Riemannian Geometry Notes

Simon Xiang

April 20, 2021

Notes for the Spring 2021 graduate section of Riemannian Geometry (Math 392C) at UT Austin, taught by Dr. Sadun. The course somewhat follows *Introduction to Riemannian Manifolds* (2nd edition), by Lee. Source files: https://git.simonxiang.xyz/math_notes/files.html

Contents

Lecture 1

March 11, 2021

1.1 General formulas for the Levi-Civita connection

Consider $g(Y,Z)$ which is just a function, so we can take its directional derivative $X(g(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$. (Denote $\langle X,Y\rangle = g(X,Y)$). So

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

is the equation for being metric. For a symmetric connection, we have

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

In a coordinate basis $[X,Y] = 0$, so $\Gamma^k_{ij} = \Gamma^k_{ji}$. Now just by changing the names of the vector fields we have $Y\langle X,Z\rangle = \langle \nabla_Y X, Z\rangle + \langle X, \nabla_Y Z\rangle$ and $Z\langle X,Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle$. Then

$$\begin{aligned}
+X\langle Y,Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \\
+Y\langle X,Z\rangle &= \langle \nabla_Y X, Z\rangle + \langle X, \nabla_Y Z\rangle \\
-Z\langle X,Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \\
&\Longrightarrow \\
X\langle Y,Z\rangle + Y\langle X,Z\rangle - Z\langle X,Y\rangle &= \langle \nabla_X Y + \nabla_Y X, Z\rangle + \langle \nabla_X Z - \nabla_Z X, Y\rangle + \langle \nabla_Y Z - \nabla_Z Y, X\rangle \\
&= 2\langle \nabla_X Y, Z\rangle - \langle [X,Y], Z\rangle + \langle [X,Z], Y\rangle + \langle [Y,Z], X\rangle \\
&\Longrightarrow
\end{aligned}$$

$$\boxed{\langle \nabla_X Y, Z\rangle = \frac{1}{2}\left(X\langle Y,Z\rangle + Y\langle X,Z\rangle - Z\langle X,Y\rangle + \langle [X,Y],Z\rangle - \langle [X,Z],Y\rangle - \langle [Y,Z],X\rangle\right).}$$

This equation is called **Koszul's formula**, which is the most general formula for the Levi-Civita connection. There are some bases we care about:

(1) If $X = \partial_i, Y = \partial_j, K = \partial_k$, then $\Gamma_{ijk} = \frac{1}{2}\left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right)$ since all the bracket terms are zero. This is the form we're familiar with of the Levi-Civita connection, and the most common one as well.

(2) Say we have an orthonormal basis $\{E_i\}$, where $X = E_i, Y = E_j, Z = E_k$. Then $[E_i, E_j] = c^k_{ij} E_k$, where the coefficients c^k_{ij} tell you to what extent are the brackets nonzero. If $c_{ijk} = \langle [E_i, E_j], E_k\rangle$, then $\Gamma_{ijk} = \Gamma^k_{ij} = \frac{1}{2}(c_{ijk} - c_{ikj} - c_{kji})$.

(3) Given a general frame, we have $\Gamma_{ijk} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) + \frac{1}{2}(c_{ijk} - c_{ikj} - c_{jki})$.

1.2 Return to geodesics

Recall that a geodesic satisfies $\nabla_{\dot{x}} \dot{x} = 0$, and $\ddot{x}^k + \Gamma^k_{jk}\dot{x}^i \dot{x}^j = 0$. Say we have a point p and a vector v_0, we want to talk about a geodesic starting at p with velocity v_0. We want to convert this second order differential equation of n variables into a first order ODE of two variables. Say $v(t) = \dot{x}(t)$, then what is $\dot{v}$? Recall $\nabla_v v = 0 \iff \dot{v}^k + \Gamma^k_{jk}(x) v^i v^j = 0$, so

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

These are our two desired differential equations. Then we have unique solutions since the Γ's are smooth functions of x (the particular smooth function is given by the Levi-Civita connection). In other words, there exists a unique solution $x(0) = p, v(0) = v_0$ for a short time (locally).

For an arbitrary amount of time, consider the line going through the origin in $\mathbb{R}^2 \setminus \{0\}$, which fails near the origin. Another example is the open interval $M = (a, b), g = dx^2$; after a certain amount of time you fall of the edge of the world. If somebody gives you a manifold with a starting point and velocity, we want to run a geodesic for as long as possible in the $\pm t$ direction. This motivates the following definition.

