CONNECTIONS AND CURVATURE NOTES

EUGENE LERMAN

Contents

1. Connections on vector bundles

1.1. **Connections.** If X is a vector field on an open subset U of $\mathbb{R}^m$, then X is determined by m-tuple $(a_1, \dots a_m)$ of functions:

$$X = \sum_i a_i \frac{\partial}{\partial x_i}$$

Therefore we know how to take directional derivatives of X at a point $q \in U$ in the direction of a vector $v \in T_q U = \mathbb{R}^m$ — we simply differentiate the coefficients:

$$(D_v X)_q = \sum_i (D_v a_i)_q \frac{\partial}{\partial x_i}|_q$$

where $D_v a_i$ is the directional derivative of the function a_i in the direction v. Consequently we know when a vector field does not change along a curve γ:

$$D_{\dot\gamma} X = 0.$$

Covariant derivatives generalize the directional derivatives allowing us to differentiate vector fields on arbitrary manifolds and, more generally, sections of arbitrary vector bundles.

Definition 1.1 (Covariant derivative of sections of a vector bundle)**.** Let $\pi : E \to M$ be a vector bundle. A *covariant derivative* (also knows as a *connection*) is an $\mathbb{R}$-bilinear map

$$\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E), \quad (X, s) \mapsto \nabla_X s$$

such that

(1) $\nabla_{fX} s = f \nabla_X s$
(2) $\nabla_X (fs) = X(f) \cdot s + f \nabla_X s$.

for all $f \in C^\infty(M)$, all $X \in \Gamma(TM)$, and all $s \in \Gamma(E)$.

Example 1.2. Let $U \subset \mathbb{R}^m$ be an open set and $E = TU \to U$ the tangent bundle. Define a connection D on $TU \to U$ by

$$D_X (\sum a_i \frac{\partial}{\partial x_i}) = \sum X(a_i) \frac{\partial}{\partial x_i}.$$

I leave it to the reader to check that this is indeed a connection.

Remark 1.3. Lie derivative $(X, Y) \mapsto L_X Y$ is *not* a connection on the tangent bundle (why not?).

Example 1.4. Let $\pi : E \to M$ be a trivial bundle of rank k. Then there exist global sections $\{s_1, \dots, s_k\}$ of E such that $\{s_j(x)\}$ is a basis for E_x for all points $x \in M$ ($\{s_i\}$ is a *frame* of $E|_U$). So for any $s \in \Gamma(E)$, we have $s = \sum_j f_j s_j$, for some C^∞ functions f_j. We define a bilinear map $\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$ by

$$\nabla_X s = \nabla_X(\sum_j f_j s_j) := \sum_j X(f_j) s_j.$$

It is easy to check that ∇ is indeed a connection on E:

$$\nabla_{fX} s = \nabla_{fX}(\sum_j f_j s_j) = \sum_j fX(f_j)s_j = f \sum_j f_j s_j = f\nabla_X s;$$

and

$$\nabla_X(fs) = \nabla_X(f \sum_j f_j s_j) = \sum_j X(ff_j)s_j = X(f) \sum_j f_j s_j + f \sum_j X(f_j)s_j = X(f)s + f\nabla_X s.$$

Lemma 1.5. *Any convex linear combination of two connections on a vector bundle $E \to M$ is a connection. More precisely, let ∇^1, ∇^2 be two connections on E and $\rho_1, \rho_2 \in C^\infty(M)$ be two functions with $\rho_1 + \rho_2 = 1$. Then*

$$\Gamma(TM) \times \Gamma(E) \ni (X, s) \mapsto \nabla_X s := \rho_1 \nabla^1_X s + \rho_2 \nabla^2_X s \in \Gamma(E)$$

is a connection.

Proof. Exercise. Check that the two properties of the connection hold. $\square$

As a corollary we get:

Proposition 1.6. *Any vector bundle $\pi : E \to M$ has connection.*

Proof. Choose a cover $\{U_\alpha\}$ on M such that $E|_{U_\alpha}$ is trivial. Let ∇^α be a connection on $E|_{U_\alpha}$, as in Example 1.4. Let $\{\rho_\beta\}$ be a partition of unity subordinate to $\{U_\alpha\}$. Then $\operatorname{supp} \rho_\beta \subset U_\alpha$ for some $\alpha = \alpha(\beta)$. Define a map $\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$ by

$$\nabla_X s = \sum_\beta \rho_\beta (\nabla^\alpha_{X_U} s|_{U_\alpha}).$$

This is indeed a connection, since a convex linear combination of any finite number of connections is a connection — see Lemma 1.5 above. $\square$

Proposition 1.7. *Let ∇ be a connection on a vector bundle $\pi : E \to M$. Then ∇ is local: for any open set U and any vector fields X and Y, and for any sections s and s' of E such that $X|_U = Y|_U$ and $s|_U = s'|_U$, we have*

$$(\nabla_X s)|_U = (\nabla_Y s')|_U.$$

Proof. Since ∇ is bilinear, it is enough to show two things:
(a) if $X|_U = 0$, then $(\nabla_X s)|_U = 0$ for any $s \in \Gamma(E)$; and
(b) if $s|_U = 0$, then $(\nabla_X s)|_U = 0$ for any $X \in \Gamma(TM)$.
Fix a point $x_0 \in U$. Then there is a smooth function $\rho : U \to [0, 1]$ with $\operatorname{supp} \rho \subset U$ and $\rho|_V = 1$ for some open neighborhood V of x_0. If $X|_U = 0$ then $\rho X = 0$, and hence for any section s of E,

$$0 = (\nabla_{\rho X} s)(x_0) = \rho(x_0)(\nabla_X s)(x_0) = (\nabla_X s)(x_0).$$

Since $x_0 \in U$ is arbitrary, (a) follows. If $s|_U = 0$ then $\rho s = 0$ on M. This in turn implies that

$$0 = (\nabla_X \rho s)(x_0) = (X(\rho)s + \rho\nabla_X s)(x_0) = 0 + \rho(x_0)(\nabla_X s)(x_0) = (\nabla_X s)(x_0).$$

$\square$

Remark 1.8. It follows that if ∇ is a connection on a vector bundle $E \to M$ then ∇ induces a connection

$$\nabla^U : \Gamma(TU) \times \Gamma(E_U) \to \Gamma(E|_U)$$

on the restriction $E|_U$ for any open set $U \subset M$. Namely, for any $x_0 \in U$ let $\rho : U \to [0,1]$ be a bump function as in the proof above. Then for any $X \in \Gamma(TU)$ and any $s \in \Gamma(E|_U)$ we have $\rho X \in \Gamma(TM)$ and $\rho s \in \Gamma(E)$ (with ρX and ρs extended to all of M by 0). We define:

$$({\nabla^U}_X s)(x_0) = (\nabla_{\rho X} \rho s)(x_0).$$

By Proposition 1.7, the right hand side does not depend on the choice of the function ρ. We leave it to the reader to check that ∇^U is a connection.

