Riemannian Geometry

UT Austin, Spring 2017

M392C NOTES: RIEMANNIAN GEOMETRY

ARUN DEBRAY
MARCH 2, 2017

These notes were taken in UT Austin's M392C (Riemannian Geometry) class in Spring 2017, taught by Dan Freed. I live-TEXed them using `vim`, so there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu. Thanks to Martin Bobb, Gill Grindstaff, and Jonathan Johnson for some corrections.

The cover image is the Cosmic Horseshoe (LRG 3-757), a gravitationally lensed system of two galaxies. Einstein's theory of general relativity, written in the language of Riemannian geometry, predicts that matter bends light, so if two galaxies are in the same line of sight from the Earth, the foreground galaxy's gravity should bend the background galaxy's light into a ring, as in the picture. The discovery of this and other gravitational lenses corroborates Einstein's theories. Source: https://apod.nasa.gov/apod/ap111221.html.

Contents

Lecture 1.

Geometry in flat space: 1/17/17

"Do you have all these equations?"

Before we begin with Riemannian manifolds, it'll be useful to do a little geometry in flat space.

Definition 1.1. Let V be a real vector space; then, an *affine space over* V is a set A with a simply transitive right V-action.

That this action is simply transitive means for any $a, b \in A$, there's a unique $\xi \in V$ such that $a \cdot \xi = b$.

Definition 1.2. A set with a simply transitive (right) V-action is called a *(right) V-torsor*.

V-torsors look like copies of V without a distinguished identity.

One of the distinct features of affine space is *global parallelism*: if I have a vector ξ at a point a, I immediately get a vector at every point, which defines a vector field on the entire space.

What is the analogue of a basis for an affine space? This is a collection of points $a_0, \dots, a_n$ such that any $a \in A$ is uniquely written as

$$a = \lambda^0 a_0 + \lambda^1 a_1 + \cdots + \lambda^n a_n \tag{1.3}$$

for some $\lambda^i \in \mathbb{R}$ with $\lambda^0 + \cdots + \lambda^n = 1$.

Equation (1.3) may be written more concisely with *index notation*: any variable written as both a superscript and a subscript is implicitly summed over. That is, we may rewrite (1.3) as

$$a = \lambda^i a_i.$$

Note that in an affine space, we don't know how to add vectors (since we don't have an origin), but we can take weighted averages.

Theorem 1.4 (Giovanni Ceva, 1678)**.** *Let A be an affine plane and $a, b, c \in A$ be a triangle (i.e. three distinct, noncollinear points). Suppose $p \in \overline{bc}$, $q \in \overline{ca}$, and $r \in \overline{ca}$. Then, $\overline{ap}$, $\overline{bq}$, and $\overline{cr}$ are coincident iff*

$$[ar : rb][bp : pc][cq : ca] = 1.$$

Typically, this is thought of as a ratio of lengths, but we don't necessarily have lengths: instead, we can use barycentric coordinates. There is a unique λ such that if $r = (1-\lambda)a + \lambda b$, then $[ar : rb] = \lambda/(1-\lambda)$.

Proof. Let

$$\begin{aligned} r &:= (1-\lambda)a + \lambda b \\ p &:= (1-\mu)b + \mu c \\ q &:= (1-\nu)c + \nu a. \end{aligned}$$

Set

$$x := \alpha a + \beta b + \gamma c, \tag{1.5}$$

where $\alpha + \beta + \gamma = 1$. Since $x \in \overline{ap}$, then

$$x = \alpha a + C((1-\mu)b + \mu c). \tag{1.6}$$

Comparing (1.5) and (1.6), $\mu/(1-\mu) = \gamma/\beta$.

⊠

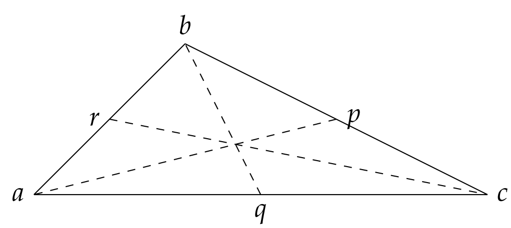

FIGURE 1. Depiction of Ceva's theorem (Theorem 1.4).

Standard affine space $\mathbb{A}^n := \{(x^1, \dots, x^n) \in \mathbb{R}^n \mid x^i \in \mathbb{R}\}$. You may complain this is the same as $\mathbb{R}^n$, but $\mathbb{A}^n$ only comes with an affine structure, not a vector-space structure.

Definition 1.7. Let A be an affine space modeled on V and B be an affine space modeled on W. Then, a map $f : A \to B$ is *affine* if there exists a linear map $T : V \to W$ such that $f(a + \xi) = f(a) + T\xi$ for all $a \in A$ and $\xi \in V$.

In other words, an affine map is a linear map plus some constant, which is not uniquely defined.

Definition 1.8. An *affine coordinate system* on A is an affine isomorphism $x = (x^1, \dots, x^n) : A \to \mathbb{A}^n$.

Them, the differentials $\mathrm{d}x_a^1, \dots, \mathrm{d}x_a^N$ are independent of basepoint a and form a basis for V^*, the dual vector space and dual basis to V and $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$, the tangent space to any $a \in A$.

But affine space is not the only flat geometry we could consider: more generally, we consider a structure on a vector space V which can be promoted to a translationally invariant structure on A. This leads to metric geometry, symplectic geometry, etc.

Definition 1.9. An *inner product* on a (finite-dimensional) vector space V is a bilinear map $\langle -,- \rangle : V \times V \to \mathbb{R}$ which is symmetric and positive definite, i.e. for all $\xi, \eta \in V$, $\langle \xi, \eta \rangle = \langle \eta, \xi \rangle$, $\langle \xi, \xi \rangle \geq 0$, and $\langle \xi, \xi \rangle = 0$ iff $\xi = 0$.

Since $\langle -,- \rangle$ is bilinear, then this can be determined in terms of n^2 numbers: let $v_1, \dots, v_n$ be a basis for V and define $g_{ij} := \langle v_i, v_j \rangle$ for $i, j = 1, \dots, n$. Of course, these numbers areen't independent: $g_{ij} = g_{ji}$, so there are really only $n(n+1)$ choices of information.

Definition 1.10. A basis $e_1, \dots, e_n$ for V is *orthonormal* if

$$\langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}.$$

Our first major result of flat Euclidean geometry is that these exist.

Theorem 1.11. *There exist orthonormal bases.*

Proof. Let $v_1, \dots, v_n$ be any basis of V. Let

$$e_1 = \frac{v_1}{\langle v_1, v_1 \rangle^{1/2}},$$

and for $i = 2, \dots, n$, let

$$v_i' = v_i - \langle v_i, e_1 \rangle e_1.$$

Then, $\langle e_1, e_1 \rangle = 1$ and $\langle e_1, v_i' \rangle = 0$. Then, repeat with $v_2', \dots, v_n'$. ⊠

This explicit algorithm is called the *Gram-Schmidt process*.

In an inner product space, we get some familiar geometric constructions: the *length* of a vector $\xi \in V$ is $|\xi| = \langle \xi, \xi \rangle^{1/2}$, and the *angle* between $\xi, \eta \in V \setminus 0$ is the θ such that

$$\cos \theta = \frac{\langle \xi, \eta \rangle}{|\xi||\eta|}.$$

Definition 1.12. A *Euclidean space* E is an affine space over an inner product space V.

This has a notion of distance: $d_E : E \times E \to \mathbb{R}^{\geq 0}$, where $a, b \mapsto |\xi|$, where $b = a + \xi$. This generalizes to notions of area, volume, etc.

Theorem 1.13 (Napoleon, 1820)**.** *Let abc be a triangle in a plane and attach an equilateral triangle to each edge. The centers of these three triangles form an equilateral triangle.*

Exercise 1.14. Prove this.

We want to understand curved analogues of this classical material, and will pick up where differential topology left off. We work on smooth manifolds: a *smooth manifold* is a space X together with an atlas of charts $U \subset X$ with homeomorphisms $x : U \to \mathbb{A}^n$ such that every point is contained in the domain of some chart and the transition maps are smooth. We do not require a manifold to have a global dimension: the different connected components may have different dimensions, e.g. $S^1 \amalg S^2$.[1]

A chart map $x : U \to \mathbb{A}^n$ is a set of n continuous maps $(x^1, \dots, x^n)$. If p is in the domain of both x and y, we can consider $x \circ y^{-1} : \mathbb{A}^n \to \mathbb{A}^n$; calculus as usual tells us what it means for this transition map to be smooth.

At any $x \in X$, we have a tangent space $T_x X$ and a cotangent space $T_x X$: a chart defines a basis of the tangent space $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ and a basis of the contangent space $\mathrm{d}x_x^1, \dots, \mathrm{d}x_x^n$. This depends strongly on x: unlike for flat space, we may not be able to parallel-transport these globally, even on something as simple as S^2.

In this course, we will study what happens when we go from a curved analogue of affine space to a curved analogue of Euclidean space, whence the following central definition.

[1]This is important for, e.g. a space of solutions of certain PDEs.

Definition 1.15. A *Riemannian metric* on a smooth manifold X is a choice of inner product $\langle -,- \rangle_x$ on T_xX for all $x \in X$ which varies smoothly in x.

Now, we can compute lengths of tangent vectors and the angle that two smooth curves intersect at (or rather, the angle their tangent vectors intersct at). We also obtain a notion of distance between points, and can develop analogues of Euclidean geometry on manifolds.

What does "varying smoothly" mean, exactly? Suppose $x^1, \ldots, x^n$ is a set of local coordinates on $U \subset X$; then, for $i, j = 1, \ldots, n$, define

$$g_{ij} := \left\langle \frac{\partial}{\partial x^i}\bigg|_x, \frac{\partial}{\partial x^j}\bigg|_x \right\rangle_{T_xX}.$$

One can check that if the metric is smoothly varying in one chart, then it's smoothly varying in all charts.

We'll write the metric as

$$g = g_{ij}\,\mathrm{d}x^i \otimes \mathrm{d}x^j.$$

This again uses the summation convention, and it's useful to think about where exactly this lives: it identifies the metric as a tensor.

Many manifolds arise as embedded submanifolds of Euclidean space, and the Whitney embedding theorem shows that all may be emnbedded. Many authors say it's best to meet manifolds as embedded submanifolds first, but there are some which arise without a natural embedding, e.g. the Grassmanian $\mathrm{Gr}_2(\mathbb{R}^4)$, the space of two-dimensional subspaces of $\mathbb{R}^4$.

In any case, if $X \subset \mathbb{E}^N$ is embedded, then X inherits a metric, since $T_xX \subset \mathbb{R}^n$ is also a subspace, and we can restrict the inner product. Classical Riemannian geometry is the study of *plane curves* (one-dimensional submanifolds of $\mathbb{R}^2$), *space curves* (one-dimensional submanifolds of $\mathbb{R}^3$), and *surfaces* (two-dimensional submanifolds of $\mathbb{R}^3$).

To study Riemannian manifolds, we should begin with the simplest cases. The zero-dimensional manifolds are disjoint unions of points with zero-dimensional tangent spaces and the trivial Riemannian metric. In the one-dimensional case, there is a little more to tell. A smooth map $X \to Y$ of Riemannian manifolds is an *isometry* if it's a map that preserves the inner product on each tangent space. This automatically implies it's injective.

Theorem 1.16. *Let C be a (complete) Riemannian 1-manifold which is diffeomorphic to $\mathbb{R}$. Then, C is isometric to $\mathbb{E}^1$.*

Before we prove this, we need a change-of-coordinates lemma. (We'll address completeness later, to avoid finite intervals.)

Remark. Let $x^1, \ldots, x^n$ and $y^1, \ldots, y^n$ be coordinate systems and suppose a metric can be written as

$$g = g_{ij}\,\mathrm{d}x^i \otimes \mathrm{d}x^j = h_{ab}\,\mathrm{d}y^a \otimes \mathrm{d}y^b.$$

Then,

$$g_{ij} = h_{ab}\frac{\partial y^a}{\partial x^i}\frac{\partial y^b}{\partial x^j}. \tag{1.17}$$

This is n^2 equations: there is no implicit summation here. ◀

Proof of Theorem 1.16. Let $x : C \to \mathbb{R}$ be a diffeomorphism, which defines a global coordinate on C. Let $g(x) = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial x} \rangle$. We seek a new coordinate $y : C \to \mathbb{R}$ such that $h(y) = \langle \frac{\partial}{\partial y}, \frac{\partial}{\partial y} \rangle = 1$ everywhere. By (1.17),

$$g = \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2, \tag{1.18}$$

so fix an $x_0 \in C$ and define

$$y(x) = \int_{x_0}^{x} \sqrt{g(t)}\,\mathrm{d}t.$$

This y satisfies (1.18) and therefore is an isometry. ⊠

The analogue to Theorem 1.16 in n dimensions (where $n > 1$) is as follows: if $x^1, \dots, x^n$ is a local coordinate system and g_{ij} is the Riemannian metric in these coordinates, is there a local change of coordinates $y^a(x^1, \dots, x^n)$ such that $h_{ab} = \delta_{ab}$? This is the analogue in Riemannian geometry to finding orthonormal coordinates, guaranteed by Theorem 1.11.

This requires solving an analogue to (1.17), but this time it's a PDE

$$g_{ij} = \sum_a \frac{\partial y^a}{\partial x^i} \frac{\partial y^a}{\partial x^j}.$$

This time, we need to ask whether there are solutions. The only thing we know how to do is differentiate:

$$\frac{\partial g_{ij}}{\partial x^k} = \sum_a \frac{\partial^2 y^a}{\partial x^k \partial x^i} \frac{\partial g^a}{\partial x^j} + \frac{\partial y^a}{\partial x^i} \frac{\partial^2 y^a}{\partial x^k \partial x^j}. \tag{1.19a}$$

By permuting indices, we obtain

$$\frac{\partial g_{ik}}{\partial x^j} = \sum_a \frac{\partial^2 y^a}{\partial x^j \partial x^i} \frac{\partial g^a}{\partial x^k} + \frac{\partial y^a}{\partial x^i} \frac{\partial^2 y^a}{\partial x^j \partial x^k} \tag{1.19b}$$

$$\frac{\partial g_{jk}}{\partial x^i} = \sum_a \frac{\partial^2 y^a}{\partial x^i \partial x^j} \frac{\partial g^a}{\partial x^k} + \frac{\partial y^a}{\partial x^j} \frac{\partial^2 y^a}{\partial x^i \partial x^k}. \tag{1.19c}$$

Taking (1.19a) + (1.19b) − (1.19c), we obtain

$$\frac{1}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right) = \sum_a \frac{\partial y^a}{\partial x^i} \frac{\partial^2 y^a}{\partial x^j \partial x^k}.$$

Now we multiply by $\frac{\partial y^b}{\partial x^\ell} g^{\ell i}$, concluding

$$\frac{\partial y^b}{\partial x^\ell} \underbrace{\frac{g^{\ell i}}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right)}_{\Gamma^\ell_{jk}} = \sum_a \frac{\partial y^a}{\partial x^i} \frac{\partial^2 y^a}{\partial x^j \partial x^k} g^{\ell i} \frac{\partial y^b}{\partial x^\ell}.$$

These Γ^ℓ_{jk} symbols therefore satisfy

$$\frac{\partial^2 y^b}{\partial x^j \partial y^k} = \Gamma^i_{jk} \frac{\partial y^b}{\partial x^i}.$$

If we differentiate once again (with respect to x^ℓ), we get

$$\begin{aligned}\frac{\partial^3 y^b}{\partial x^\ell \partial x^j \partial x^k} &= \frac{\partial \Gamma^i_{jk}}{\partial x^\ell} \frac{\partial y^b}{\partial x^i} + \Gamma^i_{jk} \frac{\partial^2 y^b}{\partial x^\ell x^i} \\ &= \left(\frac{\partial \Gamma^i_{jk}}{\partial x^\ell} + \Gamma^m_{jk}\Gamma^i_{m\ell}\right) \frac{\partial y^b}{\partial x^i}.\end{aligned}$$

Since mixed partials commute, then one discovers that if such an isometry exists, the *Riemannian curvature tensor*

$$R^i_{jk\ell} := \frac{\partial \Gamma^i_{j\ell}}{\partial x^\ell} - \frac{\partial \Gamma^i_{jk}}{\partial x^\ell} + \Gamma^m_{jk}\Gamma^i_{m\ell} - \Gamma^m_{j\ell}\Gamma^i_{mk} \tag{1.20}$$

must vanish. In simple cases, one can calculate that it's not always zero, so we don't always have global parallelism.

Riemann derived this in the middle of the 1800s. It's possible to see the glimmer of special relativity in them, though of course this was discovered later.

There's no text, though there is a website: `http://www.ma.utexas.edu/users/dafr/M392C/index.html`. There are problem sets, so undergraduates have to do some problem sets, and graduate students should. Feel free to talk to the professor about the problems, and especially to establish groups to work on the problem sets. Office hours are Wednesdays 2 to 3.

Lecture 2.

Existence of Riemannian metrics: 1/19/17

"There are so many of you... so quiet... I'll be more provocative until I get questions. Or I'll go faster."

Due to the large size of the class, it's being moved to RLM 6.104 starting next week. This means everyone who wants to sign up should be able to.

Some readings are up on the website, including a translation of Riemann's original work on curvature.

Last time, we defined affine space, which leads to the notion of a smooth manifold, and then introduced Euclidean space, an affine space over an inner product space. The curved version of that is a Riemannian manifold.

Recall that a Riemannian metric g on a smooth manifold X is a smoothly varying family of inner products on T_xX, and a Riemannian manifold is a smooth manifold together with a Riemannian metric. We also defined an isometry: if X and Y are Riemannian manifolds, then a diffeomorphism $f : X \to Y$ is an isometry if for all $x \in X$ and $\xi_1, \xi_2 \in T_xX$,

$$\langle f_*\xi_1, f_*\xi_2\rangle_{T_{f(x)}Y} = \langle \xi_1, \xi_2\rangle_{T_xX}.$$

Here, $f_* : T_xX \to T_{f(x)}Y$ is the linear pushforward of tangent vectors, also called the *differential*. If f is merely a smooth function, this is called an *isometric immersion* (the inverse function theorem automatically implies it's an immersion). If f is an embedding, this is called an *isometric embedding*.

Existence of Riemannian metrics. Suppose V is a real vector space and $g_0, g_1 : V \times V \to \mathbb{R}$ are inner products. Then for $t \in [0,1]$, $(1-t)g_0 + tg_1$ is also an inner product (you can check this directly).

The set of bilinear maps $V \times V \to \mathbb{R}$, denoted $\mathrm{Bil}(V \times V, \mathbb{R})$, is a real vector space, naturally isomorphic to $\mathrm{Hom}(V \otimes V, \mathbb{R})$ and to $V^* \otimes V^*$. Here, "natural" means this works for all finite-dimensional vector spaces at once, and commutes with linear maps.

Inner products are elements of this vector space, and our observation above means that if g_0 and g_1 are inner products, the line between them in $\mathrm{Bil}(V \times V, \mathbb{R})$ consists of inner products. In particular, *inner products form a convex set*. This only uses the affine structure on $\mathrm{Bil}(V \times V, \mathbb{R})$, since we can take convex combinations in an affine space.

This is used to generalize to the curved case, showing Riemannian metrics always exist.

Theorem 2.1. *Let X be a smooth manifold. Then, there is a Riemannian metric on X.*

Proof. Let $\mathfrak{U} = \{(U, x)\}$ be a cover of X by coordinate charts $x : U \to \mathbb{A}^n$, and let g_U denote the metric on U such that $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ are orthonormal. That is, take the standard metric on $\mathbb{A}^n$ making it into Euclidean space $\mathbb{E}^n$, and pull it back to U, where it becomes a metric (you can check that metrics pull back along closed immersions).

Now, the bases on two different charts in $\mathfrak{U}$ don't agree, and don't necessarily differ by orthonormal bases. Thus, we use a standard argument in differential geometry to globalize local objects living in a convex set: let $\{\rho_U\}_{U \in \mathfrak{U}}$ be a partition of unity subordinate to $\mathfrak{U}$; then,

$$g = \sum_{U \in \mathfrak{U}} \rho_U g_U$$

is a Riemannian metric. ⊠

Remark. Global existence is *not* assured for every geometric structure. For example, a *complex structure* on a real vector space V is an endomorphism $J : V \to V$ such that $J^2 = -\mathrm{id}_V$. This is akin to multiplication by i in a complex vector space, which squares to -1 and commutes with addition.

You can place this structure on affine space, and there's an immediate obstruction: $\dim_{\mathbb{R}} V$ must be even. Now we globalize: given an even-dimensional manifold, do we have such a structure? That is, can we place a smoothly varying complex structure on T_xX for all $x \in X$? This is called an *almost complex structure*, and not every even-dimensional manifold admits one.

Exercise 2.2. Show that S^4 has no almost complex structure.

There is an almost complex structure on S^6, and it's a famous open question whether there's a complex structure (i.e. complex coordinates with holomorphic transition functions). The known almost complex structure does not work.

Another local structure that doesn't automatically globalize is a mixed-signature metric (e.g. a Minkowski metric). In such a metric, the *null vectors*, those ξ for which $\langle \xi, \xi \rangle = 0$, form a cone whose interior is the *positive vectors* (for which the metric is positive). Trying to globalize this produces, more or less, a line in each tangent space $T_x X$. Passing to a double cover, one can choose an orientation, and therefore a nonzero vector field on X, and this can't be done in general. For example, a surface of genus 2 admits no metric of signature $(1, 1)$. These kinds of metrics arise in general relativity. ◀

In this class, we care about Riemannian metrics, which do globalize.

Let $x^1, \dots, x^n$ be local coordinates; then, we defined some local quantities in the metric in terms of these coordinates. Namely,

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle,$$

so that $g = g_{ij}\,\mathrm{d}x^i \otimes \mathrm{d}x^j = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j$. We then used this to define symbols Γ^i_{jk} and the Riemann curvature tensor $R^i_{jk\ell}$. We proved Theorem 1.16; here's a better version.

Theorem 2.3. *Let C be a Riemannian 1-manifold diffeomorphic to $\mathbb{R}$. Then, there exists an isometry $C \to I$, where $I \subset \mathbb{E}^1$ is an open interval.*

The argument we gave defining the Riemann curvature tensor generalizes this.

Theorem 2.4. *Suppose (U, g) is a Riemannian manifold and $x : U \to \mathbb{A}^n$ is a global coordinate such that*

$$g = \sum_{i=1}^{n} (\mathrm{d}x^i)^2.$$

Then, $R^i_{jk\ell} = 0$ on U.

One important thing to check here is that

$$R = R^i_{jk\ell} \frac{\partial}{\partial x^i} \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k \otimes \mathrm{d}x^\ell$$

is independent of the coordinate system (which is not clear from its definition). This means that the Riemann curvature tensor is a tensor, i.e. $R \in T_x X \otimes T^*_x X \otimes T^*_x X \otimes T^*_x X$. In the next few weeks, we will add some geometry to this discussion.

Example 2.5. Let $X = \mathbb{E}^2$ be Euclidean space with the standard metric g. Then, we have global coordinates $(x, y) : \mathbb{E}^2 \to \mathbb{A}^2$, so $g = \mathrm{d}x^2 + \mathrm{d}y^2$.

We can also introduce *polar coordinates*, another coordinate system which isn't global. This is a coordinate map $(r, \theta) : \mathbb{E}^2 \setminus \{(x, 0) : x \le 0\} \to \mathbb{A}^2$ (so $r > 0$, $-\pi < \theta < \pi$). In this case, the metric has the form

$$g = \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2.$$

This means that the vector field $\frac{\partial}{\partial r}$ has constant length 1, but the vector field $\frac{\partial}{\partial \theta}$ has length r at (r, θ). ◀

Symmetry. We've now seen vector spaces, affine spaces, Euclidean spaces, and Riemannian manifolds. As in any mathematical context, it's important to ask what the proper notion of symmetry is for these objects.

If V is a vector space, its *general linear group* is $\mathrm{GL}(V) = \mathrm{Aut}(V) := \{T : V \to V \text{ invertible}\}$. The standard example is $\mathrm{GL}_n(\mathbb{R}) := \mathrm{GL}(\mathbb{R}^n)$, the group of invertible $n \times n$ matrices, acting on the column vectors of $\mathbb{R}^n$ by scalar multiplication. For example $\mathrm{GL}_1(\mathbb{R}) = \mathbb{R}^\times$, the group of nonzero numbers under multiplication.

What about affine space? Affine space on V is a V-torsor, as V acts by translation. The symmetry group is the group of *affine transformations*

$$\mathrm{Aff}(A) := \{\alpha : A \to A \mid \alpha \text{ is invertible and affine}\}.$$

Recall that an affine map is one that preserves the affine structure: the image of a finite weighed average is the weighted average of the images. The derivative of an affine map is a linear map, so if A is an affine space modeled by V, the derivative defines a group homomorphism $\mathrm{d} : \mathrm{Aff}(A) \to \mathrm{GL}(V)$, whose kernel is the translations, a group isomorphic to V. Thus, we have a *group extension* (short exact sequence of groups)

$$1 \longrightarrow V \longrightarrow \mathrm{Aff}(A) \xrightarrow{\mathrm{d}} \mathrm{GL}(V) \longrightarrow 1. \tag{2.6}$$

The key is that in affine space, there's no canonical origin. However, (2.6) splits, if noncanonically: choose an $a \in A$. Then, any $b \in A$ can be uniquely written as $a + \xi$ for some $\xi \in V$, so for any linear transformation T, $a + \xi \mapsto a + T\xi$ is an affine transformation of A.

