset E(V') \subset W'$. Now by construction if $q \in W$ and $B_\delta(0) \subset T_q M$ then

$$\{q\} \times W \subset E(\{q\} \times B_\delta(0)) = \{q\} \times \exp_q(B_\delta(0)) \implies W \subset \exp_q(B_\delta(0)).$$

$\square$

Chapter 7

Interlude

Let's briefly recap what we've seen so far.

A topology is and how it gives us a notion of "closeness", continuity, and convergence and we saw how this allows us to: (1) make precise the idea "locally looks like Euclidean space"; and (2) provided a definition of of Topological Manifold. We saw what a diffeomorphism is and how this: (a) gives rise to the notion of a smooth structure; and (b) provided a definition of Smooth Manifold.

We defined tangent vector and tangent space at a point using derivations - a linear operator that acts on functions by differentiation. This allowed us to show that the tangent space at a point and the tangent bundle are smooth manifolds themselves; and gave rise to the differential of a map F between two manifolds, which can be interpreted as a smooth map between the respective tangent spaces. In fact, almost everything we have constructed over the past few chapters was built out of derivations and differentials.

We introduced Riemannian metrics which give us a way of measuring length and angle on a smooth manifold. We introduced connections in order to give a well-defined meaning to acceleration in a smooth manifold and showed that. Each Riemannian metric defines a unique connection called the Levi-Civita connection which best reflects the geometric properties of the metric. Riemannian geodesics are those curves whose acceleration is zero with respect to the Levi-Civita connection.

To better understand the collective behaviour of geodesics we introduced the exponential map and saw that for each $(p, v) \in TM$ the geodesic beginning at p in the direction v is given by

$$\gamma_{v_p}(t) = \exp_p(tv)$$

and the differential of the exponential map at 0 is the identity map which, together with the inverse functions theorem, implies that $\exp_p$ is a local diffeomorphism.

So where are we going from here? At no point have we talked about geodesics in terms of distance so what we're going to show is that given a metric g the curves whose acceleration is zero with respect to the Levi-Civita connection locally minimize distance between points. We'll do the converse too and show that length minimizing curves also have zero acceleration with respect to the Levi-Connection. This will take us to the Hopf-Rinow Theorem which describes the relationship between metric and geodesic properties of a manifold

Then we'll begin a study of local invariants of Riemannian metrics - this is "curvature" which measures the extent to which second covariant derivatives fail to commute. We'll show that a Riemannian manifold has zero curvature if and only if it's locally isometric to Euclidean space - the hands on definition of "flat" we saw right at the start. We'll then study the way curvature affects the behaviour of geodesics which gives an entry into the topological information contained in curvature. Finally we'll see the Cartan-Hadamard Theorem and some related results.

For the remainder of this interlude let's see how the geodesics in some familiar Riemannian spaces look...

Euclidean Space

Take $\mathbb{R}^n$ with the Euclidean metric given by the $n \times n$ identity matrix I. We saw a while back that the Levi-Civita connection is the standard directional derivative and the connection coefficients are zero

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} \left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right) = 0.$$

Consequently, the geodesic equation takes the form

$$\frac{d^2}{dt^2} x(t) = 0$$

whose solutions are straight lines.

Spheres

Recall that the round metric on the sphere $\mathbb{S}^n$ is induced by the ambient Euclidean metric (see Riemannian Submanifolds and Induced Metrics - Chapter 4). The Levi-Civita connection is given by the tangential connection

$$\nabla_X^\top Y = \pi^\top \left(\nabla_{\tilde{X}} \tilde{Y}|_{\mathbb{S}^n} \right)$$

where ∇ is the Euclidean connection, $\pi^\top$ is orthogonal projection onto the tangent space $T\mathbb{S}^n$ (see section Chapter 6). Furthermore, if $\gamma : I \to \mathbb{S}^n$ is a smooth curve and $V(t)$ is a smooth vector field along γ taking values in $T\mathbb{S}^n$ then the tangential covariant derivative of V along γ is

$$D_t^\top V(t) = \pi^\top (D_t V)$$

where D_t is the Euclidean covariant derivative of V along γ. To see this, pick any $t_o \in I$, look at a sufficiently small open neighbourhood U around $\gamma(t_o)$ in $\mathbb{R}^{n+1}$ and pick an orthonormal frame $\{E_1, \dots, E_{n+1}\}$ such that $\{E_1, \dots, E_n\}$ restricts to an orthonormal frame for $T\mathbb{S}^n$ at points in $\mathbb{S}^n \cap \mathbb{R}^{n+1}$. If $\epsilon > 0$ is small enough then $\gamma((t_o - \epsilon, t_o + \epsilon)) \subseteq U$ and for $t \in (t_o - \epsilon, t_o + \epsilon)$ we can write $V(t) = V^1(t) E_1|_{\gamma(t)} + \ldots + V^n(t) E_n|_{\gamma(t)}$ for some smooth functions

$V^1, \dots, V^n : (t_o - \epsilon, t_o + \epsilon) \to \mathbb{R}$. Then

$$\begin{aligned}\pi^\top (D_t V) &= \pi^\top \left(\sum_{i=1}^n \dot{V}^i(t) E_i|_\gamma + V^i \nabla_{\gamma'} E_i|_\gamma \right) \\ &= \left(\sum_{i=1}^n \dot{V}^i(t) E_i|_\gamma + V^i \pi^\top \nabla_{\gamma'} E_i|_\gamma \right) \\ &= D_t^\top V\end{aligned}$$

This shows that $\gamma(t)$ is a geodesic (i.e. $D_t^\top \gamma'(t)$) if and only if its ordinary acceleration is orthogonal to $T\mathbb{S}^n$. The above calculation and conclusion holds for any embedded Riemannian manifold $M \subseteq \mathbb{R}^n$.

To find the geodesics on $\mathbb{S}^n$ we do the following: the function $f(x) = \|x\|^2$ defines the unit sphere via $f^{-1}(1)$. From this we see that

$$T_p\mathbb{S}^n = \ker df_p$$

(remember f is constant on $\mathbb{S}^n$ so it's differential is zero in directions tangent to $\mathbb{S}^n$). To describe $T_p\mathbb{S}^n$ we calculate the differential using a curve $x(t) \subset \mathbb{S}^n$ passing through p in the direction $v \in T_p\mathbb{S}^n$:

$$0 = df_p v = \frac{d}{dt}|_{t=0} f(x(t)) = \frac{d}{dt}|_{t=0} \langle x(t), x(t) \rangle = 2\langle v, p \rangle.$$

So $T_p\mathbb{S}^n$ is the set of vectors orthogonal to p (thought of as a vector in $\mathbb{R}^{n+1}$). Pick $p \in \mathbb{S}^n$ and any $v \in T_p\mathbb{S}^n$ and set $a = \|v\|$ and $\hat{v} = \frac{v}{a}$. Consider the smooth curve $\gamma : \mathbb{R} \to \mathbb{R}^{n+1}$ given by

$$\gamma(t) = \cos(at)p + \sin(at)\hat{v}$$

satisfying $\gamma(0) = p$, $\gamma'(0) = v$. Since $\hat{v}$ is orthogonal to p and both have unit length its clear that $\|\gamma(t)\| = 1$ for all t so that $\gamma(t) \subset \mathbb{S}^n$. Furthermore

$$\begin{aligned}\gamma'(t) &= -a\sin(at)p + a\cos(at)\hat{v} \\ \gamma''(t) &= -a^2\cos(at)p - a^2\sin(at)\hat{v}\end{aligned}$$

which shows that γ'' is proportional to γ - considered as position vectors in $\mathbb{R}^{n+1}$ - and is therefore orthogonal to $T\mathbb{S}^n$. That is, γ is a geodesic. Each γ is periodic with period $\frac{2\pi}{\|v\|}$ and has constant speed. The image of $\gamma(t)$ is the intersection of $\mathbb{S}^n$ with the linear subspace spanned by $\{p, \hat{v}\}$. Conversely, take any orthonormal vectors $\{p, \hat{v}\}$ spanning a linear subspace Π. Then the curve C formed by $\Pi \cap \mathbb{S}^n$ has the form of γ and is a geodesic.

Hyperbolic Upper-Half Space

The Poincare upper half space model $\mathbb{U}^n(R)$ is the upper half-space in $\mathbb{R}^n$ defined in coordinates $(x^1, x^2, \dots, xYn-1, y)$ by $\mathbb{U}^n(R) = \{(x, y) : y > 0\}$ endowed with the metric

$$g = \frac{R}{y^2} I_n$$

where I_n is the $n \times n$ identity matrix. Restricting to the case $n = 2$ we can calculate the connection coefficients using

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right).$$

Applying the formula gives

$$\begin{aligned} \Gamma^x_{xx} &= \Gamma^y_{xy} = \Gamma^x_{yy} = 0, \\ \Gamma^y_{xx} &= \frac{1}{y} \\ \Gamma^x_{xy} &= \Gamma^y_{yy} = -\frac{1}{y}. \end{aligned}$$

Recall that the geodesic equation is given as a first order system

$$\begin{aligned} \dot{x}^i &= v^i \\ \dot{v}^i &= -v^i v^j \Gamma^k_{ij} \end{aligned}$$

Letting $\eta(t) = (x(t), y(t))$ and $\dot{\eta}(t) = (u(t), v(t))$ the geodesic equation can be written as...

$$\begin{aligned} \dot{x}(t) &= u(t) \quad & \dot{y}(t) &= v(t) \\ \dot{u}(t) &= \frac{2}{y} uv \quad & \dot{v} &= \frac{1}{y}\left(v^2 - u^2\right) \end{aligned}$$

and we'll assume further that this geodesic has unit speed:

$$R\frac{\dot{x}^2 + \dot{y}^2}{y^2} = 1.$$

Using some algebra we can rewrite the equation $\dot{u}(t) = \frac{2}{y} uv$ as

$$\frac{\ddot{x}y^2 - 2\dot{x}\dot{y}y}{y^4} = 0$$

which folds as

$$\frac{d}{dt}\left(\frac{\dot{x}}{y^2}\right) = 0$$

so that $\dot{x} = Cy^2$ for some constant C. So there are two possibilities:

a. $C = 0$; and

b. $C \neq 0$.