Definition 1.9 (Christoffel symbols)**.** Let $E \to M$ be a vector bundle with a connection ∇. Let $(x_1, \dots, x_n) : U \to \mathbb{R}^m$ be a coordinate chart on M small enough so that $E|_U$ is trivial. Let $\{s_\alpha\}$ be a frame of $E|_U$: for each $x \in U$ we require that $\{s_\alpha(x)\}$ is a basis of the fiber E_x. Then any local section $s \in \Gamma(E|_U)$ can be written as a linear combination of s_α's. In particular, for each index i and β

$$\nabla^U_{\frac{\partial}{\partial x_i}} s_\beta = \sum_\alpha \Gamma^\alpha_{i\beta} s_\alpha$$

for some functions $\Gamma^\beta_{i\alpha} \in C^\infty(U)$. These functions are the *Christoffel symbols* of the connection ∇ relative to the coordinates $(x_1, \dots, x_n)$ and the frame $\{s_\alpha\}$.

It follows easily that the Christoffel symbols determine the connection ∇^U on the coordinate chart U. It is customary not to distinguish between ∇ and its restriction ∇^U.

Proposition 1.10. *Let ∇ be a connection on on a vector bundle $\pi : E \to M$. For any $X \in \Gamma(TM)$, any $s \in \Gamma(E)$ and any point q the value of the connection $(\nabla_X s)(q)$ at a point $q \in M$ depends only on the vector X_q (and not on the value of X near q).*

Proof. It's enough to show that if $X_q = 0$ then $(\nabla_X s)(q) = 0$. Since connections are local we can argue in coordinates. Choose a coordinate chart $(x_1, \dots, x_n) : U \to \mathbb{R}^m$ on M with $q \in U$ such that $E|_U$ is trivial. Pick a local frame $\{s_j\}$ of $E|_U$. Then, if $X = \sum X^i \frac{\partial}{\partial x_i}$, $s = \sum f_j s_j$, and Γ^k_{ij} denote the associated Christoffel symbols,

$$\begin{aligned}
\nabla_X s = \nabla_{\sum X^i \frac{\partial}{\partial x_i}} (\sum f_j s_j) &= \sum X^i \nabla_{\frac{\partial}{\partial x_i}} (\sum f_j s_j) \\
&= \sum X^i \frac{\partial f_j}{\partial x_i} s_j + \sum X_i f_j \nabla_{\frac{\partial}{\partial x_i}} s_j \\
&= \sum X^i (\sum \frac{\partial f_j}{\partial x_i} s_j + \sum f_j \Gamma^k_{ij} s_k).
\end{aligned}$$

If $X_q = 0$ then $X^i(q) = 0$ for all i. Hence $(\nabla_X s)(q) = 0$ and we are done. $\square$

As a corollary of the proof computation above we get an expression for the connection in terms of the Christoffel symbols.

Corollary 1.10.1. *Let ∇ be a connection on on a vector bundle $\pi : E \to M$ and $(x_1, \dots, x_n) : U \to \mathbb{R}^m$ a coordinate chart on M with $E|_U$ being trivial. Let $\{s_j\}$ be a frame of $E|_U$. Then*

$$\nabla_{\sum_i X^i \frac{\partial}{\partial x_i}} (\sum_j f_j s_j) = \sum_{i,k} X^i (\frac{\partial f_k}{\partial x_i} + \sum_j f_j \Gamma^k_{ij}) s_k. \tag{1.1}$$

We note one more corollary that will be useful when we try to define connections induced on submanifolds.

Corollary 1.10.2. *Let ∇ be a connection on on a vector bundle $\pi : E \to M$. For any $X \in \Gamma(TM)$, any $s \in \Gamma(E)$ and any point q the value of the connection $(\nabla_X s)(q)$ at a point $q \in M$ depends only on the values of s along the integral curve of X through q*

Proof. By the previous corollary, for $X = \sum_i X^i \frac{\partial}{\partial x_i}$ and $s = \sum_j f_j s_j$

$$(\nabla_X s)(q) = (Xf_k)(q)\; s_k(q) + \sum_{i,k,j} X^i(q) f_j(q) \Gamma^k_{ij}(q) s_k(q).$$

And $(Xf_k)(q)$ depends only on the values of f_k along the integral curve of X. □

The proof that connections are local has an important generalization to maps of sections of vector bundles.

Definition 1.11. Let $E \to M$ and $F \to M$ be two vector bundles. We say that a map $T : \Gamma(E) \to \Gamma(F)$ is *tensorial* if T is $\mathbb{R}$-linear and for any $f \in C^\infty(M)$

$$T(fs) = fT(s)$$

for all sections $s \in \Gamma(E)$.

Lemma 1.12. *Let $E \to M$ and $F \to M$ be two vector bundles. If $T : \Gamma(E) \to \Gamma(F)$ is tensorial then there is a vector bundle map $\phi : E \to F$ so that*

$$[T(s)](x) = \phi(s(x))$$

for all $s \in \Gamma(E)$ and $x \in M$. And conversely, any vector bundle map $\phi : E \to F$ defines a tensorial map on sections $T_\phi : \Gamma(E) \to \Gamma(F)$ by $T_\phi(s) = \phi \circ s$.

Proof. The proof is in two steps. We first argue that T is local: if $s \in \Gamma(E)$ vanishes on an open set $U \subset M$ then $T(s)$ vanishes on U as well. Pick a point $x \in U$ and a smooth function $\rho \in C^\infty(M)$ with $\operatorname{supp} \rho \subset U$ and $\rho \equiv 1$ on a neighborhood V of x ($V \subset U$, of course). Then ρs is identically zero everywhere. Hence

$$0 = T(\rho s)(x) = \rho(x)\, T(s)(x) = T(s)(x).$$

Since $x \in U$ is arbitrary $T(s)|_U = 0$.

Since T is local and E, F are locally trivial, we may assume that E and F are, in fact, trivial. That is $E = M \times \mathbb{R}^k$ and $F = M \times \mathbb{R}^l$. Moreover the sections of E and F are simply k- and l-tuples of functions. We want to define a vector bundle map $\phi : E \to F$. Then $\phi : M \times \mathbb{R}^k \to M \times \mathbb{R}^l$ has to be of the form

$$\phi(x, v) = (x, A(x)v)$$

where $A : M \to \mathrm{Hom}(\mathbb{R}^k, \mathbb{R}^l)$ is smooth, with the property that

$$T(f_1, \ldots, f_k)(x) = A(x) \begin{pmatrix} f_1(x) \\ \vdots \\ f_k(x) \end{pmatrix}$$

for all $x \in M$. But this is easy: define the jth column of $A(x)$ to be the l-tuple of functions $T(e_j)$, where e_j is the section of E that assigns to every point the jth basis vector $(0, \ldots, 0, 1, 0, \ldots, 0)$ (1 in jth slot). Or, if you prefer, e_j is the k-tuple of functions with jth function being identically 1 and all the others being zero. □

Remark 1.13. Lemma 1.12 above generalizes further: let E_1, $E_2, \ldots E_k$ and F be vector bundles over a manifold M and

$$T : \Gamma(E_1) \times \cdots \times \Gamma(E_k) \to \Gamma(F)$$

a k-linear map which is tensorial in each slot:

$$T(f_1 s_1, \ldots, f_k s_k) = f_1 \ldots f_k T(s_1, \ldots, s_k)$$

for all $s_i \in \Gamma(E_i)$ and $f_j \in C^\infty(M)$. Then for every $x \in M$ there is a unique k-linear map

$$T_x : (E_1)_x \times \cdots \times (E_k)_x \to F_x$$

with

$$T_x(s_1(x), \ldots, s_k(x)) = [T(s_1, \ldots, s_k)](x).$$

Globally this means that there is a vector bundle map

$$\phi : E_1 \otimes \cdots \otimes E_k \to F$$

so that
$$T(s_1,\ldots,s_k)(x) = \phi(s_1(x)\otimes\ldots\otimes s_k(x))$$
for all $x\in M$ and all sections $s_i\in\Gamma(E_i)$.

