MATH 623: DIFFERENTIAL GEOMETRY II

DEAN BASKIN

1. Introduction

This is the second semester of a two-semester graduate course providing an introduction to differential geometry. The second semester is primarily a study of Riemannian manifolds with a focus on curvature. At the end of the course, we may go in different directions depending on the interests of the class. Possible directions include comparison theorems, principal bundles and the Atiyah–Singer index theorem, Lorentzian manifolds, the Hodge theorem, or the Chern–Gauss–Bonnet theorem.

The topics roughly covered (updated later based on course interest) include:

-

Make sure to add topics!

2. Preliminaries and review

Recall from last semester (or your previous experience) the notions of smooth manifold, the tangent and cotangent bundles of a smooth manifold, and tensor fields. Put in definitions to refresh! Maybe recall how to work with the objects!

Sections. Diffeomorphisms.

In coordinates $(x^1, \dots, x^n)$ on a patch U of M, recall that $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ form a basis for T_pM for each $p \in U$.

The following characterization of tensor fields is so useful, let's just get it out of the way.

Lemma 1. *Suppose $\overline{T} : \mathcal{X} \times \cdots \times \mathcal{X} \times \Omega^1 \times \cdots \times \Omega^1 \to C^\infty(M)$ is an $\mathbb{R}$-multillinear function of k vector fields and ℓ 1-forms (covector fields). Then $\overline{T}$ arises from a (ℓ, k)-tensor field T if and only if T is multilinear over $C^\infty(M)$ in each of its arguments, e.g.,*

$$\overline{T}\left(fX_1, \dots, X_k, \omega_1, \dots, \omega_\ell\right)(p) = f(p)\overline{T}\left(X_1, \dots, X_k, \omega_1, \dots, \omega_\ell\right)(p).$$

Proof. Given a (ℓ, k)-tensor field T, we form $\overline{T}$ from it by evaluating at each point. We then have

$$\begin{aligned}\overline{T}\left(fX_1, \dots, X_k, \omega_1, \dots, \omega_\ell\right)(p) &= T_p\left(f(p)X_{1,p}, \dots, X_{k,p}, \omega_{1,p}, \dots, \omega_{\ell,p}\right) \\ &= f(p)T\left(X_{1,p}, \dots, X_{k,p}, \omega_{1,p}, \dots, \omega_{\ell,p}\right) \\ &= f(p)\overline{T}\left(X_1, \dots, X_k, \omega_1, \dots, \omega_\ell\right)(p).\end{aligned}$$

For the other direction, let's just do the case of $\overline{T} : \mathcal{X} \to C^\infty(M)$. (This contains the main idea; the general case is an exercise in careful bookkeeping.) Suppose for all $f \in C^\infty(M)$ and $X \in \mathcal{X}(M)$ we have

$$\overline{T}(fX)(p) = f(p)\overline{T}(X)(p).$$

Our aim is to show that $\overline{T}$ arises from a 1-form ω. Let's start by defining the purported 1-form. Given $p \in M$ and $v \in T_pM$, choose an $X \in \mathcal{X}(M)$ so that $X_p = v$. We define

$$\omega_p(v) = \overline{T}(X)(p).$$

We must show that ω is well-defined, i.e., that it does not depend on the choice of vector field X. Suppose X_1 and X_2 are two vector fields with $X_{1,p} = X_{2,p} = v$. In particular, the vector field $Y = X_1 - X_2$ satisfies $Y_p = 0$ and so there are smooth vector fields $Z_1, \dots, Z_r$ and smooth functions $f_1, \dots, f_r$ with $f_1(p) = \dots = f_r(p) = 0$ so that $Y = f_1 Z_1 + \dots + f_r Z_r$. We then have

$$\overline{T}(Y)(p) = \sum_{j=1}^{r} f_j(p)\overline{T}(Z_j)(p) = 0,$$

so that $\overline{T}(X_1)(p) = \overline{T}(X_2)(p)$ and thus $\omega_p : T_pM \to \mathbb{R}$ is well-defined. Its linearity is clear from the linearity of $\overline{T}$ and its smoothness follows from the mapping properties of $\overline{T}$ so it is indeed a 1-form. □

2.1. **Notation.** Unless explicitly noted, all manifolds in this course will be smooth (i.e., C^∞) manifolds.

We use $\mathcal{X}(M)$ to denote the space of C^∞ vector fields on M, i.e., smooth sections of TM. For sections of other bundles $E \to M$ we often use $\Gamma(E)$ to denote the space of smooth sections. For vector fields along a curve α we use $\mathcal{X}(\alpha)$ and sections of E above a curve α are denoted $\Gamma(E, \alpha)$.

3. Riemannian metrics

Suppose M isa smooth manifold of dimension n (typically $n \geq 2$).

Suppose $g : \mathcal{X}(M) \times \mathcal{X}(M) \to C^\infty(M)$ is a symmetric $(0,2)$-tensor. (In other words, g is a tensor field so that $g_p : T_pM \times T_pM \to \mathbb{R}$ is symmetric at each point $p \in M$.)

Definition 2. We say that a symmetric $(0,2)$-tensor is a *Riemannian metric* if $g_p(v,v) > 0$ for all $p \in M$, $v \in T_pM$ with $v \neq 0$. The tensor g is *pseudo-Riemannian* if for all $p \in M$, if $v \in T_pM$ has $g_p(v,w) = 0$ for all $w \in T_pM$, then $v = 0$. (In other words, g is pseudo-Riemannian if it is non-degenerate and Riemannian if it is additionally an inner product as you know it from linear algebra.)

A smooth manifold M equipped with a Riemannian metric g is called a *Riemannian manifold*.