Definition 1.1. A geodesic is **maximal** if it can't be extended.

Is there a maximal geodesic? Given a starting point, take the union of the geodesics on an interval. This is unique, and we can't extend this any farther because this implies a bigger interval. So given any point and any starting vector, there is always a unique maximal geodesic.

1.3 The exponential map

Say we have a manifold M, $p \in M$, and the tangent space T_pM. Then the **exponential map** is defined as $\exp_p(v) = \gamma(1)$, where γ is a geodesic with $\gamma(0) = p, \dot{\gamma}(0) = v$. Is this defined on all of T_pM? This is defined for all sufficiently small v (for large v you may fall of the edge of the world). You can think of $\exp_p(sv) = \gamma_{sv}(1) = \gamma_v(s)$. For every p there exists a small neighborhood that looks like $\mathbb{R}^n$, so exp is well defined on the tangent space in that neighborhood. So think of the exponential map as the map

$$\exp_p : \text{Nbd of } 0 \in T_pM \to \text{Nbd of } p \in M.$$

Given a compact manifold we can extend this, so geodesics run forever and the exponential map is defined for all t. What is $d\exp_p|_{v=0}$? This is the identity, since $d\exp_p|_{v=0} : T_0(\text{Nbd of } 0 \in T_pM) \to T_pM$ which is the same as a map $T_pM \to T_pM$. At $v = 0$, this is saying "given an infinitesmally small vector, where do you wind up?" This is the same thing as saying "how fast are you moving at time zero if you have a large vector?", since how far you wind up with an infinitesmally small vector is how far you wind up with an ordinary geodesic in a short amount of time. So this asks how geodesics at small time, which is just the derivative of a geodesic at 0, which is just v!

The nice thing about the identity is that it's invertible. Suppose we have two manifolds M, N, $f : M \to N$, $p \in M$, $q = f(p) \in N$, $df : T_pM \to T_qN$. If df_p is invertible, then f is a local diffeomorphism, so $f|_V$ is a diffeomorphism $U \to V$ (where U, V are neighborhoods of p, q). So $\exp_p$ is a diffeomorphism, since it takes neighborhoods to neighborhoods. Let $r = \sup\{\text{radii } \mid \exp_p \text{ is a diffeomorphism on } B_p\}$. We say r is the **injectivity radius** at p.

Example 1.1. Consider M the torus, if we draw it as a rectangle, say it has width L_1 and height L_2, $L_2 > L_1$. Then the injectivity radius at p is $L_2/2$, since any points past $L_2/2$ wrap around. The exponential map is a local diffeomorphism, but it fails to be injective; this is why it's called the *injectivity* radius. On a sphere with p the north pole, our injectivity radius is π (since the circumference is 2π).

You might think the injectivity radius is about the topology, since $H_1(\mathbb{T}) = \mathbb{Z} \oplus \mathbb{Z}$ is nontrivial and we have a cycle to wrap around. It turns out the injectivity radius isn't just about the topology, since S^2 has no interesting first homology group. A handwavy way to think about the injectivity radius is the biggest radius such that a neighborhood of size r around p looks like a ball.

If we have orthonormal coordinates for T_pM, this induces coordinates on T_qN by the exponential map. Then $g_{ij}(p) = \delta_{ij}$, or even stronger, we have $g_{ij}(\exp_p(v)) = \delta_{ij} + O(v^2)$. Consider a ball of radius $\varepsilon < r$ in T_pM, for $p, q \in U_p$ (for U_p a neighborhood around p in M 0. Is there a geodesic connecting p and q? Sure there is, since exp is a diffeomorphism. The geodesic is also locally unique. If $q = \exp_p(v)$, then what is the distance from p to q? It's the inf of all lengths; not all geodesics globally minimize length, but a length minizming curve is certainly a geodesic. This is equivalent to the energy concept, that the variational equations for energy give us geodesics WRT the Levi-Civita connection. Since the shortest path gives us a geodesic, and there exists exactly one geodesic, then the geodesic must be the shortest path. So the distance from p to q is the magnitude of v.