Remark 1.14. We add one more layer of abstraction to the remark above: there is a bijection between vector bundle maps $\phi: E\to F$ and sections of the bundle $\mathrm{Hom}(E,F)\simeq E^*\otimes F$. Namely, if $\phi: E\to F$ is a vector bundle map, then $\phi|_{E_x}: E_x\to F_x$ is an element of $\mathrm{Hom}(E_x,F_x)=\mathrm{Hom}(E,F)_x$ for each point $x\in M$. Thus $x\mapsto \phi|_{E_x}$ is a section of the bundle $\mathrm{Hom}(E,F)\to M$.

We summarize the preceding discussion as a proposition.

Proposition 1.15. *Let E_1, $E_2,\ldots E_k$ and F be vector bundles over a manifold M. There is a bijection between k-linear tensorial maps*
$$T:\Gamma(E_1)\times\cdots\times\Gamma(E_k)\to\Gamma(F)$$
and the sections of the bundle $E_1^*\otimes\cdots\otimes E_k^*\otimes F\to M$.

Here are a few instances where the above point of view is useful.

Lemma 1.16. *Let ∇^1 and ∇^2 be two connections on a vector bundle $E\to M$. Their difference $\nabla^1-\nabla^2$ "is" a section of the bundle $T^*M\otimes E^*\otimes E\simeq Hom(TM\otimes E,E)$. Conversely, given a connection ∇ on $E\to M$ and a section A of the bundle $\mathrm{Hom}(TM\otimes E,E)$ then the map $\nabla^A:\Gamma(TM)\times\Gamma(E)\to\Gamma(E)$ given by*
$$(\nabla^A_X s)(x) := \nabla_X s(x) + A_x(X_x\otimes s(x))$$
is again a connection on E. Here, of course, $x\in M$ is a point, X a vector field on M and s is a section of E. Thus a choice of a connection on $E\to M$ defines a bijection

$\{$*space of all connections on $E\to M$*$\}\leftrightarrow\Gamma(T^*M\otimes E^*\otimes E)=\Gamma(\mathrm{Hom}(TM\otimes E,E))=\Gamma(T^*M\otimes\mathrm{Hom}(E,E))$.

Proof. In one direction it's enough to prove that $\nabla^1-\nabla^2$ is tensorial in both slots. It's obviously tensorial in the vector field slot. The tensoriality in the second slot is an easy computation.

We also leave it to the reader to check that ∇^A as defined above is a connection. □

Definition 1.17. A *connection on a manifold* M is a connection on its tangent bundle $TM\to M$.

Definition 1.18. The *torsion* T^∇ of a connection ∇ on a manifold M is a bilinear map
$$T^\nabla:\Gamma(TM)\times\Gamma(TM)\to\Gamma(TM),\quad T^\nabla(X,Y):=\nabla_XY-\nabla_YX-[X,Y].$$
If $T^\nabla=0$, the connection ∇ is called *torsion-free*.

Lemma 1.19. *The torsion of a connection is tensorial, hence corresponds to a section of the bundle $T^*M\otimes T^*M\otimes TM$.*

Proof. This is yet another computation left to the reader. □

Definition 1.20. The *curvature* R of a connection ∇ on a vector bundle $E\to M$ is a tri-linear map $\Gamma(TM)\times\Gamma(TM)\times\Gamma(E)\to\Gamma(E)$ defined by
$$R(X,Y)s=\nabla_X(\nabla_Ys)-\nabla_Y(\nabla_Xs)-\nabla_{[X,Y]}s.$$

Lemma 1.21. *Curvature is tensorial hence correspond to a section of $T^*M\otimes T^*M\otimes\mathrm{Hom}(E,E)\to M$. Moreover, since $R(X,Y)s=-R(Y,X)s$, it actually corresponds to a section of $\Lambda^2(T^*M)\otimes\mathrm{Hom}(E,E)$.*

Proof. Once again this is a computation. We check tensoriality in one slot and leave the rest to the reader. For all vector fields X,Y, sections s and functions f,
$$\begin{aligned}
R(X,Y)(fs) &= \nabla_X(\nabla_Y(fs))-\nabla_Y(\nabla_X(fs))-\nabla_{[X,Y]}(fs)\\
&= \nabla_X(Y(f)s+f\nabla_Ys)-\nabla_Y(X(f)s-f\nabla_Xs)-([X,Y]f)s-f\nabla_{[X,Y]}s\\
&= X(Y(f))s+Y(f)\nabla_Xs+X(f)\nabla_Ys+f\nabla_X(\nabla_Ys)-Y(X(f))s\\
&\quad -X(f)\nabla_Ys-Y(f)\nabla_Xs-f\nabla_X(\nabla_Ys)-([X,Y]f)s-f\nabla_{[X,Y]}s\\
&= fR(X,Y)s
\end{aligned}$$
□

1.2. **Parallel Transport.** In general there is no consistent way of identifying vectors in tangent spaces at different points of a manifold. More generally there is no consistent way of identifying vectors in fibers of a vector bundle above different points of a manifold. However we will see that given a connection ∇ on a vector bundle $\pi : E \to M$, for any curve $\gamma : [a, b] \to M$ there is a family of vector space isomorphisms

$$P_{t_1}^{t_2}(\gamma) = P_{t_1}^{t_2} : E_{\gamma(t_1)} \to E_{\gamma(t_2)},$$

depending smoothly on $t_1, t_2 \in [a, b]$. These isomorphisms $P_{t_1}^{t_2}$ are called *parallel transport* along γ. The connection can then be recovered from parallel transport. We now proceed with the construction.

Definition 1.22. Let $\pi : E \to M$ be a vector bundle and $\gamma : [a, b] \to M$ a curve. A *section σ of $E \to M$ along γ* is a smooth map $s : [a, b] \to E$ so that $\pi(\sigma(t)) = \gamma(t)$ for all $t \in [a, b]$. We denote the space of sections of E along the map γ by $\Gamma(\gamma^*E)$.

Example 1.23. If $s : M \to E$ is a section of E, then $s \circ \gamma$ is a section along γ.

Example 1.24. The derivative $\dot\gamma := d\gamma_t(\frac{d}{dt}|_t)$ is a section of the tangent bundle $TM \to M$ along γ.