If $C = 0$ then solutions to the geodesic equation are are vertical lines in the plane. If $C \neq 0$ then using (**??**) and the unit speed assumption

$$R\frac{\dot{x}^2 + \dot{y}^2}{y^2} = 1$$

$$\left(\frac{dy}{dx}\right)^2 = \frac{\dot{y}^2}{\dot{x}^2} = \frac{y^2}{R\dot{x}^2} - 1 = \frac{1}{RC^2y^2} - 1 = \frac{a - y^2}{y^2}$$

with $a = \left(RC^2\right)^{-1}$. The solution of this equation (in the upper half space) is a semi-circle centred at the origin of radius $\sqrt{a}$. Indeed, a circle of radius $\sqrt{a}$ centered at the origin is written as

$$x^2 + y^2 = a$$

and differentiating with respect to x gives

$$2x + 2y\frac{dy}{dx} = 0.$$

Solving for $\frac{dy}{dx}$, squaring, and substituting in the original equation for x^2 gives

$$\left(\frac{dy}{dx}\right)^2 = \frac{x^2}{y^2} = \frac{a - y^2}{y^2}.$$

So we have found the geodesics of in the hyperbolic upper-half plane are either vertical lines or semi-circles centered at the origin.

Chapter 8

The Hopf-Rinow Theorem

Lengths and Distances

A curve γ in M always means a continuous map

$$\gamma : I \subseteq \mathbb{R} \to M$$

To say that γ is a smooth curve is to say it is a smooth map from the manifold I into the manifold M. A smooth curve γ has a well defined velocity vector $\gamma'(t) \in T_{\gamma(t)}M$. We say that γ is a regular curve if it is smooth and $\gamma'(t) \neq 0$ for all $t \in I$.

If $[a,b] \subset \mathbb{R}$ is a closed and bounded interval, a partition of $[a,b]$ is a finite sequence $(a_o, ..., a_k)$ of real numbers such that

$$a = a_o < a_1 < \cdots < a_k = b.$$

Each interval $[a_{i-1}, a_i]$ is called a subinterval of the partition. A curve $\gamma : [a,b] \to M$ is said to be piecewise regular if there exists a partition $(a_o, \ldots, a_k)$ of $[a,b]$ such that $\gamma|_{[a_{i-1},a_i]}$ is regular (in the sense described above). We refer to a piecewise regular curve as an admissible curve and the corresponding partition as an admissible partition. It's convenient to refer to a map $\gamma : a \to M$ whose domain is a single point as an admissible curve.

Suppose γ is an admissible curve and $(a_o, \ldots, a_k)$ an admissible partition. At each of the points $a_1, \ldots, a_{k-1}$ there are two one-sided velocity vector, which we denote by

$$\gamma'(a_i^-) = \lim_{t \nearrow a_i} \gamma'(t)$$
$$\gamma'(a_i^+) = \lim_{t \searrow a_i} \gamma'(t).$$

They are both non-zero but they need not be equal.

For a smooth curve $\gamma : I \to M$ we define a reparameterisation of γ as a curve of the form

$$\tilde{\gamma} = \gamma \circ \phi : I' \to M$$

where I' is another interval and $\phi : I \to I'$ a diffeomorphism. Because intervals are connected ϕ is either strictly increasing or strictly decreasing on I'. We say that $\tilde{\gamma}$ is a forward parameterisation if ϕ is increasing and a backward parameterisation if ϕ is decreasing. For an admissible curve $\gamma : [a,b] \to M$ we define a reparameterisation of γ as a curve of the form

$$\tilde{\gamma} = \gamma \circ \phi : [c,d] \to M$$

where $\phi : [c,d] \to [a,b]$ is a homeomorphism for which there exists a partition $(c_o, \ldots, c_k)$ of $[c,d]$ such that the restriction of ϕ to each subinterval $[c_{i-1}, c_i]$ is a diffeomorphism onto its image.

If $\gamma : [a,b] \to M$ is an admissible curve in a Riemannian manifold (M,g) then the length of γ is

$$L_g(\gamma) = \int_a^b \|\gamma'(t)\|_g \, dt$$

Proposition 8.1. *Suppose (M,g) is a Riemannian manifold and $\gamma : [a,b] \to$ is an admissible curve.*

a. *ADDITIVITY of LENGTH: If $a < c < b$ then*

$$L_g(\gamma) = L_g(\gamma|_[a,c]) + L_g(\gamma|_[c,b])$$

b. *REPARAMETERISATION INVARIANCE: If $\tilde{\gamma}$ is a reparameterisation of γ, then $L_g(\tilde{\gamma}) = L_g(\gamma)$.*

c. *ISOMETRY INVARIANCE: If (M,g) and $(\tilde{M}, \tilde{g})$ are Riemannian manifolds and $\phi : M \to \tilde{M}$ is a local isometry then $L_{\tilde{g}}(\phi \circ \gamma) = L_g(\gamma)$*

Prove it yourself!

Suppose $\gamma : [a,b] \to M$ is an admissible curve. The arc-length function of γ is the function $s : [a,b] \to M$ defined by

$$s(t) = L_g(\gamma|_{[a,t]}) = \int_a^t \|\gamma'(r)\|_g \, dr.$$

If $\gamma : I \to M$ is smooth we define the speed of γ at any time $t \in I$ to be the scalar $\|\gamma'\|_g$:

a. γ has unit speed if $\|\gamma'\|_g = 1$; and

b. γ has constant speed if $\|\gamma'\|_g$ is a constant.

If $\gamma : [a,b] \to M$ is an admissible curve then we say it has unit speed if $\|\gamma'\|_g = 1$ wherever it is smooth. If $\gamma : [a,b] \to M$ is a unit speed admissible curve then its arc-length function takes the simple form

$$s(t) = L_g(\gamma|_{[a,t]}) = \int_a^t \|\gamma'(r)\|_g \, dr = \int_a^t dr = t - a.$$

If $a = 0$ then

$$s(t) = t$$

and we say that γ is parameterised by arc-length.

Proposition 8.2. *Suppose (M,g) is a Riemannian manifold.*

a. *Every regular curve in M has a unit speed forward parameterisation*

b. *Every admissible curve in M has a unique forward reparameterisation by arc-length.*

Proof. *a.* For the first part, the arc-length function is locally a diffeomorphism and s^{-1} gives a unit speed forward parameterisation.

b. The second part is proved by induction on the number of partitions - the first part being the verification step for a single partition. If ϕ and ψ are two forward reparameterisations by arc-length then the unit speed property implies $\phi^{-1} \circ \psi = Id$ - which gives uniqueness.

□

The Riemannian Distance Function

Definition 8.1. *Suppose (M, g) is a connected Riemannian manifold. For each pair of points p, $q \in M$ we define the Riemannian distance between p and q, denoted by $d_g(p, q)$, to be the infimum of the lengths of all admissible curves joining p with q.*

The function $d_g(p, q)$ is a well-defined, non-negative real number for all $p, q \in M$.

If $\phi : M \to \tilde{M}$ is a global isometry between two Riemannian manifolds (M, g) and $(\tilde{M}, \tilde{g})$ then it's fairly clear that $d_{\tilde{g}}(\phi(p), \phi(q)) = d_{\tilde{g}}(\phi(p), \phi(q))$ for all p, $q \in M$; that is, Riemannian distances are preserved by global isometries. However, if ϕ is only a local isometry then not all Riemannian distances are preserved, possibly because two distinct points may have the same image - for example $\mathbb{R}^2$ and $\mathbb{T}^2$.

The Riemannian distance function turns the Riemannian manifold (M, g) into a metric space whose metric topology is the same as the original manifold topology. To see this we need a few Lemmas...

Lemma 8.3. *Let g be a Riemannian metric on an open subset $W \subseteq \mathbb{R}^n$ and let $\bar{g}$ denote the Euclidean metric on W. For every compact subset $K \subset W$ there are positive constants c and C such that for all $x \in K$ and for $v \in T_x\mathbb{R}^n$*

$$c \|v\|_{\bar{g}} \leq \|v\|_g \leq C \|v\|_{\bar{g}}$$

This inequality allows us to say that points are close with respect to g if and only if they are close with respect to $\bar{g}$. Or, put another way, the open balls defined by g and $\bar{g}$ are "equivalent"/"comparable".