In other words, the distance function from p is just the magnitude function in T_pM. If we use polar coordinates on T_pM, the metric in the radial direction $g_{rr} = 1$, so the metric will always be dr^2+something.

1.4 Tubular neighborhoods

Suppose we have a Riemannian manifold M and a submanifold N. The tubular neighborhood theorem states that if N is compact, the set of all points within ε of N is a tubular neighborhood, and is diffeomorphic to a neighborhood of the zero section of the normal bundle. If N is not compact, then ε is not globally chosen (it varies from point to point). What does a neighborhood of $p \in N$ look like? A neighborhood of p in the big space looks like a neighborhood of the origin in T_pM. But $T_pM = T_pN \oplus N_pN$, so we parametrize by $(p, v) \to \exp_p(v)$.

Lecture 2

March 23, 2021

2.1 Calculus of variations

Suppose we have $z(t)$, where $z(0) = 0, z(2) = 0$. We want to minimize the integral $\mathscr{L} = \int_0^2 \frac{1}{2}\dot{z}^2 - 32z\,dt$. We will do this two ways: one is the "sloppy" version with our usual notation for the calculus of variations, and one by a mathematically precise method. The usual notation has a precise meaning behind it, but it just seems sloppy. If $0 = \frac{\delta \mathscr{L}}{\delta z}$, then

$$\begin{aligned}
\delta \mathscr{L} &= \int_0^2 \left(\frac{1}{2} z\dot{z}\delta\dot{z} - 32\delta z\right) dt \\
&= \int_0^2 \left(\dot{z}(\dot{\delta z}) - 32\delta z\right) dt \\
&= \cancel{\dot{z}\delta z\Big|_0^2} + \int_0^2 (-\ddot{z} - 32)\delta z\,dt \\
&= -\int_0^2 (\ddot{z} + 32)\delta z\,dt
\end{aligned}$$

since $\delta(\dot{z}) = (\dot{\delta z})$, and our boundary conditions. Since $\delta \mathscr{L} = 0$, we conclude that $\ddot{z} = -32$. Given $\ddot{z} = -32$, $z(0) = 0$, $z(2) = 0$, we conclude that $z = 32t - 16t^2$. The expression $\frac{1}{2}\dot{z}^2 - 32z$ is called the **Lagrangian**, $\mathscr{L}$ corresponds to the *action*, and this is another way to do Hamiltonian mechanics. In general, if we want to minimze $\mathscr{L} = \int_0^T \frac{1}{2}m\dot{x}^2 - V(x)\,dx$ where $V(x)$ is some *potential function*, we get that

$$\begin{aligned}
\delta \mathscr{L} &= \int m\langle \dot{x}, (\dot{\delta x})\rangle - \langle \nabla V, \delta x\rangle\,dt \\
&= \int \langle -m\ddot{x} - \nabla v, \delta x\rangle\,dt \implies \\
&\boxed{m\ddot{x} = -\nabla V(x).}
\end{aligned}$$

This expression is *Newton's law*, where $-\nabla V(x) = F$ (since the force of the gradient is the potential energy), and $m\ddot{x} = ma$. People trained in math may find this unsatisfying, particularly in the notation $\delta\mathscr{L}$ and other various grievances. So let's cover this again but more precisely. Once again, our goal is to minimize $\mathscr{L} = \int_0^2 \frac{1}{2}\dot{z}^2 - 32z\,dt$. Say we have a family of functions $z_s(t)$ for $s \in (-\varepsilon, \varepsilon)$, which in reality is a smooth function of two variables $Z(s,t)$ satisfying $z_s(0) = z_s(2) = 0$. For each s, compute

$$\mathscr{L}(s) = \int_0^2 (\frac{1}{2}(\dot{z}_s(t))^2 - 32z_s(t))\,dt.$$

We want to minimize this, so consider

$$\left.\frac{\partial \mathscr{L}}{\partial s}\right|_{=0} = \int_0^2 \frac{1}{2}\,dt \tag{1}$$

$$= \int_0^2 (\dot{z}(\dot{\partial_s z}) - 32\partial_s z)\,dt \tag{2}$$

$$= \cancel{\dot{z}(\partial_s z)\Big|_0^2} + \int_0^2 (-\ddot{z}(\partial_s z) - 32(\partial_s z))\,dt \tag{3}$$