Remark 1.25. If $E = TM$ then a section along a curve γ is also known as *a vector field along* γ. It's not true that every section σ along γ is of the form $\sigma = s \circ \gamma$ for some $s \in \Gamma(E)$: if the curve γ crosses itself than $\dot\gamma$ cannot be of the form $X \circ \gamma$ for any vector field X on M.

Remark 1.26. Here's another way to consider sections along a curve γ. Suppose $f : N \to M$ is a smooth map of manifolds and that $\pi : E \to M$ is a vector bundle. Define the *pullback of the bundle E along f* to be the set

$$f^*E = \{(n, e) \in N \times E \mid f(n) = \pi(e)\}.$$

together with the projection $\pi' : f^*E \to N$, $f^*E \ni (n, e) \mapsto n$. A transversality argument shows that f^*E is a submanifold of $N \times E$, so π' is smooth. It's not hard to see that f^*E is a vector bundle of the same rank as E. The point of this construction is that a section of a bundle $E \to M$ along a curve $\gamma : (a, b) \to M$ is simply a section of the pullback bundle $\gamma^*E \to [a, b]$.

Strictly speaking the construction above doesn't apply to maps from closed intervals, since a closed interval is not a manifold. However, a smooth map from a closed interval $[a, b]$ is, by definition, a smooth curve from a slightly larger open interval $(a', b') \supset [a, b]$ and pulling back E to a bundle over (a', b') does make sense.

Definition 1.27. Let $\pi : E \to M$ be a vector bundle and $\gamma : [a, b] \to M$ a smooth curve. A *covariant derivative $\frac{\nabla}{dt}$ along* γ is an $\mathbb{R}$-linear map

$$\frac{\nabla}{dt} : \Gamma(\gamma^*(E)) \to \Gamma(\gamma^*(E)), \quad \sigma \mapsto \frac{\nabla}{dt}\sigma$$

such that for all function $f \in C^\infty([a, b])$ and all sections $\sigma \in \Gamma(\gamma^*(E))$

$$\frac{\nabla}{dt}(f\sigma) = \frac{df}{dt}\sigma + f\frac{\nabla}{dt}\sigma. \tag{1.2}$$

Proposition 1.28. *Given a connection ∇ on a vector bundle $\pi : E \to M$ and a curve $\gamma : [a, b] \to M$, there is a unique covariant derivative $\frac{\nabla}{dt} : \Gamma(\gamma^*(E)) \to \Gamma(\gamma^*(E))$ along γ such that*

$$\frac{\nabla}{dt}(s \circ \gamma)(t) = (\nabla_{\dot\gamma(t)}s)(\gamma(t)). \tag{1.3}$$

for all sections s of the bundle E.

Proof. (Uniqueness) Arguing as in Proposition 1.7, it is not hard to show that $\frac{\nabla}{dt}$ is local: for a section σ of E along γ the value $(\frac{\nabla}{dt}\sigma)(t)$ at a point t depends only on the values of σ near t. Therefore, in order to prove uniqueness it is no loss of generality to assume that the image $\gamma([a, b])$ of γ is contained in an open set U in M with $E|_U$ trivial. Pick a frame $\{s_j\}$ of $E|_U$. Then for any $\sigma \in \Gamma(\gamma^*E)$ there are smooth functions $f_j \in C^\infty([a, b])$ so that

$$\sigma(t) = \sum f_j(t)s_j(\gamma(t))$$

for all $t \in [a, b]$. Then, using (1.2) and (1.3), we get

$$\frac{\nabla}{dt}\sigma(t) = \frac{\nabla}{dt}\left(\sum f_j\,(s_j \circ \gamma)\right)(t) = \sum \frac{df_j}{dt}(t)\, s_j(\gamma(t)) + \sum f_j(\nabla_{\dot\gamma(t)}s_j)(\gamma(t)). \tag{1.4}$$

Since the right hand side of (1.4) depends only on ∇, $\frac{\nabla}{dt}$ is unique.

(Existence) Cover $\gamma([a,b])$ with sets U_j such that $E|_{U_j}$ is trivial. It's enough to construct $\frac{\nabla}{dt}$ on each $\Gamma(\gamma^*E|_{\gamma^{-1}(U_j)})$ for by uniqueness the operators on each $\Gamma(\gamma^*E|_{\gamma^{-1}(U_j)})$ will patch together to a map $\frac{\nabla}{dt}$: $\Gamma(\gamma^*E) \to \Gamma(\gamma^*E)$. Pick a frame $\{s_k^{(j)}\}$ on $E|_{U_j}$ and define $\frac{\nabla}{dt}$ on $\gamma^*(E|_{U_j})$ by (1.4). ☐

Definition 1.29. We will refer to the covariant derivative $\frac{\nabla}{dt}$ along γ as in the Proposition 1.28 above as being *induced* by the connection ∇.

Definition 1.30. Let $E \to M$ be a vector bundle with a connection ∇, $\gamma : [a,b] \to M$ a curve. A section $\sigma \in \Gamma(\gamma^*E)$ is *parallel* if

$$\frac{\nabla}{dt}\sigma = 0,$$

where $\frac{\nabla}{dt}$ is the covariant derivative along γ induced by ∇.

To define parallel transport along a curve $\gamma : [a,b] \to M$, we want, for every vector $v \in E_{\gamma(a)}$, a section $\sigma^v \in \Gamma(\gamma^*(E))$ such that $\sigma^v(a) = v$ and $\frac{\nabla}{dt}\sigma^v = 0$. We also want the map $v \mapsto \sigma^v$ to be linear. The existence of such sections and linearity in v is the result of the next two lemmas. The first one is the standard result for linear time dependent ODE's.

Lemma 1.31. *Suppose that $B = (B_{jk}(t)) : [c,d] \to \mathbb{R}^{k^2}$ is a smooth curve in the space of $k \times k$ real matrices. Then there is a smooth curve $R : [c,d] \to \mathrm{GL}(\mathbb{R},k)$ such that $f(t) := R(t)f^0$ is a solution of the ODE*

$$\begin{pmatrix} f_1'(t) \\ \vdots \\ f_k'(t) \end{pmatrix} = B(t) \begin{pmatrix} f_1(t) \\ \vdots \\ f_k(t) \end{pmatrix}, \tag{1.5}$$

with initial conditions $f(c) = f^0$.

Lemma 1.32. *Let $E \to M$ be a vector bundle with a connection ∇ and $\gamma : [a,b] \to M$ be smooth curve. For any vector $v \in E_{\gamma(a)}$ there is a section $\sigma^v \in \Gamma(\gamma^*(E))$ such that $\sigma^v(a) = v$ and $\frac{\nabla}{dt}\sigma^v = 0$. Moreover, the map*

$$E_{\gamma(a)} \to \Gamma(\gamma^*E), \quad v \mapsto \sigma^v$$

is a linear isomorphism.