(2.6) is a sequence of manifolds with smooth group homomorphisms, making it a short exact sequence of *Lie groups*; we'll discuss Lie groups more later.

If V is an n-dimensional vector space, its bases are the set $\mathscr{B}(V) = \{b : \mathbb{R}^n \xrightarrow{\cong} V\}$. If $V = \mathbb{R}^n$, this is $\mathrm{GL}_n(\mathbb{R})$. In general, this makes $\mathscr{B}(V)$ into a right $\mathrm{GL}_n(\mathbb{R})$-torsor, defined by the simply transitive action $\mathscr{B}(V) \times \mathrm{GL}_n(\mathbb{R}) \to \mathscr{B}(V)$ sending $\beta, g \mapsto \beta \circ g$. (There is a corresponding left action by $\mathrm{GL}(V)$). The action on the right is akin to numbering elements of the basis, and the action on the left is more geometric; this is an instance of a general idea that internal actions tend to be from the right, and geometric ones from the left.

What's the analogue for an affine space A modeled on V? Let $\mathscr{B}(A)$ denote the collection of pairs (a, β) where $a \in A$ and $\beta \in \mathscr{B}(V)$, identified with the set of affine isomorphisms $\alpha : A \xrightarrow{\cong} \mathbb{A}^n$. These are the bases at specific points of A. There is a forgetful map $\pi : \mathscr{B}(A) \to A$ sending $(a, \beta) \to a$, and the fiber is $\mathscr{B}(V)$, the bases at a. In a similar way, there is a left action of $\mathrm{Aff}(A)$ on $\mathscr{B}(A)$, and a right action of $\mathrm{Aff}_n \coloneqq \mathrm{Aff}(\mathbb{A}^n)$ on $\mathscr{B}(A)$.

We'll use these torsors of bases a lot in this class. In this way, we're enacting Felix Klein's *Erlangen* program, where the kind of geometry we do is reflected by the symmetry group we place on the geometric structures.

Let's see what happens to these ideas in the Euclidean and Riemannian cases. If V is an inner product space, its *orthogonal group* $\mathrm{O}(V) \subset \mathrm{GL}(V)$ is the group of linear isomorphisms preserving the inner product, i.e. $T : V \to V$ such that $\langle T\xi_1, T\xi_2 \rangle = \langle \xi_1, \xi_2 \rangle$ for all $\xi_1, \xi_2 \in V$. For $V = \mathbb{R}^n$, we let $\mathrm{O}_n \coloneqq \mathrm{O}(\mathbb{R}^n)$.

Example 2.7. If $n = 1$, $\mathrm{O}_1 \subset \mathrm{GL}_1$ is $\{\pm 1\} \subset \mathbb{R}^\times$, so it's isomorphic to the cyclic group of order 2.

If $n = 2$, we can rotate by angles θ or reflect across lines, and playing with an orthonormal basis shows that all elements of O_2 must be rotations or reflections. Since O_2 is a Lie group, we can draw a picture as in Figure 2.

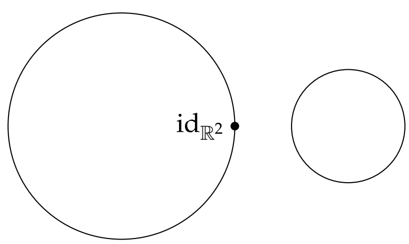

FIGURE 2. A picture of O_2. The left circle is the rotations; the right circle is the reflections, which in a sense form a circle half as long.

As with the affine symmetries, there's an extension

$$1 \longrightarrow \mathrm{SO}_2 \longrightarrow \mathrm{O}_2 \longrightarrow \{\pm 1\} \longrightarrow 1.$$

◀

Similarly, the isomorphisms of Euclidean space E, denoted $\mathrm{Euc}(E)$, are the affine isomorphisms preserving the inner product at each point. This again fits into an extension sequence

$$1 \longrightarrow V \longrightarrow \mathrm{Euc}(E) \longrightarrow \mathrm{O}(V) \longrightarrow 1.$$

All this is nice, but let's talk about manifolds. If X is a smooth manifold, we no longer have translations, and the linear symmetries talk about the tangent space. We'll see what kind of structures we get in this case.

The analogue of the torsor of bases is $\mathscr{B}(X) := \{(x,\beta) : x \in X, \beta : \mathbb{R}^n \xrightarrow{\sim} T_xX\}$. This admits a right action of $\mathrm{GL}_n(\mathbb{R})$ by precomposition, as on a vector space, and there is again a forgetful map $\pi : \mathscr{B}(X) \to X$ that ignores the basis.

If $x : U \to \mathbb{A}^n$ is a chart, then it defines a local section $U \to \mathscr{B}(X)$ sending

$$(x^1,\dots,x^n) \longmapsto \left((x^1,\dots,x^n), \left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^n}\right)\right).$$

If X is a Riemannian manifold, then we can also speak of orthonormal bases:

$$\mathscr{B}_{\mathrm{O}}(X) := \{(x,\beta) : x \in X, \beta : \mathbb{R}^n \xrightarrow{\cong} T_xX \text{ is an isometry}\}.$$

Again there is a forgetful map to X, but now a coordinate does *not* always determine a section: if the Riemann curvature tensor doesn't vanish, the image of an orthonormal basis of the tangent space at a point might not be orthonormal.

$\mathscr{B}(X)$ and $\mathscr{B}_{\mathrm{O}}(X)$ are not just sets but smooth manifolds, and the forgetful maps back to X are called fiber bundles (even principal bundles). We'll go back and discuss this in more detail.

Curvature. Let's end with something concrete. Let E be a Euclidean plane, an affine space with an underlying 2-dimensional inner product space.

Let $C \subset E$ be a 1-dimensional submanifold. Let's choose a *co-orientation* of C: an orientation of C is an orientation of its tangent bundle, so a co-orientation is an orientation of its normal bundle. In essence, this is choosing a side of the curve.[2] We'll use this to define a function $\kappa : C \to \mathbb{R}$ called the *(signed) curvature*. Intuitively, this should be positive if C is curved towards the side chosen by the co-orientation, and negative if it curves away, and a larger magnitude means a stronger curvature.

The Euclidean structure on E induces an inner product structure on T_xC for all $x \in C$ that varies smoothly, so C becomes a Riemannian manifold. Theorem 1.16 means there's nothing intrinsic about C we can measure, but the way in which it sits inside E is what κ will measure. This is an important dichotomy, between intrinsic geometry and extrinsic geometry. The Riemann curvature tensor is intrinsic, since it doesn't depend on an embedding, but the signed curvature will be extrinsic.

Lecture 3.

The curvature of a curve: 1/24/17

"And if you follow your nose... well, Euler's nose..."

In the next two lectures, we'll march through the theory of extrinsic curvature (which can fill an entire undergraduate course).

Let E be a Euclidean plane modeled on an inner product space of V, which acts on E by translations, and let $i : C \hookrightarrow E$ be an immersed 1-manifold.[3] Suppose C is co-oriented, meaning we've oriented its normal bundle (picking a side of C, so to speak). This determins a unit co-oriented normal vector e_1 at every $x \in C$, meaning the unique unit vector in $(\nu_{C\hookrightarrow E})_x$ with a positive orientation. We can also choose a unit tangent vector e_2 perpendicular to e_1, and there are two choices. Together they define an orthonormal basis at each point: $(e_1, e_2) : C \to \mathscr{B}_{\mathrm{O}}(V)$.

You learned how to do calculus with real-valued differential forms; in exactly the same way, it's possible to do calculus with vector-valued differential forms $\Omega^*_C(V)$, the forms modeled on functions $C \to V$. For $i, j \in \{1,2\}$, we can define $e_i \in \Omega^0_C(V)$ and $\mathrm{d}e_i \in \Omega^1_C(V)$, such that $\langle e_i, e_j\rangle = \delta_{ij}$ and the Leibniz rule is satisfied:

$$\langle \mathrm{d}e_i, e_j\rangle + \langle e_i, \mathrm{d}e_j\rangle = 0.$$

Thus, there exists an $\alpha \in \Omega^1_C$ such that

$$\mathrm{d}e_1 = -\alpha e_2 \qquad \text{and} \qquad \mathrm{d}e_2 = \alpha e_1.$$

[2]If $N \hookrightarrow M$ is an embedding and M is oriented, an orientation of N and a co-orientation of N determine each other.

[3]Especially if C is immersed but not embedded, it is helpful to remember i: when C self-intersects, remembering i is necessary for computing curvature.

In other words, applying d to the row vector $(e_1\ e_2)$ multiplies it by a skew-symmetric matrix:

$$\mathrm{d}\begin{pmatrix} e_1 & e_2 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 \end{pmatrix}\begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}.$$

Let $\theta^1, \theta^2 : C \to V^*$ define the dual basis at each point, i.e. at every $x \in C$, $\theta^i(e_j) = \delta^i_j$ as functions $C \to \mathbb{R}$. Then, $i^*\theta^2 \in \Omega^1_C$ and we can write

$$\alpha = k \cdot i^*\theta^2$$

for some function $k : C \to \mathbb{R}$.

Definition 3.1. The *curvature* of C is the function k.

Example 3.2. Let C denote the circle of radius R in the Euclidean plane $\mathbb{E}^2$. It's parameterized by coordinates $x = R\cos\phi$ and $y = R\sin\phi$, so

$$\begin{aligned} \mathrm{d}x &= -R\sin\phi\,\mathrm{d}\phi \\ \mathrm{d}y &= \quad R\cos\phi\,\mathrm{d}\phi. \end{aligned}$$

Let's choose the co-orientation in which the inward-pointing unit normal is positively oriented. Then,

$$e_1 = -\cos\phi\frac{\partial}{\partial x} - \sin\phi\frac{\partial}{\partial y}.$$

We also have to choose e_2: suppose it points clockwise along the circle. Then,

$$e_2 = \sin\phi\frac{\partial}{\partial x} - \cos\phi\frac{\partial}{\partial y}.$$

Thus, the dual basis is defined by

$$\begin{aligned} \theta^1 &= -\cos\phi\,\mathrm{d}x - \sin\phi\,\mathrm{d}y \\ \theta^2 &= \quad \sin\phi\,\mathrm{d}x - \cos\phi\,\mathrm{d}y, \end{aligned}$$

so $i^*\theta^2 = R\,\mathrm{d}\phi$. Then,

$$\mathrm{d}e_2 = \cos\theta\,\mathrm{d}\theta\frac{\partial}{\partial x} + \sin\theta\,\mathrm{d}\theta\frac{\partial}{\partial y} = -\mathrm{d}\theta e_1.$$

Thus, $\mathrm{d}e_2 = (1/R)i^*\theta^2(e_1)$. In particular, the curvature is $1/R$. It has units of $1/\text{length}$.

If we chose e_2 to point counterclockwise, there would be a sign change in θ^2, and another one in α, so they would cancel out to give the same result. ◂

Since the unit vector always has unit length in V, you can think of e_1 as a map $C \to S(V)$ (called the *Gauss map*), where $S(V)$ is the unit sphere inside V. At a point $p \in C$, we can define the tangent line T_pC at $i(p)$; the tangent line is a subspace of V. We can also consider the tangent line to $e_1(p) \in S(V)$, $T_{e_1(p)}S(V)$; both of these are the same space, the space of vectors in V perpendicular to $e_1(p)$.

This means the differential

$$(\mathrm{d}e_1)_p : T_pC \longrightarrow T_{e_1(p)}S(V) \tag{3.3}$$

is a map from a line to itself.

Theorem 3.4. *The map in* (3.3) *is multiplication by* $-k(p)$*.*

Proof.

$$\mathrm{d}e_1(e_2) = \alpha(e_2)\cdot d_2 = -ki^*\theta^2(e_2)e_2 = -k\cdot e_2. \qquad \boxtimes$$

Remark (History). The curvature may have been initially defined by Nicole Oresme in about 1350. It was again discovered by Huygens in c. 1650 and Newton in c. 1664. ◂

Here's a third approach to curvature. Let $i : C \hookrightarrow E$ be a co-oriented curve as usual, and assume C is embedded. For some $p \in C$, we can identify the normal line to $i(p)$ with $\mathbb{R}$, letting the positive numbers point into the positively oriented direction. Call this coordinate y. Given a choice of a unit tangent vector e_2, we can identify the tangent line with $\mathbb{R}$, again pointing the positive numbers in the x-direction. Call this coordinate x.

Lemma 3.5. *There exists an open set $U \subset E$ about p such that $C \cap U$ is the graph of a function $f : \mathbb{R} \to \mathbb{R}$ in the above xy-coordinate system such that*

- $f(0) = f'(0) = 0$, *and*
- $f''(0) = k(p)$.

Proof. The x-coordinate map $x|_C : C \to \mathbb{R}$ satisfies $\mathrm{d}x_p = \mathrm{id}_{T_pC}$; in particular, it's invertible. By the inverse function theorem, there's a local inverse $g : I \to C$, where $I \subset \mathbb{R}$ is an open interval. Define f to be $y \circ i \circ g$: since $i : C \hookrightarrow E$ and $y : E \to \mathbb{R}$, this is a map $I \to \mathbb{R}$. Write

$$e_1 = \frac{(-f', 1)}{\sqrt{1 + (f')^2}} \qquad \text{and} \qquad e_2 = \frac{(1, f')}{\sqrt{1 + (f')^2}}.$$

Then,

$$\mathrm{d}e_1 = \left(\frac{(-f'', 0)}{\sqrt{1 + (f')^2}} + \frac{(-f', 1)}{(1 + (f')^2)^{3/2}} f' \right) \mathrm{d}t.$$

At p,

$$\mathrm{d}e_1 = (-f''(0), 0)\,\mathrm{d}t = (-f''(0)\,\mathrm{d}t)e_2. \qquad \boxtimes$$

In calculus, we think of the tangent line as the best linear approximation to a function at a point, which only requires an affine space. Curvature is the process that goes one degree higher: you could ask for the *osculating parabola* to a curve at a point, the parabola that best approximates a curve at a point, or for the *osculating circle*, the circle that best approximates the curve at that point. Then, the curvature can be read off of the constants, e.g. it's 1 over the radius of the osculating circle. But knowing these parameters requires an inner product, hence a Euclidean space.

Prescribing curvature. We aim to solve the following problem: given an abstract curve C and a function $k : C \to \mathbb{R}$, construct an immersion $i : C \to E$ and a co-orientation such that k is the curvature of i.

Curvature requires thinking about a frame at each point if $i(C)$, so we should think about the bundle of orthonormal frame $\pi : \mathscr{B}_{\mathrm{O}}(E) \to E$. A point in $\mathscr{B}_{\mathrm{O}}(E)$ is a triple $(p; e_1, e_2)$, where $p \in E$ and (e_1, e_2) is an orthonormal basis of V. In particular, $\mathscr{B}_{\mathrm{O}}(E)$ is naturally a product $E \times \mathscr{B}_{\mathrm{O}}(V)$. We want to construct a lift $\widetilde{\imath} : C \to \mathscr{B}_{\mathrm{O}}(E)$ making the following diagram commute:

$$\begin{tikzcd} & \mathscr{B}_{\mathrm{O}}(E) \arrow[d, "\pi"] \\ C \arrow[ur, "\widetilde{\imath}"] \arrow[r, "i"] & E. \end{tikzcd}$$

This $\widetilde{\imath}$ is specified as a triple of functions on C, $\widetilde{\imath} = (p, e_1, e_2)$. Prescribing the curvature means we need this to satisfy

$$\mathrm{d}\begin{pmatrix} p & e_1 & e_2 \end{pmatrix} = \begin{pmatrix} p & e_1 & e_2 \end{pmatrix} \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & k\,\mathrm{d}t \\ \mathrm{d}t & -k\,\mathrm{d}t & 0 \end{pmatrix}}_{A(t)}. \tag{3.6}$$

We'll interpret A as a time-varying vector field on the manifold $\mathscr{B}_{\mathrm{O}}(E)$; then, we can evoke the basic theory of ordinary differential equations to prove there's a solution.

Digression. Let's recall what this basic theory of ordinary differential equations says. Let X be a smooth manifold and $(a, b) \subset \mathbb{R}$ be an interval. Projection onto the second factor defines a map $\pi_2 : (a, b) \times X \to X$, and we can pull the tangent bundle back along it:

$$\begin{tikzcd} \pi_2^* TX \arrow[r] \arrow[d, "p"] & TX \arrow[d] \\ (a,b) \times X \arrow[r, "\pi_2"] & X. \end{tikzcd}$$

Definition 3.7.

- A *time-varying vector field* is a section $\xi : (a,b) \times X \to \pi_2^* TX$ of $p : \pi_2^* TX \to (a,b) \times X$.
- An *integral curve* of ξ is an open interval $I \subset (a,b)$ and a function $\gamma : I \to X$ such that
$$\dot\gamma(t) = \xi_{(t,\gamma(t))}.$$

Time-varying vector fields correspond to ODEs and integral curves correspond to their solutions.

Theorem 3.8. *Given $(t_0, x_0) \in (a,b) \times X$, there exists an $\varepsilon > 0$ and an integral curve $\gamma : (t_0 - \varepsilon, t_0 + \varepsilon)$ such that $\gamma(t_0) = x_0$, and any two choices for γ agree on their common domain. Moreover, there is a maximal domain $J \subset (a,b)$ on which a solution exists and an integral curve $\gamma : J \to X$*

That is, solutions exist and are unique given an initial condition. However, they may not be globally defined.[4]

Just as $\mathscr{B}_{\mathrm{O}}(V)$ is a torsor for a right action of O_2 (an orthogonal basis composed with an orthogonal transformation is again an orthogonal basis), $\mathscr{B}_{\mathrm{O}}(E)$ is a torsor for the right action of Euc_2, the group of Euclidean transformations of $\mathbb{E}^2$. This torsor structure means the derivative of a curve in any neighborhood of the origin of the group defines a vector field on the torsor.

If $P(t)$ is a curve in O_2 such that $P(0) = \mathrm{id}$, then ${}^tP \cdot P = I$, so differentiating this condition, ${}^t \cdot P + \cdot P = 0$. That is, $T_e\mathrm{O}_2$ is the line of 2×2 skew-symmetric matrices over $\mathbb{R}$. Looking again at (3.6), the lower right entries of $A(t)$ are exactly such a matrix, so $A(t)$ is in fact a time-varying vector field on $\mathscr{B}_{\mathrm{O}}(E)$.

Corollary 3.9. *Using Theorem 3.8, given an initial $p \in E$ and an initial frame (e_1, e_2) on T_pE, there is a local and in fact a maximal solution to the prescribed curvature problem. This solution is unique up to the choice of (p, e_1, e_2).*

Uniqueness is usually expressed by saying that the group of symmetries of Euclidean space acts transitively on the solutions (so there's only one up to rotations and translations).

This is a somewhat elementary context for this material, but we'll adopt this perspective again and again. Eventually there will also be second-order conditions, e.g. when we define geodesics later.

Now, let's step up a dimension: let E be a Euclidean 3-space modeled on an inner product space V and $i : S \hookrightarrow V$ be an immersion of a 2-manifold together with a co-orientation. We can again define the unit co-oriented normal $\nu : S \to S(V)$. How can we define the curvature of this surface?

Euler solved this problem in 1760 by reducing it to something we've already done: let $L \in \mathbb{P}(T_pS)$ be a 1-dimensional subspace of the tangent space. There's a unique affine plane $\Pi(L)$ passing through p and containing L, and $\Pi(L) \cap S$ is a co-oriented curve in $\Pi(L)$. Let $k_p : \mathbb{P}(T_pS) \to \mathbb{R}$ be the function assigning to L the curvature of the curve $\Pi(L) \cap S$. Euler studied this function.

As before, locally we can write S as the graph of a function $f : T_pS \to \mathbb{R}$ with $f(0) = 0$ and $\mathrm{d}f_0 = 0$. The function k_p encodes the second derivative of f. This is expressed through the Hessian
$$\mathrm{Hess}\, f_0 : T_pS \times T_pS \longrightarrow \mathbb{R},$$
which is a symmetric bilinear form. In the context of geometry of surfaces, this Hessian is called the *second fundamental form* and denoted II_p.

Corollary 3.10. *For any $L \in \mathbb{P}(T_pS)$, $k_p(L) = II_p(\xi, \xi)$, where $|\xi| = 1$ and $\xi \in L$.*

The *first fundamental form* is the inner product
$$I_p := \langle -, - \rangle : T_pS \times T_pS \to \mathbb{R},$$
The second fundamental form may be nondegenerate (e.g. if S is flat), but we know the first is nondegenerate. This means the second fundamental form may be expressed in terms of the first fundamental form and some other operator S, called the *shape operator*:
$$II_p(\xi, \eta) = I_p(\xi, S(\eta)) = \langle \xi, S(\eta) \rangle.$$
Since II_p is symmetric, then S is self-adjoint. This means it has two real eigenvalues, so we can look at the eigenspaces, which are called the *principal lines* of S at p — unless the curvature is constant at p, in which case p is called an *umbilic point*.

Interestingly, we started with a very extrinsic notion of curvature of surfaces, but from this we've obtained some intrinsic geometry.

---

[4] In this class, we assume everything is smooth, but Theorem 3.8 is true in much greater generality, requiring only *Lipschitz continuity*, a condition slightly stronger than continuity. Many other things in this class may be relaxed, e.g. to C^2.

Lecture 4.

Curvature for surfaces: 1/26/17

"I didn't go into comedy, because I thought I would be safe here..."

Last time, we talked about the curvature of surfaces in a Euclidean plane; today, we will consider surfaces in a 3-dimensional Euclidean space E modeled on an inner product space $(V, \langle -,- \rangle)$, the vector space of translations of E.

Though E is abstractly isomorphic to $\mathbb{E}^3$, we won't fix an isomorphism by choosing coordinates; later, we'll want to pick special coordinates for E, so this would only complicate things.

Let $\Sigma \subset E$ be an embedded 2-manifold (some of our results will still apply when Σ is immersed), and assume Σ is co-oriented. Let $\nu : \Sigma \to V$ be the co-oriented positive unit normal.

Given a $p \in \Sigma$ and a plane $L \subset V$, $\Pi(L)$ denotes the plane through p containing L and ν. Then, $\Sigma \cap \Pi(L)$ is a curve, which is intuitively the curve "pointing in the L-direction at p."

The map assigning to L the curvature of $\Sigma \cap \Pi(L)$ at p is a function

$$k_p \colon \mathbb{P}(T_p\Sigma) \to \mathbb{R}.$$

Here, $\mathbb{P}(V)$ is the manifold of 1-dimensional subspaces of a vector space V.

We're going to get some information out of k_p. Let's first introduce special coordinates: choose an orthonormal basis in $\mathscr{B}_{\mathrm{O}}(E)$, so we obtain coordinats x^1 and x^2 in $T_p\Sigma$. As in the last lecture, the inverse function theorem provides for us an open set $U \subset T_p\Sigma$ containing 0, a function $f \colon U \to \mathbb{R}$, and an open $J \subset \mathbb{R}$ containing 0 such that $\Sigma \cap ((p + U) \times (p + J\nu))$ is the graph of f.

That is, there's a box inside E with an "xy-plane" $p + U$ and a "z-axis" pointing in the ν-direction, and inside this box, Σ is the graph of a function $f(x, y)$ on $p + U$. Furthermore, $f(p) = 0$ and $\mathrm{d}f_p = 0$, which is easy to check.

Last time, we defined the second fundamental form at p, $II_p = \mathrm{Hess}_p f : T_p\Sigma \times T_p\Sigma \to \mathbb{R}$. Based on what we proved last time, using the third incarnation of curvature, we got Corollary 3.10: $k_p(L) = II_p(\xi, \xi)$, where $\xi \in L$ is a unit vector.

This says the Hessian on the diagonal determines the curvature. This is because this is the second derivative of f, and we showed that if $\mathrm{d}f_p = 0$ for an f parameterizing a plane curve, then its second derivative computes the curvature.

On $T_p\Sigma$ we have two fundamental forms: the inner product, also known as the first fundamental form I_p, and the second fundamental form defined above. Since the first fundamental form is nondegenerate, then we can (and did) define the shape operator $S_p \in \mathrm{End}(T_p\Sigma)$ to satisfy the relation

$$\langle \xi, S_p(\eta) \rangle = II_p(\xi, \eta).$$

Since the inner product is nondegenerate, this uniquely defines $S_p(\eta)$. Moreover, since II_p is symmetric, then S_p is self-adjoint, i.e. $\langle \xi, S_p(\eta) \rangle = \langle S_p(\xi), \eta \rangle$ for all ξ and η. In particular, it's diagonalizable, and since $T_p\Sigma$ is two-dimensional, there are two possibilities:

(1) If there's only one eigenvalue $\lambda \in \mathbb{R}$, then $S_p = \lambda \cdot \mathrm{id}_{T_p\Sigma}$. In this case, p is called an umbilic point.
(2) If there are two eigenvalues λ_1 and λ_2 (suppose without loss of generality $\lambda_1 > \lambda_2$), then the two eigenspaces L_1 and L_2 form an orthogonal direct-sum decomposition $T_p\Sigma = L_1 \oplus L_2$. In this case, $S_p|_{L_i}$ is multiplication by λ_i. The L_i are called the *principal directions*, and the λ_i are called the *principal curvatures*. For any plane L,

$$k_p(L) = \frac{II_p(\xi, \xi)}{I_p(\xi, \xi)}.$$

The maximum of k_p is at L_1, and the minimum is at $L_2 II_p(\xi, \xi) I_p(\xi, \xi)$.