This implies $\ddot{z} = -32$. Note that

$$\frac{\partial}{\partial s}\dot{z}_s(t) = \frac{\partial}{\partial s}\frac{\partial}{\partial t}Z(s,t) = \frac{\partial^2 Z(s,t)}{\partial t \partial s} = (\partial_s \dot{z}(t)),$$

which is how we get from (1) to (2) above. This finishes the problem. These two calculations are essentially the same, but in the first one our terms weren't properly defined; we do things like taking a function z, changing it a little by δz, etc. Now we have a family of functions $Z(s,t)$, and δz means $\partial_s Z(s,t)$. Conversely, given a particular δz in mind, we can also just find a family of functions $z_s(t) = z_0(t) + s\delta z(t)$.

2.2 Return to geodesics (once more)

In the last homework, we had points p,q on a manifold in a single coordinate chart. Given a path γ such that $\gamma(0) = p, \gamma(T) = q$, $\gamma(t) = (x^1(t), x^2(t), \cdots, x^n(t))$, consider $E(\gamma) = \int \langle \dot\gamma, \dot\gamma\rangle\,dt$. Then

$$\begin{aligned}
E &= \int_0^T (g_{ij}\dot{x}^i\dot{x}^j)\,dt, \\
\delta E &= \int_0^T ((\delta g_{ij})\dot{x}^i\dot{x}^j + g_{ij}(\delta\dot{x}^i)\dot{x}^j + g_{ij}\dot{x}^i(\delta\dot{x}^j))\,dt \\
&= \int_0^T (\partial_k g_{ij}\delta x^k\dot{x}^i\dot{x}^j + 2g_{ij}(\delta\dot{x}^i)\dot{x}^j)\,dt \\
&= \int_0^T ((\partial_k g_{ij})\dot{x}^i\dot{x}^j\delta x^k - 2\frac{d}{dt}(g_{ij}\dot{x}^j)\delta x^i)\,dt \\
&= \int_0^T \partial_k g_{ij}\dot{x}^i\dot{x}^j\delta x^k - 2\ddot{x}^j g_{ij} - 2\dot{x}^j\partial_k g_{ij}\dot{x}^k\delta x^i\,dt \cdots
\end{aligned}$$

$$\downarrow \text{ (???)}$$

$$\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0, \text{ where } \Gamma^k_{ij} = \frac{1}{2}g^{k\ell}(\partial_i g_{jk} + +\partial_j g_{ik} - \partial_k g_{ij}).$$

It seems like magic that the way to minimize energy along a path is to follow a geodesic according to the Levi-Civita connection. Recall that

- If you minimize energy, then you have a geodesic.
- If you minimize energy, you minimize length and go at constant speed, where $E_{\min} = L^2_{\min}/T$.
- So if you minimize energy, then you have a constant-speed length-minimizing curve.

Theorem 2.1. *Geodesics locally minimize length, that is, the shortest path to connect two points in a neighborhood of a point is by following a geodesic.*

Length minimizing curves are geodesics, but this doesn't mean geodesics minimize length (globally). However, if there's only one geodesic, it must minimize length. So the strategy is to show that given a point p, you can get to any other point q via a unique geodesic. This is an existence-uniqueness argument.

Proof. Recall that $\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$, and after converting this second order ODE into two first order ODEs $\dot{x}^k = v^k$, $\dot{v}^k + \Gamma^k_{ij}(x)v^iv^j = 0$, we know unique solutions exist (for a short time) by the standard theory of differential equations. A *maximal geodesic* extends this time for as long as possible. Let $\gamma_v(t)$ be the geodesic at time t with initial velocity v, and note that $\gamma_{tv}(1) = \gamma_v(t)$. Our statement about existence-uniqueness implies that $\gamma_v(1)$ exists for all sufficiently small v. Define the exponential map $\exp_p(v) = \gamma_v(1)$, where $\gamma(0) = p, \dot{\gamma}(0) = v$. Since $\exp_p \colon T_pM \to M$,

Then $v_2(g) = \alpha g, \gamma(t) = e^{\alpha t}, v_a(g) = g\alpha, \dot{\gamma}(t) = \alpha e^{\alpha t} = \alpha t, \dot{\gamma}(t) = e^{\alpha t}\alpha = \gamma\alpha$. So $\gamma(1) = e^{\alpha}$. Suppose we have a Lie group $M = G$, then $T_pM = \mathfrak{g}$, a **Lie algebra**. If $G = \mathrm{SO}(n)$, then $\mathfrak{g} = \mathrm{so}(n)$ todo:? ⊠