Proof. As before, it is no loss of generality to assume the image of γ is contained in a coordinate chart $(x_1, \dots, x_m) : U \to \mathbb{R}^m$ with $E|_U$ being trivial. Let $\{s_j\}$ be a frame of $E|_U$ and Γ_{ij}^k the corresponding Christoffel symbols. Suppose σ is a section of E along γ which is parallel and satisfies $\sigma(a) = v$. Then there are smooth functions $f_j \in C^\infty([a,b])$ so that $\sigma = \sum f_j\,(s_j \circ \gamma)$. We argue that the f_j's satisfy a linear ODE as in Lemma 1.31 for some curve B. By (1.4), since $\frac{\nabla}{dt}\sigma = 0$,

$$\sum \frac{df_j}{dt}(t)\, s_j(\gamma(t)) = -\sum f_j\, (\nabla_{\dot\gamma(t)} s_j)(\gamma(t)).$$

We also have $\dot\gamma = \sum_i (\frac{d}{dt}\gamma_i)\frac{\partial}{\partial x_i}$, where $\gamma_i := x_i \circ \gamma$. Therefore

$$\nabla_{\dot\gamma} s_j = \sum \dot\gamma_i\, (\nabla_{\frac{\partial}{\partial x_i}} s_j) \circ \gamma = \sum_{i,j,k} \dot\gamma_i (\Gamma_{ij}^k s_k) \circ \gamma = \sum_k (\sum_i \dot\gamma_i\, (\Gamma_{ij}^k \circ \gamma))\, (s_k \circ \gamma)$$

We conclude that $\sigma = \sum f_j\,(s_j \circ \gamma)$ is parallel if and only if

$$\frac{df_k}{dt}(t) = -\sum_{i,j} f_j(t)\dot\gamma_i(t)\, \Gamma_{ij}^k(\gamma(t)). \tag{1.6}$$

That is, $f = (f_1, \dots, f_k)$ satisfies the ODE (1.5) with

$$B_{jk}(t) = \sum_i \dot\gamma_i(t)(\Gamma_{ij}^k(\gamma(t))$$

By Lemma 1.31 the system of linear equations (1.6) has a solution defined for all time $t \in [a,b]$ which depends linearly on the initial conditions. Therefore the desired parallel transport exists. ☐

Parallel transport leads to one definition of geodesics.

Definition 1.33. Let ∇ be a connection on the tangent bundle $TM \to M$ of a manifold M. A curve $\gamma : [a, b] \to M$ is a *geodesic* if its velocity field $\dot\gamma(t)$ is parallel:

$$\frac{\nabla}{dt}\dot\gamma = 0. \tag{1.7}$$

Remark 1.34. It will be useful to know what (1.7) means in coordinates. Let $(x_1, \ldots, x_m) : U \to \mathbb{R}^m$ be a coordinate chart on our manifold. Define $\gamma_i = x_i \circ \gamma$, $\dot\gamma_i = \frac{d}{dt}\gamma_i$ and $\ddot\gamma_i = \frac{d}{dt}\dot\gamma_i$. Then $\dot\gamma = \sum \dot\gamma_i \frac{\partial}{\partial x_i}$. Hence the functions f_k in (1.6) are $\dot\gamma_k$s. Therefore, in this case, (1.6) reads

$$\ddot\gamma_k = -\sum \dot\gamma_i\dot\gamma_j\Gamma^k_{ij}(\gamma). \tag{1.8}$$

We conclude that a curve γ is a geodesic for a connection ∇ if and only if (1.8) holds in every coordinate chart.

Exercise 1.1. Consider the manifold $\mathbb{R}^n$. We have seen that $D_X Y = \sum X(Y^i)\frac{\partial}{\partial x_i}$ is a connection. Suppose that $\gamma : \mathbb{R} \to \mathbb{R}^n$ is a curve. Let $\frac{D}{dt}$ denote the covariant derivative along γ induced by the connection D on $\mathbb{R}^n$. Show that

$$\frac{D}{dt}\dot\gamma = \ddot\gamma \;(= \frac{d^2\gamma}{dt^2}).$$

Conclude that the geodesics in $\mathbb{R}^n$ with respect to D are straight lines.

2. Riemannian geometry

2.1. **Levi-Civita connection.** We now specialize the discussion of connections and parallel transport to the case of manifolds with a choice of an inner product on each tangent space.

Definition 2.1 (Riemannian metric)**.** A *Riemannian metric* g on a manifold M assigns smoothly to each point $x \in M$ a positive definite inner product g_x on T_xM.

A *Riemannian manifold* is a manifold M together with a choice of a Riemannian metric g. In other words, it's a pair (M, g).

Remark 2.2. An inner product h on a vector space V is a bilinear map $h : V \times V \to \mathbb{R}$. Hence it is an element of the tensor product $V^* \otimes V^*$. Therefore a Riemannian metric on a manifold M is nothing but a smooth section of the bundle $(T^*M)^{\otimes 2} := T^*M \otimes T^*M \to M$. (Not all sections of $T^*M^{\otimes 2} \to M$ are Riemannian metrics. For instance, the zero section is not. But all symmetric and positive definite sections of $T^*M^{\otimes 2} \to M$ are Riemannian metrics.)

Theorem 2.3. *Any second countable manifold M has a Riemannian metric.*

Proof. Let $\{\phi_i = (x_1^{(i)}, \ldots, x_m^{(i)}) : U_i \to \mathbb{R}^m\}$ be a countable collection of coordinate charts that cover M. One each chart U_i define a metric $g^{(i)} = \sum_j dx_j^{(i)} \otimes dx_j^{(i)}$ Let $\{\rho_i\}$ be a partition of unity subordinate to this cover. Define a section g of $T^*M \otimes T^*M \to M$ by

$$g = \sum_i \rho_i g^{(i)}.$$

Then g is a Riemannian metric. ∎

Fiber metrics. The notion of a Riemannian metric generalizes to arbitrary vector bundles.

Definition 2.4. A *fiber metric* on the vector bundle $E \to M$ assigns smoothly to each point $x \in M$ a positive definite symmetric bilinear form $g_x : E_x \times E_x \to \mathbb{R}$. In particular a fiber product is a section of $E^* \otimes E^* \to M$.

Proposition 2.5. *Every vector bundle $E \to M$ over a paracompact manifold M has a fiber metric.*

Proof. If $\{s_\alpha : U \to E\}$ is a local frame, then

$$g_x(\sum a_\alpha s_\alpha(x), \sum b_\beta s_\beta(x)) = \sum a_\alpha b_\beta \delta_{\alpha\beta}$$

is a fiber metric on $E|_U$. Patch these local fiber metrics together using a partition of unity. ∎

The next theorem is the fundamental theorem of Riemannian geometry. It says that for every Riemannian manifold (M, g) there is a connection ∇ (which depends on the metric g) with two important properties. Such connection is called the *Levi-Civita connection*.

Theorem 2.6 (existence and uniqueness of the Levi-Civita connection)**.** *On every Riemannian manifold* (M, g) *there is a unique connection* $\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ *which is*
(1) Torsion-free : $\nabla_X Y - \nabla_Y X = [X, Y]$ *for all* $X, Y \in \Gamma(TM)$
(2) metric (i.e. compatible with g*) :* $X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$ *for all* $X, Y, Z \in \Gamma(TM)$*.*

Proof. (Uniqueness) The proof is a trick. Suppose that ∇ exists. Then for any $X, Y, Z \in \Gamma(TM)$,

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

since the connection is compatible with the metric. Adding up the three equations and using the fact that the connection is torsion free, we get

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

Thus, we have

$$(2.1) \quad 2g(\nabla_X Y, Z) = X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) + g([X, Y], Z) - g(Y, [X, Z]) - g(X, [Y, Z]).$$

Since Z is arbitrary and g is nondegenerate, the formula above uniquely determines $\nabla_X Y$. This proves uniqueness of a Levi-Civita connection.