Proof. Define $F : TW \to \mathbb{R}^n$ by $F(x, v) = \|v\|_g = \sqrt{g(x)(v, v)}$ for $x \in W$ and $v \in T_xW \cong T_x\mathbb{R}^n$. Let $K \subset W$ be any compact subset and define $L = \{(x, v) \in T\mathbb{R}^n : x \in K, \|v\|_{\bar{g}} = 1\} = K \times \mathbb{S}^{n-1}$ Since L is compact and F is positive on L there is a positive minimum c and a positive maximum C:

$$c \leq \|v\|_g \leq C \qquad \forall (x, v) \in L.$$

For any $x \in K$ and any $v \in T_xW \cong T_x\mathbb{R}^n$ we can write $v = \|v\|_{\bar{g}}\, \hat{v}$ and $\hat{v} \in \mathbb{S}^{n-1}$. Now $(x, \hat{v}) \in L$ so that $\hat{v}$ satisfies the previous inequality and therefore

$$c \leq \frac{1}{\|v\|_{\bar{g}}} \|v\|_g \leq C.$$

□

Now we transfer the previous Lemma to manifolds...

Lemma 8.4. *Let (M, g) be a Riemannian manifold and let d_g be its distance function. Suppose U is an open susbet of M and and $p \in U$. Then p has a coordinate neighbourhood $V \subseteq U$ with the property that there are positive constants C and D such that*

a. *If $q \in V$ then $d_g(p, q) \leq C d_{\bar{g}}(p, q)$ where $\bar{g}$ is the Euclidean metric in the given coordinates on V;*

b. *If $q \in V^c$ then $d_g(p, q) \geq D$*

Proof. Let W be any neighbourhood of p contained in U on which there exist smooth coordinates (x^i) and identify W with an open subset of $\mathbb{R}^n$. Let K be the closed Euclidean ball of radius ϵ so that $K \subset W$ and let V be the open Euclidean ball of radius ϵ. If $\bar{g}$ denotes the Euclidean metric then the previous lemma gives $c\|v\|_{\bar{g}} \leq \|v\|_g \leq C\|v\|_{\bar{g}}$ for positive constants c and C so that for any admissible curve whose image lies entirely in V we get $cL_{\bar{g}}(\gamma) \leq L_g(\gamma) \leq CL_{\bar{g}}(\gamma)$. Now if $q \in V$ and γ is a parameterisation of the straight line segment from p to q in V then

$$d_g(p, q) \leq L_g(\gamma) \leq CL_{\bar{g}}(\gamma) = Cd_{\bar{g}}(p, q).$$

For part 2. suppose $q \in V^c$. If $\gamma : [a, b] \to M$ is an admissible curve then suppose that t_o is the infimum of times $t \in [a, b]$ for which $\gamma(t) \in V$ - this means $\gamma(t_o) \in K$ and $d_{\bar{g}}(p, \gamma(t_o)) = \epsilon$. Then

$$L_g(\gamma) \geq L_g(\gamma|_{[a,t_o]}) \geq cL_{\bar{g}}(\gamma|_{[a,t_o]}) \geq c\epsilon = D.$$

□

Now we have enough

Theorem 8.5. *Let (M, g) be a Riemannian manifold. With the distance function d_g, M is a metric space whose metric topology is the same as the original manifold topology.*

Proof. It's pretty clear that $d_g(p, q) = d_g(q, p) \geq 0$ and $d_g(p, p) = 0$. Since the manifold topology is Hausdorff, for any p and $q \in M$ we can find disjoint neighbourhoods U and V such that $p \in U$ and $q \in V$. Part 2 of the last lemma then implies $d_g(p, q) > 0$ - nondegeneracy. The triangle inequality follows from the basic properties of the length function. So d_g is indeed a metric. We just need to show that d_g induces the same topology as the manifold topology. To do this part we need to show that a set U which is open in the manifold topology is also open with respect to the metric topology induced by d_g - that is, there exists an open ball $B_\epsilon(p) = \{q \in M : d_g(p,q) < \epsilon\} \subset U$ around every point $p \in U$ - and vice versa.

Let U be open in the manifold topology. For each $p \in U$ we can choose a coordinate neighbourhood $V \subset U$ around p with positive constants C and D satisfying the conclusions of the previous lemma:

a. If $q \in V$ then $d_g(p,q) \leq C d_{\bar{g}}(p,q)$ where $\bar{g}$ is the Euclidean metric in the given coordinates on V;

b. If $q \in V^c$ then $d_g(p,q) \geq D$.

Set $B_\epsilon(p) = \{q \in M : d_g(p,q) < \epsilon\}$ with $\epsilon < D$. The contrapositive of 2. implies every $q \in B_\epsilon(p)$ is also in V. Therefore $B_\epsilon(p) \subset U$ and U is open with respect to d_g.

Now suppose that U' is open in the metric topology. Given $p \in U'$ choose $\delta > 0$ such that the d_g-metric ball of radius δ around p is contained in U'. Let V be any neighbourhood of p that is open in the manifold topology and satisfies the conditions of the previous lemma with positive constants C and D:

a. If $q \in V$ then $d_g(p,q) \leq C d_{\bar{g}}(p,q)$ where $\bar{g}$ is the Euclidean metric in the given coordinates on V;

b. If $q \in V^c$ then $d_g(p,q) \geq D$

Note that we have NOT assumed $V \subset U'$. Choose ϵ small enough so that $C\epsilon < \delta$. If $q \in V$ with $d_{\bar{g}}(p,q) < \epsilon$ then

$$d_g(p,q) \leq C d_{\bar{g}}(p,q) < C\epsilon < \delta$$

and therefore q lies in the metric ball of radius δ around p and therefore in U'. Since the set $\{q \in V : d_{\bar{g}}(p,q) < \epsilon\}$ is open in the given manifold topology this shows that U' is also open in the manifold topology. □

Geodesics and Minimizing Curves

An admissible curve γ in a Riemannian manifold (M, g) is said to be length minimizing if

$$L_g(\gamma) \leq L_g(\tilde{\gamma})$$

for every other admissible curve $\tilde{\gamma}$ joining the same end-points. We're going to see that all length minimizing curves are geodesics of g. To do this we're going to think of the length L_g as a function on the set of all admissible curves in M joining a point p with a point q. We will then characterize the minima of this function. Given intervals I, $I' \subset \mathbb{R}$, a continuous map $\Gamma : I' \times I \to M$ is called a one-parameter family of curves. Such a family defines two types of curves:

a. The main curves $\Gamma_s(t) := \Gamma(s,t)$ defined for $t \in I$ by holding s constant; and

b. The transverse curves $\Gamma_t(s) := \Gamma(s,t)$ defined for $s \in I'$ by holding t constant.

If this family is smooth we denote the velocity vectors of the main and transverse curves by

$$\partial_t \Gamma(s,t) := (\Gamma_s)'(t) \in T_{\Gamma(s,t)}M \qquad \partial_s \Gamma(s,t) := (\Gamma_t)'(t) \in T_{\Gamma(s,t)}M$$

which are examples of vector fields $V : I' \times I \to T_{\Gamma(s,t)}M$ along Γ. A one parameter family Γ is called an admissible family of curves if

a. its domain is of the form $I' \times [a,b]$ for some open interval I';

b. there is a partition $(a_o, \ldots, a_k)$ of $[a,b]$ such that Γ is smooth on each rectangle $I' \times [a_{i-1}, a_i]$

c. $\Gamma_s(t)$ is an admissible curve for each $s \in I'$.

If $\gamma : [a,b] \to M$ is an admissible curve, a variation of γ is an admissible family $\Gamma : I' \times [a,b] \to M$ such that I' is an open interval containing 0 and $\Gamma_0(t) = \gamma(t)$. It is called a proper variation if all the main curves have the same starting and end points.

In the case of an admissible family, the transverse curves are smooth on I' for each fixed t but the mains curves are generally only piecewise regular on I for each fixed s. So the velocity vectors

$$\partial_t \Gamma(s,t) := (\Gamma_s)'(t) \in T_{\Gamma(s,t)}M \qquad \partial_s \Gamma(s,t) := (\Gamma_t)'(t) \in T_{\Gamma(s,t)}M$$

are smooth on each rectangle $I' \times [a_{i-1}, a_i]$ but not on the whole domain. A piecewise smooth vector field along γ is a vector field whose restriction to each rectangle $I' \times [a_{i-1}, a_i]$ is smooth. The vector field $\partial_s \Gamma$ is always a piecewise smooth vector field but $\partial_t \Gamma(s,t)$ is not continuous at the partition points a_j. A piecewise smooth vector field along γ is a continuous vector field whose restriction to each rectangle $I' \times [a_{i-1}, a_i]$ is smooth. The vector field $\partial_s \Gamma$ is always a piecewise smooth vector field but $\partial_t \Gamma(s,t)$ is not continuous at the partition points a_j. To see its continuity on the whole domain note that for each partition point a_i the value of $\partial_s \Gamma$ along $I' \times a_i$ depends only on the values of Γ on that set. Since $\partial_s \Gamma$ is smooth on $I' \times [a_{i-1}, a_i]$ and $I' \times [a_i, a_{i+1}]$, by definition, the left and right hand limits at a_i must be equal - so it's continuous. The vector field $\partial_t \Gamma$ is typically not continuous. If Γ is a variation of γ, the variation field of Γ is the piecewise smooth vector field $V(t) = \partial_s|_{s=0} \Gamma(s,t)$ along γ. We say that a vector field V along γ is proper if $V(a) = V(b) = 0$. If γ is an admissible curve and V is a piecewise smooth vector field along γ then V is the variation field of some variation of γ. If V is proper then the variation can be taken to be proper aswell.

