

MATH 765: DIFFERENTIAL GEOMETRY (UW-MADISON, SPRING 2026)

ALEX WALDRON

CONTENTS

Part 1. Introduction to Riemannian metrics	4
1. Riemannian metrics (Tu 1/20)	4
2. Basic constructions with metrics (Th 1/22)	8
Part 2. Connections	15
3. Covariant derivatives (Tu 1/27)	15
4. Local description of a connection (Wed 1/28 online)	20
5. Parallel transport (Tu 2/3)	24
6. Gauge-equivalence (Th 2/5)	27
Part 3. Curvature	30
7. The curvature operator (Tu 2/10)	30
8. Geometric interpretation of curvature (Th 2/12)	33
9. Flat connections (Th 2/12 - Tu 2/17)	37
10. Complex line bundles and the first Chern class (Tu 2/17)	40
11. Induced connections (Th 2/19)	43
12. The Differential Bianchi Identity (Th 2/19)	45
Part 4. The Levi-Civita connection	47
13. Metric compatibility (Tu 2/24)	47
14. Torsion-freeness (Tu 2/24-Th 2/26)	52
15. The Fundamental Lemma (Th 2/26)	53
16. Calculating Christoffel symbols (Tu 3/3)	55
Part 5. Geodesics	59
17. The geodesic equation (Tu 3/3)	59
18. Examples (Tu 3/3-Th 3/5)	60
19. The exponential map (Th 3/5)	64
20. Uniformly normal neighborhoods (Tu 3/10)	68
21. First variation of arclength (Tu 3/10 - Th 3/12)	69
22. Gauss's Lemma and minimization (Th 3/12)	73
23. The Hopf-Rinow Theorem (Tu 3/17)	77

Part 6. The Riemann curvature tensor	81
24. Definition and examples (Th 3/19)	81
25. Flat metrics (Th 3/19)	83
26. The Bianchi identities (Th 3/19-Tu 3/24)	84
27. Sectional curvature (Tu 3/24)	88
28. Ricci and Scalar curvature (Tu 3/24)	90
Part 7. Submanifolds and Lie groups	91
29. The Gauss equation (Th 3/26)	91
30. Geometric interpretation of sectional curvature (Tu 4/7)	95
31. Hypersurfaces and the Theorema Egregium (Tu 4/7)	96
32. Classical notation for parametrized surfaces (Tu 4/7)	98
33. Discussion of embedded versus abstract metrics on surfaces (Th 4/9)	100
34. Invariant metrics on Lie groups (Th 4/9)	101
Part 8. Curvature and topology	107
35. The Gauss-Bonnet Theorem (Tu 4/14)	107
36. Jacobi fields (Th 4/16)	110
37. First properties of Jacobi fields (Th 4/16-Tu 4/21)	112
38. Local uniqueness of constant sectional curvature (Tu 4/21)	115
39. Conjugate points (Tu 4/21)	119
40. The Cartan-Hadamard Theorem (Tu 4/21-Th 4/23)	120
41. Comparison Theorems (Tu 4/28)	129
42. Global uniqueness of constant sectional curvature (Tu 4/28)	131
43. Second variation of arclength (Th 4/30 online)	133
44. Geodesics do not minimize past conjugate points (Th 4/30 online)	136
45. The Bonnet-Myers Theorem (Th 4/30 online)	137

References:

- Lee, *(Introduction to) Riemannian manifolds*, GTM 176, 1st/2nd ed.
- Do Carmo, *Riemannian geometry*
- Tu, *Differential geometry*, GTM 275.
- Spivak, *A comprehensive introduction to differential geometry*, Vol. 2
- Kobayashi and Nomizu, *Foundations of differential geometry*

Acknowledgement: The author gratefully thanks Keyang Li, Hampe Meng, Dijia (Mark) Chen, John Cooper, Alexandros Kaiopoulos, Inbo Gottlieb Fenves, and Weijian Chen for providing first drafts of these notes after lecture.

Part 1. Introduction to Riemannian metrics

Motivation. *Topology* studies the global shape of a space. *Differential geometry* studies the local shape of a space. We want to learn what this means and see how it relates back to topology.

1. RIEMANNIAN METRICS (TU 1/20)

1.1. **The definition.** Let M be a smooth manifold. I will assume that you are very familiar with this concept from Math 761.

A **Riemannian metric**, g , is a smoothly varying real inner product on the tangent spaces of M . Formally, $g \in T^{(0,2)}M$ is a $(0,2)$ -tensor, i.e. a smooth section of $T^*M \otimes T^*M$, whose pointwise values $g_x(\cdot, \cdot)$ give an inner product on T_xM for each $x \in M$, i.e. are symmetric and positive-definite. We will sometimes use the notations

$$g(\cdot, \cdot) = \langle \cdot, \cdot \rangle_g = \langle \cdot, \cdot \rangle.$$

A smooth manifold endowed with a Riemannian metric is called a **Riemannian manifold**, (M, g) .

Given a Riemannian metric g on M , we can speak about the **length**

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

of a tangent vector $v \in T_xM$, and about the **angle** between two nonzero tangent vectors $v, w \in T_xM$:

$$\theta = \arccos g_x \left(\frac{v}{|v|}, \frac{w}{|w|} \right).$$

We can also speak about *volumes*, see next lecture.

1.2. **Local expression for a metric.** We will frequently use the **local coordinate representation** of a metric, which is just its components as a tensor. Given a local coordinate system $\{x^i\}$ on M , we have the coordinate frame $\left\{ \frac{\partial}{\partial x^i} \right\}_{i=1}^n$ for TM . (We shall also sometimes work with frames for TM that do not necessarily come from a coordinate system, for instance, an orthonormal frame). Let

$$g_{ij} = g \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)$$

be the local components; these are just n^2 smooth functions on the coordinate patch. For two vector fields $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^j \frac{\partial}{\partial x^j}$ (where we (ab)use the Einstein summation convention), the inner-product is given by

$$\begin{aligned} g(X, Y) &= g \left(X^i \frac{\partial}{\partial x^i}, Y^j \frac{\partial}{\partial x^j} \right) \\ &= X^i Y^j g \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \\ &= X^i Y^j g_{ij}. \end{aligned} \tag{1.1}$$

Note that it is sometimes most convenient to use the following type of local expression for g :

$$g = g_{ij} dx^i \otimes dx^j.$$

Sometimes one even omits the $\otimes$.

1.3. More definitions. Let's introduce a few more simple definitions before we start discussing examples.

1. Given two Riemannian manifolds (M, g) and (N, h) , a (local) diffeomorphism $f : M \rightarrow N$ is called a **(local) isometry** if

$$h(f_*X, f_*Y) = g(X, Y)$$

for all $X, Y \in \mathfrak{X}(M)$. We write $\text{Isom}(M, g)$ for the group of self-isometries of (M, g) .

2. Given an (immersed) submanifold $N \subset M$, the **induced metric** is simply the restriction of g to $T_x N \subset T_x M$ for each $x \in N$. If (N, h) has a metric h and the inclusion $\iota : N \rightarrow M$ is an isometry between h and the induced metric of g , we say that ι is an **isometric embedding**.

3. Given two Riemannian manifolds (M_1, g_1) and (M_2, g_2) , the **product metric** $g = (g_1, g_2)$ on $M_1 \times M_2$ is given by

$$g((v_1, v_2), (w_1, w_2)) = g_1(v_1, w_1) + g_2(v_2, w_2).$$

4. Suppose that a Lie group G acts smoothly, freely, and properly *by isometries* on a Riemannian manifold (M, g) , and let $N = M/G$ be the quotient manifold. (Usually this will be a known manifold, but it's also guaranteed to exist by the Quotient Manifold Theorem.) Given a vector field X on N , we can take its **horizontal lift** $\tilde{X}$, which is the unique vector field on M with $\pi_* \tilde{X} = X$ and $\tilde{X}_x \perp T_x(G \cdot x)$ for all $x \in M$. The **quotient metric** $\bar{g}$ on N is defined by

$$\bar{g}(X, Y) = g(\tilde{X}, \tilde{Y}).$$

1.4. Examples.

1.4.1. The three model spaces.

1. $(\mathbb{R}^n, \delta_{ij})$ is called **Euclidean space**. Its isometry group contains a copy of $\mathbb{R}^n$ (translations) as well as $O(n)$ (rotations), and has the structure of a semidirect product $\mathbb{R}^n \rtimes O(n)$.
2. The induced metric on the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ is called the **round metric**. The corresponding Riemannian manifold is denoted $\mathbb{S}^n$ and called the **round (unit) sphere**.

We have $\text{Isom}(\mathbb{S}^n) = O(n+1)$. To see this, note that the group acts transitively and the stabilizer of a point is $O(n)$, which is the full group of linear isometries of

an n -dimensional vector space. You cannot possibly have more isometries than this (formal justification later).

There is a famous expression for the components of the round metric in a stereographic chart:

$$g_{ij} = \frac{4\delta_{ij}}{(1 + |x|^2)^2}.$$

See homework.

3. **Hyperbolic space** $\mathbb{H}^n$ is the upper-half-space $\{(x^1, \dots, x^n) \mid x^n > 0\}$ endowed with the metric

$$g_{ij} = \frac{\delta_{ij}}{(x^n)^2}.$$

The length of a tangent vector is inversely proportional to its distance from $\mathbb{R}^{n-1}$.

Translation by $\mathbb{R}^{n-1}$ gives isometries of $\mathbb{H}^n$, while scaling by $\lambda > 0$ is also an isometry. The isometry group is therefore transitive, i.e. the space is **homogeneous**. In fact, inversion in the unit circle is also an isometry (homework exercise). With a bit more work one can see that the stabilizer of a point is again $O(n)$, so hyperbolic space also has the maximal possible group of isometries. It turns out that $\mathbb{R}^n$, $\mathbb{S}^n$, and $\mathbb{H}^n$ are the only spaces with this property, as we shall prove later in the course.

1.4.2. More examples.

1. Any submanifold of $\mathbb{R}^n$ gets an induced metric. For instance, haphazardly drawing a surface of genus g in $\mathbb{R}^3$ endows it with a Riemannian metric.
2. The torus $\mathbb{T}^n = \mathbb{S}^1 \times \dots \times \mathbb{S}^1$ can be given the product metric. This coincides with the quotient metric $\mathbb{R}^n / 2\pi\mathbb{Z}^n$.
3. We can give $\mathbb{RP}^n = \mathbb{S}^n / \pm 1$ the quotient metric. Since matrix multiplication commutes with scalar multiplication, all the isometries descend, so its isometry group is $O(n+1) / \pm 1$.
4. More interestingly, by writing $\mathbb{CP}^n = \mathbb{S}^{2n+1} / U(1)$, we get the **Fubini-Study metric**. Its isometry group contains $U(n+1)$, but $U(1)$ acts trivially, so this goes over to $U(n+1) / U(1)$. The conjugation map $\mathbb{C}^{n+1} \xrightarrow{z \rightarrow \bar{z}} \mathbb{C}^{n+1}$ also preserves the fibers, so descends to an isometry. These in fact generate all the isometries of the Fubini-Study metric.
5. Given any smooth projective variety $X \subset \mathbb{CP}^n$, we can give it the induced metric from the Fubini-Study metric.

1.5. **The Riemannian distance function.** What does a Riemannian metric buy you? Quite a number of things (more next lecture), the most fundamental of which is suggested by the double-entendre of (Riemannian) “metric.”

Definition 1.1. Given a piecewise smooth path $\gamma : [a, b] \rightarrow M$, the **length** of γ is given by

$$L(\gamma) = \int_a^b |\gamma'(t)|_g dt.$$

By the chain rule / change-of-variables, this is independent of the parametrization.

Given two points $x, y \in M$, the **Riemannian distance** between x and y is given by

$$d_g(x, y) = \inf_{\substack{\gamma \text{ connecting} \\ x \text{ and } y}} L(\gamma).$$

(If x and y are in different path-components, we take the value to be ∞ .)

Theorem 1.2. d_g is a metric, and the metric topology on M induced by d_g coincides with the topology of M .

You can prove this theorem now as an exercise, or else we will deduce it later from more precise results. Somewhat deeper is the following:

Theorem 1.3 (Hopf-Rinow). *Suppose M is connected. If (M, d_g) is complete (as a metric space) then any two points x and y are connected by a length-minimizing smooth path.*

A path that is locally length-minimizing is called a **geodesic**. Hopf-Rinow is fairly easy to prove, as we shall do later, given some basic lemmas about geodesics.

1.5.1. Examples.

- Geodesics in $\mathbb{R}^n$ are straight lines, and the metric $d_{\delta_{ij}}$ coincides with the Euclidean metric (exercise).
- Geodesics on $\mathbb{S}^n$ follow great circles. Notice that a geodesic fails to be minimizing after it passes the distance π . You don't want to fly to China through Europe.
- It is an exercise to prove that a vertical line segment in $\mathbb{H}^n$ is length-minimizing. After applying inversion in a circle, these go over to segments of circles meeting the x -axis at right angles. Notice that given a point p and a geodesic line ℓ not meeting p , there are infinitely many geodesic lines through p that do not meet ℓ . This is different from the previous two examples.

1.6. Curvature. As you can see from the examples in Section 1.5.1, the three model spaces $\mathbb{R}^n$, $\mathbb{S}^n$, and $\mathbb{H}^n$ have very different global properties. Is it possible to tell them apart without comparing their full isometry groups or examining the set of all geodesics, etc.? In particular, can one tell them apart *locally*?

The answer is yes, and requires developing the notion of **curvature**. This is the main goal of our course.

$\dim(M) = 1$. Suppose first that $M = C$ is a curve (1-manifold) embedded in $\mathbb{R}^2$, i.e. the image of map $\gamma : [a, b] \rightarrow \mathbb{R}^2$. We can take the parametrization to have unit speed, $|\gamma'(t)| \equiv 1$. Also choose a unit normal vector-field $\nu(t)$ along $\gamma(t)$. In Calculus III we teach that the *curvature* of C is given by

$$\kappa_\gamma(t) = \gamma''(t) \cdot \nu(t).$$

This indeed corresponds to the curve “curving” around a circle of radius $1/\kappa$. However, notice that the map γ is an isometry. Hence, C is isometric to $[a, b] \subset \mathbb{R}$. In particular, C has no *intrinsic curvature* but only some extrinsic curvature related to the embedding in $\mathbb{R}^2$.

$\dim(M) = 2$. Consider an embedded surface $\Sigma \subset \mathbb{R}^3$ with the induced metric, and choose a normal vector field $\nu(x)$ along Σ . Given $x \in M$ and a plane P containing both x and $\nu(x)$, the intersection

$$\Sigma \cap P =: \gamma_P$$

is a curve contained in $P \cong \mathbb{R}^2$. Euler defined the **principal curvatures**

$$\kappa_1(x) := \sup_P \kappa_{\gamma_P}(x), \quad \kappa_2(x) := \inf_P \kappa_{\gamma_P}(x),$$

where the sup and inf range over all planes P containing x and $\nu(x)$. From these we can define the **mean curvature**

$$H_x = \frac{\kappa_1(x) + \kappa_2(x)}{2}$$

and the **Gauss curvature**

$$K_x = \kappa_1(x) \cdot \kappa_2(x).$$

Example 1.4. • Let $\Sigma_1 = [-\frac{\pi}{2}, \frac{\pi}{2}] \times \mathbb{R} \times \{0\}$. This has all curvatures equal to zero.

- Let $\Sigma_2 = \{x^2 + z^2 = 1, z > 0\}$. This has mean curvature $H \equiv -\frac{1}{2}$ with respect to the upward unit normal, but Gauss curvature $K \equiv 0$.

However, notice that Σ_1 and Σ_2 are isometric. The mean curvature H is therefore only extrinsic, not intrinsic. Meanwhile:

Theorem 1.5 (Gauss’s Theorema Egregium, 1827). *The Gauss curvature K_x is intrinsic, i.e., is the same for any two isometric embeddings of a Riemannian surface.*

For another instance of the theorem, one can imagine taking a contact lens and squishing it to be long and skinny. The principal curvature in the longer direction goes down while the principal curvature in the shorter direction goes up, and it turns out that these effects exactly compensate one another.

While Gauss was able to prove this theorem by direct calculation, we will instead develop the modern theory of curvature without reference to embeddings (or even metrics, for that matter) and at some point let the Theorema Egregium fall out. The intrinsic theory of curvature goes back to Riemann himself, who introduced it along with the idea of a Riemannian metric in 1854. (See Spivak Vol. II for a translation of Riemann’s essay *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, with commentary.)

2. BASIC CONSTRUCTIONS WITH METRICS (TH 1/22)

Today we will continue discussing the question “what does a metric buy you?” We will cover a few more of the most immediate things.

2.1. Linear algebra with a metric.

2.1.1. *Musical isomorphisms.* Let V be a finite-dimensional vector space. Recall that V and V^* are not canonically isomorphic. However, after fixing an inner product (metric) $g = \langle \cdot, \cdot \rangle$, we get the obvious isomorphism

$$\begin{aligned} (\cdot)^\flat : V &\longrightarrow V^* \\ X &\longmapsto X^\flat := g(-, X). \end{aligned}$$

Let's get an explicit expression for this map.

Fix a basis $\{e_1, \dots, e_n\}$ of V , and let $\{e^1, \dots, e^n\}$ be the dual basis, satisfying

$$e^j(e_i) = \delta^j_i.$$

For $X \in V$, we will write

$$X = X^i e_i,$$

and for $\alpha \in V^*$, we will write

$$\alpha = \alpha_j e^j.$$

The components X^i or α_j can be picked out by evaluation on the basis/dual basis elements:

$$X^i = e^i(X)$$

and

$$\alpha_j = \alpha(e_j).$$

Also define

$$g_{ij} := \langle e_i, e_j \rangle.$$

Claim. $\boxed{(X^\flat)_j = X^i g_{ij}}$

Proof of claim. $(X^\flat)(Y) = \langle X, Y \rangle = (X^i g_{ij}) Y^j = (X^\flat)_j Y^j.$ ◇

Note: $(\cdot)^\flat$ is sometimes called the “lowering map.”

Denote the inverse of $(\cdot)^\flat$ by

$$(\cdot)^\sharp : V^* \longrightarrow V.$$

Let's also get a coordinate expression for this map. Let $\{g^{ij}\}$ be the inverse matrix of $\{g_{ij}\}$, defined by

$$g^{ik} g_{kj} = \delta^i_j.$$

Claim. $\boxed{(\alpha^\sharp)^i = g^{ij} \alpha_j}$

Proof of claim. We calculate:

$$g^{ij} (X^\flat)_j = g^{ij} g_{jk} X^k = \delta^i_k X^k = X^i.$$

The RHS of the boxed equation is therefore the inverse of $(\cdot)^\flat$, so coincides with $(\cdot)^\sharp$. ◇

Note: $(\cdot)^\sharp$ is sometimes called the “raising map.” Together, $(\cdot)^\flat$ and $(\cdot)^\sharp$ are called the *musical isomorphisms*.

2.1.2. *Induced metric on V^* .* The *induced metric* g^* on V^* may be defined by

$$g^*(\alpha, \beta) = g(\alpha^\sharp, \beta^\sharp).$$

We will later omit the $*$ and also refer to this metric as g .

Claim. $\boxed{g^*(\alpha, \beta) = g^{ij} \alpha_i \beta_j}$

Proof of claim. $g^*(\alpha, \beta) = g_{ij}(g^{jk} \alpha_k)(g^{i\ell} \beta_\ell) = \delta_i^k \alpha_k g^{i\ell} \beta_\ell = \alpha_i g^{i\ell} \beta_\ell.$ $\diamond$

Note: We have

$$g \in \text{Sym}^2 V^* \subset V^* \otimes V^*, \quad g^* \in \text{Sym}^2 V^{**} = \text{Sym}^2 V \subset V \otimes V.$$

Exercise 2.1. Show that $\langle e^k, e^\ell \rangle_{g^*} = g^{k\ell}$.

2.1.3. *Induced metrics on tensor products.* Let (V, g) and (W, h) be inner product spaces. We define a metric on the tensor product $V \otimes W$, denoted by $g \otimes h$.

First define a map

$$(g, h) : V \times W \times V \times W \rightarrow \mathbb{R}$$

by

$$(g, h)((v_1, w_1), (v_2, w_2)) := g(v_1, v_2) h(w_1, w_2).$$

Notice that this map is linear in all four entries. By the universal property of tensor products, it descends to a map

$$g \otimes h : V \otimes W \otimes V \otimes W \rightarrow \mathbb{R}.$$

This is the induced metric on $V \otimes W$. Explicitly, for elements $\sum_i v_i \otimes w_i$ and $\sum_j v'_j \otimes w'_j$ in $V \otimes W$, we have

$$(g \otimes h) \left(\sum_i v_i \otimes w_i, \sum_j v'_j \otimes w'_j \right) = \sum_{i,j} g(v_i, v'_j) h(w_i, w'_j).$$

Recall that the space of (k, ℓ) -tensors on V is given by

$$T^{(k, \ell)} V = \overbrace{V \otimes \cdots \otimes V}^k \otimes \overbrace{V^* \otimes \cdots \otimes V^*}^\ell.$$

Using g and g^* , we get induced metrics on $T^{(k, \ell)}$ for each pair (k, ℓ) , given explicitly as follows.

Recall that the components of a tensor $S \in T^{(k, \ell)}$ are determined by

$$S = S^{i_1 \cdots i_k}_{j_1 \cdots j_\ell} e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes e^{j_1} \otimes \cdots \otimes e^{j_\ell}.$$

Given another tensor $T \in T^{(k, \ell)}$, the induced inner product is given by

$$\boxed{\langle S, T \rangle_g = S^{i_1 \cdots i_k}_{j_1 \cdots j_\ell} T^{i'_1 \cdots i'_k}_{j'_1 \cdots j'_\ell} g_{i_1 i'_1} \cdots g_{i_k i'_k} g^{j_1 j'_1} \cdots g^{j_\ell j'_\ell}.$$

Exercise 2.2. Show that $|g|_g = \sqrt{n}$, where $n = \dim(V)$.

2.1.4. *Contractions.* Let $1 \leq a \leq k$ and $1 \leq b \leq \ell$. Given $T \in T^{(k,\ell)}V$, we can define the contraction

$$T' \in T^{(k-1,\ell-1)}V$$

by

$$(T')^{i_1 \dots \hat{i}_a \dots i_k}_{j_1 \dots \hat{j}_b \dots j_\ell} = T^{i_1 \dots i_{a-1} i_{a+1} \dots i_k}_{j_1 \dots j_{b-1} j_{b+1} \dots j_\ell}.$$

This does *not* require a metric; it is just induced by the canonical evaluation map $V \otimes V^* \rightarrow \mathbb{R}$. For $k = \ell = 1$, it is the ordinary trace map.

With a metric g , we can also contract two upper indices:

$$(T')^{i_1 \dots \hat{i}_a \dots \hat{i}_b \dots i_k}_{j_1 \dots j_\ell} = g_{ij} T^{i_1 \dots i_{a-1} i_{a+1} \dots i_{b-1} j i_{b+1} \dots i_k}_{j_1 \dots j_\ell}$$

or two lower indices.

2.1.5. *Volume Form.* Recall that the top exterior power $\Lambda^n V^*$ is 1-dimensional. An orientation on V is a choice of which connected component of $\Lambda^n V^* \setminus \{0\}$ is considered positive. A choice of metric also fixes a canonical element of $\Lambda^n V^*$:

Definition/Lemma 2.3.¹ Suppose that V is oriented and equipped with a metric g .

(a) The associated volume form is given by

$$\text{vol}_g = E^1 \wedge \dots \wedge E^n,$$

where $\{E^1, \dots, E^n\}$ is any oriented orthonormal basis of V^* .

(b) For a general oriented basis $\{e^1, \dots, e^n\}$ of V^* , we have

$$\text{vol}_g = \sqrt{\det(g_{ij})} e^1 \wedge \dots \wedge e^n,$$

where

$$g_{ij} = \langle e_i, e_j \rangle.$$

Proof. Note that since $\det(\delta_{ij}) = 1$, (b) implies that (a) is well-defined.

To prove (b), let $\{E^1, \dots, E^n\}$ be any fixed oriented orthonormal basis for V^* . Let A^i_j be the change-of-basis matrix determined by

$$E^i = A^i_j e^j,$$

which has $\det(A^i_j) > 0$ since both bases are positively oriented. Then

$$\begin{aligned} E^1 \wedge \dots \wedge E^n &= (A^1_{j_1} e^{j_1}) \wedge \dots \wedge (A^n_{j_n} e^{j_n}) \\ &= \sum \text{sgn}(j_1 \dots j_n) A^1_{j_1} \dots A^n_{j_n} e^1 \wedge \dots \wedge e^n \\ &= (\det A^i_j) e^1 \wedge \dots \wedge e^n. \end{aligned} \tag{2.1}$$

To finish the proof, we need to show:

Claim. $\det(A^i_j) = \sqrt{\det(g_{ij})}$.

Proof of claim. We calculate

$$\langle E^i, E^j \rangle = \delta^{ij} = \langle A^i_k e^k, A^j_\ell e^\ell \rangle = A^i_k g^{k\ell} A^j_\ell.$$

¹A “Definition/Lemma” is a definition where the fact that the definition is well-defined amounts to a Lemma.

Taking determinants of both sides, we have

$$1 = \det(A^i_j)^2 \det(g^{k\ell}).$$

Since $(g^{k\ell}) = (g_{ij})^{-1}$, we have $\det(g^{k\ell}) = (\det(g_{ij}))^{-1}$. Rearranging and taking square roots gives the claim. $\square$

2.2. Back to Riemannian manifolds. *All of the previous discussion extends directly to a Riemannian metric on a smooth manifold.*

For example, the lowering map

$$(\cdot)^{\flat} : \mathfrak{X}(M) \longrightarrow \Omega^1(M)$$

and raising map

$$(\cdot)^{\sharp} : \Omega^1(M) \longrightarrow \mathfrak{X}(M)$$

are given by the same formulae as above, and one just checks in local frames that they preserve smoothness. We also have induced metrics on all tensor bundles, etc.

Definition 2.4. Suppose that M is an oriented smooth manifold with a Riemannian metric g . The *Riemannian volume form* is given locally by

$$dV_g \stackrel{loc}{:=} E^1 \wedge \cdots \wedge E^n,$$

where $E^1 \wedge \cdots \wedge E^n$ is any oriented local coframe. Notice that the expression is independent of the coframe, by Definition/Lemma 2.3a; it is therefore globally well-defined.

If we use a coordinate coframe $dx^1, \dots, dx^n$ instead of an orthonormal coframe, by Definition/Lemma 2.3b, we have the expression

$$\boxed{dV_g \stackrel{loc}{=} \sqrt{\det(g_{ij})} dx^1 \wedge \cdots \wedge dx^n.} \quad (2.2)$$

Remark 2.5. In this business it is sometimes preferable to work with coordinate frames and, at other times, with an orthonormal frames. These two are mutually exclusive unless g is flat, i.e., locally isometric to Euclidean space. Meanwhile, it is worth noting that given any local frame $\{e_1, \dots, e_n\}$, the Gram–Schmidt process sends it uniquely to an orthonormal frame $\{E_1, \dots, E_n\}$. One simply observes that the formulae in the algorithm,

$$E_1 = \frac{e_1}{|e_1|_g}, \quad E_2 = \frac{e_2 - \langle e_2, E_1 \rangle_g E_1}{|e_2 - \langle e_2, E_1 \rangle_g E_1|_g}, \quad \text{etc.}$$

all preserve smoothness.

2.3. Notions from multivariable calculus. In Math 761, you generalized many things from Calculus III to smooth manifolds. Here are a few more that make sense in the presence of a Riemannian metric.

Definition 2.6. Let $f \in C^\infty(M)$. The **gradient** of f is given by

$$\nabla f := (df)^{\sharp} \in \mathfrak{X}(M).$$

Notice that by definition of the $(\cdot)^{\sharp}$ map, we have

$$X(f) = df(X) = \langle \nabla f, X \rangle_g.$$

Definition 2.7. The **divergence** of $X \in \mathfrak{X}(M)$, $\operatorname{div} X \in C^\infty(M)$, is defined by the prescription

$$(\operatorname{div} X) dV_g := d(\iota_X dV_g).$$

Claim. Writing $X = X^i \frac{\partial}{\partial x^i}$, we have
$$\operatorname{div} X = \frac{1}{\sqrt{\det g}} \frac{\partial (\sqrt{\det g} X^i)}{\partial x^i}.$$

Proof of claim. We have

$$\iota_X dV_g = \sum_{i=1}^n (-1)^{i-1} \sqrt{\det g} X^i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n.$$

Taking the exterior derivative gives

$$\begin{aligned} d(\iota_X dV_g) &= \sum_{i,j} (-1)^{i-1} \frac{\partial (\sqrt{\det g} X^i)}{\partial x^j} dx^j \wedge dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \sum_i ((-1)^{i-1})^2 \frac{\partial (\sqrt{\det g} X^i)}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n \\ &= \frac{1}{\sqrt{\det g}} \frac{\partial (\sqrt{\det g} X^i)}{\partial x^i} dV_g. \end{aligned}$$

◇

Theorem 2.8 (Divergence Theorem). *Let (M, g) be a compact oriented Riemannian manifold with boundary $(\partial M, \hat{g})$, where $\hat{g}$ is the metric induced by g on ∂M . Let N be the outward-pointing orthogonal unit normal vector field along ∂M . For any vector field $X \in \mathfrak{X}(M)$, we have*

$$\int_M \operatorname{div} X dV_g = \int_{\partial M} \langle X, N \rangle_g dV_{\hat{g}}.$$

Proof. Applying Stokes's Theorem, we have

$$\int_M \operatorname{div} X dV_g = \int_M d(\iota_X dV_g) = \int_{\partial M} \iota_X dV_g.$$

To finish the proof, it suffices to show:

Claim. $\iota_X dV_g|_{\partial M} = \langle X, N \rangle_g dV_{\hat{g}}$

Proof of claim. Complete N to an oriented orthonormal local frame for $TM|_{\partial M}$,

$$\{N, E_2, \dots, E_n\},$$

where $E_2, \dots, E_n$ are tangent to ∂M . Write the dual frame as $N^\flat, E^2, \dots, E^n$. We have

$$dV_g = N^\flat \wedge E^2 \wedge \cdots \wedge E^n, \quad dV_{\hat{g}} = E^2 \wedge \cdots \wedge E^n.$$

We get

$$\iota_X dV_g = N^\flat(X) E^2 \wedge \cdots \wedge E^n$$

because $N^\flat|_{\partial M} = 0$, so only the first term in the interior product survives. Since $N^\flat(X) = \langle N, X \rangle$, we are done. □

Theorem 2.9 (Integration by parts). *For $u \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$,*

$$\int_M \langle \nabla u, X \rangle_g dV_g = \int_{\partial M} u \langle X, N \rangle_g dV_{\hat{g}} - \int_M u \operatorname{div} X dV_g.$$

Proof. Using the identity

$$\operatorname{div}(uX) = u \operatorname{div} X + \langle \nabla u, X \rangle_g$$

(exercise) and applying the divergence theorem,

$$\int_M \operatorname{div}(uX) dV_g = \int_{\partial M} u \langle X, N \rangle_g dV_{\hat{g}},$$

we have the result. □

Definition 2.10. The **Laplace–Beltrami operator** is defined by

$$\Delta f := \operatorname{div}(\nabla f).$$

In local coordinates,

$$\Delta f = \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i} \left(\sqrt{\det g} g^{ij} \frac{\partial f}{\partial x^j} \right). \quad (2.3)$$

Notice that

$$\Delta f = g^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} + \text{lower-order terms},$$

so the exotic-looking expression (2.3) is indeed a generalization of the Laplacian.

On homework you will show that the Laplace-Beltrami operator retains several familiar analytic properties of the ordinary Laplace operator on domains in $\mathbb{R}^n$. Time permitting at the end of the class, we will develop the theory of the Laplace operator on differential forms, which has deep consequences.

Part 2. Connections

Last week we discussed Riemmanian metrics. At some point we were motivated to try and develop a concept of *curvature* for a space with a metric. It turns out that one does not need a metric to discuss curvature. The modern formalism is clearest if one (temporarily) forgets about metrics and instead focuses on a different and equally fundamental concept.

3. COVARIANT DERIVATIVES (Tu 1/27)

3.1. **Motivation.** We begin anew by discussing the following question:

What is the “derivative” of a section of a vector bundle?

In Math 761, we learned several derivatives that work on certain types of tensors, which we now recall.

1. Given a vector field $X \in \mathfrak{X}(M)$ and a function $f \in C^\infty(M)$, we know how to form the derivative

$$X(f) \in C^\infty(M).$$

In fact, a tangent vector is by definition a thing that can take the derivative of a function.

2. For a differential k -form $\omega \in \Omega^k(M)$, we learned how to form the *exterior derivative*

$$d\omega \in \Omega^{k+1}(M).$$

Proving its existence on a general smooth manifold was one of the main achievements of 761.

3. For two vector fields $X, Y \in \mathfrak{X}(M)$, we can form the *Lie derivative / Lie bracket*

$$\mathcal{L}_X Y = [X, Y] \in \mathfrak{X}(M).$$

4. For a vector field $X \in \mathfrak{X}(M)$ and a general tensor $T \in \mathcal{T}^{(k,\ell)}(M)$, we also learned how to form the Lie derivative

$$\mathcal{L}_X T \in \mathcal{T}^{(k,\ell)}(M).$$

This can either be defined using an ugly formula involving the Lie bracket, or by the elegant “derivative under pullback by the flow of X ” definition (761 notes, Section 36.3).

These derivatives suffer from the following problems/deficiencies:

- (1-2) only make sense for alternating covariant tensors. In fact, there is not really a clear “differentiating vector field” X in the exterior derivative, since all k vector fields that you could insert into $d\omega$ are on equal footing.

- (3-4) depend on the choice of the vector field X in a neighborhood. This can be seen from the formula

$$\mathcal{L}_{fX}Y = f\mathcal{L}_XY - Y(f)X.$$

In this sense, the Lie derivative is not a true derivative, since it depends on the differentiating vector field X by more than just its value $X_x \in T_xM$.

- Also notice that all four of these notions rely on the special relationship between the tangent bundle TM (or the cotangent bundle T^*M) and the smooth structure on M . We still have no particular idea how to differentiate sections of a general vector bundle $E \rightarrow M$.

3.2. The definition. Let $E \rightarrow M$ be a vector bundle of rank r over $K = \mathbb{R}$ or $\mathbb{C}$. Recall that $\Gamma(E) = \Gamma(M, E)$ denotes the space of global sections of E . A **covariant derivative** (also known as a **connection**) on E is a map

$$\begin{aligned} \nabla : \mathfrak{X}(M) \times \Gamma(E) &\rightarrow \Gamma(E) \\ (X, s) &\mapsto \nabla_X s \end{aligned}$$

that satisfies the following axioms:

1. $C^\infty(M)$ -linearity in X :

$$\nabla_{fX+gX'}s = f\nabla_X s + g\nabla_{X'}s, \quad f, g \in C^\infty(M), X, X' \in \mathfrak{X}(M)$$

2. K -linearity in s :

$$\nabla_X(as + bt) = a\nabla_X s + b\nabla_X t, \quad a, b \in K \text{ constants}, s, t \in \Gamma(E)$$

3. Leibniz rule:

$$\nabla_X(fs) = X(f)s + f\nabla_X s, \quad f \in C^\infty(M), s \in \Gamma(E).$$

Let us recall that the *tensor characterization lemma* (761 notes, Lemma 35.4) gives an equivalence:

$$\left\{ C^\infty(M)\text{-multilinear functions on } \Gamma(E_1) \times \cdots \times \Gamma(E_n) \right\} \longleftrightarrow \left\{ \text{Global sections of } E_1^* \otimes \cdots \otimes E_n^* \right\}.$$

Therefore, in view of axiom (1) above, we can equivalently think of a covariant derivative as a map

$$\begin{aligned} \nabla : \Gamma(E) &\rightarrow \Omega^1(E) = \Gamma(T^*M \otimes E) \\ s &\mapsto \nabla s. \end{aligned}$$

The value $\nabla_X s$ is simply the evaluation of ∇s on the vector-field X . This is in fact a pointwise operation. In particular, given $x \in M$, a tangent *vector* $v \in T_xM$, and a section $s \in \Gamma(E)$, the expression

$$\nabla_v s = (\nabla s)(v) \in E_x$$

makes sense. This means that a covariant derivative does not suffer from the same problem as the Lie derivative.

3.3. Examples.

1. (The directional derivative on $\mathbb{R}^n$) Let $M = \Omega \subset \mathbb{R}^n$, be an open subset, and take

$$E = T\Omega \cong \underline{\mathbb{R}^n} = \Omega \times \mathbb{R}^n.$$

Let X and Y be vector fields on Ω , written as

$$X = \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix}, \quad Y = \begin{pmatrix} Y^1 \\ \vdots \\ Y^n \end{pmatrix}.$$

Define the **directional derivative**

$$\nabla_X^\circ Y = \begin{pmatrix} X(Y^1) \\ \vdots \\ X(Y^n) \end{pmatrix}.$$

The axioms (1-2) for a covariant derivative are obvious. Let's check the Leibniz rule:

$$\nabla_X^\circ(fY) = \begin{pmatrix} X(fY^1) \\ \vdots \\ X(fY^n) \end{pmatrix} = \begin{pmatrix} X(f)Y^1 + fX(Y^1) \\ \vdots \\ X(f)Y^n + fX(Y^n) \end{pmatrix} = X(f)Y + f\nabla_X^\circ Y.$$

2. (Product connection on the product bundle) Generalizing the previous example, let M be any smooth manifold and let

$$E = \underline{K}^r = M \times K^r$$

be the product (a.k.a trivial) bundle. A section $s \in \Gamma(E)$ may be written as

$$s = (s^\alpha)_{\alpha=1}^r,$$

where s^α are smooth functions. Define the **product connection**

$$\nabla_X^\circ s = (X(s^\alpha))_{\alpha=1}^r.$$

The Leibniz rule is checked just as in the first example.

3. (Induced connection on an embedded submanifold) Let

$$M \subset \mathbb{R}^N$$

be an embedded submanifold. Let

$$E = TM \subset T\mathbb{R}^N|_M$$

be the tangent bundle of M . Denote the orthogonal projection by

$$\pi : T\mathbb{R}^N|_M \longrightarrow TM.$$

Define the **induced connection**

$$\nabla = \pi \circ \nabla^\circ.$$

We can check the Leibniz rule as follows:

$$\nabla_X(fY) = \pi(\nabla_X^\circ(fY)) = \pi(X(f)Y^i + f\nabla_X^\circ Y^i) = X(f)\pi(Y) + f\pi(\nabla_X^\circ Y).$$

Since Y is tangent to M , we have $\pi(Y) = Y$. So we arrive at

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

4. Let us revisit the first/second example with a twist. Let

$$M = \Omega \subset \mathbb{R}^n, \quad E = \underline{K}^r = \mathbb{R}^n \times K^r.$$

Let

$$A_1(x), \dots, A_n(x) \in \text{Mat}_K^{r \times r}(\Omega),$$

be any collection of n matrix-valued functions on Ω . Define

$$\nabla_X^A s = \nabla_X^\circ s + X^i A_i(x) \cdot s.$$

In components, this means:

$$(\nabla_X^A s)^\alpha = X(s^\alpha) + X^i A_{i\beta}^\alpha s^\beta.$$

Let's check the Leibniz rule:

$$\begin{aligned} (\nabla_X^A(fs))^\alpha &= X(fs^\alpha) + X^i A_{i\beta}^\alpha (fs^\beta) \\ &= X(f)s^\alpha + fX(s^\alpha) + fX^i A_{i\beta}^\alpha s^\beta \\ &= (X(f)s)^\alpha + f(\nabla_X^A s)^\alpha. \end{aligned}$$

A clear moral of this example is: *connections are not unique!*

3.4. The space of connections. Having seen that connections are not unique, we shall now describe the space of all connections on a given vector bundle $E \rightarrow M$.

Proposition 3.1.

(a) Let

$$a \in \Omega^1(\text{End}(E)) = \Gamma(T^*M \otimes \text{End}(E))$$

be 1-form valued in the endomorphism bundle of E . If ∇ is a connection on E , then $\nabla' := \nabla + a$ is also a connection. Here ∇' acts by

$$\nabla'_X s = \nabla_X s + a(X) \cdot s.$$

(b) Given any two connections ∇ and ∇' on E , their difference $\nabla' - \nabla$ is an $\text{End}(E)$ -valued 1-form:

$$\nabla' - \nabla \in \Omega^1(\text{End}(E)).$$

(c) For $\lambda \in C^\infty(M)$, $\lambda\nabla + (1 - \lambda)\nabla'$ is again a connection.

Proof. Parts (a) and (c) are left as exercises.

We check (b). By the tensor characterization lemma, it suffices to check $C^\infty(M)$ -linearity of the difference $\nabla - \nabla'$ with respect to s . (It is already linear in X .) We have

$$\begin{aligned} (\nabla_X - \nabla'_X)(fs) &= \nabla_X(fs) - \nabla'_X(fs) \\ &= X(f)s + f\nabla_X s - X(f)s - f\nabla'_X s \\ &= f(\nabla_X - \nabla'_X)s. \end{aligned}$$

□

According to (a-b), the space of all connections on E forms an *affine space modeled on* $\Omega^1(\text{End}(E))$. This means that the space of connection is basically the same as the vector space $\Omega^1(\text{End}(E))$ except that there is no distinguished origin. By (c), given any constant $\lambda \in [0, 1]$, the linear interpolation

$$\lambda\nabla + (1 - \lambda)\nabla'$$

remains a connection. Hence, while adding two connections does not make sense, linearly interpolating between them does.

Finally, we should show that the space of connections on a vector bundle is always nonempty.

Proposition 3.2. *Given any vector bundle $E \rightarrow M$, there exists a connection on E .*

Proof. Let $\{U_a\}$ be a coordinate cover such that E is trivial over each U_a . Choose local frames $\{e_\alpha^a\}$ on U_a . Let $\{\varphi_a\}$ be a partition-of-unity subordinate to $\{U_a\}$. On each U_a , define the product connection

$$\nabla^{\circ,a} : \Gamma(E|_{U_a}) \rightarrow \Omega^1(E|_{U_a})$$

according to Example 2 above. Then define globally:

$$\nabla_X s = \sum_a \varphi_a \nabla_X^{\circ,a}(s|_{U_a}).$$

We check that this satisfies the Leibniz rule:

$$\begin{aligned} \nabla_X(f \cdot s) &= \sum_a \varphi_a \nabla^{\circ,a}(fs|_{U_a}) \\ &= \sum_a \varphi_a X(f)s|_{U_a} + \sum_a \varphi_a f \nabla^{\circ,a}(s|_{U_a}) \\ &= X(f) \sum_a \varphi_a s + f \sum_a \varphi_a \nabla^{\circ,a}(s|_{U_a}) \\ &= X(f)s + f \nabla_X s. \end{aligned}$$

□

4. LOCAL DESCRIPTION OF A CONNECTION (WED 1/28 ONLINE)

4.1. **Locality.** We first clarify the following point.

Proposition 4.1. *A connection is determined locally. More precisely, given sections $s, s' \in \Gamma(E)$, if there exists an open set $V \subset M$ such that $s|_V = s'|_V$, then $\nabla s|_V = \nabla s'|_V$.*

Proof. It suffices to show that $(\nabla_X s)(p) = (\nabla_X s')(p)$ for all $p \in V$ and all $X \in \mathfrak{X}(M)$.

Fix $p \in V$ and $X \in \mathfrak{X}(M)$. Choose a cutoff function φ with $\varphi \equiv 1$ near p and $\text{supp}(\varphi) \subset V$. We have

$$\nabla_X(\varphi s)(p) = X(\varphi)(p) s(p) + \varphi(p)(\nabla_X s)(p) = (\nabla_X s)(p).$$

On the other hand, the same calculation shows that $\nabla_X(\varphi s')(p) = (\nabla_X s')(p)$. But since $s|_V = s'|_V$, we have $\varphi s = \varphi s'$ globally, so

$$\nabla_X(\varphi s)(p) = \nabla_X(\varphi s')(p).$$

Therefore $(\nabla_X s)(p) = (\nabla_X s')(p)$. Since $p \in V$ and X were arbitrary, the result follows. $\square$

Corollary 4.2. *Let $U \subset M$ be open. For any $s \in \Gamma(U, E)$, the expression $\nabla s \in \Omega_U^1(E)$ makes sense.*

Proof. Given $p \in U$, define

$$\nabla s(p) := \nabla \tilde{s}(p),$$

where $\tilde{s} \in \Gamma(E)$ is any global section that agrees with s on a neighborhood of p . That is, there exists an open set $V \ni p$ such that $s|_V = \tilde{s}|_V$ (for example, one may take $\tilde{s} = \varphi s$ as before). By the locality proposition, the value $\nabla \tilde{s}(p)$ depends only on the restriction of s near p , hence ∇s is well-defined. $\square$

4.2. **Connection matrices.** Let ∇ be any connection on $E \rightarrow M$. Given a local frame $\{e_\alpha\}_{\alpha=1}^r$ over a coordinate chart $U = \{x^i\}_{i=1}^n$, in view of Corollary 4.2, we can make the following definition.

Definition 4.3. The **connection matrices** of ∇ with respect to $\{e_\alpha\}$,

$$A_{i\alpha}^\beta, \quad i = 1, \dots, n$$

are defined by

$$\boxed{\nabla \frac{\partial}{\partial x^i} e_\alpha = A_{i\alpha}^\beta e_\beta.}$$

Example 4.4 (From HW 2, done in class 2/5). Let $\nabla = \pi \circ \nabla^\circ$ be the connection on TS^2 defined in Lecture 3. Let $U = \{z > 0\} \cap S^2$ be the upper unit hemisphere. Take the following frame over U :

$$e_1 = \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ z \\ -y \end{pmatrix}.$$

As local coordinates on U , take $x^1 = x$ and $x^2 = y$. We want to calculate the connection matrices of ∇ in these coordinates and the given frame.