It remains to prove existence. The proof is very simple, if one is willing to skip all the details. *Define* an $\mathbb{R}$-trilinear map

$$\Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to C^\infty(M)$$

by sending a triple of vector fields (X, Y, Z) to 1/2 of the right hand side of (2.1). Since g is nondegenerate this defines an $\mathbb{R}$-bilinear map

$$\Gamma(TM) \times \Gamma(TM) \to \Gamma(TM), \quad (X, Y) \mapsto \text{“}\nabla\text{”}_X Y.$$

It remains to verify that “∇” so defined is a connection, and that it is metric and torsion-free. These minor details are traditionally left to the reader. We will provide a different and more detailed proof below after a brief detour. ☐

Equation (2.1) has the following interesting consequence:

Lemma 2.7. *The Christoffel symbols of the Levi-Civita connection depend only on the metric and its first partials.*

Proof. Given a coordinate chart $(x_1, \ldots, x_m) : U \to \mathbb{R}^m$ on M, the Christoffel symbols Γ^k_{ij} of the Levi-Civita connection ∇ are defined by

$$\nabla_{\partial_i} \partial_j = \sum_k \Gamma^k_{ij} \partial_k,$$

where $\partial_i = \frac{\partial}{\partial x_i}$. Plugging $X = \partial_i$, $Y = \partial_j$ and $Z = \partial_k$ into (2.1) we get

$$2g(\nabla_{\partial_i} \partial_j, \partial_k) = \partial_i(g(\partial_j, \partial_k)) + \partial_j(g(\partial_j, \partial_i)) - \partial_k(g(\partial_i, \partial_j))$$

since $[\partial_i, \partial_j] = [\partial_j, \partial_k] = [\partial_i, \partial_k] = 0$. Writing $g_{ij} = g(\partial_i, \partial_j)$ etc., we obtain

$$(2.2) \qquad 2\sum_l \Gamma^l_{ij} g_{lk} = \partial_i g_{jk} + \partial_j g_{ji} - \partial_k g_{ij}.$$

Since g is a metric, the matrix (g_{ij}) is nondegenerate. Let (g^{rs}) denote its inverse, so that

$$\sum_s g^{rs} g_{sk} = \delta_{rk}.$$

Multiplying both sides of (2.2) by g^{sk} and summing over k we get

$$\sum_l \delta_{sl}\Gamma^l_{ij} = \frac{1}{2}\sum_k g^{sk}\left(\partial_i g_{jk} + \partial_j g_{ji} - \partial_k g_{ij}\right),$$

and simplifying

$$\Gamma^s_{ij} = \frac{1}{2}\sum_k g^{sk}\left(\partial_i g_{jk} + \partial_j g_{ji} - \partial_k g_{ij}\right). \tag{2.3}$$

This proves that the Christoffel symbols depend only on the metric and its first order partials. □

Proof of Theorem 2.6 continued. It remains to (re)prove the existence of the Levi-Civita connection. By uniqueness, it is enough to construct a Levi-Civita connection ∇ in each coordinate chart. For then by uniqueness, these coordinate chart connections patch together into a Levi-Civita connection on the whole manifold M. We have shown that if the Levi-Civita connection exists then its Christoffel symbols have to be given by (2.3). Therefore on a chart $(x_1, \ldots, x_m) : U \to \mathbb{R}^m$ we *define* a connection ∇ by

$$\nabla_{X_i\partial_i} Y_j\partial_j = X_i(\partial_i Y_j)\partial_j + X_i Y_j \Gamma^k_{ij}\partial_k$$

with Christoffel symbols Γ^k_{ij} given by (2.3). In the equation above we finally resorted to the Einstein summation convention: we sum on repeated indices and omit the symbol $\sum$. We now check that ∇ is a Levi-Civita connection.

Since $\Gamma^k_{ij} = \Gamma^k_{ji}$ (c.f. (2.3)),

$$\nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i = \Gamma^k_{ij}\partial_k - \Gamma^k_{ji}\partial_k = 0.$$

Thus, for two vector fields $X = X_i\partial_i$ and $Y = Y_j\partial_j$, we have

$$\begin{aligned}
\nabla_X Y - \nabla_Y X &= \nabla_{X_i\partial_i}(Y_j p_j) - \nabla_{Y_j p_j}(X_i\partial_i) \\
&= X_i(\partial_i Y_j)\partial_j + X_i Y_j \nabla_{\partial_i}\partial_j - Y_j(\partial_j X_i)\partial_i - Y_j X_i \nabla_{\partial_j}\partial_i \\
&= X_i(\partial_i Y_j)\partial_j - Y_j(\partial_j X_i)\partial_i \\
&= [X_i\partial_i, Y_j\partial_j].
\end{aligned}$$

Thus, ∇ is torsion-free. Compatibility with g is a somewhat longer computation. First, note that

$$\begin{aligned}
g(\nabla_{\partial_i}\partial_j, \partial_k) + g(\partial_j, \nabla_{\partial_i}\partial_k) &= g(\Gamma^l_{ij}\partial_l, \partial_k) + g(\partial_j, \Gamma^m_{ik}\partial_m) \\
&= \Gamma^l_{ij} g_{lk} + \Gamma^m_{ik} g_{jm} \\
&= \partial_i g_{jk},
\end{aligned}$$

where the last equality follows from (2.3). Thus, we have for vector fields $X = X_i\partial_i$,$Y = Y_j\partial_j$ and $Z = Z_k\partial_k$,

$$\begin{aligned}
(X_j\partial_j) g(Y_i\partial_i, Z_k\partial_k) &= X_j\partial_j(Y_i Z_k g_{ik}) \\
&= X_j(\partial_j Y_i) Z_k g_{ik} + X_j Y_i(\partial_j Z_k) g_{ik} + X_i Y_j Z_k(\partial_j g_{ik}) \\
&= g(X_j(\partial_j Y_i)\partial_i, Z_k\partial_k) + g(Y_i\partial_i, X_j(\partial_j Z_k)\partial_k) \\
&\quad + X_j Y_i Z_k(g(\nabla_{\partial_j}\partial_i, \partial_k) + g(\partial_i, \nabla_{\partial_j}\partial_k)) \\
&= g((X_j\partial_j)Y_i\partial_i, Z_k\partial_k) + g(Y_i\nabla_{X_j\partial_j}\partial_i, Z_k\partial_k) \\
&\quad + g(Y_i\partial_i, (X_j\partial_j)Z_k\partial_k) + g(Y_i\partial_i, Z_k\nabla_{X_j\partial_j}\partial_k) \\
&= g(\nabla_{X_j\partial_j}(Y_i\partial_i), Z_k\partial_k) + g(Y_i\partial_i, \nabla_{X_j\partial_j}(Z_k\partial_k)).
\end{aligned}$$

That is, the connection ∇ is compatible with the metric g. Therefore the connection with Christoffel symbols defined by (2.3) is a Levi-Civita connection. This finishes the proof of existence and uniqueness of the Levi-Civita connection. □

Example 2.8. Consider the manifold $\mathbb{R}^n$. We have seen that $D_X Y = \sum X(Y_i)\frac{\partial}{\partial x_i}$ is a connection. An easy computation shows D is the Levi-Civita connection on $\mathbb{R}^n$ with respect to the standard inner product on $\mathbb{R}^n$.