Lecture 3

March 25, 2021

Picking off from yesterday, say $\exp_p \colon T_pM \to M$. Then $(x^1, \cdots, x^n) \mapsto \exp_p(x^1\mathbf{b}_1 + \cdots + x^n\mathbf{b}_n)$, where $\{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ are orthonormal coordinates for T_pM. Define $r = \sqrt{(x^1)^2 + \cdots + (x^n)^2}$, $\theta = \mathbf{x}/r$. We want to figure out the values of the metric $g_{rr}, g_{r\theta}, g_{\theta\theta}$.

Gauss Lemma. *We have $g_{rr} = 1$ and $g_{r\theta} = 0$.*

todo:

3.1 The Riemann curvature tensor

We finally arrive at a big topic in Riemannian geometry. First we give a handwavy explanation. Say we have a manifold M, and we want to do parallel transport from p to q along a path. What if we go along a different path? Say γ is a closed loop with $\gamma(0) = p$ and $\gamma(t) = p$, then parallel transport is a map $P_\gamma \colon T_pM \to T_pM$.If $\widetilde{v}, \widetilde{w}$ are parallel tranports of v, w, then

$$\frac{d}{dt}\langle\widetilde{v}, \widetilde{w}\rangle = \langle D_t\widetilde{v}, \widetilde{w}\rangle + \langle\widetilde{v}, D_t\widetilde{w}\rangle = 0.$$

Since parallel transport preserves inner products, this map P_γ around a loop is an isometry. Given two “transverse” vector fields X, Y, take a point p and flow in the direction of X for a little while, then Y, then $-X$, then $-Y$. The distance this flow will be off is $st[X, Y]$ as we have shown.

Definition 3.1. We define a $(1, 3)$-tensor field that takes three inputs $R(X, Y)Z = \nabla_x\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$. This tensor is called the **Riemann curvature tensor**.

This looks like a differential operator. A tensor satisfies $R(X, Y)(fZ) = fR(X, Y)Z$. If this were a differential operator, it would pick up extra terms involving f. Let’s check to see R satisfies this.

$$\begin{aligned}
\nabla_Y(fZ) &= Y(f)Z + f\nabla_Y Z,\\
\nabla_X(\nabla_Y(fZ)) &= X(Y(f))Z + Y(f)\nabla_X Z + X(f)\nabla_Y Z + f\nabla_X\nabla_Y Z,\\
\nabla_Y\nabla_X(fZ) &= Y(X(f))Z + X(f)\nabla_Y Z + Y(f)\nabla_X Z + f\nabla_Y\nabla_X Z,\\
(\nabla_X\nabla_Y - \nabla_Y\nabla_X)(fZ) &= [X, Y]f + f(\nabla_X\nabla_Y - \nabla_y\nabla_X)Z,\\
\nabla_{[X,Y]}(fZ) &= ([X, Y]f)Z + f\nabla_{[X,Y]}Z,\\
R(X, Y)(fZ) &= fR(X, Y)Z.
\end{aligned}$$

So we’re left with something linear in Z, and this is a tensor.

Example 3.1. Let's work out an example. We have

$$\begin{aligned}R(X,fY)Z &= \nabla_X\nabla_{fY}(Z)-\nabla_{fY}\nabla_X Z-\nabla_{[X,fY]}Z\\ &= \nabla_X f(\nabla_Y Z)-f\nabla_Y\nabla_X Z-\nabla_{f[X,Y]+X(f)Y}Z\\ &= X(f)\nabla_Y Z+f\nabla_X\nabla_Y Z-f\nabla_Y\nabla_X Z-f\nabla_{[X,Y]}Z-X(f)\nabla_Y Z\\ &= f(\nabla_X\nabla_Y Z-\nabla_Y\nabla_X Z-\nabla_{[X,Y]}Z)\\ &= fR(X,Y)Z.\end{aligned}$$

This shows it works for Y, which implies it works for X (doesn't matter which order we defined it in), so the Riemann curvature tensor is indeed a tensor.