On the sphere we have $z = \sqrt{1 - x^2 - y^2}$, so $\frac{\partial z}{\partial x} = -\frac{x}{z}$ and $\frac{\partial z}{\partial y} = -\frac{y}{z}$. We calculate

$$\nabla_{\frac{\partial}{\partial x^1}}^\circ e_1 = \frac{\partial}{\partial x}(z, 0, -x)^T = \left(-\frac{x}{z}, 0, -1\right)^T.$$

For vector fields X, Y tangent to S^2 , we have:

$$\nabla_X Y = \pi(\nabla_X^\circ Y) = \nabla_X^\circ Y - \langle \nabla_X^\circ Y, \vec{r} \rangle \vec{r}.$$

We calculate

$$\left\langle \nabla_{\frac{\partial}{\partial x^1}}^\circ e_1, \vec{r} \right\rangle = -\frac{x^2}{z} - z = -\frac{x^2 + z^2}{z} = \frac{y^2 - 1}{z}.$$

Plugging in, we get

$$\begin{aligned} \nabla_{\frac{\partial}{\partial x^1}} e_1 &= \left(-\frac{x}{z}, 0, -1\right)^T + \left(\frac{1 - y^2}{z}\right)(x, y, z)^T \\ &= \left(\frac{x(-1 + 1 - y^2)}{z}, \frac{y(1 - y^2)}{z}, -1 + 1 - y^2\right)^T \\ &= \left(-\frac{xy^2}{z}, \frac{y(1 - y^2)}{z}, -y^2\right)^T. \end{aligned}$$

By examining the first two coefficients, we see that the answer must be

$$\nabla_{\frac{\partial}{\partial x^1}} e_1 = \frac{-xy^2}{z^2} e_1 + \frac{y(1 - y^2)}{z^2} e_2.$$

This determines A_{11}^1 and A_{11}^2 , the first column of A_1 . As a reality check, let's make sure the third component is also right:

$$\frac{1}{z^2} (x^2 y^2 - y^2(1 - y^2)) = \frac{y^2}{z^2} (x^2 + y^2 - 1) = -y^2 \checkmark.$$

Computing the other components in similar fashion, one can get

$$\begin{aligned} A_1 &= \frac{1}{1 - x^2 - y^2} \begin{pmatrix} -xy^2 & x^2 y \\ y(1 - y^2) & x(y^2 - 1) \end{pmatrix} \\ A_2 &= \frac{1}{1 - x^2 - y^2} \begin{pmatrix} y(x^2 - 1) & x(1 - x^2) \\ xy^2 & -x^2 y \end{pmatrix}. \end{aligned}$$

Proposition 4.5. *The connection matrices determine ∇ completely, by the local formula*

$$\boxed{\nabla_X s \stackrel{loc}{=} X^i \left(\frac{\partial s^\alpha}{\partial x^i} + A_{i\beta}^\alpha s^\beta \right) e_\alpha.} \quad (4.1)$$

Proof. Let $s = s^\alpha e_\alpha \in \Gamma(U, E)$. For a vector field $X = X^i \frac{\partial}{\partial x^i}$, we compute:

$$\begin{aligned} \nabla_X s &= \nabla_{X^i \frac{\partial}{\partial x^i}} (s^\alpha e_\alpha) = X^i \nabla_{\frac{\partial}{\partial x^i}} (s^\alpha e_\alpha) \\ &= X^i \left(\frac{\partial s^\alpha}{\partial x^i} e_\alpha + s^\alpha \nabla_{\frac{\partial}{\partial x^i}} e_\alpha \right). \end{aligned}$$

Using the definition $\nabla_{\frac{\partial}{\partial x^i}} e_\alpha = A_{i\alpha}^\beta e_\beta$, we obtain

$$\nabla_X s = X^i \left(\frac{\partial s^\alpha}{\partial x^i} e_\alpha + s^\alpha A_{i\alpha}^\beta e_\beta \right)$$

Exchanging the index labels α and β in the second term and factoring out e_α , we obtain the desired formula (4.1) $\square$

Remark 4.6. The formula (4.1) should be compared with Example 4 of the last class. What we have just shown is that every connection is *locally* of the same form as Example 4!

Notation: We define

$$\left(\nabla_{\frac{\partial}{\partial x^i}} s\right)^\alpha = \boxed{\frac{\partial s^\alpha}{\partial x^i} + A_{i\beta}^\alpha s^\beta =: \nabla_i s^\alpha.}$$

Equivalently, we have

$$\nabla s = \nabla_i s^\alpha dx^i \otimes e_\alpha \in \Omega^1(E),$$

i.e., $\nabla_i s^\alpha$ are the tensor components of ∇s as an E -valued 1-form. For a vector field X , the formula (4.1) becomes

$$\boxed{\nabla_X s = X^i \nabla_i s^\alpha e_\alpha.}$$

Warning: The component function $\nabla_i s^\alpha$ should not be confused with a derivative of the single component function s^α . Evidently, it also depends on the other components s^β , $\beta \neq \alpha$. As long as one keeps this in mind, it is perfectly good notation.

4.3. The transformation rule. How do the connection matrices change when we change the frame?

Let $\{e_\alpha\}$ and $\{e'_\beta\}$ be two local frames over U , with corresponding connection matrices $A_{i\beta}^\alpha$ and $A'_{i\beta}{}^\alpha$. Define the change-of-frame matrix σ by

$$\boxed{e_\alpha = \sigma^\beta{}_\alpha e'_\beta.}$$

We compute how the connection matrices transform under a change of frame. On one hand, we have

$$\begin{aligned} \nabla_{\frac{\partial}{\partial x^i}} e_\alpha &= \nabla_{\frac{\partial}{\partial x^i}} (\sigma^\beta{}_\alpha e'_\beta) \\ &= \frac{\partial \sigma^\beta{}_\alpha}{\partial x^i} e'_\beta + \sigma^\beta{}_\alpha \nabla_{\frac{\partial}{\partial x^i}} e'_\beta \\ &= \left(\frac{\partial \sigma^\gamma{}_\alpha}{\partial x^i} + \sigma^\beta{}_\alpha A'_{i\beta}{}^\gamma \right) e'_\gamma. \end{aligned}$$

On the other hand, we have

$$\nabla_{\frac{\partial}{\partial x^i}} e_\alpha = A_{i\alpha}^\beta e_\beta = A_{i\alpha}^\beta \sigma^\gamma{}_\beta e'_\gamma.$$

Comparing coefficients of e'_γ , we get

$$A_{i\alpha}^\beta \sigma^\gamma{}_\beta = \frac{\partial \sigma^\gamma{}_\alpha}{\partial x^i} + \sigma^\beta{}_\alpha A'_{i\beta}{}^\gamma.$$

In matrix notation, this becomes

$$\sigma \cdot A_i = \frac{\partial \sigma}{\partial x^i} + A'_i \cdot \sigma.$$

Multiplying on the right by σ^{-1} and rearranging, we obtain the **transformation law for connection matrices under a change of frame**:

$$A'_i = \sigma \cdot A_i \cdot \sigma^{-1} - \frac{\partial \sigma}{\partial x^i} \cdot \sigma^{-1}. \quad (4.2)$$

Note that the extra term corresponds to the fact that connection matrices *do not* transform like tensors.

4.4. Coordinate / frame description. Let $E \rightarrow M$ be a vector bundle. Recall that there are two ways of thinking about E : as an abstract topological space together with a map to M satisfying some axioms, or as a collection of trivial bundles patched together. For the second point of view, one chooses an open cover $\{U_a\}$ of M such that each restriction $E|_{U_a}$ is trivial. On each U_a , fix a local frame $\{e_\alpha^{(a)}\}_{\alpha=1}^r$. On overlaps $U_a \cap U_b$, the frames are related by the transition functions

$$\sigma_{ab} : U_a \cap U_b \rightarrow \text{GL}(r)$$

defined by

$$e_\alpha^{(a)} = \sigma_{ab}{}^\beta{}_\alpha e_\beta^{(b)}.$$

The σ_{ab} 's satisfy *cocycle conditions*, one of which is $\sigma_{ba} = \sigma_{ab}^{-1}$. Conversely, recall that given any collection of transition functions satisfying the cocycle conditions, one can make a vector bundle just by gluing together trivial bundles (761 notes, Section 31).

We now describe connections from this point of view. Let $(A_i^{(a)})$ denote the connection matrices on U_a with respect to the local frames $\{e_\alpha^{(a)}\}$. The above transformation law becomes the **compatibility condition**

$$A_i^{(b)} = \sigma_{ab} A_i^{(a)} \sigma_{ba} - \frac{\partial \sigma_{ab}}{\partial x^i} \sigma_{ba}$$

on $U_a \cap U_b$. Conversely, given any collection of matrix-valued functions $\{A_i^{(a)}\}$ satisfying the compatibility condition on the overlaps $U_a \cap U_b$, there exists a unique globally defined connection ∇ on E whose connection matrices are exactly $\{A_i^{(a)}\}$.

Remark 4.7. If one is using different coordinates $U_a = \{x^i\}$ and $U_b = \{y^j\}$ on the two charts, then the compatibility condition reads

$$A_j^{(b)} = \sigma_{ab} \frac{\partial x^i}{\partial y^j} A_i^{(a)} \sigma_{ba} - \frac{\partial \sigma_{ab}}{\partial y^j} \sigma_{ba}.$$

Remark 4.8. Sooner or later we need to answer the question: why do we keep calling a covariant derivative a “connection”?

5. PARALLEL TRANSPORT (TU 2/3)

5.1. Covariant derivative along a path. Let $\pi : E \rightarrow M$ be a vector bundle and $\gamma : [a, b] \rightarrow M$ a path. A **section of E along γ** is a map $s : [a, b] \rightarrow E$ such that

$$\pi \circ s = \gamma.$$

Equivalently, $s(t) \in E_{\gamma(t)}$ for all $t \in [a, b]$. The following definition/lemma should be no surprise.

Definition/Lemma 5.1 (Covariant derivative along a path). *Fix a connection ∇ on $E \rightarrow M$. Given a piecewise C^1 section $s : [a, b] \rightarrow E$ along γ , there exists a well-defined **covariant derivative of s along γ** , written $\frac{Ds}{dt}$, which is again a section of E along γ ; in other words,*

$$\frac{Ds}{dt} \in E_{\gamma(t)}$$

for $t \in [a, b]$. It has the following properties:

(A) (Additivity) For sections s_1 and s_2 along γ ,

$$\frac{D(s_1 + s_2)}{dt} = \frac{Ds_1}{dt} + \frac{Ds_2}{dt}.$$

(B) (Leibniz rule) For $f \in C^\infty([a, b])$ and a section s along γ ,

$$\frac{D(f \cdot s)}{dt} = \frac{df}{dt} s + f \frac{Ds}{dt}.$$

(C) (Relationship with ∇) If $s(t) = \tilde{s}(\gamma(t))$, where $\tilde{s} \in \Gamma(E)$, then

$$\frac{Ds}{dt} = \nabla_{\gamma'(t)} \tilde{s}(\gamma(t)).$$

(D) (Local expression) Given a local frame $\{e_\alpha\}$, let $A_{i\beta}^\alpha$ be the connection matrices of ∇ and write $s(t) = s^\alpha(t) e_\alpha|_{\gamma(t)}$. We then have

$$\boxed{\frac{Ds}{dt} = \left(\frac{ds^\alpha(t)}{dt} + \frac{d\gamma^i}{dt} A_{i\beta}^\alpha(\gamma(t)) s^\beta(t) \right) e_\alpha|_{\gamma(t)}}.$$

Proof. We take (D) as our definition. One should then check that the result is independent of changes of frame. This is a good exercise using the transformation law (4.2).

We also check that (C) will be satisfied with this definition. Suppose $s(t) = \tilde{s}(\gamma(t))$ for some section $\tilde{s} \in \Gamma(U, E)$ defined on an open set U containing $\gamma([a, b])$. Write $\tilde{s} = \tilde{s}^\alpha e_\alpha$ in a local frame, so $s^\alpha(t) = \tilde{s}^\alpha(\gamma(t))$. In local coordinates $\{x^i\}$ we have $\gamma'(t) = \frac{d\gamma^i}{dt}(t) \frac{\partial}{\partial x^i}$, and we use the local formula $\nabla_i \tilde{s}^\alpha = \frac{\partial \tilde{s}^\alpha}{\partial x^i} + A_{i\beta}^\alpha \tilde{s}^\beta$. Then

$$\nabla_{\gamma'(t)} \tilde{s} = \frac{d\gamma^i}{dt} \nabla_i \tilde{s}^\alpha(\gamma(t)) e_\alpha|_{\gamma(t)} = \frac{d\gamma^i}{dt} \left(\frac{\partial \tilde{s}^\alpha}{\partial x^i} + A_{i\beta}^\alpha \tilde{s}^\beta \right) (\gamma(t)) e_\alpha|_{\gamma(t)}.$$

By the chain rule,

$$\frac{d\gamma^i}{dt} \frac{\partial \tilde{s}^\alpha}{\partial x^i}(\gamma(t)) = \frac{d}{dt}(\tilde{s}^\alpha(\gamma(t))) = \frac{ds^\alpha}{dt},$$

hence

$$\frac{Ds}{dt} = \nabla_{\gamma'(t)} \tilde{s}(\gamma(t)) = \left(\frac{ds^\alpha}{dt} + A_{i\beta}^\alpha(\gamma(t)) s^\beta(t) \frac{d\gamma^i}{dt} \right) e_\alpha|_{\gamma(t)}.$$

□

5.2. Parallel transport. A section s is called **parallel** along γ if

$$\frac{Ds(t)}{dt} = 0.$$

Examples 5.2 (From 2/5 class). 1. For ∇° on a trivial bundle $E \cong U \times \mathbb{R}^r$, the equation becomes

$$\frac{Ds(t)}{dt} = 0 \iff \dot{s}(t) = 0$$

in the trivialization; hence parallel sections are constant. We drew a picture in class.

2. For TS^2 with $\nabla = \pi \circ \nabla^\circ$ (where $\pi_p : \mathbb{R}^3 \rightarrow T_p S^2$ is orthogonal projection), a section $v(t) \in T_{\gamma(t)} S^2$ is parallel iff

$$0 = \nabla_{\dot{\gamma}} v = \pi_{\gamma(t)}(\dot{v}(t)).$$

Equivalently, $\dot{v}(t) \perp T_{\gamma(t)} S^2$, i.e. $\dot{v}(t)$ is proportional to $\gamma(t)$. We drew a picture in class of a parallel vector moving along the great circle that it is tangent to.

Proposition 5.3. *Let $\gamma : [a, b] \rightarrow M$ be a piecewise C^1 path and let $s_0 \in E_{\gamma(a)}$. There exists a unique parallel section s along γ such that $s(a) = s_0$.*

Proof. Break up the path γ into a concatenation $\gamma_1 * \gamma_2 * \dots * \gamma_n$ of finitely many differentiable paths, each of whose image lies within a single coordinate chart over which E is trivial. If we prove the proposition (existence and uniqueness) for each path γ_i , we will have proven it for γ . It therefore suffices to assume that $\gamma = \gamma_1$.

In a coordinate chart, the equation $\frac{Ds}{dt} = 0$ amounts to the system

$$\frac{ds^\alpha}{dt} + \frac{d\gamma^i}{dt} A_{i\beta}^\alpha(\gamma(t)) s^\beta(t) = 0, \quad \alpha = 1, \dots, r.$$

Since the path $\gamma(t)$ is fixed, the coefficients do not depend on $s^\alpha(t)$. This is therefore a first-order linear ODE system. As with any system, Picard's local existence theorem allows us to solve the ODE uniquely on short time intervals. It is an exercise (on homework) to prove that *linear* ODEs in fact enjoy global solutions. (In class we sketched two proofs of this fact, which I omit in order not to trivialize your homework.) □

Definition 5.4 (Parallel transport). Let $\gamma : [a, b] \rightarrow M$ be a piecewise C^1 path, and let ∇ be a connection on $E \rightarrow M$. The associated **parallel transport map** along γ is the map

$$P_{a,b}^\gamma : E_{\gamma(a)} \longrightarrow E_{\gamma(b)}$$

defined by

$$P_{a,b}^\gamma(s_0) := s(b),$$

where s is the unique parallel section along γ satisfying $\frac{Ds}{dt} = 0$ and $s(a) = s_0$.

Lemma 5.5 (Properties of parallel transport). *Let $\gamma : [a, b] \rightarrow M$ be a piecewise C^1 path.*

(1) (Invertibility) *The map $P_{a,b}^\gamma : E_{\gamma(a)} \rightarrow E_{\gamma(b)}$ is an isomorphism.*

(2) (Linearity) *For $s_0, s_1 \in E_{\gamma(a)}$ and constants $r_0, r_1 \in K$, we have*

$$P_{a,b}^\gamma(r_0 s_0 + r_1 s_1) = r_0 P_{a,b}^\gamma(s_0) + r_1 P_{a,b}^\gamma(s_1).$$

(3) (Concatenation) *For $c \in [a, b]$, we have*

$$P_{a,b}^\gamma = P_{c,b}^\gamma \circ P_{a,c}^\gamma.$$

Proof. These properties follow directly from uniqueness (1 & 3) and linearity (2) of the ODE system. For instance, the isomorphism property (1) can be seen by considering parallel transport by the reversed path. In class I sketched the proofs verbally and it is a good exercise to think them through again. $\square$

5.3. Recovering the connection. The connection ∇ can be recovered from its parallel transport map.

Claim.

$$\boxed{\frac{Ds}{dt} = \lim_{h \rightarrow 0} \frac{s(t+h) - P_{t,t+h}^\gamma s(t)}{h}}.$$

Remark 5.6. Observe that the formula makes sense: the elements in the numerator belong to the same fiber $E_{\gamma(t+h)}$, so we can subtract them and divide by the constant h . The result is still a point in the total space of E , and the limit can be understood in the topology of E , which holds no surprises in local coordinates.

Proof of claim. The formula can be checked in local coordinates (exercise). It is instructive, however, to give the following alternative argument.

Define an operator on sections along γ by

$$(Ts)(t) := \lim_{h \rightarrow 0} \frac{s(t+h) - P_{t,t+h}^\gamma s(t)}{h},$$

whenever the limit exists. Then the claim is that

$$\frac{Ds}{dt} = (Ts)(t).$$

Observe that for a parallel section s the formula is true, as both sides are zero.

By parallel transport, one can construct a parallel frame along γ . Since $\frac{D}{dt}$ is characterized by its values on a frame, linearity, and the Leibniz rule, it will suffice to check that T satisfies the Leibniz rule. This will take us right back to Calculus I.

Let $f \in C^\infty([a, b])$ and let s be a section along γ . Since parallel transport is linear on each fiber, we have

$$P_{t,t+h}^\gamma(f(t) s(t)) = f(t) P_{t,t+h}^\gamma s(t).$$

Therefore

$$\frac{f(t+h)s(t+h) - P_{t,t+h}^\gamma(f(t)s(t))}{h} = \frac{f(t+h)s(t+h) - f(t)P_{t,t+h}^\gamma s(t)}{h}.$$

Add and subtract $f(t)s(t+h)$ to rewrite this as

$$\frac{f(t+h) - f(t)}{h} s(t+h) + f(t) \frac{s(t+h) - P_{t,t+h}^\gamma s(t)}{h}.$$

Taking $h \rightarrow 0$ gives $s(t+h) \rightarrow s(t)$ and $\frac{f(t+h)-f(t)}{h} \rightarrow \frac{df}{dt}$, hence in the limit we obtain

$$T(fs)(t) = \frac{df}{dt} s(t) + f(Ts)(t),$$

which is exactly the Leibniz rule for T . $\diamond$

Remark 5.7. We now know why a covariant derivative is called a “connection:” it is because the associated parallel transport operator “connects” different fibers along a path. In fact, by the previous claim, the parallel transport operator is sufficient to recover the covariant derivative; so the two notions really are equivalent.

An alternative approach (see e.g. my “Vector bundles and gauge theory” notes, from Math 865) is to first define a connection as a collection of parallel transport operators subject to some reasonable axioms, then to construct the covariant derivative using the formula in the previous claim.

6. GAUGE-EQUIVALENCE (TH 2/5)

Today we went back and did Examples 4.4 and 5.2 to make sure everything was clear. We then reviewed all of the equivalent pictures that we have of a connection. Later we discussed the following

Question. When should two connections A and B be considered isomorphic?

6.1. Equivalent definitions of a connection. We now change notation: rather than always using the symbol “ ∇ ,” a connection should really have its own name: for instance, A . Recall that a connection A can be thought of in the following equivalent ways:

1. A covariant derivative

$$\nabla = \nabla^A : \Gamma(E) \rightarrow \Gamma(T^*M \otimes E) = \Omega^1(E),$$

i.e. a K -linear map such that $\nabla^A(fs) = df \otimes s + f\nabla^A s$.

2. A collection of matrix-valued functions $A_i^{(a)}$ (the connection matrices) on U_a satisfying the compatibility conditions on overlaps $U_a \cap U_b$.
3. For each piecewise C^1 path $\gamma : [a, b] \rightarrow M$, a covariant derivative along γ ,

$$\frac{D^A}{dt} : \Gamma(\gamma^* E) \rightarrow \Gamma(\gamma^* E),$$

satisfying linearity and the Leibniz rule with respect to functions on $[a, b]$.

4. For each piecewise C^1 path $\gamma : [a, b] \rightarrow M$, a family of parallel transport operators: given $x := \gamma(a)$ and $s_0 \in E_x = E_{\gamma(a)}$, one obtains

$$P_{a,b}^{A,\gamma}(s_0) \in E_{\gamma(b)}$$

and hence a linear map

$$P_{a,b}^{A,\gamma} : E_{\gamma(a)} \rightarrow E_{\gamma(b)}.$$

satisfying the properties in Lemma 5.5; from these one recovers $\frac{D^A}{dt}$ by the formula in the claim at the end of last class.

5. There is another approach that you can think about if you are interested. A connection can be encoded by a choice of **horizontal distribution** on the total space E , namely a smooth distribution $H \subset TE$ (a smooth subbundle) such that for each $e \in E$

$$T_e E = H_e \oplus V_e, \quad V_e := \ker(d\pi_e).$$

In particular, given $s_0 \in E_x$ (so $x = \pi(s_0)$), this determines a splitting of the short exact sequence

$$0 \longrightarrow E_x \longrightarrow T_{s_0} E \xrightarrow[\underset{\substack{\longleftarrow A}}{d\pi_{s_0}}]{} T_x M \longrightarrow 0.$$

Specifying horizontal lifts of tangent vectors tells you how to do parallel transport.

6.2. Isomorphism = gauge-equivalence. Two connections A and B are called **gauge-equivalent** (or **isomorphic**) over an open set U if there exists an invertible section $\sigma \in \Gamma(U, \text{End}(E))$ such that, for every section $s \in \Gamma(U, E)$,

$$\nabla^B(\sigma(s)) = \sigma \nabla^A(s).$$

Equivalently,

$$\nabla^B(s) := \sigma \nabla^A(\sigma^{-1}s).$$

In this case, we write $B = \sigma(A)$.

In terms of connection matrices A_i and B_i in local coordinates $\{x^i\}$ and a fixed frame $\{e_\alpha\}$, this condition is equivalent to

$$B_i = \sigma A_i \sigma^{-1} - \frac{\partial \sigma}{\partial x^i} \sigma^{-1}$$

(exercise on HW2). Note that this is just the same formula as for a change-of-frame!

Exercise 6.1 (HW3). Show that gauge-equivalence respects parallel transport, in the following sense:

$$P_{a,b}^{\sigma(A),\gamma} = \sigma_{\gamma(b)} P_{a,b}^{A,\gamma} \sigma_{\gamma(a)}^{-1} : E_{\gamma(a)} \rightarrow E_{\gamma(b)}.$$

This was sketched in class. The point is that if a path $s(t)$ is covariantly constant for A , then $\sigma(s(t))$ is covariantly constant for $\sigma(A)$.

6.3. Flatness. A connection is called **flat** if it is locally gauge-equivalent to the product connection.

Example 6.2. Consider the product bundle $\underline{K}^r := M \times K^r \rightarrow M$, and let A° denote the product connection. The local connection matrices of A° in the product frame are all zero. Moreover, parallel transport along any path $\gamma : [a, b] \rightarrow M$ is

$$P_{a,b}^{A^\circ, \gamma} = \mathbb{1}_{K^r}.$$

Given any invertible section

$$\sigma(x) \in \Gamma(M, K^{r \times r})$$

(i.e. a smooth map $\sigma : M \rightarrow GL(r, K)$), we can define a connection on $\underline{K}^r$ by

$$A = \sigma(A^\circ).$$

This is by definition a flat connection. Its connection matrices in any coordinate chart, and with respect to the product frame, are

$$A_i = -\frac{\partial \sigma}{\partial x^i} \sigma^{-1}.$$

The moral of this example is that not only does every bundle have a plethora of connections, but some bundles even have a plethora of *flat* connections.

Question. How can one tell whether a given connection is flat?

Example 6.3. Consider $\nabla = \pi \circ \nabla^\circ$ on $TS^2 \rightarrow S^2$. Let γ be a closed loop based at the north pole, defined as a concatenation

$$\gamma = \gamma_1 * \gamma_2 * \gamma_3,$$

where γ_1 and γ_3 run along meridians, and γ_2 runs along the equator.

We drew a picture showing that parallel transport satisfies $P_\gamma(v) = w$ for some vectors $v, w \in T_{\text{north pole}} S^2$ with $v \neq w$. Compare with the product connection, for which $P_{a,b}^\gamma = \text{Id}$; and if we change frame by σ , then

$$\sigma_{\gamma(b)} P_{a,b}^\gamma \sigma_{\gamma(a)}^{-1} = \sigma_{\gamma(b)} \text{Id} \sigma_{\gamma(a)}^{-1} = \text{Id},$$

so parallel transport around any loop must be trivial. Hence, at least over the upper hemisphere, ∇ is *not* gauge-equivalent to ∇° .

In the next part, we will completely solve the problem of determining when a connection is flat.

Part 3. Curvature

7. THE CURVATURE OPERATOR (Tu 2/10)

7.1. The definition. Let A be a connection on a vector bundle $E \rightarrow M$. The **curvature** F_A of a connection A is the map

$$F_A : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(E),$$

defined by

$$F_A(X, Y)s = \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]} s.$$

Theorem 7.1 (Tensoriality of curvature). F_A is $C^\infty(M)$ -linear in each argument. Hence the curvature defines a smooth endomorphism-valued 2-form

$$F_A \in \Omega^2(\text{End } E),$$

i.e. an alternating section of $\Lambda^2 T^*M \otimes \text{End}(E)$.

Proof. Write $F = F_A$ and $\nabla = \nabla_A$.

First, we check tensoriality in the first argument:

$$\begin{aligned} F(fX, Y)s &= \nabla_{fX} \nabla_Y s - \nabla_Y \nabla_{fX} s - \nabla_{[fX, Y]} s \\ &= f \nabla_X \nabla_Y s - \nabla_Y (f \nabla_X s) - \nabla_{f[X, Y] - Y(f)X} s \\ &= f \nabla_X \nabla_Y s - (Y(f) \nabla_X s + f \nabla_Y \nabla_X s) - f \nabla_{[X, Y]} s + Y(f) \nabla_X s \\ &= f (\nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]} s) \\ &= f F(X, Y)s. \end{aligned}$$

Second argument:

$$F(X, fY) = -F(fY, X) = -fF(Y, X) = fF(X, Y).$$

Third argument:

$$\begin{aligned} F(X, Y)(fs) &= \nabla_X (\nabla_Y (fs)) - \nabla_Y (\nabla_X (fs)) - \nabla_{[X, Y]} (fs) \\ &= \nabla_X (Y(f)s + f \nabla_Y s) - \nabla_Y (X(f)s + f \nabla_X s) - ([X, Y](f)s + f \nabla_{[X, Y]} s) \\ &= (XY(f) - YX(f) - [X, Y](f))s \\ &\quad + Y(f) \nabla_X s - Y(f) \nabla_X s + X(f) \nabla_Y s - X(f) \nabla_Y s \\ &\quad + f (\nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X, Y]} s) \\ &= f F(X, Y)s. \end{aligned}$$

Hence $F_A \in \Omega^2(\text{End } E)$. □

Corollary 7.2. Let $\sigma \in \Gamma(\text{Aut}(E))$ be a gauge transformation, and define the transformed connection

$$\nabla_X^{\sigma(A)} s := \sigma(\nabla_X^A(\sigma^{-1}s))$$

If F_A and $F_{\sigma(A)}$ denote the curvatures of ∇^A and $\nabla^{\sigma(A)}$, then

$$F_{\sigma(A)} = \sigma F_A \sigma^{-1}.$$

Proof. Let $X, Y \in \mathfrak{X}(M)$ and $s \in \Gamma(E)$. By definition of the transformed connection,

$$\nabla_X^{\sigma(A)} = \sigma \circ \nabla_X^A \circ \sigma^{-1}.$$

Hence

$$\begin{aligned} \nabla_X^{\sigma(A)} \nabla_Y^{\sigma(A)} s &= (\sigma \nabla_X^A \sigma^{-1})(\sigma \nabla_Y^A (\sigma^{-1} s)) \\ &= \sigma \nabla_X^A (\nabla_Y^A (\sigma^{-1} s)), \\ \nabla_Y^{\sigma(A)} \nabla_X^{\sigma(A)} s &= \sigma \nabla_Y^A (\nabla_X^A (\sigma^{-1} s)), \\ \nabla_{[X,Y]}^{\sigma(A)} s &= \sigma \nabla_{[X,Y]}^A (\sigma^{-1} s). \end{aligned}$$

Therefore, using the curvature definition,

$$\begin{aligned} F_{\sigma(A)}(X, Y)s &= \nabla_X^{\sigma(A)} \nabla_Y^{\sigma(A)} s - \nabla_Y^{\sigma(A)} \nabla_X^{\sigma(A)} s - \nabla_{[X,Y]}^{\sigma(A)} s \\ &= \sigma \left(\nabla_X^A \nabla_Y^A (\sigma^{-1} s) - \nabla_Y^A \nabla_X^A (\sigma^{-1} s) - \nabla_{[X,Y]}^A (\sigma^{-1} s) \right) \\ &= \sigma (F_A(X, Y) \sigma^{-1} s). \end{aligned}$$

Since this holds for all s ,

$$F_{\sigma(A)}(X, Y) = \sigma F_A(X, Y) \sigma^{-1}.$$

Hence, as an $\text{End}(E)$ -valued 2-form, $F_{\sigma(A)} = \sigma F_A \sigma^{-1}$. $\square$

Corollary 7.3. *For some point $p \in M$, if $F_A(p) \neq 0$, then A is not flat on any neighborhood of p .*

Proof. If A is flat then one can choose a frame in which it is equal to the product connection, whose curvature tensor obviously vanishes. By the previous corollary, if the curvature at some point vanishes in some frame, then it still vanishes after conjugation by any endomorphism. This proves the contrapositive of the statement. $\square$

7.2. Example: $\mathbb{S}^n$. On $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, let $\nabla = \pi \circ \nabla^\circ$, where $\nabla_X^\circ Y = X(Y)$ is the ambient flat connection and π is the orthogonal projection onto the tangent space TS^n .

Lemma 7.4. *Suppose X, Y, Z are tangent vector fields on $\mathbb{S}^n$.*

1. $\nabla_X Y = X(Y) + \langle X, Y \rangle \vec{r}$.
2. $X(\vec{r}) = X$.
3. $\nabla_X Y - \nabla_Y X = [X, Y]$ (*torsion freeness*).
4. $X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$ (*Leibniz rule for $\langle \cdot, \cdot \rangle$*).

Proof. These were explained briefly in class; we leave the details as a homework exercise (HW3). $\square$

Now extend X, Y, Z to a neighborhood of $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$. Then

$$\begin{aligned}
\nabla_X \nabla_Y Z &= X(\nabla_Y Z) + \langle X, \nabla_Y Z \rangle \vec{r} \\
&= X(Y(Z) + \langle Y, Z \rangle \vec{r}) + \langle X, \nabla_Y Z \rangle \vec{r}, \\
&= X(Y(Z)) + \langle Y, Z \rangle X + X \langle Y, Z \rangle \vec{r} + \langle X, \nabla_Y Z \rangle \vec{r}, \\
\nabla_Y \nabla_X Z &= Y(X(Z)) + \langle X, Z \rangle Y + Y \langle X, Z \rangle \vec{r} + \langle Y, \nabla_X Z \rangle \vec{r}, \\
\nabla_{[X, Y]} Z &= [X, Y](Z) + \langle [X, Y], Z \rangle \vec{r}.
\end{aligned}$$

Substituting into the curvature formula and noticing already that the first three terms cancel, we have

$$F(X, Y)Z = \langle Y, Z \rangle X - \langle X, Z \rangle Y + \left(X \langle Y, Z \rangle - Y \langle X, Z \rangle - \langle [X, Y], Z \rangle + \langle X, \nabla_Y Z \rangle - \langle Y, \nabla_X Z \rangle \right) \vec{r}.$$

Now, *the coefficient of $\vec{r}$ must vanish*, because we know that the answer is tangent to the sphere! So we arrive at the simple formula

$$\boxed{F(X, Y)Z = \langle Y, Z \rangle X - \langle X, Z \rangle Y.} \quad (7.1)$$

As a reality check, we should make sure that the coefficient of $\vec{r}$ indeed vanishes. Using (4) and (3), we have

$$\begin{aligned}
&X \langle Y, Z \rangle - Y \langle X, Z \rangle - \langle [X, Y], Z \rangle + \langle X, \nabla_Y Z \rangle - \langle Y, \nabla_X Z \rangle \\
&\stackrel{(4)}{=} \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle - \langle \nabla_Y X, Z \rangle - \langle X, \nabla_Y Z \rangle - \langle [X, Y], Z \rangle + \langle \nabla_Y Z, X \rangle - \langle Y, \nabla_X Z \rangle \\
&= \langle \nabla_X Y - \nabla_Y X - [X, Y], Z \rangle \stackrel{(3)}{=} 0.
\end{aligned}$$

7.3. Local formula for the curvature. As with any tensor, the local components of $F = F_A$ are determined by

$$F_{ij}{}^\beta{}_\alpha e_\beta = F\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) e_\alpha.$$

Proposition 7.5. *We have*

$$\boxed{F_{ij}{}^\alpha{}_\beta = \frac{\partial A_{j\beta}^\alpha}{\partial x^i} - \frac{\partial A_{i\beta}^\alpha}{\partial x^j} + A_{i\gamma}^\alpha A_{j\beta}^\gamma - A_{j\gamma}^\alpha A_{i\beta}^\gamma.}$$

Proof. We compute

$$\nabla_{\frac{\partial}{\partial x^j}} e_\beta = A_{j\beta}^\alpha e_\alpha,$$

hence

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial x^i}} \nabla_{\frac{\partial}{\partial x^j}} e_\beta &= \nabla_{\frac{\partial}{\partial x^i}} (A_{j\beta}^\alpha e_\alpha) \\
&= \frac{\partial A_{j\beta}^\alpha}{\partial x^i} e_\alpha + A_{j\beta}^\alpha \nabla_{\frac{\partial}{\partial x^i}} e_\alpha \\
&= \frac{\partial A_{j\beta}^\alpha}{\partial x^i} e_\alpha + A_{j\beta}^\alpha A_{i\alpha}^\gamma e_\gamma.
\end{aligned}$$

Similarly,

$$\nabla_{\frac{\partial}{\partial x^j}} \nabla_{\frac{\partial}{\partial x^i}} e_\beta = \frac{\partial A_{i\beta}^\alpha}{\partial x^j} e_\alpha + A_{i\beta}^\alpha A_{j\alpha}^\gamma e_\gamma.$$

Subtracting,

$$F_{ij}{}^\gamma{}_\beta e_\gamma = \left(\frac{\partial A_{j\beta}^\gamma}{\partial x^i} - \frac{\partial A_{i\beta}^\gamma}{\partial x^j} + A_{i\alpha}^\gamma A_{j\beta}^\alpha - A_{j\alpha}^\gamma A_{i\beta}^\alpha \right) e_\gamma.$$

Therefore the curvature components are as in the boxed formula above. $\square$

Note that in matrix form, the above formula takes the nicer shape

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} + [A_i, A_j].$$

Example 7.6. Let $\{x^i\}$ be any local coordinate system on $\mathbb{S}^n$. Since $E = TM$, we can use the coordinate frame $\{\frac{\partial}{\partial x^i}\}$ for both. According to (7.1), we have

$$\begin{aligned} F\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \frac{\partial}{\partial x^\ell} &= \left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^\ell} \right\rangle \frac{\partial}{\partial x^i} - \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^\ell} \right\rangle \frac{\partial}{\partial x^j} \\ &= g_{j\ell} \frac{\partial}{\partial x^i} - g_{i\ell} \frac{\partial}{\partial x^j}. \end{aligned}$$

So we must have

$$F_{ij}{}^k{}_\ell = g_{j\ell} \delta_i^k - g_{i\ell} \delta_j^k. \quad (7.2)$$

It is customary to lower the third index, to obtain

$$F_{ijk\ell} = g_{j\ell} g_{ki} - g_{i\ell} g_{kj}. \quad (7.3)$$

Notice that this has a number of unexpected symmetries.

8. GEOMETRIC INTERPRETATION OF CURVATURE (Th 2/12)

8.1. Motivation. Last time we defined the curvature operator and saw, as if by a miracle, that it is a tensor. One motivation might have been to consider whether, as with ordinary derivatives, second covariant partial derivatives in coordinate directions commute with one-another. Although this is not true, it turns out to be almost true, in the sense that the commutator between second partials is actually a zero-th order differential operator (whereas a priori it would be a second-order differential operator).

Personally, I do not find this motivation very compelling, so I would like to show you a geometric way that you could have discovered the curvature (and even shown that it is tensorial).

8.2. The (restricted) holonomy group. Let $E \rightarrow M$ be a vector bundle with connection A . Fix a base point $x_0 \in M$. For a piecewise smooth path $\gamma : [0, 1] \rightarrow M$, write

$$P_{0,1}^{A,\gamma} : E_{\gamma(0)} \rightarrow E_{\gamma(1)}$$

for parallel transport along γ .

Definition 8.1 (Holonomy group). The **(full) holonomy group** of A at x_0 is given by

$$\mathrm{Hol}_{x_0}(A) := \{P_{0,1}^{A,\gamma} \in \mathrm{GL}(E_{x_0}) \mid \gamma \text{ is a loop based at } x_0\}.$$

The **restricted holonomy group** of A at x_0 is

$$\mathrm{Hol}_{x_0}^\circ(A) := \{P_{0,1}^{A,\gamma} \in \mathrm{GL}(E_{x_0}) \mid \gamma \text{ is a contractible loop based at } x_0\}.$$

We have the following strengthened version of Corollary 7.3.

Proposition 8.2. *If A is flat (locally gauge-equivalent to the trivial connection) then*

$$\mathrm{Hol}_{x_0}^\circ(A) = \{\mathbb{1}\}$$

for each $x_0 \in M$.

Proof. (Sketch) We imitate the proof of the Seifert-van Kampen theorem. A convincing picture was drawn in class.

Let γ be any contractible loop based at x_0 . There exists a map from the unit square that contracts γ . Cover the image of the square (which is compact) with a finite collection of open sets over which E is gauge-equivalent to the product collection. Then subdivide the square into small squares, each of whose images is contained in one of these open sets (using the Lebesgue number lemma). Since A is equivalent to the product connection on each of these small squares, the holonomy around each boundary is the identity. But the holonomy around the boundary of the full square is just a composition of the holonomies around these small squares, which is therefore equal to the identity. This shows that P^γ is the same as the holonomy of the constant path, which is the identity. $\square$

Example 8.3. A simple example from your previous homework will show that while a flat connection has trivial restricted holonomy, it can have nontrivial holonomy around a non-contractible loop. On your homework you (presumably) constructed a flat connection on the Möbius bundle, i.e., the real line bundle over S^1 which is trivial over $U_0 = (0, 2\pi)$ and $U_1 = (-\pi, \pi) \bmod 2\pi\mathbb{Z}$, and has the transition function

$$\sigma_{01} = \begin{cases} 1 & 0 < \theta < \pi \\ -1 & \pi < \theta < 2\pi. \end{cases}$$

The connection is defined by taking $A_0 = 0$ and $A_1 = 0$, so is flat. Its parallel transport is defined by keeping the absolute value of $v \in \mathbb{R}$ constant. As you well know, the holonomy around S^1 is equal to $-\mathbb{1}$.

8.3. Adapted frame on a coordinate ball. The following construction is very useful for calculations.

Let $\{x^1, \dots, x^n\}$ be a coordinate *ball* in M , which we identify as $B_r(0) \subset \mathbb{R}^n$. Define the radial vector field

$$\vec{r} = x^i \frac{\partial}{\partial x^i}.$$

Definition 8.4 (Adapted frame). A frame $\{e_\alpha\}_{\alpha=1}^m$ for $E|_{B_r(0)}$ is called an **adapted frame** (a.k.a. “radial gauge”) if

$$\nabla_{\bar{r}} e_\alpha(x) \equiv 0 \quad \text{on } B_r(0),$$

for $x \in B_r(0)$, and at the origin

$$\nabla e_\alpha(0) = 0.$$

Note that the second condition is equivalent to $A_i(0) = 0$ for $i = 1, \dots, r$.

Lemma 8.5. Fix a basis $\{v_\alpha\}_{\alpha=1}^m$ of E_0 . For $x \in B_r(0)$ define the radial path $\gamma_x(t) = tx$ and set

$$e_\alpha(x) := P_{0,1}^{A, \gamma_x}(v_\alpha).$$

Then $\{e_\alpha\}$ is an adapted frame.

Proof. We will give an equivalent construction of e_α that is manifestly smooth, then check that it is identical to the frame in the statement.

Let $\{f_\alpha\}_{\alpha=1}^r$ be any smooth frame over $B_r(0)$. By a constant change-of-frame, we may assume that $f_\alpha(0) = v_\alpha$ for $\alpha = 1, \dots, r$. Now, for each $x \in B_r(0)$, define $(\sigma_t)^\beta_\alpha(x)$, $t \in [0, 1]$, to be the solution of the matrix ODE

$$\frac{d\sigma_t}{dt} + x^i A_i(tx) \cdot \sigma_t(x) = 0, \quad \sigma_0(x) = \mathbb{1}. \quad (8.1)$$

Since $x^i A_i(tx)$ is smooth in x and t , $\sigma_t(x)$ is smooth in both entries. We define

$$e_\alpha(x) = (\sigma_1)^\beta_\alpha(x). \quad (8.2)$$

This is a smooth frame; it remains to check that it satisfies the conclusions of the lemma.

*Claim.*² Fix $a \geq 1$. For each $t \leq \frac{1}{a}$, we have $\sigma_t(ax) = \sigma_{at}(x)$.

Proof of claim. We have

$$\frac{d\sigma_{at}(x)}{dt} = a \frac{d\sigma_u(x)}{du} \Big|_{at} = a \left(-x^i A_i(atx) \sigma_{at}(x) \right) = -(ax^i) A_i(tx) \sigma_{at}(x).$$

Therefore $\sigma_{at}(x)$ solves the same ODE as $\sigma_t(ax)$. By uniqueness, they must be equal. $\diamond$

For $t > 0$, we can apply the claim with $a = \frac{1}{t}$, to get

$$\sigma_t(x) = \sigma_1(tx) \quad (8.3)$$

for each $x \in B_r$ and $0 < t \leq 1$. The identity (8.3) is also true at $t = 0$, where both sides are equal to $\mathbb{1}$. For $e_\alpha(x)$ as defined in (8.2), we now compute

$$\begin{aligned} \frac{De_\alpha(tx)}{dt} &= \frac{d(\sigma_1)^\beta_\alpha(tx)}{dt} f_\beta(tx) + (\sigma_1)^\beta_\alpha(tx) x^i A_{i\beta}^\gamma(tx) f_\gamma(tx) \\ &= \left(\frac{d(\sigma_t)^\beta_\alpha(x)}{dt} + x^i A_{i\beta}^\gamma(tx) (\sigma_t)^\beta_\alpha(x) \right) f_\beta(tx) = 0 \end{aligned} \quad (8.4)$$

by (8.1). This shows that $e_\alpha(x)$ is parallel along radial paths, so satisfies the first condition to be an adapted frame.

²Thanks to Ziji Ma for fixing the claim and its proof after class.

To see that $\nabla e_\alpha(0) = 0$, or equivalently $A_i(0) = 0$ for the connection matrices in the frame $\{e_\alpha\}$, fix $i \in \{1, \dots, n\}$ and consider the connection matrix along the x^i -axis:

$$A_i(0, \dots, s, \dots, 0).$$

Since $\vec{r} = s \frac{\partial}{\partial x^i}$ along the x^i -axis, we have

$$A_i(0, \dots, s, \dots, 0) \equiv 0$$

for $s \neq 0$. But we have just shown that the frame $\{e_\alpha\}$ is smooth, so the connection matrices are continuous. Hence also $A_i(0) = 0$. Since i was arbitrary, we're done. $\square$

8.4. Holonomy interpretation of curvature. Let 0 be the center of coordinates and fix $1 \leq i < j \leq n$. Let $\gamma_{s,t}$ be the path around the small rectangle based at 0 with side lengths s in the x^i direction and t in the x^j direction. We write

$$\gamma_{s,t} = \gamma_1 \star \gamma_2 \star \gamma_3 \star \gamma_4.$$

In the following theorem and its proof, $O(d)$ denotes a smooth function that is bounded by a polynomial of degree d in s and t .