In terms of the coordinate basis induced by a coordinate patch $(x^1, \dots, x^n)$ in M, we set

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

so that $g_{ij} \in C^\infty(U)$, $g_{ji} = g_{ij}$, and if $v = v^i \frac{\partial}{\partial x^i}$ and $w = w^j \frac{\partial}{\partial x^j}$, then

$$\langle v, w\rangle_p := g_p(v,w) = \sum_{i,j} g_{ij} v^i w^j |_p.$$

A Riemannian metric g then gives an inner product on each tangent space. It also induces metrics on associated bundles (in what follows g is always a metric on the tangent spaces; notation for the induced metrics varies wildly by source and eventually we'll just use g to denote all of them unless there can be confusion as to where various objects live).

One way to view the mechanism for this induction is that g gives a way to "raise and lower indices". More precisely, a Riemannian metric g provides an isomorphism between the tangent and cotangent spaces at each point. Given $\omega \in T_p^*M$, we associate to it a vector $w_\omega \in T_pM$ by demanding that

$$g_p(w_\omega, v) = \omega(v)$$

for all $v \in T_pM$. Because g is non-degenerate, this uniquely defines the vector w_ω. In local coordinates, the displayed equation reads

$$\sum_{i,j} g_{ij}(w_\omega)^i v^j = \sum_j \omega_j v^j,$$

so that

$$(w_\omega)^i = \sum_j g^{ij}\omega_j,$$

where g^{ij} are the components of the matrix inverse of (g_{ij}). Similarly, if $v \in T_pM$, one can identify it with the one-form $\omega_v \in T_p^*M$ so that

$$\omega_v(u) = g_p(v, u)$$

for all $u \in T_pM$. In local coordinates, $(\omega_v)_i = \sum_j g_{ij}v^j$.

(1) Cotangent bundle. Given $\omega, \eta \in T_p^*M$, we define the metric G (sometimes denoted g^{-1}, sometimes still just g) by

$$G(\omega, \eta) = g(w_\omega, w_\eta),$$

where w_ω is the vector associated to ω and w_η is the one associated to η. In coordinates, we have that the (i,j)-component of the metric G is the same as the (i,j) component of the matrix $g^{-1} = (g_{k\ell})^{-1}$, i.e., g^{ij}. To check this, observe that

$$\begin{aligned} g(w_\omega, w_\eta) &= \sum_{i,j} g_{ij}(w_\omega)^i(w_\eta)^j \\ &= \sum_{i,j,k,\ell} g_{ij}g^{ik}\omega_k g^{j\ell}\eta_\ell \\ &= \sum_{jk\ell} \delta_j^k g^{j\ell}\omega_k\eta_\ell = \sum_{i,j} g^{ij}\omega_i\eta_j. \end{aligned}$$

(2) Tensor bundles. For $T_pM \otimes T_pM$ (and higher powers), say that

$$g^{\otimes}(v_1 \otimes v_2, w_1 \otimes w_2) = g(v_1, w_1)g(v_2 w_2),$$

and extend linearly. For factors of T_p^*M, also use the raising/lower operator.

(3) Exterior powers. Use that $\Lambda^k(TM)$ is a sub-bundle of $(T^*M)^{\otimes k}$ and use above.

(4) Endomorphism bundle. Identify $\mathrm{End}(TM)$ with $T^*M \otimes TM$.

As an example of a Riemannian metric, suppose $F : M \to \mathbb{R}^N$ is an immersion (so that dF_p is injective for all p and thus has rank $\dim M$). The immersion F (and the ambient inner product on $\mathbb{R}^N$) induces a Riemannian metric on M by

$$g_p(v, w) = \langle dF_p(v), dF_p(w)\rangle_{\mathbb{R}^N}$$

for all $v, w \in T_pM$. Exercise: Check that this is a Riemannian metric on M.

Definition 3. Two Riemannian manifolds (M, g_M) and (N, g_N) are *isometric* if there is a diffeomorphism $F : M \to N$ so that

$$\langle dF_p(v), dF_p(w)\rangle_{g_N} = \langle v, w\rangle_{g_M}$$

for all $p \in M$ and $v, w \in T_pM$.

In other words, (M, g_M) is isometric to (N, g_N) if there is a diffeomorphism $F : M \to N$ for which $F^*g_N = g_M$.

More examples:

(1) $\mathbb{R}^n$ equipped with the dot product. Here $T_p\mathbb{R}^n \cong \mathbb{R}^n$ canonically, and $g(v, w) = v \cdot w$. Once we have the machinery to make this precise, we'll see that this is our model of a *flat* space.

(2) $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ with the metric induced by the inclusion map. Concretely, we can use coordinates $(\theta^1, \dots, \theta^n) \in [0, \pi)^{n-1} \times [0, 2\pi)$ on a large patch of $\mathbb{S}^n$ together with the map F given by

$$F(\theta^1, \dots, \theta^n) = \begin{pmatrix} \cos\theta^1 \\ \sin\theta^1\cos\theta^2 \\ \sin\theta^1\sin\theta^2\cos\theta^3 \\ \vdots \\ \sin\theta^1\sin\theta^2\dots\sin\theta^{n-1}\cos\theta^n \\ \sin\theta^1\sin\theta^2\dots\sin\theta^{n-1}\sin\theta^n \end{pmatrix}$$

A straightforward computation shows that

$$dF_{(\theta^1,\dots,\theta^n)} = \begin{pmatrix} -\sin\theta^1 & 0 & \dots & 0 \\ \cos\theta^1\cos\theta^2 & -\sin\theta^1\sin\theta^2 & \dots & 0 \\ \cos\theta^1\sin\theta^2\cos\theta^3 & \sin\theta^1\cos\theta^2\cos\theta^3 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \cos\theta^1\sin\theta^2\dots\sin\theta^n & \sin\theta^1\cos\theta^2\dots\sin\theta^n & \dots & \sin\theta^1\sin\theta^2\dots\cos\theta^n \end{pmatrix}$$

In particular, we have

$$g\left(\frac{\partial}{\partial\theta^i}, \frac{\partial}{\partial\theta^j}\right) = dF\left(\frac{\partial}{\partial\theta^i}\right)\cdot dF\left(\frac{\partial}{\partial\theta^j}\right) = \begin{cases} 0 & i \neq j \\ 1 & i = j = 1 \\ \sin^2\theta^1\sin^2\theta^2\dots\sin^2\theta^{j-1} & i = j \neq 1 \end{cases}.$$

(This in fact homogeneous, isotropic, etc., but we haven't defined these terms.) The sphere is the nicest example of a *positively curved* space (to be made more precise later).