For tensors, we usually break them down in coordinates, like saying

$$R(X,Y)Z = R^\ell_{ijk}x^i y^j z^k e_\ell, \text{ where } R(e_i,e_j)e_k = R^\ell_{ijk}e_\ell.$$

We might also want to consider $\langle R(X,Y)Z,W\rangle = R_{ijkl\ell}x^i y^j z^k w^\ell$,or

$$R_{ijk\ell} = \langle R(e_i,e_j)e_k, e_\ell\rangle.$$

We claim this has something to do with parallel transport along small loops. To see this, suppose we have a function $f(x)$, and we want to move a point $p\mapsto p+a$. We know that

$$\begin{aligned}f(x+a) &= \sum_{n=0}^{\infty}\frac{f^{(n)}(x)}{n!}a^n\\ &= e^{a\frac{d}{dx}}f.\end{aligned}$$

For functions, flowing around the box means we're interested in $e^{-st[X,Y]}e^{sY}e^{tX}e^{-sY}e^{-tX}$ as flows of vector fields, which turns out to be the identity since mixed partials commute. If you want to push vector fields around, we need to consider $e^{-st\nabla_{[X,Y]}}e^{s\nabla_Y}e^{t\nabla_X}e^{-s\nabla_Y}e^{-t\nabla_X}$, and the extent to which this fails to be the identity is measured by the Riemann curvature tensor. The reason why this isn't the identity outright is because the covariant derivatives ∇_X,∇_Y don't commute. Calculating the previous expression by performing a Taylor expansion results in $1-stR(X,Y)$ plus higher order terms.

Lecture 4

March 30, 2021

4.1 More on the Riemann curvature tensor

Recall that $R(X,Y)Z = X^iY^jZ^kR(e_i,e_j)e_k$, where $R(e_i,e_j)e_k = R^\ell_{ijk}e_\ell$. We also write $\langle R(e_i,e_j)e_k,e_\ell\rangle = g(R(e_i,e_j)e_k,e_\ell) = R_{ijk\ell} = g_{\ell m}R^m_{ijk}$. Why do we care? We say

(1) A manifold is **flat** if it is locally isometric to Euclidian $\mathbb{R}^n$ (that is, $g_{ij}=\delta_{ij}$).

(2) A manifold meets the **flatness criterion**[1] if it has zero curvature, or $R=0$.

Theorem 4.1. *A manifold M is flat if and only if it meets the flatness criterion.*

Example 4.1. Some examples of flat manifolds include the torus $\mathbb{T}^n=\mathbb{R}^n/\mathbb{Z}^n$ and the Möbius strip. Note that curvature in one dimension is vacuous, since the tensor is antisymmetric, and we only have one choice of basis vector so everything cancels out. This fails for a sphere since even though you can write it as a topological quotient, you can't write the metric as a quotient metric.

[1]Nobody actually says this, we just say *flat*. However, we need a separate term to distinguish this from flatness before we prove the following theorem.

Example 4.2. Let's compute the curvature of $\mathbb{R}^n$ and show that $R = 0$, implying that anything with nonzero curvature cannot be $\mathbb{R}^n$. Recall that $g_{ij} = \delta_{ij}$, and $\Gamma^k_{ij} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) = 0$. Raising the index Γ^K_{ij} also gives zero, telling us that $\nabla_{e_i} e_j = \Gamma^k_{ij} e_k = 0$. So

$$\begin{aligned} R(e_i, e_j)e_k &= \nabla_i \nabla_j (e_k) - \nabla_j \nabla_i (e_k) - \nabla_{[e_i, e_j]} e_k \\ &= \nabla_i(0) - \nabla_j(0) - 0 \\ &= 0. \end{aligned}$$

Example 4.3. Now let's compute the curvature of a sphere. Let's use our standard parametrization $(x^1, x^2) = (\varphi, \theta)$ with the metric $g_{11} = 1, g_{12} = 0$, and $g_{22} = \sin^2 \varphi$. We know $R_{11jk} = 0$ since $R(e_1, e_1)e_j = \nabla_1 \nabla_1 e_j - \nabla_1 \nabla_1 e_j - \nabla_{[e_1, e_1]} e_j = 0$. Similarly $R_{22jk} = 0$, and $R_{21jk} = -R_{12jk}$. So the only interesting this is R_{12jk}, where $j, k = 1, 2$; that means we only compute $R(e_1, e_2)e_1$. Along the way we'll need to compute Γ_{ijk} and Γ^k_{ij}; since $\partial_1 g_{22} = 2 \sin \varphi \cos \varphi$, we have

$$\begin{aligned} \Gamma_{122} = \Gamma_{212} &= \sin \varphi \cos \varphi, \\ \Gamma_{221} &= -\sin \varphi \cos \varphi, \\ \Gamma^2_{12} = \Gamma_{21} &= \cos \varphi. \end{aligned}$$