Theorem 8.6 (Holonomy interpretation of curvature). *We have*

$$P^{A, \gamma_{s,t}} = \mathbb{1} - st F_{ij}(0) + O(3).$$

Proof. We are free to work in an adapted frame. Here we have

$$A_i(0, \dots, 0, s, 0, \dots, 0) \equiv 0,$$

where s is in the i -th place, and

$$A_j(0, \dots, 0, t, 0, \dots, 0) \equiv 0,$$

where t is in the j -th place. By Taylor's theorem, we have

$$\begin{aligned} A_i(0, \dots, s, \dots, t, \dots, 0) &= 0 + t \left(\frac{\partial A_i}{\partial x^j}(0) + O(1) \right) + O(2) \\ &= t \frac{\partial A_i}{\partial x^j}(0) + O(2) \end{aligned} \tag{8.5}$$

and

$$A_j(0, \dots, s, \dots, t, \dots, 0) = s \frac{\partial A_j}{\partial x^i}(0) + O(2). \tag{8.6}$$

Now, for the paths along the axis, we clearly have

$$P^{\gamma_1} = \mathbb{1} = P^{\gamma_4}.$$

Claim. $P^{\gamma_2} = \mathbb{1} - st \frac{\partial A_j}{\partial x^i}(0) + O(3)$.

Proof of claim. To prove the claim, we need only check that the quadratic part of the RHS,

$$P(s, t) := \mathbb{1} - st \frac{\partial A_j}{\partial x^i}(0),$$

satisfies the parallel-transport ODE up to higher-order terms. We have:

$$\begin{aligned}
 \frac{dP}{dt} &= -s \frac{\partial A_j}{\partial x^i}(0) \\
 &= \left(-s \frac{\partial A_j}{\partial x^i}(0) + O(2) \right) \left(\mathbb{1} - st \frac{\partial A_j}{\partial x^i}(0) \right) + O(2) \\
 &= -A_j(0, \dots, s, \dots, t, \dots, 0)P + O(2).
 \end{aligned} \tag{8.7}$$

Therefore $P^{\gamma_2} = P(s, t) + O(3)$, as claimed. $\diamond$

Similarly, since γ_3 goes in the reverse x^i -direction, we have

$$P^{\gamma_3} = \mathbb{1} + st \frac{\partial A_i}{\partial x^j}(0) + O(3).$$

Finally, we compute the composition

$$\begin{aligned}
 P^{\gamma_{s,t}} &= P^{\gamma_4} \circ P^{\gamma_3} \circ P^{\gamma_2} \circ P^{\gamma_1} \\
 &= \mathbb{1} \circ \left(\mathbb{1} + st \frac{\partial A_i}{\partial x^j}(0) + O(3) \right) \circ \left(\mathbb{1} - st \frac{\partial A_j}{\partial x^i}(0) + O(3) \right) \circ \mathbb{1} \\
 &= \mathbb{1} + st \left(\frac{\partial A_i}{\partial x^j} - \frac{\partial A_j}{\partial x^i} \right) + O(3).
 \end{aligned} \tag{8.8}$$

Since we are in an adapted frame, we have $[A_i(0), A_j(0)] = 0$. So the expression in parentheses is simply $F_{ji}(0) = -F_{ij}(0)$. $\square$

Example 8.7. In the picture of holonomy on the 2-sphere that we have drawn several times, we saw that if we follow a path whose initial vector is X and final vector is $-Y$, then X turns in the direction of Y . By the formula (7.1) for the curvature of the sphere, we have

$$-F(X, Y)X = Y,$$

which is consistent with the picture. (This is known as “positive curvature.”)

Example 8.8. I also drew a picture of holonomy around the x^i/x^j rectangle in an abstract vector bundle. The theorem just says that multiplication by the matrix $-F_{ij}(0) = (-F_{ij}^{\alpha\beta}(0))$ corresponds to the leading order of the holonomy map!

9. FLAT CONNECTIONS (TH 2/12 - TU 2/17)

Theorem 9.1 (Characterization of flatness). *TFAE:*

1. A is flat (see Section 6.3)
2. $F_A \equiv 0$
3. $\text{Hol}_x^\circ(A) = \{\mathbb{1}\}$ for each $x \in M$.

Proof. ((1) $\Rightarrow$ (2)) The contrapositive is Corollary 7.3.

((2) $\Rightarrow$ (1)) Assume $F_A \equiv 0$ on a coordinate ball $B_r(0)$. In an adapted frame, we have

$$0 \equiv \nabla_{x^i} \frac{\partial}{\partial x^i} e_\alpha = x^i \nabla \frac{\partial}{\partial x^i} e_\alpha,$$

so

$$x^i A_i \equiv 0.$$

Differentiating this formula, we get

$$\frac{\partial(x^i A_i)}{\partial x^j} = 0 = A_j + x^i \frac{\partial A_i}{\partial x^j}. \quad (9.1)$$

By assumption, we have

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} + [A_i, A_j] \equiv 0.$$

We contract this with x^i , to obtain

$$x^i \frac{\partial A_j}{\partial x^i} - x^i \frac{\partial A_i}{\partial x^j} + x^i [A_i, A_j] = 0.$$

Because $x^i A_i = 0$, also $x^i [A_i, A_j] = [x^i A_i, A_j] = 0$, so

$$x^i \frac{\partial A_j}{\partial x^i} - x^i \frac{\partial A_i}{\partial x^j} = 0.$$

Inserting (9.1), we get

$$x^i \frac{\partial A_j}{\partial x^i} + A_j = 0.$$

Now fix $x \in B_r(0)$ and $j \in \{1, \dots, r\}$, and define

$$\phi(t) := t A_j(tx).$$

Using the product rule and the chain rule, we have

$$\phi'(t) = A_j(tx) + t x^i \frac{\partial A_j}{\partial x^i}(tx) = 0.$$

Hence $\phi(t)$ is constant. Since $\phi(0) = 0$, we get $\phi(1) = A_j(x) = 0$. Since x and j were arbitrary, we have shown that the connection matrices all vanish in this frame, i.e. A is flat.

((1) $\Rightarrow$ (3)) Proposition 8.2.

((3) $\Rightarrow$ (2)) To prove the contrapositive, Theorem 8.6 implies that if $F_A(p) \neq 0$ at any point then the holonomy at that point is nontrivial. $\square$

Remark 9.2. During 2/17 class I also showed you a “quick and dirty” proof of (2) $\Rightarrow$ (1) using polar coordinates.

Example 9.3 (Möbius bundle again, cf. Example 8.3). Any connection on the Möbius bundle is flat, since the base space is one-dimensional (F_{ij} is anti-symmetric and the only curvature term in a 1-dimensional bundle is $F_{11} = 0$). From being flat, the restricted holonomy $\text{Hol}_x^\circ(A) = \{\mathbb{1}\}$. The full holonomy group is $\text{Hol}_x(A) = \{\pm \mathbb{1}\}$ though, as a result of the twist.

Theorem 9.4. *Let M be a connected manifold. The map*

$$\left\{ \begin{array}{l} \text{Flat connections on bundles} \\ \text{of rank } r \text{ over } M \end{array} \right\} / \text{gauge-equivalence} \longrightarrow \left\{ \begin{array}{l} \text{Representations of} \\ \pi_1(M) \text{ in } GL(r, K) \end{array} \right\} / \text{conjugacy},$$

$$A \mapsto ([\gamma] \rightarrow P^{A,\gamma}),$$

is an isomorphism.

Proof. First, we need to check that the map from flat connections to homotopy classes is well-defined as written. This is because two homotopic loops have the same holonomy under a flat connection, by the proof of Proposition 8.2.

Next, we want to check that gauge-equivalence translates to conjugacy of the corresponding representations. By a homework exercise: if $B = \sigma(A)$, then $P^{B,\gamma} = \sigma_{x_0} P^{A,\gamma} \sigma_{x_0}^{-1}$, so the corresponding representations of $\pi_1(M, x_0)$ are conjugate. This shows that gauge-equivalent connections A, B create the same parallel transport maps up to conjugation, so the map descends to a map on the quotient spaces.

To show isomorphism, we construct an inverse map. Let $\rho : \pi_1(M) \rightarrow GL(r, K)$ be a representation. Let $\tilde{M} \rightarrow M$ be the universal cover. Endow the product bundle $\tilde{M} \times K^r \rightarrow \tilde{M}$ with the product connection. Let

$$E = \tilde{M} \times K^r / \pi_1(M),$$

where the action of $\gamma \in \pi_1(M)$ is given by

$$\gamma \cdot (\tilde{x}, v) = (\gamma(x), \rho(\gamma)v).$$

Under this quotient, the product connection descends to a connection on E , which remains flat. The construction can be summed up in the following diagram:

$$\begin{array}{ccc} \tilde{M} \times K^r & \xrightarrow{\quad / \pi_1(M) \quad} & E \\ \downarrow & & \downarrow \\ \tilde{M} & \xrightarrow{\quad / \pi_1(M) \quad} & M. \end{array}$$

□

Exercise 9.5. Mull over the details of the previous proof.

Example 9.6 (The Möbius bundle, again²). $M = S^1$, so $\tilde{M} = \mathbb{R}$. Any $n \in \mathbb{Z} \cong \pi_1(M)$ acts on $\tilde{M} \cong \mathbb{R}$ as

$$x \mapsto x + n.$$

By picking the representation $\rho : \pi_1(M) \rightarrow GL(1, \mathbb{R})$ as

$$n \mapsto (-1)^n,$$

we can quotient by $\pi_1(M)$ and obtain the Möbius bundle with a flat connection from the product connection. (HW exercise.) A picture was drawn in class.

10. COMPLEX LINE BUNDLES AND THE FIRST CHERN CLASS (Tu 2/17)

We will now discuss how the above theory plays out in the case of complex line bundles. Assume that $K = \mathbb{C}$ and $r = 1$.

10.1. Topology of complex line bundles. First of all, we should compare with the case of real line bundles. The Möbius bundle gave us an example of a flat but nontrivial real line bundle. One thing to notice is that if you tensor the Möbius bundle with $\mathbb{C}$, to obtain a complex line bundle, this will be trivial.

Claim. All complex line bundles on S^1 are topologically trivial.

Proof of claim. Let $L \rightarrow S^1$ be a complex line bundle. To show that L is trivial, one just has to write down a nonvanishing section. Take any connection on L . (Since S^1 is 1-dimensional, this has no curvature and the connection is flat.) Parallel transport around S^1 gives a nonvanishing section that may not match up at the fiber you started from. But since $\mathbb{C} \setminus \{0\}$ is connected, you can just fudge this section a little bit at the end to make it close up. $\diamond$

More generally, one has the following fact:

Theorem 10.1. *Assume that $H^2(M, \mathbb{Z})$ is torsion-free.³ Then any $\mathbb{C}$ line bundle on M that carries a flat connection is topologically trivial.*

We will not prove this, but you may want to try and think about it as an exercise.

Example 10.2. Recall that

$$\mathcal{O}(-1) = \{(v, \ell) \in \mathbb{C}^2 \times \mathbb{CP}^1 \mid v \in \ell\},$$

$$\mathcal{O}(1) = \mathcal{O}(-1)^*,$$

and

$$\mathcal{O}(d) = \begin{cases} \mathcal{O}(1)^{\otimes d} & d \geq 0 \\ \mathcal{O}(-1)^{\otimes |d|} & d < 0 \end{cases}$$

In my 761 notes, Proposition 34.3, we proved that $\mathcal{O}(m) \cong \mathcal{O}(n) \iff m = n$. In particular, $\mathcal{O}(n)$ is trivial only if $n = 0$. We will see another proof of this fact below.

Example 10.3. The torsion-freeness assumption is necessary. In fact, one can even make an example of a nontrivial, flat, real line bundle that remains nontrivial after complexification: namely, the real tautological bundle $\mathcal{O}(-1) \rightarrow \mathbb{RP}^2$. This has a flat connection (as does any real line bundle), which remains flat after tensoring with $\mathbb{C}$. But it is an interesting exercise to show explicitly that it remains nontrivial.

³Omitted during class!

10.2. Curvature of complex line bundles and the first Chern class. We will now discuss the curvature of a connection on a complex line bundle in general. Let A be a connection on L .

Observation: The endomorphism bundle of a line bundle is canonically trivial:

$$\begin{aligned} \text{End}(L) &= L \otimes L^* \cong \underline{K} \\ (v, \alpha) &\mapsto \alpha(v). \end{aligned} \tag{10.1}$$

In particular, *the curvature F_A of any connection on L is just a $\mathbb{C}$ -valued 2-form!* I.e.

$$F_A \in \Omega_{\mathbb{C}}^2(M).$$

Now, fix an open set U over which L is trivialized by a frame $\{e_1\}$ with a single element. The connection matrices $A_i \in C_{\mathbb{C}}^{\infty}(U)$ are just smooth $\mathbb{C}$ -valued scalar functions over U . We can define the “**connection form**”, which depends on the frame $\{e_1\}$, as

$$A \stackrel{\text{loc}}{=} A_i dx^i \in \Omega_{\mathbb{C}}^1(U).$$

Here we abuse notation and denote the connection form by the same letter, A , as the connection itself. In coordinates, the curvature components are

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} + [A_i, A_j].$$

But 1×1 complex matrices (scalars) commute, so $[A_i, A_j] = 0$. The remaining terms are just the exterior derivative:

$$\begin{aligned} F_A &\stackrel{\text{loc}}{=} F_{ij} dx^i \otimes dx^j \\ &= \frac{1}{2} \left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} \right) dx^i \wedge dx^j, \\ &= \frac{\partial A_j}{\partial x^i} dx^i \wedge dx^j, \\ &= dA. \end{aligned}$$

Proposition 10.4. $F_A \in \Omega_{\mathbb{C}}^2$ is closed.

Proof. The above calculation showed that F_A is locally exact; it is therefore closed. In other words,

$$dF_A \stackrel{\text{loc}}{=} d(dA) = 0.$$

□

Definition/Lemma 10.5. For a $\mathbb{C}$ -line-bundle $L \rightarrow M$, the **first Chern class**⁴ is given by

$$c_1(L) = \left[\frac{i}{2\pi} F_A \right] \in H^2(M, \mathbb{C})$$

where A is any connection on L .

⁴Strictly speaking, this is *not* really the first Chern class but only the *image* of the first Chern class under the DeRham map $H^2(M, \mathbb{Z}) \rightarrow H_{dR}^2(M, \mathbb{R})$. When there is torsion, these may actually be different. For instance, in the case of $\mathcal{O}(-1) \otimes \mathbb{C} \rightarrow \mathbb{RP}^2$ discussed above, the bundle is flat, so c_1 as above vanishes by definition, but the genuine (integral) c_1 actually does not vanish (since this is a nontrivial complex line bundle!). I thank Aleksander Doan for pointing this out to me.

Proof. We showed above that F_A is closed, so defines a cohomology class for each connection A . We need to check that the cohomology class is independent of A .

Let A, B be any two connections on L . Recall that $B - A \in \Omega^1(\text{End}L) = \Omega^1_{\mathbb{C}}(M)$ is a globally well-defined endomorphism-valued 1-form, by Proposition 3.1b. So

$$F_B - F_A = dB - dA = d(B - A)$$

This makes $F_B - F_A$ exact, so $[F_A] = [F_B]$. □

Exercise 10.6. On homework you will define a connection on $\mathcal{O}(-1)$ with the connection matrices

$$A^{(0)} = \frac{dz \bar{z}}{1 + |z|^2}, \quad A^{(1)} = \frac{dw \bar{w}}{1 + |w|^2}$$

in the standard charts. You can calculate its curvature and integrate, to prove directly that $c_1(\mathcal{O}(-1)) = -1$. Hence, $\mathcal{O}(-1)$ does *not* admit a flat connection. This gives a proof of a special case of Theorem 10.1.

10.3. Restricted holonomy of a complex line bundle. Last, I want to mention the fact that the holonomy interpretation of curvature (Theorem 9.4) can be made into a stronger statement in the case of complex line bundles. We plan to use the following result later in the course to deduce the so-called “Local Gauss-Bonnet Theorem.”

Theorem 10.7. *Let L be a complex line bundle with a connection A . Suppose that $\gamma = \partial D$ is an oriented curve that is a homological boundary. Then*

$$P^{A, \gamma} = e^{-\int_D F_A} \cdot \mathbb{1}.$$

Here e^z is the complex exponential function.

Proof. By subdividing, we can assume without loss of generality that the image of γ is contained in a coordinate chart over which L is trivial. Write $A = A_i dx^i$ for the connection form, as above. Given $s_0 \in E_{\gamma(0)}$, we solve the parallel transport ODE for a section of the form

$$s = e^{\varphi} s_0, \quad \varphi(0) = 0.$$

The ODE reads

$$\frac{Ds}{dt} = 0 = \frac{ds}{dt} + \frac{d\gamma^i}{dt} A_i \cdot s = \left(\frac{d\varphi}{dt} + \frac{d\gamma^i}{dt} A_i \right) s,$$

in other words

$$\frac{d\varphi}{dt} = -\frac{d\gamma^i}{dt} A_i.$$

We therefore have

$$\varphi(1) - \varphi(0) = - \int \frac{d\gamma^i}{dt} A_i dt = - \int_{\gamma} A.$$

Applying Stokes’s Theorem, we have

$$\varphi(1) - \varphi(0) = - \int_D dA = - \int_D F_A.$$

This gives the desired formula. □

11. INDUCED CONNECTIONS (TH 2/19)

To generalize some of the observations from the previous section to rank $r > 1$, we must introduce a bit of formalism that we did not need until now. We return to considering a vector bundle E of general rank r over $K = \mathbb{R}$ or $\mathbb{C}$. As with a Riemannian metric, a connection A naturally induces a connection on the dual bundle E^* and on tensor bundles.

Definition/Lemma 11.1. *A connection A on $E \rightarrow M$ induces a connection on E^* by the rule*

$$X(t(s)) = (\nabla_X t)(s) + t(\nabla_X s), \quad (11.1)$$

where $t \in \Gamma(E^*)$, $s \in \Gamma(E)$, $X \in \mathfrak{X}(M)$.

Proof. We first show that if such a connection exists on E^* , it is unique. Connections are local, so we work with $\{e_\alpha\}$ and $\{e^\alpha\}$ as local frames for E and E^* respectively. Take sections $s = s^\gamma e_\gamma$ on E and $t = t_\alpha e^\alpha$ on E^* . We have

$$t_\alpha = t(e_\alpha).$$

Differentiating this expression, we get

$$\begin{aligned} \frac{\partial t_\alpha}{\partial x^i} &= (\nabla_{\frac{\partial}{\partial x^i}} t)(e_\alpha) + t(\nabla_{\frac{\partial}{\partial x^i}} e_\alpha) \\ &= (\nabla_{\frac{\partial}{\partial x^i}} t)(e_\alpha) + t(A_{i\alpha}^\beta e_\beta) \\ &= (\nabla_{\frac{\partial}{\partial x^i}} t)_\alpha + A_{i\alpha}^\beta t_\beta. \end{aligned}$$

Rearranging, we obtain

$$\left(\nabla_{\frac{\partial}{\partial x^i}} t \right)_\alpha = \frac{\partial t_\alpha}{\partial x^i} - A_{i\alpha}^\beta t_\beta.$$

This shows that the connection is uniquely determined by (11.1). Indeed, we must have

$$\nabla_{\frac{\partial}{\partial x^i}} t = (\nabla_i t_\alpha) e^\alpha,$$

where

$$\nabla_i t_\alpha = \frac{\partial t_\alpha}{\partial x^i} - A_{i\alpha}^\beta t_\beta.$$

Now: on an open neighborhood with a coordinate frame, suppose that we define ∇ on E^* by the above formula. We must now check that this actually satisfies the Leibniz rule (11.1) (locally):

$$\begin{aligned} LHS &= \frac{\partial(t_\alpha s^\alpha)}{\partial x^i} \\ RHS &= (\nabla_{\frac{\partial}{\partial x^i}} t)(s) + t(\nabla_{\frac{\partial}{\partial x^i}} s) \\ &= \frac{\partial t_\alpha}{\partial x^i} s^\alpha - A_{i\alpha}^\beta t_\beta s^\alpha + t_\alpha \frac{\partial s^\alpha}{\partial x^i} + t_\alpha A_{i\beta}^\alpha s^\beta \\ &= \frac{\partial t_\alpha}{\partial x^i} s^\alpha + t_\alpha \frac{\partial s^\alpha}{\partial x^i} = LHS. \end{aligned}$$

So such a locally defined connection does indeed satisfy the Leibniz rule (11.1).

But since the local connection so defined satisfies (11.1), it is unique, by what we showed at the beginning. It is therefore globally well-defined! $\square$

Remark 11.2. To explain the end of the proof a bit more, think of taking a system of trivializations of $E \rightarrow M$ and defining connections on each one by the above formula. To check that these give a globally well-defined connection, one just has to check that they agree on overlaps. But this is true by the uniqueness property.

Remark 11.3. The rule for the dual connection could be written “ $A_i^* = -A_i^\top$.” This is not too different from the transition functions on E^* with respect to dual frames, where for $\sigma = \sigma_{ab}$ for E , then $\sigma^* = (\sigma^\top)^{-1}$ for E^* (exercise).

Definition/Lemma 11.4. *Let E and F be bundles with connections A and B , respectively. There exists a unique induced connection on $E \otimes F$ determined by the Leibniz rule*

$$\nabla_X(s \otimes t) = (\nabla_X s) \otimes t + s \otimes (\nabla_X t).$$

Proof. First check that, if such a connection exists on $E \otimes F$, then it is unique. Let $\{e_\alpha\}$ and $\{f_\beta\}$ be local frames for E and F . Then

$$\begin{aligned} \nabla_{\frac{\partial}{\partial x^i}}(s \otimes t) &= \nabla_{\frac{\partial}{\partial x^i}}(s^\alpha e_\alpha \otimes t^\beta f_\beta) \\ &= \left(\frac{\partial s^\alpha}{\partial x^i} + A_{i\gamma}^\alpha s^\gamma\right) e_\alpha \otimes t^\beta f_\beta + s^\alpha e_\alpha \otimes \left(\frac{\partial t^\beta}{\partial x^i} + B_{i\gamma}^\beta t^\gamma\right) f_\beta \\ &= \left(\frac{\partial(s^\alpha t^\beta)}{\partial x^i} + A_{i\gamma}^\alpha s^\gamma t^\beta + s^\alpha B_{i\gamma}^\beta t^\gamma\right) e_\alpha \otimes f_\beta \end{aligned}$$

So we must define

$$\nabla_{\frac{\partial}{\partial x^i}}(\omega^{\alpha\beta} e_\alpha \otimes e_\beta) = \nabla_i \omega^{\alpha\beta} e_\alpha \otimes e_\beta,$$

where

$$\nabla_i \omega^{\alpha\beta} = \frac{\partial \omega^{\alpha\beta}}{\partial x^i} + A_{i\gamma}^\alpha \omega^{\gamma\beta} + B_{i\gamma}^\beta \omega^{\alpha\gamma}$$

To complete the proof, as before, we must check that the connection so defined (locally) satisfies the Leibniz rule — this is left as a homework exercise. Since it does, it is locally unique, hence globally well-defined. $\square$

Lemma 11.5. *Let $T \in \Gamma(E \otimes \cdots \otimes E \otimes E^* \otimes \cdots \otimes E^*)$ be a section, written locally as*

$$T \stackrel{loc}{=} T^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_\ell} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \otimes e^{\beta_1} \otimes \cdots \otimes e^{\beta_\ell}.$$

The induced connection acts by

$$\nabla_{\frac{\partial}{\partial x^i}} T = (\nabla_i T^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_\ell}) e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \otimes e^{\beta_1} \otimes \cdots \otimes e^{\beta_\ell}$$

where

$$\begin{aligned} \nabla_i T^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_\ell} &:= \frac{\partial T^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_\ell}}{\partial x^i} + A_{i\alpha}^{\alpha_1} T^{\alpha_2 \dots \alpha_k}_{\beta_1 \dots \beta_\ell} + \cdots + A_{i\alpha}^{\alpha_k} T^{\alpha_1 \dots \alpha_{k-1} \alpha}_{\beta_1 \dots \beta_\ell} \\ &\quad - A_{i\beta_1}^{\beta_1} T^{\alpha_1 \dots \alpha_k}_{\beta_2 \dots \beta_\ell} - \cdots - A_{i\beta_\ell}^{\beta_\ell} T^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_{\ell-1} \beta} \end{aligned}$$

Proof. This follows by combining the above constructions of induced connections on dual and tensor bundles. $\square$

Example 11.6. $\sigma \in \Gamma(\text{End} E) = \Gamma(E \otimes E^*)$.

$$\sigma \stackrel{\text{loc}}{=} \sigma^\alpha{}_\beta e_\alpha \otimes e^\beta$$

$$\nabla_i \sigma^\alpha{}_\beta = \frac{\partial \sigma^\alpha{}_\beta}{\partial x^i} + A_{i\gamma}^\alpha \sigma^\gamma{}_\beta - \sigma^\alpha{}_\gamma A_{i\beta}^\gamma$$

In matrix form,

$$\nabla_{\frac{\partial}{\partial x^i}} \sigma = \frac{\partial \sigma}{\partial x^i} + [A_i, \sigma]$$

Proposition 11.7. $X(\text{Tr } \sigma) = \text{Tr}(\nabla_X \sigma)$.

Proof. We use matrix notation in a local frame. The RHS is

$$\text{Tr}(X(\sigma) + X^i[A_i, \sigma]) = \text{Tr } X(\sigma)$$

by the cyclic property of the trace. But we clearly have

$$\text{Tr } X(\sigma) = X(\text{Tr } \sigma).$$

To spell this out:

$$X(\text{Tr } \sigma) = X\left(\sum_\alpha \sigma^\alpha{}_\alpha\right) = \sum_\alpha X(\sigma^\alpha{}_\alpha) = \text{Tr}(X(\sigma)).$$

□

12. THE DIFFERENTIAL BIANCHI IDENTITY (TH 2/19)

We can now generalize some of the results shown above for line bundles to vector bundles of general rank.

Theorem 12.1 (Differential Bianchi Identity, a.k.a. “2nd Bianchi Identity”). *For a connection A on E , let $F = F_A$, and write $\nabla = \nabla^A$ for the induced connection on $\text{End}(E)$. Then, for $X, Y, Z \in \mathfrak{X}(M)$, we have*

$$\begin{aligned} & \nabla_X F(Y, Z) + \nabla_Y F(Z, X) + \nabla_Z F(X, Y) \\ & + F(X, [Y, Z]) + F(Y, [Z, X]) + F(Z, [X, Y]) = 0. \end{aligned}$$

Proof. Denote the LHS of the above expression by $B(X, Y, Z)$.

Claim. B is tensorial.

Proof of claim. We calculate

$$\begin{aligned} B(fX, Y, Z) &= f \nabla_X F(Y, Z) + \nabla_Y (f F(Z, X)) + \nabla_Z (f F(X, Y)) \\ &\quad + f F(X, [Y, Z]) + F(Y, [Z, fX]) + F(Z, [fX, Y]) \\ &= f B(X, Y, Z) + Y(f) F(Z, X) + Z(f) F(X, Y) + Z(f) F(Y, X) - Y(f) F(Z, X) \\ &= f B(X, Y, Z), \end{aligned}$$

where we used $F(Y, X) = -F(X, Y)$. Since B is invariant under cyclic permutations of the arguments, the same holds for $Y \rightarrow fY$ and $Z \rightarrow fZ$. $\diamond$

Since B is tensorial, it suffices to check that $B(X, Y, Z) = 0$ for $X = \frac{\partial}{\partial x^i}$, $Y = \frac{\partial}{\partial x^j}$, and $Z = \frac{\partial}{\partial x^k}$. In fact, it suffices to check at a single point $x_0 \in M$. We work in an adapted frame where $x_0 = 0$. We have

$$B\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right)(0) = \frac{\partial}{\partial x^i} \left(\frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k} + [A_j, A_k] \right)(0) + [A_i(0), F_{jk}] + \text{cyclic}.$$

Now, the term $[A_i(0), F_{jk}]$ vanishes because $A_i(0) = 0$. We also have

$$\frac{\partial}{\partial x^i} [A_j, A_k](0) = \left[\frac{\partial A_j}{\partial x^i}, A_k(0) \right] + \left[A_j(0), \frac{\partial A_k}{\partial x^i} \right] = 0.$$

So we are left simply with

$$\begin{aligned} B\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right)(0) &= \frac{\partial^2 A_k}{\partial x^i \partial x^j} - \frac{\partial^2 A_j}{\partial x^i \partial x^k} \\ &\quad + \frac{\partial^2 A_i}{\partial x^j \partial x^k} - \frac{\partial^2 A_k}{\partial x^j \partial x^i} \\ &\quad + \frac{\partial^2 A_j}{\partial x^k \partial x^i} - \frac{\partial^2 A_i}{\partial x^k \partial x^j} = 0, \end{aligned}$$

since the terms cancel in pairs. □

Remark 12.2. In slightly fancier language (see 865 notes), the differential Bianchi identity just states that F_A is *covariantly closed*. In that sense it is a strict generalization of Proposition 10.4.

Definition/Lemma 12.3. *We may define*

$$c_1(E) = \left[\frac{i}{2\pi} \text{Tr } F_A \right],$$

where A is any connection on E .

Proof. First, we need to check that $\text{Tr } F_A$ is closed. Take the trace of the differential Bianchi identity:

$$\begin{aligned} 0 &= \text{Tr}(B(X, Y, Z)) = \text{Tr } \nabla_X F(Y, Z) + \text{Tr } F(X, [Y, Z]) + \text{cyclic} \\ &= X(\text{Tr } F(Y, Z)) + \text{Tr } F(X, [Y, Z]) + \text{cyclic}. \end{aligned}$$

We have used Proposition 11.7. But we can recognize the expression on the RHS as the invariant form of the exterior derivative—see 761 notes, Theorem 38.7. We obtain

$$(d\text{Tr } F)(X, Y, Z) = 0.$$

Since X, Y, Z were arbitrary, we're done.

We also need to check that if A, B are both connections on E , then we get the same cohomology class. This goes as in the proof of Definition/Lemma 10.5, where one also uses $\text{Tr}[A_i, A_j] = 0$. □

Remark 12.4. These constructions belong to the subject of **Chern-Weil Theory**, see e.g. the 865 notes.

Part 4. The Levi–Civita connection

Recall that our goal has always been to develop a notion of curvature for a Riemannian manifold. We now know that curvature is actually an object belonging to a *connection*. To define the curvature of a Riemannian *metric*, our approach will be to attach a canonical choice of connection to any Riemannian metric. Later, we will declare the curvature of the metric to simply be the curvature of this canonical connection.

Now, since every vector bundle has a plethora of connections to choose from, we have to figure out what requirements to throw in to get a *unique* connection attached to a metric on TM .

13. METRIC COMPATIBILITY (Tu 2/24)

While we are ultimately interested in Riemannian metrics (i.e. metrics on TM), it's best if we continue working on an abstract vector bundle $E \rightarrow M$ for today.

13.1. Metrics on general vector bundles. ($K = \mathbb{R}$) A **metric** h on a real vector bundle $E \rightarrow M$ is a smoothly varying inner product on the fibers E_x , i.e. a section $h \in \Gamma(E^* \otimes E^*)$ such that $h_x(\cdot, \cdot)$ is symmetric and positive-definite for each $x \in M$. In local coordinates,

$$h_{\alpha\beta} := h(e_\alpha, e_\beta)$$

is a symmetric, positive definite matrix-valued function.

Note: If $E = TM$, this is exactly a Riemannian metric.

Remark 13.1. • Every vector bundle carries a metric (using a partition of unity).

- Given any frame $\{e_\alpha\}$, Gram–Schmidt takes this to an orthonormal frame $\{\tilde{e}_\alpha\}$.
- The transition functions in an orthonormal frame are orthogonal-matrix-valued functions ($\sigma(x) \in O(n)$).

Proof. If $\{e_\alpha\}$ and $\{\tilde{e}_\beta\}$ are orthonormal frames, then

$$\begin{aligned} \delta_{\alpha\beta} &= \tilde{h}_{\alpha\beta} = h(\sigma(e_\alpha), \sigma(e_\beta)) \\ &= h(\sigma^\gamma_\alpha e_\gamma, \sigma^\eta_\beta e_\eta) \\ &= \sigma^\gamma_\alpha \sigma^\eta_\beta h_{\gamma\eta}. \end{aligned}$$

But also $h_{\gamma\eta} = \delta_{\gamma\eta}$, so the last expression is just

$$\delta_{\alpha\beta} = (\sigma^\top \sigma)^\alpha_\beta.$$

□

Remark 13.2. (E, h) still has no local structure! We can just convert any frame to an orthonormal one and achieve $h_{\alpha\beta} = \delta_{\alpha\beta}$.

Definition 13.3. Let $E \rightarrow M$ be a complex vector bundle. We define the *conjugate bundle* $\overline{E}$ to be the same underlying real bundle E , but with the complex scalar multiplication twisted by conjugation:

$$\lambda \cdot_{\overline{E}} s = \overline{\lambda} \cdot_E s$$

for $\lambda \in \mathbb{C}$ and $s \in \overline{E}_x = E_x$.

Remark 13.4. If $\{e_\alpha\}$ is a local frame for E , this it is also one for $\overline{E}$. The transition functions for σ become $\overline{\sigma}$ for $\overline{E}$:

$$\begin{aligned} e_\alpha^a &= \sigma^\beta_\alpha \cdot_E e_\beta^b \\ &= \overline{\sigma}^\beta_\alpha \cdot_{\overline{E}} e_\beta^b \end{aligned}$$

Definition 13.5 ($K = \mathbb{C}$). A **Hermitian metric** h is a smoothly varying Hermitian inner product on E_x , i.e. a tensor

$$h \in \Gamma(E^* \otimes_{\mathbb{C}} \overline{E}^*)$$

such that $h_x(\cdot, \cdot)$ is a Hermitian inner product on E_x for each x .

Remark 13.6. • Every complex bundle carries a Hermitian metric.

- Given local frame $\{e_\alpha\}$, applying Gram–Schmidt (fibrewise) makes it a local orthonormal frame (canonically).
- In an orthonormal frame, transition functions are unitary-matrix-valued functions.

Lemma 13.7. *Every oriented, real, rank-2 vector bundle is isomorphic to the underlying real bundle of a (unique) rank-1 complex vector bundle. In other words, the isomorphism classes of each are in natural bijection.*

Proof. (Sketch) Given a complex frame, we can take that frame and i times that frame to get a frame on the 2-dimensional vector bundle. This map will be oriented because the transition functions are multiplications by nonzero complex scalars $z \in \mathbb{C}^*$, which, viewed as real-linear maps on $\mathbb{R}^2$, have positive determinant $|z|^2 > 0$ and thus preserve orientation.

Going the other way, if we start with an oriented, rank-2, real bundle with frame $\{e_1, e_2\}$, i.e. all transition functions have $\det \sigma > 0$. Choose at random a metric on this bundle. (Since all metrics are equivalent by local scaling, this does not affect the outcome.) After doing Gram–Schmidt, the transition functions belong to

$$\mathrm{SO}(2) \cong \mathrm{U}(1).$$

So this can now be viewed as a complex line bundle. More explicitly, the complex structure is simply counterclockwise rotation by $\pi/2$ in the real bundle (which is oriented and has a metric, so this makes sense). Locally, given an oriented frame $\{e_1, e_2\}$ for a fiber, define the bundle endomorphism J by $J(e_1) = e_2$ and $J(e_2) = -e_1$. This J satisfies $J^2 = -\mathrm{Id}$, as $J^2(e_1) = J(e_2) = -e_1$ and $J^2(e_2) = J(-e_1) = -e_2$. Complex scalar multiplication is then defined by $(a + bi) \cdot v = av + bJ(v)$ for $a, b \in \mathbb{R}$ and v in the bundle. $\square$

Exercise 13.8. Mull over the details of the foregoing proof.

13.2. Metric-compatible connections. A connection A on E is said to be **compatible** with a metric h on E if for all vector fields X and all sections $s, u \in \Gamma(E)$,

$$\boxed{X\langle s, u \rangle = \langle \nabla_X s, u \rangle + \langle s, \nabla_X u \rangle}$$

where $\nabla = \nabla^A$ and $\langle \cdot, \cdot \rangle = h(\cdot, \cdot)$.

Proposition 13.9. *Let $K = \mathbb{R}$. The following are equivalent:*

1. ∇ is metric compatible.
2. (“**Ricci’s Lemma**”) *The metric $h \in \Gamma(E^* \otimes E^*)$ is covariantly constant with respect to the induced connection on $E^* \otimes E^*$, i.e. $\nabla h = 0$.*
3. *With respect to an arbitrary local frame $\{e_\alpha\}$,*

$$\boxed{\frac{\partial h_{\alpha\beta}}{\partial x^i} = A_{i\alpha}^\gamma h_{\gamma\beta} + A_{i\beta}^\gamma h_{\alpha\gamma}}$$

4. *With respect to an orthonormal frame, the connection matrix $A_i = (A_{i\alpha}^\gamma)$ is skew-symmetric:*

$$A_i^\top = -A_i, \forall i$$

5. *For all sections s, u along a curve γ ,*

$$\frac{d}{dt} \langle s, u \rangle = \left\langle \frac{Ds}{dt}, u \right\rangle + \left\langle s, \frac{Du}{dt} \right\rangle$$

6. *Parallel transport*

$$P_a^{A,\gamma} : (E_{\gamma(a)}, h_{\gamma(a)}) \rightarrow (E_{\gamma(b)}, h_{\gamma(b)})$$

is an isometry for any curve γ and a, b in its domain.

7. *There exist orthonormal adapted frames (see Section 8.2).*

Proof. ((1) $\Rightarrow$ (3))

$$\begin{aligned} \frac{\partial h_{\alpha\beta}}{\partial x^i} &= \frac{\partial}{\partial x^i} \langle e_\alpha, e_\beta \rangle \\ &= \langle \nabla_{\frac{\partial}{\partial x^i}} e_\alpha, e_\beta \rangle + \langle e_\alpha, \nabla_{\frac{\partial}{\partial x^i}} e_\beta \rangle \\ &= \langle A_{i\alpha}^\gamma e_\gamma, e_\beta \rangle + \langle e_\alpha, A_{i\beta}^\gamma e_\gamma \rangle \\ &= A_{i\alpha}^\gamma h_{\gamma\beta} + A_{i\beta}^\gamma h_{\alpha\gamma} \end{aligned}$$

((2) $\Leftrightarrow$ (3)) Assume (3) holds in some local frame $\{e_\alpha\}$, so that for all i, α, β ,

$$\frac{\partial h_{\alpha\beta}}{\partial x^i} - A_{i\alpha}^\gamma h_{\gamma\beta} - A_{i\beta}^\gamma h_{\alpha\gamma} = 0$$

Consider the induced connection on $E^* \otimes E^*$, writing $h_{\alpha\beta} = h(e_\alpha, e_\beta)$ and $\nabla_{\frac{\partial}{\partial x^i}} e_\alpha = A_{i\alpha}^\gamma e_\gamma$,

$$(\nabla_{\frac{\partial}{\partial x^i}} h)_{\alpha\beta} = \frac{\partial}{\partial x^i} (h_{\alpha\beta}) - A_{i\alpha}^\gamma h_{\gamma\beta} - A_{i\beta}^\gamma h_{\alpha\gamma},$$

Hence the assumption (3) says precisely that

$$(\nabla_{\frac{\partial}{\partial x^i}} h)_{\alpha\beta} = 0$$

for all i, α, β , i.e. $\nabla_{\frac{\partial}{\partial x^i}} h = 0$ for all i . Therefore $\nabla h = 0$, so h is covariantly constant.

((3) $\Rightarrow$ (1)) Left as a homework exercise.

((3) $\iff$ (4)) Choose a local orthonormal frame $\{e_\alpha\}$, so that $h_{\alpha\beta} = h(e_\alpha, e_\beta) = \delta_{\alpha\beta}$. Then for each i we have

$$\frac{\partial \delta_{\alpha\beta}}{\partial x^i} = 0.$$

Applying (3),

$$\frac{\partial \delta_{\alpha\beta}}{\partial x^i} = 0 = A_{i\alpha}^\beta + A_{i\beta}^\alpha.$$

Therefore $A_{i\alpha}^\beta = -A_{i\beta}^\alpha$ for all α, β , i.e. $A_i^\top = -A_i$.

The remaining items are pretty obvious. $\square$

So when we are differentiating things involving a metric, we can just slide connections through the metric.

Remark 13.10. • When $K = \mathbb{C}$ with a Hermitian metric, (4) becomes $A_i^\dagger = -A_i$ (“skew-Hermitian”).

- Every vector bundle (real or complex) equipped with a metric (Riemannian or Hermitian) carries (many) compatible connections! So we have not solved our uniqueness problem.

Example 13.11. We saw in Lemma 7.4.3 that the induced connection on S^n is metric-compatible. We will see below (or you can prove it now as an exercise) that the induced connection on any submanifold of $\mathbb{R}^n$ is metric-compatible.

Remark 13.12 (From class on Th 2/26). There is an alternate definition of the covariant derivative of a tensor: for $\omega \in \Gamma(\overbrace{E^* \otimes \cdots \otimes E^*}^n)$, simply define

$$\begin{aligned} (\nabla_X \omega)(s_1, \dots, s_n) &:= X(\omega(s_1, \dots, s_n)) - \omega(\nabla_X s_1, \dots, s_n) \\ &\quad - \omega(s_1, \nabla_X s_2, \dots, s_n) \\ &\quad \vdots \\ &\quad - \omega(s_1, \dots, \nabla_X s_n). \end{aligned} \tag{13.1}$$

Let’s check that this agrees with Definitions 11.1-11.4 above. For $n = 1$, (13.1) is exactly the Leibniz rule defining the induced connection on E^* . For $n = 2$, to show that this connection is the same as the one above, by uniqueness, it suffices to check the Leibniz rule on $E^* \otimes E^*$. By $\mathbb{R}$ -linearity, we can assume without loss of generality that $\omega = t_1 \otimes t_2$ is decomposable. We compute

$$\begin{aligned} (\nabla_X(t_1 \otimes t_2))(s_1, s_2) &\stackrel{(13.1)}{=} X(t_1(s_1)t_2(s_2)) - (t_1 \otimes t_2)(\nabla_X s_1, s_2) - (t_1 \otimes t_2)(s_1, \nabla_X s_2) \\ &= X(t_1(s_1))t_2(s_2) + t_1(s_1)X(t_2(s_2)) - t_1(\nabla_X s_1)t_2(s_2) - t_1(s_1)t_2(\nabla_X s_2) \\ &= (X(t_1(s_1)) - t_1(\nabla_X s_1))t_2(s_2) + t_1(s_1)(X(t_2(s_2)) - t_2(\nabla_X s_2)) \\ &= (\nabla_X t_1)(s_1)t_2(s_2) + t_1(s_1)(\nabla_X t_2)(s_2) \\ &= (\nabla_X t_1 \otimes t_2 + t_1 \otimes \nabla_X t_2)(s_1, s_2). \end{aligned}$$

So the connection defined by (13.1) does indeed satisfy the Leibniz rule on $E^* \otimes E^*$, and therefore agrees with the one defined above.

Notice that with this definition, in the above Proposition, Ricci's lemma $(1) \iff (2)$ is true by definition, whereas with the earlier definition, $(2) \iff (3)$ is true by definition.

We can now give the following simple application of the above discussion.

Proposition 1. *Let $E \rightarrow M$ be a complex vector bundle and let A be any connection on E . Then*

$$c_1(E) = \frac{i}{2\pi} [\text{Tr } F_A]$$

is a real cohomology class, i.e.

$$c_1(E) \in H^2(M, \mathbb{R}).$$

Proof. Take a Hermitian metric, h , on E , and take A to be compatible with h . Use a local orthonormal frame, so the connection matrices are skew-Hermitian

$$A_i^\dagger = -A_i$$

The trace is indifferent to transpose, i.e.

$$\text{Tr } A_i^\dagger = \overline{\text{Tr } A_i} = -\text{Tr } A_i$$

$$\Rightarrow \text{Tr } A_i \in i\mathbb{R} \subset \mathbb{C}$$

is purely imaginary.

The same is true for any skew-Hermitian endomorphism (in the local frame). The curvature components are

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} + [A_i, A_j]$$

But note that

$$\{\text{skew-Hermitian}\} = \mathfrak{u}(n) = \text{Lie}(U(n))$$

is a Lie algebra, it is closed under brackets. In particular, if each A_i is $\mathfrak{u}(n)$ -valued (skew-Hermitian), then $[A_i, A_j]$ is also $\mathfrak{u}(n)$ -valued, hence skew-Hermitian.

So $F_{ij} \in \mathfrak{u}(n)$ and $\text{Tr } F_{ij} \in i\mathbb{R}$, making c_1 real. □

Example 13.13. If $\text{rk}_{\mathbb{C}} E = 1$,

$$\mathfrak{u}(1) = \text{Lie}(U(1)) = i\mathbb{R}.$$

In case I forgot to mention, our notation comes from electromagnetism: up to a factor of i , A is the (magnetic) vector potential and F_A is the electromagnetic field-strength tensor.

14. TORSION-FREENESS (TU 2/24-TH 2/26)

Recall the goal: attach a canonical connection to a Riemannian metric g on TM . We need to involve the fact that TM knows about the smooth structure of M .