If you reverse the co-orientation, then $k \mapsto -k$ and $\lambda_i \mapsto -\lambda_i$. From this we get the *mean curvature* (named after one Mr. Mean)

$$H := \frac{\lambda_1 + \lambda_2}{2} = \frac{1}{2}\,\mathrm{Tr}(S_p),$$

a function $\Sigma \to \mathbb{R}$. Reversing the co-orientation sends $H \mapsto -H$. The *Gauss curvature* (named after Gauss) is

$$K := \lambda_1 \lambda_2 = \det S,$$

also a function $\Sigma \to \mathbb{R}$. This is unchanged when you reverse the co-orientation, which suggests that it comes from an intrinsic invariant! The units of the Gauss curvature has units $1/\text{length}^2$.

We also have the unit normal vector field $\nu \colon \Sigma \to S(V) \subset V$, and it tells us things about the curvature too.

Proposition 4.1. $\mathrm{d}\nu_p \colon T_p\Sigma \to T_p\Sigma$ *equals* $-S_p$.

Proof. Introduce "Euclidean coordinates" x^1, x^2 on $p + T_p\Sigma$, and let $f = f(x^1, x^2)$ be such that near p, Σ is the graph of f. Then,

$$\nu = \nu(x^1, x^2) = \frac{(-f_1, -f_2, 1)}{\sqrt{1 + f_1^2 + f_2^2}},$$

where $f_i = \frac{\partial f}{\partial x^i}$.

Exercise 4.2. Check that this is in fact a unit normal vector.

You can then calculate

$$\mathrm{d}\nu_p = \begin{pmatrix} -\partial_{11} f & -\partial_{12} f \\ -\partial_{21} f & -\partial_{22} f \end{pmatrix}\bigg|_p,$$

and this is $-\operatorname{Hess}_p f = -II_p$ as desired. (Here, it may help to remember that p is identified with $(0,0)$.) ⊠

Many people bemoan computations and coordinates, but certainly computations are useful, and coordinates are useful for computations. The solution is to judiciously choose coordinates to make computations simpler.

Now we can cover two beautiful theorems of Gauss, one global, one local.

Theorem 4.3 (Gauss-Bonnet). *Let* $\Sigma \subset E$ *be a closed, co-oriented surface and* $K : \Sigma \to \mathbb{R}$ *be its Gauss curvature. Let* $|\mathrm{d}A|$ *denote its Riemannian measure. Then,*

$$\int_\Sigma K\,|\mathrm{d}A| = 2\pi\chi(\Sigma). \tag{4.4}$$

Some of these words merit an explanation.

- A *closed manifold* is not the same thing as a closed subset: it means Σ is compact and has no boundary. It turns out all closed surfaces in E are co-orientable, but this is not necessarily true for immersed surfaces (e.g. the standard immersion of the Klein bottle).
- The Riemannian measure is discussed in the homework, but the essential idea is that on a Riemannian manifold, we know the lengths and angles of vectors, and therefore of the volume of the parallelogram that a basis $v_1, \dots, v_n$ of a tangent space spans, namely $|\det(\langle v_i, v_j\rangle_{ij})|$. Thus, we know how to compute volumes, which defines a measure that we can use to integrate functions.
- $\chi(\Sigma)$ is the Euler characteristic of Σ.

Though the proof we'll see uses the embedding (and implicitly the fact that Σ is orientable), all of the notions in (4.4) turn out to be extrinsic, and the theorem holds for abstract closed surfaces with a Riemannian metric, orientable or not.

Example 4.5. Consider a sphere $S^2(R)$ of radius R inside E. Then, every point is umbilic, and the Gauss curvature is $1/R^2$ everywhere. The surface area of the sphere is $4\pi R^2$, so

$$\int_{S^2} K\,|\mathrm{d}A| = 4\pi = 2\pi \cdot 2,$$

and indeed $\chi(S^2) = 2$. ◀

Theorem 4.3 is the first of many theorems which relate local and global geometry. It can be used to calculate global quantities, and to constrain local ones: for example, the sphere cannot have a metric with negative curvature, because its Euler characteristic is positive. The torus T^2 has Euler characteristic $\chi(T^2) = 0$, so any metric on it is either everywhere flat (no curvature) or has points of both positive and negative curvature. The standard embedding into $\mathbb{E}^3$ has points of both positive and negative curvature, but the flat torus can't be embedded isometrically into $\mathbb{E}^3$. It can be embedded into $\mathbb{E}^4$, as the product of two copies of the unit circle in $\mathbb{E}^2$.

Proof of Theorem 4.3. The proof will use the language of differential topology. Recall that if M and M' are oriented manifolds of the same dimension n, we can define the degree of a smooth map $\nu : M' \to M$, and if $\omega \in \Omega^n_M$, then

$$\int_{M'} \nu^*\omega = (\deg \nu) \int_M \omega.$$

In our case, ν is the unit vector map $\nu : \Sigma \to S(V)$; we computed that $\mathrm{d}\nu = -S$ (where S is the shape operator) in Proposition 4.1. Thus,

$$\det(\mathrm{d}\nu) = \det(-S) = K.$$

Let $\omega \in \Omega^2_{S(V)}$ be the area form; then,

$$v^*\omega = (\det \mathrm{d}\nu) \cdot \mathrm{d}A = K\,\mathrm{d}A.$$

Thus, when we integrate,

$$\int_\Sigma K\,\mathrm{d}A = \int_\Sigma \nu^*\omega = (\deg \nu) \int_{S(V)} \omega = 4\pi \deg \nu,$$

since the area of the unit sphere is 4π. Thus, it suffices to show $\deg \nu = \chi(\Sigma)/2$.

The Euler number emerges from the Poincaré-Hopf theorem, that if $\mathbf{v}$ is a vector field with isolated zeroes on Σ, the sum of the indices of $\mathbf{v}$ at its zeroes produces $\chi(\Sigma)$.

Compose ν with the quotient map $S(V) \twoheadrightarrow \mathbb{P}(V)$, and let q be a regular value of this composition, with two preimages $\pm\eta \in S(V)$. η pulls back to a vector field on Σ (constantly pointing in the direction η with unit length). Let ξ_p denote the vector field produced by projecting η onto $T\Sigma$; this has isolated zeros $x_1, \dots, x_n$.

You can do the computation without coordinates, but it's not hard in them: if $\eta = (0, 0, 1)$ (which is true up to a rotation), then at any x_i,

$$\xi = \frac{(f_1, f_2, f_1^2 + f_2^2)}{1 + f_1^2 + f_2^2},$$

and you don't have to worry about the denominator in the derivative, so

$$\mathrm{d}\nu_p = \mathrm{d}\xi_p = \begin{pmatrix} \partial_{11} f & \partial_{12} f \\ \partial_{21} f & \partial_{22} f \end{pmatrix}\Bigg|_p. \qquad \boxtimes$$

This is the first connection between topology and geometry.

You might wonder how this can be generalized. In odd dimensions, the Euler characteristic is zero, but for even dimensions, Chern proved the Gauss-Bonnet-Chern theorem in the 1940s which expresses the Euler characteristic in more complicated terms involving the Riemann curvature tensor.

Lecture 5.

Extrinsic and intrinsic curvature: 1/31/17

On the first day, we derived some equations as to when a Riemannian manifold is locally isometric to Euclidean space. Namely, if

$$A_{ijk} := \frac{\partial g_{\ell j}}{\partial x^k} + \frac{\partial g_{\ell k}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^\ell}$$

and

$$\Gamma^i_{jk} := \frac{1}{2} g^{i\ell} A_{\ell jk},$$

then we derived in (1.20)

$$E^i_{jk\ell} = \frac{\partial \Gamma^i_{j\ell}}{\partial x^k} - \frac{\partial \Gamma^i_{jk}}{\partial x^\ell} + \Gamma^m_{j\ell}\Gamma^i_{mk} - \Gamma^m_{jk}\Gamma^i_{m\ell},$$

and the Riemann curvature tensor

$$R = R^i_{ijk\ell} \frac{\partial}{\partial x^i} \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k \otimes \mathrm{d}x^\ell$$

is an obstruction to a Riemannian manifold being locally isometric to flat, Euclidean space. There's an exercise in the homework to show this is invariant under change of coordinates, and therefore R is an intrinsic object.

Today, we will tie this to the study of curvature of a surface Σ embedded in Euclidean 3-space E. Suppose Σ is co-oriented; then, at any $p \in \Sigma$, we defined the second fundamental form $II_p : T_p\Sigma \times T_p\Sigma \to \mathbb{R}$ and the shape operator $S_p : T_p\Sigma \to T_p\Sigma$ satisfying $II_p(\xi, \eta) = \langle \xi, S_p(\eta) \rangle$. The Gauss curvature is $k_p = \det S_p$, and the normal curvature is $II_p(\xi, \xi)/I_p(\xi, \xi)$.

Locally, Σ is the graph of a function $f = f(x^1, x^2)$ defined on an open neighborhood U in the x^1x^2-plane; here, x^1 and x^2 are special coordinates determined up to an element of O_2.

Theorem 5.1 (Gauss' *Theorema egregium*, c. 1823)**.** *In any of these special local coordinates at p,*

$$R^1_{212}(p) = k_p.$$

The right-hand side is defined extrinsically, determining how curves contained in orthogonal planes bend when embedded in the surface. But the left-hand side is defined intrinsically, depending only on the metric. Thus, the Gauss curvature is an intrinsic quantity, and does not depend on the co-orientation or embedding.

Corollary 5.2. *If Σ, Σ' are two surfaces embedded in E and $\varphi : \Sigma' \to \Sigma$ is an isometry, then $\varphi^* k = k'$.*

This is because the isometry preserves the metric, and the Gauss curvature can be computed only from the metric. This version is closer to how Gauss stated it.

Looking at Corollary 5.2, we know one embedding of the sphere of radius R into E such that the Gauss curvature is $k = 1/R^2$, and that the flat plane has curvature 0. Thus, map projections must be inaccurate: there's no way to map a plane onto any part of the sphere without distorting some length or angle.

The Riemannian curvature tensor on a Riemannian manifold X has a lot of symmetry. From (1.20), one can show that $R^i_{jk\ell} = -R^i_{j\ell k}$: it's skew-symmetric in these arguments. Thus,

$$R = \frac{1}{2} R^i_{jk\ell} \left(\frac{\partial}{\partial x^i} \otimes \mathrm{d}x^j \right) \otimes \mathrm{d}x^k \wedge \mathrm{d}x^\ell.$$

That is, $R \in \Omega^2_X(\operatorname{End} TX)$: the i and j indices give you an endomorphism of each tangent space. In fact, $R \in \Omega^2(\operatorname{SkewEnd} TX)$: the endomorphism is skew-symmetric.

Applying this to when $\dim X = 2$, if $V := T_pX$, then $R_p \in \operatorname{SkewEnd}(V) \otimes \Lambda^2 V^*$. The second component is the top exterior power, hence the *determinant line* $\operatorname{Det} V^*$. Moreover, $\operatorname{SkewEnd}(V) \xrightarrow{\cong} \Lambda^2 V^*$ through the map sending

$$T \longmapsto (\xi, \eta \longmapsto \langle \xi, T\eta \rangle).$$

This is akin to the way we got the shape operator out of the second fundamental form.

Anyways, this means $R_p \in (\operatorname{Det} V^*)^{\otimes 2} = (\operatorname{Det} V^{\otimes 2})^*$. What is this determinant line? The idea is that for every pair of vectors ξ, η, $\xi \wedge \eta$ can be identified with its area. We don't know what area 1 is *per se*, but we know given ξ', η' how to figure out the ratio of the area of $\xi' \wedge \eta'$ to that of $\xi \wedge \eta$, giving us a one-dimensional subspace.

But we do have an orthonormal basis produced by the metric, so we obtain a distinguished unit vector $e \in \operatorname{Det} V$. Thus, we can express $R^1_{212}(p)$ coordinate-independently, by evaluating $R_p \in ((\operatorname{Det} V)^{\otimes 2})^*$ on $e \otimes e \in (\operatorname{Det} V)^{\otimes 2}$.

Proof of Theorem 5.1. Near p, the surface is the graph of a function $(x^1, x^2) \mapsto (x^1, x^2, f(x^1, x^2))$. Let $f_i := \frac{\partial f}{\partial x^i}$, so

$$\left.\frac{\partial}{\partial x^1}\right|_{(x^1,x^2)} = (1, 0, f_1) \in T_{(x^1,x^2,f(x^1,x^2))}\Sigma \subset V$$

$$\left.\frac{\partial}{\partial x^2}\right|_{(x^1,x^2)} = (0, 1, f_2).$$

Let $\Delta := 1 + f_1^2 + f_2^2$. Then, you can calculate that the metric and its inverse satisfy

$$\begin{aligned} g_{11} &= 1 + f_1^2 & g^{11} &= \frac{1+f_2^2}{\Delta} \\ g_{12} &= f_1 f_2 & g^{12} &= -\frac{f_1 f_2}{\Delta} \\ g_{22} &= 1 + f_2^2 & g^{22} &= \frac{1+f_1^2}{\Delta}. \end{aligned}$$

The right-hand side is obtained from the left by inverting the 2×2 matrix for g_{ij}.

Exercise 5.3. Check that $A_{\ell jk} = 2 f_\ell f_{jk}$.

Recall that $f(0,0) = f_\ell(0,0) = 0$, so $A_{\ell ij}(0) = 0$ and $\Gamma^i_{jk}(0) = 0$. Thus,

$$R^1_{212}(0,0) = \left.\frac{\partial \Gamma^1_{22}}{\partial x^1}\right|_{(0,0)} - \left.\frac{\partial \Gamma^1_{21}}{\partial x^2}\right|_{(0,0)}.$$

Another plug-and-chug shows that

$$\begin{aligned} \Gamma^1_{22} &= \frac{1}{2} g^{11} A_{122} + \frac{1}{2} g^{12} A_{222} \\ &= \frac{2}{2\Delta}\Big((1+f_2^2) f_1 f_{22} - f_1 f_2 f_2 f_{22}\Big) \\ &= \frac{f_1 f_{22}}{\Delta}. \end{aligned}$$

A similar calculation shows

$$\Gamma^1_{21} = \frac{f_1 f_{21}}{\Delta}.$$

Therefore

$$\begin{aligned} R^1_{212}(0,0) &= (f_{11} f_{22} - f_{12} f_{21})|_{(0,0)} \\ &= \det \operatorname{Hess}_{(0,0)} f \\ &= k_p. \end{aligned}$$ ⊠

You should run through these calculations to make sure you understand them.

This provides us an interpretation of R, measuring curvature in different directions on the manifold. If it's equal to 0, the manifold is flat. We'd also like to interpret the Γ^i_{jk} symbols. This should be easier because they're built from first derivatives, whereas R was built from second derivatives.

Let's think about parallelism. In the Euclidean plane E, we have *global parallelism*, that given a vector field $\eta; E \to V$, we can compute its directional derivatives by considering the function $t \mapsto p + t\xi_p$ along a direction ξ_p (thought of as rooted at p). That is, the directional derivative of η in the direction ξ_p is

$$D_{\xi_p} \eta := \lim_{t \to 0} \frac{\eta(p + t\xi_p) - \eta(p)}{t}.$$

If $\gamma : (-\varepsilon, \varepsilon) \to E$ is a curve with $\gamma(0) = p$ and $\dot\gamma(0) = \xi_p$, then

$$D_{\xi_p} \eta = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \eta(\gamma(t)).$$

This doesn't work quite so well on embedded surfaces $\Sigma \hookrightarrow E$. There's a "poor man's parallelism" that translates a vector using the ambient parallelism on E, but there are lots of issues with this: it does not preserve tangency. So you project down onto $T\Sigma$, you say, but then sometimes you get the zero vector, and it feels like parallelism should preserve lengths and angles, right?

Let's ask a smaller question: given an immersed curve $\gamma\colon (-\varepsilon, \varepsilon) \to \Sigma$ with $\gamma(0) = p$ and $\dot\gamma(0) = \xi_p$, can we parallelize?

Definition 5.4. The *covariant derivative* $\nabla_{\xi_p} \eta$ is the orthogonal projection of $D_{\xi_p} \eta \in V$ onto $T_p\Sigma$.

Here, η is a section of the vector bundle $T\Sigma \to \Sigma$, and $\xi_p \in T_p\Sigma$, so $D_{\xi_p}\eta$ is in $T_pE = V$.

Definition 5.5. We say η is *parallel along* $\gamma : (a,b) \to \Sigma$ if $\nabla_{\dot\gamma}\eta = 0$ for all $t \in (a,b)$. If $\nabla_{\dot\gamma}\dot\gamma = 0$ (i.e. $\dot\gamma$ is parallel along γ), then γ is called a *geodesic*.

Here, η is a *vector field along* γ, meaning a section of the pullback bundle $\gamma^*T\Sigma \to (a,b)$. That is, at each t, $(\gamma^*T\Sigma)_t := T_{\gamma(t)}\Sigma$, and these fit together smoothly. So at each t, η chooses a tangent vector in $T_{\gamma(t)}\Sigma$. Thus, if γ is self-intersecting, we get a different tangent vector each time $\gamma(t)$ reaches the intersection point, so everything is still well-behaved.

Geodesics are the curves which have no acceleration along the curve, so the only acceleration is normal to the surface. For example, if you have a geodesic on a sphere (which is a great circle), it's only accelerating perpendicular to the sphere, the minimal acceleration necessary to stay on the sphere.

One of the first things we prove in multivariable calculus is that the directional derivative is linear in the direction. This is still true here, where we derived it from parallelism, among the oldest notions in geometry.

Lemma 5.6.

(1) *$\nabla_{\xi_p}\eta$ is linear in ξ_p, i.e. $\nabla\eta \in T_p^*\Sigma$.*

(2) *∇_{ξ_p} satisfies a Leibniz rule:*
$$\nabla_{\xi_p}(f\eta) = (\xi_p \cdot f)\eta + f\nabla_{\xi_p}\eta.$$

(3)
$$\nabla_{\xi_p}(\eta + \eta') = \nabla_{\xi_p}\eta + \nabla_{\xi_p}\eta'.$$

(4)
$$\xi_p\langle \eta, \eta'\rangle = \langle \nabla_{\xi_p}\eta, \eta'\rangle + \langle \eta, \nabla_{\xi_p}\eta'\rangle.$$

Though we've defined geodesics extrinsically, they are intrinsic, and we'll be able to describe them using the symbols Γ^i_{jk}.

Theorem 5.7. *Let η be a vector field on Σ. Then, $\nabla\eta$ is intrinsic, i.e. determined solely by the metric.*

In particular, $\nabla\eta \in \Omega^1_\Sigma(T\Sigma)$.

Proof. Use coordinates $(x^1, x^2, f(x^1,x^2))$ as before, so Σ is the graph of f. A basis for the tangent space is $\frac{\partial}{\partial x^1} = (1,0,f_1)$ and $\frac{\partial}{\partial x^2} = (0,1,f_2)$ as before.

Write $\eta = \eta^i\frac{\partial}{\partial x^i}$ with $\eta^i = \eta^i(x^1,x^2)$ for $i = 1,2$. Thus, $\eta = (\eta^1, \eta^2, \eta^i f_i)$, so by a Leibniz rule
$$D\eta = (\mathrm{d}\eta^2 1, \mathrm{d}\eta^2, f_i\,\mathrm{d}\eta^i + \eta^i\,\mathrm{d}f_i).$$
In particular, $D\eta_p = (\mathrm{d}\eta^1_p, \mathrm{d}\eta^p_2, *)$ and $\nabla\eta_p = (\mathrm{d}\eta^1_p, \mathrm{d}\eta^2_p, 0)$, or
$$\nabla\eta = \mathrm{d}\nabla^i \cdot \frac{\partial}{\partial x^i},$$
so $\nabla\frac{\partial}{\partial x^i} = 0$ at p.

We used special coordinates x^1, x^2; let's change to arbitrary coordinates y^1, y^2. Calculus on manifolds (or, for grade students, canceling fractions) shows that
$$\frac{\partial}{\partial y^a} = \frac{\partial x^i}{\partial y^a}\frac{\partial}{\partial x^i},$$
so at p,
$$\begin{aligned}\nabla\frac{\partial}{\partial y^a} &= \frac{\partial^2 x^i}{\partial y^b \partial y^a}\,\mathrm{d}y^b \cdot \frac{\partial}{\partial x^i}\\ &= \underbrace{\frac{\partial^2 x^i}{\partial y^b \partial y^a}\frac{\partial y^c}{\partial x^i}}_{Q^c_{ab}}\,\mathrm{d}y^b \cdot \frac{\partial}{\partial y^c}.\end{aligned}$$

We'll finish the proof by showing $Q_{ab}^c = \Gamma_{ab}^c$ as computed in the (y^1, y^2)-coordinate system. Since Γ_{ab}^c doesn't depend on the metric, neither can $\nabla\eta$.

At p,

$$\begin{aligned} g_{ab} &= \left\langle \frac{\partial}{\partial y^a}, \frac{\partial}{\partial y^b} \right\rangle = \frac{\partial x^i}{\partial y^a}\frac{\partial x^j}{\partial y^b} g_{ij} \\ &= \sum_i \frac{\partial x^i}{\partial y^a}\frac{\partial x^i}{\partial y^b}, \end{aligned}$$

so (again at p),

$$\frac{\partial g_{ab}}{\partial y^c} = \sum_i \frac{\partial^2 x^i}{\partial y^c \partial y^a}\frac{\partial x^i}{\partial y^b} + \frac{\partial x^i}{\partial y^a}\frac{\partial^2 x^i}{\partial y^c \partial y^b}.$$

Therefore

$$A_{dab} = 2\sum_i \frac{\partial^2 x^i}{\partial y^a \partial y^b}\frac{\partial x^i}{\partial y^d}$$

and

$$g^{cd} = \sum_j \frac{\partial y^c}{\partial x^j}\frac{\partial y^d}{\partial x^j}.$$

Thus,

$$\Gamma_{ab}^c = \frac{1}{2}\gamma^{cd} A_{dab} = \sum_{i,j} \frac{\partial y^c}{\partial x^j} \underbrace{\frac{\partial y^d}{\partial x^j}\frac{\partial x^i}{\partial y^d}}_{\delta_j^i} \frac{\partial^2 x^i}{\partial y^a \partial y^b}.$$

Thus, we can collapse to when $i = j$, which recovers Q_{ab}^c. ⊠

Embedded in this proof is the calculation as to how the Γ_{ij}^k change when the coordinates change.

This allows us to define a differential equation for geodesics: if $\eta = \eta^a \frac{\partial}{\partial y^a}$, so that

$$\nabla\eta = \left(\frac{\partial \eta^c}{\partial y^b} + \Gamma_{ab}^c \eta^a\right) \mathrm{d}y^b \frac{\partial}{\partial y^c},$$

then the *geodesic equation* for $\dot\gamma = \xi = \xi^b \frac{\partial}{\partial y^b}$ is

$$\nabla_\xi \xi = \left(\ddot y^a + \Gamma_{ab}^c \dot y^a \dot y^b\right)\frac{\partial}{\partial y^c} = 0. \tag{5.8}$$

That is, for surfaces, we have intrinsic notions of parallelism and geodesics. This holds in more generality. Next time, we'll say one more thing about surfaces in space (looking at the normal component of the directional derivative), and recover the second fundamental form on it. Then, we'll do some background lectures on differential geometry.

Lecture 6.

Vector fields and integral curves: 2/2/17

We've talked about how for surfaces, the sectional curvature at a point p is a map $k_p : \mathbb{P}(V) \to \mathbb{R}$. More generally, the Riemann curvature tensor is $R \in \Omega_X^2(\mathrm{SkewEnd}\, TX)$, so for any $x \in X$, if $V = T_xX$, $R_x : \Lambda^2 V \to \mathrm{SkewEnd}(V) \cong \Lambda^2 V^*$, hence determined by a bilinear map $\Lambda^2 V \times \Lambda^2 V \to \mathbb{R}$. If $\Pi \subset V$ is a two-dimensional subspace, we can evaluate $R_x(\Pi, \Pi) \in \mathbb{R}$, so letting Π vary, we obtain the *sectional curvature* $K_x : \mathrm{Gr}_2(T_xX) \to \mathbb{R}$. Here, $\mathrm{Gr}_2(V)$ is the *Grassmannian*, the manifold of 2-dimensional subspaces of V.

Let's return to the case of a co-oriented surface Σ embedded in a 3-dimensional Euclidean space E, and let η be a vector field on Σ. Then, the directional derivative in the direction ξ_p (a vector ξ rooted at p) is $D^{(E)}_{\xi_p}\eta \in \mathbb{R}^3$. This has tangential and normal components:

$$D_{\xi_p}\eta = \underbrace{\nabla_{\xi_p}\eta}_{\text{tangential}} + \underbrace{B(\xi_p, \eta) \cdot \nu}_{\text{normal}}.$$

Last time, we showed in Theorem 5.7 that the tangential part is intrinsic to Σ. If $\gamma : (-\varepsilon, \varepsilon) \to \Sigma$ is a curve, then $\nabla_{\dot\gamma}\eta$ is the covariant derivative of η along γ. We said η is parallel along γ if $\nabla_{\dot\gamma}\eta = 0$, and γ is a geodesic if $\dot\gamma$ is parallel along γ.