So $\nabla_1 e_2 = \nabla_2 e_1 = \cos(\varphi) e_2$, and $\nabla_2 e_2 = -\sin(\varphi)\cos(\varphi) e_1$. So

$$\begin{aligned} R(e_1, e_2)e_1 &= \nabla_1 \nabla_2 e_1 - \cancel{\nabla_2 \nabla_1 e_1} = \nabla_1 \cos(\varphi) e_2 \\ &= -\csc^2(\varphi) e_2 + \cot(\varphi) \nabla_1 e_2 \\ &= -\csc^2(\varphi) e_2 + \cot^2(\varphi) e_2 \\ &= -e_2. \end{aligned}$$

This implies $R^2_{121} = -1$, and $R^1_{121} = 0$. The negative result may seem strange, but soon we'll talk about how to get *scalars* out of the Riemann tensor, which results in a plus sign for S^2. In terms of other components, $R_{1212} = -\sin^2(\varphi), R_{1221} = \sin^2(\varphi), R^1_{122} = \sin^2(\varphi)$, because $R^k_{21j} = -R_{12} j^k$. These are all the nonzero elements of the curvature tensor. So the sphere is not flat, and cannot possibly be isometric to Euclidian space.

Proof of Theorem 4.1. Now that we've sufficiently motivated Theorem 4.1, let's prove it. Suppose $R = 0$. Pick an orthonormal basis $e_1, \cdots, e_n$ at $p = 0$. We parallel transport along the e_1 direction, resulting in a frame at every point along e_1. Then we parallel transport along e_2, then e_3, and so on. (For now we work in three dimensions.) We need to show that

(1) The frame is covariantly constant, or $\nabla_i e_j = 0$,
(2) The frame is a coordinate frame, or $[e_i, e_j] = 0$.[2]
(3) $g_{ij} = \delta_{ij}$.

The interesting argument is (1), where we really use the lack of curvature. We do this inductively; we show on the x^1-axis, $\nabla_1 e_1 = 0$. Then we show on the x^1x^2-plane, $\nabla_1 e_1 = \nabla_1 e_2 = \nabla_2 e_1 = \nabla_2 e_2 = 0$. Finally we show that on xyz spaces, $\nabla_1 e_j = 0$. The statement $\nabla_1 e_1$ is true because we're parallel transporting. The fact that $\nabla_2 e_1 = \nabla_2 e_2 = 0$ are obvious by the way we parallel transported. Compute

$$\begin{aligned} \nabla_2(\nabla_1 e_j) &= \cancel{R(e_2, e_1)e_j} + \nabla_1(\nabla_2 e_j) \\ &= \nabla_1(\nabla_2 e_j) = \nabla_1(0) \\ &= 0. \end{aligned}$$

Since this is 0 when time $s = 0$ (along the x^1-axis) and the derivative is zero (no change), this must be zero everywhere. This shows $\nabla_1 e_1 = \nabla_1 e_2 = 0$. To show $\nabla_i e_j = 0$, start with our plane and move things in the third direction. We have already shown in the plane that $\nabla_1 e_1 = \nabla_1 e_2 = \nabla_2 e_1 = \nabla_2 e_2 = 0$, and we know that $\nabla_1 e_3 = \nabla_2 e_3 = 0$ since our argument works for *any* j. Take the derivative with respect to the third coordinate, then everything is zero in $\mathbb{R}^3$, and so on. ⊠

[2] The Frobenius theorem says that if you have a bunch of vector fields that commute, they're really derivatives with respect to some coordinates. But we never proved this.

4.2 Symmetries

We have

(1) $R_{ijk\ell} = -R_{jik\ell}$,

(2) $R_{ij\ell k} = -R_{ijk\ell}$,

(3) $R_{ijk\ell} + R_{jki\ell} + R_{kij\ell} = 0$,

(4) $R_{ijk\ell} = R_{k\ell ij}$.