Lemma 8.6. *If γ is an admissible curve and V is a piecewise smooth vector field along γ then V is the variation field of some variation of γ. If V is proper then the variation can be taken to be proper aswell.*

Proof. $\Gamma(s,t) = \exp_{\gamma(t)}(sV(t))$ is the variation we want. Since the exponential map is smooth, the composition with V is smooth wherever V is smooth and it's continuous on the whole domain. Since $D\exp_p(0) = I$, V is the variation field of this variation and if V is proper then $\Gamma(s,a) = \exp_{\gamma(a)}(0) = \gamma(a)$ and similarly for $t = b$ so that Γ is proper. □

If V is piecewise smooth along Γ we can compute the covariant derivatives along either the main curves (wherever they are smooth) or along the transverse curves. The resulting covariant derivatives will be denoted D_tV and D_sV, respectively. A key ingredient in the proof that length minimizing curves are geodesics is the symmetry of the Levi-Civita connection. In particular...

Lemma 8.7. *Let $\Gamma : I' \times [a,b] \to M$ be an admissible family of curves in a Riemannian manifold. On every rectangle $I' \times [a_{i-1}, a_i]$ where Γ is smooth we have*

$$D_t\partial_s\Gamma = D_s\partial_t\Gamma.$$

Proof. This is a local statement so pick coordinates around a point $\Gamma(s_o, t_o)$ and in these coordinates write $\Gamma(s,t) = (x^1(s,t), \ldots, x^n(s,t))$. Then we have

$$\partial_s\Gamma = \frac{\partial x^k}{\partial s}\partial_k \qquad \partial_t\Gamma = \frac{\partial x^k}{\partial t}\partial_k$$

and in coordinate the covariant derivatives are

$$\begin{aligned} D_t\partial_s\Gamma &= \left(\frac{\partial^2 x^k}{\partial t\partial s} + \frac{\partial x^i}{\partial t}\frac{\partial x^j}{\partial s}\Gamma^k_{ij}\right)\partial_k, \quad D_s\partial_t\Gamma \\ &= \left(\frac{\partial^2 x^k}{\partial s\partial t} + \frac{\partial x^i}{\partial s}\frac{\partial x^j}{\partial t}\Gamma^k_{ij}\right)\partial_k \end{aligned}$$

Symmetry of the connection coefficients means I can swap i for j in one of the expressions and then you see the two quantities are the same. □

To show that length minimizing curves are geodesics we're going to understand critical points of the length function L_g. This means taking a variation Γ of an admissible curve γ and differentiating $L_g(\Gamma)$ with respect to the variation parameter s at $s = 0$. If γ is length minimizing it must be a local minimum of the length-function and the derivative from the previous step is zero. This then gives a condition on the acceleration of γ which turns out to be the geodesic condition.

Theorem 8.8. *Let (M,g) be a Riemannian manifold. Suppose γ is a unit speed admissible curve, $\Gamma : I' \times [a,b] \to M$ is a variation of γ, and V is its variation field. Then $L_g(\Gamma_s)$ is a smooth function of s and*

$$\frac{d}{ds}|_{s=0} L_g(\Gamma_s) = -\int_a^b g\left(V, D_t\gamma'\right) dt - \sum_{i=1}^{k-1} g\left(V(a_i), \Delta_i\gamma'\right) + g\left(V(b), \gamma'(b)\right) - g\left(V(a), \gamma'(a)\right)$$

where $(a_o, \ldots, a_k)$ is an admissible partition for V, and for each $i = 1, \ldots, k-1$, $\Delta_i\gamma' = \gamma'(a_i^+) - \gamma'(a_i^-)$ is the "jump" in the velocity vector γ at a_i. If Γ is a proper variation then

$$\frac{d}{ds}|_{s=0} L_g(\Gamma_s) = -\int_a^b g\left(V, D_t\gamma'\right)\, dt - \sum_{i=1}^{k-1} g\left(V(a_i), \Delta_i\gamma'\right)$$

Proof. Let

$$T = \partial_t\Gamma(s,t) \qquad S = \partial_s\Gamma(s,t).$$

Differentiating on the interval $[a_{i-1}, a_i]$, using the chain rule, compatibility of the metric and Levi-connection, and finally symmetry:

$$\begin{aligned} \frac{d}{ds}|_{s=0} L_g(\Gamma_s) = \int_{a_{i-1}}^{a_i} \frac{\partial}{\partial s} g(T,T)^{\frac{1}{2}}\, dt &= \int_{a_{i-1}}^{a_i} \frac{1}{2g(T,T)^{\frac{1}{2}}} \cdot \frac{\partial}{\partial s} g(T,T)\, dt \\ &= \int_{a_{i-1}}^{a_i} \frac{1}{2\,\|T\|} 2g(D_sT, T)\, dt \\ &= \int_{a_{i-1}}^{a_i} \frac{1}{\|T\|} g(D_tS, T)\, dt \end{aligned}$$

Setting $s = 0$ and noting that $S(0,t) = V(t)$ and $T(0,t) = \gamma'(t)$ has length 1, we proceed to integrate by parts and then apply the fundamental theorem of calculus.

$$\begin{aligned} \frac{d}{ds}|_{s=0} L_g(\Gamma_s) &= \int_{a_{i-1}}^{a_i} g(D_tV, \gamma')\, dt \\ &= \int_{a_{i-1}}^{a_i} \left(\frac{d}{dt} g(V,\gamma') - g(V, D_t\gamma')\right) dt \\ &= -\int_{a_{i-1}}^{a_i} g\left(V, D_t\gamma'\right)\, dt \\ &\quad + g(V(a_i, \gamma'(a_i)) - g(V(a_{i-1}, \gamma'(a_{i-1})). \end{aligned}$$

Summing over i and collecting terms gives the formula in the Theorem. □

Theorem 8.9. *In a Riemannian manifold, every length minimizing curve is a geodesic when it is given a unit-speed parameterisation.*

Proof. Suppose γ is minimizing and of unit speed, and $(a_o, \dots, a_k)$ is an admissible partition of γ. If Γ is any proper variation of γ, then $L_g(\Gamma_s)$ achieves its minimum at $s = 0$ so that

$$\frac{d}{ds}|_{s=0} L_g(\Gamma_s) = 0$$

This means that

$$-\int_a^b g\left(V, D_t\gamma'\right)\, dt - \sum_{i=1}^{k-1} g\left(V(a_i), \Delta_i\gamma'\right) = 0.$$

for every proper variation field.

We first show that $D_t\gamma' = 0$ on each subinterval $[a_{i-1}, a_i]$ so that γ is a broken geodesic. Let $\phi \in C^\infty$ be a smooth function with $\phi > 0$ on (a_{i-1}, a_i) and $\phi = 0$ everywhere else. Picking $V = \phi D_t\gamma'$ we get

$$0 = -\int_a^b g\left(\phi D_t\gamma', D_t\gamma'\right)\, dt - \sum_{i=1}^{k-1} g\left(0, \Delta_i\gamma'\right) = -\int_a^b \phi \left\|D_t\gamma'\right\|\, dt$$

so that $D_t\gamma' = 0$ on each sub-interval $[a_{i-1}, a_i]$.

Next we show that γ has no corners, i.e. $\Delta_i\gamma' = 0$ for each i between 0 and k.

We already know that $D_t\gamma' = 0$ on each sub-interval $[a_i, a_{i+1}]$ so for each i take a coordinate chart around a_i and a function $\phi \in C^\infty$ with $\phi > 0$ in a very small neighbourhood of a_i and 0 outside.

Set $V(a_i) = \Delta_i\gamma'$ and define $V(t) = \phi \cdot \tau(t) \cdot V(a_i)$ - parallel transport along γ.

Then

$$\begin{aligned} 0 = -\int_{a_i-\epsilon}^{a_i+\epsilon} g(V(t), D_t\gamma')\, dt - \sum_{i=1}^{k-1} g\left(V(a_i), \Delta_i\gamma'\right) \\ = \phi(a_i) g\left(\Delta_i\gamma', \Delta_i\gamma'\right) \end{aligned}$$

This implies that $\|\Delta_i\gamma'\| = 0$ and hence $\Delta_i\gamma' = 0$ and shows that the two one-sided velocity vectors of γ' match up at the partition points. From the uniqueness of geodesics this shows that $\gamma|_{[a_i,a_{i+1}]}$ is the continuation of the geodesic $\gamma|_{[a_{i-1},a_i]}$ and therefore γ is smooth. □

Geometrically, the above proof says that if $D_t\gamma' \neq 0$ on $[a_i, a_{i+1}]$ then if we take the variation field $V = \phi D_t\gamma'$, which means we deform γ in the direction of its acceleration, then the length function L_g is decreasing (its derivative is negative). That is, deforming curves in the direction of their acceleration decreases their length. Similarly, the length of a "broken" geodesic is decreased if we deform it in the direction of a vector field with $V(a_i) = \Delta_i\gamma'$ - which corresponds to smoothing off, or rounding, corners.