Last time, we saw that geodesics are the solutions to the ODE (5.8); by the general theory of ODEs, solutions exist and are unique. Given a smooth curve $\gamma\colon (a, b) \to \Sigma$, a $t_0 \in (a, b)$, and an $\eta_0 \in T_{\gamma(t_0)}\Sigma$, there exists a unique parallel vector field η along γ such that $\eta_{\gamma(t_0)} = \eta_0$.

But local parallelism doesn't imply global parallelism. Consider a geodesic triangle on a sphere,[5] as in Figure 3. If you start with a vector tangent along the upper left piece and parallel-transport it to the lower

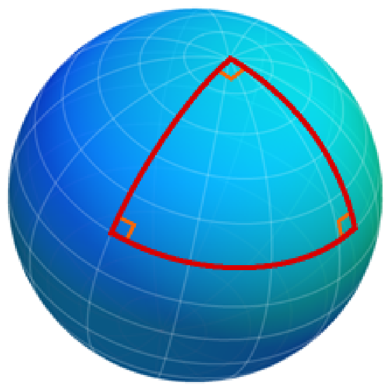

FIGURE 3. A geodesic triangle on the sphere. Each line is a piece of a great circle, and all three angles are right angles. Source: http://world.mathigon.org/Dimensions_and_Distortions.

left corner, then parallel-transport it to the lower-right corner, then parallel-transport it back upm to the pole, you'll end up with a different vector than you started with.

Parallel-transport can be thought of as a way of isometrically identifying tangent spaces (which we can canonically do in Euclidean space, but not always on manifolds).

Proposition 6.1. *Let $\gamma\colon (a, b) \to \Sigma$ be a curve and η_1, η_2 be parallel vector fields along γ. Then, $\gamma\eta_1, \eta_2\colon \gamma \to \mathbb{R}$ is constant.*

Proof. Let $\xi = \dot\gamma$, so

$$\xi\langle \eta_1, \eta_2\rangle = \langle \nabla_\xi \eta_1, \eta_2 + \eta_1, \nabla_\xi \eta_2\rangle = 0.$$ ⊠

If γ is closed, traveling along γ from a point p to itself will produce an isometry of $T_p\Sigma$, but not always the identity: in the above example, it was a nontrivial rotation. This is called the *holonomy* around the loop. If the top angle is θ (and the sphere has radius 1), the area of the triangle is θ.

Now let's discuss the geometry of the normal component $B(\xi_p, \eta)$.

Lemma 6.2.

(1) *If $f \in \Omega^0_\Sigma$, $B(\xi_p, f\eta) = fB(\xi_p, \eta)$.*
(2) *B is the second fundamental form.*

[5]There are other surfaces than spheres, of course! Check out the homework for some examples.

Remark. If X is a manifold, $\mathcal{X}(X)$ denotes the space of vector fields on X. We say T is linear if it's $\mathbb{R}$-linear, i.e. for all $\xi, \xi' \in \mathcal{X}(X)$ and $\lambda \in \mathbb{R}$,

$$T(\lambda\eta + \eta') = \lambda T(\eta) + T(\eta').$$

We say T is *linear over functions* if it's Ω^0_X-linear, i.e. for any $f \in \Omega^0_X$ (i.e. $C^\infty(X)$), $T(f\eta) = fT(\eta)$. This means T doesn't differentiate f or anything like that, e.g. $\nabla_{\xi_p}(f\eta) = (\xi_p f)\eta + f\nabla_{\xi_p}\eta$ is not linear over functions. If T is linear over functions, it defines a cotangent vector field. ◂

Proof of Lemma 6.2. For the first part, $B(\xi_p, \eta) = \langle D_{\xi_p}\eta, \nu\rangle$, so at p,

$$\begin{aligned} B(\xi_p, f\eta) &= \langle D_{\xi_p} f\eta, \nu\rangle = \langle (\xi_p f)\eta + fD_{\xi_p}\eta, \nu\rangle \\ &= f(p)\langle D_{\xi_p}\eta, \nu\rangle = f(p)B(\xi_p, \eta). \end{aligned}$$

So B determines a bilinear map $B : T_p\Sigma \times T_p\Sigma \to \mathbb{R}$.[6] Let's see that it agrees with II.

Choose coordinates (x^1, x^2) such that Σ is the graph of $f(x^1, x^2)$. Then, $\frac{\partial}{\partial x^1} = (1, 0, f_1)$ and $\frac{\partial}{\partial x^2} = (0, 1, f_2)$ as normal. Write

$$\xi_p = \xi^i \frac{\partial}{\partial x^i}\bigg|_p \qquad \text{and} \qquad \eta = \eta^i \frac{\partial}{\partial x^i}\bigg|_p,$$

for $\xi^i, \eta^i \in \mathbb{R}$, we want to extend η_p to a map on local vector fields. We have liberty in this extension, so let's make our life easier and set $\eta^i(x^1, x^2) = \eta^i$, so it's constant. Thus, $\eta = (\eta^1, \eta^2, \eta^i f_i)$, so at $(x^1, x^2) = (0, 0)$,

$$D_{\xi_p}\eta = (0, 0, \eta^i(\xi_p f_i)) = (0, 0, \eta^i\xi^j f_{ij}).$$

Since $\nu_p = (0, 0, 1)$,

$$B(\xi_p, \eta_p) = f_{ij}\xi^i\eta^j = II(\xi_p, \eta_p). \qquad \boxtimes$$

This provides a coordinate-free interpretation of the second fundamental form, which is nice.

The first chapter of Warner's "Foundations on Differentiable Manifolds and Lie Groups" is a good reference for a lot of this material.

Anyways, this means the directional derivative is

$$D_{\xi_p}\eta = \nabla_{\xi_p}\eta + II_p(\xi_p, \eta_p)\nu_p.$$

We can use this to derive a coordinate-free interpretation of the shape operator:

$$\begin{aligned} II(\xi, \eta) &= \langle D_\xi\eta, \nu\rangle = \xi\langle\eta, \nu\rangle - \langle\eta, D_\xi\nu\rangle \\ &= -\langle\eta, D_\xi\nu\rangle = -\langle\eta, \mathrm{d}\nu(\xi)\rangle, \end{aligned}$$

so $S_p = -\mathrm{d}\nu_p$.

Though we'll begin talking about abstract Riemannian manifolds, these concrete examples, which you can draw, make these ideas clearer, and have fairly direct analogues in the abstract setting.

The choice of a unit tangent vector on a curve is discrete: on each connected component, you can flip between e_1 and $-e_1$. On a surface, though, it's possible to rotate a local frame $\{e_1, e_2, e_3\}$ (where e_1 is normal and e_2 and e_3 are tangential), so there's a continuous choice, and this continues to be true in higher dimensions.

In mathematics, a common approach to studying a situation where one needs to make a choice is to study all choices (sometimes you can make a convenient choice). Thus, we'll have to study this and other structures attached to smooth manifolds, including Lie groups, Lie derivatives, and a little geometry of smooth manifolds. But the payoff is that understanding how the frames change determines a lot of the Riemannian geometry.

[6]Equivalently, $B(\xi_p, \eta)$ only depends on the value of η at p.

Vector fields. Let X be a smooth manifold and ξ be vector field on X.

Definition 6.3. A piecewise-smooth curve $\gamma : (a, b) \to X$ is an *integral curve* of ξ if for all $t \in (a, b)$, $\dot{\gamma}(t) = \xi_{\gamma(t)}$.

That is, ξ is tangent along γ. We'd like to impose some constant-velocity constraint on this, but need a Riemannian metric to do that. It's possible to show that integral curves always exist, by starting with a finite approximation, iterating in a nice manner, and using some soft analysis (the contraction mapping theorem) to show there's a solution.

Theorem 6.4. *Given an $x_0 \in X$ and a vector field ξ on X, there's a unique maximal integral curve $\gamma : (a(x_0), b(x_0)) \to x$ (where $a, b \in [-\infty, \infty]$), such that $\gamma(0) = x_0$ and if $\mu(a, b) \to X$ is an integral curve for ξ with $\mu(0) = x_0$, then $a(x_0) \le a < 0 < b \le b(x_0)$ and $\mu = \gamma|_{(a,b)}$.*

This curve will be called γ_{x_0}.

Definition 6.5. With notation as in the above definition, γ is *complete* if for all $x_0 \in X$, $a(x_0) = -\infty$ and $b(x_0) = \infty$.

Here's a useful sufficient condition:

Theorem 6.6. *If for some Riemannian metric $\langle -, - \rangle$, $\|\xi\| : X \to \mathbb{R}^{\ge 0}$ is bounded, then ξ is complete.*

Corollary 6.7. *If X is a compact manifold, all ξ are complete.*

We would like to travel along a vector field. Let $\varphi(t, x) := \gamma_x(t)$; if ξ is complete, then $\varphi\colon \mathbb{R} \times X \to X$ is well-defined. Otherwise, you may find yourself flowing off the end of the world!

Definition 6.8. A *flow* on a manifold X is a (discrete) group homomorphism $\widehat{\varphi} : \mathbb{R} \to \mathrm{Diff}(X)$ (the latter group is under composition) such that the action map $\varphi\colon \mathbb{R} \times X \to X$ is C^∞.

That is, we ask for $\widehat{\varphi}(t_1 + t_2) = \widehat{\varphi}(t_2) \circ \widehat{\varphi}(t_1)$ and $\varphi(t_1 x) = \widehat{\varphi}(t)(x)$. We don't know how to express smoothness on $\mathrm{Diff}(X)$, so the smoothness criterion is stated in terms of the finite-dimensional manifolds $\mathbb{R}$ and X.

Given a vector field ξ, define for some $t \in \mathbb{R}$ the set

$$\mathcal{D}_t := \{x \in X \mid t \in (a(x), b(x))\}.$$

The following theorem rests on a proof of Theorem 6.4. This is often left unproven in geometry textbooks, but can be found, e.g. in Lang's ODE book or in Coddington-Levinson.

Theorem 6.9.

(1) *$\mathcal{D}_t$ is open.*
(2) *The map $\varphi_t : \mathcal{D}_t \to \mathcal{D}_{-t}$ is a diffeomorphism.*
(3) *The domain of $\varphi_{t_2} \circ \varphi_{t_1}$ is a subset of the domain of $\varphi_{t_1+t_2}$, and on that domain, $\varphi_{t_2} \circ \varphi_{t_1} = \varphi_{t_1+t_2}$.*
(4) *If $x \in X$ and $U \subset X$ is an open set containing x, then there's a $V \subset U$ and an $\varepsilon > 0$ such that $\varphi(-\varepsilon, \varepsilon) \times V$ maps into U.*
(5) *If ξ is complete, then $\mathcal{D}_t = X$ for all t, and φ is a global flow.*

This all relies on ξ being fixed with time (an *autonomous system*). If $\xi = \xi(t)$ varies with time, then a lot of these arguments don't work; in particular $\varphi_{t_2} \circ \varphi_{t_1} \ne \varphi_{t_1+t_2}$. Fortunately, we can use a neat technique to dispatch these.

A vector field ξ is a section of the tangent bundle $p : TX \to X$, and a time-varying vector field $\xi(t)$ is a section of the pullback: if $\pi_2\colon (a, b) \times X \to X$ denotes projection onto the second component, $\widetilde{\xi}$ is a section of the pullback $\pi_2^* TX \to (a, b) \times X$. That is, on each time-slice, you get a section of TX, and these vary smoothly.

Let $\widehat{\xi} = \frac{\partial}{\partial t} + \widetilde{\xi}$, so $\widehat{\xi}$ is a vector field on $(a, b) \times X$. Given an initial condition (t_0, x_0), Theorem 6.4 says there's an integral curve $\widehat{\gamma}\colon (\widehat{a}, \widehat{b}) \to (a, b) \times X$. Letting $\widehat{\gamma}(t) = (t, \gamma(t))$, then $\dot{\widehat{\gamma}}(t) = \widehat{\xi}_{\widehat{\gamma}(t)}$. This is $(1, \dot{\gamma}(t)) = (1, \widetilde{\xi}(t))$, so $\gamma(t)$ is what we were looking for, and the solution exists, at least locally!

Once we have this flow, we're going to look at what happens if you carry various objects along the flow, e.g. vector fields or differential forms.

Lecture 7.

Tangential structures of manifolds: 2/7/17

Today, we'll talk about tangential structures: vector fields, subspaces of the tangent bundle, etc. We'll later dualize and look at functions and differential forms, and still later use this to understand Lie groups.

Let X be a smooth manifold and $\xi \in \mathcal{X}(X)$ be a vector field on it. Let φ_t be the (local) flow generated by ξ: given a vector field, traveling along its integral curves moves points along the flow, and we can therefore flow all sorts of other objects: functions, vectors, differential forms, connections...

The Lie derivative is the instantaneous change in a quantity as you flow it. By a *covariant* object we mean something which pushes forward along maps, e.g. vectors. Similarly, *contravariant* things are those which pull back under maps.

Definition 7.1. Let T be a covariant object. Then, the *Lie derivative* of T is

$$\mathcal{L}_\xi T \coloneqq \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_{-t})_* T.$$

If T is a contravariant object, then the Lie derivative of T is

$$\mathcal{L}_\xi T \coloneqq \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_t)^* T.$$

Example 7.2. Let $f \in \Omega^0_X$ be a function. Functions pull back, so this is contravariant: $\varphi_t^* f(p) = f(\varphi_t(p))$. Thus, the Lie derivative of f is

$$\mathcal{L}_\xi f(p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \varphi_t^* f(p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(\varphi_t(p)).$$

Thus, this is the directional derivative along a curve $\gamma : t \mapsto \varphi_t(p)$, since $\gamma(0) = p$ and $\dot\gamma(p) = \xi_p$. In symbols, $\mathcal{L}_\xi f(p) = \mathrm{d}f|_p(\xi_p)$. ◀

Example 7.3. Let $\eta \in \mathcal{X}(X)$ be a vector field. To compute its Lie derivative, let's introduce local coordinates $x^1, \dots, x^n$, so $\xi = \xi^i \frac{\partial}{\partial x_i}$, $\eta = \eta^j \frac{\partial}{\partial x^j}$, and $\varphi(t,x) = (\varphi^1(t,x), \dots, \varphi^n(t,x))$. Since φ_t is the flow, it satisfies

$$\dot\varphi(t,x) = \xi_{\varphi(t,x)}, \quad \left.\frac{\partial \varphi^i}{\partial x^j}\right|_{t=0} = \delta^i_j, \quad \text{and} \quad \left.\frac{\partial^2 \varphi^i}{\partial x^k \partial x^k}\right|_{t=0} = 0.$$

Therefore we compute

$$(\varphi_{-t})\eta = (\varphi_{-t})_* \left(\eta^i \frac{\partial}{\partial x^i}\right) = \left(\varphi_t^* \eta^i\right)(\varphi_{-t})_* \frac{\partial}{\partial x^i}.$$

This pullback and pushforward are

$$\begin{aligned}
\varphi_t^* \eta^i &= \eta^i\left(\varphi^1(t,x), \dots, \varphi^n(t,x)\right)\\
(\varphi_{-t})_* \frac{\partial}{\partial x^i} &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{t=0} \varphi_{-t}\left(x^1, \dots, x^i + s, \dots, x^n\right)\\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \left(\varphi^j\left(-t, x^1, \dots, x^i + s, \dots, x^n\right)\right)_j\\
&= \frac{\partial \varphi^j}{\partial x^i}(\varphi(-t,x)) \cdot \frac{\partial}{\partial x^j}.
\end{aligned}$$

Putting these together,

$$(\varphi_{-t})\eta = \eta^i(\varphi(t,x)) \cdot \frac{\partial \varphi^j}{\partial x^i}(\varphi(-t,x)) \frac{\partial}{\partial x^j}$$

and

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(\varphi_{-t})_*\eta &= \left.\left(\frac{\partial\xi^i}{\partial\eta^j}\dot{\varphi}^k\frac{\partial\varphi^j}{\partial x^i}\frac{\partial}{\partial x^j} - \eta^i\frac{\partial^2\varphi^j}{\partial x^k\partial x^i}\dot{\varphi}^k\frac{\partial}{\partial x^j} - \eta^i\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\varphi^j}{\partial x^i}\right)\frac{\partial}{\partial x^j}\right)\right|_{t=0}\\ &= \frac{\partial\eta^i}{\partial x^k}\xi^k\delta^j_i\frac{\partial}{\partial x^j} - \eta^i\frac{\partial\xi^j}{\partial x^i}\frac{\partial}{\partial x^j}\\ &- \left(\xi^i\frac{\partial\eta^j}{\partial x^i} - \eta^i\frac{\partial\xi^j}{\partial x^i}\right)\frac{\partial}{\partial x^j}.\end{aligned}$$

You might think this formula in coordinates is boring; if so, do it yourself. ◀

Therefore we have proven:

Proposition 7.4 (Lie derivative of vector fields in local coordinates)**.**

$$\mathcal{L}_{\xi^i\partial/\partial x^i}\eta^j\frac{\partial}{\partial x^j} = \xi^i\frac{\partial\eta^j}{\partial x^i}\frac{\partial}{\partial x^j} - \eta^j\frac{\partial\xi^i}{\partial\eta^j}\frac{\partial}{\partial x^i}.$$

Notice the almost symmetry between ξ and η in the above result.

Corollary 7.5. $\mathcal{L}_\xi\eta = -\mathcal{L}_\eta\xi$.

Let's take a different point of view. Recall that the vector fields on X are alternatively the *derivations* $\xi : \Omega^0_X \to \Omega^0_X$, i.e. linear maps satisfying the Leibniz rule

$$\xi(fg) = (\xi f)g + f(\xi g).$$

Hopefully you proved this in a differential topology class; if not, it's an exercise! You could also use the more general criterion that anything tensorial over functions is a tensor.

Definition 7.6. If $\xi, \eta \in \mathcal{X}(X)$, their *Lie bracket* $[\xi, \eta]$ is the derivation

$$f \longmapsto \xi(\eta f) - \eta(\xi f). \tag{7.7}$$

Lemma 7.8. (7.7) *satisfies the Leibniz rule and hence is actually a derivation.*

Proof. Applying it to fg, we get

$$\begin{aligned}\xi(\eta(fg)) - \eta(\xi(fg)) &= \xi(\eta(f)\cdot g - f\eta(g)) - \eta(\xi(f)\cdot g - f\xi(g))\\ &= \xi\eta f\cdot g + \eta f\cdot\xi g + \xi f\cdot\eta g + f\xi\eta g - \eta\xi f\cdot g - \xi f\cdot\eta g - \eta f\cdot\xi g - f\eta\xi g.\end{aligned}$$

The cross terms cancel, so this is

$$= ([\xi,\eta]f)\cdot g + f\cdot[\xi,\eta]g.$$ ☒

This is in fact the same thing as the Lie derivative!

Proposition 7.9. *Let $\xi, \eta \in \mathcal{X}(X)$.*

(1) *The Lie bracket is the Lie derivative: $[\xi,\eta] = \mathcal{L}_\xi\eta$.*
(2) *The Lie bracket is antisymmetric: $[\eta,\xi] = -[\xi,\eta]$.*
(3) *(Jacobi identity)*
$$[[\xi,\eta],\zeta] + [[\eta,\zeta],\xi] + [[\zeta,\xi],\eta] = 0.$$
(4)
$$[f\xi, g\eta] = fg[\xi,\eta] + f(\xi g)\eta - g(\eta f)\xi.$$

Parts (2) and (3) follow formally from properties of the commutator in an associative algebra, and parts (1) and (4) can be checked in local coordinates. This uses the fact that mixed partials commute, or for all i and j,

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

So the Lie bracket vanishes for coordinate vector fields, but may not vanish in general, reflecting yet another obstruction to global parallelism; the Lie bracket measures the failure of vector fields to commute as derivations.

There are a couple of other forms of the Jacobi identity. Some of them say that $\mathcal{L}_\xi = [\xi, –]$ is a derivation, meaning it satisfies the Leibniz rule:

$$\mathcal{L}_\xi[\eta,\zeta] = [\mathcal{L}_\xi\eta,\zeta] + [\eta,\mathcal{L}_\xi\zeta].$$

Let $\xi,\eta \in \mathcal{X}(X)$, and let φ_t, ψ_s be their local flows. Consider flowing along a rectangle: t and s are small, and we flow by $\psi_{-s}\varphi_{-t}\psi_s\varphi_t$. That the Lie bracket isn't always zero means this flow might not get back to where it started.

Specifically, choose local coordinate $x^1,\dots,x^n$, and let $(t,s) \mapsto x^i(t,s)$ be the map sending $(t,s) \mapsto \psi_{-s}\varphi_{-t}\psi_s\varphi_t(p)$. This maps the rectangle into the manifold. The two axes (where $s = 0$ or $t = 0$) are collapsed onto p.

Proposition 7.10. *In this case,*

$$\mathcal{L}_\xi\eta(p) = \left.\frac{\partial^2 x^i}{\partial t\partial s}\right|_{\substack{s=0\\t=0}} \frac{\partial}{\partial x^i}.$$

Proof. The proof is, again, a calculation.

$$\left.\frac{\partial x^i}{\partial s}\frac{\partial}{\partial x^i}\right|_{s=0} = -\eta = (\varphi_{-t})_*\eta.$$

The first term is constant with respect to t, so disappears when we differentiate with respect to r, and the second term becomes $\mathcal{L}_\xi\eta(p)$ by the definition of the Lie derivative. ⊠

These computations aren't just a nuisance: there's lots of ways to think about them or to choose notation, and these choices of notation are particularly nice for not making mistakes, etc. They also demonstrate how to think about local coordinates.

Now, suppose ξ and η are complete (so the flow exists for all time), so we can make a global statement. Pushing forward a vector field along a map $X' \to X$ doesn't in general define a vector field: if the map isn't surjective, there's not a vector at every point, and if it's not injective, there may be multiple choices for the vector at a given point. If $\psi\colon X \to X$ is a diffeomorphism, however, you can push vector fields forward, and $\psi_*\xi$ generates the flow $\psi\circ\varphi_t\circ\psi^{-1}$. (This is a nice exercise using the existence and uniqueness of ODEs.) Thus, $\psi_*\xi = \xi$ iff $\psi\varphi_t\psi^{-1} = \varphi_t$ for all t, and therefore $(\varphi_s)_*\xi = \xi$ for all s iff $\psi_s\varphi_t\psi_{-s}\varphi_{-t} = 0$ for all s and t. This is one direction of the following.

Proposition 7.11. $[\xi,\eta] = 0$ *iff for all s and t, $\psi_s\varphi_t\psi_{-s}\varphi_{-t} = \mathrm{id}$, where ψ_s and φ_t are the flows associated with η and φ, respectively.*

The direction we didn't prove follows from observing that $[\eta,\xi] = \frac{\mathrm{d}}{\mathrm{d}s}(\psi_{-s})_*\xi = 0$. The situation in this proposition might hold more generally, however.

Definition 7.12. Let $\psi\colon X' \to X$ be a smooth map, $\xi' \in \mathcal{X}(X')$, and $\xi \in \mathcal{X}(X)$. Then, ξ' and ξ are *ψ-related* if $(\psi_*)_{p'}\xi'_{p'} = \xi_{\psi(p')}$ for all $p' \in X'$.

The consequence is that this preserves Lie bracket data.

Proposition 7.13. *Suppose ξ' and ξ are ψ-related and η' and η are ψ-related. Then, $[\xi',\eta']$ and $[\xi,\eta]$ are ψ-related.*

You can (and should) prove this yourself, using the formula for the Lie bracket as the commutator of the derivatives of the flows. This is part of a general principle, that to prove something about vector fields, it's useful to think of them as infinitesimally small curves. Alternatively, you could translate this into the language of derivations and prove it that way.

Another useful thing you can do with vector fields is use them to define local coordinate systems.

Theorem 7.14. *Let $\xi \in \mathcal{X}(X)$ and $p \in X$ be such that $\xi_p \neq 0$. Then, there exists a local chart U containing p and local coordinates $(x^1,\dots,x^n)\colon U \to X$ such that $\frac{\partial}{\partial x^1} = \xi$ on all of U.*

Proof sketch. Let $(y^1, \dots, y^n)$ be local coordinates about p such that p maps to the origin, so $y^i(p) = 0$, and $\frac{\partial}{\partial y^1}\Big|_p = \xi_p$. Let φ_t be the flow generated by ξ. Then, define

$$(x^1, \dots, x^n) \longmapsto \varphi_{x^1}(0, x^2, \dots, x^n), \tag{7.15}$$

where on the right-hand side, the argument of φ_{x^1} is written in y^i-coordinates. Then, using y^i-coordinates, you can check that the differential at 0 is invertible. This means (7.15) defines a local coordinate system, and the computation will show that $\frac{\partial}{\partial x^1} = \xi$. ⊠

Now let's do this with multiple vector fields.