Since we will now consider connections A on the tangent bundle TM , it is customary to use *the same* frame for both the differentiating vector field and the vector field being differentiated. In fact, we shall almost always use a coordinate frame $\{\frac{\partial}{\partial x^i}\}$ for both. In other words, we no longer need Greek indices but only Latin indices, and will denote the local components of A by A_{ij}^k .

Definition 14.1. For a connection A on TM , the **torsion** of A is the function

$$\begin{aligned}\tau_A : \mathfrak{X}(M) \times \mathfrak{X}(M) &\rightarrow \mathfrak{X}(M) \\ \tau_A(X, Y) &= \nabla_X^A Y - \nabla_Y^A X - [X, Y].\end{aligned}\tag{14.1}$$

A is called **torsion-free** if $\tau_A \equiv 0$.

Lemma 14.2. τ_A is tensorial in X and Y , so defines an element of $\Omega^2(TM)$.

Proof. The claim is that

$$\tau_A(fX, Y) = f\tau_A(X, Y) \quad \text{and} \quad \tau_A(X, fY) = f\tau_A(X, Y).$$

Notice $\tau_A(X, Y) = -\tau_A(Y, X)$, so it suffices to check the first identity. Indeed,

$$\begin{aligned}\tau_A(fX, Y) &= \nabla_{fX} Y - \nabla_Y(fX) - [fX, Y] \\ &= f\nabla_X Y - (Y(f)X + f\nabla_Y X) - (f[X, Y] - Y(f)X) \\ &= f\nabla_X Y - Y(f)X - f\nabla_Y X - f[X, Y] + Y(f)X \\ &= f\nabla_X Y - f\nabla_Y X - f[X, Y] \\ &= f\tau_A(X, Y).\end{aligned}$$

□

Examples 14.3. • If A° is the directional derivative on $\mathbb{R}^n$, then $\tau_{A^\circ} \equiv 0$.

- If $M = \mathbb{S}^n \subset \mathbb{R}^{n+1}$ with the induced connection, then $\tau \equiv 0$. This is precisely Lemma 7.4.4 above.

Proposition 14.4. For $\nabla = \nabla^A$ on TM , the following are equivalent:

1. ∇ is torsion-free.
2. ∇ is symmetric, i.e. with respect to any coordinate frame, $A_{ij}^k = A_{ji}^k$.
3. The diagram

$$\begin{array}{ccc}\Omega^k(M) & \xrightarrow{\nabla} & \Gamma(T^*M \otimes \Lambda^k(T^*M)) \\ & \searrow d & \downarrow \wedge \\ & & \Omega^{k+1}(M)\end{array}$$

commutes, for all nonnegative integers k . Here $\nabla : \Omega^p(M) \rightarrow \Gamma(T^*M \otimes \Lambda^p(T^*M))$ is the induced covariant derivative of a tensor and $\wedge$ is the anti-symmetrization map $T^*M \otimes \Lambda^k(T^*M) \rightarrow \Lambda^{k+1}(T^*M)$.

4. Given any point $p \in M$, there exist local coordinates $\{x^i\}$ with $p = 0$ such that, with respect to $\{\frac{\partial}{\partial x^i}\}$, we have $A_{ij}^k(x_0) = 0$.

Proof. For (1) $\Leftrightarrow$ (2), the torsion tensor is $\tau(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$. In a coordinate frame, $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$, so

$$\tau\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} = (A_{ij}^k - A_{ji}^k) \frac{\partial}{\partial x^k}.$$

Thus $\tau \equiv 0$ iff $A_{ij}^k - A_{ji}^k = 0$ for all i, j, k .

For (2) $\Leftrightarrow$ (3), for $k = 1$, it suffices to check decomposable 1-forms. Let $\alpha \in \Omega^1(M)$ with $\alpha = \alpha_j dx^j$. The covariant derivative of α is:

$$\nabla \alpha = \left(\frac{\partial \alpha_j}{\partial x^i} - A_{ij}^k \alpha_k \right) dx^i \otimes dx^j$$

Applying $\wedge$, we get

$$(\wedge \circ \nabla) \alpha = \left(\frac{\partial \alpha_j}{\partial x^i} - A_{ij}^k \alpha_k \right) dx^i \wedge dx^j = d\alpha - \frac{1}{2} (A_{ij}^k - A_{ji}^k) \alpha_k dx^i \wedge dx^j.$$

Thus $\wedge \circ \nabla = d$ iff $A_{ij}^k = A_{ji}^k$.

For (4) $\Rightarrow$ (1), if $A_{ij}^k(p) = 0$, then

$$\tau_{ij}(p) = \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} - \overbrace{\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right]}^{=0} = A_{ij}^k(p) - A_{ji}^k(p) = 0.$$

For (1) $\Rightarrow$ (4), we will do this later—indeed, we will construct a coordinate system called *geodesic normal coordinates* which will have many nice properties, including this one. $\square$

15. THE FUNDAMENTAL LEMMA (TH 2/26)

We can now fix our uniqueness problem by requiring both metric compatibility and torsion-freeness.

Theorem 15.1 (Fundamental Lemma of Riemannian Geometry). *Given a Riemannian manifold (M, g) , there exists a **unique** torsion-free, metric-compatible connection*

$$\nabla^g = \nabla^{L.C.}$$

on TM , called the **Levi-Civita connection**.

Remark 15.2. In local coordinates, the connection coefficients of the Levi–Civita connection are known as **Christoffel symbols**, and are denoted with the special letter Γ . We shall see in the proof that these are determined by the local formula

$$A_{ij}^k = \Gamma_{ij}^k = \frac{1}{2}g^{k\ell} \left(\frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{i\ell}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell} \right).$$

Proof. Step 1. Assume that a metric compatible and torsion-free connection ∇ exists locally; we prove it is unique.

Metric compatibility reads

$$\frac{\partial}{\partial x^i} \left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k} \right\rangle = \left\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k} \right\rangle + \left\langle \frac{\partial}{\partial x^j}, \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^k} \right\rangle.$$

We write $g_{jk} = \left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k} \right\rangle$ and $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}$, and apply cyclic permutations, to obtain the three equations

$$\frac{\partial g_{jk}}{\partial x^i} = \Gamma_{ij}^m g_{mk} + \Gamma_{ik}^m g_{jm}, \quad (1)$$

$$\frac{\partial g_{ki}}{\partial x^j} = \Gamma_{jk}^m g_{mi} + \Gamma_{ji}^m g_{km}, \quad (2)$$

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ki}^m g_{mj} + \Gamma_{kj}^m g_{im}. \quad (3)$$

We add (1) and (2) and subtract (3). Using torsion-freeness $\Gamma_{ij}^k = \Gamma_{ji}^k$ and $g_{ij} = g_{ji}$, etc., we see that the first term on the RHS of (1) and the second term on the RHS of (2) combine and the other four terms cancel in pairs. We obtain

$$\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} = 2\Gamma_{ij}^m g_{mk}$$

Multiplying by the inverse matrix g^{kn} and using $g_{mk}g^{kn} = \delta_m^n$ gives

$$\Gamma_{ij}^n = \frac{1}{2}g^{kn} \left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right).$$

Hence, Γ_{ij}^k is uniquely determined by the requirements of metric compatibility and torsion-freeness.

Step 2. We now define a connection locally by the above formula. It is evident from the formula that $\Gamma_{ij}^k = \Gamma_{ji}^k$, so the connection is symmetric, i.e. torsion-free. We still need to check compatibility with the metric. This is left as a HW exercise. Now, as usual, since a metric-compatible and torsion-free connection exists locally and is unique on overlaps, it is globally well-defined. $\square$

Corollary 15.3. *Let $u : (M, g) \rightarrow (N, h)$ be an isometry. Then the Levi–Civita connections ∇^g and ∇^h satisfy*

$$u_*(\nabla_X^g Y) = \nabla_{u_*X}^h(u_*Y).$$

Proof. By uniqueness of the Levi–Civita connection, it suffices to check that the pushforward connection defined by $\tilde{\nabla}_X Y = (u^{-1})_*(\nabla_{u_*X}^h(u_*Y))$ is torsion-free and compatible with g . This is a routine exercise. $\square$

Proposition 15.4. *Let (M, g) be a Riemannian manifold and $N \subset M$ a submanifold. Let h be the induced metric on N . Then*

$$\nabla^h = \pi \circ \nabla^g,$$

where $\pi : TM|_N \rightarrow TN$ is orthogonal projection. In other words, the induced connection is exactly the Levi-Civita connection of the induced metric.

Proof. By uniqueness, it suffices to check that $\pi \circ \nabla^g$ is metric-compatible and torsion free. For X, Y, Z tangent to N , metric compatibility on M gives

$$X\langle Y, Z \rangle = \langle \nabla_X^g Y, Z \rangle + \langle Y, \nabla_X^g Z \rangle.$$

Since X, Y, Z are tangent to N , this is

$$\begin{aligned} X\langle Y, Z \rangle &= \langle \pi \circ \nabla_X^g Y, Z \rangle + \langle Y, \pi \circ \nabla_X^g Z \rangle \\ &= \langle \nabla_X^h Y, Z \rangle + \langle Y, \nabla_X^h Z \rangle, \end{aligned} \tag{15.1}$$

as desired. For torsion, let X, Y be tangent to N and extend to $\tilde{X}, \tilde{Y}$ on M . Then

$$\pi \circ \nabla_X^g Y - \pi \circ \nabla_Y^g X - [X, Y]_N = \pi(\nabla_{\tilde{X}}^g \tilde{Y} - \nabla_{\tilde{Y}}^g \tilde{X} - [\tilde{X}, \tilde{Y}]_M) = 0.$$

Here $[\tilde{X}, \tilde{Y}]_M$ is tangent to N (see 761 notes, Corollary 17.3), so $\pi[\tilde{X}, \tilde{Y}]_M = [X, Y]_N$. $\square$

16. CALCULATING CHRISTOFFEL SYMBOLS (TU 3/3)

Example 16.1 (Stereographic chart on $\mathbb{S}^n$). Let $\mathbb{S}^n = \{(X^0, \dots, X^n) \in \mathbb{R}^{n+1} : |X| = 1\}$ and let $N = (1, 0, \dots, 0)$ be the north pole. The stereographic chart from N is

$$(X^0, \dots, X^n) \mapsto \left(\frac{X^1}{1 - X^0}, \dots, \frac{X^n}{1 - X^0} \right).$$

Its inverse is given by

$$(x^1, \dots, x^n) \mapsto X(x) := \frac{1}{|x|^2 + 1} (|x|^2 - 1, 2x^1, \dots, 2x^n).$$

We checked in class that $|X(x)|^2 = 1$, as you have surely done before.

Let's change notation to make the calculation easier. Writing $E_0, E_1, \dots, E_n$ as the standard basis of $\mathbb{R}^{n+1}$, we have

$$X(x) = \frac{1}{|x|^2 + 1} ((|x|^2 - 1)E_0 + 2x^i E_i) \in \mathbb{R}^{n+1}.$$

We calculate

$$\frac{\partial X}{\partial x^i} = \frac{2}{1 + |x|^2} (x^i E_0 + E_i - x^i X), \quad 1 \leq i \leq n.$$

Hence

$$\begin{aligned} g_{ij} &= \left\langle \frac{\partial X}{\partial x^i}, \frac{\partial X}{\partial x^j} \right\rangle \\ &= \frac{4}{(1 + |x|^2)^2} \langle x^i E_0 + E_i - x^i X, x^j E_0 + E_j - x^j X \rangle \end{aligned}$$

Since $\langle E_0, E_j \rangle = 0$, $\langle E_i, E_j \rangle = \delta_{ij}$, $\langle E_0, X \rangle = X^0$, $\langle E_i, X \rangle = X^i$, and $|X| = 1$, we get

$$g_{ij} = \frac{4}{(1 + |x|^2)^2} (x^i x^j + \delta_{ij} + x^i x^j - x^i x^j X^0 - x^j X^i - x^i x^j X^0 - x^i X^j)$$

Since $x^i = \frac{X^i}{1 - X^0}$, we have $x^i X^0 + X^i = x^i$, the cross-terms simplify as follows:

$$\begin{aligned} -x^j (x^i X^0 + X^i) &= -x^i x^j \\ -x^i (x^j X^0 + X^j) &= -x^i x^j \end{aligned}$$

Placing these back into the sum gives

$$\begin{aligned} &= \frac{4}{(1 + |x|^2)^2} (2x^i x^j + \delta_{ij} - 2x^i x^j) \\ &= \frac{4\delta_{ij}}{(1 + |x|^2)^2}, \end{aligned}$$

which is the formula from Lecture 1 that you proved on Homework 1. As we already discussed, this shows the round metric $g_{ij} = \lambda^2 \delta_{ij}$ and the Euclidean metric are conformally related, with the conformal factor being $\lambda = \frac{2}{1 + |x|^2}$. I sketched for you a geometric proof of this fact on the board, which I encourage you to try and replicate yourself.

Let's now calculate the Christoffel symbols. We have

$$\begin{aligned} \frac{\partial g_{jk}}{\partial x^i} &= -\frac{16x^i}{(1 + |x|^2)^3} \delta_{jk} \\ &= -\frac{4x^i}{1 + |x|^2} g_{jk}. \end{aligned}$$

Using

$$\Gamma_{ij}^k = \frac{1}{2} g^{k\ell} \left(\frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{i\ell}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell} \right),$$

we compute

$$\begin{aligned} \Gamma_{ij}^k &= \frac{1}{2} g^{k\ell} \left(-\frac{4x^i}{1 + |x|^2} g_{j\ell} - \frac{4x^j}{1 + |x|^2} g_{i\ell} + \frac{4x^\ell}{1 + |x|^2} g_{ij} \right) \\ &= -\frac{2x^i}{1 + |x|^2} g^{k\ell} g_{j\ell} - \frac{2x^j}{1 + |x|^2} g^{k\ell} g_{i\ell} + \frac{2x^\ell}{1 + |x|^2} g^{k\ell} g_{ij} \\ &= -\frac{2x^i}{1 + |x|^2} \delta^k_j - \frac{2x^j}{1 + |x|^2} \delta^k_i + \frac{2x^k}{1 + |x|^2} \delta^i_j \\ &= -\frac{2}{1 + |x|^2} (x^i \delta^j_k + x^j \delta^i_k - x^k \delta^i_j). \end{aligned}$$

In particular,

$$\begin{aligned}
 \Gamma_{ij}^k &= 0 \text{ if } i, j, k \text{ are distinct} \\
 \Gamma_{ij}^j &= \Gamma_{ji}^j = -\frac{2}{1+|x|^2} (x^i \delta_{jj}^j + x^j \delta_{ji}^j - x^j \delta_{ij}^j) \text{ (no sum on } j) \\
 &= -\frac{2x^i}{1+|x|^2}, \\
 \Gamma_{ii}^k &= -\frac{2}{1+|x|^2} (2x^i \delta_{ik}^i - x^k) \text{ (no sum on } i) \\
 &= \frac{2x^k}{1+|x|^2} \text{ if } i \neq k
 \end{aligned} \tag{16.1}$$

Example 16.2 (The hyperbolic plane). Let

$$\mathbb{H}^2 = \{(x, y) \in \mathbb{R}^2 : y > 0\}$$

with coordinates $x^1 = x$, $x^2 = y$. The hyperbolic metric is

$$g_{ij} = \frac{1}{y^2} \delta_{ij}, \quad i, j \in \{1, 2\}.$$

Equivalently,

$$g = \frac{dx^2 + dy^2}{y^2}.$$

So the matrix form is

$$(g_{ij}) = \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} y^2 & 0 \\ 0 & y^2 \end{pmatrix}.$$

Since $g_{ij} = y^{-2} \delta_{ij}$ depends only on y ,

$$\frac{\partial g_{ij}}{\partial x} = 0, \quad \frac{\partial g_{ij}}{\partial y} = -\frac{2}{y^3} \delta_{ij}.$$

Using

$$\Gamma_{ij}^k = \frac{1}{2} g^{k\ell} \left(\frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{i\ell}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell} \right)$$

and $g^{k\ell} = y^2 \delta^{k\ell}$, we get

$$\begin{aligned}
 \Gamma_{11}^k &= \frac{y^2}{2} \left(\frac{\partial g_{1k}}{\partial x^1} + \frac{\partial g_{1k}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^k} \right) \\
 &= -\frac{y^2}{2} \frac{\partial g_{11}}{\partial x^k} \\
 &= \frac{\delta_{k2}}{y},
 \end{aligned}$$

so $\Gamma_{11}^1 = 0$ and $\Gamma_{11}^2 = \frac{1}{y}$. Similarly,

$$\begin{aligned}
 \Gamma_{12}^k &= \frac{y^2}{2} \left(\frac{\partial g_{2k}}{\partial x^1} + \frac{\partial g_{1k}}{\partial x^2} - \frac{\partial g_{12}}{\partial x^k} \right) \\
 &= \frac{y^2}{2} \frac{\partial g_{1k}}{\partial x^2}
 \end{aligned}$$

$$= -\frac{\delta_{k1}}{y},$$

so $\Gamma_{12}^1 = \Gamma_{21}^1 = -\frac{1}{y}$ and $\Gamma_{12}^2 = \Gamma_{21}^2 = 0$. Finally,

$$\begin{aligned}\Gamma_{22}^k &= \frac{y^2}{2} \left(\frac{\partial g_{2k}}{\partial x^2} + \frac{\partial g_{2k}}{\partial x^2} - \frac{\partial g_{22}}{\partial x^k} \right) \\ &= \frac{y^2}{2} \left(-\frac{4}{y^3} \delta_{2k} + \frac{2}{y^3} \delta_{2k} \right) \\ &= -\frac{\delta_{k2}}{y},\end{aligned}$$

so $\Gamma_{22}^1 = 0$ and $\Gamma_{22}^2 = -\frac{1}{y}$.

In summary, the Christoffel symbols for the hyperbolic metric are:

$$\Gamma_{11}^1 = 0, \quad \Gamma_{11}^2 = \frac{1}{y}, \quad (16.2)$$

$$\Gamma_{12}^1 = \Gamma_{21}^1 = -\frac{1}{y}, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = 0, \quad (16.3)$$

$$\Gamma_{22}^1 = 0, \quad \Gamma_{22}^2 = -\frac{1}{y}. \quad (16.4)$$

Part 5. Geodesics

Recall the following fundamental definitions from Lecture 1. The *length* of a piecewise smooth curve γ in a Riemannian manifold (M, g) is

$$L_g(\gamma) = \int_a^b |\gamma'(t)|_g dt.$$

Given two points $p, q \in M$, the associated *Riemannian distance* is given by

$$d_g(p, q) = \inf \{L_g(\gamma) : \gamma \text{ piecewise smooth, } \gamma(a) = p, \gamma(b) = q\}.$$

Goal: Use the Levi–Civita connection to investigate $d_g(\cdot, \cdot)$.

These have no relation *a priori*, but will be connected through the following definition.

17. THE GEODESIC EQUATION (TU 3/3)

17.1. **The definition.** A smooth curve $\gamma : I \rightarrow M$ is called a **geodesic**⁵ if:

$$\boxed{\frac{D\dot{\gamma}(t)}{dt} = 0,}$$

where $\frac{D}{dt}$ is the Levi–Civita covariant derivative along γ , i.e. it differentiates the section $\dot{\gamma}(t) \in T_{\gamma(t)}M$ of TM along the curve γ .

In local coordinates $\{x^i\}_{i=1}^n$, write

$$\gamma(t) = (\gamma^1(t), \dots, \gamma^n(t)), \quad \dot{\gamma}^i = \frac{d\gamma^i}{dt}.$$

Then

$$\left(\frac{D\dot{\gamma}}{dt}\right)^k = \frac{d}{dt} \left(\frac{d\gamma^k}{dt}\right) + \frac{d\gamma^i}{dt} \Gamma_{ij}^k(\gamma(t)) \frac{d\gamma^j}{dt}.$$

So the geodesic equation is

$$\boxed{\ddot{\gamma}^k + \Gamma_{ij}^k(\gamma(t)) \dot{\gamma}^i \dot{\gamma}^j = 0,} \quad k = 1, \dots, n.$$

Proposition 17.1. *Given $x \in M$ and $v \in T_x M$, there exists a unique geodesic*

$$\gamma : (a, b) \rightarrow M$$

such that

$$\gamma(0) = x, \quad \dot{\gamma}(0) = v, \quad a < 0 < b.$$

Moreover, (a, b) can be chosen maximal among all intervals containing 0 on which this geodesic is defined.

Proof. This is true of any second-order ODE system. □

Note: Since the system is *nonlinear*, we cannot always take $a = -\infty$ and $b = \infty$! This is a bit different than with the parallel-transport ODE studied above. Meanwhile, the following proposition shows that the geodesic equation has far better properties than a general second-order ODE.

⁵In Lecture 1 we gave a different definition of geodesic; we will show later that the two definitions are equivalent.

Proposition 17.2. *Geodesics have constant speed:*

$$|\dot{\gamma}(t)|_g = \text{const.}$$

Proof. By metric compatibility,

$$\frac{d}{dt}|\dot{\gamma}(t)|_g^2 = \frac{d}{dt}\langle \dot{\gamma}, \dot{\gamma} \rangle = 2 \left\langle \frac{D\dot{\gamma}}{dt}, \dot{\gamma} \right\rangle.$$

If γ is geodesic, then $\frac{D\dot{\gamma}}{dt} = 0$, hence

$$\frac{d}{dt}|\dot{\gamma}(t)|_g^2 = 0.$$

So $|\dot{\gamma}(t)|_g^2$ (and therefore $|\dot{\gamma}(t)|_g$) is constant. $\square$

Equivalently: if γ is a geodesic, then its velocity field $\dot{\gamma}$ is parallel along γ . Since the Levi-Civita connection is metric-compatible, parallel transport preserves inner products (hence length), which gives the same conclusion.

Definition 17.3. A geodesic $\gamma : (a, b) \rightarrow M$ is called **complete** if

$$a = -\infty, \quad b = \infty.$$

18. EXAMPLES (TU 3/3-TH 3/5)

18.1. Trivial examples.

- $M = \mathbb{R}^n$: Geodesics are straight lines, so all geodesics are complete.
- $M = B_1(0) \subseteq \mathbb{R}^n$: No complete, nonconstant geodesics (any straight line hits the boundary in finite time).
- $M = \mathbb{R}^n \setminus \{0\}$: Most, but not all, geodesics are complete (a straight line aimed at the origin ends in finite time).

18.2. Nontrivial examples.

18.2.1. *The sphere.* All geodesics are complete—in fact every geodesic is a great circle, defined for all $t \in \mathbb{R}$. This is easy to see by parameterizing the great circle as $\gamma(t) = \cos tv + \sin tw$, where v and w are orthogonal unit vectors, and observing that $\ddot{\gamma} = -\gamma(t)$. This shows that the acceleration of γ is orthogonal to the tangent space of S^n , so dies under the orthogonal projection.

Lee's book also gives a “clever” proof that the geodesic starting at $x \in \mathbb{S}^n$ with initial velocity v must lie in the 2-plane through the origin spanned by x and v . Indeed, reflection across this plane is an isometry that fixes x and v . If the geodesic were to leave the plane, reflecting it would produce a different geodesic with the same initial data, contradicting uniqueness.

Let's now try to recover this answer by direct calculation in stereographic coordinates. Recall from (16.1) that the nonzero Christoffel symbols are

$$\boxed{\Gamma_{ij}^i = -\frac{2x^j}{1+|x|^2}, \quad \Gamma_{ii}^k = \frac{2x^k}{1+|x|^2} \text{ for } i \neq k.} \quad (18.1)$$

Fix $i \in \{1, \dots, n\}$ and consider

$$\gamma(t) = (0, \dots, 0, f(t), 0, \dots, 0),$$

with $f(t)$ in the i -th slot. The geodesic equation is

$$\ddot{\gamma}^k + \Gamma_{ab}^k(\gamma(t))\dot{\gamma}^a\dot{\gamma}^b = 0.$$

Along this curve, the only nonzero velocity component is $\dot{\gamma}^i = f'(t)$, so this reduces to

$$\ddot{\gamma}^k + \Gamma_{ii}^k(\gamma(t))(f'(t))^2 = 0.$$

If $k \neq i$, then $\gamma^k \equiv 0$, so $\ddot{\gamma}^k = 0$, and also

$$\Gamma_{ii}^k(\gamma(t)) = \frac{2\gamma^k}{1+|\gamma(t)|^2} = 0$$

because $\gamma^k = 0$. Hence the $k \neq i$ equations are automatically satisfied.

If $k = i$, then $\ddot{\gamma}^i = f''(t)$ and

$$\Gamma_{ii}^i(\gamma(t)) = -\frac{2\gamma^i}{1+|\gamma(t)|^2} = -\frac{2f(t)}{1+f(t)^2}.$$

So the $k = i$ equation becomes

$$f''(t) - \frac{2f(t)}{1+f(t)^2}(f'(t))^2 = 0. \quad (18.2)$$

How on earth does one solve this?

Instead of solving a difficult 2nd-order ODE, let's just recall that $|\dot{\gamma}|_g = \text{const}$ (let's call it $2c$ for convenience, so the square is $4c^2$). So we have

$$|\dot{\gamma}|_g^2 = \frac{4}{(1+f(t)^2)^2}(f'(t))^2 = \text{const} = 4c^2.$$

Hence

$$f'(t) = \pm c(1+f(t)^2).$$

After choosing sign in c , this is

$$\frac{f'(t)}{1+f(t)^2} = c. \quad (18.3)$$

To see that the equation (18.3) implies (18.2), simply take the derivative of both sides.

Now, to solve for $f(t)$, we integrate (18.3), to get

$$\arctan(f(t)) = ct + d,$$

so

$$f(t) = \tan(ct + d).$$

Hence the stereographic-coordinate expression is only valid while $ct + d \in (-\frac{\pi}{2}, \frac{\pi}{2})$ (more generally, within any interval of length π avoiding poles of $\tan$). As $ct + d$ approaches a pole,

the coordinate value runs off to infinity, meaning the geodesic exits this coordinate chart, while the geodesic itself remains continuous (indeed smooth) on $\mathbb{S}^n$.

18.2.2. *The hyperbolic plane.* Consider the upper-half-plane model

$$\mathbb{H}^2 = \{(x, y) \in \mathbb{R}^2 : y > 0\}, \quad g = \frac{dx^2 + dy^2}{y^2}.$$

Recall from (16.2) that the Christoffel symbols are:

$$\boxed{\Gamma_{11}^1 = 0, \quad \Gamma_{11}^2 = \frac{1}{y}, \quad \Gamma_{12}^1 = \Gamma_{21}^1 = -\frac{1}{y}, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = 0, \quad \Gamma_{22}^1 = 0, \quad \Gamma_{22}^2 = -\frac{1}{y}.} \quad (18.4)$$

Consider a curve

$$\gamma(t) = (0, f(t)).$$

Then

$$\dot{\gamma}^1 = 0, \quad \dot{\gamma}^2 = f'(t), \quad \ddot{\gamma}^1 = 0, \quad \ddot{\gamma}^2 = f''(t).$$

The geodesic equation is

$$\ddot{\gamma}^k + (\dot{\gamma}^1)^2 \Gamma_{11}^k(\gamma(t)) + 2\dot{\gamma}^1 \dot{\gamma}^2 \Gamma_{12}^k(\gamma(t)) + (\dot{\gamma}^2)^2 \Gamma_{22}^k(\gamma(t)) = 0, \quad (18.5)$$

for $k = 1, 2$.

The $k = 1$ equation is

$$0 + \Gamma_{22}^1(\gamma(t))(f'(t))^2 = 0,$$

which is automatic since $\Gamma_{22}^1 = 0$. The $k = 2$ equation is

$$f''(t) + \Gamma_{22}^2(\gamma(t))(f'(t))^2 = 0.$$

Along $\gamma(t)$, we have $\gamma^2(t) = f(t)$, hence

$$\Gamma_{22}^2(\gamma(t)) = -\frac{1}{f(t)}.$$

So

$$f''(t) - \frac{(f'(t))^2}{f(t)} = 0.$$

Equivalently,

$$\frac{d}{dt} \left(\frac{f'}{f} \right) = 0,$$

so $\frac{f'}{f} = c$ and therefore

$$f(t) = e^{ct+d}.$$

Hence the vertical rays $\gamma(t) = (0, e^{ct+d})$ are geodesics.

Now check a semicircle:

$$\boxed{\gamma(t) = (\sin \theta(t), \cos \theta(t)),} \quad \theta(t) \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right).$$

Then

$$\dot{\gamma}(t) = \dot{\theta}(\cos \theta, -\sin \theta),$$

and

$$\ddot{\gamma}(t) = \ddot{\theta}(\cos \theta, -\sin \theta) - \dot{\theta}^2(\sin \theta, \cos \theta).$$

So

$$\dot{\gamma}^1 = \dot{\theta} \cos \theta, \quad \dot{\gamma}^2 = -\dot{\theta} \sin \theta, \quad \ddot{\gamma}^1 = \ddot{\theta} \cos \theta - \dot{\theta}^2 \sin \theta, \quad \ddot{\gamma}^2 = -\ddot{\theta} \sin \theta - \dot{\theta}^2 \cos \theta.$$

Using

$$\Gamma_{12}^1 = \Gamma_{21}^1 = -\frac{1}{y}, \quad \Gamma_{11}^2 = \frac{1}{y}, \quad \Gamma_{22}^2 = -\frac{1}{y},$$

the geodesic equations (18.5) become

$$\ddot{\gamma}^1 - \frac{2}{y} \dot{\gamma}^1 \dot{\gamma}^2 = 0, \quad \ddot{\gamma}^2 + \frac{1}{y} ((\dot{\gamma}^1)^2 - (\dot{\gamma}^2)^2) = 0.$$

The $k = 1$ equation is

$$\ddot{\theta} \cos \theta - \dot{\theta}^2 \sin \theta - \frac{2}{\cos \theta} (\dot{\theta} \cos \theta)(-\dot{\theta} \sin \theta) = 0,$$

which simplifies to

$$\boxed{\ddot{\theta} \cos \theta + \dot{\theta}^2 \sin \theta = 0.} \quad (18.6)$$

The $k = 2$ equation is

$$-\ddot{\theta} \sin \theta - \dot{\theta}^2 \cos \theta + \frac{1}{\cos \theta} (\dot{\theta}^2 \cos^2 \theta - \dot{\theta}^2 \sin^2 \theta) = 0,$$

which simplifies to

$$\tan \theta (\ddot{\theta} \cos \theta + \dot{\theta}^2 \sin \theta) = 0.$$

This is equivalent to (18.6). We can already conclude that the semicircle is the image of a geodesic.

Let's now solve (18.6) to see whether or not the semicircular geodesic is complete. A very clever person might observe directly that

$$\ddot{\theta} \cos \theta + \dot{\theta}^2 \sin \theta = \frac{d}{dt} \left(\frac{\dot{\theta}}{\cos \theta} \right) \cos^2 \theta = 0.$$

Or recall that $|\dot{\gamma}|_g = \text{const}$:

$$|\dot{\gamma}|_g^2 = \frac{(\dot{\gamma}_1)^2 + (\dot{\gamma}_2)^2}{y^2} = \frac{\dot{\theta}^2}{\cos^2 \theta} = c^2,$$

so

$$\frac{\dot{\theta}}{\cos \theta} = c.$$

Integrating, one gets

$$\log |\sec(\theta) + \tan(\theta)| = ct + d$$

and

$$\sec(\theta) + \tan(\theta) = e^{ct+d}.$$

One can examine the LHS, $\frac{1+\sin \theta}{\cos \theta}$, and see that this is a diffeomorphism $(-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow (0, \infty)$. So the solution $\theta(t)$ is in fact defined for $t \in (-\infty, \infty)$, giving a complete geodesic.

With some half-angle formulas (courtesy of google), one can actually determine the inverse function of $x = \sec(\theta) + \tan(\theta)$ to be $\theta = 2 \arctan(x) - \frac{\pi}{2}$. So the explicit formula is

$$\boxed{\theta(t) = 2 \arctan(e^{ct+d}) - \frac{\pi}{2}.$$

One can translate and scale the semicircular and vertical geodesics to get a complete geodesic in $\mathbb{H}^n$ with any initial position and velocity. In other words, all geodesics in hyperbolic space are complete. Also notice that there exists a unique geodesic between any two points, by contrast with the sphere.

19. THE EXPONENTIAL MAP (Th 3/5)

Goal: Use geodesics to construct a “good” local coordinate system near $p \in M$, analogous to the adapted frames we studied before.

Notation: Fix $p \in M$ and $v \in T_p M$. Write

$$\gamma_v^{(p)}(t) = \gamma_v(t)$$

for the unique geodesic starting at $\gamma_v(0) = p$ with initial velocity

$$\dot{\gamma}_v(0) = v.$$

Lemma 19.1. *For each $p \in M, v \in T_p M$, and $c, t \in \mathbb{R}$,*

$$\gamma_{cv}(t) = \gamma_v(ct)$$

whenever either side is defined.

Proof. Define

$$\tilde{\gamma}(t) := \gamma_v(ct).$$

Then

$$\tilde{\gamma}(0) = x, \quad \dot{\tilde{\gamma}}(t) = c \dot{\gamma}_v(ct), \quad \ddot{\tilde{\gamma}}(0) = c \dot{\gamma}_v(0) = cv, \quad \ddot{\tilde{\gamma}}(t) = c^2 \ddot{\gamma}_v(ct).$$

Hence

$$\begin{aligned} \ddot{\tilde{\gamma}}^k + \Gamma_{ij}^k(\tilde{\gamma}) \dot{\tilde{\gamma}}^i \dot{\tilde{\gamma}}^j &= c^2 \ddot{\gamma}_v^k(ct) + \Gamma_{ij}^k(\gamma_v(ct)) c \dot{\gamma}_v^i(ct) c \dot{\gamma}_v^j(ct) \\ &= c^2 \left(\ddot{\gamma}_v^k(ct) + \Gamma_{ij}^k(\gamma_v(ct)) \dot{\gamma}_v^i(ct) \dot{\gamma}_v^j(ct) \right) = 0. \end{aligned}$$

So $\tilde{\gamma}$ is a geodesic. By uniqueness of geodesics with given initial data, we are done. $\square$

Definition 19.2. The **exponential map** is defined by

$$\exp : \mathcal{E} \subset TM \rightarrow M$$

$$\exp(v) := \gamma_v^{(x)}(1),$$

where $x \in M$ is arbitrary, $v \in T_x M$, and

$$\mathcal{E} := \{v \in TM \mid \gamma_v^{(x)}(t) \text{ is defined on an open interval containing } [0, 1]\}.$$

For each $p \in M$, set

$$\mathcal{E}_p := \mathcal{E} \cap T_p M.$$

Observe that $\mathcal{E}_p$ is open and star-shaped about $0 \in T_p M$. In fact, by smooth dependence of ODE solutions on initial data, the full domain $\mathcal{E} \subset TM$ is open and $\exp$ is a smooth map.

We also write

$$\exp_p : \mathcal{E}_p \rightarrow M, \quad \exp_p(v) := \gamma_v(1).$$

Proposition 19.3. *For each $p \in M$, the exponential map satisfies:*

1. $\exp_p(tv) = \gamma_v(t)$ for all $v \in \mathcal{E}_p$ and all t for which either side is defined.
2. $d(\exp_p)_0 : T_0(T_p M) \rightarrow T_p M$ is the identity map.
3. $\exp_p$ is a local diffeomorphism near $0 \in T_p M$.

Proof. (1) Since $\exp_p(tv) = \gamma_{tv}(1) \stackrel{19.1}{=} \gamma_v(t)$.

(2) Identify $T_0(T_p M) \cong T_p M$ via the vector-space structure on $T_p M$. For $w \in T_p M$, run along the path $t \mapsto tw$. Differentiating at $t = 0$,

$$d(\exp_p)_0(w) = \left. \frac{d}{dt} \right|_0 \exp_p(tw) \stackrel{(1)}{=} \left. \frac{d}{dt} \right|_0 \gamma_w(t) = \dot{\gamma}_w(0) = w.$$

Hence $d(\exp_p)_0 = \text{Id}$.

(3) Since $d(\exp_p)_0$ is an isomorphism, the inverse function theorem implies that $\exp_p$ is a diffeomorphism from a neighborhood of $0 \in T_p M$ onto a neighborhood of $p \in M$. $\square$

Terminology:

- **Geodesic ball** centered at p , of radius ε : the diffeomorphic image $\exp_p(B_\varepsilon(0))$, sometimes written $B_\varepsilon(p)$.
- **Geodesic sphere** $S_\varepsilon^{n-1}(p)$: the diffeomorphic image $\exp_p(\partial B_\varepsilon(0))$.

- **Pointwise injectivity radius:**

$$\text{inj}_p(M) = \sup\{\varepsilon > 0 \mid \exp_p : B_\varepsilon(0) \rightarrow M \text{ is a diffeomorphism onto its image}\}.$$

- **Global injectivity radius:**

$$\text{inj}(M) = \inf_{p \in M} \text{inj}_p(M) \in [0, \infty].$$

- **Normal neighborhood of p :** an open set $W \ni p$ such that

$$W = \exp_p(W_0),$$

where $W_0 \subset T_p M$ is open, star-shaped with respect to 0 , and $\exp_p : W_0 \rightarrow W$ is a diffeomorphism.

- **(δ -)uniformly normal neighborhood:** a normal neighborhood $W \ni p$ such that there exists $\delta > 0$ with

$$W \subset B_\delta(q) \quad \text{for all } q \in W.$$

Definition 19.4. Fix $p \in M$ and an orthonormal basis $\{E_1, \dots, E_n\}$ of $T_p M$. Define the linear isomorphism

$$E : \mathbb{R}^n \rightarrow T_p M, \quad E(x^1, \dots, x^n) := x^i E_i.$$

If U is any normal neighborhood of p , the map

$$\varphi := E^{-1} \circ \exp_p^{-1} : U \xrightarrow{\sim} \varphi(U) \subset \mathbb{R}^n$$

is called a **normal coordinate chart** centered at p (also called **geodesic normal coordinates**).

Definition 19.5. In normal coordinates $(x^1, \dots, x^n)$ centered at p , define

$$r(x) = \sqrt{\sum_{i=1}^n (x^i)^2}.$$

This definition is independent of the choice of orthonormal basis $\{E_1, \dots, E_n\}$.

The vector field on $\{r > 0\}$

$$\frac{\partial}{\partial r} = \frac{x^i}{r} \frac{\partial}{\partial x^i}.$$

is called the **radial vector field**.

Proposition 19.6. Let $(x^1, \dots, x^n)$ be normal coordinates centered at p .

1. For $v \in T_p M$, the geodesic starting at p with initial velocity v is the line

$$\gamma_v(t) = t v$$

2. The components of the metric at p are

$$g_{ij}(p) = \delta_{ij}.$$

3. If $|v| = 1$, then along the geodesic γ_v ,

$$\dot{\gamma}_v(t) = \frac{\partial}{\partial r}.$$

4. The Christoffel symbols vanish at p :

$$\Gamma_{ij}^k(p) = 0 \quad \text{for all } i, j, k.$$

5. All of the first partial derivatives of g_{ij} vanish at p :

$$\left. \frac{\partial g_{ij}}{\partial x^k} \right|_p = 0 \quad \text{for all } i, j, k.$$

Proof. (1) By definition of normal coordinates,

$$\varphi = E^{-1} \circ \exp_p^{-1},$$

so for $v \in T_p M$ we have

$$\varphi(\gamma_v(t)) = \varphi(\exp_p(tv)) = E^{-1}(tv) = t E^{-1}(v).$$

Thus geodesics through p are $t \mapsto tv$ in normal coordinates.

(2) The basis $\{E_i\}$ at p is orthonormal, and the coordinate vector fields $\frac{\partial}{\partial x^i}|_p$ agree with this basis, so

$$g_{ij}(p) = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle_p = \delta_{ij}.$$

(3) In normal coordinates, $r(x) = \sqrt{\sum_i (x^i)^2}$ and

$$\frac{\partial}{\partial r} = \frac{x^i}{r} \frac{\partial}{\partial x^i}.$$

Along $\gamma_v(t) = tv$ with $|v| = 1$, we have $x^i = tv^i$ and $r = t$, the radial vector field

$$\frac{\partial}{\partial r} = \frac{tv^i}{t} \frac{\partial}{\partial x^i} = v^i \frac{\partial}{\partial x^i} = \dot{\gamma}_v(t).$$

(4) For any $v = (v^1, \dots, v^n)$, the geodesic $\gamma_v(t) = tv$ has $\ddot{\gamma}_v^k = 0$. The geodesic equation simplifies to

$$\Gamma_{ij}^k(tv) v^i v^j = 0.$$

At $t = 0$,

$$\Gamma_{ij}^k(0) v^i v^j = 0 \quad \text{for all } k.$$

In particular, take $v = \frac{\partial}{\partial x^i}$. Then

$$\Gamma_{ii}^k(0) = 0 \quad \text{for all } i, k.$$

Now take $v = \frac{\partial}{\partial x^i} + \frac{\partial}{\partial x^j}$ for $i \neq j$. Using $\Gamma_{ii}^k(0) = \Gamma_{jj}^k(0) = 0$,

$$\Gamma_{ij}^k(0) + \Gamma_{ji}^k(0) = 0.$$

By torsion freeness, $\Gamma_{ij}^k = \Gamma_{ji}^k$, therefore $\Gamma_{ij}^k(0) = 0$ for all i, j, k .

(5) Use metric compatibility in coordinates:

$$\Gamma_{ki}^l g_{lj} + \Gamma_{kj}^l g_{il} = \frac{\partial g_{ij}}{\partial x^k}.$$

By (4), $\Gamma_{ki}^l(p) = 0$ for all k, i, l , hence

$$\frac{\partial g_{ij}}{\partial x^k} \Big|_p = 0$$

for all i, j, k . □

Remark 19.7. Combining items (2) and (5) of the last proposition, we see that

$$g_{ij} = \delta_{ij} + O(2)$$

in normal coordinates.

20. UNIFORMLY NORMAL NEIGHBORHOODS (TU 3/10)

Recall from above that a neighborhood $W \ni p$ is called $(\delta-)$ uniformly normal if there exists $\delta > 0$ such that

$$W \subset B_\delta(q) \quad \text{for all } q \in W.$$

Here, as before, $B_\delta(q)$ denotes the diffeomorphic image of $B_\delta(0) \subset T_q M$ under the exponential map. We will now see that uniformly normal neighborhoods always exist and derive a simple consequence of this fact.

Proposition 20.1. *Every $p \in M$ has a uniformly normal neighborhood.*

Proof. Consider the smooth map

$$F : \mathcal{E} \subset TM \rightarrow M \times M,$$

$$(x, v) \mapsto (x, \exp_x(v)).$$

Recall that the tangent bundle is constructed so that given any coordinate system $U = \{x^i\} \subset M$ centered at p , we have the following coordinate system on $\pi^{-1}(U) \cong U \times \mathbb{R}^n \subset M$:

$$(x^i, v^j) \mapsto (x^i, x^j \frac{\partial}{\partial x^j}).$$

In this coordinate system, the derivative of F at $p = (0, 0)$ is given by

$$(dF)_{(0,0)} = \left(\begin{array}{cc} \frac{\partial x^i}{\partial x^j} & \frac{\partial x^i}{\partial v^j} \\ \frac{\partial \exp_x(v)^i}{\partial x^j} & \frac{\partial \exp_x(v)^i}{\partial v^j} \end{array} \right) \Big|_{(p,0)}$$

The upper-left block is the identity. Since x^i does not depend on v^j , the upper-right block is zero. By Proposition 19.3.2, the lower-left block is also the identity. We obtain

$$(dF)_{(0,0)} = \begin{pmatrix} I & 0 \\ * & I \end{pmatrix}.$$

This is an invertible matrix; by the inverse function theorem, F is a local diffeomorphism near p . In fact, we take the domain to be an open set of the form

$$U' = \{(q, v) \mid q \in B_\delta(p) \subset M, v \in B_\delta(0) \subset T_p M\}.$$

For δ sufficiently small, F induces an isomorphism

$$F : U' \xrightarrow{\sim} W' \subset M \times M$$

from U' onto its image W' .

Now, by the definition of the product topology, we can choose $W \ni p$ such that

$$W \times W \subset W'.$$

For each $q \in W$, we have

$$\begin{aligned} F(\{q\} \times B_\delta(0)) &= (\{q\} \times M) \cap W' \\ &\supset \{q\} \times W. \end{aligned} \tag{20.1}$$

Hence

$$\exp_q(B_\delta(0)) \supset W$$

for each $q \in W$, as desired. $\square$

Corollary 20.2. *Suppose M is compact. Then $\text{inj}(M) > 0$.*

Proof. Cover M by finitely many δ_i -uniformly normal neighborhoods W_i . For $q \in W_i$, we have

$$\text{inj}_q(M) \geq \delta_i.$$

Hence

$$\text{inj}(M) = \inf_{q \in M} \text{inj}_q(M) \geq \min_i \delta_i > 0.$$

$\square$

21. FIRST VARIATION OF ARCLENGTH (TU 3/10 - TH 3/12)

We will now delve deeper into the relationship between geodesics and arclength.

21.1. The first-variation formula. In studying the arc-length functional, it is beneficial to work with the following class of curves on M .

Definition 21.1. • An **admissible curve** is a continuous, piecewise smooth curve

$$\gamma : I = [a, b] \rightarrow M$$

with $\dot{\gamma}(t) \neq 0$ for all $t \in I$. I.e., there exists a partition

$$a = a_0 < a_1 < \cdots < a_k = b$$

of I such that γ is smooth on each interval $[a_i, a_{i+1}]$.