The literal converse of the previous theorem is false: just think about the sphere $\mathbb{S}^{n-1}$. But what is true is that geodesics locally minimize length. What we're going to do is show that in a sufficiently small ball $\mathcal{B}$ around a point p, the geodesics are the curves which minimize the length between p and any point $q \in \mathcal{B}$.

A curve $\gamma : I \to M$ is said to be locally minimizing if every $t_o \in I$ has a neighbourhood $I_o \subseteq I$ such that whenever $a, b \in I_o$ with $a < b$ the restriction of γ to $[a, b]$ is length minimizing.

If ϵ is a positive number such that $\exp_p$ is a diffeomorphism from the ball $B_\epsilon(0) \subset T_pM$ to its image then $\exp_p(B_\epsilon(0))$ is a normal neighbourhood called a geodesic ball. If the closed ball $\bar{B}_\epsilon(0)$ is contained in an open set $V \subset T_pM$ on which $\exp_p$ is a diffeomorphism onto its image then $\exp_p\left(\bar{B}_\epsilon(0)\right)$ is a closed geodesic ball and $\exp_p(\partial B_\epsilon(0))$ is called a geodesic sphere.

Lemma 8.10. *Let $p \in M$ and $v \in T_pM$ such that $\exp_p(v)$ is defined and let $w \in T_vT_pM \equiv T_pM$. Then*

$$g_{\exp_p(v)}(D\exp_p(v) \cdot v, D\exp_p(v) \cdot w) = g_p(v, w)$$

Remark 8.1. *This says that $D\exp_p$ is a "radial" isometry meaning the angle between vectors w and v are preserved along geodesics $\exp_p(tv)$.*

Proof. Decompose w as $w = w_T + w_N$ where $w_T = \lambda \cdot v$ is parallel to v and w_N is orthogonal to v. We're first going to show that

$$g_{\exp_p(v)}(D\exp_p(v) \cdot v, D\exp_p(v) \cdot w_T) = g_p(v, w_T).$$

If $\gamma_{v_p}(t) = exp_p(tv)$ is the geodesic beginning at p in the direction v then

$$D\exp_p(v) \cdot v = \frac{d}{dt}|_{t=1} \exp_p(tv) = \frac{d}{dt}|_{t=1}\gamma_{v_p}(t) = \gamma'_{v_p}(1)$$

Since $\left\|\gamma'_{v_p}(t)\right\|^2$ is constant

$$\begin{aligned}
g_{\exp_p(v)}&(D\exp_p(v) \cdot v, D\exp_p(v) \cdot w_T) \\
&= \lambda \cdot g_{\exp_p(v)}(D\exp_p(v) \cdot v, D\exp_p(v) \cdot v) \\
&= \lambda \cdot g_{\exp_p(v)}(\gamma'_{v_p}(1), \gamma'_{v_p}(1)) \\
&= \lambda \cdot g_p(\gamma'_{v_p}(0), \gamma'_{v_p}(0)) \\
&= \lambda \cdot g_p(v, v) \\
&= g_p(v, w_T).
\end{aligned}$$

Let $\Sigma_v = \{x \in T_pM : \|x\| = \|v\|\}$. Choose a smooth curve $\sigma : (-\epsilon, \epsilon) \to \Sigma_v$ satisfying $\sigma(0) = v$ and $\sigma'(0) = w_N$, and consider the parameterised surface

$$f(s, t) = \exp_p(t\sigma(s))$$

For each fixed s_o the curve $f(s_o,t)$ is a geodesic and

$$\frac{\partial}{\partial s}|_{s=0} f(s,t) = D\exp_p(tv)\cdot tw_N \quad \frac{\partial}{\partial t}|_{t=1} f(0,t) = D\exp_p(v)\cdot v$$

$$\frac{\partial}{\partial s}|_{s=0} f(s,0) = 0 \quad \frac{\partial}{\partial t}|_{t=0} f(0,t) = v$$

Therefore

$$g_{\exp_p(v)}(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial s})|_{(0,1)} = g_{\exp_p(v)}(D\exp_p(v)\cdot v, D\exp_p(v)w_N)$$

and

$$g_p(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t})|_{(0,0)} = 0 = g_p(v, w_N).$$

Moreover, $g_{\exp_p(t)}(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial s})$ is independent of t so the above quantities are equal (which proves the lemma):

$$\begin{aligned}\frac{\partial}{\partial t} g_{\exp_p(v)}(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial s}) &= g_{\exp_p(v)}(D_t\frac{\partial f}{\partial t}, \frac{\partial f}{\partial s}) + g_{\exp_p(v)}(\frac{\partial f}{\partial t}, D_t\frac{\partial f}{\partial s})\\ &= g_{\exp_p(v)}(\frac{\partial f}{\partial t}, D_s\frac{\partial f}{\partial t}) = \frac{1}{2}\frac{\partial}{\partial s}\left\|\frac{\partial f}{\partial t}\right\|^2 = 0.\end{aligned}$$

□

Theorem 8.11. *Let $p \in M$, U a normal neighbourhood of p and $B \subset U$ a geodesic ball centered at p. Let $\gamma : [0,1] \to B$ be a geodesic parameterised by arc-length with $\gamma(0) = p$. If $c : [0,1] \to M$ is any piecewise differentiable curve joining $\gamma(0)$ with $\gamma(1)$ then $L_g(\gamma) \le L_g(c)$ and if equality holds then $\gamma = c$.*

Proof. Write $\gamma(t) = \exp_p(tv)$ for some unit vector $v \in T_pM$. Then $L_g(\gamma) = 1$ since γ has unit speed. Suppose first that $c([0,1]) \subset B$ and joins p with $\gamma(1)$ and assume this to be parameterised by arc-length aswell. There exists a curve $\sigma(t) \in T_pM$ with $\sigma(0) = 0$ and $\sigma(1) = v$ such that $c(t) = \exp_e(\sigma(t))$. Introduce polar coordinates and write $\sigma(t) = r(t)\cdot\tilde{\sigma}(t)$ where $r(t) = \|\sigma(t)\|$ and $\|\tilde{\sigma}(t)\| = 1$. Consider the two parameter family of curves

$$f(s,t) = \exp_p(s\cdot\tilde{\sigma}(t))$$

and observe that f(r(t),t)=c(t). So

$$\dot{c}(t) = \frac{\partial f}{\partial t}(r(t),t) + \frac{\partial f}{\partial s}(r(t),t)\cdot\dot{r}(t).$$

Also

$$\frac{\partial f}{\partial s} = D\exp_e(s\cdot\tilde{\sigma}(t))\cdot\tilde{\sigma}(t) \quad \frac{\partial f}{\partial t} = D\exp_e(s\cdot\tilde{\sigma}(t))\cdot s\dot{\tilde{\sigma}}(t)$$

and since $\dot{\tilde{\sigma}}(t)$ is orthogonal to $\tilde{\sigma}(t)$ the Gauss lemma gives

$$g\left(\frac{\partial f}{\partial t}(r(t),t), \frac{\partial f}{\partial s}(r(t),t)\right) = 0.$$

Using the above and observing that $\left\|\frac{\partial f}{\partial s}(r(t),t)\right\|^2 = \|\dot{c}(t)\|^2 = 1$

$$\|\dot{c}(t)\|^2 = \left\|\frac{\partial f}{\partial t}(r(t),t)\right\|^2 + \left\|\frac{\partial f}{\partial s}(r(t),t)\dot{r}(t)\right\|^2 \geq |\dot{r}(t)|^2$$

and integrating we obtain

$$L(c(t)) \geq \int_0^1 |\dot{r}(t)|\, dt \geq r(1) - r(0) = 1 = \int_0^1 \|\dot{\gamma}(t)\|\, dt = L(\gamma)$$

so that γ is minimizing.

If $L(c) = L(\gamma)$ then the above inequalities are equalities. In particular, $\frac{\partial f}{\partial t} = 0$ which can only happen if $\dot{\tilde{\sigma}}(t) = 0$ so that $\tilde{\sigma}(t) = v$ and

$$c(t) = \exp_p(r(t) \cdot v)$$

Also $|\dot{r}(t)| = \dot{r}(t) > 0$ which means the $r(t)$ is an increasing diffeomorphism and therefore c is a forward reparameterisation of γ. □

The Hopf-Rinow Theorem

We introduced the length of a curve between two points. We showed that length minimizing curves are geodesics and geodesics locally minimize length. We introduced the Riemannian distance between two points. We showed that this defines a metric on M so that we can think of a Riemannian manifold as a metric space. A natural question is: "Is M complete as a metric space?". The answer is: "Generally not"

> The open unit ball is a Riemannian manifold in $\mathbb{R}^n$ with the Euclidean metric. Fix a a vector v in the ball and let $\{c_n\}$ be any cauchy sequence in $[0, \frac{1}{\|v\|}]$ converging to $\frac{1}{\|v\|}$. Then $\{x_n\} = \{c_n v\}$ is a Cauchy sequence converging to a vector of length one which lies in the boundary of the ball - i.e. the limit point is NOT contained in the ball. What's the problem?