Theorem 7.16. *Let $k \le \dim X$ and $\xi_1, \dots, \xi_k \in \mathcal{X}(X)$. If $\{\xi_1|_p, \dots, \xi_k|_p\}$ are linearly independent at a $p \in X$, then there exists a local chart U containing p and local coordinates $(x^1, \dots, x^n) : U \to X$ with $\frac{\partial}{\partial x^i} = \xi_i$ for $i = 1, \dots, k$ iff $[\xi_i, \xi_j] = 0$ for all $1 \le i, j \le k$.*

Proof sketch. Let $y^1, \dots, y^n$ be local coordinates about p such that $y^i(p) = 0$ for all i and $\frac{\partial}{\partial y^i}\Big|_p = \xi_p$ for $i = 1, \dots, k$. Let $\varphi_t^{(i)}$ be the flow generated by ξ_i, and let

$$(x^1, \dots, x^n) \longmapsto \varphi_{x^k}^{(k)} \cdots \varphi_{x^2}^{(2)} \varphi_{x^1}^{(1)}(0, \dots, 0, x^{k+1}, \dots, x^n).$$

Again, the argument of φ is in y^i-coordinates. Now, check that the differential is the identity (in y^i-coordinates) and $\frac{\partial}{\partial x^i} = \xi_i$ in the same way. However, this will use that $\varphi^{(i)}$ and $\varphi^{(j)}$ commute, which is true iff the Lie brackets vanish. ⊠

So defining a coordinate system through vector fields only works if they commute. Think back to Riemannian geometry: we studies surfaces by introducing special coordinates at each point, but we may not be able to promote this to an orthonormal frame. Being a coordinate system and being orthonormal are in tension, and what measures this tension is the Riemann curvature tensor.

You might want to generalize this to plane fields or hyperplane fields at a point instead of just vector fields. Then, you'll get integral manifolds, and this perspective can be useful to define maps between manifolds. Again, there will be a condition about commutators, encoded in a theorem due to Klebsch (and attributed to Frobenius).

Lecture 8.

Distributions and Foliations: 2/9/17

Today, we'll work through distributions, the local Frobenius theorem, foliations, and some other things preparing us for Lie groups. After that, we'll be able to return to Riemannian geometry. Throughout today's lecture, X is a smooth manifold.

Definition 8.1.

(1) A *distribution* is a vector subbundle $E \subset TX \to X$.
(2) If E is a distribution, a vector field $\xi \in \mathcal{X}(X)$ *belongs to* E if ξ is a section of $E \to X$, i.e. $\xi_p \in E_p$ for all $p \in X$.
(3) E is *involutive* (or *integrable*) if whenever $\xi, \eta \in E$, $[\xi, \eta] \in E$ as well.
(4) A submanifold $Y \subset X$ is an *integral manifold* of E if for all $p \in Y$, $T_pY = E_p$.

You can think of a distribution as a hyperplane field at every point, distinguishing the directions in T_pX that are contained in E_p. If ξ belongs to E, then at every point it's contained in the hyperplane defined at that point. So a distribution is first-order information for constructing a manifold: we've specified the tangent space, and want to find a curved manifold which satisfies it.

Proposition 8.2. *If* $\operatorname{rank} E = 1$, *$E$ is involutive.*

These E are also called *line fields*.

Proof. We can work locally: choose an open $U \subset X$ and a nonzero section e of $E|_U \to U$, so that any $\xi, \eta \in \Gamma(E|_U)$ can be written as $\xi = fe$ and $\eta = ge$ for $f, g \in \Omega^0_U$. Then, compute:

$$\begin{aligned}[\xi, \eta] &= [fe, ge] = fg[e,e] + f(e \cdot g)e - g(e \cdot f)e \\ &= [f(e \cdot g) - g(e \cdot f)]e,\end{aligned}$$

which is also a section of $E|_U$. ☒

Integral manifolds to line fields also exist: locally choose a nonvanishing ξ belonging to E and choose its integral curve.

Remark. However, not every line field admits a global nonvanishing section. Let X be the Möbius band, the total space of a line bundle $\pi \colon X \to S^1$ defined by gluing $\mathbb{R} \times [0,1] \to [0,1]$ with degree -1. We'll let E be the copy of X inside the tangent bundle. That is, π is a submersion, so we have a short exact sequence of vector bundles

$$0 \longrightarrow T(X/S^1) \longrightarrow TX \longrightarrow \pi^* TS^1 \longrightarrow 0.$$

This has no natural splitting (splittings will be called covariant derivatives later). Since the fiber X/S^1 is a vector space V, then its tangent space at each $v \in V$ is identified with V again. We choose E to be this subspace of TX, and ultimately because the Möbius strip isn't orientable, E admits no global nonvanishing section. ◭

Example 8.3. Let $X = \mathbb{A}^3_{x,y,z}$ and

$$E := \operatorname{span}\left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y} + x\frac{\partial}{\partial z} \right\}.$$

Then, for any $(x_0, y_0, z_0) \in \mathbb{A}^3$, there's a piecewise smooth map γ from $(0,0,0)$ to (x_0, y_0, z_0) with $\dot\gamma \in E$. The idea is to zigzag from varying x to varying y and z.

However, E is not involutive:

$$\left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y} + x\frac{\partial}{\partial z} \right] = \frac{\partial}{\partial z} \notin E.$$ ◭

Checking involutivity seems kind of hard, but there's a nice criterion. We know $\operatorname{span}\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\}$ is involutive, and this turns out to be the key.

Theorem 8.4 (Local Frobenius). *Let $E \subset TX$ be a distribution of rank k.*

(1) *E is involutive iff about every $p \in X$, there exist local coordinates $x^1, \dots, x^n$ on some $U \subset X$ containing p such that*

$$E|_U = \operatorname{span}\left\{ \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k} \right\}. \tag{8.5}$$

(2) *If this is true, then any connected integral manifold $Y \subset U$ has the form $x^m = c^m$, for $m = k+1, \dots, n$, for some $c^{k+1}, \dots, c^n \in \mathbb{R}$.*

One can imagine a structure on a manifold which is an atlas only of charts satisfying (8.5). This is a geometric structure in the same way that a Riemannian metric is a geometric structure. Part (2) says that in this case, the integral manifolds are parallel to the $x^1 \cdots x^k$-plane on U.

Proof. It suffices to prove the forward direction; the reverse direction is a computation. Choose local coordinates $y^1, \dots, y^n$ about p such that

(1) $y^i(p) = 0$ for each i,
(2) $y(U)$ has the form $-\varepsilon < y^i < \varepsilon$ for some $\varepsilon > 0$ (i.e. it's a cube),[7] and
(3) $E_p = \operatorname{span}\left\{ \frac{\partial}{\partial y^1}, \dots, \frac{\partial}{\partial y^k} \right\}$.

[7] This is easy to accomplish: given any local coordinates, we can pick the open cube of side length ε around 0 and restrict y to it.

Let $\pi : \mathbb{A}^n \to \mathbb{A}^k$ be projection onto the first k coordinates; then, shrinking ε if necessary, $\pi_* : \mathbb{R}^n \to \mathbb{R}^k$ sends E_y isomorphically onto $\mathbb{R}^k$ for all $y \in y(U)$. Let $\xi_1, \dots, \xi_k$ be defined to satisfy $\pi_*\xi_i = \frac{\partial}{\partial y^i}$.

Now we have a basis $\xi^1, \dots, \xi^k$ in E that is π_*-related to $\frac{\partial}{\partial y^i}$. In particular,

$$\pi_*[\xi_i, \xi_j] = \left[\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\right] = 0.$$

If E is involutive, then we conclude that $[\xi_i, \xi_j] = 0$, so by Theorem 7.16, we can find local coordinates $x^1, \dots, x^n$ (shrinking ε if necessary) such tht $\frac{\partial}{\partial x^i} = \xi_i$ for $i = 1, \dots, k$. ⊠

This is a very geometric approach to a question originally motivated by systems of differential equations; involutivity is the analogue of mixed partials commuting.

The coordinate system guaranteed by Theorem 8.4 for an involutive E is called an *E-coordinate system* or a *slice*.

Example 8.6. Consider $\mathbb{A}^2$ and E be a one-dimensional constant distribution. Then, the integral submanifolds of $\mathbb{A}^2$ are the lines parallel to E. Quotient out by $\mathbb{Z}^2 \subset \mathbb{R}^2$ and let $X = \mathbb{A}^2/\mathbb{Z}^2$ be the torus. The integral manifolds project to integral submanifolds of the torus, but their structure depends on the embedding $\mathbb{Z}^2 \subset \mathbb{R}^2$.

If the images of $(1,0)$ and $(0,1)$ are related by rational numbers, then each integral manifold has finite length and closes up to a circle. For example, it could be a circle that traverses each direction of the torus once, or more than once. If the integral manifolds traverse one direction p times and the other q times, this characterizes the homology class $(p,q) \in H_1(\mathbb{A}^2/\mathbb{Z}^2) \cong \mathbb{Z}^2$. This is a *foliation*, and the integral submanifolds are called *leaves*. In this case, the quotient space parameterizing the leaves is a manifold, specifically S^1.

If instead the images of $(1,0)$ and $(0,1)$ aren't related by rational numbers, then each integral manifold has infinite length, and is in fact dense in the torus! So we get an immersed submanifold, not an embedded one. This is a more pathological case, and is one of the reasons we require manifolds to be second countable. The quotient space is still a topological space, but not Hausdorff. ◂

Example 8.7. The *Hopf fibration* is a map $\pi\colon S^3 \to S^2$. Identify S^3 as the unit sphere in $\mathbb{C}^2$, the set $\{(\xi^1, \xi^2) \in \mathbb{C}^2 \mid |\xi^1|^2 + |\xi^2|^2 = 1\}$. Then, identify $S^2 = \mathbb{CP}^2$, the space of 1-dimensional subspaces in $\mathbb{C}^2$. The Hopf fibration sends an $x \in S^2$ to the line containing it.

Each fiber of π is a circle $S^1 \subset S^3$, which is actually unknotted. Thinking of $S^3 = \mathbb{A}^3 \cup \{*\}$, there's one fiber which is a straight line, and the rest are circles winding around this line. Any two circlular fibers are linked with linking number 1. There's a distribution $E = \ker \pi_* \subset TS^3$. The fibers of the Hopf fibration are integral manifolds for E. ◂

The leaf space of a foliation is sometimes a manifold, but not always; in fact, it's an example of a *noncommutative space*, a significant example in Alain Connes' noncommutative geometry.

Definition 8.8. A *k-dimensional foliation* on X is a decomposition of X as

$$X = \coprod_{\alpha \in A} \mathscr{F}_\alpha,$$

where each $\mathscr{F}_\alpha$ is a k-dimensional immersed submanifold, such that for every $p \in X$, there exist local coordinates $x^1, \dots, x^n$ on a $U \subset X$ such that $\mathscr{F}_\alpha \cap U$ is a union of slices $x^m = c^m$ (for $m = k+1, \dots, n$). These $\mathscr{F}_\alpha$ are called the *leaves* of the foliation $\mathscr{F}$.

The *tangent bundle* to a foliaton $\mathscr{F}_\alpha$ is

$$T\mathscr{F} := \coprod_{p \in X} T_p\mathscr{F}_{\alpha(p)},$$

where $p \in \mathscr{F}_{\alpha(p)}$.

$T\mathscr{F} \to X$ is locally trivial; in particular, it's a vector bundle. Foliations are a richly studied subject.

Theorem 8.9. *Let $E \subset TX$ be an involutive distribution. Then, there exists a foliation $\mathscr{F}$ such that $T\mathscr{F} = E$. In particular, the leaves are maximal integral manifolds for E, and there is a unique such maximal integral submanifold though every point.*

So this is the solution to the differential equation that E defines; rather than solve it at each point, we solve it at all points at once.

We will often only need one leaf when we apply this, but the full generality of the theorem is too beautiful to pass up.

Proof sketch of Theorem 8.9. Second-countability of X is important in this theorem: we cover X by a *sequence* of charts $U_1, U_2, \dots$ of E-coordinate systems. If $S \subset U_i$ is a slice, then $\pi_0(S \cap U_j)$ is either empty, finite, or countably infinite.

Introduce an equivalence relation on slices where $S \subset U_i$ and $T \subset U_j$ are equivalent if there's a fintie sequence $i_0 = i, i_1, i_2, \dots, i_N = j$ and slices $S_{i_j} \subset U_{i_j}$ such that $S_{i_j} \cap S_{i_{j+1}} \neq \varnothing$, $S_0 = S$, and $S_N = T$. Let $\mathscr{F}_\alpha$ be the union of the slices S in an equivalence class; these will be the leaves in the foliation.

Clearly $\{\mathscr{F}_\alpha\}$ is a decomposition of X into a disjoint union of pieces, but we need each inclusion map $\mathscr{F}_\alpha \to X$ to be an immersion, which is data. The construction we've defined already does this: it's covered by these U_{i_j}. Then, since there's only countably many charts, $\mathscr{F}_\alpha$ must be second countable. ☒

This implies the distribution in Example 8.3 isn't involutive: we saw how to get from the origin to any point, but if it were involutive, there would have to be a foliation, and there would be no way to travel between different leaves in a foliation.

We'll now pass to something more formal, which involves functions on X. This involves drawing fewer pictures, which is maybe a little unfortunate, but the calculations we can make are useful.

Let $\Omega_X^\bullet$ denote the algebra of differential forms on X, which is a $\mathbb{Z}$-graded algebra, graded by

$$\Omega_X^\bullet = \bigoplus_{k\in\mathbb{Z}} \Omega_X^k.$$

The algebra structure is wedge product: $\alpha, \beta \mapsto \alpha \wedge \beta$. It's commutative, in the sense of a $\mathbb{Z}$-graded algebra: if α and β are homogeneous elements of degrees $|\alpha|$, resp. $|\beta|$, then

$$\alpha \wedge \beta - (-1)^{|\alpha||\beta|} \beta \wedge \alpha. \tag{8.10}$$

This is called the *Koszul sign rule*.

Definition 8.11. Let $A^\bullet$ be a commutative (meaning (8.10) holds) $\mathbb{Z}$-graded algebra and $T\colon A^\bullet \to A^\bullet$. T is a *derivation* if it's a linear map of degree s[8] and T satisfies the Liebniz rule

$$T(\alpha \wedge \beta) = T\alpha \wedge \beta + (-1)^{s|\alpha|} \alpha \wedge T\beta.$$

Example 8.12.

(1) The *Cartain-de Rham differential* $\mathrm{d} : \Omega_X^\bullet \to \Omega_X^{\bullet+1}$ is a derivation of degree 1.

(2) Let $\xi \in \mathcal{X}(X)$. If $\alpha \in \Omega_X^\bullet$ and ξ generates the flow φ_t, then the Lie derivative is a derivation of degree 0. This boils down to it satisfying the Leibniz rule: to verify this, use the fact that the pullback commutes with wedge product:

$$\varphi_t^*(\alpha \wedge \beta) = \varphi_t^* \alpha \wedge \varphi_t^* \beta.$$

Now, you can differentiate this with the usual Leibniz rule. There's no sign here, but that's okay, because the degree is 0. ◀

Given a vector field ξ, there's a derivation of degree -1 called contraction ι_ξ, which can be defined somewhat algebraically. We'll talk about this next time; so we now have three operators on differential forms of degrees 1, 0, and -1; under Lie bracket, we get some interesting relations between them, namely

$$\mathscr{L}_\xi = [\mathrm{d}, \iota_\xi].$$

There's one more operation on differential forms, integration over the fiber in the context of a fiber bundle. If you know all the brackets between these four operators, you can recover the calculus of differential forms and therefore calculus in the usual sense on X.

[8]A linear map of $\mathbb{Z}$-graded algebras has *degree s* if it sends homogenenous elements of degree n to homogeneous elements of degree $n+s$ for all $n \in \mathbb{Z}$.

Lecture 9.

Lie algebras and Lie groups: 2/14/17

Let $A = \bigoplus_{n\in\mathbb{Z}} A_n$ be a $\mathbb{Z}$-graded algebra. (That is, multiplication sends $A_m \cdot A_n$ into A_{m+n}.) For example, if V is a real vector space, $\Lambda^\bullet V$ is a commutative $\mathbb{Z}$-graded algebra, because we specified commutativity by the Koszul sign rule

$$z \wedge w = (-1)^{|z||w|} w \wedge z.$$

Last time, we defined a derivation of degree d to be a linear map $T : A \to A$ of degree d (so $A_n \to A_{n+d}$ on homogeneous elements) such that

$$T(\alpha \wedge \beta) = T\alpha \wedge \beta + (-1)^{|T||\alpha|} \alpha \wedge T\beta.$$

Definition 9.1. Let V be a vector space and $\xi \in V$. Define $\iota_\xi \colon \Lambda^k V^* \to \Lambda^{k-1} V^*$ by

$$(\iota_\xi \alpha)(\xi_2, \dots, \xi_k) := \alpha(\xi, \xi_2, \xi_3 \dots, \xi_k).$$

This is called *contracting with the first index*.

There is a natural identification $\Lambda^k V^*$ with $(\Lambda^k V)^*$ defined through a pairing $\Lambda^k V^* \otimes \Lambda^k V \to \mathbb{R}$ sending

$$\alpha_1 \wedge \cdots \wedge \alpha_k \otimes \xi, \dots, \xi \longmapsto \det(\alpha_i(\xi_j)_{i,j}),$$

for $\alpha_i \in V^*$ and $\xi_j \in V$.

Proposition 9.2.

(1) *ι_ξ is adjoint to left exterior multiplication $\varepsilon_\xi \colon \Lambda^{k-1} V \to \Lambda^k V$, the map sending $z \mapsto \xi \wedge z$.*
(2) *$\iota_\xi : \Lambda^\bullet V^* \to \Lambda^\bullet V^*$ is a derivation of degree -1.*
(3) *If $\xi, \eta \in V$, then $\iota_\xi \iota_\eta + \iota_\eta \iota_\xi = 0$.*

If we define the *commutator* of $a, b \in A$ to be

$$[a, b] := ab - (-1)^{|a||b|} ba,$$

then (3) can be restated as $[\iota_\xi, \iota_\eta] = 0$. For any odd-degree element a, $[a, a] = 2a^2$.

In a basis, $\xi = \xi^i e_i$ and

$$\alpha = \frac{1}{k!} \alpha_{i_1 \dots i_k} e^{i_1} \wedge \cdots \wedge e^{i_k},$$

so that contraction is

$$\iota_\xi \alpha = \frac{1}{(k-1)!} \xi^i \alpha_{i_1 i_2 \dots i_k} e^{i_2} \wedge \cdots \wedge e^{i_k}.$$

Proof sketch of Proposition 9.2. Part (1) is immediate from the definition.

For part (2), let $\alpha_1, \dots, \alpha_{k+\ell} \in V^*$, let $\alpha = \alpha_1 \wedge \cdots \wedge \alpha_k$ and $\beta = \alpha_{k+1} \wedge \cdots \wedge \alpha_{k+\ell}$. Then,

$$\iota_{\xi_1}(\alpha \wedge \beta)(\xi_2 \wedge \cdots \wedge \xi_{k+\ell}) = \det(\alpha_i(\xi_j)),$$

and you can finish the proof by expanding along $j = 1$.

For part (3),

$$(\iota_\xi \iota_\eta \alpha)(z) = \iota_\eta \alpha(\xi \wedge z) = \alpha(\eta \wedge \xi \wedge z).$$
$$(\iota_\eta \iota_\xi \alpha)(z) = \iota_\xi \alpha(\eta \wedge z) = \alpha(\xi \wedge \eta \wedge z) = -\alpha(\eta \wedge \xi \wedge z). \qquad \boxtimes$$

Now suppose X is a smooth manifold and $\xi \in \mathcal{X}(X)$. Then, $\iota_\xi \colon \Omega_X^\bullet \to \Omega_X^{\bullet-1}$ defined pointwise is a derivation of degree -1. Now we have three kinds of derivations on differential forms: d has degree 1, ι_ξ has degree -1, and $\mathcal{L}_\xi$ has degree 0.

The vector space of derivations on $\Omega_X^\bullet$ is closed under commutators, so it's worth asking what derivations we obtain from commutators of our three amigos. By symmetry, there are six ones to consider.

Theorem 9.3.

- $[\iota_\xi, \iota_\eta] = 0$.
- $[\mathcal{L}_\xi, \iota_\eta] = \iota_{[\xi,\eta]}$.
- $[\mathrm{d}, \iota_\eta] = \mathcal{L}_\eta$ *(the Cartan formula).*

- $[\mathcal{L}_\xi, \mathcal{L}_\eta] = \mathcal{L}_{[\xi,\eta]}$.
- $[\mathcal{L}_\xi, \mathrm{d}] = 0$.
- $[\mathrm{d}, \mathrm{d}] = 0$.

Some of these we've proven before; others are new, but can be computed from the definitions. You can quickly check that the degree of a commutator is the sum of the degrees of its two arguments.

The calculations above are in a sense generators and relations for a $\mathbb{Z}$-graded Lie algebra (with the Koszul sign rule).

Definition 9.4. A *Lie algebra* is a (real) vector space $\mathfrak{g}$ together with a bilinear form $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that $[\xi, \eta] = -[\eta, \xi]$ and the *Jacobi identity* holds:

$$[[\xi, \eta], \zeta] + [[\eta, \zeta], \xi] + [[\zeta, \xi], \eta] = 0$$

for all $\xi, \zeta, \eta \in \mathfrak{g}$.

To define a graded Lie algebra, one must insert signs corresponding to the Koszul sign rule, e.g. the bracket satisfies

$$[\xi, \eta] = (-1)^{|\xi||\eta|}[\eta, \xi],$$

and there are additional signs in the graded Jacobi identity.

The commutation relations either follow directly or have been already proven, except for the second and third ones. Let's prove those. To check that two derivations on $\Omega_X^\bullet$ agree, it suffices to check on Ω_X^0 and $\mathrm{d}\Omega_X^0$, since every element of $\Omega_X^\bullet$ is (locally) a sum of wedges of these forms, and the Leibniz rule behaves the same for both derivations.

Let's prove that $[\mathcal{L}_\xi, \iota_\eta] = \iota_{[\xi,\eta]}$. If $f \in \Omega_X^0$, the right-hand side is 0 and the left-hand side is

$$[\mathcal{L}_\xi, \iota_\eta]f = -\iota_\eta \mathcal{L}_\xi f = 0.$$

Now let's check on $\mathrm{d}f$:

$$\begin{aligned}
[\mathcal{L}_\xi, \iota_\eta]\mathrm{d}f &= \mathcal{L}_\xi \iota_\eta \mathrm{d}f - \iota_\eta \mathcal{L}_\xi \mathrm{d}f \\
&= \mathcal{L}_\xi(\eta \cdot f) - \iota_\eta\, \mathrm{d}\mathcal{L}_\xi f \\
&= \xi\eta \cdot f - \eta\xi \cdot f \\
&= \iota_{[\xi,\eta]}\mathrm{d}f.
\end{aligned}$$

The Cartan formula is proved in a similar manner.

Corollary 9.5. *Let $\alpha \in \Omega_X^1$ and $\xi, \eta \in \mathcal{X}(X)$. Then,*

$$\mathrm{d}\alpha(\xi, \eta) = \xi \cdot \alpha(\eta) - \eta \cdot \alpha(\xi) - \alpha([\xi, \eta]).$$

This is a very useful formula, and sometimes is used to define d!

Proof.

$$\begin{aligned}
\mathrm{d}\alpha(\xi, \eta) &= \iota_\eta \iota_\xi \mathrm{d}\alpha \\
&= -\iota_\eta\, \mathrm{d}\iota_\xi \alpha + \iota_\eta \mathcal{L}_\xi \alpha \\
&= -\eta \cdot \alpha(\xi) + \mathcal{L}_\xi \iota_\eta \alpha - \iota_{[\xi,\eta]}\alpha \\
&= -\eta \cdot \alpha(\xi) + \xi \cdot \alpha(\eta) - \alpha([\xi, \eta]).
\end{aligned}$$

⊠

Let X be a smooth manifold and $E \subset TX$ be a distribution as last time. Let

$$\mathscr{I}(E) := \{\alpha \in \Omega_X^\bullet : \alpha|_E = 0\}.$$

That is, $\mathscr{I}(E)$ is the annihilator of E. For example, on $\mathbb{A}^3$, if

$$E = \mathrm{span}\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}\right\},$$

$\mathscr{I}(E) = \mathrm{span}\{x\,\mathrm{d}y - \mathrm{d}z\}$.

Lemma 9.6. *$\mathscr{I}(E)$ is closed under* d *iff E is involutive.*

We defined integrability in terms of vector fields, and the dual condition on differential forms is a little simpler. The key step in the proof is showing that if ξ and η belong to E and $\alpha \in \Omega^1_X \cap \mathscr{I}(E)$, Corollary 9.5 shows that

$$\mathrm{d}\alpha(\xi,\eta) = \xi \cdot \alpha(\eta) - \eta \cdot \alpha(\xi) - \alpha([\xi,\eta]) = -\alpha([\xi,\eta]).$$

In the example, it's easy to check that E isn't involutive: $\mathrm{d}(x\,\mathrm{d}y - \mathrm{d}z) = \mathrm{d}x \wedge \mathrm{d}y \notin \mathscr{I}(E)$.

Let's talk about Lie groups.

Definition 9.7. A *Lie group* G is a smooth manifold equipped with a group structure such that multiplication $G \times G \to G$ and inversion $G \to G$ are smooth.

Like many definitions in mathematics, this is a marriage of two structures, with compatibility conditions (the group structure is smooth). There should be a compatibility condition for the identity, that inclusion of the identity is smooth, but this follows automatically for manifolds. It is needed for defining topological groups more generally, however.

Example 9.8. We have already seen many examples of Lie groups.