- The derivative from the left at a_i is denoted $\dot{\gamma}(a_i^-)$ and the derivative from the right is denoted $\dot{\gamma}(a_i^+)$. Define

$$\Delta_i \dot{\gamma} = \dot{\gamma}(a_i^+) - \dot{\gamma}(a_i^-).$$

- An **admissible family** is a map

$$\Gamma : J \times I \rightarrow M,$$

where J is also an interval, such that

$$\Gamma_s(t) := \Gamma(s, t)$$

is an admissible curve for each s with respect to a fixed partition $\{a_i\}_{i=0}^n$. We also require that Γ is smooth on $J \times (a_i, a_{i+1})$.

- A **variation** of γ is any admissible family Γ such that

$$\Gamma_0(t) = \gamma(t).$$

A **proper variation** is further required to satisfy

$$\Gamma(s, a) \equiv \gamma(a), \quad \Gamma(s, b) \equiv \gamma(b).$$

- The **variation field** of Γ is

$$V(t) = \frac{\partial \Gamma}{\partial s}(0, t),$$

which is a vector field along γ .

Remark 21.2. Given any vector field $V(t)$ along γ which is continuous and smooth when restricted to each $[a_i, a_{i+1}]$, it is always possible to construct a variation Γ with $V = \frac{\partial \Gamma}{\partial s}|_{s=0}$. For instance, one can just take

$$\Gamma(s, t) = \exp_{\gamma(t)}(sV(t)),$$

which is defined on an open interval $J \ni 0$.

Lemma 21.3. *Let Γ be any variation of γ . Then*

$$\frac{D}{\partial s} \frac{\partial \Gamma}{\partial t} = \frac{D}{\partial t} \frac{\partial \Gamma}{\partial s}.$$

Proof. Locally write

$$\Gamma(s, t) = (x^1(s, t), \dots, x^n(s, t)).$$

Then

$$\frac{\partial \Gamma}{\partial t} = \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i}, \quad \frac{\partial \Gamma}{\partial s} = \frac{\partial x^i}{\partial s} \frac{\partial}{\partial x^i}.$$

We compute

$$\left(\frac{D}{\partial s} \frac{\partial \Gamma}{\partial t} \right)^k = \frac{\partial^2 x^k}{\partial s \partial t} + \frac{\partial x^i}{\partial s} \frac{\partial x^j}{\partial t} \Gamma_{ij}^k.$$

Similarly,

$$\left(\frac{\partial D}{\partial t} \frac{\partial \Gamma}{\partial s} \right)^k = \frac{\partial^2 x^k}{\partial t \partial s} + \frac{\partial x^i}{\partial t} \frac{\partial x^j}{\partial s} \Gamma_{ij}^k.$$

Since the mixed partial derivatives commute and Γ is symmetric, the two expressions are equal. $\square$

Theorem 21.4 (First variation of arclength). *Suppose γ is an admissible curve with unit speed and*

$$\Gamma : J \times [a, b] \rightarrow M$$

is an admissible variation. Let

$$V(t) = \frac{\partial \Gamma}{\partial s}(0, t).$$

Then

$$\frac{dL(\Gamma_s)}{ds} \Big|_{s=0} = - \int_a^b \langle V, D_t \dot{\gamma} \rangle dt - \sum_{i=1}^{k-1} \langle V(a_i), \Delta_i \dot{\gamma} \rangle + \langle V(b), \dot{\gamma}(b) \rangle - \langle V(a), \dot{\gamma}(a) \rangle.$$

Remark 21.5. Notice that the RHS turns out to only involve V and not the higher orders of Γ_s , which is not surprising since this is the first derivative. Traditionally, the LHS of the formula would be written as $\delta_V L$ or $\delta L(V)$ or something similar, to represent a “variational derivative” in the direction V . This is also the modern point of view: one sets up the space of paths as a Banach or Fréchet manifold on which $L_g(\cdot)$ is a (Fréchet-)differentiable function. For our purposes in this course, the “admissible path” setup is perfectly adequate.

Proof. We have

$$L(\Gamma_s) = \sum_{i=1}^k L(\Gamma_s|_{[a_{i-1}, a_i]}),$$

each term being smooth. Let

$$T = \frac{\partial \Gamma}{\partial t}, \quad S = \frac{\partial \Gamma}{\partial s}.$$

Then

$$\begin{aligned} \left. \frac{d}{ds} L(\Gamma_s|_{[a_{i-1}, a_i]}) \right|_{s=0} &= \int_{a_{i-1}}^{a_i} \frac{\partial}{\partial s} \langle T, T \rangle^{1/2} dt. \\ &= \int_{a_{i-1}}^{a_i} \frac{1}{2|T|} \cdot 2 \left\langle \frac{DT}{ds}, T \right\rangle dt \Big|_{s=0}. \end{aligned} \quad (21.1)$$

Since γ has unit speed, this becomes

$$\left. \frac{d}{ds} L(\Gamma_s|_{[a_{i-1}, a_i]}) \right|_{s=0} = \int_{a_{i-1}}^{a_i} \left\langle \frac{DT}{ds}, T \right\rangle dt \Big|_{s=0}.$$

Using Lemma 21.3, this becomes

$$\left. \frac{d}{ds} L(\Gamma_s|_{[a_{i-1}, a_i]}) \right|_{s=0} = \int_{a_{i-1}}^{a_i} \left\langle \frac{DV}{dt}, T \right\rangle dt.$$

Using integration by parts (using metric compatibility), we obtain

$$\begin{aligned} \left. \frac{d}{ds} L(\Gamma_s|_{[a_{i-1}, a_i]}) \right|_{s=0} &= \int_{a_{i-1}}^{a_i} \left\langle \frac{DV}{dt}, T \right\rangle dt \\ &= \langle V, T \rangle \Big|_{a_{i-1}}^{a_i} - \int_{a_{i-1}}^{a_i} \left\langle V, \frac{DT}{dt} \right\rangle dt \\ &= \langle V(a_i), \dot{\gamma}(a_i^-) \rangle - \langle V(a_{i-1}), \dot{\gamma}(a_{i-1}^+) \rangle - \int_{a_{i-1}}^{a_i} \left\langle V, \frac{D\dot{\gamma}}{dt} \right\rangle dt. \end{aligned} \quad (21.2)$$

Summing over i , we obtain

$$\left. \frac{d}{ds} L(\Gamma_s) \right|_{s=0} = \sum_{i=1}^k (\langle V(a_i), \dot{\gamma}(a_i^-) \rangle - \langle V(a_{i-1}), \dot{\gamma}(a_{i-1}^+) \rangle) - \int_a^b \left\langle V, \frac{D\dot{\gamma}}{dt} \right\rangle dt.$$

We can rearrange the terms to obtain

$$\left. \frac{d}{ds} L(\Gamma_s) \right|_{s=0} = \langle V(b), \dot{\gamma}(b) \rangle - \sum_{i=1}^{k-1} \langle V(a_i), \Delta_i \dot{\gamma} \rangle - \langle V(a), \dot{\gamma}(a) \rangle - \int_a^b \left\langle V, \frac{D\dot{\gamma}}{dt} \right\rangle dt,$$

which is the desired formula. $\square$

21.2. Critical points of L are geodesics. We make the following further definitions.

Definition 21.6. • An admissible curve γ is called a **critical point** of the length functional $L(\cdot)$ if, for all proper admissible variations Γ ,

$$\left. \frac{d}{ds} L(\Gamma_s) \right|_{s=0} = 0.$$

- γ from p to q is **minimizing** if

$$L(\gamma) \leq L(\gamma_1)$$

for all admissible curves γ_1 with

$$\gamma_1(a) = p, \quad \gamma_1(b) = q.$$

- γ is **locally minimizing** if for each $c \in [a, b]$ there exists $\varepsilon > 0$ such that $\gamma|_{[c-\varepsilon, c+\varepsilon] \cap [a, b]}$ is minimizing.

We have the following consequence of Theorem 21.4.

Corollary 21.7. *An admissible curve γ is a critical point of $L_g(\cdot)$ if and only if γ is a smooth geodesic.*

Proof. ($\Leftarrow$) Suppose that γ is a geodesic. Then $\frac{D\dot{\gamma}}{dt} = 0$ on each interval and there is no jump at the partition points. Therefore the right-hand side of the first-variation formula is zero for any proper admissible variation.

($\Rightarrow$) **Step 1.** We first show that

$$\frac{D\dot{\gamma}}{dt} = 0 \quad \text{on } [a, b].$$

(At a point a_i where γ may not be smooth, this refers to both the covariant derivative from the left and from the right.)

Let $c \in (a_{i-1}, a_i)$. Choose a bump function φ such that

$$\varphi(c) \neq 0, \quad \text{supp}(\varphi) \subset (a_{i-1}, a_i).$$

Take

$$V = \varphi D_t \dot{\gamma}, \quad \gamma_s = \exp(sV).$$

This is a natural choice because it tests the first variation directly against the quantity $\frac{D\dot{\gamma}}{dt}$ that we want to prove is zero.

Since γ is a critical point of the length functional,

$$0 = \left. \frac{d}{ds} L(\gamma_s) \right|_{s=0} = - \int_{a_{i-1}}^{a_i} \varphi |D_t \dot{\gamma}|^2 dt.$$

The right-hand side is nonpositive, so the only way it can be 0 is that the integrand must vanish on the support of φ . Hence

$$D_t \dot{\gamma}(c) = 0.$$

Since c was arbitrary, we conclude that

$$D_t \dot{\gamma} = 0 \quad \text{on } (a_{i-1}, a_i).$$

By smoothness, this is also true at the endpoints. Since i was arbitrary, we are done.

Step 2. We now show that

$$\Delta_i \dot{\gamma} = 0 \quad \text{for } i = 1, \dots, k-1.$$

This step checks that the velocity also matches across the break points, so that the geodesic pieces fit together smoothly.

Choose a bump function φ supported on (a_{i-1}, a_{i+1}) such that

$$\varphi(a_i) = 1.$$

We want $\varphi(a_i) = 1$ so that the variation remembers exactly the jump of the velocity at the point a_i . Take

$$V = \varphi \Delta_i \dot{\gamma} \quad (\text{in a coordinate chart}), \quad \gamma_s = \exp(sV).$$

Then by the first variation formula,

$$0 = \left. \frac{d}{ds} L(\gamma_s) \right|_{s=0} = -\langle V(a_i), \Delta_i \dot{\gamma} \rangle.$$

Here the integral term disappears because Step 1 already showed that $D_t \dot{\gamma} = 0$ on each open subinterval, so only the jump term remains. Since $\varphi(a_i) = 1$, we have

$$V(a_i) = \Delta_i \dot{\gamma}.$$

Therefore

$$0 = -\langle \Delta_i \dot{\gamma}, \Delta_i \dot{\gamma} \rangle = -|\Delta_i \dot{\gamma}|^2.$$

Since a norm square is always nonnegative, this implies

$$\Delta_i \dot{\gamma} = 0.$$

Thus the velocity has no jump at any partition point. Together with Step 1, we conclude that γ satisfies

$$\frac{D\dot{\gamma}}{dt} = 0$$

everywhere and is smooth across the partition points. Therefore γ is a smooth geodesic. $\square$

22. GAUSS'S LEMMA AND MINIMIZATION (TH 3/12)

Recall that in general a critical point of a functional need not be an extremum. In general, a geodesic does not always minimize length over all possible paths from p to q , as seen from the example of flying to China through Australia. However, we do have the following characterization.

Theorem 22.1. *An admissible path γ is locally minimizing (see Definition 21.6) if and only if it is a geodesic.*

($\Rightarrow$) This direction is an exercise on HW7.

To prove ($\Leftarrow$), we will show more: namely that radial paths in geodesic balls are minimizing.

Lemma 22.2. (*Gauss*) Let $p \in M$ and let $v, w \in T_p M$. Suppose $v \in \mathcal{E}_p \subset T_p M$ is in the domain of the exponential map. Then

$$\langle (d \exp_p)_v(v), (d \exp_p)_v(w) \rangle_{g_{\exp_p(v)}} = \langle v, w \rangle_{g_p}.$$

In particular, in geodesic coordinates we have

$$\left| \frac{\partial}{\partial r} \right| = 1 \quad \text{and} \quad \frac{\partial}{\partial r} \perp S_r(p) \quad (r > 0).$$

Proof. **Case 1.** $v = w$. Let

$$\exp_p(tv) = \gamma_v(t).$$

Then

$$\gamma'_v(1) = (d \exp_p)_v(v).$$

Since geodesics have constant speed, we have

$$|\gamma'_v(1)|^2 = |\gamma'_v(0)|^2 = |v|^2.$$

Hence

$$\langle (d \exp_p)_v(v), (d \exp_p)_v(v) \rangle = \langle v, v \rangle.$$

Case 2. $\langle v, w \rangle = 0$. Choose a curve

$$\sigma(s) \in T_p M$$

such that

$$\sigma(0) = v, \quad \sigma'(0) = w, \quad |\sigma(s)| = \text{const.}$$

Define

$$\Gamma(s, t) = \exp_p(t\sigma(s)).$$

Then for each fixed s ,

$$\Gamma_s(t)$$

is a geodesic with speed $|v|$.

Let

$$S = \frac{\partial \Gamma}{\partial s}, \quad T = \frac{\partial \Gamma}{\partial t}.$$

Then

$$S(0, 0) = 0, \quad T(0, 0) = v,$$

and

$$\begin{aligned} S(0, 1) &= (d \exp_p)_v(w), \\ T(0, 1) &= (d \exp_p)_v(v). \end{aligned}$$

In particular,

$$\langle S(0, 1), T(0, 1) \rangle = \langle v, w \rangle = 0.$$

We compute

$$\frac{\partial \langle S, T \rangle}{\partial t} = \left\langle \frac{DS}{\partial t}, T \right\rangle + \left\langle S, \frac{DT}{\partial t} \right\rangle.$$

Since the variation is by geodesics, $\frac{DT}{\partial t} = 0$, so

$$\frac{\partial \langle S, T \rangle}{\partial t} = \left\langle \frac{DS}{\partial t}, T \right\rangle \stackrel{\text{L21.3}}{=} \left\langle \frac{DT}{\partial s}, T \right\rangle.$$

Hence

$$= \frac{1}{2} \frac{\partial}{\partial s} |T|^2.$$

But the speed of the geodesic is constant, so

$$|T|^2 = |v|^2,$$

which is independent of s , therefore

$$\frac{\partial \langle S, T \rangle}{\partial t} = \frac{1}{2} \frac{\partial |T|^2}{\partial s} = 0.$$

Integrating in t gives

$$0 = \langle S(0, 1), T(0, 1) \rangle - \langle S(0, 0), T(0, 0) \rangle.$$

Since

$$S(0, 0) = 0, \quad T(0, 0) = v,$$

we obtain

$$\langle S(0, 1), T(0, 1) \rangle = \langle (d \exp_p)_v(w), (d \exp_p)_v(v) \rangle.$$

Thus in the orthogonal case

$$\langle (d \exp_p)_v(v), (d \exp_p)_v(w) \rangle = 0.$$

General case. This follows by observing that both sides of the main formula are linear in w . $\square$

Proposition 22.3. *Suppose $p \in M$ and q is contained in a geodesic ball around p . Then the radial geodesic from p to q is the unique minimizing path from p to q in M .*

Proof. Let

$$B = B_\varepsilon(p) \ni q, \quad \gamma : I = [0, \rho] \rightarrow B$$

be the radial geodesic

$$\gamma(t) = \exp_p(tv), \quad |v| = 1.$$

Consider a competing admissible curve

$$c(t), \quad c(0) = p, \quad c(b) = q.$$

Case 1. Assume

$$\text{Im}(c) \subset B.$$

We can then write

$$c(t) = \exp_p(r(t)v(t)), \quad |v(t)| = 1.$$

By deleting any initial segment that returns to the point p , we may assume WLOG that

$$c(t) \neq p \quad \text{for } t > 0.$$

Then $r(t), v(t)$ are differentiable. With this same choice of $v(t)$, define the two-variable function

$$f(r, t) = \exp_p(rv(t)),$$

so that

$$c(t) = f(r(t), t).$$

By the chain rule, we have

$$c'(t) = \frac{\partial f}{\partial r} r'(t) + \frac{\partial f}{\partial t}.$$

Since

$$|v(t)| = 1,$$

we have

$$\langle v'(t), v(t) \rangle = 0.$$

Applying Gauss's Lemma, we have

$$\langle d(\exp_p)_v(v'(t)), d(\exp_p)_v(v(t)) \rangle = \left\langle \frac{\partial f}{\partial t}, \frac{\partial f}{\partial r} \right\rangle = 0.$$

This gives us

$$\begin{aligned} |c'(t)|^2 &= |r'(t)|^2 \left| \frac{\partial f}{\partial r} \right|^2 + \left| \frac{\partial f}{\partial t} \right|^2 \\ &= |r'(t)|^2 + \left| \frac{\partial f}{\partial t} \right|^2 \\ &\geq |r'(t)|^2, \end{aligned} \tag{22.1}$$

with equality if and only if $\frac{\partial f}{\partial t} \equiv 0$, i.e. $v'(t) \equiv 0$, i.e. the path is radial.

We now take square roots on both sides of (22.1) and integrate, to obtain

$$\int_0^b |c'(t)| dt \geq \int_0^b |r'(t)| dt \geq \int_0^b r'(t) dt = r(b) = \rho,$$

with equality if and only if $c(t)$ is radial. This completes Case 1.

Case 2. Suppose that $c([0, T])$ is not contained in $B_\varepsilon(p)$, and let $q = c(T)$. Let $t_1 \in (0, T]$ be the first time such that

$$c(t_1) \in \partial B_\varepsilon(p) = S_\varepsilon^{n-1}(p).$$

Set $x := c(t_1)$. Since $c([0, t_1]) \subset \overline{B_\varepsilon(p)}$, Case 1 implies that

$$c([0, t_1]) = \gamma_x([0, t_1]),$$

where γ_x is the radial minimizing geodesic from p to x . In particular,

$$L(c|_{[0, t_1]}) = d_g(p, x) = \varepsilon.$$

Now decompose c as

$$c = c|_{[0, t_1]} * c|_{[t_1, T]}.$$

Then

$$L(c) = L(c|_{[0, t_1]}) + L(c|_{[t_1, T]}) \geq \varepsilon + d_g(x, q).$$

Hence

$$L(c) \geq \varepsilon + d_g(x, q).$$

On the other hand, since $q \in B_\varepsilon(p)$, the minimizing radial geodesic γ from p to q has length

$$L(\gamma) = d_g(p, q) < \varepsilon.$$

Therefore

$$L(c) \geq \varepsilon + d_g(x, q) > \varepsilon > d_g(p, q) = L(\gamma).$$

So c cannot be minimizing. This proves that any minimizing curve from p to q must remain entirely inside $B_\varepsilon(p)$, and hence by Case 1 it must coincide with γ . $\square$

23. THE HOPF-RINOW THEOREM (TU 3/17)

Definition 23.1. For $p \in M$ and $r > 0$, the *open geodesic ball* centered at p of radius r is

$$B_r(p) := \{q \in M : d_g(p, q) < r\},$$

provided that $r < \text{Inj}_p(M)$. The *closed geodesic ball* is

$$\overline{B_r(p)} := \{q \in M : d_g(p, q) \leq r\},$$

and the *geodesic sphere* is

$$S_r(p) := \{q \in M : d_g(p, q) = r\}.$$

Remark 23.2. If $d_g(p, q) > \text{Inj}_p(M)$, then the local argument using normal coordinates no longer implies uniqueness of a minimizing geodesic.

Question. Suppose $d_g(p, q) > \text{Inj}_p(M)$. Does a minimizing geodesic from p to q still exist?

Proposition 23.3. Suppose M is connected and every geodesic starting from a fixed point $p \in M$ is complete, that is, defined on $(-\infty, \infty)$. Then for every $q \in M$, there exists a minimizing geodesic from p to q .

Lemma 23.4. Assume that every geodesic in M starting from p is complete. Let γ_i be a sequence of unit speed geodesics with $\gamma_i(0) = p$ for all i . Then there exist a subsequence, still denoted by γ_i , and a geodesic γ starting from p such that $\gamma_i \rightarrow \gamma$ uniformly on every compact interval $[0, r]$.

Proof. Let

$$v_i := \gamma'_i(0) \in S_p M.$$

Since the unit sphere $S_p M \subset T_p M$ is compact, after passing to a subsequence we may assume that

$$v_i \rightarrow v \in S_p M.$$

Let γ be the geodesic determined by the initial conditions

$$\gamma(0) = p, \quad \gamma'(0) = v.$$

By smooth dependence of solutions of the geodesic equation on the initial data, $\gamma_i \rightarrow \gamma$ uniformly on every compact interval. If one wants a single subsequence that works for all $r > 0$, one may use a diagonal argument. $\square$

Proof of Proposition 23.3. Define

$$A := \{q \in M : \text{there exists a minimizing geodesic from } p \text{ to } q\}.$$

We will show that $A = M$.

First, A is nonempty. Indeed, if $\varepsilon < \text{inj}_p(M)$, then every point of $B_\varepsilon(p)$ is joined to p by a unique minimizing geodesic, so

$$B_\varepsilon(p) \subset A.$$

For $r > 0$, let

$$D_r := \{q \in M : d_g(p, q) \leq r\}.$$

We claim that $D_r \cap A$ is compact.

Let $q_i \in D_r \cap A$. For each i , choose a unit speed minimizing geodesic

$$\gamma_i : [0, r_i] \rightarrow M$$

from p to q_i , where

$$r_i = d_g(p, q_i) \leq r.$$

By assumption, each γ_i extends to a complete geodesic on $[0, \infty)$. After passing to a subsequence, we may assume that

$$r_i \rightarrow r_\infty \leq r$$

and, by the lemma, that $\gamma_i \rightarrow \gamma$ uniformly on $[0, r]$ for some geodesic γ starting at p .

Then

$$q_i = \gamma_i(r_i) \rightarrow \gamma(r_\infty) =: q.$$

Since distance is continuous,

$$d_g(p, q) = \lim_{i \rightarrow \infty} d_g(p, q_i) = \lim_{i \rightarrow \infty} r_i = r_\infty.$$

But $\gamma|_{[0, r_\infty]}$ has length r_∞ , so it is minimizing. Therefore $q \in D_r \cap A$. This proves that $D_r \cap A$ is compact for all $r > 0$. Therefore, A is a closed set.

Now suppose for contradiction that $A^c \neq \emptyset$. Since A contains a neighborhood of p and M is connected, the set

$$A \cap \overline{A^c}$$

is nonempty. Define

$$r_0 := \inf\{d_g(p, x) : x \in A \cap \overline{A^c}\}.$$

Then $r_0 > 0$. By the compactness established above, this infimum is attained at some point $q_0 \in A \cap \overline{A^c}$, and

$$d_g(p, q_0) = r_0.$$

Choose a uniformly normal neighborhood W of q_0 , and choose $\delta > 0$ such that $0 < \delta < r_0/2$ and $\overline{B_\delta(q)} \subset W$ for every $q \in W$.

Claim. $W \subset A$.

Proof of claim. Let $q \in W$, and set $r := d_g(p, q)$. Since $q \in W$ and $d_g(p, q_0) = r_0$, we have $r < r_0 + \delta$. Also, since $\delta < r_0/2$,

$$r \geq r_0 - \delta > \delta,$$

so $p \notin \overline{B_\delta(q)}$.

Since $\overline{B_\delta(q)}$ is compact, there exists $z \in \overline{B_\delta(q)}$ such that

$$d_g(p, z) = \min_{x \in \overline{B_\delta(q)}} d_g(p, x).$$

We now show that

$$d_g(p, z) = r - \delta.$$

First, by the triangle inequality,

$$d_g(p, z) + d_g(z, q) \geq d_g(p, q) = r.$$

Since $d_g(z, q) \leq \delta$, it follows that

$$d_g(p, z) \geq r - \delta.$$

Conversely, let σ be any piecewise smooth curve from p to q . Since $p \notin \overline{B_\delta(q)}$ and $q \in B_\delta(q)$, the curve σ must meet the sphere $S_\delta(q)$ at some point y . Hence

$$L(\sigma) \geq d_g(p, y) + d_g(y, q) \geq d_g(p, z) + \delta.$$

Taking the infimum over all such curves σ , we obtain

$$r = d_g(p, q) \geq d_g(p, z) + \delta.$$

Therefore

$$d_g(p, z) = r - \delta.$$

Since $r < r_0 + \delta$, we get

$$d_g(p, z) = r - \delta < r_0.$$

By the definition of r_0 , this implies $z \in A$. Thus there exists a minimizing geodesic from p to z .

Because $z, q \in W$ and W is uniformly normal, there exists a unique minimizing geodesic from z to q , and its length is $d_g(z, q) \leq \delta$. Since $z \in \overline{B_\delta(q)}$, in fact $d_g(z, q) = \delta$. Therefore there exists a piecewise smooth curve from p to q of length

$$d_g(p, z) + d_g(z, q) = (r - \delta) + \delta = r = d_g(p, q).$$

Hence $q \in A$. ◇

Thus $W \subset A$, which contradicts $q_0 \in \overline{A^c}$. Therefore $A^c = \emptyset$, and so $A = M$. □

Definition 23.5. A Riemannian manifold (M, g) is called **geodesically complete** if every maximal geodesic is defined on all of $\mathbb{R}$.

Theorem 23.6 (Hopf–Rinow). *A Riemannian manifold (M, g) is geodesically complete if and only if the metric space (M, d_g) is complete.*

Proof. ($\Rightarrow$) Let $\{q_i\}$ be a Cauchy sequence in (M, d_g) . Without loss of generality, all q_i lie in the same connected component M_0 of M . Fix $p \in M_0$.

Since $\{q_i\}$ is Cauchy, it is bounded, so the numbers $r_i := d_g(p, q_i)$ are bounded. By geodesic completeness, every geodesic starting from p is complete. Hence, by Proposition 23.3, for each i there exists a minimizing unit-speed geodesic

$$\gamma_i : [0, r_i] \rightarrow M$$

from p to q_i .

After passing to a subsequence, we may assume that $r_i \rightarrow r < \infty$, and by the previous lemma that $\gamma_i \rightarrow \gamma$ uniformly on compact intervals for some geodesic γ starting at p . Then

$$q_i = \gamma_i(r_i) \rightarrow \gamma(r) =: q.$$

Thus a subsequence of $\{q_i\}$ converges to q . Since the original sequence is Cauchy, the whole sequence converges to q . Therefore (M, d_g) is metrically complete.

($\Leftarrow$) Assume that (M, d_g) is complete. Suppose for contradiction that (M, g) is not geodesically complete. Then there exists a unit speed maximal geodesic

$$\gamma : [0, b) \rightarrow M$$

with $b < \infty$ which cannot be continued beyond b .

Let $t_i \uparrow b$. Since γ has unit speed,

$$d_g(\gamma(t_i), \gamma(t_j)) \leq |t_i - t_j|,$$

so $\{\gamma(t_i)\}$ is a Cauchy sequence. By metric completeness, $\gamma(t_i) \rightarrow q \in M$.

Let W be a δ -uniformly-normal neighborhood of q . For i sufficiently large, $\gamma(t_i) \in W$. Choose j such that $b - t_j < \delta$ and $t_j \in W$. Since W is δ -uniformly normal, every unit-speed geodesic starting from a point of W exists for time at least δ . In particular, this is true of the geodesic $\sigma(t)$ with $\sigma(0) = \gamma(t_j)$ and $\sigma'(0) = \gamma'(t_j)$. But by uniqueness, σ is just a reparametrization of γ . We may therefore continue γ by setting

$$\tilde{\gamma}(t) = \begin{cases} \gamma(t) & t \in [0, b) \\ \sigma(t - t_j) & t \in (t_j - \delta, t_j + \delta), \end{cases}$$

which is a contradiction.

Therefore (M, g) is geodesically complete. $\square$

Corollary 23.7. *If M is connected and there exists $p \in M$ such that $\mathcal{E}_p = T_p M$, then M is complete.*

Proof. This follows by observing that the ($\Rightarrow$) direction of the proof of Hopf-Rinow only used completeness of all geodesics starting at a point p . $\square$

Corollary 23.8. *If M is metrically complete and connected, then any two points $p, q \in M$ can be joined by a minimizing geodesic.*

Proof. This follows from the Hopf-Rinow Theorem and Proposition 23.3. $\square$

Corollary 23.9. *If M is compact then any Riemannian metric on M is geodesically complete.*

Proof. Any compact metric space is metrically complete, therefore geodesically complete. $\square$

Note: This last fact is also easy to see directly from Corollary 20.2.

Corollary 23.10. *The model spaces $\mathbb{R}^n$, $\mathbb{H}^n$, and $\mathbb{S}^n$ are all metrically complete with respect to d_g .*

Proof. We checked earlier that all geodesics in these spaces are complete. $\square$

Part 6. The Riemann curvature tensor

We have shown that a Riemannian metric g determines a unique torsion-free, metric compatible connection, namely the Levi-Civita connection ∇^g . The next geometric object we want to study is its curvature.

24. DEFINITION AND EXAMPLES (TH 3/19)

Definition 24.1. Let (M, g) be a Riemannian manifold. The *Riemann curvature operator* associated to g is the curvature operator (see Section 7) of the associated Levi-Civita connection, i.e., the map

$$\mathcal{R} : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$$

defined by

$$\mathcal{R}(X, Y)Z = \nabla_X^g \nabla_Y^g Z - \nabla_Y^g \nabla_X^g Z - \nabla_{[X, Y]}^g Z.$$

By Theorem 7.1 $\mathcal{R}$ is tensorial in X, Y , and Z , i.e. is given by an $\text{End}(TM)$ -valued 2-form, i.e. a tensor of type $(1, 3)$.

Working in a coordinate frame $\{\partial_i\}$, the tensor components of $\mathcal{R}$ are determined by

$$\mathcal{R}(\partial_i, \partial_j)\partial_l = R_{ij}{}^k{}_l \partial_k,$$

so that for arbitrary vector fields $X = X^i \partial_i$, $Y = Y^j \partial_j$, $Z = Z^l \partial_l$, we have

$$\mathcal{R}(X, Y)Z = R_{ij}{}^k{}_l X^i Y^j Z^l \partial_k.$$

By Proposition 7.5, we have

$$R_{ij}{}^k{}_l = \partial_i \Gamma_{jl}^k - \partial_j \Gamma_{il}^k + \Gamma_{im}^k \Gamma_{jl}^m - \Gamma_{jm}^k \Gamma_{il}^m. \quad (24.1)$$

Definition 24.2. The *Riemann curvature tensor* Rm is the $(0, 4)$ tensor defined by

$$Rm(X, Y, Z, W) = \langle \mathcal{R}(X, Y)W, Z \rangle.$$

(Notice the change of order in W and Z).

In local coordinates, we have

$$\begin{aligned} Rm(\partial_i, \partial_j, \partial_k, \partial_l) &= \langle \mathcal{R}(\partial_i, \partial_j)\partial_l, \partial_k \rangle \\ &= \langle R_{ij}{}^m{}_l \partial_m, \partial_k \rangle \\ &= R_{ij}{}^m{}_l g_{mk} = R_{ijkl}. \end{aligned} \quad (24.2)$$

Hence, the Riemann curvature tensor is just the Riemann curvature operator with the third index lowered. This justifies using the same letter R for the tensor components of both.

Example 24.3. In (7.1), we determined the curvature operator of the round sphere $\mathbb{S}^n$ to be

$$\mathcal{R}(X, Y)W = \langle Y, W \rangle X - \langle X, W \rangle Y.$$

The Riemann tensor is therefore given by

$$Rm(X, Y, Z, W) = \langle X, Z \rangle \langle Y, W \rangle - \langle X, W \rangle \langle Y, Z \rangle.$$

In components, this reads

$$R_{ijkl}^{\mathbb{S}^n} = g_{ik}g_{jl} - g_{il}g_{jk}$$

as in (7.3) above. This is known as “constant positive sectional curvature,” see below. Indeed, since R_{ijkl} is built out of the metric, by Ricci’s Lemma 13.9.2 and the Leibniz rule, we have

$$\nabla Rm = 0.$$

Let us now check this calculation using the Christoffel symbols in stereographic coordinates. We do this for three reasons: 1. It is always worth checking a calculation via a second method, 2. To demonstrate that brute-force calculation in local coordinates is not always the best method, and 3. Because I assigned you a similar calculation on HW8.

We shall use the formulae (18.1). First observe that if $|\{i, j, k, l\}| = 4$ then no terms will survive in (24.1). If $|\{i, j, k, l\}| = 3$, one has to make a calculation to show that $R_{ijkl} = 0$, which we will skip. Up to permutations, the only nonzero component is the following one (here we do *not* use Einstein summation):

$$\begin{aligned} R_{ij}{}^i{}_j &= \partial_i \Gamma_{jj}^i - \partial_j \Gamma_{ij}^i + \sum_m \Gamma_{im}^i \Gamma_{jj}^m - \Gamma_{jm}^i \Gamma_{ij}^m \\ &= \partial_i \left(\frac{2x^i}{1+|x|^2} \right) + \partial_j \left(\frac{2x^j}{1+|x|^2} \right) + \Gamma_{ii}^i \Gamma_{jj}^i + \Gamma_{ij}^i \Gamma_{jj}^j - \Gamma_{ji}^i \Gamma_{ij}^i - \Gamma_{jj}^i \Gamma_{ij}^j \\ &\quad + \sum_{m \neq i, j} \Gamma_{im}^i \Gamma_{jj}^m - \Gamma_{jm}^i \Gamma_{ij}^m \\ &= \frac{4}{1+|x|^2} - \frac{4}{(1+|x|^2)^2} ((x^i)^2 + (x^j)^2) \\ &\quad + \frac{4}{(1+|x|^2)^2} ((-x^i)x^i + (-x^j)(-x^j) - (-x^j)(-x^j) - x^i(-x^i)) \\ &\quad + \frac{4}{(1+|x|^2)^2} \sum_{k \neq i, j} (-x^k)x^k \\ &= \frac{4}{(1+|x|^2)^2} (1+|x|^2 - |x|^2) \\ &= g_{jj} \\ &= g_{jj} \delta^i{}_i - g_{ij} \delta^i{}_j. \end{aligned} \tag{24.3}$$

This agrees with (7.2).

Example 24.4. For hyperbolic space $\mathbb{H}^2$, the local calculation is much simpler. Referring back to the formulae (18.4), we calculate

$$\begin{aligned} R_{12}{}^1{}_2 &= \partial_1 \Gamma_{22}^1 - \partial_2 \Gamma_{12}^1 + \Gamma_{11}^1 \Gamma_{22}^1 + \Gamma_{12}^1 \Gamma_{22}^2 - \Gamma_{21}^1 \Gamma_{12}^1 - \Gamma_{22}^1 \Gamma_{12}^2 \\ &= -\frac{1}{y^2} + \frac{1}{y^2} - \frac{1}{y^2} \\ &= -\frac{1}{y^2}. \end{aligned} \tag{24.4}$$

Lowering the index, we have

$$R_{1212} = -\frac{1}{y^4} = -g_{11}g_{22}.$$

With some further simple reasoning, one can see that the curvature tensor of hyperbolic n -space must be

$$R_{ijkl}^{\mathbb{H}^n} = -g_{ik}g_{jl} + g_{il}g_{jk}.$$

This is called “constant negative sectional curvature,” see more below.

25. FLAT METRICS (TH 3/19)

Theorem 25.1. *Let (M, g) be a Riemannian manifold. The following are equivalent:*

1. $Rm \equiv 0$ on M .
2. For every $p \in M$, $g_{ij} = \delta_{ij}$ in normal coordinates centered at p .

Equivalently, (M, g) is locally isometric to Euclidean space if and only if its Riemann curvature tensor vanishes identically.

Remark 25.2. A manifold as in the theorem is called a **flat (Riemannian) manifold**. Note that 1. is just the statement that the Levi-Civita connection ∇^g of g is a *flat connection*, in the sense of Subsection 6.3-Theorem 9.1. But since this is the Levi-Civita connection, one gets a stronger conclusion (existence of a coordinate system, not merely of a frame).

First proof. The implication (2) $\Rightarrow$ (1) is immediate. If in some local coordinates $g_{ij} = \delta_{ij}$, then all Christoffel symbols vanish, hence the curvature tensor vanishes.

Now assume $Rm \equiv 0$. Fix $p \in M$. By Theorem 9.1, there exists a local orthonormal frame $\{e_1, \dots, e_n\}$ for TM near p such that $\nabla_{e_\alpha} e_\beta = 0$ for all α, β .

Since the Levi-Civita connection is torsion free,

$$0 = \tau(e_\alpha, e_\beta) = \nabla_{e_\alpha} e_\beta - \nabla_{e_\beta} e_\alpha - [e_\alpha, e_\beta] = -[e_\alpha, e_\beta].$$

In other words, $\{e_\alpha\}$ is a commuting local frame for TM . By Theorem 22.5 of the 761 notes, there exist local coordinates $(x^1, \dots, x^n)$ such that $e_i = \frac{\partial}{\partial x^i}$ for $i = 1, \dots, n$. Moreover, since the frame is orthonormal, we have

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = \langle e_i, e_j \rangle = \delta_{ij}.$$

Therefore the metric is Euclidean in these coordinates. Clearly these are none other than normal coordinates. $\square$

Second proof. We shall give a longer proof via a method that we intend to generalize later.

We show that $\exp_p$ is a local isometry near $0 \in T_p M$. Let $v \in T_p M$, $q = \exp_p(v)$, and let $w \in T_v(T_p M) \cong T_p M$. Assume without loss of generality that $\langle w, v \rangle = 0$. Also assume for notational simplicity that $|v| = 1$. Set

$$X = d(\exp_p)_v(w) \in T_q M.$$

The goal is to determine $|X|_{g_q}$.

We follow the same setup as in the proof of the Gauss Lemma. Choose a smooth curve $\sigma(s)$ in $T_p M$ such that $\sigma(0) = v$, $\sigma'(0) = w$, and $|\sigma(s)| \equiv 1$. Define a variation by radial geodesics

$$\Gamma(s, t) = \exp_p(t \sigma(s)).$$

Let $T = \partial_t \Gamma$ and $S = \partial_s \Gamma$. For each fixed s , the curve $t \mapsto \Gamma(s, t)$ is a geodesic, so $D_t T = 0$. We have $\langle T, T \rangle \equiv 1$ since this is a variation by unit-speed geodesics. Moreover, By Gauss's lemma, $\langle S, T \rangle = 0$.

We now calculate the second covariant derivative of S in the t -direction:

$$D_t^2 S = D_t D_t S = D_t D_s T$$

by Lemma 21.3. Now, by definition of the curvature, we have

$$[D_t, D_s] = \mathcal{R}(T, S) = 0$$

since $Rm_g \equiv 0$ and S, T are pushforwards of commuting vector fields (see Lemma 36.1 below for a careful proof). Hence, we can commute D_t and D_s in the above, to obtain

$$D_t^2 S = D_t D_s T = D_s D_t T = 0$$

since $D_t T \equiv 0$.

We have shown that $D_t S$ is covariantly constant in the t direction. We therefore have

$$|D_t S(0, t)| \equiv |D_t S(0, 0)| = 1.$$

This gives

$$\partial_t^2 |S|^2 = 2 \partial_t \langle D_t S, S \rangle = 2 (\langle D_t^2 S, S \rangle + |D_t S|^2) = 2.$$

We clearly have $|S(0, 0)|^2 = 0 = \partial_t |S(0, t)|^2$. So we may conclude that

$$|S(0, t)|^2 = t^2.$$

Evaluating at $t = 1$, we get

$$|X|_{g_q}^2 = |S(0, 1)|^2 = 1 = |w|^2.$$

Since w was arbitrary, we have shown that $d(\exp_p)_v$ preserves lengths. By polarizing, we can conclude that it preserves the metric. $\square$

26. THE BIANCHI IDENTITIES (TH 3/19-TU 3/24)

26.1. Symmetries of the Riemann tensor.

Theorem 26.1. *The Riemann curvature tensor always obeys the following symmetries:*

$$R_{ijkl} = -R_{jikl} \tag{1}$$

$$R_{ijkl} = -R_{ijlk} \tag{2}$$

$$R_{ijkl} = R_{klij} \tag{3}$$

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0. \tag{4}$$

Remark 26.2. The first identity follows directly from the definition of curvature and the second one follows from tensoriality of curvature and metric compatibility. However, the third and fourth are not obvious; especially the fourth, which is called the **algebraic Bianchi identity**.

Remark 26.3. In dimension 2, these symmetries imply that the curvature has essentially only one independent component, namely R_{1212} . Hence, the Riemann curvature tensor carries only the information of a single scalar quantity, which we will see is proportional to the Gaussian curvature.

Proof. Let $p \in M$; we will prove the identities at p . In normal coordinates centered at p , we have

$$g_{ij}(p) = \delta_{ij}, \quad \Gamma_{ij}^k(p) = 0, \quad \partial_i g_{jk}(p) = \partial_i g^{jk}(p) = 0.$$

Then

$$\begin{aligned} R_{ijkl}(p) &= \partial_i \Gamma_{j\ell}^k(p) - \partial_j \Gamma_{i\ell}^k(p) + \text{commutator at } p \\ &= \frac{1}{2} \partial_i (g^{kn} (\partial_j g_{\ell n} + \partial_\ell g_{jn} - \partial_n g_{j\ell})) - \frac{1}{2} \partial_j (g^{kn} (\partial_i g_{\ell n} + \partial_\ell g_{in} - \partial_n g_{i\ell}))(p). \end{aligned}$$

Passing the derivatives through at p , we obtain the following useful formula for the Riemann tensor in geodesic coordinates:

$$\boxed{R_{ijkl}(p) = \frac{1}{2} \begin{pmatrix} \partial_i \partial_\ell g_{jk} - \partial_i \partial_k g_{j\ell} \\ + \partial_j \partial_k g_{i\ell} - \partial_j \partial_\ell g_{ik} \end{pmatrix} (p).} \quad (26.1)$$

The identities (1)-(3) can be checked by inspecting this expression. For (4), we just have to write it out:

$$\begin{aligned} (R_{ijkl} + R_{jkil})(p) &= \frac{1}{2} \begin{pmatrix} \partial_i \partial_\ell g_{jk} - \partial_i \partial_k g_{j\ell} \\ + \partial_j \partial_k g_{i\ell} - \partial_j \partial_\ell g_{ik} \end{pmatrix} (p) + \frac{1}{2} \begin{pmatrix} \partial_j \partial_\ell g_{ki} - \partial_j \partial_i g_{k\ell} \\ + \partial_k \partial_i g_{j\ell} - \partial_k \partial_\ell g_{ji} \end{pmatrix} (p) \\ &= \frac{1}{2} \begin{pmatrix} \partial_i \partial_\ell g_{jk} - \partial_i \partial_j g_{k\ell} \\ + \partial_k \partial_j g_{i\ell} - \partial_k \partial_\ell g_{ij} \end{pmatrix} (p) \\ &= R_{ikjl}(p) \\ &= -R_{kijl}(p). \end{aligned}$$

□

Alternate proof of (4). Without loss of generality, assume that X, Y , and Z commute. Then:

$$\begin{aligned}
\mathcal{R}(X, Y)Z + \mathcal{R}(Y, Z)X + \mathcal{R}(Z, X)Y &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z \\
&\quad + \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X \\
&\quad + \nabla_Z \nabla_X Y - \nabla_X \nabla_Z Y \\
&= \nabla_X (\nabla_Y Z - \nabla_Z Y) \\
&\quad + \nabla_Y (\nabla_Z X - \nabla_X Z) \\
&\quad + \nabla_Z (\nabla_X Y - \nabla_Y X) \\
&= \nabla_X [Y, Z] + \nabla_Y [Z, X] + \nabla_Z [X, Y] \\
&= 0.
\end{aligned}$$

□

26.2. The differential Bianchi identity, revisited. Now recall from Theorem 12.1 that the curvature operator of any connection satisfies the *Differential Bianchi identity*:

$$\begin{aligned}
&\nabla_X \mathcal{R}(Y, Z) + \nabla_Y \mathcal{R}(Z, X) + \nabla_Z \mathcal{R}(X, Y) \\
&+ \mathcal{R}(X, [Y, Z]) + \mathcal{R}(Y, [Z, X]) + \mathcal{R}(Z, [X, Y]) = 0.
\end{aligned}$$

Here ∇_X is the covariant derivative of $\mathcal{R}(Y, Z)$ as an element of $\Gamma(\text{End}(TM))$. With the Levi-Civita connection, we can now reformulate the Differential Bianchi identity in a more commonly-used form. We use the notation

$$\nabla Rm(U, X, Y, Z, W) := (\nabla_U Rm)(X, Y, Z, W),$$

where $\nabla_U Rm$ means the covariant derivative as a $(0, 4)$ -tensor. We have:

Theorem 26.4 (Differential Bianchi identity for Rm).

$$\boxed{\nabla Rm(U, X, Y, Z, W) + \nabla Rm(X, Y, U, Z, W) + \nabla Rm(Y, U, X, Z, W) = 0.}$$

Proof. Since we can work with the coordinate vector fields in geodesic normal coordinates, it suffices to check the identity for commuting vector fields all of which are covariantly constant at a point p :

$$\nabla_U X(p) = \nabla_U Y(p) = \dots = 0.$$

Then, by the Leibniz rule:

$$\begin{aligned}
U(Rm(X, Y, Z, W))(p) &= (\nabla_U Rm)(X, Y, Z, W)(p) + Rm(\nabla_U X(p), Y, Z, W) + \\
&\quad + Rm(X, \nabla_U Y(p), Z, W) + Rm(X, Y, \nabla_U Z(p), W) + Rm(X, Y, Z, \nabla_U W(p)) \\
&= (\nabla_U Rm)(X, Y, Z, W)(p).
\end{aligned}$$

However, we also have

$$U(Rm(X, Y, Z, W)) = U(\langle \mathcal{R}(X, Y)W, Z \rangle)(p) = \langle \nabla_U (\mathcal{R}(X, Y))W, Z \rangle(p).$$

Therefore,

$$\begin{aligned}
\nabla Rm(U, X, Y, Z, W) + \text{cyclic in } U, X, Y &= \langle \nabla_U (\mathcal{R}(X, Y))W, Z \rangle + \text{cyclic in } U, X, Y \\
&= \langle (\nabla_U (\mathcal{R}(X, Y)) + \text{cyclic})W, Z \rangle = 0,
\end{aligned}$$

by the Differential Bianchi identity, as the vector fields all commute. $\square$

Remark 26.5. The above identities are in fact all of those satisfied by Rm and ∇Rm at a point p . In other words, any two tensors at a point satisfying those identities are in fact the curvature of some metric and its first derivative. I've never sat down and proved this, but I think it's actually not too hard.