Earlier we introduced the idea of geodesic completeness. A Riemannian manifold was said to be geodesically complete if all of its geodesics are defined for all time. The problem with the ball example is that geodesics of the Euclidean metric do not exist for all time - they hit the boundary and leave the manifold. The Hopf-Rinow theorem basically says a connected Riemannian manifold is metrically complete if and only if it is geodesically complete.

To prove the Hopf-Rinow Theorem we'll begin with a Lemma which takes us half-way there.

Lemma 8.12. *Let (M,g) be a connected Riemannian manifold and suppose there is a point p such that $\exp_p$ is defined on all of T_pM (geodesics emanating from p exist for all time). Then*

- *Given any other point $q \in M$ there exists a minimizing geodesic segment joining p with q; and*
- *M is metrically complete.*

Proof. Let $q \in M$ and say that a geodesic $\gamma : [a, b] \to M$ "aims at q" if

$$d_g(p, q) = d_g(p, \gamma(b)) + d_g(\gamma(b), q)$$

To prove 1. we're going to construct a geodesic segment which begins at p, aims at q, and has length equal to $d_g(p, q)$ - which implies γ is minimizing. Consequently,

$$d_g(p, q) = d_g(p, q) + d_g(\gamma(b), q)$$

which means $\gamma(b) = q$. Choose an $\epsilon > 0$ so that there exists a closed geodesic ball $B_\epsilon(p)$ around p that doesn't contain q. Since the distance function on a metric space is continuous and the sphere $S_\epsilon(p)$ is compact there exists a point $x \in S_\epsilon(p)$ where $d_g(x, q)$ achieves its minimum. Let γ be the maximal unit speed geodesic whose restriction to $[0, \epsilon]$ is the geodesic segment joining p with x - by assumption γ is defined for all $t \in \mathbb{R}$. We now show that $\gamma|_{[0,\epsilon]}$ "aims at q":

$$d_g(p, q) = d_g(p, x) + d_g(x, q)$$

and we already know that $d_g(p, x) = L_g(\gamma|_{[0,\epsilon]})$. Let $\sigma : [a_o, b_o] \to M$ be any admissible curve joining p with q and let t_o be the first time at which σ hits $S_\epsilon(p)$. Let σ_1 and σ_2 denote the restriction of σ to $[a_o, t_o]$ and $[t_o, b_o]$, respectively. Now

$$L(\sigma_1) \geq d_g(p, \sigma(t_o)) = d_g(p, x)$$

since every point on $S_\epsilon(p)$ is a distance ϵ from p. By our choice of x we have

$$L(\sigma_2) \geq d_g(\sigma(t_o), q) \geq d_g(x, q)$$

Putting these together we get

$$L(\sigma) = L(\sigma_1) + L(\sigma_2) \geq d_g(p, x) + d_g(x, q)$$

Taking the infimum over all σ we get

$$d_g(p, q) \geq d_g(p, x) + d_g(x, q)$$

while the opposite inequality is just the triangle inequality so that

$$d_g(p, q) = d_g(p, x) + d_g(x, q)$$

and γ aims at q.

Let $T = d_g(p, q)$ and set

$$\mathcal{A} = \{b \in [0, T] : \gamma|_{[0,b]} \,\text{aims at,}\, q\}.$$

We've just shown that $\epsilon \in \mathcal{A}$. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(c) = d_g(p, \gamma(c)) + d_g(\gamma(c), q)$$

which is continuous since d_g and γ are both continuous. Then $f^{-1}(d_g(p,q)) = \mathcal{A}$ so that $\mathcal{A}$ is closed by continuity. Therefore $A = \sup \mathcal{A} \in \mathcal{A}$ and is at least as big as ϵ.

If $A = T$ then $\gamma|_{[0,T]}$ is a geodesic of length $T = d_g(p,q)$ which aims at q and we are done. So suppose $A < T$ and let's derive a contradiction. Let $y = \gamma(A)$ and choose a $\delta > 0$ so that $B_\delta(y)$ is a closed geodesic ball of radius δ around y that does not contain q. Since $A \in \mathcal{A}$ we have

$$d_g(y,q) = d_g(p,q) - d_g(y,p) = T - A$$

Let $z \in S_\delta(y)$ be a point where $d_g(z,q)$ attains its minimum and let $\tau : [0,\delta] \to M$ be the unit speed geodesic from y to z. By exactly the same argument τ aims at q so

$$d_g(z,q) = d_g(y,q) - d_g(y,z) = (T - A) - \delta$$

By the triangle inequality

$$d_g(p,z) \geq d_g(p,q) - d_g(z,q) = T - (T - A - \delta) = A + \delta = L(\tau \circ \gamma|_{[0,A+\delta]})$$

Therefore the admissible curve consisting of $\gamma|_{[0,A]}$ followed by τ is a minimizing curve from p to z. This means that it has no corners so z must lie on γ and $z = \gamma(A + \delta)$. Then

$$d_g(p,q) = T = (T - A - \delta) + A + \delta = d_g(z,q) + A + \delta = d_g(z,q) + d_g(z,p)$$

so that $\gamma|_{[0,A+\delta]}$ aims at q and $A + \delta \in \mathcal{A}$ - but this is a contradiction since $A = \sup \mathcal{A}$... This proves 1.

To prove 2. we have to prove that every Cauchy sequence in M converges in M. Let $\{q_i\}$ be a Cauchy sequence in M. For each i let $\gamma_i(t) = \exp_p(tv_i)$ be a unit speed minimizing geodesic from p to q_i and let $d_i = d_g(p, q_i)$ so that $q_i = \exp_e(d_i v_i)$. The sequence $\{d_i\}$ is bounded in $\mathbb{R}$. Since $\{q_i\}$ is Cauchy in M, $\{d_i\}$ is Cauchy in $\mathbb{R}$ and therefore bounded. The sequence $\{v_i\}$ consists of unit vectors in T_pM and therefore the sequence $\{d_i v_i\}$ is bounded in T_pM. Therefore there exists a subsequence $\{d_{i_k} v_{i_k}\}$ converging to some $v \in T_pM$. Since the exponential map is smooth

$$q_{i_k} = \exp_p(d_{i_k} v_{i_k}) \to \exp_p(v) = q \in M$$

and since the original sequence is Cauchy it converges to the same limit. □

We're now ready to prove the full Hopf-Rinow Theorem. The proof makes use of uniformly normal neighbourhoods whose definition and existence was given at the end of Chapter 6.

Theorem 8.13. *A connected Riemannian manifold is metrically complete if and only if it's Geodesically complete.*

Proof. Lemma 8.12 said geodesic completeness implies metric completeness. So let's show that metric completeness implies geodesic completeness. Suppose it's metrically complete but NOT geodesically complete. Then there is some unit speed geodesic $\gamma : [0, b) \to M$ that has no extension past b. Let t_i be any increasing sequence in $[0, b)$ which approaches b and set $q_i = \gamma(t_i)$. Since γ is parameterised by arc length, the length of $\gamma|_{[t_i, t_j]}$ is exactly $|t_i - t_j|$ and

$$d_g(q_i, q_j) \leq |t_i - t_j|$$

so that q_i is Cauchy in M and converges to some point $q \in M$. Let W be a uniformly normal δ-neighbourhood of q. Choose j large enough so that $t_j > b-\delta$ and $q_j \in W$. Since $B_\delta(q_j)$ is a geodesic ball, every unit speed geodesic starting at q_j exists for at least $t \in [0, \delta)$. In particular, this true for the geodesic σ with $\sigma(0) = q_j$ and $\sigma'(0) = \gamma'(t_j)$. Define $\tilde{\gamma} : [0, t_j + \delta) \to M$

$$\tilde{\gamma}(t) = \begin{cases} \gamma(t) & t \in [0, b) \\ \sigma(t - t_j) & t \in (t_j - \delta, t_j + \delta) \end{cases}$$

Each part of $\tilde{\gamma}$ is a geodesic and the two parts agree at t_j - by uniqueness the two parts are the same geodesic and since $t_j + \delta > b$, $\tilde{\gamma}$ is an extension of γ; a contradiction. □

We immediately obtain three useful corollaries

Corollary 8.14. *If M is a connected Riemannian manifold and there exists a point p such that $\exp_p$ is defined on all of T_pM then M is metrically and geodesically complete.*

Proof. Lemma 8.12 implies M is metrically complete and Hopf-Rinow implies it is geodesically complete. □

Corollary 8.15. *If M is a connected Riemannian manifold which is either metrically or geodesically complete then any two points can be joined by a minimizing geodesic*

Proof. Hopf-Rinow implies that there exists a point p such that $\exp_p$ is defined on all of T_pM and part 1. of Lemma 8.12 gives the conclusion. □

Corollary 8.16. *If M is a compact Riemannian manifold then every maximal geodesic in M is defined for all time.*

Proof. Since M is compact it contains all of its limit points as a topological space. Consequently it is metrically complete and therefore geodesically complete by the Hopf-Rinow Theorem. □

Chapter 9

curvature

As is typical in mathematics we are interested in knowing if all the examples of a structure on a smooth manifold are equivalent (or which subset of examples are equivalent). For example, every non-vanishing vector field X can be written as $X = \frac{\partial}{\partial x^1}$ in some suitable coordinate system - so all non-vanishing vector fields are locally equivalent. Every one dimensional Riemannian manifold is locally isometric to $\mathbb{R}$ with its Euclidean metric - so all 1-dimensional Riemannian manifolds are locally equivalent. A symplectic form ω on a smooth manifold M is a smooth bi-linear map whose total derivative is zero and satisfies $\omega(w, v) = 0$ for all $v \in T_pM$ only if $w = 0$. By a Theorem of Darboux all symplectic forms have the same coordinate expressions and are therefore all locally equivalent.