- If V is a finite-dimensional, real vector space, $(V,+)$ is an abelian Lie group.
- $\mathrm{GL}(V) = \mathrm{Aut}(V)$ is a Lie group, in general nonabelian. If $V = \mathbb{R}^n$, $\mathrm{GL}(\mathbb{R}^n)$ is called $\mathrm{GL}_n(\mathbb{R})$, and is the group of invertible $n \times n$ matrices. As a manifold, it's an open subset of the space of all $n \times n$ matrices (inversion is the preimage of an open subset of $\mathbb{R}$ under a polynomial, hence smooth, map).
- If V has an inner product, we can look at its orthogonal group $\mathrm{O}(V)$. If $V = \mathbb{R}^n$ and the inner product is standard, this is called O_n.
- If A is an affine space modeled on V, then the group of affine transformations $\mathrm{Aff}(A)$ is a Lie group, and fits into a short exact sequence of Lie groups

$$1 \longrightarrow V \longrightarrow \mathrm{Aff}(A) \longrightarrow \mathrm{GL}(V) \longrightarrow 0.$$

Similarly, if E is a Euclidean space modeled on V, the Euclidean transformations $\mathrm{Euc}(E)$ form a Lie group. If $A = \mathbb{A}^n$, $\mathrm{Aff}(\mathbb{A}^n)$ is denoted Aff_n, and similarly $\mathrm{Euc}(\mathbb{E}^n)$ is called Euc_n. ◀

Observe that on any manifold X, $\mathcal{X}(X)$ is a Lie algebra under Lie bracket, but it's infinite-dimensional, which makes life hairy.

Lie groups have extra geometric structure: the identity is a distiguished point, and there are distinguished symmetries given by left and right multiplication and conjugation: for each $g \in G$, let

- $L_g \colon G \to G$ send $x \mapsto gx$,
- $R_g \colon G \to G$ send $x \mapsto xg$, and
- $A_g \colon G \to G$ send $x \mapsto gxg^{-1}$.

Definition 9.9.

- Let $\xi \in \mathcal{X}(G)$ be a vector field. Then, ξ is *left-invariant* if $(L_g)_*\xi = \xi$ for all $g \in G$.
- Let $\alpha \in \Omega^\bullet_G$ be a differential form. Then, α is *right-invariant* if $(L_g)^*\alpha = \alpha$ for all $g \in G$.

In the same way, one can define left-invariance of Riemannian metrics, etc. Replacing L_g with R_g defines *right-invariance* of vector fields, forms, etc. And requiring invariance under any two of L_g, R_g, and A_g forces invariance under the third, and this is called *bi-invarinace*.

Since the differential of L_g is an isomorphism (since $L_{g^{-1}}$ is an inverse), then G acts on G transitively by left multiplication. This means in particular the dimension of G is constant (on arbitrary manifolds, such as $S^1 \amalg S^2$, the dimension is only locally constant).

Definition 9.10. Let G be a Lie group. Then, its Lie algebra $\mathfrak{g} \subset \mathcal{X}(G)$ is the subspace of left-invariant vector fields on G.

Evaluation at the identity defines a linear map $\mathrm{ev}_e \colon \mathfrak{g} \to T_eG$, sending $\xi \mapsto \xi|_e$. This map is an isomorphism, so $\dim \mathfrak{g} = \dim G$.

Proposition 9.11. $\mathfrak{g}$ *is closed under* $[–,–]$.

Proof. Let $\xi, \eta \in \mathfrak{g}$ and $g \in G$. Since ξ is left-invariant, it's L_g-related to itself, and similarly for η. We showed in Proposition 7.13 that if ξ and ξ' are related under ψ and η and η' are related under ψ, then so are $[\xi, \eta]$ and $[\xi', \eta']$. Thus, for our ξ and η, $[\xi, \eta]$ is L_g-related to itself for all g, hence is L_g-invariant. ⊠

In a very weak sense, the bracket is dual to d, so let's see what the corresponding statement is for differential forms.

Proposition 9.12. *The subspace $(\Omega_G^\bullet)^G$ of left-invariant forms is closed under* d.

Proof. Let $\omega \in (\Omega_G^\bullet)^G$, i.e. $L_g^*\omega = \omega$ for all $g \in G$. Then, by the commutation relations,

$$L_g^* \mathrm{d}\omega = \mathrm{d}L_g^*\omega = \mathrm{d}\omega.$$ ⊠

Observe that evaluation at the identity defines a map

$$\mathrm{ev}_e \colon (\Omega_G^\bullet)^G \longrightarrow \Lambda^\bullet T_e^* G = \Lambda^\bullet \mathfrak{g}^*.$$

Thus one obtains a complex

$$0 \longrightarrow \mathfrak{g} \xrightarrow{\mathrm{d}} \Lambda^2\mathfrak{g}^* \xrightarrow{\mathrm{d}} \Lambda^3\mathfrak{g}^* \longrightarrow \cdots \tag{9.13}$$

The map $\mathrm{d} \colon \mathfrak{g} \to \Lambda^2\mathfrak{g}^*$ can be computed with Corollary 9.5 to be the map

$$\alpha \longmapsto (\xi, \eta \longmapsto -\alpha([\xi, \eta])). \tag{9.14}$$

But this makes sense for any abstract Lie algebra, whether or not it came from G, and so one could define the complex (9.13) through the map (9.14). In this way one can talk about Lie algebra cohomology.

Suppose that $\xi_1, \dots, \xi_n$ is a basis for $\mathfrak{g}$, so that there exist $C_{jk}^i \in \mathbb{R}$ such that

$$[\xi_j, \xi_k] = C_{jk}^i \xi_i.$$

These are called the *structure constants* for $\mathfrak{g}$ in this basis. Skew-symmetry implies that $C_{jk}^i + C_{kj}^i = 0$, and there's a similar formula that comes from the Jacobi identity.

Exercise 9.15. Suppose that $\theta^1, \dots, \theta^n$ is the dual basis to $\xi_1, \dots, \xi_n$ for $\mathfrak{g}^* \subset \Omega_G^1$. Show that

$$\mathrm{d}\theta^i + \frac{1}{2} C_{jk}^i \theta^j \wedge \theta^k = 0.$$

If $\theta := \theta^i \xi_i \in \Omega_G^1(\mathfrak{g})$, then θ does not depend on the choice of basis, coming from $\mathfrak{g}^* \otimes \mathfrak{g} \cong \mathrm{End}(\mathfrak{g})$ (as vector spaces): θ maps to $\mathrm{id}_{\mathfrak{g}}$. This θ is a canonical 1-form that restricts to the form corresponding to the identity at each point; θ is called the *Maurer-Cartan 1-form*.

Proposition 9.16. *θ satisfies the* Maurer-Cartan equation

$$\mathrm{d}\theta + \frac{1}{2}\theta \wedge \theta = 0. \tag{9.17}$$

Lecture 10.

The Maurer-Cartan form: 2/16/17

Let's continue where we left off, letting G be a Lie group. There are a lot of nice facts about Lie groups, some of which are homework problems (e.g. using the inverse function theorem to prove that if multiplication is smooth, inversion is automatically smooth). The Lie algebra of G, denoted $\mathfrak{g}$, is the Lie algebra of left-invariant vector fields on G. We also defined a $\mathfrak{g}$-valued 1-form $\theta \in \Omega_G^1(\mathfrak{g})$, called the Mauer-Cartan 1-form.

If G is a Lie group, the path component of the identity G_e is a subgroup: for any $g, h \in G_e$, there's a path through the identity to g and one from the identity to h, so since multiplication is continuous, we get a path from g to gh. Moreover, all components are diffeomorphic: left translation $L_g : G_e \xrightarrow{\cong} g \cdot G_e$ defines a diffeomorphism from G_e to the component of G containing g. Many Lie groups that we care about aren't connected, e.g. the orthogonal groups O_n and the general linear groups $\mathrm{GL}_n(\mathbb{R})$, as well as any nontrivial finite group.

The Maurer-Cartan equation (9.17) encodes a lot of structure about G. For example, if $\xi_1, \dots, \xi_n$ is a basis of $\mathfrak{g}$ and $\theta = \theta^i \xi_i$, so $\theta^i \in \Omega^1_G$ is an ordinary 1-form, then

$$[\theta \wedge \theta] = [\theta^i \xi_i \wedge \theta^j \xi_j] = \theta^i \wedge \theta^j [\xi_i, \xi_j].$$

Proof of Proposition 9.16. We'll compute at $x \in G$, evaluating on $\xi_x, \eta_x \in T_x G$. We can extend them to left-invariant vector fields ξ, η on G; since the Maurer-Cartan equation is tensorial, its truth or falsity doesn't depend on the choice of extension. Then,

$$\begin{aligned}
\mathrm{d}\theta_x[\xi_x, \eta_x] &= \xi_x \cdot \theta(\eta) - \eta_x \cdot \theta(\xi) - \theta_x([\xi, \eta]_x) \\
&= -\theta_x([\xi, \eta]_x) = -[\xi, \eta]_x. \\
\frac{1}{2}[\theta \wedge \theta]_x(\xi_x, \eta_x) &= \frac{1}{2}([\theta_x(\xi_x), \theta_x(\eta_x)] - [\theta_x(\eta_x), \theta_x(\xi_x)]) \\
&= [\theta_x(\xi_x), \theta_x(\eta_x)] \\
&= [\xi, \eta_x].
\end{aligned}$$
⊠

It's locally possible to get from the Lie algebra to the Lie group.

Definition 10.1. If $\xi \in \mathfrak{g}$, then there's a unique integral curve $\gamma_\xi(t)$ such that $\gamma_\xi(0) = e$ and $\cdot\gamma_\xi(0) = \xi$. Define $\exp\colon \mathfrak{g} \to G$ by

$$\exp \xi := \gamma_\xi(1).$$

That is, flow along the curve in the direction of ξ for time 1. In particular, $\gamma_\xi(t) = \exp(t\xi)$, which will also be written $e^{t\xi}$. You can check that the integral curve with initial position $x \in G$ is $t \mapsto xe^{t\xi}$, so the flow φ_t generated by ξ is $\varphi_t = R_{e^{t\xi}}$, i.e. right translation by $\exp(t\xi)$.

A lot of Lie groups arise as matrix groups, in which the exponential map is just the matrix exponential.

Example 10.2. The *general linear group* $\mathrm{GL}_n(\mathbb{R})$ is the group of $n \times n$ invertible matrices. This is an open condition in all $n \times n$ matrices, since it's asking the determinant to be nonzero, so this is a Lie group. Inclusion defines a canonical function $g\colon G \hookrightarrow M_n(\mathbb{R})$. ◀

Since $\mathrm{GL}_n(\mathbb{R})$ is an open subset of the vector space $M_n(\mathbb{R})$, then its Lie algebra $\mathfrak{gl}_n(\mathbb{R})$ can be canonically identified with $M_n(\mathbb{R})$.

Theorem 10.3. *For $G = \mathrm{GL}_n(\mathbb{R})$,*

(1) *the Lie bracket on $\mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R})$ is $[A, B] = AB - BA$,*
(2) *the exponential is the matrix exponential $\exp\colon M_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ defined by*

$$e^A = I + A + \frac{A^2}{2!} + \cdots,$$

(3) *and the Maurer-Cartan form is*

$$\theta = g^{-1}\,\mathrm{d}g.$$

As most Lie groups we think about are matrix groups, hence subgroups of $\mathrm{GL}_n(\mathbb{R})$ for some n, this makes all of these notions nicely concrete, and Theorem 10.3 holds for them too.

Proof sketch of Theorem 10.3. For (2), first prove that for an $A \in M_n(\mathbb{R})$, the function

$$f(t) := I + tA + \frac{(tA)^2}{2!} + \cdots + \frac{(tA)^N}{N!} + \cdots$$

converges, and then differentiate term-by-term to show that $\dot{f}(t) = f(t) \cdot A$, which is what we needed.

For (1), the flow generated by ξ_A is $\varphi_t := R_{e^{tA}}$ and for ξ_B it's $\psi_s := R_{e^{sB}}$. Thus,

$$\begin{aligned}
[\xi_A, \xi_B]_I &= \left.\frac{\partial^2}{\partial s \partial t}\right|_{s,t=0} \psi_{-s}\varphi_{-t}\psi_s\varphi_t(I) \\
&= \left.\frac{\partial^2}{\partial s \partial t}\right|_{s,t=0} e^{tA}e^{sB}e^{-tA}e^{-sB} \\
&= \left.\frac{\partial^2}{\partial s \partial t}\right|_{s,t=0} (I+tA)(I+sB)(I-tA)(I-sB) \\
&= AB - BA.
\end{aligned}$$

The calculation of the Maurer-Cartan form comes directly from its definition. ⊠

The Maurer-Cartan equation $\mathrm{d}\theta + \theta \wedge \theta = 0$ turns into matrix multiplication for G a matrix group; in particular, if θ^i_j denotes the coefficients of the matrix θ in some basis, then matrix multiplication in coordinates implies that

$$\mathrm{d}\theta^i_j + \theta^i_k \wedge \theta^k_j = 0.$$

This will be important to us.

Example 10.4. Let's compute the Maurer-Cartan form for Aff_n. The trick is to embed $\mathrm{Aff}_n \hookrightarrow \mathrm{GL}_{n+1}(\mathbb{R})$ as follows: any affine transformation is of the form $x \mapsto Mx + \xi$ for some $M \in \mathrm{GL}_n(\mathbb{R})$ and $\xi \in \mathbb{R}^n$, so it can be represented in $\mathrm{GL}_{n+1}(\mathbb{R})$ as the block matrix

$$\begin{pmatrix} M & \xi \\ 0 & 1 \end{pmatrix},$$

which acts on $x = (x^1, \dots, x^n, 1)$. In other words, looking at the line $\{x^{n+1} = 1\}$ in $\mathbb{R}^{n+1}$ recovers $\mathbb{A}^n$ and Aff_n.

Therefore the Lie algebra also embeds: $\mathfrak{aff}_n \hookrightarrow \mathfrak{gl}_{n+1}(\mathbb{R}) = M_{n+1}(\mathbb{R})$: the affine Lie algebra consists of transformations $x \mapsto Mx + \xi$ where M need not be invertible, and this is sent to the matrix

$$\begin{pmatrix} M & \xi \\ 0 & 0 \end{pmatrix}.$$

With this embedding, the Maurer-Cartan form is

$$\begin{pmatrix} \theta^i_j & \theta^i \\ 0 & 0 \end{pmatrix},$$

and the equations of matrix multiplication inform us that

$$\begin{aligned}
\mathrm{d}\theta^i + \theta^i_j \wedge \theta^j &= 0 \\
\mathrm{d}\theta^i_j + \theta^i_k \wedge \theta^k_j &= 0.
\end{aligned}$$ ◀

Example 10.5. For curvd space, we care about the orthogonal group $\mathrm{O}_n \hookrightarrow M_n(\mathbb{R})$ of orthogonal $n \times n$ matrices, those matrices M such that $M^{\mathrm{T}}M = I$. The Lie algebra $\mathfrak{o}_n$ is the vector space of skew-symmetric matrices, i.e. $M^{\mathrm{T}} + M = 0$. In this case, the Maurer-Cartan form satisfies

$$\theta^i_j + \theta^j_i = 0.$$

◀

Similarly, we can embed $\mathrm{Euc}_n \hookrightarrow M_{n+1}(\mathbb{R})$ through the embedding $\mathrm{Euc}_n \hookrightarrow \mathrm{Aff}_n \hookrightarrow \mathrm{GL}_{n+1}(\mathbb{R}) \hookrightarrow M_{n+1}(\mathbb{R})$, and obtain its Maurer-Cartan formula in that way. The same is true for the *conformal group* $\mathrm{CO}_n \subset \mathrm{GL}_n(\mathbb{R})$ of matrices M such that $\langle M\xi, M\eta\rangle = c\langle \xi, \eta\rangle$ for some $c > 0$ and all $\xi, \eta \in \mathbb{R}^n$. You can play the same game for any kind of geometric structure: symplectic geometry with the symplectic group, spin geometry with the spin group, and so forth.

The adjoint action. Let $g \in G$. Then, the conjugation action $A_g = L_g \circ R_{g^{-1}} = R_{g^{-1}} \circ L_g$ sends $x \mapsto gxg^{-1}$. Therefore A defines a map $G \to \mathrm{Aut}(G)$.

Definition 10.6. Identifying $\mathfrak{g} \cong T_eG$, the *adjoint action* $\mathrm{Ad}_g \in \mathrm{End}(\mathfrak{g})$ is defined by $\mathrm{Ad}_g = \mathrm{d}(A_g)_e$.

Lemma 10.7.

(1) *The adjoint action preserves the Lie algebra, hence is a Lie algebra endomorphism.*
(2) *The differential of* $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g})$ *(sending $g \mapsto A_g$) at the identity is*

$$d\mathrm{Ad}_g|_e(\xi) = [\xi, -].$$

The Maurer-Cartan form behaves nicely under the left and right actions.

Proposition 10.8. *In $\Omega^1_G(\mathfrak{g})$, $L_g^*\theta = \theta$, but $R_g^*\theta = \mathrm{Ad}_{g^{-1}}\theta$.*

These equations are simple, but will come up again.

Proof. Let $x \in G$, so that if $\xi_x \in T_eG$,

$$(R_g^*\theta)_x(\xi_x) = \theta_{xg}(R_{g*}\xi_x).$$

Let ξ be a left-invariant extension of $\xi_x = \theta_x(\xi_x)$. Then, at the identity,

$$\begin{aligned}(R_g^*\theta)_x(\xi_x) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \theta_{xg}\left(xe^{t\xi}g\right) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (xg)^{-1}\left(xe^{t\xi}g\right) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g^{-1}e^{t\xi}g \\ &= \mathrm{Ad}_{g^{-1}}\xi.\end{aligned}$$

⊠

We've met the Maurer-Cartan form on groups, but it will also arise on G-torsors. This is an instance of the notion that left-invariant objects on G extend to right G-torsors, and right-invariant objects extend to left torsors.

Recall that a *right G-torsor G* is a manifold with a simply transitive right action $T \times G \to T$, sending $t, g \mapsto t \cdot g$, so the map $T \times G \to T \times T$ sending $t, g \mapsto t, t \cdot g$ is a diffeomorphism.

Torsors are like groups with multiplication but no preferred origin. Examples include affine space $\mathbb{A}^n$, a torsor over $\mathbb{R}^n$; and $\mathscr{B}(V)$, the space of bases for a vector space V, which is a $\mathrm{GL}(V)$-torsor.

If $t \in T$, $\varphi_t\colon GT \to T$ sending $g \mapsto t \cdot g$ is a diffeomorphism, in fact an isomorphism of right G-torsors. If I have two of these, φ_{t_0} and φ_{t_1}, then the map $\varphi_{t_1}^{-1} \circ \varphi_{t_0}\colon G \to G$ is a right-invariant diffeomorphism, and in fact is left multiplication by some $h \in G$, such that $t_1(hg) = t_0 g$ for all $g \in G$; in particular, $t_0 = t_1 h$. This means that the Maurer-Cartan form, which is left-invariant, can be pulled back along some (well, any) φ_t^{-1} and therefore makes sense on a right G-torsor.

This concludes our crash course on Lie groups; we'll see this stuff again later.

Plane curves. Let E be a Euclidean plane modeled on an inner product space V.[9] Let $\mathscr{B}_{\mathrm{O}}(E)$ denote the space of affine isometries $\alpha\colon \mathbb{E}^2 \to E$. These are ways of choosing an origin and orthonormal coordinates for E. That is, α can be identified with pairs (p, b) where $p \in E$ and b is a basis of V. Forgetting the basis defines a map $\pi\colon \mathscr{B}_{\mathrm{O}}(E) \to E$, called the *frame bundle*; it is a right Euc_2-torsor, where Euc_2 acts by changing the coordinates of b.

Therefore all left-invariant information on Euc_2 comes over to $\mathscr{B}_{\mathrm{O}}(E)$. In particular, there are Maurer-Cartan forms θ^1 and θ^2 with $\theta^1_2 = -\theta^2_1$, coming from the same equation on Euc_2. This will be useful: differential forms make computations particularly easy compared to coordinate systems and vector fields (though those are not bad). Figuring out what this form actually calculates on a pair (p, b) is the first step towards understanding the Riemann curvature tensor: θ^i is the component of translation in the direction e^i.

Let $\gamma\colon (a, b) \to E$ be an embedded curve, which lifts to a map $\widetilde{\gamma}$ of frames, so e_1 is normal and e_2 is tangent. What is the pullback $\gamma^*\theta^i$ and $\gamma^*\theta^i_j$? Since there's no translation in direction e^1, $\gamma^*\theta^1 = 0$, and

[9] Some of what we do here will work for more general Euclidean spaces.

consequently $\gamma^*\theta^1 = \mathrm{d}t$. The pullback $\gamma^*\theta^1_2 = -k\,\mathrm{d}t$. So the Maurer-Cartan forms encode curvature on the torsor of frames, even in the simple situation of a plane curve.

Lecture 11.

Characterizing the Maurer-Cartan form: 2/21/17

"General confusion reigns in the land here..."

Recall that the Maurer-Cartan form for a Lie group G is the canonical 1-form $\theta \in \Omega^1_G(\mathfrak{g})$ that at f pulls back a vector field ξ to $L_{g^{-1}}(\xi_g)$, i.e. produces the unique left-invariant vector field extending ξ_g. This uses the fact that evaluation defines an isomorphsim $\mathfrak{g} \to T_xG$ for each $x \in G$.

Since $\mathrm{GL}_n(\mathbb{R})$ is an open submanifold of $M_n(\mathbb{R})$, then $T_A\mathrm{GL}_n(\mathbb{R})$ is canonically identified with $M_n(\mathbb{R})$ again. In particular, this means a vector field is a function $\xi\colon \mathrm{GL}_n(\mathbb{R}) \to M_n(\mathbb{R})$, and ξ is left-invariant means $\xi_{AB} = A \cdot \xi_B$.

Let $g\colon \mathrm{GL}_n(\mathbb{R}) \to M_n(\mathbb{R})$ be the embedding, so $\mathrm{d}g \in \Omega^1_{\mathrm{GL}_n(\mathbb{R})}(M_n\mathbb{R})$, and $g^{-1}\,\mathrm{d}g$ is also a $\mathfrak{gl}_n(\mathbb{R})$-valued 1-form. It operates on a matrix $\dot{A}$ as $g^{-1}\,\mathrm{d}g_A(\dot{A}) = A^{-1}\dot{A}$, which is indeed pulling back to the identity, then plugging in the vector $(\dot{A})$ in question.

This applies just as well to any matrix group, i.e. a group with an embedding $G \hookrightarrow M_n(\mathbb{R})$. This is equivalent to the data of a faithful (real) representation of G.

Remark. Not every Lie group is a matrix group. For example, the double cover of $\mathrm{SL}_2(\mathbb{R})$ isn't. This is most easily shown using representation theory. ⊲

Some of the following proposition was in Proposition 10.8, but not all of it. In any case, checking these tells you how the Maurer-Cartan form behaves under various pullbacks.

Proposition 11.1. *Let $g \in G$.*

(1) $L_g^*\theta = \theta$ *and* $R_g^*\theta = \mathrm{Ad}_{g^{-1}}\theta$.
(2) *If $i\colon G \to G$ is the inversion map $g \mapsto g^{-1}$, $(i^*\theta)_g = -\mathrm{Ad}_g\theta_g$.*
(3) *If $m\colon G \times G \to G$ is the multiplication map $g, h \mapsto gh$, then $m^*\theta_{(g,h)} = \mathrm{Ad}_{h^{-1}}\pi_1^*\theta + \pi_2^*\theta$, where π_i is the projection onto the i^{th} component.*

Proof of parts (2) and (3) for matrix groups. When G is a matrix group, $\theta = g^{-1}\,\mathrm{d}g$, so we can calculate:

$$i^*\theta = (g^{-1})^{-1}\,\mathrm{d}(g^{-1}) = gg^{-1}\,\mathrm{d}gg^{-1} = -\mathrm{Ad}_g(g^{-1}\,\mathrm{d}g).$$

For part (3), it comes from the calculation

$$(g_1g_2)^{-1}\,\mathrm{d}(g_1g_2) = g_2^{-1}g_1^{-1}(\,\mathrm{d}g_1g_2 + g_1\,\mathrm{d}g_2) = g_2^{-1}(g_1^{-1}\,\mathrm{d}g_1)g_2 + g_2^{-1}\,\mathrm{d}g_2. \qquad \boxtimes$$

We're going to return to curves and surfaces, and next time generalize to higher dimensions, but before we do that, we need to discuss a theorem that we'll return to several times.

Theorem 11.2. *Let Y be a manifold, G be a Lie group, and $\theta_Y \in \Omega^1_Y(\mathfrak{g})$.*

(1) *If Y is connected and $F, F'\colon Y \to G$ are such that $F^*\theta = (F')^*\theta = \theta_Y$, then there exists a $g \in G$ such that $F' = L_g \circ F$.*
(2) *If*

$$\mathrm{d}\theta_Y + \frac{1}{2}[\theta_Y \wedge \theta_Y] = 0, \tag{11.3}$$

then for any $y_0 \in Y$ and $g_0 \in G$, there's a neighborhood $U \subset Y$ containing y_0 and an $F\colon U \to G$ such that $F(y_0) = g_0$ and $F^*\theta = \theta_Y|_U$, and if U is connected F is unique.