(3) Hyperbolic space $\mathbb{H}^n$. We'll describe three models. (But there are more! Hyperbolic space just keeps on giving.) Fix some $R > 0$ (this is a parameter going into the metric, just as we could have changed the radius of our sphere in the previous example).

(a) The upper half-space $U_R = \{(x, y^1, \dots, y^{n-1}) \in \mathbb{R}^n \mid x > 0\}$ equipped with the metric

$$g = R^2\frac{dx^2 + (dy^1)^2 + \dots + (dy^{n-1})^2}{x^2},$$

i.e., we take the inner product of two vectors $v, w \in T_{(x,y)}U_R$ by

$$\langle v, w\rangle_{(x,y)} = R^2\frac{v\cdot w}{x^2}.$$

(b) The Poincaré ball model: Let $B^n_R \subset \mathbb{R}^n$ denote the ball of radius R and equip it with the metric

$$g = 4R^2 \frac{(du^1)^2 + \cdots + (du^n)^2}{(R^2 - |u|^2)^2}.$$

(c) The hyperboloid model: Consider $\mathbb{R}^{n+1}$ equipped with the pseudo-Riemannian metric $\eta_{(t,x^1,\ldots,x^n)} = -(dt)^2 + \sum (dx^j)^2$ and let $\mathbb{H}^n_R$ denote one sheet of the two sheeted hyperboloid:

$$\mathbb{H}^n_R = \{t > 0\} \cap \left\{-t^2 + |x|^2 = -R^2\right\},$$

and let $g = i^*\eta$, where $i : \mathbb{H}^n_R \to \mathbb{R}^{n+1}$ is the inclusion.

Theorem 4. *All three of the above models of hyperbolic space are isometric.*

Proof. You should fill in most of this proof yourself! I'll give you the maps to show it for the hyperboloid and ball models.

Let $S = (-R, 0, \ldots, 0) \in \mathbb{R}^{n+1}$ and let $P \in \mathbb{H}^n_R$, say $P = (t, x^1, \ldots, x^n)$. Define the map $\pi : \mathbb{H}^n_R \to B^n_R$ by letting $\pi(P) = u \in \mathbb{R}^n$, where $(0, u)$ is the point where the line from S to P intersects $\{t = 0\}$.

Note that this line is given by

$$(-R, 0, \ldots, 0) + s(t + R, x^1, \ldots, x^n),$$

which hits $t = 0$ when $st + sR - R = 0$, i.e., when $s = \frac{R}{t+R}$, so that

$$\pi\left(t, x^1, \ldots, x^n\right) = \frac{R}{t+R}(x^1, \ldots, x^n).$$

As $P \in \mathbb{H}^n_R$, we have that $|x|^2 = t^2 - R^2$ and thus

$$|\pi(P)|^2 = \frac{R^2}{(t+R)^2}|x|^2 = \frac{R^2(t^2 - R^2)}{(t+R)^2} = R^2\frac{t-R}{t+R} < R^2$$

and thus $\pi(P) \in B^n_R$.

The inverse map $\pi^{-1} : B^n_R \to \mathbb{H}^n_R$ is given by

$$\pi^{-1}(z^1, \ldots, z^n) = \left(R\frac{R^2 + |z|^2}{R^2 - |z|^2}, 2\frac{R^2 z}{R^2 - |z|^2}\right).$$

You should check that this is the correct form of the inverse and that both π preserves the inner product. $\square$

4. Covariant differentiation and connections

Recall that a (smooth) k-dimensional vector bundle over a smooth manifold M consists of the data $\pi : E \to M$ so that

(1) π is surjective,
(2) $\pi^{-1}(p)$ is a k-dimensional vector space for each $p \in M$, and
(3) for each $p \in M$, there is a chart (x, U) around p in M and a diffeomorphism $\varphi : \pi^{-1}(U) \to x(U) \times \mathbb{R}^k$ that restricts to a vector space isomorphism on each fiber.

One of the first challenges in differential geometry is to determine how to differentiate sections of a vector bundle. For a trivial vector bundle in Euclidean space, you have the "constant sections" and so you just differentiate their coefficients. In general, however, there is no constant section because there is no canonical way of identifying the different fibers of the bundle. One way to get around this is to define the notion of "parallel transport" along a curve. In this view, for each smooth path $\gamma : [a,b] \to M$, we equip the bundle E with linear maps $P(\gamma)_s^t : E_{\gamma(s)} \to E_{\gamma(t)}$ depending smoothly on $s,t \in [a,b]$ (and also on γ in an appropriate sense). We further demand that

$$P(\gamma)_r^t \circ P(\gamma)_s^r = P(\gamma)_s^t.$$

The maps P provide a way of performing parallel translation along the curve γ. (If E were equipped with a way of measuring distance or angles, we'd also demand that the parallel translation preserve this.) Given these maps, we could differentiate a section Y of E at $p \in M$ in the direction $v \in T_pM$ by taking a curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ and $\gamma'(0) = v$ and then finding

$$\nabla_v Y = \lim_{s\to 0} \frac{P(\gamma)_s^0 (Y_{\gamma(s)}) - Y_p}{s}.$$

One can check that this is a derivation, but showing that in fact it depends only on v and not on the extension γ doesn't quite follow without a more careful accounting of hypotheses.