Remark 26.6. In physics, you may encounter these identities in the form

$$\begin{aligned} R_{[ijk]\ell} &= 0 \text{ (algebraic Bianchi identity)} \\ \nabla_{[i} R_{jk]\ell m} &= 0 \text{ (differential Bianchi identity).} \end{aligned}$$

26.3. Reformulation of the algebraic identities. We can reformulate the identities (1)-(4) as follows. The Riemann tensor is given locally by

$$Rm = R_{ijkl} dx^i \otimes dx^j \otimes dx^k \otimes dx^\ell.$$

The identities (1)-(3) are equivalent to the statement

$$Rm \in \Gamma(\text{Sym}^2(\Lambda^2 M)) \subset T^{(0,4)}(M).$$

Now consider the map

$$\begin{aligned} \Theta : \text{Sym}^2(\Lambda^2 M) &\rightarrow \Lambda^4 M \\ \alpha \cdot \beta &\mapsto \alpha \wedge \beta. \end{aligned}$$

This map is well-defined, because $\beta \cdot \alpha$ is mapped to $\beta \wedge \alpha = (-1)^{2 \cdot 2} \alpha \wedge \beta = \alpha \wedge \beta$. Also, it is not hard to check that Θ is surjective. By the 4th identity above (the algebraic Bianchi identity), we have:

$$\begin{aligned} \Theta(Rm) &= R_{ijk\ell} dx^i \wedge dx^j \wedge dx^k \wedge dx^\ell \\ &= \frac{1}{3} (R_{ijk\ell} + R_{jkil} + R_{kij\ell}) dx^i \wedge dx^j \wedge dx^k \wedge dx^\ell \\ &= 0. \end{aligned}$$

In other words, the algebraic Bianchi identity is equivalent to the statement

$$\boxed{Rm \in \ker \Theta =: \mathcal{R}.}$$

This is perhaps a more attractive way to think about the symmetries of the Riemann tensor.

The space $\mathcal{R}$ is called the *space of algebraic curvature tensors*. Using the rank-nullity theorem, we calculate

$$\begin{aligned} \dim \mathcal{R} &= \frac{\binom{n}{2} \left(\binom{n}{2} + 1 \right)}{2} - \binom{n}{4} \\ &= \frac{n(n-1)}{4} \left(\frac{n(n-1)}{2} + 1 \right) - \frac{n(n-1)(n-2)(n-3)}{24} \\ &= \frac{n^2(n^2-1)}{12}. \end{aligned}$$

In particular, for $n = 1$, $\dim(\mathcal{R}) = 0$. For $n = 2$, $\dim(\mathcal{R}) = 1$ (as we mentioned earlier, the only nonzero component for Rm up to permutations is R_{1212}). For $n = 3$, the dimension is 6 and for $n = 4$, it equals 20, and increases rapidly thereafter.

For this reason, we want to break the Riemann tensor into “more digestible” pieces.

27. SECTIONAL CURVATURE (TU 3/24)

Definition/Lemma 27.1. *For a 2-plane $\Pi \subset T_p M$, the **sectional curvature** of Π is given by*

$$\sec_p(\Pi) = \frac{Rm(v, w, v, w)}{|v \wedge w|^2},$$

where $\{v, w\}$ is any basis for Π and $|v \wedge w| = \sqrt{|v|^2|w|^2 - \langle v, w \rangle^2}$.

Proof. Let $\{e_1, e_2\}$ be any orthonormal basis for Π and

$$v = ae_1 + be_2$$

$$w = ce_1 + de_2$$

any other basis of Π . Then

$$|v \wedge w| = \sqrt{(a^2 + b^2)(c^2 + d^2) - (ac + bd)^2} = \sqrt{a^2d^2 + b^2c^2 - 2abcd} = |ad - bc|$$

Also,

$$Rm(v, w, v, w) = Rm(ae_1 + be_2, ce_1 + de_2, v, w) = (ad - bc)Rm(e_1, e_2, v, w),$$

where we used the skew-symmetry in the first 2 entries. Similarly, by the skew-symmetry in the last 2 entries,

$$Rm(e_1, e_2, v, w) = (ad - bc)Rm(e_1, e_2, e_1, e_2).$$

Overall

$$\frac{Rm(v, w, v, w)}{|v \wedge w|^2} = \frac{(ad - bc)^2 Rm(e_1, e_2, e_1, e_2)}{(ad - bc)^2} = Rm(e_1, e_2, e_1, e_2)$$

which proves well-definedness. □

Example 27.2. For $M = S^n$, we have

$$Rm(e_1, e_2, e_1, e_2) = \langle e_1, e_1 \rangle \langle e_2, e_2 \rangle - \langle e_1, e_2 \rangle \langle e_2, e_1 \rangle = 1$$

for any two orthogonal vectors e_1, e_2 . In particular, we have

$$\sec_p(\Pi) \equiv 1$$

for all points p and planes Π . In other words, *the sectional curvature is constant and equal to 1.*

Example 27.3. Similarly, hyperbolic space $\mathbb{H}^n$ has *constant sectional curvature* -1 .

Proposition 27.4. *The sectional curvatures at p determine the full Riemann tensor at p .*

Proof. (Omitted in class) Let $R_1, R_2 \in \mathcal{R}$ be algebraic curvature tensors. The claim is that if $R_1(x, y, x, y) = R_2(x, y, x, y)$ for all vectors x and y , then $R_1 = R_2$. Let

$$D = R_1 - R_2 \in \mathcal{R}.$$

We are assuming that $D(x, y, x, y) = 0$ for all x, y , and we need to prove that $D \equiv 0$.

Claim. For all x, v, w , $D(x, v, x, w) = 0$.

Proof of claim. We have

$$0 = D(x, v + w, x, v + w) = \cancel{D(x, v, x, v)}^0 + \cancel{D(x, w, x, w)}^0 + D(x, v, x, w) + D(x, w, x, v).$$

Exchanging the first and last pairs, we get

$$0 = 2D(x, v, x, w).$$

◇

Claim. $D(x, v, u, w) = -D(u, v, x, w)$.

Proof of claim. By the previous claim, we have

$$0 = D(x + u, v, x + u, w) = \cancel{D(x, v, x, w)}^0 + \cancel{D(u, v, u, w)}^0 + D(x, v, u, w) + D(u, v, x, w).$$

Rearranging gives the claim.

◇

To complete the proof, we apply the algebraic Bianchi identity:

$$0 = D(x, u, v, w) + D(u, v, x, w) + D(v, x, u, w). \quad (27.1)$$

We have

$$D(u, v, x, w) = -D(v, u, x, w) \stackrel{\text{Claim 2}}{=} D(x, u, v, w)$$

and

$$D(v, x, u, w) \stackrel{\text{Claim 2}}{=} -D(u, x, v, w) = D(x, u, v, w).$$

So the three terms in (27.1) are equal, and we get

$$0 = 3D(x, u, v, w).$$

□

28. RICCI AND SCALAR CURVATURE (TU 3/24)

Definition 28.1. The **Ricci curvature** is given by

$$Rc_p(X, Y) = \sum_i Rm(e_i, X, e_i, Y),$$

where $\{e_i\}$ is any orthonormal basis for T_pM .

Note that Rc is just the trace of Rm in the 1st and 3rd indices, i.e. Rc is independent of the choice of the basis $\{e_i\}$ and thus well-defined. In local coordinates, we write

$$Rc(\partial_i, \partial_j) =: R_{ij} = g^{k\ell} R_{kilj} = R_{ki}{}^k{}_j.$$

Notice that by the identity (3) above, we have

$$R_{ij} = R_{ji},$$

in other words Rc is symmetric.

To help interpret the Ricci curvature, we observe that $Rc_p(X, X)$ is the sum of the sectional curvature over all planes containing X_p . Since Rc is symmetric, $Rc(X, Y)$ is determined by $Rc(X, X)$ by polarization.

Definition 28.2. The **scalar curvature** is given by

$$S = \sum_i Rc(e_i, e_i) = g^{ij} R_{ij} = g^{ij} R_{ki}{}^k{}_j.$$

Example 28.3. You showed on homework that

$$Rc^{\mathbb{S}^n} = (n-1)g, \quad S^{\mathbb{S}^n} = n(n-1).$$

We have seen that for $\dim(\Sigma) = 2$, the Riemann tensor is determined by the single number R_{1212} . In terms of sectional curvature, there is only one choice of 2-plane to make at any point, namely $\Pi = T_pM$. We can therefore speak unambiguously about the *curvature* of a Riemannian 2-manifold, which is a single number:

Definition 28.4. Suppose $\dim(\Sigma) = 2$. The **(Gauss) curvature** of Σ at p is given by

$$K_\Sigma(p) = \sec_p(T_p\Sigma) = \frac{S(p)}{2}.$$

We will see below that in the case that Σ is embedded in $\mathbb{R}^3$, this agrees with the Gauss curvature defined in Lecture 1.

Part 7. Submanifolds and Lie groups

Having defined and described the Riemann tensor and computed it for the model spaces, we now want to compute it in some more cases. We first do the case of submanifolds, then of invariant metrics on Lie groups.

29. THE GAUSS EQUATION (TH 3/26)

29.1. The second fundamental form. Let (M, g) be an (embedded) Riemannian submanifold of an ambient Riemannian manifold $(\tilde{M}, \tilde{g})$; i.e. g is the induced metric by $\tilde{g}$. Recall from Proposition 15.4 that for the respective Levi-Civita connections, it holds

$$\nabla^g = \pi \circ \nabla^{\tilde{g}},$$

where π is the orthogonal projection onto the tangent space of M .

Consider the *ambient tangent bundle* $T\tilde{M}|_M$. If $X \in \Gamma(T\tilde{M}|_M)$, we can extend it to a vector field $\tilde{X}$ in a neighborhood of $M \subset \tilde{M}$ (this can be proved using partitions of unity). The metric $\tilde{g}$ on $\tilde{M}$ restricts to a bundle metric on $T\tilde{M}|_M$.

Define the *normal bundle*

$$NM := TM^\perp \subset T\tilde{M}|_M.$$

We clearly have

$$T\tilde{M}|_M = TM \oplus NM.$$

Lemma 29.1. *On $T\tilde{M}|_M$, there exist orthonormal local frames of the form*

$$\left\{ \underbrace{E_1, \dots, E_n}_{\text{o.n. frame for } TM}, \underbrace{E_{n+1}, \dots, E_m}_{\text{o.n. frame for } NM} \right\}.$$

Remark 29.2. Note that the term “adapted frame” doesn’t have the same meaning as above.

Proof. Apply Gram-Schmidt, first on the initial n -tuple and then on the rest. $\square$

Now recall that covariant differentiation is independent of the choice of extensions. In other words, if $X, Y \in \mathfrak{X}(M)$ and $\tilde{X}, \tilde{Y} \in \mathfrak{X}(\tilde{M})$ are any two respective extensions, then $\tilde{\nabla}_{\tilde{X}}\tilde{Y}|_M$ is independent of the choice of $\tilde{X}, \tilde{Y}$, so we can unambiguously write

$$\tilde{\nabla}_X Y = \tilde{\nabla}_{\tilde{X}} \tilde{Y}.$$

Also, let

$$\tilde{\nabla}_X Y = (\tilde{\nabla}_X Y)^\top + (\tilde{\nabla}_X Y)^\perp,$$

where $(\tilde{\nabla}_X Y)^\top \in TM$ and $(\tilde{\nabla}_X Y)^\perp \in NM$.

Definition 29.3. The **second fundamental form** of M inside $\tilde{M}$ is

$$\boxed{\text{II}(X, Y) := (\tilde{\nabla}_X Y)^\perp.}$$

Remark 29.4. The metric tensor g of an embedded submanifold used to be known as the “first fundamental form.”

Lemma 29.5. *II is bilinear and symmetric.*

Proof. Choose any extensions $\tilde{X}, \tilde{Y}$. Then by torsion-freeness,

$$\begin{aligned} \text{II}(X, Y) - \text{II}(Y, X) &= (\tilde{\nabla}_{\tilde{X}} \tilde{Y})^\perp - (\tilde{\nabla}_{\tilde{Y}} \tilde{X})^\perp \\ &= (\tilde{\nabla}_{\tilde{X}} \tilde{Y} - \tilde{\nabla}_{\tilde{Y}} \tilde{X})^\perp \\ &= [\tilde{X}, \tilde{Y}]^\perp = [X, Y]^\perp = 0. \end{aligned}$$

So II is symmetric. Since II is linear in X by definition, it is linear in Y as well. $\square$

Example 29.6. $M = \mathbb{S}^n \subset \mathbb{R}^{n+1}$. Then

$$\begin{aligned} 0 &= \tilde{\nabla}_X \langle \vec{r}, Y \rangle \\ &= \langle \tilde{\nabla}_X \vec{r}, Y \rangle + \langle \vec{r}, \tilde{\nabla}_X Y \rangle \\ &= \langle X, Y \rangle + \langle \vec{r}, \nabla_X Y \rangle \\ \implies \langle \vec{r}, \tilde{\nabla}_X Y \rangle &= -\langle X, Y \rangle. \end{aligned}$$

Therefore,

$$\text{II}(X, Y) = (\tilde{\nabla}_X Y)^\perp = \langle \vec{r}, \tilde{\nabla}_X Y \rangle \vec{r} = -\langle X, Y \rangle \vec{r}.$$

In other words,

$$\boxed{\text{II}^{\mathbb{S}^n} = -g\vec{r}}.$$

Lemma 29.7 (Gauss's formula). $\boxed{\tilde{\nabla}_X Y = \nabla_X Y + \text{II}(X, Y)}$

Proof. By definition. $\square$

Example 29.8. In our previous example $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, this becomes

$$\nabla_X Y = X(Y) + \langle X, Y \rangle \vec{r},$$

which is just the formula of Lemma 7.4.1!

It turns out that the second fundamental form also determines the covariant derivatives of normal vector fields.

Lemma 29.9 (Weingarten equation). *Let $X, Y \in \mathfrak{X}(M)$, $N \in \Gamma(NM)$. Then*

$$\boxed{\langle \tilde{\nabla}_X N, Y \rangle = -\langle N, \text{II}(X, Y) \rangle}.$$

Proof. Extend X, Y, N in a neighborhood of $\tilde{M}$. Then

$$0 = X \langle N, Y \rangle = \langle \tilde{\nabla}_X N, Y \rangle + \langle N, \tilde{\nabla}_X Y \rangle.$$

Since only the normal component survives in the second term, we have

$$\langle \tilde{\nabla}_X N, Y \rangle = -\langle N, (\tilde{\nabla}_X Y)^\perp \rangle = -\langle N, \text{II}(X, Y) \rangle.$$

$\square$

Here is a related object that we could have introduced much sooner.

Definition 29.10. The **Hessian** of a smooth function $f : M \rightarrow \mathbb{R}$ is

$$\boxed{\text{Hess} f = \nabla df}$$

as a $(0, 2)$ -tensor.

Lemma 29.11. $\text{Hess } f$ is symmetric, i.e. $\text{Hess}_{X,Y} f = \text{Hess}_{Y,X} f$.

Proof. By the Leibniz rule,

$$\text{Hess}_{X,Y} f = (\nabla df)(X, Y) = \nabla_X(df(Y)) - df(\nabla_X Y) = X(Yf) - (\nabla_X Y)(f)$$

At the same time,

$$\text{Hess}_{Y,X} f = Y(Xf) - (\nabla_Y X)(f).$$

Therefore,

$$\text{Hess}_{X,Y} f - \text{Hess}_{Y,X} f = [X, Y]f - (\nabla_X Y - \nabla_Y X)(f) = 0,$$

by torsion-freeness. $\square$

Remark 29.12. It is important for your general knowledge to know that while the Hessian at a general point depends on the Levi-Civita connection, at a *critical* point of f , it does not. Indeed, if $df(p) = 0$, then by definition we have

$$(\text{Hess}_{X,Y} f)(p) = X(Yf)(p).$$

In other words, the Hessian is just the “matrix of second partials,” which in fact transforms as a (pointwise) $(0, 2)$ -tensor under any coordinate change. In this sense it is sometimes said that “the Hessian at a critical point does not depend on the coordinate system.” (Really this means it does not depend on whether or not the metric is Euclidean in the coordinate system.)

Example 29.13. Suppose $M = \{F = 0\} \subset \tilde{M}$ is a smooth hypersurface, i.e. F vanishes transversely. Then $N = \tilde{\nabla} F$ is a normal vector field to M . For X, Y tangent to M , the Weingarten equation yields

$$\text{Hess}_{X,Y} F = (\tilde{\nabla} dF)(X, Y) = \langle \tilde{\nabla}_X \tilde{\nabla} F, Y \rangle = -\langle \tilde{\nabla} F, \text{II}(X, Y) \rangle.$$

Therefore

$$\boxed{\text{II}(X, Y) = -\text{Hess}_{X,Y} F \frac{\tilde{\nabla} F}{|\tilde{\nabla} F|^2}.} \quad (29.1)$$

Example 29.14. Consider the case that $M \subset \mathbb{R}^{n+1}$ is the graph of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, i.e. the zero set of the function

$$F(x^1, \dots, x^{n+1}) := x^{n+1} - f(x^1, \dots, x^n).$$

We have

$$\tilde{\nabla} F = \frac{\partial}{\partial x^{n+1}} - \nabla f$$

and

$$\text{Hess } F = -\text{Hess } f.$$

Here we think of f as a function of $n+1$ variables which is independent of x^{n+1} . The formula (29.1) reads

$$\text{II}(X, Y) = \text{Hess}_{X,Y} f \frac{\tilde{\nabla} F}{|\tilde{\nabla} F|^2} = \text{Hess } f \frac{\frac{\partial}{\partial x^{n+1}} - \nabla f}{1 + |\nabla f|^2}. \quad (29.2)$$

Note that if we rotate our coordinate system so that p is a critical point

$$df(p) = 0, \quad (29.3)$$

then the normal vector is

$$\tilde{\nabla} F(p) = \frac{\partial}{\partial x^{n+1}}$$

and we have

$$\text{II}(X, Y)(p) = (\text{Hess}_{X, Y} f)(p) \frac{\partial}{\partial x^{n+1}}.$$

So, up to rotations, the second fundamental form of a graph of a function is just its Hessian. Note that by Remark 29.12, assuming (29.3), this is just the matrix of second partials.

29.2. Curvature of a submanifold. We can now compare the curvature of the submanifold with that of the ambient manifold. It turns out that the difference depends only on the second fundamental form.

Theorem 29.15 (Gauss Equation).

$$\tilde{Rm}(X, Y, Z, W) = Rm(X, Y, Z, W) - \langle \text{II}(X, Z), \text{II}(Y, W) \rangle + \langle \text{II}(X, W), \text{II}(Y, Z) \rangle$$

Proof. Assume, without loss of generality, that the vector fields are defined in a neighborhood of M (if not, extend). Applying the Gauss formula, we have

$$\begin{aligned} \tilde{Rm}(X, Y, Z, W) &= \langle \tilde{\nabla}_X \tilde{\nabla}_Y W - \tilde{\nabla}_Y \tilde{\nabla}_X W - \tilde{\nabla}_{[X, Y]} W, Z \rangle \\ &= \left\langle \begin{aligned} &\tilde{\nabla}_X (\nabla_Y W + \text{II}(Y, W)) \\ &- \tilde{\nabla}_Y (\nabla_X W + \text{II}(X, W)) \\ &- (\nabla_{[X, Y]} W + \text{II}([X, Y], W)) \end{aligned}, Z \right\rangle. \end{aligned} \quad (29.4)$$

We have $\langle \text{II}([X, Y], W), Z \rangle = 0$, since II outputs a normal vector field. Now, we may apply the Weingarten equation on the first term,

$$\langle \tilde{\nabla}_X \text{II}(Y, W), Z \rangle = -\langle \text{II}(Y, W), \text{II}(X, Z) \rangle,$$

and similarly on the second term. We get

$$\begin{aligned} \tilde{Rm}(X, Y, Z, W) &= \langle \tilde{\nabla}_X \overset{\nabla_X}{\nearrow} \tilde{\nabla}_Y W, Z \rangle - \langle \text{II}(Y, W), \text{II}(X, Z) \rangle \\ &\quad - \langle \tilde{\nabla}_Y \overset{\nabla_Y}{\nearrow} \tilde{\nabla}_X W, Z \rangle + \langle \text{II}(X, W), \text{II}(Y, Z) \rangle \\ &\quad - \langle \nabla_{[X, Y]} W, Z \rangle \\ &= Rm(X, Y, Z, W) - \langle \text{II}(Y, W), \text{II}(X, Z) \rangle + \langle \text{II}(X, W), \text{II}(Y, Z) \rangle, \end{aligned}$$

as desired. □

Example 29.16. For $M = \mathbb{S}^n$, we computed $\text{II} = -g\tilde{r}$. Since $\tilde{Rm} = 0$, we get

$$Rm(X, Y, Z, W) = \langle X, Z \rangle \langle Y, W \rangle - \langle X, W \rangle \langle Y, Z \rangle.$$

which agrees with what we have seen many times above.

Example 29.17. Consider the graph of $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, i.e. the zero set of $F(x, y, z) = z - f(x, y)$. By (29.2), we have $\Pi(X, Y) = \text{Hess}_{X,Y} f \frac{\tilde{\nabla} F}{|\tilde{\nabla} F|^2}$. Applying our calculation (29.4), we have

$$\begin{aligned} Rm(X, Y, X, Y) &= \langle \Pi(X, X), \Pi(Y, Y) \rangle - \langle \Pi(X, Y), \Pi(X, Y) \rangle \\ &= \frac{\text{Hess}_{X,X} f \cdot \text{Hess}_{Y,Y} f - (\text{Hess}_{X,Y} f)^2}{|\tilde{\nabla} F|^2}. \end{aligned}$$

We have $\tilde{\nabla} F = \hat{k} - f_x \hat{i} - f_y \hat{j}$. Let $\sigma(x, y) = (x, y, f(x, y))$ be the standard parameterization with tangent vectors $X = \sigma_* \partial_x$ and $Y = \sigma_* \partial_y$. Then

$$|X \wedge Y| = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = \sqrt{1 + f_x^2 + f_y^2} = |\tilde{\nabla} F|.$$

Therefore, for the sectional curvature, we have (where $\Pi = T_p M = \text{Span}\{X_p, Y_p\}$)

$$\sec_p(\Pi) = \frac{Rm(X, Y, X, Y)}{|X \wedge Y|^2} = \frac{\text{Hess}_{X,X} f \cdot \text{Hess}_{Y,Y} f - (\text{Hess}_{X,Y} f)^2}{|\nabla F|^4}$$

and

$$\boxed{\sec_p(\Pi) = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2}}.$$

This is a useful special case of the classical formulae that we will give in Section 32.

30. GEOMETRIC INTERPRETATION OF SECTIONAL CURVATURE (Tu 4/7)

We can use the Gauss equation to elucidate the meaning of sectional curvature, as follows:

Proposition 30.1 (Geometric interpretation of sectional curvature). *Let (M, g) be a Riemannian manifold, let $p \in M$, and let $\Pi \subset T_p M$ be a 2-plane. Set $\Sigma := \exp_p(\Pi)$. Then*

$$\boxed{K_\Sigma(p) = \sec_p(\Pi)}.$$

Proof. By the Gauss equation, it suffices to show that $\Pi_p = 0$, where Π is the second fundamental form of the embedded surface $\Sigma \subset M$. Let $v \in \Pi$, and let $\gamma_v \subset \Sigma$ be the geodesic with $\gamma_v(0) = p$ and $\dot{\gamma}_v(0) = v$. Since γ_v is a geodesic in the ambient manifold,

$$0 = \tilde{\nabla}_{\dot{\gamma}_v} \dot{\gamma}_v = \nabla_{\dot{\gamma}_v} \dot{\gamma}_v + \Pi(\dot{\gamma}_v, \dot{\gamma}_v),$$

by the Gauss formula. Since the first term is tangent to Σ and the second is normal to Σ , they must vanish individually:

$$\nabla_{\dot{\gamma}_v} \dot{\gamma}_v = 0, \quad \Pi(\dot{\gamma}_v, \dot{\gamma}_v) = 0.$$

Evaluating at $t = 0$ gives $\Pi_p(v, v) = 0$. Since $v \in \Pi$ was arbitrary, by polarization, we conclude that $\Pi_p = 0$. Hence $K_\Sigma(p) = \sec_p(\Pi)$. $\square$

31. HYPERSURFACES AND THE THEOREMA EGREGIUM (TU 4/7)

31.1. Hypersurfaces and the shape operator. We now specialize to the case that M is a hypersurface, i.e. a submanifold of codimension one:

$$(M^n, g) \subset (\widetilde{M}^{n+1}, \widetilde{g}).$$

We can then choose a unit normal vector field, N , which is unique up to sign.

Definition 31.1. The **scalar second fundamental form** is

$$h(X, Y) := \langle \text{II}(X, Y), N \rangle.$$

Thus $\text{II}(X, Y) = h(X, Y)N$.

Definition 31.2. The **shape operator** $s = h^\sharp : TM \rightarrow TM$ is defined by

$$h(X, Y) = \langle X, sY \rangle = \langle sX, Y \rangle.$$

Since h is symmetric, s is self-adjoint.

In terms of h , the Gauss formula becomes

$$\widetilde{\nabla}_X Y = \nabla_X Y + h(X, Y)N.$$

Using $\langle N, N \rangle = 1$ and the Weingarten equation, one finds that $\widetilde{\nabla}_X N$ is tangent and in fact

$$\widetilde{\nabla}_X N = -sX.$$

Note that $-s(\cdot) = \widetilde{\nabla}_{(\cdot)} N : TM \rightarrow TM$ is called the *Weingarten map*.

The Gauss equation becomes:

Proposition 31.3 (Gauss equation for hypersurfaces). *For tangent vectors X, Y, Z, W ,*

$$\widetilde{R}m(X, Y, Z, W) = Rm(X, Y, Z, W) - h(X, Z)h(Y, W) + h(X, W)h(Y, Z).$$

Also note that the geodesic equation in M becomes

$$0 = D_t \dot{\gamma} = \widetilde{D}_t \dot{\gamma} - \text{II}(\dot{\gamma}, \dot{\gamma})$$

and

$$\widetilde{D}_t \dot{\gamma} = h(\dot{\gamma}, \dot{\gamma})N. \tag{31.1}$$

Definition 31.4. The **mean curvature** of $M^n \subset \widetilde{M}^{n+1}$ is given by

$$H := \frac{1}{n} \text{Tr } s.$$

The **Gauss–Kronecker curvature** is given by

$$K := \det s.$$

31.2. Theorema Egregium. In dimension two, the Gauss equation shows that the Gaussian curvature depends only on the induced metric.

Theorem 31.5 (Theorema Egregium). *For $M^2 \subset \mathbb{R}^3$, then the Gauss-Kronecker curvature (Definition 31.4) is intrinsic, and equal to the curvature of M^2 (Definition 28.4).*

Proof. We work in an orthonormal frame e_1, e_2 of $T_p M$, so that $h_{ij}(p) = h(e_i, e_j)(p) = s(e^i, e_j)(p) = s^i_j(p)$, and $\det s(p) = \det h_{ij}(p)$. Since $\mathbb{R}^3$ is flat, the Gauss equation reads

$$0 = \widetilde{Rm}(e_1, e_2, e_1, e_2) = Rm(e_1, e_2, e_1, e_2) - h_{11}h_{22} + h_{12}^2.$$

Hence

$$K_\Sigma(p) = Rm(e_1, e_2, e_1, e_2) = h_{11}(p)h_{22}(p) - h_{12}(p)^2 = \det(h_{ij}(p)) = K(p).$$

□

31.3. Principal Curvatures. To close the circle of ideas from the first lecture, we should also show that the Gauss curvature just defined agrees with the one defined geometrically in Section 1.6.

Definition 31.6. Let $M \subset \widetilde{M}$ be a hypersurface. The **principal curvatures** of M at p ,

$$\kappa_1 \leq \kappa_2 \leq \dots \leq \kappa_n,$$

are the eigenvalues of the shape operator $s_p : T_p M \rightarrow T_p M$. Since s_p is self-adjoint, these are real, and there is an orthonormal basis $\{e_1, \dots, e_n\}$ at p such that $s(e_i) = \kappa_i e_i$ (no sum on i). For tangent vectors

$$X = \sum_i X^i e_i, \quad Y = \sum_i Y^i e_i$$

at p written in this basis, we get

$$\boxed{h(X, Y) = \kappa_1 X^1 Y^1 + \dots + \kappa_n X^n Y^n.}$$

In particular,

$$\boxed{H = \frac{1}{n} \sum_{i=1}^n \kappa_i,} \quad \boxed{K = \prod_{i=1}^n \kappa_i.}$$

As stated in Lecture 1, the principal curvatures admit a geometric interpretation in terms of 2-plane sections of M .

Proposition 31.7. *Let $M^n \subset \mathbb{R}^{n+1}$ be a hypersurface, let $p \in M$, let $N \in N_p M$ be a unit normal, and let $X \in T_p M$ be a tangent vector with $|X| = 1$. Let Q be the plane in $\mathbb{R}^{n+1}$ through p spanned by X and N , and let γ_X be a unit-speed parametrization of the curve $M \cap Q$ through p with $\gamma_X(0) = p$ and $\dot{\gamma}_X(0) = X$. Then the acceleration of γ_X at p is given by*

$$\ddot{\gamma}_X(0) = h(X, X)N = \left(\sum_{i=1}^n (X^i)^2 \kappa_i \right) N.$$

Proof. Since $|\dot{\gamma}_X| = 1$, we have $\langle \ddot{\gamma}_X, \dot{\gamma}_X \rangle = 0$. Also $\gamma_X \subset Q$, so $\ddot{\gamma}_X \in Q$; hence $\ddot{\gamma}_X$ is proportional to N .

Differentiating $\langle \dot{\gamma}_X, N \rangle = 0$ gives

$$\langle \ddot{\gamma}_X, N \rangle + \langle \dot{\gamma}_X, D_t N \rangle = 0.$$

Using $D_t N = \widetilde{\nabla}_{\dot{\gamma}_X} N = -s(\dot{\gamma}_X)$, we obtain

$$\langle \ddot{\gamma}_X, N \rangle = \langle s(\dot{\gamma}_X), \dot{\gamma}_X \rangle = h(\dot{\gamma}_X, \dot{\gamma}_X).$$

Therefore $\ddot{\gamma}_X(0) = h(X, X)N$. Writing $X = \sum_i X^i e_i$ in a frame as above, we get

$$h(X, X) = \sum_{i=1}^n (X^i)^2 \kappa_i.$$

□

We now recover the definition from Lecture 1.

Corollary 31.8.

$$\kappa_1 = \inf_{|X|=1} \langle \ddot{\gamma}_X(0), N \rangle, \quad \kappa_n = \sup_{|X|=1} \langle \ddot{\gamma}_X(0), N \rangle.$$

32. CLASSICAL NOTATION FOR PARAMETRIZED SURFACES (Tu 4/7)

We now switch to the classical notation used for surfaces in $\mathbb{R}^3$. In my opinion, this is still the best way to do concrete calculations of mean and Gauss curvatures.

Let

$$\Sigma = \text{Im}(\sigma(u, v)) \subset \mathbb{R}^3.$$

Write $\sigma_u = \frac{\partial \sigma}{\partial u}$ and $\sigma_v = \frac{\partial \sigma}{\partial v}$. The first fundamental form is

$$g = |\sigma_u|^2 du^2 + \langle \sigma_u, \sigma_v \rangle du dv + |\sigma_v|^2 dv^2.$$

(Here we use the convention, different from the classical one, that

$$dudv = du \otimes dv + dv \otimes du,$$

i.e.

$$dudv(\partial_u, \partial_v) = 1 = dudv(\partial_v, \partial_u).$$

This is for consistency with our convention for the wedge product.) Set

$$E = |\sigma_u|^2, \quad F = \langle \sigma_u, \sigma_v \rangle, \quad G = |\sigma_v|^2.$$

Then

$$g = E du^2 + F du dv + G dv^2.$$

In matrix notation,

$$g = \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Choose the unit normal

$$\nu = \frac{\sigma_u \times \sigma_v}{|\sigma_u \times \sigma_v|}.$$

The scalar second fundamental form is $h(X, Y) = \langle \widetilde{\nabla}_X Y, \nu \rangle$. Its coefficients are

$$\boxed{L = \langle \sigma_{uu}, \nu \rangle, \quad M = \langle \sigma_{uv}, \nu \rangle, \quad N = \langle \sigma_{vv}, \nu \rangle,}$$

so

$$h = L du^2 + M du dv + N dv^2, \quad h = \begin{pmatrix} L & M \\ M & N \end{pmatrix}.$$

In these coordinates, the Gaussian curvature is

$$\det(s) = \det(g^{-1}h) = \frac{\det(h)}{\det(g)} = \boxed{\frac{LN - M^2}{EG - F^2} = K}.$$

Also,

$$H = \frac{1}{2} \text{Tr } S = \frac{1}{2} \text{Tr}(g^{-1}h), \quad g^{-1} = \frac{1}{EG - F^2} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix},$$

hence

$$\boxed{H = \frac{1}{2} \frac{LG - 2MF + NE}{EG - F^2}}.$$

Definition 32.1. A surface with $H = 0$ is called a **minimal surface**.

Example 32.2 (The Helicoid). Consider the parametrized surface

$$\sigma(r, \theta) = (r \cos \theta, r \sin \theta, \theta).$$

We compute

$$\sigma_r = (\cos \theta, \sin \theta, 0), \quad \sigma_\theta = (-r \sin \theta, r \cos \theta, 1),$$

and

$$\sigma_{rr} = 0, \quad \sigma_{r\theta} = (-\sin \theta, \cos \theta, 0), \quad \sigma_{\theta\theta} = (-r \cos \theta, -r \sin \theta, 0).$$

Therefore

$$E = 1, \quad F = 0, \quad G = 1 + r^2, \quad EG - F^2 = 1 + r^2.$$

Also

$$\sigma_r \times \sigma_\theta = (\sin \theta, -\cos \theta, r), \quad \nu = \frac{1}{\sqrt{1 + r^2}} (\sin \theta, -\cos \theta, r).$$

The second fundamental form coefficients are

$$L = \langle \sigma_{rr}, \nu \rangle = 0, \quad M = \langle \sigma_{r\theta}, \nu \rangle = -\frac{1}{\sqrt{1 + r^2}}, \quad N = \langle \sigma_{\theta\theta}, \nu \rangle = 0.$$

Hence

$$K = \frac{LN - M^2}{EG - F^2} = -\frac{1}{(1 + r^2)^2}, \quad H = \frac{1}{2} \frac{\cancel{L}G - 2M\cancel{F} + \cancel{N}E}{EG - F^2} = 0.$$

Thus **the helicoid is a minimal surface of negative curvature**. You will do another example on homework.

33. DISCUSSION OF EMBEDDED VERSUS ABSTRACT METRICS ON SURFACES (Th 4/9)

We have seen that for a parametrized surface $\Sigma \subset \mathbb{R}^3$, the Gauss and mean curvatures are given by

$$K = \frac{LN - M^2}{EG - F^2} = \kappa_1 \kappa_2, \quad H = \frac{1}{2} \frac{LG - 2MF + NE}{EG - F^2} = \frac{1}{2}(\kappa_1 + \kappa_2),$$

where κ_1, κ_2 are the principal curvatures. The sign of K reflects the local shape of the surface:

$$\begin{aligned} K > 0 &\implies \text{locally elliptic / sphere-like,} \\ K < 0 &\implies \text{locally hyperbolic / saddle-like.} \end{aligned}$$

A surface is called *minimal* if

$$H \equiv 0.$$

In that case $\kappa_1 + \kappa_2 = 0$, so

$$K = \kappa_1 \kappa_2 = -\kappa_1^2 \leq 0.$$

Hence a minimal surface in $\mathbb{R}^3$ cannot have positive Gaussian curvature. We saw an example of such a surface (the helicoid) and you will do another one on homework (the catenoid).

Question. Does there exist a *closed* surface $\Sigma \subset \mathbb{R}^3$ such that $K \leq 0$ everywhere? Can such a surface be minimal?

Answer. No and no. This was a problem on HW1. At the point where the maximum value of the distance to the origin is attained, both principal curvatures must be strictly positive.

So we see that a closed embedded surface in $\mathbb{R}^3$ must have positive Gauss curvature at some points. In fact, we shall see later (and it's not hard to believe) that if $K > 0$ at all points then the surface is homeomorphic to the 2-sphere.

On the other hand, notice that the usual drawings of surfaces of genus $g \geq 1$ in $\mathbb{R}^3$, while of course having positive curvature at the extreme points, always have *some* negative curvature on the “inside parts” of the “donut holes.” We will see next week that there must be a substantial amount of negative curvature to “cancel out” the positive curvature.

Example. The torus $T^2 = S^1 \times S^1$ admits a flat metric, namely the product metric. However, as we have just seen, this metric is not isometrically embeddable in $\mathbb{R}^3$! Hence, the optimal/most obvious metric on the torus is not even attainable by any embedding in $\mathbb{R}^3$. (It can of course be realized in $\mathbb{R}^4$:

$$S^1 \times S^1 \subset \mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4.)$$

Another moral of this example is that both the sphere ($g = 0$) and the torus ($g = 1$) admit metrics of *constant* curvature.

Question. Given a closed surface of genus $g > 1$, does it admit an intrinsic metric of strictly negative curvature? How about constant negative curvature?

Answer. Yes and yes. In class, I drew for you a well-known picture of a regular geodesic $4g$ -gon in hyperbolic space with angle $\frac{2\pi}{4g} = \frac{\pi}{2g}$ at each corner, which (with a little bit more thought) shows that a smooth surface of genus $g \geq 2$ admits at least *one* Riemannian metric of constant curvature -1 (a.k.a. a *hyperbolic metric*).

In fact, every conformal class contains such a metric: this is known as the *Uniformization Theorem* for closed Riemann surfaces. A consequence is that a surface of genus g carries a $6g-6$ -dimensional family of distinct hyperbolic metrics ! We won't discuss this here, but you are certain to learn much more about hyperbolic metrics on closed surfaces if you continue in our department.

34. INVARIANT METRICS ON LIE GROUPS (TH 4/9)

We close this part of the class by touching on a class of examples of a very different flavor from embedded surfaces.

34.1. Left-invariant metrics. Let G be a Lie group, with identity element e . For each $g \in G$, let

$$L_g : G \rightarrow G, \quad h \mapsto gh$$

be left translation.

Start with any inner product $\langle \cdot, \cdot \rangle_e$ on $T_e G$. We extend it to each tangent space $T_g G$ by left translation:

$$\langle v, w \rangle_g := \langle (dL_{g^{-1}})_g v, (dL_{g^{-1}})_g w \rangle_e, \quad v, w \in T_g G.$$

This defines a Riemannian metric on G .

Proposition 34.1. *The metric defined above is left-invariant, i.e.*

$$\langle (dL_h)_g v, (dL_h)_g w \rangle_{hg} = \langle v, w \rangle_g, \quad \forall g, h \in G, \quad \forall v, w \in T_g G.$$

Equivalently, each left-translation

$$L_h : (G, \langle \cdot, \cdot \rangle) \rightarrow (G, \langle \cdot, \cdot \rangle)$$

is an isometry.

Proof. By definition,

$$\langle (dL_h)_g v, (dL_h)_g w \rangle_{hg} = \langle (dL_{(hg)^{-1}})_{hg} (dL_h)_g v, (dL_{(hg)^{-1}})_{hg} (dL_h)_g w \rangle_e.$$

Since

$$L_{(hg)^{-1}} \circ L_h = L_{g^{-1}},$$

we get

$$(dL_{(hg)^{-1}})_{hg} \circ (dL_h)_g = (dL_{g^{-1}})_g.$$

Therefore

$$\langle (dL_h)_g v, (dL_h)_g w \rangle_{hg} = \langle (dL_{g^{-1}})_g v, (dL_{g^{-1}})_g w \rangle_e = \langle v, w \rangle_g.$$

□

34.2. Left-invariant vector fields. Recall that a vector field X on G is called *left-invariant* if

$$(dL_g)_h(X_h) = X_{gh}, \quad \forall g, h \in G.$$

In particular, a left-invariant vector field is completely determined by its value at the identity:

$$X \longleftrightarrow X_e \in T_e G = \mathfrak{g}.$$

Thus the space of left-invariant vector fields, $\mathfrak{g}$, is naturally identified with $T_e G$. Moreover, if X and Y are left-invariant, then so is their Lie bracket $[X, Y]$. This is because pushforward preserves the Lie bracket.

34.3. The Lie-group exponential map. Recall the Lie-group exponential map

$$\text{Exp} : \mathfrak{g} = T_e G \rightarrow G.$$

Given $v \in \mathfrak{g}$, let $X_v \in \mathfrak{g}$ be the left-invariant vector field such that

$$(X_v)_e = v.$$

Let $\gamma_v(t)$ be the integral curve of X_v starting at e . Then

$$\text{Exp}(v) := \gamma_v(1).$$

Equivalently, the curve

$$t \mapsto \text{Exp}(tv)$$

is the integral curve of the left-invariant vector field X_v through e .

For matrix Lie groups, this agrees with the usual matrix exponential:

$$\text{Exp}(tA) = e^{At}.$$

Remark 34.2 (Warning). In general, left-invariant vector fields need not be Killing fields (i.e. ones for which the Lie derivative of the metric is zero) for a left-invariant metric. In fact, the right-invariant ones are. This is because if X_v is left-invariant then one has:

$$\left. \frac{d}{dt} \right|_{t=0} R_{\text{Exp}(tv)} = X_v.$$

See 761 notes, (27.1) on p. 85. However, for a bi-invariant metric, both left- and right-translations are isometries, and this leads to a much stronger structure.

34.4. Bi-invariant metrics.

Definition 34.3. A Riemannian metric on G is called *bi-invariant* if it is both left- and right-invariant.

To characterize bi-invariance, recall that conjugation by $g \in G$ is the map

$$C_g := L_g \circ R_{g^{-1}} : G \rightarrow G, \quad h \mapsto ghg^{-1}.$$

Its differential at the identity defines

$$\text{Ad}_g := (dC_g)_e : \mathfrak{g} \rightarrow \mathfrak{g}.$$

For matrix Lie groups, this is again just the conjugation action on matrices.

Proposition 34.4. *A left-invariant metric on G is bi-invariant if and only if the inner product*

$$\langle \cdot, \cdot \rangle_e$$

on $\mathfrak{g}$ is Ad_g -invariant for every $g \in G$, i.e.

$$\langle \text{Ad}_g v, \text{Ad}_g w \rangle_e = \langle v, w \rangle_e, \quad \forall g \in G, \quad \forall v, w \in \mathfrak{g}.$$

Proof. Right-invariance of the metric is equivalent to saying that each conjugation map $C_g = L_g \circ R_{g^{-1}}$ is an isometry fixing e . Hence its differential at e , namely Ad_g , must preserve the inner product on $\mathfrak{g}$. Conversely, if $\langle \cdot, \cdot \rangle_e$ is Ad_g -invariant for every g , then the corresponding left-invariant metric is also right-invariant. $\square$

Remark 34.5. One can also see the converse directly from the identity

$$R_g = L_g \circ C_{g^{-1}}.$$

Indeed, differentiating at the identity gives

$$(dR_g)_e = (dL_g)_e \circ \text{Ad}_{g^{-1}}.$$

If the metric is left-invariant and the inner product at e is Ad -invariant, then both $(dL_g)_e$ and $\text{Ad}_{g^{-1}}$ preserve the inner product. Hence $(dR_g)_e$ does as well, so R_g is an isometry. Therefore the metric is right-invariant.

34.5. Examples. Consider the following metric on the space $\mathbb{R}^{n \times n}$, of square matrices:

$$\langle X, Y \rangle := \text{Tr}(X^\top Y).$$

Observe that this is nothing but the Euclidean metric!

Given any matrix Lie group $G \subset \text{GL}(n, \mathbb{R}) \subset \mathbb{R}^{n \times n}$, we can endow G with the induced metric.

1. For $G = \text{O}(n)$, this inner product is bi-invariant. Indeed,

$$\text{Tr}((AX)^\top (AY)) = \text{Tr}(X^\top A^\top AY) = \text{Tr}(X^\top Y),$$

so it is left-invariant, and similarly

$$\text{Tr}((XA)^\top (YA)) = \text{Tr}(A^\top X^\top YA) = \text{Tr}(X^\top YAA^\top) = \text{Tr}(X^\top Y),$$

so it is also right-invariant.

2. Likewise, on $\text{U}(n)$ one may use the Hermitian metric

$$\langle X, Y \rangle = \text{Tr}(X^\dagger Y),$$

which is again bi-invariant. In fact this is just the metric as above, where one considers $\text{U}(n) \subset \text{SO}(2n)$.