(3) *If Y is simply connected and (11.3) holds, then there's a unique $F\colon Y \to G$ such that $F(y_0) = g_0$ and $F^*\theta = \theta_Y$.*

Example 11.4.

(1) Let $Y = \mathbb{R}$ with a t-coordinate and $\xi \in \mathfrak{g}$, and let $\theta_Y = \xi\,\mathrm{d}t \in \Omega^1_{\mathbb{R}}(\mathfrak{g})$. Then, Theorem 11.2 shows there's a unique $F\colon \mathbb{R} \to G$ such that $F(0) = e$ and $F^*\theta = \xi\,\mathrm{d}t$. This is the exponential map $F(t) = e^{t\xi}$ (which requires a check).

(2) Theorem 11.2 can be used to lift some morphisms of Lie algebras to Lie groups. Let G' be a simply connected Lie group and $\dot\phi\colon \mathfrak{g}' \to \mathfrak{g}$ be a Lie algebra homomorphism. Let $\theta' \in \Omega^1_{G'}(\mathfrak{g}')$ be the Maurer-Cartain form and $\theta_{G'} := \dot\phi\theta'$. Then, (11.3) holds:

$$\begin{aligned} \mathrm{d}\theta_{G'} + \frac{1}{2}[\theta_{G'} \wedge \theta_{G'}] &= \mathrm{d}\dot\phi\theta' + \frac{1}{2}[\dot\phi\theta' \wedge \dot\phi\theta'] \\ &= \dot\phi\,\mathrm{d}\theta' + \frac{1}{2}\dot\phi[\theta' \wedge \theta'] \\ &= \dot\phi\left(\mathrm{d}\theta' + \frac{1}{2}[\theta' \wedge \theta']\right) \\ &= 0. \end{aligned}$$

By Theorem 11.2 we conclude that there's a unique $\phi\colon G \to G$ such that $\phi(e') = e$ and $\phi^*\theta_{G'} = \theta$. That G' is simply connected is essential. ◄

Proof of Theorem 11.2. Part (1) is a calculation. If G is a matrix group,

$$\begin{aligned} (F'F^{-1})^*\theta = (F'F^{-1})^{-1}\,\mathrm{d}(F'F^{-1}) &= (F(F')^{-1})\Big(\mathrm{d}F'F^{-1} - F'F^{-1}\,\mathrm{d}FF^{-1}\Big) \\ &= F((F')^{-1}\,\mathrm{d}F')F^{-1} - F(F^{-1}\,\mathrm{d}F)F^{-1} \\ &= F\Big((F')^{-1}\,\mathrm{d}F' - F^{-1}\,\mathrm{d}F\Big)F^{-1} \\ &= F\big((F')^*\theta - F^*\theta\big)F^{-1} = 0. \end{aligned}$$

For a general group, one must decompose $F'F^{-1}$ as the composition

$$Y \xrightarrow{(F',F)} G \times G \xrightarrow{(\mathrm{id},i)} G \times G \xrightarrow{m} G.$$

For part (2) we use an old tactic: using the graph of a function to pass between sets and functions. Let π_i be the projection onto the i^{th} factor out of $Y \times G$. Let $\xi_1, \dots, \xi_n$ be a basis of $\mathfrak{g}$, $\theta_Y = \theta^i_Y\xi_i$, and $\theta = \theta^i\xi_i$. Then, let $\mathscr{I} \subset \Omega^\bullet_{Y\times G}$ be the ideal generated by $\pi_1^*\theta^i_Y - \pi_2^*\theta^i_Y$. This is the vanishing ideal for a distribution $E \subset T(Y \times G) \to Y \times G$ of codimension n, so $\dim E = \dim Y$.

We'll check that $\mathrm{d}(\pi_1^*\theta_Y - \pi_2^*\theta)$ is an ideal. For ease of writing, this will be written $\mathrm{d}(\theta_Y - \theta)$, with the projections implicit. Since

$$\mathrm{d}\theta_Y - \mathrm{d}\theta = \mathrm{d}\theta_Y + \frac{1}{2}[\theta \wedge \theta]$$

and

$$\frac{1}{2}[(\theta - \theta_Y) \wedge (\theta + \theta_Y)] = \frac{1}{2}[\theta \wedge \theta] - \frac{1}{2}[\theta_Y \wedge \theta_Y],$$

then $\mathrm{d}\theta_Y - \mathrm{d}\theta \in \mathscr{I}$ iff (11.3) holds. Then, Theorem 8.4 guarantees the existence of a graph locally, proving (2).

For part (3), first observe that E is left-invariant under G:[10] if $g \in G$, then $L_{g*}E_{(y,h)} = E_{(y,gh)}$. Equivalently, $L_g^*\mathscr{I} = \mathscr{I}$, which is true because $L_g^*\theta = \theta$. The global Frobenius theorem (Theorem 8.9) says that around each $(y,h) \in Y \times G$ there's a neighborhood $U \times V \subset Y \times G$ containing (y,h) and local coordinates such that $E|_{U\times V}$ is a product of a distribution that's constant on U and one that's constant on V.

By left-invariance, we can assume $V = G$. Let Γ be the leaf of the foliation through (y_0, g_0); then, $\Gamma \cap (U \times G)$ has at most countably many components which are horizontal. It follows that π_1 is both open and closed, so $\mathrm{Im}(\pi_1)$ is both open and closed, hence connected; since Y is connected, then $\mathrm{Im}(\pi_1) = Y$. Furthermore, the local form (that U is evenly covered) shows that $\pi_1|_\Gamma$ is a covering map. Since Y is simply connected and Γ is connected, then $\pi_1|_\Gamma$ has to be a diffeomorphism. Therefore we have the graph, and recover the function as $F = \pi_2 \circ (\pi_1|_\Gamma)^{-1}$. ☒

The key here (aside from the local Frobenius theorem) is the left-invariance, which is what guarantees the leaves of the foliation can't do anything funny.

[10] It's also right-invariant, but we don't need that.

This fancy technology of differential forms and Lie groups takes us a long way quickly even just in the case of curves and surfaces.

Let E be a Euclidean plane, modeled on a two-dimensional inner product space V, and let $\mathscr{B}_{\mathrm{O}}(E)$ denote the space of isometries $\mathbb{E}^2 \to E$, i.e. pairs (p, b) with $p \in E$ and $b \colon \mathbb{R}^2 \to V$ a basis. As we've discussed before, $\mathscr{B})_{\mathrm{O}}(E)$ is a right Euc_2-torsor.

In this context and in a basis, the Maurer-Cartan form is a matrix

$$\begin{pmatrix} 0 & \theta_2^1 & \theta^1 \\ \theta_1^2 & 0 & \theta^2 \\ 0 & 0 & 0 \end{pmatrix}$$

Here $\theta_2^1 = -\theta_1^2$ and $\theta^i, \theta_j^i \in \Omega^1_{\mathscr{B}_{\mathrm{O}}(E)}$. The Maurer-Cartan equation imposes some important relations between these forms:

$$\mathrm{d}\theta^1 + \theta_2^1 \wedge \theta^2 = 0 \tag{11.5a}$$

$$\mathrm{d}\theta^2 + \theta_1^2 \wedge \theta^1 = 0 \tag{11.5b}$$

$$\mathrm{d}\theta_1^2 = 0. \tag{11.5c}$$

These are also called the *Maurer-Cartan equations*, and we're about to get a lot more familiar with them. In curved space, the matrix form of θ will be different, and things will be different.

Let $i \colon (a, b) \hookrightarrow E$ be a co-oriented curve, and lift it across $\pi \colon \mathscr{B}_{\mathrm{O}}(E) \to E$ to a map $\widetilde{\imath}$ such that e_2 is the oriented unit normal. This is a choice of what happens to e_1.

In general, a short exact sequence of vector bundles $0 \to V' \to V \to V'' \to 0$ splits, so one of V' is a quotient and the other is a subspace. In the absence of orientations, which is which doesn't really matter, but if V, V', and V'' are oriented vector spaces, then the induced orientations on $V \cong V' \oplus V''$ force V' to be the quotient and V'' to be the subspace. Thus, the tangent-normal sequence of oriented vector bundles on C is

$$0 \longrightarrow TC \longrightarrow TE|_C \longrightarrow \nu \longrightarrow 0.$$

Then, we can calculate components of the Maurer-Cartan form: $\widetilde{\imath}^*\theta^1 = \mathrm{d}t$ and $\widetilde{\imath}^*\theta^2 = 0$, based on how e^1 and e^2 change with time. Then, $\widetilde{\imath}^*\theta_1^2$ measures the rate of turning of e_1 in the direction of e_2, which is precisely $k\,\mathrm{d}t$.

We can also revisit the problem of prescribing curvature: given $Y = (a, b)$ and a function $k \colon Y \to \mathbb{R}$, is there an immersed curve with curvature k? We saw in Corollary 3.9 that this is (locally) possible and the curve is unique up to a Euclidean motion. Here's another proof.

Another proof of Corollary 3.9. Define $\theta_Y \in \Omega^1_Y(\mathfrak{euc}_2)$ to satisfy the identities we just calculated:

$$\theta_Y := \begin{pmatrix} 0 & k\,\mathrm{d}t & \mathrm{d}t \\ -k\,\mathrm{d}t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

By Theorem 11.2 we get a map $\widetilde{\imath} \colon Y \to \mathscr{B}_{\mathrm{O}}(E)$ such that $\widetilde{\imath}\theta = \theta_Y$ and if we fix an initial condition, $\widetilde{\imath}$ is unique. ⊠

For a co-oriented surface Σ embedded in Euclidean 3-space E, things work differently: instead of a discrete set of choices of lift of the embedding to a map $\Sigma \to \mathscr{B}_{\mathrm{O}}(E)$, there's an O_2 worth of them, where O_2 acts as the automorphisms of a framing in which e_3 is fixed.

Now the Maurer-Cartan form contain six pieces of information: $\theta^1, \theta^2, \theta^3, \theta_2^1, \theta_3^1, \theta_3^2$. Here $\theta^3 = 0$ and (θ^1, θ^2) is a local orthonormal (co)framing of E, which is called a *moving frame* (or *repère mobile* in French). Elie Cartan discovered moving frames and used them to make calculations.

The Maurer-Cartan equations take the form

$$\mathrm{d}\theta^i + \theta_j^i \wedge \theta^j = 0$$

$$\mathrm{d}\theta_j^i + \theta_k^i \wedge \theta_j^k = 0.$$

Next time, we'll write these out explicitly, and they will cause results like Gauss' theorema egregium to fall out in the blink of an eye! We'll also see the second fundamental form and the Gauss curvature and find some relations between them. This leads to a version of the prescribed curvature problem for surfaces, which involves solving a PDE instead of an ODE.

Lecture 12.

Applications to immersed surfaces: 2/23/17

Let E be Euclidean 3-space modeled on a 3-dimensional inner product space V. Then, $\mathscr{B}_{\mathrm{O}}(E)$, the space of isometries $\mathbb{E}^3 \to E$, is a right Euc_3-torsor, and the map $\pi\colon \mathscr{B}_{\mathrm{O}}(E) \to E$ realizes it as a principal O_3-bundle: the fibers are acted on by the isometries of $\mathbb{R}^3$. This describes a section to the short exact sequence

$$1 \longrightarrow \mathbb{R}^3 \longrightarrow \mathrm{Euc}_3 \longrightarrow \mathrm{O}_3 \longrightarrow 1.$$

The Maurer-Cartan form on $\mathscr{B}_{\mathrm{O}}(E)$ is determined by the 1-forms θ^1, θ^2, θ^3, θ^2_1, θ^3_1, and θ^3_2 such that $\theta^i_j = -\theta^j_i$ and the Maurer-Cartan equations are satisfied:

$$\mathrm{d}\theta^1 + \theta^1_2 \wedge \theta^2 + \theta^1_3 \wedge \theta^3 = 0 \tag{12.1a}$$

$$\mathrm{d}\theta^2 + \theta^2_1 \wedge \theta^1 + \theta^2_3 \wedge \theta^3 = 0 \tag{12.1b}$$

$$\mathrm{d}\theta^3 + \theta^3_1 \wedge \theta^1 + \theta^3_2 \wedge \theta^2 = 0 \tag{12.1c}$$

$$\mathrm{d}\theta^2_1 + \theta^2_3 \wedge \theta^3_1 = 0 \tag{12.1d}$$

$$\mathrm{d}\theta^3_1 + \theta^3_2 \wedge \theta^2_1 = 0 \tag{12.1e}$$

$$\mathrm{d}\theta^3_2 + \theta^3_1 \wedge \theta^1_2 = 0. \tag{12.1f}$$

Let $i\colon \Sigma \hookrightarrow E$ be an immersed surface, and choose a lift $\widetilde{\imath}\colon \Sigma \to \mathscr{B}_{\mathrm{O}}(E)$, an orthonormal frame on Σ. Let e_3 be the unit normal to Σ. We'll restrict the pieces of the Maurer-Cartan form to Σ via $\widetilde{\imath}$, though we'll leave the $\widetilde{\imath}^*$ out of the equation.

Suppose $\widetilde{\gamma}\colon (-\varepsilon,\varepsilon) \to \mathscr{B}_{\mathrm{O}}(E)$ is the lift of a curve $\gamma\colon (-\varepsilon,\varepsilon) \to E$ such that $\dot\gamma_0 = \xi$ and $\dot{\widetilde{\gamma}} = \widetilde{\xi}$. Then, $\theta^i(\widetilde{\xi})$ is the e_i-component of $\xi = \pi_*\widetilde{\xi}$, and $\theta^i_j(\widetilde{\xi}) = \langle \dot e_j(0), e_i\rangle$.

Let $U \subset \Sigma$ be a neighborhood. Then, on U, $\theta^3 = 0$ and $\{\theta^1, \theta^2\}$ is a basis for Ω^1_U. For $\mu, \nu \in \{1,2\}$, write $\theta^3_\mu = h_{\mu\nu}\theta^\nu$ for some $h_{\mu\nu}\colon U \to \mathbb{R}$, which defines a 2×2 matrix $h := (h_{\mu\nu})$.

Lemma 12.2. *$h_{12} = h_{21}$ and h is the second fundamental form in the basis $\{e_1, e_2\}$, i.e. $II = h_{\mu\nu}\theta^\mu \otimes \theta^\nu$.*

Proof. By (12.1c),

$$0 = h_{12}\theta^2 \wedge \theta^1 + h_{21}\theta^1 \wedge \theta^2 = (h_{12} - h_{21})\theta^2 \wedge \theta^1,$$

so $h_{21} - h_{12} = 0$.

Recall that in these coordinates, we have $e_3\colon U \to V$, and the shape operator is $-\mathrm{d}e_3\colon TU \to V$. Since $\theta^i_j(\widetilde{\xi}) = \langle \dot e_j(0), e_i\rangle$, then $-\mathrm{d}e_3 = -\theta^\mu_3 e_\mu = h_{\mu\nu}\theta^\nu e_\mu$, and $II(\xi_1, \xi_2) = \langle \xi_1, S(\xi_2)\rangle$, so h describes II. ⊠

This is part of a theme: once you write down what the Maurer-Cartan form actually is, everything falls out, and the objective is to recognize it before it falls past you.

Proposition 12.3. $\mathrm{d}\theta^2_1 = -K\theta^1 \wedge \theta^2$*, where K is the Gauss curvature.*

Proof. From (12.1d),

$$\begin{aligned} 0 &= \mathrm{d}\theta^2_1 - (h_{21}\theta^1 + h_{22}\theta^2) \wedge (h_{11}\theta^1 + h_{12}\theta^2) \\ &= \mathrm{d}\theta^2_1 + \underbrace{(h_{11}h_{22} - h_{12}h_{21})}_{\det(h)}\theta^1 \wedge \theta^2 \\ &= \mathrm{d}\theta^2_1 + K\theta^1 \wedge \theta^2. \end{aligned}$$

⊠

Proposition 12.4 (Gauss' *Theorema egregium*)**.** *θ^2_1 is determined by θ^1 and θ^2.*

Proof. Suppose $\theta_1^2 = a\theta^1 + b\theta^2$ for some $a, b \in \Omega_U^0$. By (12.1a) and (12.1b),

$$\mathrm{d}\theta^1 + a\theta^1 \wedge \theta^2 = 0$$
$$\mathrm{d}\theta^2 - b\theta^1 \wedge \theta^2 = 0.$$

This means that a and b are determined by computing $\mathrm{d}\theta^1$ and $\mathrm{d}\theta^2$. ⊠

The relation with the more conventional statement of Theorem 5.1 is that θ_1^2 is intrinsic, and therefore so is $\mathrm{d}\theta_1^2$, hence also K.

The last two equations, (12.1e) and (12.1f), called the *Codazzi-Mainardi equations*, haven't been used yet, but they are constraints on the first and second fundamental forms of an immersed surface. You can ask, given an abstract surface and choices of the first and second fundamental form, is there an immersion such that the induced metric produces the chosen first and second fundamental forms? This is the surface-level analogue of the prescription of curvature problem for plane curves. The fact that the Gauss curvature matches the second fundamental form forces a relation between the first and second fundamental form, and the derivative of the second fundamental form is constricted by the Codazzi-Mainardi equations.

Older proofs of this boil everything down to solutions of systems of partial differential equations, and the solutions exist because mixed partials commute. However, we've managed to take a more geometric viewpoint which encodes everything into symmetries of the Maurer-Cartan form.

Example 12.5. Suppose r, θ are local coordinates on a two-dimensional Riemannian manifold akin to polar coordinates, in that the metric is

$$\mathrm{d}s^2 = \mathrm{d}r^2 + G(r)\mathrm{d}\theta^2,$$

where G is some positive function. One can then show that $\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \rangle = 1$ and $\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta} \rangle = 0$, so these are always perpendicular. One can show that such a coordinate system exists locally around any point in any Riemannian surface, and an analogous theorem is true in higher dimension.

For example, on $\mathbb{E}^2$ (i.e. $\mathbb{R}^2$ with the standard Euclidean metric), one can choose (x, y) (so $G = 1$) or polar coordinates (r, θ) where the metric is $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2$. On the sphere (with the induced metric as the unit sphere in $\mathbb{E}^3$), we have spherical coordinates (ϕ, θ) and the metric is $\mathrm{d}s^2 = \mathrm{d}\phi^2 + \sin^2\phi\,\mathrm{d}\theta$.

We can compute the Gauss curvature K in terms of G. Namely, if $g(r)$ is such that $g^2 = G$, then $\theta^1 = \mathrm{d}r$ and $\theta^2 = g\,\mathrm{d}\theta$. Therefore $\mathrm{d}\theta^1 = 0$ and $\mathrm{d}\theta^2 = g'\,\mathrm{d}r \wedge \mathrm{d}\theta = (g'/g)\theta^1 \wedge \theta^2$. Thus, $\theta_1^2 = -(g'/g)\theta^2 = -g'\,\mathrm{d}\theta$ and $\mathrm{d}\theta_1^2 = -g''\,\mathrm{d}r \wedge \mathrm{d}\theta = -(g''/g)\theta^1 \wedge \theta^2$, so we conclude

$$K = -g''/g.$$

If you plug this into (x, y) or (r, θ) on $\mathbb{E}^2$, g'' vanishes, so the Gauss curvature is 0; for the sphere, the second derivative of $\sin\phi$ is $-\sin\phi$, so $K = 1$. Thus, we have a surface of constant flat curvature and one of constant positive curvature; negative curvature is missing from this list, but one can realize it using hyperbolic space, replacing $\sin^2\phi$ with $\sinh^2\phi$. ◀

In the next few lectures, we'll continue on to higher dimensions. Suppose X is an n-dimensional Riemannian manifold, and let $\pi\colon \mathscr{B}_{\mathrm{O}}(X) \to X$ be the bundle of pairs (x, b) with $x \in X$ and $b\colon \mathbb{R}^n \xrightarrow{\cong} T_xX$ an isometry. This means we've switched to an abstract, intrinsic story: one can set up the extrinsic story again, and there are a few differences, e.g. in higher dimensions there are extra normal directions.

Anyways, $\mathscr{B}_{\mathrm{O}}(X)$ is called the *bundle of orthonormal frames* of X, and has a free right O_n-action, and π is the quotient map. Therefore it's possible to construct the pieces of the Maurer-Cartan forms $\theta^1, \dots, \theta^n$ on $\mathscr{B}_{\mathrm{O}}(X)$ and θ_j^i from the structure equations. Then, the equations $\theta_j^i = -\theta_i^j$ and $\mathrm{d}\theta + \theta \wedge \theta = 0$. This will define the Levi-Civita equations for us.

The orthogonal group is not the only choice here: you could ask for bases for T_xX that preserve a $\mathrm{GL}_n(\mathbb{R})$ action, which is weaker; in this case, you get θ^i but not unique θ_i^j. If you ask for a complex structure on the tangent space, this leads to the notion of a complex structure and local holomorphic coordinates. It's possible to develop a general theory for these θ_j^i for general structure groups.

In any case, the existence *and* uniqueness in the case of Riemannian manifolds, which is the fundamental theorem of Riemannian geometry, is completely mysterious: except in a few cases, such as Kähler manifolds, where there are beautiful formulas, it's completely unclear *why* the unique connection compatible with the metric should exist.

In the general setting, we'll need a definition.

Definition 12.6. Let X be a smooth manifold and G be a Lie group. Then, a *principal G-bundle over X* is a manifold P together with a free right G-action and quotient map $\pi\colon P \to X$ such that π admits local smooth sections.

That π is a quotient means that for every $x \in X$, the fiber $P_x := \pi^{-1}(x)$ is an orbit of G, so for any $p_1, p_2 \in P_x$, there's a unique $g \in G$ such that $g \cdot p_1 = p_2$. The condition of local smooth sections means that for every $x \in X$, there's a neighborhood $U \subset X$ of x and a section $s\colon U \to P$ such that $\pi \circ s = \mathrm{id}_U$.

Intuitively, the local smooth sections criterion says that the fibers are "locally constant," and don't move too much if x doesn't.

Keep in mind that P is not the principal bundle: we need the data of the base X and the quotient π.

Example 12.7. Let $P = X \times G$, with the action on G by right multiplication and π projection onto the first component. This principal G-bundle is called the *trivial bundle*. ◂

Lemma 12.8. *If $\pi\colon P \to X$ is a principal G-bundle, then π admits* local trivializations. *That is, for any $x \in X$, there's a neighborhood U of x and an isomorphism $\pi^{-1}(U) \xrightarrow{\cong} U \times G$ that commutes with the projection back to U.*

Proof. Given x, choose a local section $s\colon U \to P$, and define $\varphi\colon U \times G \to P$ to send $y, g \mapsto s(y) \cdot g$. You can check φ is a diffeomorphism $U \times G \to P|_U := \pi^{-1}(U)$, but better is to show it commutes with the right G-actions, and therefore is an *isomorphism of principal G-bundles*. ⊠

We'll let $R_g\colon P \to P$ denote the action of a $g \in G$ on the principal G-bundle $P \to X$.

Proposition 12.9. *$\mathscr{B}_{\mathrm{O}}(X) \to X$ is a principal O_n-bundle.*

Proof. First, let X be any manifold and let $\pi\colon \mathscr{B}(X) \to X$ be the bundle of all frames, the pairs (x, b) such that $b\colon \mathbb{R}^n \to T_xX$ is a linear isomorphism. This is a principal $\mathrm{GL}_n(\mathbb{R})$-bundle — you should check that it's a manifold, e.g. by producing a chart $U \times \mathrm{GL}_n(\mathbb{R})$ for $\mathscr{B}(X)$ for every chart U, and use gluing on X and local sections to glue (there's more to show here). To obtain the local section near x, choose local coordinates $x^1, \dots, x^n$ neat x; then, the local section is given by $\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\}$.

Great, so how about orthonormal frames? If X has a Riemannian metric, then the orthonormal frames $\mathscr{B}_{\mathrm{O}}(X)$ form a submanifold of $\mathscr{B}(X)$, and the quotient by O_n is X, but we need to check that there's a local section. Given a local section of $\mathscr{B}(X)$, one can use the Gram-Schmidt process to smoothly deform it into a section of $\mathscr{B}_{\mathrm{O}}(X)$. ⊠

Next time we'll talk about connections in this context.

Lecture 13.

Principal G-bundles: 2/28/17

Recall that if X is a smooth manifold and G is a Lie group, then a principal G-bundle over X is a map $P \to X$ such that P is a smooth manifold equipped with a free right G-action, such that π is the quotient map, and π admits local sections.[11]

Local sections are maps $s\colon U \to P$ for a chart $U \subset X$ such that $\pi \circ s = \mathrm{id}$. This is equivalent to a local trivialization, a commutative diagram

$$\begin{tikzcd} U \times G \arrow[rr, "\varphi"] \arrow[dr, "\pi_1"'] & & P|_U \arrow[dl, "\pi|_U"] \\ & U. & \end{tikzcd}$$

Here, $\varphi(x, g) := s(x) \cdot g$.

[11] More on principal bundles can be found at http://www.ma.utexas.edu/users/dafr/M392C/Notes/lecture13.pdf.

If G is compact, then it suffices to specify its free action on P, as the quotient of a manifold by a free action of a compact Lie group is again a manifold. However, this is not true for G noncompact: let $\mathbb{R}$ act on the torus $\mathbb{R}^2/\mathbb{Z}^2$ by translation by $(1/2, a)$ where a is an irrational number. Then, the orbits are dense, and in fact the quotient space isn't even Hausdorff!