Instead of defining parallel translation directly, we instead recover it from one of the other related quantities. As is common in many differential geometry texts (especially those focusing on vector bundles like the tangent and cotangent bundles), we'll use the notion of a *Koszul connection*, which we'll just call a *connection*.

Definition 5. A connection ∇ on the vector bundle $\pi : E \to M$ is an $\mathbb{R}$-linear map $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$ so that the product rule holds, i.e.,

$$\nabla(fs) = df \otimes s + f\nabla s,$$

for all smooth functions $f \in C^\infty(M)$ and smooth sections $s \in \Gamma(E)$.

Unwinding this definition, it's the same as providing, for each section $s \in \Gamma(E)$ and $p \in M$, an $\mathbb{R}$-linear map $(\nabla s)_p : T_pM \to E_p$ so that

1. $(\nabla s)_p$ depends smoothly on p,
2. for all $a, b \in \mathbb{R}$ and $s_1, s_2 \in \Gamma(E)$, $\nabla(as_1 + bs_2)_p = a(\nabla s_1)_p + b(\nabla s_2)_p$, and
3. ∇ satisfies a product rule, so for all smooth functions f on M and $v \in T_pM$,

$$\nabla(fs)_p(v) = df_p(v)s_p + f(p)(\nabla s)_p(v).$$

From now on we'll drop the p subscript and let it be implicit (as $v \in T_pM$). We also typically write $(\nabla s)(v)$ as $\nabla_v s$. When v is the value of a vector field $X \in \mathcal{X}(M)$, we also write $\nabla_X s$, which is the section with value $\nabla_{X_p} s$ at p.

The following lemma tells us that connections are local and so we will not need to worry about whether sections are defined globally or only locally.

Lemma 6. *If ∇ is a connection on E and $s_1, s_2 \in \Gamma(E)$ are such that $s_1 \equiv s_2$ in a neighborhood of $p \in M$, then for all $v \in T_pM$, $\nabla_v s_1 = \nabla_v s_2$.*

Proof. By linearity it suffices to show that if $s \equiv 0$ in a neighborhood of p then $\nabla_v s = 0$. This statement follows from the product rule. Indeed, for any $f \in C^\infty(M)$ so that $\operatorname{supp} f \subset \{s = 0\}$, we have $fs \equiv 0$ on M, and

$$0 = \nabla_v(fs) = df_p(v)s_p + f(p)\nabla_v s,$$

so that $f(p)\nabla_v s = 0$. This is true for any such f, so $\nabla_v s = 0$. $\square$

A word of warning: you might think that because a connection eats vector fields and gives you vector fields that it is a tensor, but the product rule (and Lemma 1) tells you that it's not. Connections do, however, lie in an affine space whose underlying linear space is the space of tensors.

Lemma 7. *If ∇ and $\widetilde{\nabla}$ are two connections on the tangent bundle TM then the difference $\nabla - \widetilde{\nabla}$ is a $(1,2)$-tensor.*

Proof. Let $T = \nabla - \widetilde{\nabla}$ be the $\mathbb{R}$-multilinear object $\mathcal{X} \times \mathcal{X} \times \Omega^1 \to C^\infty$ given by

$$T(X, Y, \omega) = \omega\left(\nabla_X Y - \widetilde{\nabla}_X Y\right).$$

By Lemma 1, we must only check that T is multilinear over C^∞. As it is already multilinear over C^∞ in X and ω, we need only check in Y, but this follows from the product rule. $\square$

We can think of $\nabla_v s$ as denoting a directional derivative of s in the direction of v. Just as we did in calculus (and in the previous semester of this course), we'd like to also differentiate along curves. Let's fix a curve $\alpha : (a, b) \to M$.

Definition 8. A section along the curve α is a map $t \mapsto s(t) \in E_{\alpha(t)}$ depending smoothly on t. In an abuse of notation we'll denote the set of smooth sections along α by $\Gamma(E, \alpha)$.

Proposition 9. *There is a unique map $\Gamma(E, \alpha) \to \Gamma(E, \alpha)$, denoted $s \mapsto \frac{D}{dt}s$ and called the covariant derivative of s along α, so that*

(1) $\dfrac{D}{dt}\left(s_1(t) + s_2(t)\right) = \dfrac{D}{dt}s_1(t) + \dfrac{D}{dt}s_2(t)$,
(2) $\dfrac{D}{dt}\left(f(t)s(t)\right) = f(t)\dfrac{D}{dt}s(t) + f'(t)s(t)$, *and*
(3) If $\widetilde{s} \in \Gamma(E)$ satisfies $s(t) = \widetilde{s}_{\alpha(t)} \in E_{\alpha(t)}$, then $\dfrac{D}{dt}s(t) = \nabla_{\alpha'(t)}\widetilde{s}$.

Proof. By localization we can assume that $\alpha(I)$ is contained in a single coordinate chart on which the bundle E is trivial. We then take a local basis $e_1, \ldots, e_k$ for all E_p for p contained in this chart and write $s(t) = \sum_{j=1}^k s^j(t)e_j$.

For uniqueness, we observe that if $\frac{D}{dt}$ satisfies all three conditions, we must have

$$\begin{aligned} \frac{D}{dt}s(t) &= \sum_{j=1}^k \frac{D}{dt}(s^j(t)e_j) \\ &= \sum_{j=1}^k \left((s^j)'(t)e_j + s^j(t)\nabla_{\alpha'(t)}e_j\right), \end{aligned} \tag{1}$$

as e_j are defined in a neighborhood. The right side does not depend on $\frac{D}{dt}$ so we have uniqueness.