3. In the case

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} \mid |z|^2 + |w|^2 = 1 \right\} \subset \mathbb{C}^{2 \times 2} \cong \mathbb{R}^{4 \times 4},$$

we obtain a bi-invariant metric on $\mathrm{SU}(2)$. But considering the columns shows that this is simply the round metric on $\mathrm{SU}(2) \cong \mathbb{S}^3$. So with the structure of $\mathbb{S}^3$ as a Lie group, the round metric is bi-invariant.

34.6. Non-examples.

1. Fix a positive-definite symmetric matrix $Z > 0$ that is not a multiple of the identity. Define

$$\langle X, Y \rangle_Z := \mathrm{Tr}(ZX^\top Y).$$

This gives a Riemannian metric on $\mathrm{GL}(n)$ that is left-invariant but not right-invariant.

2. The Killing form on $\mathfrak{g}$ is always Ad-invariant, so gives a bi-invariant bilinear form on the tangent spaces of any Lie group. However, it need not be positive- or negative-definite, so can only be used to define a bi-invariant Riemannian metric in specific cases. (In fact, only for compact semisimple groups; we will see below why this is the case.)

34.7. Existence on compact Lie groups.

Proposition 34.6. *Every compact Lie group admits a bi-invariant Riemannian metric.*

Proof. Start with any inner product $\langle \cdot, \cdot \rangle_0$ on $\mathfrak{g}$.

Since G is compact, it admits a bi-invariant Haar measure (equivalently, a bi-invariant volume form), which we denote by ω —see 761 notes, Example 45.5. We average over the group and define

$$\langle v, w \rangle_e := \int_G \langle \mathrm{Ad}_g v, \mathrm{Ad}_g w \rangle_0 \omega.$$

This gives a new inner product on $\mathfrak{g}$.

It is Ad-invariant. Indeed, for any $h \in G$,

$$\langle \mathrm{Ad}_h v, \mathrm{Ad}_h w \rangle_e = \int_G \langle \mathrm{Ad}_g \mathrm{Ad}_h v, \mathrm{Ad}_g \mathrm{Ad}_h w \rangle_0 \omega.$$

Using

$$\mathrm{Ad}_g \mathrm{Ad}_h = \mathrm{Ad}_{gh}$$

and the right-invariance of Haar measure, this becomes

$$\int_G \langle \mathrm{Ad}_{gh} v, \mathrm{Ad}_{gh} w \rangle_0 \omega = \int_G \langle \mathrm{Ad}_g v, \mathrm{Ad}_g w \rangle_0 \omega = \langle v, w \rangle_e.$$

Therefore the inner product $\langle \cdot, \cdot \rangle_e$ is Ad-invariant, and hence the associated left-invariant metric is bi-invariant. $\square$

34.8. Consequences of bi-invariance. Now assume G carries a bi-invariant metric $\langle \cdot, \cdot \rangle$, and let X, Y, Z be left-invariant vector fields.

Proposition 34.7. *For left-invariant vector fields X, Y, Z on a Lie group with a bi-invariant metric, we have:*

1. $\langle [X, Y], Z \rangle = \langle X, [Y, Z] \rangle$.
2. $\nabla_X Y = \frac{1}{2}[X, Y]$.
3. $\text{Exp}_e(tv) = \exp(tv), \forall v \in \mathfrak{g} = T_e G$, i.e., the Lie-group exponential map agrees with the Riemannian exponential map.

Proof. Since the metric is bi-invariant and Y, Z are left-invariant, the function

$$g \mapsto \langle Y_g, Z_g \rangle_g = \langle dL_g Y_e, dL_g Z_e \rangle_e$$

is constant on G . Consider this constant function along the one-parameter subgroup

$$t \mapsto g(t) = \text{Exp}(tX).$$

Translating back by right-invariance, we get

$$\langle Y_{g(t)}, Z_{g(t)} \rangle_{g(t)} = \langle dR_{g^{-1}(t)} Y_{g(t)}, dR_{g^{-1}(t)} Z_{g(t)} \rangle_e = \langle dC_{g(t)} Y_e, dC_{g(t)} Z_e \rangle_e = \langle \text{Ad}_{g(t)} Y_e, \text{Ad}_{g(t)} Z_e \rangle_e.$$

Differentiating at $t = 0$ and using the Leibniz rule, we obtain

$$\begin{aligned} 0 &= \left\langle \frac{d}{dt} \Big|_{t=0} \text{Ad}_{\text{Exp}(tX)} Y_e, \text{Ad}_{\text{Exp}(tX)} Z_e \right\rangle_e + \left\langle \text{Ad}_{\text{Exp}(tX)} Y_e, \frac{d}{dt} \Big|_{t=0} \text{Ad}_{\text{Exp}(tX)} Z_e \right\rangle_e \\ &= \langle \text{ad}_X(Y), Z \rangle + \langle Y, \text{ad}_X(Z) \rangle. \end{aligned}$$

But recall from 761 notes, Definition/Lemma 27.2, that $\text{ad}_X(Y) = [X, Y]$ is merely the Lie bracket. Thus

$$\langle [X, Y], Z \rangle = -\langle Y, [X, Z] \rangle = \langle X, [Y, Z] \rangle,$$

which proves (1).

For (2), apply the Koszul formula (HW6 #1):

$$2\langle \nabla_X Y, Z \rangle = X\langle Y, Z \rangle + Y\langle Z, X \rangle - Z\langle X, Y \rangle + \langle [X, Y], Z \rangle - \langle [Y, Z], X \rangle - \langle [X, Z], Y \rangle.$$

Since $\langle X, Y \rangle$, $\langle Y, Z \rangle$, and $\langle Z, X \rangle$ are constant, the three derivative terms vanish. Hence

$$2\langle \nabla_X Y, Z \rangle = \langle [X, Y], Z \rangle - \langle [Y, Z], X \rangle - \langle [X, Z], Y \rangle.$$

Using part (1), this simplifies to

$$2\langle \nabla_X Y, Z \rangle = \langle [X, Y], Z \rangle.$$

Since this holds for every left-invariant Z , we conclude that

$$\nabla_X Y = \frac{1}{2}[X, Y].$$

For (3), let $v \in \mathfrak{g}$, and let X be the corresponding left-invariant vector field, so that $X_e = v$. The integral curve of X through e is

$$\gamma(t) = \text{Exp}(tv).$$

We have

$$\nabla_{\dot{\gamma}} \dot{\gamma} = \nabla_X X \stackrel{(2)}{=} \frac{1}{2}[X, X] = 0.$$

So γ is a geodesic with initial velocity v . Therefore

$$\text{Exp}(tv) = \gamma(t) = \exp(tv).$$

□

Remark 34.8. Thus, for a bi-invariant metric, the Levi-Civita connection is especially simple on left-invariant vector fields, and geodesics through the identity are exactly one-parameter subgroups.

34.9. Curvature formula.

Theorem 34.9. *For a Lie group with a bi-invariant metric and left-invariant vector fields X, Y, Z , one has*

$$\mathcal{R}(X, Y)Z = -\frac{1}{4}[[X, Y], Z]$$

and

$$Rm(X, Y, Z, W) = \frac{1}{4} \langle [X, Y], [Z, W] \rangle.$$

Proof. Using

$$\nabla_X Y = \frac{1}{2}[X, Y],$$

we compute

$$\mathcal{R}(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z.$$

Thus

$$\mathcal{R}(X, Y)Z = \frac{1}{4}[X, [Y, Z]] - \frac{1}{4}[Y, [X, Z]] - \frac{1}{2}[[X, Y], Z].$$

By the Jacobi identity,

$$[X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z].$$

Therefore

$$\mathcal{R}(X, Y)Z = \frac{1}{4}[[X, Y], Z] - \frac{1}{2}[[X, Y], Z] = -\frac{1}{4}[[X, Y], Z].$$

The formula for the Riemann tensor follows by pairing with W and using the identity of Proposition 34.7.1. □

Corollary 34.10. *The sectional curvature of a bi-invariant Riemannian metric on a Lie group is nonnegative!*

Example 34.11. On HW3 #7, I tricked you into using a basis of left-invariant vector fields to compute the curvature of S^3 .

Part 8. Curvature and topology

In the previous Part we saw the following facts:

- Minimal surfaces in $\mathbb{R}^3$ have nonpositive curvature everywhere.
- A *closed* (i.e. compact without boundary) embedded surface in $\mathbb{R}^3$ must have positive curvature somewhere. In particular, it can never be minimal.
- There do exist (abstract) metrics of constant negative curvature on closed surfaces of genus two or more.
- A bi-invariant metric on a Lie group has nonnegative curvature.
- Every compact Lie group has a bi-invariant metric, hence a metric with nonnegative sectional curvature.
- (HW) Quotienting *increases* sectional curvature. We discussed the example of the Hopf fibration

$$S^1 = U(1) \rightarrow S^3 = SU(2) \rightarrow S^2_{1/2} = \mathbb{CP}^1.$$

The (constant) sectional curvature of S^3 is 1. Meanwhile, as I explained, the quotient (Fubini-Study) metric on S^2 actually corresponds to a sphere of radius $1/2$, which has constant curvature equal to 4.

Rather than collecting more isolated facts, we will now try to prove some general theorems about the interaction between curvature and topology. It will be interesting to see how each of these examples fits into the general picture.

35. THE GAUSS-BONNET THEOREM (TU 4/14)

Today we will give a highly satisfactory answer to the question of how topology and curvature interact in the case of closed surfaces. We begin with the following Lemma, which as we shall see, is just Theorem 10.7 in different language.

Lemma 35.1. *Let $\gamma = \partial D$ be a contractible closed piecewise smooth loop in an orientable surface Σ , oriented counterclockwise. Then $P^{\nabla^g, \gamma}$ is counterclockwise rotation by angle*

$$\theta = \int_D K_g dV_g.$$

Proof. Since Σ is oriented, we can view $T\Sigma = L$ as a complex line bundle, where multiplication by i is rotation by $\pi/2$ counterclockwise. The (real) metric $\langle \cdot, \cdot \rangle$ on $T\Sigma$ gives a Hermitian metric $(\cdot, \cdot)$ on L , by the formula

$$(v, w) = \langle v, w \rangle + i\langle v, iw \rangle.$$

Moreover $\langle \cdot, \cdot \rangle = \text{Re}(\cdot, \cdot)$. The Levi-Civita connection ∇^g on $T\Sigma$ gives a (Hermitian-)metric-compatible connection A on L via the identification $T\Sigma = L$.

Claim. $\boxed{F_A = -iK_g dV_g.}$

Proof of claim. Let $\{e_1, e_2\}$ be an oriented orthonormal basis at a point. Then

$$K_g = \langle \mathcal{R}(e_1, e_2)e_2, e_1 \rangle.$$

Since this is an oriented basis, we in fact have $e_2 = ie_1$, so

$$\begin{aligned} K_g &= \text{Re}(F_A(e_1, e_2)ie_1, e_1) \\ &= \text{Re}(iF_A(e_1, e_2)(e_1, e_1)) \\ &= iF_A(e_1, e_2). \end{aligned}$$

Here we have used $\mathbb{C}$ -linearity in the first factor, $(e_1, e_1) = 1$, and the fact that $iF_A(e_1, e_2)$ is real since A is compatible with a Hermitian metric. Since $dV_g(e_1, e_2) = 1$, this is equivalent to the claim. $\diamond$

Recall from Theorem 10.7 that for any connection A on a complex line bundle L , we have

$$P^{A, \gamma} = e^{-\int_D F_A}.$$

We therefore have

$$P^{\nabla^g, \gamma} = P^{A, \gamma} = e^{-\int_D F_A} = e^{i \int_D K_g dV_g} = e^{i\theta}.$$

□

Theorem 35.2 (Local Gauss-Bonnet Formula). *Let Δ be a geodesic triangle in a uniformly normal neighborhood on a surface Σ . Let $\theta_1, \theta_2, \theta_3$ be the angles of Δ . Then*

$$\boxed{\theta_1 + \theta_2 + \theta_3 = \pi + \int_{\Delta} K_g dV_g.}$$

We call $\int_{\Delta} K_g dV_g$ the **angle defect**.

Proof. Assume without loss of generality that one of the vertices is the origin in normal coordinates. Let v, w be unit vectors along the sides beginning at 0. Let γ be its third side. Then γ stays in the (coordinate) sector between v and w . Take $\theta_1 = \angle(v, w)$, θ_2 the angle between the side corresponding to v and γ , and θ_3 the angle between the side corresponding to w and γ . Moreover, let $u = \frac{\gamma'(0)}{|\gamma'(0)|}$ be the initial unit vector of γ . Applying Gauss's lemma and the fact that parallel transport preserves the metric (with reference to the gorgeous picture drawn in class), we obtain the following facts.

- $\angle(v, u) = \angle(P^\gamma v, P^\gamma u)$.
- $\angle(v, P^\gamma v) = \int_{\Delta} K_g dV_g = \theta$ (by the previous Lemma).
- $\angle(P^\gamma v, w) = \angle(P^\gamma v, v) + \angle(v, w) = -\theta + \theta_1 = \theta_1 - \theta$.
- $\angle(v, u) = \pi - \theta_2$.
- $\angle(w, P^\gamma u) = \theta_3$.

(Note that we always take angles using the metric at the point in question, which only necessarily agrees at the origin with angles in the coordinate chart.)

We can then calculate

$$\pi - \theta_2 = \angle(v, u) = \angle(P^\gamma v, P^\gamma u) = \angle(P^\gamma v, w) + \angle(w, P^\gamma u) = \theta_1 - \theta + \theta_3.$$

Rearranging gives the desired formula. $\square$

Theorem 35.3 (Global Gauss-Bonnet Theorem). *The Euler characteristic of a closed orientable surface Σ is given by*

$$2 - 2g(\Sigma) = \chi(\Sigma) = \frac{1}{2\pi} \int_{\Sigma} K_g dV_g.$$

First proof. Triangulate Σ by geodesic triangles each of which is contained in a uniformly normal neighborhood (HW exercise). Since this is a triangulation, each face has three edges but shares each edge with another face, so $2E = 3F$. Meanwhile, the angles at each vertex all sum to 2π . Therefore

$$2\pi V = \sum_{\Delta \in F} (\pi + \theta_{\Delta}) = \pi F + \sum_{\Delta \in F} \int_{\Delta} K_g dV_g = \pi F + \int_{\Sigma} K_g dV_g.$$

Adding $\pi(3F - 2E) = 0$, we get

$$2\pi(V - E + F) = 2\pi\chi(\Sigma) = \int_{\Sigma} K_g dV_g.$$

$\square$

Second proof. This proof gives the global formula directly. Recall from Definition/Lemma 10.5 that the first Chern class $c_1(L) = \left[\frac{i}{2\pi} F_A\right]$ of a complex line bundle is independent of the connection. By the previous discussion, given any metric g on Σ , we obtain a connection A on $L \cong T\Sigma$ for which, by the Claim on p. 108, we have

$$iF_A = K_g dV_g.$$

We conclude that the number

$$\xi(\Sigma) := \frac{1}{2\pi} \int_{\Sigma} K_g dV_g$$

is independent of the metric g on Σ !

It remains to show that $\xi(\Sigma) = \chi(\Sigma)$. We proceed by induction on the genus. For $g(\Sigma) = 0$, plugging in the round metric, we get

$$\xi(S^2) = \frac{4\pi}{2\pi} = 2 = \chi(S^2).$$

For $g(\Sigma) = 1$, T^2 admits a flat metric, so $\xi(T^2) = 0 = \chi(T^2)$. We have proven the Gauss-Bonnet theorem for genus zero and one.

Assume that the theorem has been proven for genus g ; we prove it for genus $g+1$. We are free to choose the metrics exactly as we please. Take a surface Σ' of genus g , and choose a metric with two perfectly round hemispherical caps.

Chop off the caps, and glue on half of a $\mathbb{Z}_2$ -symmetric torus instead. This produces from Σ' a surface Σ of genus $g + 1$, which we write schematically as

$$\Sigma = \Sigma' - S^2 + \frac{1}{2}T^2.$$

Another gorgeous picture was drawn in class. Integrating the curvature of this metric, we get

$$\xi(\Sigma) = \xi(\Sigma') - 2 + 0 = 2 - 2g - 2 = 2 - 2(g + 1) = \chi(\Sigma).$$

Here we have used the fact that two round hemispheres together form a round sphere and that half of a $\mathbb{Z}_2$ -symmetric torus has zero total curvature. This completes the induction step. $\square$

We state some immediate corollaries of Gauss-Bonnet. Here g is a metric on a closed, orientable surface Σ .

Corollary 35.4. *If $K_g \geq 0$ everywhere and $K_g > 0$ somewhere then $\Sigma \simeq \mathbb{S}^2$.*

Corollary 35.5. *If $K_g \leq 0$ everywhere and $K_g < 0$ somewhere then $g(\Sigma) \geq 2$.*

Corollary 35.6. *Suppose $\Sigma = T^2$. If $K_g \geq 0$ or $K_g \leq 0$ everywhere then $K_g \equiv 0$.*

36. JACOBI FIELDS (TH 4/16)

36.1. Overview. Last time we were able to give a remarkably complete description of the relationship between curvature and topology in dimension two, based on the interpretation of curvature as a differential form. This approach related the *integral* of curvature to the topology of the space. Gauss-Bonnet can be generalized to higher dimensions, but neither the general formula (due to Chern) nor its consequences are quite as clear and dramatic as in the original version.

In the remaining lectures we will study the effect of a stronger assumption, namely *point-wise* positivity or negativity of the curvature, on the shape of the manifold. Our approach will be to study the effect of curvature on geodesics. *Spoiler alert:*

- Positive curvature makes geodesics converge, whereas
- Negative curvature makes geodesics diverge.

Notice that we have already seen something along these lines with the Local Gauss-Bonnet Formula (Theorem 35.2): positive curvature makes for fat triangles (as in the sphere), while negative curvature makes for skinny triangles (as in hyperbolic space).

36.2. The Jacobi equation. Let $\Gamma(s, t)$ be a smooth map from a product of two intervals $J \times I$ into M , and let $\gamma(t) = \Gamma(0, t)$. We say that Γ is a **variation of γ through geodesics** if each of the curves $\Gamma_s(t) = \Gamma(s, t)$ is a geodesic.

Recall that given a variation $\Gamma(s, t)$, we use the notation

$$S = \partial_s \Gamma, \quad T = \partial_t \Gamma.$$

These are vector fields along Γ . The condition to be a variation through geodesics is equivalent to

$$D_t T \equiv 0. \quad (36.1)$$

Also recall that the *variation field* of Γ is the vector field

$$V(t) = S(0, t) = \partial_s \Gamma|_{s=0}$$

along γ .

We wish to derive a necessary equation on V to be the variation field of a variation through geodesics. (The equation will also turn out to be sufficient, in our situation.) The following lemma was used without justification in the second proof of Theorem 25.1 above; it is more-or-less obvious from the definition of curvature and the fact that S and T are pushforwards of the commuting vector field ∂_s and ∂_t on $J \times I$. Just as a sanity check, let's give a full John-Lee-style proof.

Lemma 36.1. *For any vector field W along Γ , we have*

$$D_s D_t W - D_t D_s W = \mathcal{R}(S, T)W.$$

Proof. We work in local coordinates $\{x^i\}$, in which $W = W^i \partial_i$ etc. We have

$$D_t W = \partial_t W^i \partial_i + W^i D_t \partial_i,$$

$$D_s D_t W = \partial_s \partial_t W^i \partial_i + \partial_t W^i D_s \partial_i + \partial_s W^i D_t \partial_i + W^i D_s D_t \partial_i$$

and

$$\begin{aligned} D_s D_t W - D_t D_s W &= (\partial_s \partial_t W^i - \partial_t \partial_s W^i) \partial_i \\ &\quad + \partial_t W^i D_s \partial_i - \partial_s W^i D_t \partial_i \\ &\quad + \partial_s W^i D_t \partial_i - \partial_t W^i D_s \partial_i \\ &\quad + W^i (D_s D_t \partial_i - D_t D_s \partial_i) \\ &= W^i (D_s D_t \partial_i - D_t D_s \partial_i). \end{aligned} \quad (36.2)$$

It remains to expand out the commutator term:

$$\begin{aligned} D_s D_t \partial_i - D_t D_s \partial_i &= D_s (T^j \nabla_{\partial_j} \partial_i) - D_t (S^j \nabla_{\partial_j} \partial_i) \\ &= S^k T^j (\nabla_{\partial_k} \nabla_{\partial_j} - \nabla_{\partial_j} \nabla_{\partial_k}) \partial_i + (\partial_s T^j \nabla_{\partial_j} \partial_i - \partial_t S^j \nabla_{\partial_j} \partial_i). \end{aligned} \quad (36.3)$$

But, writing $\Gamma^j(s, t)$ for the j -coordinate of $\Gamma(s, t)$, we have

$$\partial_s T^j = \partial_s \partial_t \Gamma^j$$

and

$$\partial_t S^j = \partial_s \partial_t \Gamma^j = \partial_t \partial_s \Gamma^j = \partial_s T^j.$$

So we get

$$(\partial_s T^j \nabla_{\partial_j} \partial_i - \partial_t S^j \nabla_{\partial_i} \partial_j) = \partial_t \partial_s \Gamma^j (\nabla_{\partial_j} \partial_i - \nabla_{\partial_i} \partial_j) = 0.$$

Hence (36.3) simplifies to

$$D_s D_t \partial_i - D_t D_s \partial_i = S^k T^j (\nabla_{\partial_k} \nabla_{\partial_j} - \nabla_{\partial_j} \nabla_{\partial_k}) \partial_i = \mathcal{R}(S, T) \partial_i.$$

Returning to (36.2), we get

$$D_s D_t W - D_t D_s W = W^i \mathcal{R}(S, T) \partial_i = \mathcal{R}(S, T) W,$$

by linearity of the curvature operator. $\square$

Theorem 36.2 (Jacobi equation). *Let γ be a geodesic in M and Γ a variation of γ through geodesics. Then the variation field $V = \partial_s \Gamma|_{s=0}$ satisfies the **Jacobi equation***

$$D_t^2 V + \mathcal{R}(V, \dot{\gamma}) \dot{\gamma} = 0.$$

Such a V is called a **Jacobi field** along γ .

Proof. We take the s -derivative of (36.1) and use the above Lemma:

$$\begin{aligned} 0 &= D_s D_t T \\ &= D_t D_s T + \mathcal{R}(S, T) T. \end{aligned}$$

Next, we use the Symmetry Lemma 21.3, to obtain

$$0 = D_t D_t S + \mathcal{R}(S, T) T.$$

Specializing to $s = 0$, we obtain the Jacobi equation. $\square$

Remark 36.3. The Jacobi equation is the “linearization” of the geodesic equation, $D_t \frac{d\gamma}{dt} = 0$. While the geodesic equation does not explicitly involve the curvature, it is interesting that the Jacobi equation does. This will allow us to make the connection between geodesics and curvature.

37. FIRST PROPERTIES OF JACOBI FIELDS (Th 4/16-Tu 4/21)

37.1. Existence and uniqueness.

Proposition 37.1. *Let $\gamma : I \rightarrow M$ be a geodesic. Let $c \in I$ and take $v, w \in T_{\gamma(c)}M$. There exists a unique Jacobi field $J : I \rightarrow M$ with $J(c) = v$ and $D_t J(c) = w$.*

Proof. Via parallel transport, we can choose a parallel orthonormal frame $E_\alpha(t)$ over γ . Writing $J = J^\eta E_\eta$, $\dot{\gamma} = \dot{\gamma}^\eta E_\eta$, and $\mathcal{R}(E_\alpha, E_\beta) E_\delta = R_{\alpha\beta}{}^\eta{}_\delta E_\eta$, the Jacobi equation becomes

$$\frac{d^2 J^\eta}{dt^2} + R_{\alpha\beta}{}^\eta{}_\delta (\gamma(t)) J^\alpha \dot{\gamma}^\beta \dot{\gamma}^\delta = 0, \quad \eta = 1, \dots, n.$$

We get local existence near c as a consequence of the theory of second-order ordinary differential equations. Notice that this is a linear equation in J (everything else is fixed). Therefore the local solution extends as long as γ is defined. $\square$

Proposition 37.2. *The space of Jacobi fields is a $2n$ -dimensional subspace of the space of vector fields along γ .*

Remark 37.3. The last two propositions should be expected: v corresponds to varying the initial position of γ and w corresponds to varying the initial velocity of γ .

Proposition 37.4. *Suppose either that I is compact or M is complete. Then every Jacobi field is the variation field of a variation through geodesics.*

Proof. Assume wlog that $0 \in I$. Let $p = \gamma(0) \in M$. Choose a curve $\sigma(s) \in M$ with $\sigma(0) = p$ and $\dot{\sigma}(0) = J(0)$. Next, choose a vector field W along σ with $W(0) = \dot{\gamma}(0)$ and $D_s W(0) = D_t J(0)$. Then define

$$\Gamma(s, t) = \exp_{\sigma(s)}(tW(s)).$$

If I is compact or M is complete then the exponential map is defined on all of I for $\sigma(s)$ near $\sigma(0) = p$ and $W(s)$ near $W(0) = \dot{\gamma}(0)$, so Γ is defined on $J \times I$ for some interval J . Hence Γ is a variation through geodesics.

Let $V = \partial_s \Gamma|_{s=0}$ be the variation field of Γ . By Theorem 36.2, V is a Jacobi field. It is an exercise to check that with the above choices, we have

$$V(0) = J(0), \quad D_t V(0) = D_s W(0) = D_t J(0).$$

By uniqueness of Jacobi fields, we conclude that $V(t) = J(t)$ for all $t \in I$. □

Remark 37.5. A geodesic $\gamma : \mathbb{R} \rightarrow M$ is called **closed** if it is periodic. Equivalently, it is given by a map $S^1 \rightarrow M$ solving the geodesic equation. The study of closed geodesics is a classical topic in differential geometry and was the subject of problem 7 on HW7. It is interesting to note that the last Proposition *fails* for closed geodesics: a closed geodesic can have “non-integrable Jacobi fields,” i.e. periodic solutions of the Jacobi equation that do not come from a bona-fide variation of *closed* geodesics. This is a daily headache for anyone who studies second-order equations coming from geometry.

37.2. Jacobi fields vanishing at a point. Having defined Jacobi fields and proved their existence, I should point out that in the case of greatest interest, they are “trivial.”

Proposition 37.6. *Suppose that a Jacobi field vanishes at $\gamma(0) = p$, i.e. $J(0) = 0$. Suppose that γ has unit speed, and let $w = D_t J(0)$. Then, in exponential coordinates centered at p , we have*

$$J(t)^i = tw^i \partial_i|_{\gamma(t)}.$$

Proof. This is clear because $J(t)$ is the variation field of the variation

$$\Gamma(s, t) = \exp_p(t(v + sw)),$$

where $v = \gamma'(0)$, by the proof of Proposition 37.4. □

The last proposition can be considered to furnish lots of examples of Jacobi fields. On the other hand, these are not really explicit examples, because *exponential coordinates are not explicit!* In fact, the whole point of studying Jacobi fields is to get a handle on the

exponential map. We will have occasion to solve the Jacobi equation explicitly in the model spaces next class.

37.3. Tangential and normal Jacobi fields (Tu 4/21). There are two trivial Jacobi fields along a geodesic, coming from reparametrizing the geodesic. Indeed, if γ is a geodesic, then for fixed s the curves

$$t \mapsto \gamma(t + s), \quad t \mapsto \gamma((1 + s)t)$$

are also geodesics. Therefore these define variations through geodesics. By chain rule, their variation fields are

$$J_0(t) = \dot{\gamma}(t), \quad J_1(t) = t\dot{\gamma}(t).$$

In particular, by linearity of the Jacobi equation, every vector field of the form

$$J(t) = (a + bt)\dot{\gamma}(t)$$

is a Jacobi field. Such Jacobi fields are called **tangential Jacobi fields**.

Definition 37.7. A Jacobi field J along γ is called **normal**, written as $J \in \mathcal{J}^\perp(\gamma)$, if

$$\langle J(t), \dot{\gamma}(t) \rangle = 0$$

for all t .

Proposition 37.8. *Let J be a Jacobi field along a geodesic γ . The following are equivalent:*

1. J is normal.
2. J is orthogonal to $\dot{\gamma}$ at two distinct points.
3. At one point t_0 , both $J(t_0)$ and $D_t J(t_0)$ are orthogonal to $\dot{\gamma}(t_0)$.
4. For every t , both $J(t)$ and $D_t J(t)$ are orthogonal to $\dot{\gamma}(t)$.

Proof. Set

$$f(t) = \langle J(t), \dot{\gamma}(t) \rangle.$$

Since γ is a geodesic, $D_t \dot{\gamma} = 0$, so

$$f'(t) = \langle D_t J, \dot{\gamma} \rangle.$$

Differentiating once more and using the Jacobi equation, we get

$$\begin{aligned} f''(t) &= \langle D_t^2 J, \dot{\gamma} \rangle \\ &= -\langle \mathcal{R}(J, \dot{\gamma})\dot{\gamma}, \dot{\gamma} \rangle \\ &= -\text{Rm}(J, \dot{\gamma}, \dot{\gamma}, \dot{\gamma}) = 0. \end{aligned}$$

Thus f is an affine linear function of t .

Now (1) says exactly that $f \equiv 0$. Since an affine linear function which vanishes at two distinct points is identically zero, (2) implies (1). The implication (1) $\Rightarrow$ (2) is obvious.

Next, (3) says precisely that $f(t_0) = f'(t_0) = 0$. But an affine linear function with vanishing value and derivative at one point must be identically zero, so (3) implies (1). Conversely,

if $f \equiv 0$, then certainly f' vanishes everywhere, so (1) implies (4), and (4) trivially implies (3). $\square$

Corollary 37.9. *The space of normal Jacobi fields along γ is a $(2n-2)$ -dimensional subspace of the space of all Jacobi fields along γ .*

Proof. Fix t_0 . By Proposition 37.8, a Jacobi field is normal if and only if both $J(t_0)$ and $D_t J(t_0)$ lie in $\dot{\gamma}(t_0)^\perp$, which is an $(n-1)$ -dimensional subspace of $T_{\gamma(t_0)}M$. By existence and uniqueness of Jacobi fields, such a field is uniquely determined by these two vectors. Hence the space of normal Jacobi fields has dimension $2n-2$. $\square$

38. LOCAL UNIQUENESS OF CONSTANT SECTIONAL CURVATURE (Tu 4/21)

We now explain why the three model spaces are locally the only examples of metrics of constant sectional curvature.

Proposition 38.1. *If (M, g) has constant sectional curvature equal to c , then*

$$Rm(v, w, x, y) = c(\langle v, x \rangle \langle w, y \rangle - \langle v, y \rangle \langle w, x \rangle).$$

Equivalently,

$$\mathcal{R}(v, w)x = c(\langle w, x \rangle v - \langle v, x \rangle w).$$

Proof. Fix a point $p \in M$. We must show that the curvature tensor at p is given by the claimed formula. Define a $(0, 4)$ -tensor on $T_p M$ by

$$Rm_c(x, y, z, w) := c(\langle y, w \rangle \langle x, z \rangle - \langle x, w \rangle \langle y, z \rangle).$$

It is easy to check that Rm_c has all of the symmetries of a Riemann tensor, i.e. it belongs to the space of algebraic curvature tensors $\mathcal{R}$.

We claim that Rm_c has sectional curvature equal to c . Indeed, if x, y are orthonormal vectors in $T_p M$, then

$$\begin{aligned} Rm_c(x, y, x, y) &= c(\langle y, y \rangle \langle x, x \rangle - \langle x, y \rangle \langle y, x \rangle) \\ &= c. \end{aligned}$$

On the other hand, since (M, g) has constant sectional curvature equal to c , we also have

$$Rm(x, y, x, y) = c$$

for every orthonormal pair $x, y \in T_p M$.

Thus Rm and Rm_c have the same sectional curvatures at the point p . By Proposition 27.4, it follows that

$$Rm = Rm_c$$

at p . Since p was arbitrary, the formula holds at every point of M .

The formula for the curvature operator is just the raised-index version of this identity. $\square$

For each $c \in \mathbb{R}$, define the function

$$s_c(t) = \begin{cases} t & c = 0 \\ R \sin \frac{t}{R} & c = \frac{1}{R^2} > 0 \\ R \sinh \frac{t}{R} & c = -\frac{1}{R^2} < 0. \end{cases}$$

Note that $s_c(t)$ is just the solution of the ODE

$$u'' + cu = 0, \quad u(0) = 0, \quad u'(0) = 1.$$

Proposition 38.2. *Suppose that (M, g) has constant sectional curvature equal to c , and let γ be a unit-speed geodesic. Then every normal Jacobi field J along γ with $J(0) = 0$ is of the form*

$$J(t) = k s_c(t) E(t),$$

where $E(t)$ is a parallel unit normal vector field along γ and

$$D_t J(0) = k E(0).$$

In particular,

$$|J(t)| = |s_c(t)| |D_t J(0)|.$$

Proof. By Proposition 38.1,

$$\mathcal{R}(v, w)x = c(\langle w, x \rangle v - \langle v, x \rangle w).$$

Hence, if J is normal and γ has unit speed, then

$$\begin{aligned} \mathcal{R}(J, \dot{\gamma})\dot{\gamma} &= c(\langle \dot{\gamma}, \dot{\gamma} \rangle J - \langle J, \dot{\gamma} \rangle \dot{\gamma}) \\ &= cJ. \end{aligned}$$

Therefore the Jacobi equation becomes

$$D_t^2 J + cJ = 0.$$

If $D_t J(0) = 0$, then uniqueness for the Jacobi equation gives $J \equiv 0$, so the conclusion is clear. Otherwise, let

$$k = |D_t J(0)|$$

and let $E(t)$ be the parallel unit vector field along γ with

$$E(0) = \frac{D_t J(0)}{|D_t J(0)|}.$$

Since $D_t E = 0$, writing ansatz

$$J(t) = u(t)E(t)$$

gives

$$D_t^2 J = u''(t)E(t).$$

Thus u satisfies

$$u'' + cu = 0, \quad u(0) = 0, \quad u'(0) = k.$$

By the definition of s_c , we conclude that

$$u(t) = k s_c(t),$$

so

$$J(t) = k s_c(t) E(t).$$

□

Example 38.3 ($c = 0$). Let

$$\pi : \mathbb{R}^n \setminus \{0\} \rightarrow S^{n-1}, \quad x \mapsto \frac{x}{|x|},$$

and let

$$\hat{g} := \pi^* g_{S^{n-1}},$$

where $g_{S^{n-1}}$ is the round metric on the unit sphere. Then the Euclidean metric can be written in polar coordinates as

$$g_{\text{Euc}} = dr^2 + r^2 \hat{g}.$$

Theorem 38.4. *Suppose that (M, g) has constant sectional curvature equal to c , and let $p \in M$. Then, in normal coordinates at p , the metric is given by*

$$\boxed{g = dr^2 + s_c(r)^2 \hat{g}.}$$

where $\hat{g}$ is the standard round metric on S^{n-1} pulled back to $\mathbb{R}^n \setminus \{0\}$.

Proof. Work in normal coordinates centered at p . By the Gauss Lemma, we already know that

$$g = dr^2 + \tilde{g},$$

where $\tilde{g}$ is the restriction of g to the tangent spaces of the geodesic spheres centered at p . Thus it remains to check that

$$\tilde{g} = s_c(r)^2 \hat{g}.$$

Fix a point q in the normal coordinate neighborhood, and let

$$b = d_g(p, q).$$

Let $w \in T_q M$ satisfy

$$\langle w, \partial_r \rangle = 0,$$

so that w is tangent to the geodesic sphere of radius b through q .

Write the radial geodesic from p to q as

$$\gamma(t) = \left(\frac{t}{b} q^1, \dots, \frac{t}{b} q^n \right), \quad 0 \leq t \leq b.$$

Then

$$\gamma(0) = p, \quad \gamma(b) = q.$$

Now let J be the normal Jacobi field along γ such that

$$J(b) = w.$$

Since the variation is through radial geodesics, we have

$$D_t J(0) = \frac{1}{b} w,$$

viewing w at p via the normal coordinates. In particular,

$$D_t J(0) \perp \dot{\gamma}(0).$$

Because M has constant sectional curvature c , Proposition 38.2 implies that

$$J(t) = s_c(t)P_t(D_t J(0)),$$

where P_t denotes parallel transport along γ . Hence

$$|J(t)|_g^2 = s_c(t)^2 |D_t J(0)|_{g_p}^2.$$

Evaluating at $t = b$ gives

$$|w|_{g_q}^2 = |J(b)|_{g_q}^2 = |D_t J(0)|_{g_p}^2 s_c(b)^2.$$

Since

$$D_t J(0) = \frac{1}{b} w,$$

we get

$$|D_t J(0)|_{g_p}^2 = \frac{1}{b^2} |w|_{g_p}^2.$$

Therefore

$$|w|_{g_q}^2 = |J(b)|_{g_q}^2 = |D_t J(0)|_{g_p}^2 |s_c(b)|^2 = \frac{|w|_{g_p}^2}{b^2} |s_c(b)|^2.$$

By definition of $\hat{g}$,

$$\hat{g}(w, w) = |d\pi(w)|_{S^{n-1}}^2 = \left(\frac{1}{b} |w|_{g_p} \right)^2 = \frac{1}{b^2} |w|_{g_p}^2$$

Multiply both sides by b^2 ,

$$|w|_{g_p}^2 = b^2 \hat{g}(w, w).$$

Substituting this into the previous identity yields

$$|w|_{g_q}^2 = s_c(b)^2 \hat{g}(w, w).$$

Since this holds for every $w \perp \partial_r$, we conclude that on the tangent space of the geodesic sphere of radius b ,

$$\tilde{g} = s_c(b)^2 \hat{g}.$$

Replacing b by the radial variable r , we obtain

$$g = dr^2 + s_c(r)^2 \hat{g}.$$

□

Corollary 38.5. *Let (M, g) and $(\widetilde{M}, \tilde{g})$ be n -dimensional Riemannian manifolds with the same constant sectional curvature c . Then g and $\tilde{g}$ are locally isometric.*

Proof. Fix points $p \in M$ and $\tilde{p} \in \widetilde{M}$. Let $U \ni p$ and $\tilde{U} \ni \tilde{p}$ be geodesic balls of the same radius ε . Then the exponential maps give diffeomorphisms

$$\varphi = \exp_p^{-1} : U \rightarrow B_\varepsilon(0), \quad \tilde{\varphi} = \exp_{\tilde{p}}^{-1} : \tilde{U} \rightarrow B_\varepsilon(0).$$

By the theorem, both metrics are given on $B_\varepsilon(0)$ by the same formula

$$dr^2 + s_c(r)^2 \hat{g}.$$

In other words,

$$\varphi_*g = \tilde{\varphi}_*\tilde{g}.$$

Therefore

$$\tilde{\varphi}^{-1} \circ \varphi : U \rightarrow \tilde{U}$$

is an isometry. □

Corollary 38.6. *A space of constant sectional curvature is locally isometric to the sphere ($c > 0$), Euclidean space ($c = 0$), or hyperbolic space ($c < 0$).*

Remark 38.7. We already proved the case $c = 0$ in Theorem 25.1 above; the second proof secretly used Jacobi fields.

39. CONJUGATE POINTS (TU 4/21)

Question. How far out is the exponential map a local diffeomorphism? (The global question has an extra layer of subtlety.)

Definition 39.1. Let $\gamma : [a, b] \rightarrow M$ be a geodesic, and write $p = \gamma(a)$, $q = \gamma(b)$. We say that p and q are **conjugate along** γ if there exists a nonzero normal Jacobi field

$$J \in \mathcal{J}^\perp(\gamma)$$

such that

$$J(a) = 0 = J(b).$$

The **order of conjugacy** of p and q along γ is the dimension of the space of such Jacobi fields.

Proposition 39.2. *Let $p \in M$ and let $v \in \mathcal{E}_p$. Then v is a critical point of $\exp_p$ if and only if*

$$q = \exp_p(v)$$

is conjugate to p along γ_v .

Proof. Suppose first that v is a critical point of $\exp_p$. Then there exists a nonzero vector

$$w \in T_v\mathcal{E}_p \cong T_pM$$

such that

$$d(\exp_p)_v(w) = 0.$$

Decompose $w = av + w^\perp$, where $w^\perp \perp v$. By Proposition 19.3,

$$d(\exp_p)_v(v) = \dot{\gamma}_v(1) \neq 0.$$

Moreover, by Gauss's Lemma, $d(\exp_p)_v(v)$ is orthogonal to $d(\exp_p)_v(w^\perp)$. Since

$$0 = d(\exp_p)_v(w) = a d(\exp_p)_v(v) + d(\exp_p)_v(w^\perp),$$

we must have $a = 0$. Hence we may assume from the start that $w \perp v$.

Consider the variation through geodesics

$$\Gamma(s, t) = \exp_p(t(v + sw)).$$

Its variation field

$$J(t) = \partial_s \Gamma(0, t)$$

is a Jacobi field along γ_v . We have

$$J(0) = 0, \quad D_t J(0) = w, \quad J(1) = d(\exp_p)_v(w) = 0.$$

Since $w \perp v = \dot{\gamma}_v(0)$, Proposition 37.8 implies that J is normal. Thus p and q are conjugate along γ_v .

Conversely, suppose that $q = \exp_p(v)$ is conjugate to p along γ_v . Then there exists a nonzero normal Jacobi field J along γ_v such that

$$J(0) = 0 = J(1).$$

Since the interval $[0, 1]$ is compact, Proposition 37.4 implies that J is the variation field of a variation through geodesics. Because $J(0) = 0$, we may take this variation to be of the form

$$\Gamma(s, t) = \exp_p(tW(s))$$

for some curve $W(s)$ in $T_p M$ with $W(0) = v$. Setting

$$w = W'(0),$$

we obtain

$$J(1) = d(\exp_p)_v(w) = 0.$$

Since J is nonzero, uniqueness for the Jacobi equation implies that $w = D_t J(0)$ is nonzero. Therefore v is a critical point of $\exp_p$. $\square$

Corollary 39.3. *Let $p \in M$ and let $v \in \mathcal{E}_p$. Then $\exp_p$ is a local diffeomorphism near v if and only if p and*

$$q = \exp_p(v)$$

are not conjugate along γ_v .

Proof. By Proposition 39.2, the points p and q are not conjugate along γ_v if and only if v is not a critical point of $\exp_p$, i.e. if and only if

$$d(\exp_p)_v : T_v \mathcal{E}_p \rightarrow T_q M$$

is an isomorphism. By the inverse function theorem, this is equivalent to saying that $\exp_p$ is a local diffeomorphism near v . $\square$

40. THE CARTAN-HADAMARD THEOREM (TU 4/21-TH 4/23)

Lemma 40.1. *Suppose that (M, g) is complete with*

$$\sec_p(\Pi) \leq 0$$

for all $p \in M$ and all 2-planes $\Pi \subset T_p M$. Then for every $p \in M$,

$$\exp_p : T_p M \rightarrow M$$

is a local diffeomorphism.

Proof. Since (M, g) is complete, every geodesic starting at p is defined for all time, so $\mathcal{E}_p = T_p M$. By the previous corollary, it suffices to show that there are no conjugate points along any geodesic issuing from p .

Let γ be a unit-speed geodesic with $\gamma(0) = p$, and let $J \in \mathcal{J}^\perp(\gamma)$ be a normal Jacobi field with

$$J(0) = 0, \quad D_t J(0) \neq 0.$$

Set

$$f(t) = |J(t)|^2.$$

Then

$$f'(t) = 2 \langle D_t J, J \rangle$$

and hence

$$\begin{aligned} f''(t) &= 2 \langle D_t^2 J, J \rangle + 2 |D_t J|^2 \\ &= -2 \langle \mathcal{R}(J, \dot{\gamma}) \dot{\gamma}, J \rangle + 2 |D_t J|^2 \\ &= -2 Rm(J, \dot{\gamma}, J, \dot{\gamma}) + 2 |D_t J|^2. \end{aligned}$$

Since $J \perp \dot{\gamma}$ and the sectional curvature is nonpositive, we have

$$Rm(J, \dot{\gamma}, J, \dot{\gamma}) \leq 0.$$

Therefore

$$f''(t) \geq 0.$$

Also,

$$f(0) = 0, \quad f'(0) = 2 \langle D_t J(0), J(0) \rangle = 0.$$

Because $D_t J(0) \neq 0$, we have

$$f''(0) = 2 |D_t J(0)|^2 > 0.$$

Hence $f'(t) > 0$ for all sufficiently small $t > 0$. Since $f'' \geq 0$, the function f' is nondecreasing, so in fact

$$f'(t) > 0 \quad \text{for all } t > 0.$$

It follows that $f(t) > 0$ for all $t > 0$, i.e.

$$J(t) \neq 0 \quad \text{for all } t > 0.$$

Thus there are no conjugate points to p along γ . Since γ was arbitrary, the result follows. $\square$

Definition 40.2. A Riemannian covering map

$$\pi : (N, h) \rightarrow (M, g)$$

is a local isometry that is also a topological covering map; i.e., for every $p \in M$, there exists an open neighborhood $U \ni p$, called an **evenly covered neighborhood**, such that

$$\pi^{-1}(U) = \bigsqcup_{\alpha} \tilde{U}_{\alpha}$$

is a disjoint union of open sets, and each restriction

$$\pi|_{\tilde{U}_{\alpha}} : \tilde{U}_{\alpha} \rightarrow U$$

is a diffeomorphism.

Warning: Not every local isometry is a covering map. For example, let

$$M = (-1, 1), \quad N = \bigsqcup_{i \in \mathbb{N}} \left(-\frac{1}{i}, \frac{1}{i} \right),$$

and define $\pi : N \rightarrow M$ by $\pi(x) = x$ on each connected component. Then π is a local isometry, since on each component it is just the identity map onto an open interval in M . However, π is not a covering map: for any neighborhood U of $0 \in M$, infinitely many components of N map into U , but only finitely many of them map “onto” U , so U is not evenly covered.