On the other hand, we saw a while back that the round 2-sphere and the Euclidean plane $\mathbb{R}^2$ are not locally isometric - so these two Riemannian manifolds are not locally equivalent. What we're going to do is construct a local invariant of Riemannian metrics - a quantity that is preserved by local equivalences (like isometry) - called curvature. Recall that a Riemannian manifold (M, g) is flat if there exists a local isometry between it and $\mathbb{R}^n$ with the Euclidean dot-product (see Isometries of Chapter 4). The goal of this section is to prove that (M, g) is flat if and only if $R \equiv 0$.

Motivating Heuristics

We know from our experience with calculus in $\mathbb{R}^n$ that partial derivatives commute. Since covariant differentiation in $\mathbb{R}^n$ is just the ordinary directional derivative we know that covariant derivatives commute as-well. As a consequence, parallel transport along curves in Euclidean space commutes and preserves vectors around closed loops. However, this need not be true on more general spaces...

> On the sphere parallel transport around closed loops does not preserve vectors. For example, consider parallel transport on parallelograms defined by coordinate axes.

Recall that covariant differentiation is completely determined by parallel transport: if γ is a smooth curve and if V is a smooth vector field along γ then for each t_o

$$D_t V(t_o) = \lim_{t \to t_o} \frac{P_{t,t_o} V(t) - V(t_o)}{t}.$$

We can use this formula to see that the general failure of parallel transport to commute implies the general failure of covariant differentiation to commute... If $\Gamma(s,t)$ is a smooth one parameter variation of curves and if $V(s,t)$ is a smooth vector field along Γ then

$$\begin{aligned}
&D_s D_t V(s_o,t_o) = \\
&\quad \lim_{s\to s_o}\lim_{t\to t_o} \frac{P_{s,s_o}P_{t,t_o}V(s,t) - P_{s,s_o}V(s,t_o) - P_{t,t_o}V(s_o,t) + V(s_o,t_o)}{(s-s_o)(t-t_o)} \\
&\quad \neq \\
&\quad \lim_{t\to t_o}\lim_{s\to s_o} \frac{P_{t,t_o}P_{s,s_o}V(s,t) - P_{s,s_o}V(s,t_o) - P_{t,t_o}V(s_o,t) + V(s_o,t_o)}{(s-s_o)(t-t_o)} \\
&\quad = D_t D_s V(s_o,t_o).
\end{aligned}$$

Going back to $\mathbb{R}^n$ and looking at the behaviour of covariant differentiation there we first see that

$$\begin{aligned}
\nabla_X \nabla_Y Z &= \left(XY(Z^k)\right)\partial_k \\
\nabla_Y \nabla_X Z &= \left(YX(Z^k)\right)\partial_k
\end{aligned}$$

while on the other hand

$$\nabla_{[X,Y]}Z = \left(XY(Z^k) - YX(Z^k))\right)\partial_k.$$

Putting these together we see that

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \nabla_{[X,Y]}Z$$

Motivated by these considerations and the above calculation we will make a definition and show that it captures and formalizes these heuristics.

The Curvature Tensor

Definition 9.1. *Let (M,g) be a Riemannian manifold and define a map $R : \mathcal{X}(M)\times\mathcal{X}(M)\times\mathcal{X}(M)\to\mathcal{X}(M)$ by*

$$R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$$

Let's just check that this definition captures the heuristics...

Observe that if $M=\mathbb{R}^n$ then $R(X,Y)Z=0$ for all vector fields X, Y, and Z. Indeed, since the Levi-Civita connection is the usual directional derivative

$$\begin{aligned}
\nabla_X \nabla_Y Z &= \left(XY(Z^k)\right)\partial_k \\
\nabla_Y \nabla_X Z &= \left(YX(Z^k)\right)\partial_k \\
\nabla_{[X,Y]}Z &= \left(XY(Z^k - YX(Z^k))\right)\partial_k
\end{aligned}$$

Putting these together gives $R(X,Y)Z=0$. So we can heuristically interpret R as measuring how much M deviates from $\mathbb{R}^n$.

Putting coordinates on M and considering the coordinate vector fields ∂_{x_i} with $[\partial_{x_i}, \partial_{x_j}] = 0$ we have

$$R(\partial_{x_i}, \partial_{x_j})\partial_{x_k} = \left(\nabla_{\partial_{x_i}}\nabla_{\partial_{x_j}} - \nabla_{\partial_{x_j}}\nabla_{\partial_{x_i}}\right)\partial_{x_k}$$

and we see that R heuristically measures the failure of covariant derivatives (along coordinate axes) to commute.

We're going to give some formal properties of R and describe some of its symmetries and then we're going to give proofs of the above heuristics which will lead to a qualitative geometric interpretation of R.

Proposition 9.1. *The curvature R of a Riemannian manifold satisfies the following properties*

a. *R is bilinear:*

$$R(fX_1 + gX_2, Y) = fR(X_1, Y) + gR(X_2, Y)$$
$$R(X, fY_1 + gY_2) = fR(X, Y_Y) + gR(X, Y_2)$$

for all f, $g \in C^\infty(M)$ and any X_1, X_2, $Y_1\, Y_2 \in \mathcal{X}(M)$.

b. *For any X, $Y \in \mathcal{X}(M)$ the curvature operator $R(X,Y) : \mathcal{X}M \to \mathcal{X}M$ is linear:*

$$R(X,Y)(fZ + gW) = fR(X,Y)Z + gR(X,Y)W$$

for all f, $g \in C^\infty(M)$ and any Z, $W \in \mathcal{X}(M)$

Proof. Basically follows from linearity and derivation properties of the connection. It's an algebraic exercise - you do it. □

Recall that a smooth one-parameter family of curves $\Gamma : J\times \to M$ has smooth velocity vectors $\partial_t\Gamma(s,t) = (\Gamma_s)'(t)$ and $\partial_s\Gamma(s,t) = (\Gamma_t)'(s)$ along Γ

Proposition 9.2. *Let (M, g) be a smooth Riemannian manifold and $\Gamma : J \times I \to M$ a smooth one-parameter family of curves in M. For every smooth vector field V along Γ*

$$D_sD_tV - D_tD_sV = R(\partial_s\Gamma, \partial_t\Gamma)V.$$

In particular, R measures the failure of covariant derivatives to commute.

Proof. This is a local question so we can choose smooth coordinates (x^i) on a neighbourhood of $\Gamma(s,t)$ and write

$$\Gamma(s,t) = \left(\gamma^1(s,t), \ldots, \gamma^n(s,t)\right) \qquad V(s,t) = V^j(s,t)\partial_j|_{\Gamma(s,t)}.$$

Applying the formula for the covariant derivative we have

$$D_tV = \frac{\partial V^i}{\partial t}\partial_i + V^iD_t\partial_i.$$

Applying it again in s gives

$$D_s D_t V = \frac{\partial^2 V^i}{\partial_s \partial t}\partial_i + \frac{\partial V^i}{\partial t} D_s \partial_i + \frac{\partial V^i}{\partial s} D_t \partial_i + V^i D_s D_t \partial_i.$$

Doing the opposite we get

$$D_t D_s V = \frac{\partial^2 V^i}{\partial_t \partial s}\partial_i + \frac{\partial V^i}{\partial s} D_t \partial_i + \frac{\partial V^i}{\partial t} D_s \partial_i + V^i D_t D_s \partial_i.$$

and after subtracting the two quantities we see that everything cancels except the last term:

$$D_s D_t V - D_t D_s V = V^i \left(D_s D_t - D_t D_s \right) \partial_i.$$

So we now compute the term in parentheses: Write

$$S = \partial_s \Gamma = \frac{\partial \gamma^k}{\partial_s} \partial_k \qquad T = \partial_t \Gamma = \frac{\partial \gamma^j}{\partial_t} \partial_j.$$

Since ∂_i is extendible we have

$$D_t \partial_i = \nabla_T \partial_i = \frac{\partial \gamma^j}{\partial_t} \nabla_{\partial_j} \partial_i$$

and since $\nabla_{\partial_j} \partial_i$ is also extendible we have

$$\begin{aligned} D_s D_t \partial_i &= D_s \left(\frac{\partial \gamma^j}{\partial_t} \nabla_{\partial_j} \partial_i \right) \\ &= \frac{\partial^2 \gamma^j}{\partial_s \partial_t} \nabla_{\partial_j} \partial_i + \frac{\partial \gamma^j}{\partial_t} \nabla_S \partial_i \\ &= \frac{\partial^2 \gamma^j}{\partial_s \partial_t} \nabla_{\partial_j} \partial_i + \frac{\partial \gamma^j}{\partial_t} \frac{\partial \gamma^k}{\partial_s} \nabla_{\partial_k} \nabla_{\partial_j} \partial_i \end{aligned}$$

Swapping s and t and swapping j and k we see that the first terms cancel and we are left with

$$D_s D_t \partial_i - D_t D_s \partial_i = \frac{\partial \gamma^j}{\partial_t} \frac{\partial \gamma^k}{\partial_s} \left(\nabla_{\partial_k} \nabla_{\partial_j} - \nabla_{\partial_j} \nabla_{\partial_k} \right) \partial_i$$

and using linearity of the curvature tensor and our previous heuristic calculation for coordinate vector fields we obtain

$$D_s D_t \partial_i - D_t D_s \partial_i = \frac{\partial \gamma^j}{\partial_t} \frac{\partial \gamma^k}{\partial_s} R(\nabla_{\partial_k}, \nabla_{\partial_j}) \partial_i = R(S, T) \partial_i.$$

Using linearity once more we see that

$$D_s D_t V - D_t D_s V = V^i R(S, T) \partial_i = R(S, T) V$$

□

Here's one reason for the interest: Curvature, as we have defined it, is a local isometry invariant.