Example 13.1. The *trivial G-bundle* is $P = X \times G$, with G acting by right multiplication on G and trivially on X. Here π is projection onto the first factor. ◂

Definition 13.2. A *morphism of principal G-bundles* is a G-equivariant map $\varphi\colon P \to P'$ that commutes with projection to the base:

$$\begin{tikzcd} P \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & P' \arrow[dl, "\pi'"] \\ & X. & \end{tikzcd}$$

Local triviality means that principal G-bundles are examples of *fiber bundles* as defined by Norman Steenrod.

Example 13.3. In this class, we care the most about frame bundles, but there are lots of other examples.

(1) Let $G = \mathrm{GL}_n(\mathbb{R})$, so, as we discussed last time, $\mathscr{B}(X) \to X$ is a principal $\mathrm{GL}_n(\mathbb{R})$-bundle. The fiber over an $x \in X$ is the space of bases $b\colon \mathbb{R}^n \to T_xX$.
(2) Similarly, if X is a Riemannian manifold, we can restrict to orthonormal frames, which defines a principal O_n-bundle $\mathscr{B}_\mathrm{O}(X) \to X$.
(3) In specific cases, you can say more. For example, if $X = \mathbb{E}^n$ (so Euclidean space with the standard metric), $\mathscr{B}_\mathrm{O}(\mathbb{E}^n)$ is a right Euc_n-torsor: any (x, b) and (x', b') are related by a unique Euclidean transformation.
(4) If S^n carries the usual metric, $\mathscr{B}_\mathrm{O}(S^n)$ is an O_{n+1}-torsor, as it's determined by $n+1$ unit vectors: the first determines the point $e_0 \in S^n$, and the rest determine the frame $e_1, \dots, e_n$.
(5) Let $X = \mathbb{H}^n$ be hyperbolic space, e.g. a hyperboloid of two sheets in $\mathbb{R}^{n,1}$ inheriting a Riemannian metric (even though the metric on $\mathbb{R}^{n,1}$ has signature $(n, 1)$, and in particular is not Riemannian). Then, $\mathscr{B}_\mathrm{O}(\mathbb{H}^n)$ is an $\mathrm{O}^+_{n,1}$-torsor. By $\mathrm{O}_{n,1}$ we mean the group of matrices preserving the (Lorentzian) metric on $\mathbb{R}^{n,1}$, and then we choose the connected component containing the identity.
(6) There's a $\mathbb{T}$-bundle $\pi\colon S^3 \to S^2$ which is the restriction of the projection $\mathbb{C}^2 \twoheadrightarrow \mathbb{CP}^1$ to $S^3 \subset \mathbb{C}^2$, and using the identification $S^2 \cong \mathbb{CP}^1$. This is called the *Hopf fibration*. The same construction more generally defines a principal $\mathbb{T}$-bundle $\pi\colon S^{2n+1} \twoheadrightarrow \mathbb{CP}^n$.
(7) Let G be a Lie group and $H \subset G$ be a closed subgroup. Then, G/H is a manifold, and the quotient map $\pi\colon G \to G/H$ is a principal H-bundle. Verifying this is on the homework. ◂

Though we've just seen some examples where the group of isometries is transitive, this is not true for every Riemannian manifold. For example, the curved torus (with the standard embedding in $\mathbb{R}^3$) doesn't have a transitive group of isometries.

Definition 13.4. Let F be a smooth manifold with a left G-action and $P \to X$ be a principal G-bundle. Then, the *mixing construction* or *associated bundle* is the fiber bundle $F_P = P \times_G F \to X$, where

$$P \times_G F := P \times F/((pg, f) \sim (p, gf)).$$

The right G-action is $(p, f) \cdot g = (p \cdot g, g^{-1}f)$, which you can check is well-defined in the quotient.

The idea is that the quotient tells you how far the G-action is from being a product action: this is always true locally, so the mixing construction is locally trivial, and therefore a fiber bundle. The principal bundle controls everything: the fibers look like F, but they're twisted in a way dictated by P.

This is how Steenrod originally defined fiber bundles, and in fact every fiber bundle arises in this way. This perspective means that if P or F has extra structure, so does any fiber bundle obtained by mixing them.

Example 13.5. Consider the frame $\mathrm{GL}_n(\mathbb{R})$-bundle $\mathscr{B}(X) \to X$ over an n-dimensional manifold X. The tangent bundle $TX \to X$ is the result of the mixing construction applied to $F = \mathbb{R}^n$ with the usual left $\mathrm{GL}_n(\mathbb{R})$-action. If one takes $F = (\mathbb{R}^n)^*$ instead, the result is the cotangent bundle. The extra structure on F

in both cases carries over to the mixing construction, which is a vector bundle. Similarly, one can look at $\mathrm{End}(\mathbb{R}^n)$ or the space of inner products in $\mathrm{Sym}^2 \mathbb{R}^n$; the latter mixes to the bundle of Riemannian metrics (which is not a vector bundle).

$\mathrm{GL}_n(\mathbb{R})$ also acts on the set of two points $\{\pm 1\}$ by $g \mapsto \operatorname{sign} \det g$. The associated bundle is a principal $\mathbb{Z}/2$-bundle, also known as a double cover, and specifically is the orientation double cover of X. ◀

Example 13.6. Let G be a Lie group and $H \subset G$ be a closed subgroup, so there is a principal H-bundle $G \to G/H$. There is an adjoint representation of H on $\mathfrak{g}/\mathfrak{h}$, and the mixing construction can be canonically identified with $T(G/H) \to G/H$ — but with structure group H, rather than $\mathrm{GL}_n(\mathbb{R})$. If H is small, this is a lot of extra information.

For example, let $X = S^6$ with the round metric. It can be written as the homogeneous space $\mathrm{O}_7/\mathrm{O}_6$, or more exotically $\mathrm{G}_2/\mathrm{SU}_3$ (which is smaller: $\dim \mathrm{O}_7 = 15$, but $\dim \mathrm{G}_2 = 8$). Thus we obtain an O_6- or SU_3-structure on TS^6. ◀

This can be thought of as a differential-geometric realization of Felix Klein's *Erlangen* program, which says that geometric properties of an object should be understood in terms of the symmetries of that object.

Let $\pi\colon P \to X$ be a principal G-bundle and $p \in P$. Then, one can push forward along $\pi_{p*}\colon T_pP \to T_{\pi(p)}X$, which defines a short exact sequence of vector spaces

$$0 \longrightarrow \ker(\pi_{*p}) \longrightarrow T_pP \longrightarrow T_{\pi(p)}X \longrightarrow 0,$$

or, doing this for all $p \in P$ simultaneously,

$$0 \longrightarrow \ker(\pi_*) \longrightarrow TP \longrightarrow \pi^*TX \longrightarrow 0, \tag{13.7}$$

a short exact sequence of vector bundles over P. The kernel of π_* is the bundle of vectors tangent to the G-orbits, and is called the *vertical vector bundle*, denoted $T(P.X)$ or $T(\pi)$.

Lemma 13.8. *There's a canonical identification $T(P/X) \cong P \times \mathfrak{g}$ as vector bundles, with the projection $P \times \mathfrak{g} \to P$ onto the first factor.*

Proof. Given $(p, \xi) \in P \times \mathfrak{g}$, the isomorphism sends it to $t \mapsto pe^{t\xi}$. That is, we use the G-action to define a map $\mathfrak{g} \to \mathcal{X}(P)$; with a right G-action this preserves Lie bracket, but for a left G-action there would have to be a sign. Then, evaluation at p defines a map $\mathcal{X}(P) \to T_pP$. ☒

Anyways, the point is that the vertical tangent bundle is trivializable, and is trivialized by $\mathfrak{g}$.

The frame bundle $\pi\colon \mathscr{B}(X) \to X$ has extra structure, a canonical form $\theta \in \Omega^1_{\mathscr{B}(X)}(\mathbb{R}^n)$ called the *soldering form*. Let's fix some notation: let $e_1, \dots, e_n$ be the standard basis of $\mathbb{R}^n$ and $e^1, \dots, e^n$ be the dual absis for $(\mathbb{R}^n)^*$. Let $e_i^j := e_i \otimes e^j$ in $\mathrm{End}(\mathbb{R}^n) \cong (\mathbb{R}^n)^* \otimes \mathbb{R}^n$, i.e. $e_i^j(e_k) = \delta_k^j e_i$: that is, $e_j \mapsto e_i$ and $e_k \mapsto 0$ for $k \neq j$.

The soldering form isdefined by the formula $\theta_p(\eta) = b(p)^{-1}\pi_*\eta$, where $\eta \in T_p\mathscr{B}(X)$ and $b(p)\colon \mathbb{R}^n \to T_{\pi(p)}X$ is the basis associated to $p = (x, b) \in \mathscr{B}(X)$. Another way to say this is that $\pi_*\eta = b(\theta^i(\eta)e_i)$.

Vertical vector fields $\widehat{\xi} \in \mathcal{X}_{\mathscr{B}(X)}(\mathfrak{g}^*)$ are killed by θ, so θ is "horizontal" in a sense. You might imagine that there's a horizontal vector field $\widehat{\zeta} \in \mathcal{X}_{\mathscr{B}(X)}((\mathbb{R}^n)^*)$ and a $\Theta \in \Omega^1_{\mathscr{B}(X)}(\mathfrak{g})$ that kills the horizontal vector field. You could get that information if you had a distribution that's complimentary to the vertical bundle, equivalent to a section for (13.7), which would express TP as a direct sum of $T(P/X)$ and π^*TX. This structure is called a connection.

Definition 13.9. A *connection* on a principal G-bundle $\pi\colon P \to X$ is a G-invariant distribution $H \subset TP$ complementary to the vertical $\ker(\pi_*)$.

A vector in $H_p \subset T_pP$ is called *horizontal*, and a vector in $\ker(\pi_*)$ is called *vertical*.

What this means is that if $\xi \in T_{\pi(p)}X$ (a vector downstairs), the connection determines a horizontal lift of it, a $\widehat{\xi} \in H_p$. We hope to integrate that to convert paths on X to paths on P, and if P is the frame bundle, we get something beautiful: one gets a horizontal lift of basis elements and obtains a vector field for each basis element, at least locally. So on the frame bundle, these horizontal spaces are identified with $\mathbb{R}^n$, and the vertical is already identified with $\mathfrak{o}_n$. The existence of a connection (and the integrability condition we'll get back to) parallelizes the neighborhood of a point!

The integral curves in the frame bundle project down to particular curves on X. What's special about these curves? Tune in next time to find out.

Anyways, we also have the form Θ, which can be thought of as splitting (13.7) as a map $TP \to \ker(\pi_*)$. The data of the connection is determined by Θ, but since $\ker(\pi_*) \cong P \times \mathfrak{g}$, this means the connection is determined by a Lie-algebra-valued 1-form $\Theta \in \Omega^1_{\mathscr{B}(X)}(\mathfrak{g})$.

When we discussed distributions, we asked whether they were integrable. We know they're always locally integrable, but what about globally? We'll introduce curvature on a general Riemannian manifold as the obstruction to global integrability of the distribution.

Meanwhile, let's discuss the geometry that a connection buys. Recall that covering spaces $\pi\colon \widetilde{X} \to X$ have *path lifting* (so are an example of a fibration in homotopy theory): if $\gamma\colon (a,b) \to X$ is a path sending $0 \mapsto x_0$ and $\widetilde{x}_0 \in \pi^{-1}(x_0)$, then there's a unique path $\widetilde{\gamma}\colon (a,b) \to \widetilde{X}$ sending $0 \mapsto \widetilde{x}_0$.

If G is a discrete group, principal G-bundles are Galois covering spaces with covering group G. But more generally, we need a connection H to do path lifting on a principal G-bundle $P \to X$.

Definition 13.10. A curve $\widetilde{\gamma}\colon (a,b) \to P$ is *horizontal* if $\dot{\widetilde{\gamma}} \in H_{\gamma(t)}$ for all t.

Theorem 13.11. *Given a connection H on a principal bundle $\pi\colon P \to X$, a path $\gamma\colon (a,b) \to X$ with $\gamma(0) = x_0$, and a $\widetilde{x}_0 \in \pi^{-1}(x_0)$, there is a unique horizontal lift $\widetilde{\gamma}\colon (a,b) \to P$ such that $\widetilde{\gamma}(0) = \widetilde{x}_0$.*

If one specifies that the curve must begin and end at the same points, so the curve closes up, its lift need not close up; its *holonomy* measures the difference (in the fiber, as a G-torsor) between its starting and ending points.

Proof. You can check that the pullback of a principal G-bundle $P \to X$ by a map $f\colon Y \to X$ is again a principal G-bundle $f^*P \to Y$. So let's pull back $P \to X$ by γ, producing a principal G-bundle $\gamma^*P \to (a,b)$, and a map $\widehat{\gamma}\colon \gamma^*P \to P$. Concretely, $\gamma^*P = \{(t,p) \in (a,b) \times P \mid \gamma(t) = \pi(p)\}$.

The connection also pulls back, just as one-forms pull back: $\gamma^*H_p := \{\eta \in T_p\gamma^*P \mid \widehat{\gamma}_*\eta \in H_{\widehat{\gamma}(p)}\}$. This is a rank-1 distribution, hence integrable (or involutive), so let Γ be the maximal leaf of the foliation through $\widetilde{x}_0$. Then, one can show that $\pi|_\Gamma\colon \Gamma \to (a,b)$ is a diffeomorphism, and we can define $\widetilde{\gamma} = (\pi|_\Gamma)^{-1}$, which is unique by the general theory of integrating distributions.

The argument that $\pi|_\Gamma$ is a diffeomorphism is the same as above: it's a covering map where the cover is connected, but the base (a,b) is simply connected. The G-invariance is what keeps it from going to infinity. ⊠

The connection defines an isomorphism of fibers $P_{x_0} \to P_{x_1}$ given a path $x_0 \to x_1$, which is called *parallel transport*. This comes along for all associated bundles, and in particular it's posible to parallel-transport vectors, covectors, etc. Unfortunately, we can only do this along curves, not globally.

Next time, we'll return to the Riemannian situation, and see that in Riemannian geometry, there is a distinguished connection[12] that satisfies the first Maurer-Cartan equation.

Lecture 14.

: 3/2/17

Let G be a Lie group and X be a smooth manifold. Last time, we talked about the mixing construction: if F is a smooth manifold with a left G-action, then $F_P = P \times_G F$, which forms a principal G-bundle over X. A section of $F_P \to X$ is a G-equivariant map $\psi\colon P \to F$, i.e. $\psi(pg) = g^{-1}\psi(p)$.

We also defined a connection on a principal G-bundle $\pi\colon P \to X$ to be a G-invariant distribution $H \subset TP$ such that $H \oplus T(P/X) \cong TP$. By G-invariance, we mean under the right G-actioin: $(R_g)_*H_p = H_{p\cdot g}$. We proved that a connection induces a unique lift of horizontal paths, which in particular induces parallel transport in every associated fiber bundle $F_P \to X$. Specifically, if γ is a path from x_0 to x_1, the connection induces a path γ' from p_0 to p_1 in P, and we want to lift this to F_P. We do this by making the F-component constant: on F_P, the path lift starting at (p_0, f) is $(\gamma'(t), f)$. Since there's a quotient by an equivalence relation here, one should check that this behaves well under the G-action, which it does.

[12] A typical partition-of-unity argument shows that connections always exist on a manifold.

Remark. Another way to think about this is that H defines a distribution on $P \times F$ which is G-invariant, and therefore descends to a distribution on F_P. The parallel transport on F_P is horizontal with respect to this distribution. This is an instance of the idea that additional geometric structure on a principal bundle carries over to all of its associated fiber bundles, where this geometric structure is the connection. ◂

The main case for us is where $P \to X$ is a frame bundle, orthonormal or not. The parallel transport we recover resembles the parallelism that exists in an affine space — but here, we can only transport along curves, and there may be holonomy.

You could take subspaces of tangent spaces, symmetric bilinear forms, and any object that defines a fiber bundle can be parallel-transported using the associated bundle construction. This will also enable us to define a derivative: derivatives require subtraction of values obtained from nearby points, and this requires parallel transport. This doesn't require the bundle of frames, as it can be done more generally.

If $H \subset TP$ is a connection, where $\pi\colon P \to X$ is a principal G-bundle, we get a short exact sequence of vector bundles

$$0 \longrightarrow T(P/X) \longrightarrow TP \xrightarrow{\pi^*} \pi^*TX \longrightarrow 0,$$

and $T(P/X) \cong P \times \mathfrak{g} = \underline{\mathfrak{g}}$ (i.e. the constant vector bundle with fiber $\mathfrak{g}$). Given an $\eta \in T_pP$, let $\Theta_p(\eta) \in T_p(P/X)\colon \mathfrak{g}$ be the vertical projection of η along H_p, so Θ_p is the identity when restricted to $T_p(P/X)$ and is 0 on H_p. This defines a map $\Theta_p\colon T_pP \to \mathfrak{g}$, hence a $\mathfrak{g}$-valued 1-form $\Theta \in \Omega^1_P(\mathfrak{g})$.

This notation looks familiar, and that's no coincidence.

Lemma 14.1. *For any $x \in X$, $\Theta|_{P_x} = \theta_G$ is the Maurer-Cartan form for G. Moreover, for any $g \in G$, $R_g^*\Theta = \mathrm{Ad}_{g^{-1}}\Theta$. Conversely, any $\Theta \in \Omega^1_P(\mathfrak{g})$ satisfying these two properties determines a connection.*

Proof. The first part comes from unwinding the definition: the tangent space of any G-torsor can be identified with $\mathfrak{g}$, which is how we wrote down the Maurer-Cartan form on a G-torsor. Thus, $\Theta|_{P_x} = \theta$.

For the second part, $(R_g^*\Theta)_p(\eta) = \Theta_{pg}(R_{g*}\eta)$ when $\eta \in T_pP$. Choose a curve $p_t\colon (-\varepsilon, \varepsilon) \to P$ with $p(0) = p$ and $\eta = \cdot p(0)$, and write $\eta = \eta_H + \eta_V$, with η_H and η_V denoting the horizontal and vertical components of η, respectively. Then, $\eta_V = \widehat{\xi}_p$ for some $\xi \in \mathfrak{g}$ such that $\xi = \Theta_p(\eta)$, and in particular

$$\eta_V = \widehat{\xi}_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pe^{t\xi}.$$

Then, $(R_g)_*\eta = (R_g)_*\eta_H + (R_g)_*\eta_V$. The first part is in $H_{p\cdot g}$, and

$$\begin{aligned}
(R_g)_*\eta_V &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} R_g\left(pe^{t\xi}\right)\\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pe^{t\xi}g\\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} pg\left(g^{-1}e^{t\xi}g\right)\\
&= \left(\mathrm{Ad}_{g^{-1}\xi}\right)\Big|_{pg}.
\end{aligned}$$

Then, compare this with

$$(\mathrm{Ad}_{g^{-1}}\Theta)_p(\eta) = \mathrm{Ad}_{g^{-1}}\Theta_p\eta = \mathrm{Ad}_{g^{-1}}\eta. \qquad \boxtimes.$$

The two equations in (14.1) are an affine equation: a constant value through P_x and a linear equation. This implies that the space of solutions, namely connections on X, is an affine space: the difference of any two connections is a vector space.

Recall that $\mathscr{B}(X) \to X$ is the $\mathrm{GL}_n(\mathbb{R})$-bundle of frames on a smooth n-manifold X, and there's a soldering form $\theta = \theta^- e_i \in \Omega^1_{\mathscr{B}(X)}(\mathbb{R}^n)$, which transforms under the equation $R_g^*\theta = g^{-1}\cdot\theta$, where the action is matrix multiplication. Moreover, if ζ is vertical, then $\iota_\zeta\theta = 0$, so you might want to push θ down to the base, but the action of G on $\mathbb{R}^n$ is nontrivial.

However, you can bring it over to the associated fiber bundle modeled on $\mathbb{R}^n$, i.e. the tangent bundle, so we obtain a form $\theta \in \Omega^1_X(TX)$. This construction is canonical, and so the only choice we have is for it to be id_{TX}. It's a good exercise to check that what you actually get is the identity.

Now suppose $\Theta \in \Omega^1_{\mathscr{B}(X)}(\mathfrak{gl}_n(\mathbb{R}))$ is a connection (since the frame bundle is an example of a principal bundle). Then, we can write $\Theta = \Theta^i_j e^j_i$, where $\{e^j_i\}$ is the basis for the Lie algebra consisting of matrices with a 1 in entry (i, j) and 0s elsewhere. Then, the forms θ^i and Θ^i_j give $n^2 + n$ linearly independent forms which give a global trivialization of $T^*\mathscr{B}(X) \to \mathscr{B}(X)$. This gloabl parallelism makes this a very nice place to do calculus.

You could also take the dual trivialization: dual to θ^i is ∂_i, the horizontal component, which we've seen before; and dual to Θ^i_j is the vertical component $\tilde{e}^j_i$, which is something new. More explicitly, $\partial_i|_p$ is the horizontal lift of the i^{th} basis element of the basis $b(p)$ of $T_{\pi(p)}X$.

Definition 14.2. A curve $\gamma\colon (a, b) \to X$ is a *geodesic* (relative to Θ) if $\dot\gamma$is parallel.

So geodesics are those which aren't turning: there's no acceleration. To know whether the velocity is changing, you have to compute instantaneous change through parallel transport, which requires the *affine connection* on $\mathscr{B}(X)$.

Proposition 14.3. *Integral curves of* ∂_1 *project under* π *to geodesics on* X*.*

The idea is that the acceleration of an integral curve for ∂_1 is only in the vertical direction.

Proof. Let γ be such a curve, so that the horizontal lift of $\dot\gamma$ is $\dot{\tilde\gamma} = \partial_1$. Writing $\dot\gamma$ as a function $\gamma^*\mathscr{B}(X) \to \mathbb{R}^n$, it's the constant function with value e_1; since it's constant along a horizontal curve, then $\dot\gamma$ is parallel, and hence γ is a geodesic. ⊠

This perspective gives you all of the usual theorems on geodesics: for example, given a point and an initial velocity, one finds a unique parallel solution starting at a given point in the frame bundle, hence a unique geodesic with that initial position and velocity data on X.

Torsion. There are lots of possible connections on a manifold. But we're going to impose a condition – that the torsion vanishes – which singles out a unique connection in the Riemannian case.

Recall that when $X = \mathbb{A}^n$, $\mathscr{B}(X) = \mathrm{Aff}_n$, and if the Maurer-Cartan forms define the soldering form and connection, then we had equations

$$\begin{aligned} d\theta + \Theta \wedge \theta &= 0 \\ \mathrm{d}\Theta + \Theta \wedge \Theta &= 0, \end{aligned}$$

or in indices,

$$\begin{aligned} \mathrm{d}\theta^i + \Theta^i_j \wedge \theta^j &= 0 \\ \mathrm{d}\Theta^i_j + \Theta^i_k \wedge \Theta^k_j &= 0. \end{aligned} \tag{14.4}$$

However, this is *not* true for general X! Instead, we give them names.

Definition 14.5. Let Θ be a connection on the frame bundle. Then, the *curvature* is

$$\Omega := \mathrm{d}\Theta + \Theta \wedge \Theta \in \Omega^2_{\mathscr{B}(X)}(\mathfrak{gl}_n(\mathbb{R}))$$

and the *torsion* is

$$\tau := \mathrm{d}\theta + \Theta \wedge \theta \in \Omega^2_{\mathscr{B}(X)}(\mathbb{R}^n).$$

Remark. The curvature can be defined more generally, in the context of a principal G-bundle $P \to X$, in which case we would say

$$\Omega := \mathrm{d}\Theta + \frac{1}{2}[\Theta \wedge \Theta] \in \Omega^2_P(\mathfrak{g}),$$

which agrees with our definition when we pass to $\mathscr{B}(X)$. ◂

To interpret the torsion, let's compute it on basis vectors.

$$\begin{aligned} \tau(\partial_k, \partial_\ell) &= \mathrm{d}\theta(\partial_k, \partial_\ell) + (\Theta \wedge \theta)(\partial_k, \partial_\ell) \\ &= \partial_k\theta(\partial_\ell) - \partial_\ell\theta(\partial_k) - \theta([\partial_k, \partial_\ell]) + \Theta(\partial_k)\theta(\partial_\ell) - \Theta(\partial_\ell)\theta(\partial_k) \\ &= -\theta([\partial_k, \partial_\ell]). \end{aligned}$$

To figure out what this is, let ∂_k generate the flow φ_t and ∂_ℓ generate the flow ψ_s. Then,

$$[\partial_k, \partial_\ell] = \left.\frac{\partial^2}{\partial s \partial t}\right|_{s,t=0} \psi_{-s}\varphi_{-t}\psi_s\varphi_t.$$

The idea is that, as $s, t \to 0$, flowing in the x^k-direction, then the x^j-direction, then back along the $-x^k$-direction, then back along the $-x^j$-direction. You don't always end up back where you started, though you do in affine space. If the connection is *torsion-free*, meaning infinitesimally the connection looks a bit like affine space, the geometry is very nice, and in general the torsion provides a way to quantify how differently X and its connection behave from affine space.

We'll restrict to the Riemannian case soon, but the existence of a torsion-free connection compatible with a geometric structure – complex structure, symplectic structure, etc. – is an integrability condition, and such connections may or may not exist. One of the beautiful aspects of Riemannian geometry is that there always exists a unique connection that's compatible with the Riemannian metric and that is torsion-free.