For existence, we now have a formula: we write s in terms of a local basis for the sections and use equation (1) to define the covariant derivative along α. $\square$

Another notion of connection is called an *Ehresmann connection* and involves a splitting of the tangent bundle of E into "horizontal" and "vertical" sub-bundles. In particular, there is always a canonical sub-bundle V of TE (called the "vertical bundle") given by the kernel of the pushfoward map (i.e., the differential of the projection) $\pi_* : TE \to TM$. An Ehresmann connection is the data of a "horizontal" sub-bundle complementary to the vertical one, i.e., a sub-bundle $H \subset TE$ so that $TE = H \oplus V$. The notion of connection above induces such a splitting. (To ensure that it is equivalent to the definition above involves another condition that we omit here.)

Proposition 10. *A connection ∇ on E induces a splitting $TE = H \oplus V$.*

Proof. We must show that ∇ defines a horizontal sub-bundle $H \subset TE$ and that TE splits as the direct sum $H \oplus V$. We start by noting that, for each $e \in E$, $V_e = T_e E_{\pi(e)} \cong E_{\pi(e)}$ because the tangent space of a vector space is canonically isomorphic to the vector space.

We now aim to define the horizontal subspace. We first define a map $K : T_e E \to E_{\pi(e)}$ and then define H_e to be the kernel of K. Given $e \in E$ and $v \in T_e E$, choose $\gamma : (-\epsilon, \epsilon) \to E$ so that $\gamma(0) = e$ and $\gamma'(0) = v$. We now regard γ as a section of E over $\pi \circ \gamma$, i.e., $\gamma \in \Gamma(E, \pi \circ \gamma)$ and set

$$Kv = \frac{D}{dt}\gamma(t)|_{t=0}.$$

We claim that Kv is independent of the choice of γ. By linearity it suffices to show that if $\gamma : (-\epsilon, \epsilon) \to E$ has $\gamma(0) = e$ and $\gamma'(0) = 0$, then $\frac{D}{dt}\gamma(t)|_{t=0} = 0$. We then note that $(\pi \circ \gamma)'(0) = 0$ and appeal to equation (1) after writing γ in terms of a local frame for E to see that indeed $\frac{D}{dt}\gamma(t)|_{t=0} = 0$.

Now, equipped with the map $K : T_e E \to E_{\pi(e)}$, we define $H_e = \ker K$. We now claim that $T_e E \cong H_e + V_e$. Indeed, note that if $v \in V_e$ is a vertical vector, we use the identification $V_e \cong E_{\pi(e)}$ to construct the curve $\gamma : (-\epsilon, \epsilon) \to E_{\pi(e)}$ given by $\gamma(t) = e + tv$. This curve satisfies $\gamma(0) = e$ and $\gamma'(0) = v$ and $\frac{D}{dt}\gamma(t)|_{t=0} = v$, so $Kv = v$ for vertical vectors. The operator K can therefore be regarded as a projection onto V_e and so $T_e E \cong H_e \oplus V_e$.

The smoothness of the sub-bundles follows from the smoothness of the maps $(\pi)_*$ and K; that K depends smoothly on e is a consequence of the identity (1). $\square$

We now return to parallel transport. Given a curve $\alpha : [0, 1] \to M$ so that $\alpha(0) = p$ and $\alpha(1) = q$, we can construct the parallel translation of a vector $v \in E_p$ along α in two related ways. One way is by solving a differential equation: we say that a section $s \in \Gamma(E, \alpha)$ is parallel if and only if $\frac{D}{dt}s(t) = 0$ for all $t \in [0, 1]$.

Lemma 11. *For every $v \in E_p$, there is a unique $s \in \Gamma(E, \alpha)$ so that $s(0) = v$ and s is parallel.*

Proof. Working locally in charts where the bundle is trivial, this again follows from the identity (1), this time interpreted as a linear system of differential equations for the coefficients of the frame. It is not hard to check that existence and uniqueness for ODEs then guarantees a solution. $\square$

We then define the parallel translate of v by $P(\alpha)_0^1 v = s(1)$. Note that this value typically depends on the choice of path!

Another way to define the parallel translate of v along α to lift α to a curve $\widetilde{\alpha} : [0,1] \to E$ so that $\pi \circ \widetilde{\alpha} = \alpha$, $\alpha(0) = (p, v)$ and so that $\widetilde{alpha}'(t) \in H_{\widetilde{\alpha}(t)}$ for all t. The existence of such a lift follows from the observation that $H_e \cong T_{\pi(e)}M$ for all $e \in E$ and the decomposition of $TE = V \oplus H$. (You need $\pi_* \widetilde{\alpha}'(t) = \alpha'(t)$ and you need $\widetilde{\alpha}'(t)$ to be horizontal.)

We'll return to the question of why parallel transport is so called once we start talking about specific connections.