The following Lemma shows that this does not happen in the present case. The proof is elementary but involves a long sequence of claims.

Lemma 40.3. *Suppose (M, g) and (N, h) are nonempty connected manifolds and let*

$$\pi : (N, h) \rightarrow (M, g)$$

be a local isometry. If (N, h) is complete, then (M, g) is complete and π is a Riemannian covering map.

Proof. Claim. Given a geodesic $\gamma : I \rightarrow M$, a point $a \in I$, and a point

$$x \in \pi^{-1}(\gamma(a)),$$

there exists a unique geodesic

$$\tilde{\gamma} : I \rightarrow N$$

such that

$$\tilde{\gamma}(a) = x, \quad \pi \circ \tilde{\gamma} = \gamma.$$

Proof of claim. Since π is a local isometry, $d\pi_x : T_x N \rightarrow T_{\gamma(a)} M$ is an isomorphism. Hence there is a unique vector

$$v \in T_x N \quad \text{with} \quad d\pi_x(v) = \dot{\gamma}(a).$$

By Hopf–Rinow, (N, h) is geodesically complete, so there exists a geodesic

$$\sigma : \mathbb{R} \rightarrow N$$

with

$$\sigma(a) = x, \quad \dot{\sigma}(a) = v.$$

Restrict σ to I , and call the restriction $\tilde{\gamma}$. Then $\pi \circ \tilde{\gamma}$ is a geodesic in M with the same initial position and velocity at time a as γ , so uniqueness of geodesics gives

$$\pi \circ \tilde{\gamma} = \gamma.$$

If $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ are two such lifts, then both are geodesics with

$$\tilde{\gamma}_1(a) = x = \tilde{\gamma}_2(a),$$

and since

$$d\pi_x(\dot{\tilde{\gamma}}_1(a)) = \dot{\gamma}(a) = d\pi_x(\dot{\tilde{\gamma}}_2(a)),$$

the injectivity of $d\pi_x$ implies

$$\dot{\tilde{\gamma}}_1(a) = \dot{\tilde{\gamma}}_2(a).$$

Hence $\tilde{\gamma}_1 = \tilde{\gamma}_2$. ◇

Claim. (M, g) is complete.

Proof of claim. Choose a point

$$p \in \pi(N),$$

and let $x \in N$ satisfy $\pi(x) = p$. Let

$$\gamma : [0, b) \rightarrow M$$

be any maximal geodesic with $\gamma(0) = p$. By Claim 1, there is a unique geodesic

$$\tilde{\gamma} : [0, \infty) \rightarrow N$$

such that

$$\tilde{\gamma}(0) = x, \quad \pi \circ \tilde{\gamma} = \gamma$$

on $[0, b)$. Since $\pi \circ \tilde{\gamma}$ is defined for all $t \geq 0$, it extends γ for all time. Thus every geodesic starting at p is complete, i.e.

$$\mathcal{E}_p = T_p M.$$

Since M is connected, Corollary 23.7 implies that (M, g) is complete. $\diamond$

Claim. π is surjective.

Proof of claim. Let $p \in \pi(N)$. Since (M, g) is complete and connected by Claim 2, Proposition 23.3 implies that for every $q \in M$ there exists a minimizing geodesic

$$\gamma : [0, L] \rightarrow M$$

from p to q . Choose $x \in N$ with $\pi(x) = p$. By Claim 1, there is a unique lift

$$\tilde{\gamma} : [0, L] \rightarrow N$$

such that

$$\tilde{\gamma}(0) = x, \quad \pi \circ \tilde{\gamma} = \gamma.$$

Therefore

$$\pi(\tilde{\gamma}(L)) = \gamma(L) = q.$$

Since $q \in M$ was arbitrary, π is surjective. $\diamond$

We will now proceed to show that π is a Riemannian covering map. Fix $p \in M$, and choose

$$U = B_\varepsilon(p), \quad 0 < \varepsilon < \text{inj}_p(M),$$

so that U is a normal neighborhood of p . Write

$$\pi^{-1}(p) = \{\tilde{p}_\alpha\}_{\alpha \in A}.$$

For each $\alpha \in A$, let

$$\tilde{U}_\alpha = B_\varepsilon(\tilde{p}_\alpha).$$

In the remaining claims, we check that these choices fulfill the requirements in the definition of a covering map. We first check the fact that the union is disjoint.

Claim. If $\tilde{p}_\alpha \neq \tilde{p}_\beta$, then

$$\tilde{U}_\alpha \cap \tilde{U}_\beta = \emptyset.$$

Proof of claim. Suppose instead that

$$\tilde{U}_\alpha \cap \tilde{U}_\beta \neq \emptyset.$$

Then by the triangle inequality,

$$d_h(\tilde{p}_\alpha, \tilde{p}_\beta) < 2\varepsilon.$$

Since (N, h) is complete, Proposition 23.3 gives a minimizing unit-speed geodesic

$$\tilde{\gamma}: [0, L] \rightarrow N$$

from $\tilde{p}_\alpha$ to $\tilde{p}_\beta$, where

$$L = d_h(\tilde{p}_\alpha, \tilde{p}_\beta) < 2\varepsilon.$$

Set

$$\gamma = \pi \circ \tilde{\gamma}.$$

Then γ is a unit-speed geodesic from p to p . If its image were contained in U , then since U is a normal neighborhood, the only geodesic in U joining p to p would be the constant one, contradiction. Hence γ must leave U . Choose $t_0 \in (0, L)$ such that

$$\gamma(t_0) \notin U.$$

Then

$$t_0 = L_g(\gamma|_{[0, t_0]}) \geq d_g(p, \gamma(t_0)) \geq \varepsilon,$$

and similarly

$$L - t_0 = L_g(\gamma|_{[t_0, L]}) \geq d_g(\gamma(t_0), p) \geq \varepsilon.$$

Therefore

$$L \geq 2\varepsilon,$$

contradicting $L < 2\varepsilon$. ◇

Claim.

$$\bigcup_{\alpha \in A} \tilde{U}_\alpha \subseteq \pi^{-1}(U).$$

Proof of claim. Let $\widehat{q} \in \tilde{U}_\alpha$. Then there exists a minimizing geodesic

$$\widehat{\delta}: [0, r] \rightarrow N$$

from $\tilde{p}_\alpha$ to $\widehat{q}$, with

$$r = d_h(\tilde{p}_\alpha, \widehat{q}) < \varepsilon.$$

Thus

$$L_g(\pi \circ \widehat{\delta}) = L_h(\widehat{\delta}) = r < \varepsilon.$$

Since $\pi(\tilde{p}_\alpha) = p$, it follows that

$$\pi(\widehat{q}) \in U.$$

Hence

$$\widehat{q} \in \pi^{-1}(U).$$

◇

Claim.

$$\pi^{-1}(U) \subseteq \bigcup_{\alpha \in A} \tilde{U}_\alpha.$$

Proof of claim. Let $\widehat{q} \in \pi^{-1}(U)$, and write

$$q = \pi(\widehat{q}).$$

Then

$$r = d_g(p, q) < \varepsilon.$$

Let $\gamma : [0, r] \rightarrow U$ be the unique unit-speed radial minimizing geodesic from q to p . By Claim 1, there is a unique lift

$$\widehat{\delta} : [0, r] \rightarrow N$$

such that

$$\widehat{\delta}(0) = \widehat{q}, \quad \pi \circ \widehat{\delta} = \gamma.$$

Then

$$\pi(\widehat{\delta}(r)) = \gamma(r) = p,$$

so

$$\widehat{\delta}(r) = \widetilde{p}_\alpha$$

for some $\alpha \in A$. Therefore

$$d_h(\widehat{q}, \widetilde{p}_\alpha) \leq L_h(\widehat{\delta}) = L_g(\gamma) = r < \varepsilon,$$

and hence

$$\widehat{q} \in \tilde{U}_\alpha$$

for the same α . ◇

Claim. For each $\alpha \in A$, the restriction

$$\pi|_{\tilde{U}_\alpha} : \tilde{U}_\alpha \rightarrow U$$

is a diffeomorphism.

Proof of claim. It is a local diffeomorphism because π is a local isometry. Surjectivity and injectivity were proved above. Explicitly, if

$$q = \gamma(r) \in U,$$

where $\gamma : [0, r] \rightarrow U$ is the unique radial unit-speed minimizing geodesic from p to q , then the inverse image of q in $\tilde{U}_\alpha$ is

$$\widetilde{\gamma}(r),$$

where $\widetilde{\gamma}$ is the lift of γ with

$$\widetilde{\gamma}(0) = \widetilde{p}_\alpha.$$

Hence

$$(\pi|_{\tilde{U}_\alpha})^{-1}(\gamma(r)) = \widetilde{\gamma}(r).$$

Therefore

$$\pi|_{\tilde{U}_\alpha} : \tilde{U}_\alpha \rightarrow U$$

is a diffeomorphism. ◇

By Claims 4, 5, 6, and 7, we have

$$\pi^{-1}(U) = \bigsqcup_{\alpha \in A} \tilde{U}_\alpha,$$

and each restriction

$$\pi|_{\tilde{U}_\alpha} : \tilde{U}_\alpha \rightarrow U$$

is a diffeomorphism. Therefore U is evenly covered. Since $p \in M$ was arbitrary, π is a Riemannian covering map. $\square$

Proposition 40.4. *Let*

$$\pi : (N, h) \rightarrow (M, g)$$

be a Riemannian covering map between connected Riemannian manifolds. Then

$$(N, h) \text{ is complete} \iff (M, g) \text{ is complete}.$$

Proof. $(\Rightarrow)$ This is Lemma 40.3.

$(\Leftarrow)$ Assume (M, g) is complete, and let

$$\tilde{\gamma} : [0, b) \rightarrow N$$

be a maximal geodesic. Then

$$\gamma = \pi \circ \tilde{\gamma}$$

is a geodesic in M , so by completeness it extends for all time. Since π is a covering map, the extended geodesic lifts uniquely to a path

$$\hat{\gamma} : [0, \infty) \rightarrow N$$

starting at $\tilde{\gamma}(0)$. On $[0, b)$, both $\hat{\gamma}$ and $\tilde{\gamma}$ are lifts of γ with the same initial point, so they agree. Since π is locally a diffeomorphism and local isometry, the lifted extension is again a geodesic. Thus every geodesic in N extends for all time. By Hopf–Rinow, (N, h) is complete. $\square$

We can now give the following topological characterization of negatively curved spaces.

Theorem 40.5 (Cartan–Hadamard). *Let (M^n, g) be a complete, connected Riemannian manifold with*

$$\sec_p(\Pi) \leq 0$$

for every $p \in M$ and every 2-plane $\Pi \subset T_p M$. Then the universal cover $\tilde{M}$ of M is diffeomorphic to $\mathbb{R}^n$.

Proof. Choose $p \in M$. Since

$$T_p M \cong \mathbb{R}^n, \quad \pi_1(T_p M) = \{e\},$$

it suffices to show that

$$\exp_p : T_p M \rightarrow M$$

is a covering map.

By Lemma 40.1,

$$\exp_p : T_p M \rightarrow M$$

is a local diffeomorphism. Define a Riemannian metric on $T_p M$ by

$$\tilde{g} := \exp_p^* g.$$

Then

$$\exp_p : (T_p M, \tilde{g}) \rightarrow (M, g)$$

is a local isometry.

Claim. $(T_p M, \tilde{g})$ is complete.

Proof of claim. Let $v \in T_0(T_p M) \cong T_p M$. The line

$$\ell_v(t) = tv$$

satisfies

$$\exp_p(\ell_v(t)) = \exp_p(tv) = \gamma_v(t),$$

where γ_v is the geodesic in M with

$$\gamma_v(0) = p, \quad \dot{\gamma}_v(0) = v.$$

Since (M, g) is complete, γ_v is defined for all $t \in \mathbb{R}$. Because $\exp_p$ is a local isometry, the curve

$$t \mapsto tv$$

is a geodesic in $(T_p M, \tilde{g})$. Thus every geodesic starting at the origin is defined for all time. By Corollary 23.7, $(T_p M, \tilde{g})$ is complete. $\diamond$

Therefore Lemma 40.3 implies that

$$\exp_p : T_p M \rightarrow M$$

is a Riemannian covering map. Since $T_p M \cong \mathbb{R}^n$ is a simply-connected covering space of M , it is the universal cover. $\square$

Corollary 40.6. *Let (M, g) be a complete, connected Riemannian manifold with nonpositive sectional curvature. Then M is diffeomorphic to a quotient of $\mathbb{R}^n$ by a free and proper group action.*

Proof. By the Cartan–Hadamard theorem, the universal cover $\tilde{M}$ is diffeomorphic to $\mathbb{R}^n$. The deck transformation group (isomorphic to $\pi_1(M)$) acts freely and properly on $\tilde{M}$, and the quotient is M . $\square$

Example 40.7. Suppose (M, g) has constant sectional curvature

$$\text{sec} \equiv -1.$$

Choose $p \in M$. Then the universal cover $(T_p M, \tilde{g})$ also has constant sectional curvature -1 , and $\tilde{g}$ is in normal coordinates centered at 0. Theorem 38.4 shows that

$$\tilde{g} = dr^2 + \sinh^2(r)\hat{g},$$

which is exactly the hyperbolic metric. In particular, for a closed surface Σ_g of genus $g \geq 2$, the uniformization metric has universal cover isometric to $\mathbb{H}^2$.

Corollary 40.8. *A complete, connected Riemannian manifold with nonpositive sectional curvature is aspherical; that is,*

$$\pi_i(M) = 0 \quad \text{for all } i > 1.$$

Proof. The universal cover of M is diffeomorphic to $\mathbb{R}^n$ by Cartan–Hadamard, hence contractible. Therefore all higher homotopy groups of M vanish. $\square$

Corollary 40.9. *Let (M, g) be a compact, connected Riemannian manifold of positive dimension with nonpositive sectional curvature. Then $\pi_1(M)$ is infinite.*

Proof. If $\pi_1(M)$ were finite, then the universal cover $\widetilde{M}$ would also be compact (exercise). But by Cartan–Hadamard,

$$\widetilde{M} \cong \mathbb{R}^n,$$

which is noncompact for $n > 0$. This is a contradiction. $\square$

Example 40.10. We saw above that compact surfaces of nonpositive curvature do exist. However, by the last corollary, these must have infinite fundamental group. By the classification of surfaces, this is equivalent to having genus $g \geq 1$. (We already knew this from Gauss–Bonnet, of course.)

Corollary 40.11. *Let M and N be compact, connected manifolds of positive dimension, and assume that at least one of them is simply connected. Then $M \times N$ admits no Riemannian metric of nonpositive sectional curvature.*

Proof. Without loss of generality, assume M is simply connected. Since M is compact and simply connected, it is orientable. If $M \times N$ admitted a metric of nonpositive sectional curvature, then by the previous corollary its universal cover would be contractible. On the other hand, the universal cover of $M \times N$ is

$$M \times \widetilde{N},$$

since M is simply connected. Fix a point $\tilde{q} \in \widetilde{N}$. Then

$$M \cong M \times \{\tilde{q}\} \subset M \times \widetilde{N}$$

is a retract via projection onto the first factor. Hence M would be contractible. This is impossible, because M has nonzero top homology. $\square$

Alternative proof of the previous corollary. Without loss of generality, assume M is simply connected. Since M is compact and simply connected, it is orientable. Let

$$\widetilde{N} \rightarrow N$$

be the universal cover, and consider the covering map

$$\pi : M \times \widetilde{N} \rightarrow M \times N.$$

If $M \times N$ admitted a metric g with

$$\sec \leq 0,$$

then g would be complete by compactness. Hence the pullback metric

$$\pi^*g$$

on $M \times \tilde{N}$ is also complete, and still has nonpositive sectional curvature. By Cartan–Hadamard,

$$M \times \tilde{N} \cong \mathbb{R}^n.$$

Let μ be a volume form on M . Then $d\mu = 0$ for degree reasons, and after choosing the orientation we may arrange that

$$\int_M \mu > 0.$$

Let

$$\rho : M \times \tilde{N} \rightarrow M$$

be projection onto the first factor. Then $\rho^*\mu$ is a closed form on $M \times \tilde{N} \cong \mathbb{R}^n$, so by the Poincaré lemma it is exact. Fix $y_0 \in \tilde{N}$, and define

$$\sigma : M \rightarrow M \times \tilde{N}, \quad \sigma(x) = (x, y_0).$$

Then

$$\int_{\sigma(M)} \rho^*\mu = \int_M \sigma^* \rho^*\mu = \int_M (\rho \circ \sigma)^*\mu = \int_M \mu > 0.$$

On the other hand, $\rho^*\mu$ is exact and $\sigma(M)$ is closed, so Stokes’s theorem gives

$$\int_{\sigma(M)} \rho^*\mu = 0,$$

a contradiction. □

41. COMPARISON THEOREMS (TU 4/28)

Question: What happens if the sectional curvature has a positive upper bound, i.e.

$$\sec \leq C = 1/R^2 > 0?$$

In this case, the geodesics are allowed to come together, but not too fast. For example, consider the sphere of radius R . The larger its radius (hence the smaller the curvature), the longer it takes for two distinct geodesics starting at the same point to meet.

In fact, comparing a Riemannian manifold to one with constant curvature is a good strategy to address this question. In that direction, we have the so-called “comparison theorems.” We will base all of our comparison theorems on the following simple lemma due to Sturm, sometimes called “ODE comparison.”

Lemma 41.1 (Sturm comparison theorem). *Let $u, v : [0, T] \rightarrow \mathbb{R}$ with $u(t) > 0 \forall 0 < t < T$. If*

$$\ddot{u} + a(t)u = 0$$

$$\ddot{v} + a(t)v \geq 0,$$

$u(0) = v(0) = 0$ and $\dot{u}(0) = \dot{v}(0) > 0$, then $u(t) \leq v(t) \forall t \in [0, T]$.

Proof. Let

$$f(t) = \frac{v(t)}{u(t)}.$$

By L'Hôpital's rule,

$$\lim_{t \rightarrow 0} f(t) = \lim_{t \rightarrow 0} \frac{\dot{v}(t)}{\dot{u}(t)} = 1.$$

We also compute

$$\dot{f} = \frac{\dot{v}u - v\dot{u}}{u^2}.$$

Differentiating the numerator,

$$(\dot{v}u - v\dot{u}) = \ddot{v}u - v\ddot{u} \geq 0$$

And also $(\dot{v}u - v\dot{u})(0) = 0$, which implies that $\dot{f}(t) \geq 0$, thus $f(t) \geq 1 \forall t$. $\square$

Theorem 41.2 (Jacobi-field comparison theorem). *Assume $\sec \leq c$ and $J(t) \in \mathcal{J}_0^\perp(\gamma)$ (i.e. it vanishes at $t = 0$: $J(0) = 0$). Also assume $|\dot{\gamma}| \equiv 1$. Then*

$$|J(t)| \geq s_c(t)|D_t J(0)|.$$

Proof. $|J(t)|$ is smooth wherever it is nonzero, therefore, at all these points,

$$\begin{aligned} \frac{d}{dt}|J(t)| &= \frac{\langle D_t J, J \rangle}{|J|}, \\ \frac{d^2}{dt^2}|J(t)| &= \frac{\langle D_t^2 J, J \rangle}{|J|} + \frac{|D_t J|^2}{|J|} - \frac{\langle D_t J, J \rangle^2}{|J|^3}. \end{aligned}$$

By the Cauchy-Schwarz inequality, $\langle D_t J, J \rangle^2 \leq |D_t J|^2 |J|^2$. Also, since J is normal,

$$\sec\{J, \dot{\gamma}\} = \frac{\langle \mathcal{R}(J, \dot{\gamma})\dot{\gamma}, J \rangle}{|J|^2}$$

Consequently,

$$\frac{d^2}{dt^2}|J(t)| \geq -\sec\{J, \dot{\gamma}\}|J| \geq -c|J| \implies \frac{d^2}{dt^2}|J(t)| + c|J| \geq 0.$$

Also, since $J(0) = 0$,

$$\left. \frac{d}{dt}|J(t)| \right|_{t=0} = |D_t J(0)|.$$

By the Sturm comparison theorem, $|J(t)| \geq s_c(t)|D_t J(0)|$ as long as $s_c(t) \geq 0$. $\square$

Corollary 41.3 (Conjugate-point comparison theorem). *If $\sec \leq c$, then*

- *If $c \leq 0$, there are no conjugate points (restatement of the Cartan-Hadamard theorem).*
- *If $c = 1/R^2 > 0$, there are no conjugate points before distance πR .*

Corollary 41.4 (Metric comparison theorem). *For a given Riemannian manifold (M, g) , let $g_c = dr^2 + s_c(r)^2 \hat{g}$ and $\sec \leq c$. Then, in normal coordinates,*

$$g(v, v) \geq g_c(v, v).$$

Proof. It is essentially the same as the proof of the uniqueness of constant sectional curvature spaces. $\square$

One might ask whether one can do comparison theorems with *lower* bounds on the sectional curvature instead of upper bounds. In this case, we have the following theorem, however the proof is more elaborate:

Theorem 41.5 (Rauch comparison theorem). *If $0 < L \leq \sec$, then the distance along a geodesic to the first conjugate point is at most $\pi/\sqrt{L}$.*

We will actually prove a version of this result using a different method, next class. You also have a homework exercise that proves it in dimension two.

42. GLOBAL UNIQUENESS OF CONSTANT SECTIONAL CURVATURE (Tu 4/28)

As an application of the conjugate point comparison, we can prove:

Theorem 42.1 (Hopf). *Let (M, g) be complete with $\pi_1(M) = \{e\}$ and constant sectional curvature $\sec \equiv c$. Then M is isometric to*

$$\begin{cases} \mathbb{R}^n, & \text{if } c = 0 \\ \mathbb{S}_R^n, & \text{if } c = 1/R^2 > 0 \\ \mathbb{H}_R^n, & \text{if } c = -1/R^2 < 0. \end{cases}$$

Proof. If $c \leq 0$, by the Cartan-Hadamard theorem, $\exp_p : T_p M \cong \mathbb{R}^n \rightarrow M$ is an isometry, with respect to $\tilde{g} = \exp_p^* g$. Also, by construction, euclidean coordinates on $\mathbb{R}^n$ are normal coordinates for $\tilde{g}$, centered at 0. By local uniqueness, $g = g_c$.

If $c = 1/R^2 > 0$, let $\{N, -N\} \subset S_R^n$ and consider the exponential map

$$\exp_N : B_{\pi R}(0) \rightarrow S_R^n \setminus \{-N\}.$$

Choose $\phi : T_N S_R^n \rightarrow T_p M$ to be a linear isometry and consider on the ball $B_{\pi R}(0) \subset T_N S_R^n$ the metrics

$$\exp_N^* g_{S_R^n} \text{ and } (\exp_p \circ \phi)^* g.$$

Both have constant sectional curvature $\equiv 1/R^2$ and, in normal coordinates, they are both equal to g_c .

Now we construct

$$\Phi := \exp_p \circ \phi \circ \exp_N^{-1} : S_R^n \setminus \{-N\} \rightarrow \text{Im}(\Phi) \subset M,$$

which, by construction, it is a local isometry. Pick any point $Q \in S_R^n \setminus \{N, -N\}$. Let $q = \Phi(Q)$ and consider

$$\tilde{\phi} = d\Phi_Q : T_Q S_R^n \rightarrow T_q M.$$

Since Φ is a local isometry, $\tilde{\phi}$ is a linear isometry. Construct

$$\tilde{\Phi} := \exp_p \circ \tilde{\phi} \circ \exp_Q^{-1} : S_R^n \setminus \{-Q\} \rightarrow \text{Im}(\tilde{\Phi}) \subset M,$$

which is a local isometry. We have:

$$\tilde{\Phi}(Q) = \Phi(Q) = q \text{ and } d\Phi_Q = d\tilde{\Phi}_Q$$

and both $\Phi, \tilde{\Phi}$ are local isometries. Moreover, $S_R^n \setminus \{-N, -Q\}$ is connected (if we assume, without loss of generality, that $n \geq 2$). By one of your homework exercises, $\Phi = \tilde{\Phi}$ on $S_R^n \setminus \{-N, -Q\}$. Because of that, we can define $\Psi : S_R^n \rightarrow M$, with

$$\Psi = \begin{cases} \Phi, & \text{on } S_R^n \setminus \{-N\} \\ \tilde{\Phi}, & \text{on } S_R^n \setminus \{-Q\}, \end{cases}$$

which is a globally defined local isometry from S_R^n to M . Since S_R^n is complete, Ψ is a covering map. Lastly, since $\pi_1(M) = \{e\}$, it follows that Ψ is an isometry. $\square$

Recall that on the first day we discussed spaces with maximal isometry groups (i.e. homogeneous spaces with $O(n)$ stabilizer at each point), and I claimed to you that the model spaces are the only such spaces. We can now prove this; in fact a weaker condition on the stabilizers is sufficient.

Corollary 42.2. *A homogeneous Riemannian manifold (M, g) such that the stabilizer Isom_p acts transitively on 2-planes in $T_p M$, is locally isometric to $S_R^n, \mathbb{R}^n$, or $\mathbb{H}_R^n$. If $\pi_1(M) = \{e\}$, then it is globally isometric.*

Proof. At each point $p \in M$, the sectional curvatures $\sec_p \Pi$ are constant with respect to any $\Pi \subset T_p M$, by the transitivity condition. $\square$

Remark 42.3. Let's take stock of what we have proven so far about sectional curvature. Suppose that $\pi_1(M) = \{e\}$ and M is complete.

- The Cartan-Hadamard theorem states that if $\sec \leq 0$ then $M \cong \mathbb{R}^n$, and in particular, is noncompact.
- We have just shown that if $\sec \equiv c > 0$ is constant and positive, then $M \cong S_R^n$, and in particular, is compact.
- We know from homework that for positive sectional curvature, we cannot completely remove the assumption of *constancy*: indeed, $\mathbb{CP}^n$ is simply connected, compact, and has positive sectional curvature, but is not homeomorphic to a sphere.
- Nevertheless, next time we will see that (not necessarily constant) strictly positive sectional curvature still implies compactness.

43. SECOND VARIATION OF ARCLength (TH 4/30 ONLINE)

Theorem 43.1 (Second variation of arclength). *Let $\gamma : [a, b] \rightarrow M$ be a unit-speed geodesic, and let Γ be a proper admissible variation of γ . Write*

$$V(t) = \frac{\partial \Gamma}{\partial s}(0, t) = V^\top(t) + V^\perp(t), \quad V^\top = \langle V, \dot{\gamma} \rangle \dot{\gamma}.$$

Then

$$\left. \frac{d^2}{ds^2} L(\Gamma_s) \right|_{s=0} = \int_a^b (|D_t V^\perp|^2 - \text{Rm}(V^\perp, \dot{\gamma}, V^\perp, \dot{\gamma})) dt.$$

Proof. Let

$$T = \partial_t \Gamma, \quad S = \partial_s \Gamma.$$

Fix one of the smooth subintervals $[a_{i-1}, a_i]$. As in the proof of Theorem 21.4,

$$\frac{d}{ds} L(\Gamma_s|_{[a_{i-1}, a_i]}) = \int_{a_{i-1}}^{a_i} \frac{\langle D_t S, T \rangle}{|T|} dt.$$

Differentiating once more and evaluating at $s = 0$, we obtain

$$\begin{aligned} \left. \frac{d^2}{ds^2} L(\Gamma_s|_{[a_{i-1}, a_i]}) \right|_{s=0} &= \int_{a_{i-1}}^{a_i} \left(\frac{\langle D_s D_t S, T \rangle}{|T|} + \frac{|D_t S|^2}{|T|} - \frac{\langle D_t S, T \rangle^2}{|T|^3} \right) \Big|_{s=0} dt \\ &= \int_{a_{i-1}}^{a_i} (\langle D_s D_t S, T \rangle + |D_t V|^2 - \langle D_t V, \dot{\gamma} \rangle^2) dt, \end{aligned}$$

since $|T|_{s=0} = |\dot{\gamma}| = 1$.

Next, by the same curvature identity used in the proof of Theorem 36.2,

$$D_s D_t S = D_t D_s S + \mathcal{R}(S, T)S.$$

Also, since γ is a geodesic,

$$D_t \dot{\gamma} = 0.$$

Therefore, at $s = 0$,

$$\begin{aligned} \langle D_s D_t S, T \rangle &= \langle D_t D_s S, \dot{\gamma} \rangle + \langle \mathcal{R}(V, \dot{\gamma})V, \dot{\gamma} \rangle \\ &= \frac{d}{dt} \langle D_s S, T \rangle - \text{Rm}(V, \dot{\gamma}, V, \dot{\gamma}). \end{aligned}$$

Hence

$$\begin{aligned} \left. \frac{d^2}{ds^2} L(\Gamma_s|_{[a_{i-1}, a_i]}) \right|_{s=0} &= \langle D_s S, T \rangle \Big|_{a_{i-1}}^{a_i} \\ &\quad + \int_{a_{i-1}}^{a_i} (|D_t V|^2 - \langle D_t V, \dot{\gamma} \rangle^2 - \text{Rm}(V, \dot{\gamma}, V, \dot{\gamma})) dt. \end{aligned}$$

We now rewrite the integrand in terms of $V^\perp$. Since

$$V^\top = \langle V, \dot{\gamma} \rangle \dot{\gamma}$$

and $D_t \dot{\gamma} = 0$, we have

$$D_t(V^\top) = \langle D_t V, \dot{\gamma} \rangle \dot{\gamma} = (D_t V)^\top.$$

Therefore,

$$D_t(V^\perp) = D_t V - D_t(V^\top) = (D_t V)^\perp,$$

and so

$$|D_t V|^2 = \langle D_t V, \dot{\gamma} \rangle^2 + |D_t V^\perp|^2.$$

Also, by the symmetries of Rm ,

$$Rm(V, \dot{\gamma}, V, \dot{\gamma}) = Rm(V^\perp, \dot{\gamma}, V^\perp, \dot{\gamma}).$$

Summing over i , the interior boundary terms cancel. Since Γ is a proper variation,

$$S(s, a) \equiv 0 = S(s, b),$$

so

$$D_s S|_{t=a} = 0 = D_s S|_{t=b}.$$

Therefore

$$\frac{d^2}{ds^2} L(\Gamma_s) \Big|_{s=0} = \int_a^b (|D_t V^\perp|^2 - Rm(V^\perp, \dot{\gamma}, V^\perp, \dot{\gamma})) dt.$$

□

Remark 43.2. It makes sense that only $V^\perp$ appears in the formula, since $V^\top$ corresponds to reparametrization.

Definition 43.3. Let $\gamma : [a, b] \rightarrow M$ be a geodesic. The **index form** of γ is the bilinear form on piecewise smooth vector fields V, W along γ with

$$V(a) = V(b) = W(a) = W(b) = 0$$

defined by

$$I(V, W) = \int_a^b (\langle D_t V, D_t W \rangle - Rm(V, \dot{\gamma}, W, \dot{\gamma})) dt.$$

It is symmetric by Theorem 26.1. By Theorem 43.1, if Γ is a proper admissible variation of γ , then

$$\frac{d^2}{ds^2} L(\Gamma_s) \Big|_{s=0} = I(V^\perp, V^\perp),$$

where $V = \partial_s \Gamma|_{s=0}$ is the variation field.

Remark 43.4. If γ is minimizing from $\gamma(a)$ to $\gamma(b)$, then

$$I(V, V) \geq 0$$

for every piecewise smooth vector field V along γ with

$$V(a) = V(b) = 0.$$

Indeed, by the remark before Lemma 21.3, there exists a proper admissible variation with variation field V . Since γ minimizes arclength among curves with the same endpoints, the function

$$s \mapsto L(\Gamma_s)$$

has a minimum at $s = 0$. Therefore the one-variable second derivative test and Theorem 43.1 imply that

$$0 \leq \frac{d^2}{ds^2} L(\Gamma_s) \Big|_{s=0} = I(V, V).$$

Proposition 43.5. *Let $\gamma: [a, b] \rightarrow M$ be a geodesic, and let*

$$a = a_0 < a_1 < \cdots < a_s = b$$

be a subdivision such that V is smooth on each interval $[a_{i-1}, a_i]$. Suppose also that $V(a) = V(b) = 0$, and let W be a smooth vector field along γ with $W(a) = W(b) = 0$. Then

$$I(V, W) = - \int_a^b \langle D_t^2 V + \mathcal{R}(V, \dot{\gamma})\dot{\gamma}, W \rangle dt - \sum_{i=1}^{s-1} \langle \Delta_i D_t V, W(a_i) \rangle,$$

where

$$\Delta_i D_t V := D_t V(a_i^+) - D_t V(a_i^-).$$

Proof. On each interval $[a_{i-1}, a_i]$, integration by parts gives

$$\int_{a_{i-1}}^{a_i} \langle D_t V, D_t W \rangle dt = \langle D_t V, W \rangle \Big|_{a_{i-1}}^{a_i} - \int_{a_{i-1}}^{a_i} \langle D_t^2 V, W \rangle dt.$$

Therefore

$$\begin{aligned} I(V, W) &= \sum_{i=1}^s \langle D_t V, W \rangle \Big|_{a_{i-1}}^{a_i} \\ &\quad - \int_a^b \langle D_t^2 V + \mathcal{R}(V, \dot{\gamma})\dot{\gamma}, W \rangle dt. \end{aligned}$$

Since $W(a) = W(b) = 0$, the endpoint terms vanish. At each interior point a_i , the remaining contribution is

$$\langle D_t V(a_i^-), W(a_i) \rangle - \langle D_t V(a_i^+), W(a_i) \rangle = - \langle \Delta_i D_t V, W(a_i) \rangle.$$

Summing over $i = 1, \dots, s-1$ yields the formula. $\square$

Corollary 43.6. *Let $\gamma: [a, b] \rightarrow M$ be a geodesic, and let V be a smooth vector field along γ with*

$$V(a) = V(b) = 0.$$

Then the following are equivalent:

1. $I(V, W) = 0$ for every smooth vector field W along γ with $W(a) = W(b) = 0$.
2. V is a Jacobi field along γ .

Proof. Since V is smooth, Proposition 43.5 gives

$$I(V, W) = - \int_a^b \langle D_t^2 V + \mathcal{R}(V, \dot{\gamma})\dot{\gamma}, W \rangle dt$$

for every such W . Hence (1) holds if and only if

$$D_t^2 V + \mathcal{R}(V, \dot{\gamma})\dot{\gamma} = 0,$$

which is equivalent to saying that V is a Jacobi field by Theorem 36.2. $\square$

44. GEODESICS DO NOT MINIMIZE PAST CONJUGATE POINTS (TH 4/30 ONLINE)

Theorem 44.1. *Let $\gamma : [0, b] \rightarrow M$ be a geodesic. Suppose that for some*

$$a \in (0, b)$$

the points $\gamma(0)$ and $\gamma(a)$ are conjugate along γ . Then γ does not minimize arclength from $\gamma(0)$ to $\gamma(b)$.

Proof. Since $\gamma(0)$ and $\gamma(a)$ are conjugate along γ , there exists a nonzero normal Jacobi field J along $\gamma|_{[0, a]}$ such that

$$J(0) = 0 = J(a).$$

Since J is nonzero, we must have

$$D_t J(a) \neq 0,$$

for otherwise uniqueness for the Jacobi equation with initial data at $t = a$ would imply that $J \equiv 0$.

Define a piecewise smooth vector field V along γ by

$$V(t) = \begin{cases} J(t), & 0 \leq t \leq a, \\ 0, & a \leq t \leq b. \end{cases}$$

Then $V(0) = V(b) = 0$, and V is smooth except possibly at $t = a$. Also,

$$\Delta_a D_t V = D_t V(a^+) - D_t V(a^-) = -D_t J(a).$$

Applying Proposition 43.5 with $W = V$, we obtain

$$I(V, V) = - \int_0^b \langle D_t^2 V + \mathcal{R}(V, \dot{\gamma})\dot{\gamma}, V \rangle dt - \langle \Delta_a D_t V, V(a) \rangle.$$

The boundary term vanishes because $V(a) = J(a) = 0$, and the integral vanishes because V satisfies the Jacobi equation on $[0, a]$ and is identically zero on $[a, b]$. Hence

$$I(V, V) = 0.$$

Now choose a smooth vector field W along γ such that

$$W(0) = 0 = W(b), \quad W(a) = \Delta_a D_t V = -D_t J(a).$$

This is possible because we may prescribe the value of a smooth vector field at a point. By Proposition 43.5,

$$I(V, W) = - \int_0^b \langle D_t^2 V + \mathcal{R}(V, \dot{\gamma})\dot{\gamma}, W \rangle dt - \langle \Delta_a D_t V, W(a) \rangle.$$

Again the integral vanishes, so

$$I(V, W) = -|W(a)|^2 < 0.$$

For $\varepsilon \in \mathbb{R}$, set

$$X_\varepsilon = V + \varepsilon W.$$

Then $X_\varepsilon(0) = X_\varepsilon(b) = 0$, so if γ were minimizing from $\gamma(0)$ to $\gamma(b)$, the previous remark would imply that

$$I(X_\varepsilon, X_\varepsilon) \geq 0.$$

But by bilinearity,

$$\begin{aligned} I(X_\varepsilon, X_\varepsilon) &= I(V, V) + 2\varepsilon I(V, W) + \varepsilon^2 I(W, W) \\ &= -2\varepsilon |W(a)|^2 + \varepsilon^2 I(W, W). \end{aligned}$$

Since $|W(a)|^2 > 0$, this is strictly negative for all sufficiently small positive ε , a contradiction. Therefore γ is not minimizing from $\gamma(0)$ to $\gamma(b)$. $\square$

45. THE BONNET-MYERS THEOREM (TH 4/30 ONLINE)

Definition 45.1. The **diameter** of a Riemannian manifold (M, g) is

$$\text{diam}(M) = \sup\{d_g(p, q) : p, q \in M\}.$$

Lemma 45.2 (From HW). *If (M, g) is complete and has finite diameter, then M is compact.*

Proof. Choose $p \in M$ and let $D = \text{diam}(M) < \infty$. Then

$$M = \overline{B_D(p)}.$$

Since M is complete, every $q \in M$ can be joined to p by a minimizing geodesic, so

$$M = \exp_p(\overline{B_D(0)}),$$

where the closed ball $\overline{B_D(0)} \subset T_p M \cong \mathbb{R}^n$ is compact. Since the image of a compact set under a continuous map is compact, M is compact. $\square$

Theorem 45.3 (Bonnet). *Let (M^n, g) be complete and connected with*

$$\text{sec} \geq C > 0.$$

Then

$$\text{diam}(M) \leq \frac{\pi}{\sqrt{C}}.$$

In particular, M is compact and $\pi_1(M)$ is finite.

Proof. Suppose for contradiction that

$$\text{diam}(M) > \frac{\pi}{\sqrt{C}}.$$

Then there exist points $p, q \in M$ with

$$d_g(p, q) > \frac{\pi}{\sqrt{C}}.$$

Since M is complete and connected, Theorem 23.6 and Proposition 23.3 imply that there exists a unit-speed minimizing geodesic

$$\gamma : [0, L] \rightarrow M$$

from p to q , where

$$L = d_g(p, q) > \frac{\pi}{\sqrt{C}}.$$

Choose a parallel unit vector field E along γ such that

$$E \perp \dot{\gamma}.$$

Define

$$V(t) = \sin\left(\frac{\pi t}{L}\right) E(t).$$

Then

$$V(0) = V(L) = 0,$$

so since γ is minimizing, the previous remark implies that

$$I(V, V) \geq 0.$$

Because E is parallel, we have

$$D_t V = \frac{\pi}{L} \cos\left(\frac{\pi t}{L}\right) E(t),$$

and therefore

$$|D_t V|^2 = \frac{\pi^2}{L^2} \cos^2\left(\frac{\pi t}{L}\right).$$

Also,

$$|V|^2 = \sin^2\left(\frac{\pi t}{L}\right),$$

and since $V \perp \dot{\gamma}$, the sectional-curvature lower bound gives

$$\begin{aligned} Rm(V, \dot{\gamma}, V, \dot{\gamma}) &= |V|^2 \sec(V \wedge \dot{\gamma}) \\ &\geq C \sin^2\left(\frac{\pi t}{L}\right). \end{aligned}$$

Hence

$$\begin{aligned} I(V, V) &= \int_0^L (|D_t V|^2 - Rm(V, \dot{\gamma}, V, \dot{\gamma})) dt \\ &\leq \int_0^L \left(\frac{\pi^2}{L^2} \cos^2\left(\frac{\pi t}{L}\right) - C \sin^2\left(\frac{\pi t}{L}\right) \right) dt \\ &= \frac{L}{2} \left(\frac{\pi^2}{L^2} - C \right). \end{aligned}$$

Since $L > \pi/\sqrt{C}$, we have

$$\frac{\pi^2}{L^2} - C < 0,$$

so

$$I(V, V) < 0,$$

contradicting the minimizing property of γ . Therefore

$$\text{diam}(M) \leq \frac{\pi}{\sqrt{C}}.$$

Since M is complete and has finite diameter, the previous corollary implies that M is compact.

Finally, let $\widetilde{M}$ be the universal cover of M , equipped with the pullback metric. It has the same sectional-curvature lower bound, and it is complete by Proposition 40.4. Applying the diameter estimate to $\widetilde{M}$, we see that $\widetilde{M}$ is compact. Therefore each fiber of the universal

covering map $\widetilde{M} \rightarrow M$ is a discrete subset of a compact space, hence finite. Since each fiber is in bijection with $\pi_1(M)$, we conclude that $\pi_1(M)$ is finite. $\square$

It is interesting to observe that the previous argument goes through under a weaker hypothesis on the curvature.

Theorem 45.4 (Myers). *Let (M^n, g) be complete and connected, and suppose that for some $C > 0$,*

$$Rc(X, X) \geq (n-1)C|X|^2$$

for all tangent vectors X . Then

$$\text{diam}(M) \leq \frac{\pi}{\sqrt{C}}.$$

In particular, M is compact and $\pi_1(M)$ is finite.

Proof. Let $p, q \in M$. By Theorem 23.6 and Proposition 23.3, there exists a unit-speed minimizing geodesic

$$\gamma : [0, \ell] \rightarrow M$$

from p to q , where $\ell = d_g(p, q)$.

Choose a parallel orthonormal frame

$$E_1, \dots, E_{n-1}, E_n = \dot{\gamma}$$

along γ . For each $i = 1, \dots, n-1$, define

$$V_i(t) = \sin\left(\frac{\pi t}{\ell}\right) E_i(t).$$

Then $V_i(0) = V_i(\ell) = 0$, so by the remark before Lemma 21.3, there exists a proper admissible variation with variation field V_i . Since γ is minimizing, $s = 0$ is a minimum of $L(\Gamma_s)$, hence

$$\left. \frac{d^2}{ds^2} L(\Gamma_s) \right|_{s=0} \geq 0.$$

By Theorem 43.1,

$$\begin{aligned} 0 &\leq \int_0^\ell (|D_t V_i|^2 - Rm(V_i, \dot{\gamma}, V_i, \dot{\gamma})) dt \\ &= \int_0^\ell \left(\frac{\pi^2}{\ell^2} \cos^2\left(\frac{\pi t}{\ell}\right) - \sin^2\left(\frac{\pi t}{\ell}\right) Rm(E_i, \dot{\gamma}, E_i, \dot{\gamma}) \right) dt. \end{aligned}$$

Summing over $i = 1, \dots, n-1$, we obtain

$$\begin{aligned} 0 &\leq \int_0^\ell \left((n-1) \frac{\pi^2}{\ell^2} \cos^2\left(\frac{\pi t}{\ell}\right) - \sin^2\left(\frac{\pi t}{\ell}\right) Rc(\dot{\gamma}, \dot{\gamma}) \right) dt \\ &\leq (n-1) \int_0^\ell \left(\frac{\pi^2}{\ell^2} \cos^2\left(\frac{\pi t}{\ell}\right) - C \sin^2\left(\frac{\pi t}{\ell}\right) \right) dt. \end{aligned}$$

Since

$$\int_0^\ell \cos^2\left(\frac{\pi t}{\ell}\right) dt = \int_0^\ell \sin^2\left(\frac{\pi t}{\ell}\right) dt = \frac{\ell}{2},$$

this gives

$$0 \leq \frac{n-1}{2} \left(\frac{\pi^2}{\ell} - C\ell \right).$$

Hence

$$\ell \leq \frac{\pi}{\sqrt{C}}.$$

Since p, q were arbitrary, we conclude that

$$\text{diam}(M) \leq \frac{\pi}{\sqrt{C}}.$$

The rest of the argument goes as before. $\square$

Remark 45.5. Notice that if there were a conjugate point before πR , this would imply the theorem (since geodesics do not minimize past conjugate points). The proof we gave did not quite show this, but it is in fact true, and follows from the Rauch comparison theorem.

Remark 45.6. Recall that an **Einstein metric** satisfies

$$Rc = \lambda g.$$

If $n > 2$, then λ is automatically constant; indeed,

$$\lambda = \frac{S}{n}.$$

Therefore, if $\lambda > 0$, then Theorem 45.4 applies with

$$C = \frac{\lambda}{n-1},$$

and we conclude that M is compact and $\pi_1(M)$ is finite.

Remark 45.7. The Bonnet-Myers theorem is consistent with the examples that we've seen above (summarized on p. 107):

- $\mathbb{S}^n$ and $\mathbb{CP}^n$ both have positive sectional curvature, and are compact and simply connected.
- *Compact* Lie groups admit bi-invariant metrics, which have nonnegative sectional curvature. As a partial converse, if the Ricci curvature is positive, then the group must be compact.