Proposition 9.3. *If (M, g) and (N, h) are Riemannian manifolds and $\phi : M \to N$ is a local isometry then $\phi^* R^h = R^g$.*

Here, $\phi^* R$ means

$$\phi^* R(X, Y)Z = d\phi^{-1} \cdot R(d\phi \cdot X, d\phi \cdot Y) d\phi \cdot Z$$

This Proposition shows that if (M, g) is locally flat, i.e. there exists a local isometry between M with g and $\mathbb{R}^n$ with its Euclidean metric then $R = 0$.

Proof. Recall that

$$\left(\phi^* \nabla^h\right)_X Y = d\phi^{-1} \cdot \nabla^h_{d\phi \cdot X} d\phi \cdot Y$$

We've already shown (see Proposition 6.2 in the notes) that if $\phi : M \to N$ is a local isometry then locally $\nabla^g = \phi^* \nabla^h$. Since ϕ is also a local diffeomorphism the Lie bracket also transforms like $d\phi \cdot [X, Y] = [d\phi \cdot X, d\phi \cdot Y]$. Putting this together, for any X, Y, $Z \in \mathcal{X}(M)$...

$$\begin{aligned}
R^g(X, Y)Z =& \nabla^g_X \nabla^g_Y Z - \nabla^g_Y \nabla^g_X Z - \nabla^g_{[X,Y]} Z \\
=& (\phi^* \nabla^h)_X \phi^* (\nabla^h)_Y Z - (\phi^* \nabla^h)_Y (\phi^* \nabla^h)_X Z - (\phi^* \nabla^h)_{[X,Y]} Z \\
=& d\phi^{-1} \cdot \nabla^h_{d\phi \cdot X} \nabla^h_{d\phi \cdot Y} d\phi \cdot Z - d\phi^{-1} \cdot \nabla^h_{d\phi \cdot Y} \nabla^h_{d\phi \cdot X} d\phi \cdot Z \\
& - d\phi^{-1} \cdot \nabla^h_{d\phi \cdot [X,Y]} d\phi \cdot Z \\
=& d\phi^{-1} \cdot \nabla^h_{d\phi \cdot X} \nabla^h_{d\phi \cdot Y} d\phi \cdot Z - d\phi^{-1} \cdot \nabla^h_{d\phi \cdot Y} \nabla^h_{d\phi \cdot X} d\phi \cdot Z \\
& - d\phi^{-1} \cdot \nabla^h_{[d\phi \cdot X, d\phi \cdot Y]} d\phi \cdot Z \\
=& d\phi^{-1} R^h \left(d\phi \cdot X, d\phi \cdot Y\right) d\phi \cdot Z \\
=& \left(\phi^* R^h\right)(X, Y) Z
\end{aligned}$$

□

To prove that if $R = 0$ then there exists a local isometry with $\mathbb{R}^n$ we need to do a little more work.

Lemma 9.4. *Suppose M is a smooth manifold and ∇ is any connection satisfying*

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \nabla_{[X,Y]} Z.$$

Given $p \in M$ and $v \in T_p M$ there exists a parallel vector field on a neighbourhood of p with $V_p = v$.

Proof. Let $p \in M$ and $v \in T_p M$ and let $(x^1, \ldots, x^n)$ be smooth coordinates centered at p. By shrinking the coordinate neighbourhood if necessary we can assume the image of the coordinate map is the open cube

$$C_\epsilon = \{x_i : |x_i| < \epsilon, \, i = 1, \ldots, n\}.$$

Begin by parallel transporting v along the x^1 axis; then from each point on the x^1-axis, parallel transport along the coordinate line parallel to the x^2-axis; then successively parallel transport along lines parallel to the x^3 through x^n-axes. The result is a smooth vector field V in C_ϵ.

Since $\nabla_X V$ is linear over C^∞ in X to show that V is parallel it suffices to show that $\nabla_{\partial_i} V = 0$ for each i. By construction $\nabla_{\partial_1} V = 0$ on the x^1-plane; $\nabla_{\partial_2} V = 0$ on the (x^1, x^2)-plane (but $\nabla_{\partial_1} V$ is not necessarily zero here); and in general $\nabla_{\partial_k} V = 0$ on the slice $M_k \subseteq C_\epsilon$ defined by $x^{k+1} = \cdots = x^n = 0$ (but $\nabla_{\partial_i} V$, $i < k$, is not necessarily zero here). BUT we will prove by induction that actually

$$\nabla_{\partial_1} V = \nabla_{\partial_2} V = \cdots = \nabla_{\partial_k} V = 0$$

on M_k under the assumption

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \nabla_{[X,Y]} Z.$$

For $k = 1$ it's true by construction and for $k = n$ it means that V is parallel on all of C_ϵ

So assume

$$\nabla_{\partial_1} V = \nabla_{\partial_2} V = \cdots = \nabla_{\partial_k} V = 0$$

for some k.

By construction $\nabla_{\partial_{k+1}} V = 0$ on all of M_{k+1} and for $i \leq k$ the inductive hypothesis guarantees that $\nabla_{\partial_i} V = 0$ on all of $M_k \subseteq M_{k+1}$.

Since $[\partial_{k+1}, \partial_i] = 0$ the assumption gives

$$\nabla_{\partial_{k+1}} (\nabla_{\partial_i} V) = \nabla_{\partial_i} \left(\nabla_{\partial_{k+1}} V \right) = 0$$

on M_{k+1}. This shows that $\nabla_{\partial_i} V$ is parallel along x^{k+1}-curves starting on M_k. Since $\nabla_{\partial_i} V = 0$ on M_k and 0 is the unique parallel transport of 0 we conclude that $\nabla_{\partial_i} V$ is zero on each x^{k+1}-curve. Since every point of M_{k+1} is on one of these curves it follows that $\nabla_{\partial_i} V = 0$ on M_{k+1}. This completes the inductive step and therefore V is parallel on C_ϵ. $\square$

Theorem 9.5. *A Riemannian manifold is flat if and only if its curvature tensor is identically zero.*

Remark 9.1. *The proof of the Theorem relies on a fact about vector fields which we will state without proof. Recall that a local frame on M is a set of vector fields E_i on an open subset U such that $E_i|_p$ form a basis for T_pM. A smooth local frame is said to commute if $[E_i, E_j] = 0$ for all i and j:*

Theorem 9.6. *Suppose E_i is a commuting local frame on an open subset U. Then, for each $p \in U$ there exists a smooth coordinate chart $\left(W, y^i\right)$ such that $E_i = \frac{\partial}{\partial y^i}$.*

Proof. Recall that a Riemannian manifold (M, g) is flat if there exists a local isometry from M with g to $\mathbb{R}^n$ with the Euclidean metric. The Proposition above show that if (M, g) is flat then its curvature tensor is identically zero. Now suppose that the curvature tensor of (M, g) is zero. Then

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \nabla_{[X,Y]} Z.$$

We'll start by showing that g admits a parallel orthonormal frame in a neighbourhood of each point. Choose a point $p \in M$ and let $\{e_1, \dots, e_n\}$ be an orthonormal basis for T_pM. By the Previous Lemma there exist parallel vector fields $E_1, \dots, E_n$ in a neighbourhood of p obtained by parallel transporting the $e_i's$. Since parallel transport is an isometry of g the fields $E_1, \dots, E_n$ are orthonormal at every point in the neighbourhood on which they are defined. Since the Levi-Civita connection is symmetric and the E_i's are parallel

$$[E_i, E_j] = \nabla_{E_i} E_j - \nabla_{E_j} E_i = 0$$

Therefore $\{E_1, \dots, E_n\}$ are a commuting orthonormal frame of vector fields. The canonical form theorem for commuting vector fields says that there exists coordinates y^i such that $E_i = \frac{\partial}{\partial y^i}$. Now in these coordinates the metric takes the form

$$g_{ij} = g(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}) = g(E_i, E_j) = \delta_{ij}$$

so that g is the Euclidean metric on a neighbourhood of p. This means the map $y = (y^1, \dots, y^n)$ is an isometry from a neighbourhood of p to an open subset of $\mathbb{R}^n$. $\square$