4.1. **Induced connections.** A connection ∇ on a vector bundle E over M induces connections over other bundles formed from E. A few examples:

(1) Dual bundles. If ∇ is a connection on E, we get a connection ∇^* (we'll later just call this ∇) on the dual bundle E^* by duality. Indeed, if $\xi \in \Gamma(E^*)$, we define $\nabla_v \xi$ by demanding that, for all $s \in \Gamma(E)$,
$$d\left(\langle \xi, s\rangle\right)(v) = \langle \nabla^*_v \xi, s\rangle + \langle \xi, \nabla_v s\rangle.$$
(2) If ∇^E and ∇^F are connections on vector bundles E and F over M, then we get a connection $\nabla^E \otimes \nabla^F$ on the vector bundle $E \otimes F$ over M by demanding it satisfy a product rule:
$$\left(\nabla^E \otimes \nabla^F\right)_v (s \otimes t) = \nabla^E_v s \otimes t + s \otimes \nabla^F_v t.$$
(3) A similar construction gives a connection on the exterior powers $\Lambda^k E$ by a product rule:
$$\nabla_v \left(s_1 \wedge \cdots \wedge s_k\right) = \nabla_v s_1 \wedge \cdots \wedge s_k + \cdots + s_1 \wedge \cdots \wedge \nabla_v s_k.$$
(4) Similarly we get a connection $\nabla^E \oplus \nabla^F$ on the direct sum bundle $E \oplus F$ by linearity:
$$\left(\nabla^E \oplus \nabla^F\right)_v (s \oplus t) = \left(\nabla^E_v s\right) \oplus \left(\nabla^F_v t\right).$$
(5) By identifying the endomorphisms of E with $E^* \otimes E$ we also get a connection on the endomorphism bundle $\mathrm{End}(E)$.

As a result, if we have a connection on the vector bundle TM then we in fact have a connection on all of the tensor bundles.

4.2. **The Levi-Civita connection.** We now specialize to the case where $E = TM$ and its associated vector bundles.

Definition 12. Given a connection ∇ on TM, its *torsion* is given by
$$T(\nabla)(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \in C^\infty(M)$$
where $X, Y \in \mathcal{X}(M)$ and $[X, Y]$ is the Lie bracket.

Lemma 13. *The torsion of a connection on TM is an antisymmetric $(0,2)$-tensor.*

Proof. It's clearly antisymmetric; that it is a tensor follows from the characterization of tensor fields given by Lemma 1. Indeed, we have to check that it is multilinear over $C^\infty(M)$. Take $f \in C^\infty(M)$, $X, Y \in \mathcal{X}(M)$ and consider
$$\begin{aligned} T(\nabla)(fX, Y) &= \nabla_{fX} Y - \nabla_Y (fX) - [fX, Y] \\ &= f\nabla_X Y - (Y(f)X + f\nabla_Y X) - (f[X, Y] - Y(f)X) \\ &= fT(\nabla)(X, Y). \end{aligned}$$
□

Theorem 14. *Suppose (M, g) is a Riemannian manifold. There is a unique connection ∇ so that*

(1) ∇ is torsion-free (i.e., $\nabla_X Y - \nabla_Y X - [X, Y] = 0$), and
(2) ∇ is compatible with the metric g, i.e., for all $X, Y, Z \in \mathcal{X}(M)$,

$$X\left(g(Y, Z)\right) = g\left(\nabla_X Y, Z\right) + g\left(Y, \nabla_X Z\right).$$

Definition 15. The unique torsion-free connection compatible with g is called the *Levi-Civita connection.*

Proof. To prove uniqueness, suppose that ∇ is torsion-free and compatible with g. Then compatibility yields

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

while ∇ being torsion-free lets us write, e.g., $\nabla_X Z - \nabla_Z X = [X, Z]$, so that the above expression is equal to (here using that g is symmetric)

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

In other words, we end up with the equality

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

Because g is non-degenerate and this relationship must hold for all Z, any two choices of torsion-free metric-compatible connection must agree.

Similarly, because g is non-degenerate, the formula above also defines the connection ∇, so we also get existence. (You should check that it is tensorial in X and Z and satisfies a product rule in Y and should also check that it is torsion-free and metric-compatible.) $\square$

How does this connection look in local coordinates? Let $(x^1, \dots, x^n)$ be coordinates in a chart on M and ∂_j denote the corresponding basis for the tangent space. We then have

$$\begin{aligned} \langle \nabla_{\partial_i} \partial_j, \partial_k\rangle &= \frac{1}{2}\left[\partial_i\langle \partial_j, \partial_k\rangle + \partial_j\langle \partial_i, \partial_k\rangle - \partial_k\langle \partial_i, \partial_j\rangle\right] \\ &= \frac{1}{2}\left[\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right]. \end{aligned}$$

If we write $\nabla_{\partial_i}\partial_j = \sum_\ell \Gamma_{ij}^\ell \partial_\ell$, then

$$\begin{aligned} \sum_{\ell=1}^n \Gamma_{ij}^\ell g_{\ell k} &= \sum_{\ell=1}^n \Gamma_{ij}^\ell \langle \partial_\ell, \partial_k\rangle \\ &= \frac{1}{2}\left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right). \end{aligned}$$

This is an $n \times n$ system of linear equations, which we solve to find

$$\Gamma_{ij}^k = \sum_{\ell=1}^n \frac{1}{2} g^{k\ell}\left(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}\right),$$

where $g^{k\ell}$ are the components of the inverse metric (i.e., the inverse of the matrix (g_{ij}) or the components of the induced metric on the cotangent bundle, etc.).

Definition 16. The Γ_{ij}^k are called *Chrisoffel symbols* or *connection coefficients.*

Word of warning: Γ_{ij}^k are NOT the components of a tensor (they lie in an affine space).

As described above, ∇ induces a connection on the cotangent bundle, given for $\omega \in \Omega^1(M)$ by

$$(\nabla_v \omega)(X) = v(\omega(X)) - \omega(\nabla_v X).$$

For more general tensors it has an analogous form, with $\nabla_v A$ of a (k, ℓ)-tensor given by

$$\begin{aligned}(\nabla_v A)(\omega_1, \ldots, \omega_k, X_1, \ldots, X_\ell) = v\left(A(\omega_1, \ldots, \omega_k, X_1, \ldots, X_\ell)\right) \\ - \sum_{i=1}^k A\left(\omega_1, \ldots, \omega_{i-1}, \nabla_v \omega_i, \omega_{i+1}, \ldots, \omega_k, X_1, \ldots, X_\ell\right) \\ - \sum_{j=1}^\ell A\left(\omega_1, \ldots, \omega_k, X_1, \ldots, X_{j-1}, \nabla_v X_j, X_{j+1}, \ldots, X_\ell\right).\end{aligned}$$

Proposition 17. *Parallel transport using the Levi-Civita connection is an isometry.*

Proof. Suppose $\alpha : [0, 1] \to M$ is smooth and $V, W \in \mathcal{X}(\alpha)$. Then

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

so that if V and W are parallel then their inner product is preserved. □

5. Geodesics and Hamiltonian flows

Most differential geometry textbooks use the connection directly to define and reason about geodesics. We'll instead take a Hamiltonian approach. To do that, we need some preliminaries about the symplectic structure on the cotangent bundle.