We end this section with a brief discussion of the geometric meaning of a connection being metric.

Definition 2.9. Let $E \to M$ be a vector bundle with a fiber metric g. A connection ∇ on E is *metric* if

$$X(g(s,s')) = g(\nabla_X s, s') + g(s, \nabla_X s')$$

for all vector fields X and sections $s, s' \in \Gamma(E)$.

Definition 2.10. Let V_1, V_2 be two vector spaces with inner products g_1, g_2 respectively. A linear map $A : V_1 \to V_2$ is an *isometry* if

$$g_2(Av, Aw) = g_1(v, w)$$

for all $v, w \in V_1$.

Lemma 2.11. *If a connection ∇ is metric then the associated parallel transport is an isometry.*

Proof. We will only prove the lemma for embedded curves and leave the general case as an exercise. If $\gamma : [a,b] \to M$ is an embedded curve, then locally any section $\sigma : [a,b] \to E$ is of the form $s \circ \gamma$. Let $v, w \in E_{\gamma(a)}$ be two vectors and $\sigma^v, \sigma^w : [a,b] \to E$ two parallel sections with $\sigma^v(a) = a$ and $\sigma^w(a) = w$. We want to prove that the function $t \mapsto g_{\gamma(t)}(\sigma^v(t), \sigma^w(t))$ is constant. For this it's enough to prove that its derivative is zero for all t. This condition is local in t, so we may assume, by above remark, that $\sigma^v = s^v \circ \gamma$ and $\sigma^w = s^w \circ \gamma$ for some (local) sections s^v, s^w of E. Then

$$g_{\gamma(t)}(\sigma^v(t), \sigma^w(t)) = [g(s^v, s^w)](\gamma(t)).$$

Hence

$$\begin{aligned} \frac{d}{dt}\bigg|_t g_{\gamma(t)}(\sigma^v(t), \sigma^w(t)) =& \dot\gamma(g(s^v, s^w)) \\ &= g(\nabla_{\dot\gamma} s^v, s^w) + g(s^v, \nabla_{\dot\gamma} s^w) \\ &= g(0, s^w) + g(s^v, 0) = 0. \end{aligned}$$

□

2.2. **Connections induced on submanifolds.** Let (M, g) be a Riemannian manifold and $N \hookrightarrow M$ an embedded submanifold (think of a surface in $\mathbb{R}^3$). We'll see that the embedding induces a Levi-Civita connection on N in two ways that turn out to be equivalent. It will also turn out that for surfaces in $\mathbb{R}^3$ the curvature of the induced connection is intimately related to Gauss curvature.

Suppose $f : N \to M$ is a map of manifolds. Then we can use f to pull back a metric g on M to a positive semi-definite symmetric bilinear form on N:

$$(f^*g)_x(v, w) = g_{f(x)}(df_x v, df_x w)$$

for all $x \in N$, $v, w \in T_xN$. Moreover, if df_x is *injective* then $(f^*g)_x$ is non-degenerate. Therefore if $f : N \to M$ is an immersion then $g^N := f^*g$ is a metric on N. The metric g^N defines a Levi-Civita connection ∇^N on N.

Suppose now that $f : N \hookrightarrow M$ is an *embedding*. Then there is another way to induce a connection on N from a connection M. First of all, for all point $x \in N$ the tangent space T_xM splits as an orthogonal direct sum with respect to g_x:

$$T_xM = T_xN \oplus (T_xN)^\perp.$$

Hence there is an orthogonal projection

$$\Pi_x : T_xM \to T_xN.$$

Globally $\nu := \sqcup_x (T_xN)^\perp$ is a vector bundle, the *normal bundle* of the embedding of N into M. Hence globally the first equation says that the restriction $TM|_N$ is a direct sum of two bundles:

$$TM|_N = TN \oplus \nu$$

and the second equation says that we have a bundle map

$$\Pi : TM|_N \to TN.$$

Here is how one can see that Π_x depends smoothly on x: Choose coordinates $\phi = (x_1, \ldots, x_n, \ldots, x_m) : U \to \mathbb{R}^m$ on M near a point $x \in N$ that are adapted to N. That is, $\phi(N \cap U) = \phi(U) \cap \{x_{n+1} = 0, \ldots, x_m = 0\}$. Apply Gram-Schmidt to the basis vectors $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}, \ldots, \frac{\partial}{\partial x_m}\}$ to obtain an orthonormal frame $\{e_1(x), \ldots, e_n(x), \ldots, e_m(x)\}$ on TU. Remember that every tangent space

T_xM has an inner product g_x that depends smoothly on x. The Gram-Schmidt is smooth in the inner product. Define the projection Π by

$$\Pi_x(v) = \sum_{i=1}^{n} g_x(v, e_i(x))e_i(x)$$

Definition 2.12. Let $N \subset M$ be an embedded submanifold. A vector field $\tilde{X} \in \Gamma(TM)$ is an *extension* of a vector field $X \in \Gamma(TN)$ if

$$X_x = \tilde{X}_x$$

for all $x \in N$. We will also say that $\tilde{X}$ is *tangent to* N.

Lemma 2.13. *Let $N \subset M$ be an embedded submanifold and $X \in \Gamma(TN)$ a vector field. Then for any $x \in N$ there is a neighborhood $U \subset M$ and an extension $\tilde{X} \in \Gamma(TM|_U)$ of $X|_{N\cap U}$.*

Proof. Let $(x_1, \dots, x_n, \dots, x_m) : U \to \mathbb{R}^m$ be coordinates on M adapted to N. Then $X = \sum_{i=1}^n X_i \frac{\partial}{\partial x_i}$, with X_i being smooth functions on $U \cap N$. Extend X_i to all of U by making them constant in $x_{n+1}, \dots, x_m$. This extends X to all of U. □

Lemma 2.14. *Let $N \subset M$ be an embedded submanifold, $X, Y \in \Gamma(TN)$ be two vector fields and $\tilde{X}, \tilde{Y} \in \Gamma(TM)$ their extensions. Then their Lie bracket $[\tilde{X}, \tilde{Y}]$ is tangent to N, hence is an extension of $[X, Y]$.*

Proof. We give two proofs. The first is computational. In coordinates $(x_1, \dots, x_n, \dots, x_m)$ on M adapted to N, $\tilde{X} = \sum_{i=1}^m \tilde{X}_i \frac{\partial}{\partial x_i}$ with $\tilde{X}_i(x) = 0$ for $i > n$ for all $x \in N$. Similarly $\tilde{Y} = \sum_{i=1}^m \tilde{Y}_i \frac{\partial}{\partial x_i}$ with $\tilde{Y}_i(x) = 0$ for $i > n$ for all $x \in N$. Since

$$[\tilde{X}, \tilde{Y}] = \sum_{i,j} \tilde{X}_i \frac{\partial \tilde{Y}_j}{\partial x_i} \frac{\partial}{\partial x_j} - \sum_{i,j} \tilde{Y}_j \frac{\partial \tilde{X}_i}{\partial x_j} \frac{\partial}{\partial x_i}$$

for $i > n$ the coefficient in front of $\frac{\partial}{\partial x_i}$ vanishes at the points of N.