(1) follows from the definition, (2) can be shown by taking derivations of $g(e_k, e_\ell)$. (3) is called the *algebraic Bianchi identity* or *1st Bianchi identity*, and to show this we need to use facts about symmetry like $\nabla_j(e_k) = \nabla_k(e_j)$, and (4) follows from the rest.

4.3 The Ricci tensor and scalar curvature

What are some nice invariants of a linear transformation M_i^j? A simple one is the *trace*, where $\operatorname{Tr} M = M_i^i$; recall that the trace is invariant under change of basis. We define a *partial* trace of the Riemann tensor R_{ij},[3] which we call the **Ricci tensor**. We can write the Ricci tensor as

$$R_{ij} = R^k_{kij} = -R^k_{ikj} = -R^k_{kij}.$$

Furthermore, $R_{ij} = R_{kijm} g^{km}$. Finally, we have **scalar curvature** given by

$$R = g^{ij} R_{ij}.$$

Example 4.4. Let us return to the sphere S^2. What is the Ricci tensor and scalar curvature of the sphere? On the sphere,

$$\begin{aligned} R_{11} &= R^1_{111} + R^2_{211} = 1 = g_{11}, \\ R_{22} &= R^1_{112} + R^2_{222} = \sin^2\varphi = g_{22}, \\ R_{12} &= R^1_{112} + R^2_{212} = 0 = g_{12}. \end{aligned}$$

Note that the Ricci tensor is identically the metric. Spaces whose Ricci tensor is proportional to the metric have a special name, called an **Einstein metric**. For the scalar curvature,

$$\begin{aligned} R &= \frac{g^{11}R_{11} + g^{22}R_{21}}{n} \\ &= \frac{1 \cdot 1 + \sin^{-2}\varphi \sin^2\varphi}{n} \\ &= \frac{2}{2} = 1. \end{aligned}$$

Example 4.5. Let us compute things in hyperbolic space. Recall that $g_{11} = g_{22} = \frac{1}{y^2}$, $g_{12} = 0$, $\partial_2 g_{11} = \partial_2 g_{22} = -\frac{2}{y^3}$. We also have

$$\begin{aligned} \Gamma_{112} &= \frac{1}{y^3}, \\ \Gamma_{121} &= \frac{1}{y^3}, \\ \Gamma_{122} &= \frac{1}{2}(\partial_1 g_{22} + \partial_2 g_{12} - \partial_2 g_{12}) = 0, \\ \Gamma_{211} &= \frac{1}{y^3}, \\ \Gamma_{222} &= -\frac{1}{y^3}, \end{aligned}$$

[3] Unfortunately they use the same notation. R_{ij} refers to the Ricci tensor, while $R_{ijk\ell}$ refers to the Riemann tensor.

$$\begin{aligned}\nabla_1 e_1 &= \frac{1}{y} e_2,\\ \nabla_2 e_2 &= -\frac{1}{y} e_2,\\ \nabla_1 e_2 &= \frac{1}{y} e_1,\\ \nabla_2 e_1 &= -\frac{1}{y} e_1.\end{aligned}$$

Now we compute $R(e_1, e_2)e_1$, since R_{1212} is the only thing we're interested in (has to be antisymmetric in both). So

$$\begin{aligned}R(e_1,e_2)e_1 &= \nabla_1\nabla_2 e_1 - \nabla_2\nabla_1 e_1\\ &= \nabla_1\left(-\frac{1}{y}e_1\right) - \nabla_2\left(\frac{1}{y}e_2\right)\\ &= -\frac{1}{y}(\nabla_1 e_1) + \frac{1}{y^2}e_2 - \frac{1}{y}\nabla_2 e_2\\ &= \frac{1}{y^2}e_2.\end{aligned}$$

So $R^2_{121} = \frac{1}{y^2}$,

$$\begin{aligned}R_{11} &= R^2_{211} + R^1_{111} = -\frac{1}{y^2} = -g_{11},\\ R_{22} &= R^1_{122} + R^2_{222} = -\frac{1}{y^2} = -g_{22},\\ R_{12} &= 0 = -g_{12}.\end{aligned}$$

In hyperbolic space, the Ricci curvature is the negative of the metric. This works for S^n and $\mathbb{H}^n$ in general.

Lecture 5

April 1, 2021