5.1. Symplectic manifolds.

Definition 18. A manifold (M, ω) is called a symplectic manifold if M is a smooth manifold and ω is a non-degenerate closed 2-form on M.

In other words, at each point ω is an alternating $(0, 2)$-tensor so that $d\omega = 0$ and, if $v \in T_p M$ satisfies

$$\omega(v, u) = 0 \text{ for all } u \in T_p M,$$

then $v = 0$.

Lemma 19. *A symplectic manifold must be even-dimensional.*

Proof. Working in local coordinates, this reduces to the statement that if n is odd, then any skew-symmetric $n \times n$ real matrix must have a kernel. To see this, a skew-symmetric matrix A has

$$\det A = \det(A^{\intercal}) = \det(-A) = (-1)^n \det A,$$

so that $\det A = 0$ if n is odd and thus 0 is an eigenvalue of A, i.e., A must have a kernel. □

There is a lot to say about symplectic manifolds, most of which we omit here. One of the most famous is Darboux's theorem, which essentially says that there are no local invariants of symplectic manifolds, i.e., one can always find local coordinates (q, p) so that $\omega = \sum_i dq^i \wedge dp_i$. There's a lovely proof of this using Moser's trick which I'm happy to talk about if you like:

Theorem 20 (Darboux)**.** *Suppose (M, ω) is a symplectic $2k$-dimensional manifold. Around any point $p \in M$ there is a coordinate chart U and coordinates $(x^1, \dots, x^k, y^1, \dots, y^k)$ so that*

$$\omega|_U = \sum_{i=1}^k dx^i \wedge dy^i.$$

Proof. You can always find a coordinate system $(\widetilde{x}^1, \dots, \widetilde{x}^k, \widetilde{y}^1, \dots, \widetilde{y}^k)$ achieving this exactly at the point p. We claim then that there is a local diffeomorphism ϕ of a smaller neighborhood so that

$$\phi^* \left(\sum d\widetilde{x}^i \wedge d\widetilde{y}^i \right) = \omega|_U,$$

and then the desired coordinate system is $x = \widetilde{x} \circ \phi$ and $y = \widetilde{y} \circ \phi$.

It remains to prove the claim. We first claim that if ω_t is a family of symplectic forms so that $\frac{d}{dt}\omega_t = d\sigma_t$ for some 1-forms σ_t, then there is a family of diffeomorphisms ψ_t so that $\psi_t^* \omega_t = \omega_0$. Indeed, this follows from Lemma 21 below by using Cartan's magic formula

$$\mathcal{L}_X \omega = d \circ i_X + i_X \circ d,$$

where i_X denotes the interior product (i.e., plugging X into the first slot) and observing that the differential equation in Lemma 21 becomes

$$0 = \frac{d}{dt}\omega_t + d(i_X \omega_t) + i_X(d\omega_t) = d(\sigma_t + i_X \omega_t).$$

As ω_t is non-degenerate, one can find X so that $\sigma_t = i_X \omega_t$ and so Lemma 21 applies.

Finally, by the Poincaré lemma (which says that on $\mathbb{R}^n$, closed forms are exact), the difference $\omega_t - \omega_0 = d\sigma_t$ and so we can apply the claim. □

Lemma 21 (Moser's trick)**.** *Suppose ω_t, $t \in [0, 1]$ is a family of differential forms on M. If there is a solution $X_t \in \mathcal{X}(M)$, $t \in [0, 1]$ to the differential equation*

$$\frac{d}{dt}\omega_t + \mathcal{L}_{X_t}\omega_t = 0,$$

where $\mathcal{L}_X$ denotes the Lie derivative with respect to X, then there exists a family of diffeomorphisms ψ_t on M so that $\psi_t^* \omega_t = \omega_0$ and $\psi_0 = \mathrm{Id}$.

Proof. Given X_t, let ψ_t be the flow it generates, so

$$\frac{d}{dt}(\psi_t^* \omega_t) = \psi_t^* \left(\frac{d}{dt}\omega_t + \mathcal{L}_{X_t}\omega_t \right) = 0,$$

so $\psi_t^* \omega_t = \psi_0^* \omega_0 = \omega_0$. □

Now, back on track. One of the main things we want to use symplectic structures for is to get *Hamilton vector fields*. Just as we had with Riemannian metrics, we can use the nondegenerate 2-form ω to identify the tangent and cotangent spaces at each point. Indeed, given any 1-form σ, we can find some vector field X associated to it by demanding that

$$\omega(v, X) = \sigma_p(v)$$

for all vectors $v \in T_pM$. In particular, if we have a real-valued function H (called a Hamiltonian) on M, we can find the *Hamilton vector field* of H, which we'll denote X_H, by demanding that X_H be the vector field associated to dH by ω.

As smooth vector fields yield flows, we therefore also obtain a flow ϕ_t from the Hamilton vector field X_H.

Lemma 22. *The Hamiltonian H is conserved by the flow ϕ_t.*

Proof. This is an exercise in unraveling the definitions and using the chain rule. Indeed, suppose $p \in M$ and let $\gamma(t) = \phi_t(p)$, so that

$$\begin{aligned} \gamma'(t) &= (X_H)_{\gamma(t)}, \\ \gamma(0) &= p. \end{aligned}$$

We then differentiate $H(\gamma(t))$:

$$\begin{aligned} \frac{d}{dt} H(\gamma(t)) &= dH(\gamma'(t)) \\ &= dH(X_H) = \omega(X_H, X_H) = 0, \end{aligned}$$

so that $H(\gamma(t)) = H(\gamma(0))$. □

5.2. **A very brief foray into Hamiltonian mechanics.** One of the most important examples of a symplectic manifold is the cotangent bundle T^*M of a smooth manifold M. (Note that the dimension of T^*M is always twice the dimension of M and therefore even.) If $\pi : T^*M \to M$ denotes the projection, we can define the *canonical* 1*-form* α on T^*M by

$$\alpha_{(x,\xi)}(v) = \xi(\pi_* v)$$

where $(x, \xi) \in T^*M$ and $v \in T_{(x,\xi)}(T^*M)$. In other words, the form α acts on a vector v at a point (x, ξ) in the cotangent bundle by evaluating the covector ξ (a covector on M) on the pushvorward of v. In terms of local coordinates (x, ξ),[1] you can check that

$$\alpha = \sum_j \xi_j dx^j.$$

We then define a symplectic form ω by $\omega = d\alpha$. In a local coordinate system ω has the form

$$\omega = \sum_J d\xi_j \wedge dx^j.$$

It is plainly a 2-form, $d\omega = d(d\alpha) = 0$, and there are several ways to check that it is non-degenerate. One way is to observe that ω^n, the n-th wedge power of ω, is a non-vanishing volume form. We can also check it directly by taking $v \in T_{(x,\xi)}(T^*M)$ with $\omega(v, \bullet) = 0$. Writing v in terms of the basis given by the coordinate system, we write

$$v = \sum_j v^j \frac{\partial}{\partial x^j} + \sum_j w^j \frac{\partial}{\partial \xi_j},$$

so that

$$\omega(v, \bullet) = -\sum_j v^j d\xi_j + \sum_j w^j dx^j,$$

so that we must have $v^j = w^j = 0$, i.e., $v = 0$.

[1] Recall that a coordinate system x on M induces a coordinate chart (x, ξ) on T^*M by writing covectors in terms of the basis dx^J of each cotangent space; ξ_j are the coefficients here.

From a physical perspective, the symplectic manifold (T^*M, ω) is thought of as a "phase space" for a physical system taking place on the "configuration space" M while a Hamiltonian H is the total energy of the system. If $M = \mathbb{R}$ and

$$H(x,\xi) = \frac{1}{2m}\left|\xi\right|^2 + V(x),$$

(i.e., H is kinetic energy plus potential energy), then

$$dH = \frac{1}{m}\xi\, d\xi + V'(x)dx,$$

so that

$$X_H = \frac{1}{m}\xi\frac{\partial}{\partial x} - V'(x)\frac{\partial}{\partial \xi},$$

and thus the integral curves of the flow generated by X_H satisfy

$$\frac{dx}{dt} = \frac{1}{m}\xi, \quad \frac{d\xi}{dt} = -V'(x).$$

In particular, x satisfies the second-order differential equation

$$x''(t) = -\frac{1}{m}V'(x(t)),$$

which you might recognize as Newton's second law for the conservative force given by the potential $V(x)$.

5.3. **Geodesics.** Suppose now that (M, g) is a Riemannian manifold and (in a mild abuse of notation) let g^{-1} denote the induced inner product on each cotangent space. Recall from the last section that T^*M is always a symplectic manifold and consider the Hamiltonian function

$$H(x,\xi) = \frac{1}{2}\left|\xi\right|^2_{g^{-1}} = \frac{1}{2}\sum_{i,j} g^{ij}(x)\xi_i\xi_j.$$

By the discussion above, we associate to H a Hamilton vector field X_H and a (very important!) flow ϕ_t.

Definition 23. We say that ϕ_t is the *geodesic flow on the cotangent bundle* and the integral curves of X_H are called *lifted geodesics*. If $\gamma(t)$ is a lifted geodesic, its projection $\pi\circ\gamma$ to M is called a *geodesic*.

In local coordinates (x, ξ) on the cotangent bundle, the Hamilton vector field of H is given by

$$\begin{aligned} X_H &= \sum_i g^{ii}(x)\xi_i\partial_{x_i} + \frac{1}{2}\sum_{i\neq j} g^{ij}(x)\xi_i\partial_{x_j} - \frac{1}{2}\frac{\partial g^{ij}(x)}{\partial x_k}\xi_i\xi_j\partial_{\xi_k} \\ &= w_\xi - \frac{1}{2}\frac{\partial g^{ij}(x)}{\partial x_k}\xi_i\xi_j\partial_{\xi_k}, \end{aligned}$$

where, in another abuse of notation, w_ξ is the vector field on M (regarded as a vector field on T^*M) associated to ξ by the metric g. In particular, the integral curves $(x(t), \xi(t))$ of X_H satisfy

$$\frac{dx}{dt} = w_\xi, \quad \frac{d\xi}{\tilde{d}} = -\frac{1}{2}\sum \frac{\partial g^{ij}(x(t))}{\partial \xi}\xi_i\xi_j.$$

Before we get to examples, let's connect[2] this computation with our discussion of connections from earlier.

Lemma 24. *If $\widetilde{\gamma} : (a, b) \to T^*M$ is a lifted geodesic, then $\gamma = \pi \circ \widetilde{\gamma}$ satisfies*

$$\frac{D}{dt}\gamma'(t) = 0$$

along γ.

Proof. □

[2] HA!