Here is a geometric proof. If $\tilde{X}$ is tangent to N, its flow ϕ_t preserves N (maps it into itself). Hence its differential $d\phi_t$ maps vectors tangent to N to vectors tangent to N. But $\tilde{Y}$ is tangent to N. Hence for any $x \in N$

$$(d(\phi_{-t})\tilde{Y})_x \in T_xN$$

for all t. Differentiating with respect to t we get

$$[\tilde{X}, \tilde{Y}]_x \in T_xN.$$

□

We now define a connection $\bar{\nabla}$ on a manifold N induced by its embedding into a Riemannian manifold (M, g) by

$$\bar{\nabla}_X Y(x) := \Pi_x(\nabla_{\tilde{X}} \tilde{Y}(x)),$$

where $x \in N$ is a point, $X, Y \in \Gamma(TN)$ are two vector fields, $\tilde{X}, \tilde{Y}$ their (local) extensions to M, $\Pi_x : T_xM \to T_xN$ is the orthogonal projection and ∇ is the Levi-Civita connection on (M, g).

We need to make sure that $\bar{\nabla}$ is well-defined, that is, that $\bar{\nabla}_X Y(x)$ does not depend on the choice of the local extensions $\tilde{X}, \tilde{Y}$. By Corollary 1.10.2 $\nabla_{\tilde{X}} \tilde{Y}(x)$ depends only on $\tilde{X}_x = X_x$ and the values of $\tilde{Y}$ along the integral curve of $\tilde{X}$ through x. Therefore $\nabla_{\tilde{X}} \tilde{Y}(x)$ depends only on X_x and the values of Y along the integral curve of X through x. Hence $\bar{\nabla}$ is well-defined. Moreover, $\bar{\nabla}_X Y$ is clearly tensorial in the X slot. To see that it is a connection, let $f \in C^\infty(N)$ be a function and $\tilde{f}$ its (local) extension to M. Then, at the points of N,

$$\begin{aligned}\bar{\nabla}_X(fY) &= \Pi(\nabla_{\tilde{X}}(\tilde{f}\tilde{Y}) = \Pi((\tilde{X}\tilde{f})\tilde{Y} + \tilde{f}\nabla_{\tilde{X}}\tilde{Y}) \\ &= (\tilde{X}\tilde{f})\Pi(\tilde{Y}) + \tilde{f}\Pi(\nabla_{\tilde{X}}\tilde{Y}) = (Xf)Y + f\bar{\nabla}_X Y.\end{aligned}$$

We conclude that the induced connection $\bar{\nabla}$ is indeed a connection.

Remark 2.15. The projection Π is really necessary in the definition of the induced connection. This is because even if vector fields $\tilde{X}$ and $\tilde{Y}$ are tangent to a submanifold N there is no reason for their covariant derivative $\nabla_{\tilde{X}}\tilde{Y}$ to be tangent to N. Here is an example:

Let $W = Z = x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}$, two vector fields on $M = \mathbb{R}^2$. Let D denote the Levi-Civita connection on $\mathbb{R}^2$ for the standard metric $dx \otimes dx + dy \otimes dy$. Then

$$D_W Z = (Wx_2)\frac{\partial}{\partial x_1} + (W(-x_1))\frac{\partial}{\partial x_2} = -x_1\frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_2}.$$

Let $N = S^1$. Then W and Z are tangent to N, hence are extensions of a vector field on N. But $D_W Z$ is orthogonal to S^1.

Lemma 2.16. *Let (M, g) be a Riemannian manifold and $i : N \hookrightarrow M$ an embedded submanifold. Then the connection $\bar{\nabla}$ induced on N by the Levi-Civita connection ∇ on M is the Levi-Civita connection for the pullback metric $g^N := i^*g$.*

Proof. It is enough to check that

(1) $\bar{\nabla}$ is torsion-free and that
(2) $\bar{\nabla}$ is metric.

For all $X, Y \in \Gamma(TN)$ and their local extensions $\tilde{X}, \tilde{Y} \in \Gamma(TM)$

$$\bar{\nabla}_X Y - \bar{\nabla}_Y X = \Pi(\nabla_{\tilde{X}}\tilde{Y} - \nabla_{\tilde{Y}}\tilde{X}) = \Pi([\tilde{X}, \tilde{Y}]) = \Pi([X, Y]) = [X, Y].$$

To show that $\bar{\nabla}$ is metric we need to check that

$$Z(g^N(X, Y)) = g^N(\bar{\nabla}_Z X, Y) + g^N(X, \bar{\nabla}_Z Y)$$

for any vector fields X, Y, Z on N. At any point of N,

$$\begin{aligned} Z(g^N(X,Y)) &= \tilde{Z}(g(\tilde{X}, \tilde{Y})) \\ &= g(\nabla_{\tilde{Z}}\tilde{X}, \tilde{Y}) + g(\tilde{X}, \nabla_{\tilde{Z}}\tilde{Y}) \\ &= g(\bar{\nabla}_Z X + (\nabla_{\tilde{Z}}\tilde{X} - \bar{\nabla}_Z X), Y) + g(X, \bar{\nabla}_Z Y + (\nabla_{\tilde{Z}}\tilde{Y} - \bar{\nabla}_Z Y)) \\ &= g(\bar{\nabla}_Z X, Y) + g(X, \bar{\nabla}_Z Y). \end{aligned}$$

since $\nabla_{\tilde{Z}}\tilde{Y} - \bar{\nabla}_Z Y$ and $\nabla_{\tilde{Z}}\tilde{X} - \bar{\nabla}_Z X$ are perpendicular to N. □

2.3. **The second fundamental form of an embedding.** As before let $N \hookrightarrow M$ be an embedded submanifold of a Riemannian manifold (M, g). We want to understand how much N curves in M. We define a tensor, the second fundamental form $I\!I_x : T_xN \times T_xN \to (T_xN)^\perp$ to measure the extrinsic geometry of N in M. We first define

$$I\!I : \Gamma(TN) \times \Gamma(TN) \to \Gamma(TN^\perp)$$

by

$$I\!I(X, Y) = \nabla_{\tilde{X}}\tilde{Y} - \bar{\nabla}_X Y,$$

where, as before, ∇ is the Levi-Civita connection on M, $\bar{\nabla}$ is the induced Levi-Civita connection on N, $\tilde{X}, \tilde{Y} \in \Gamma(TM)$ are local extensions of the vector fields $X, Y \in \Gamma(TN)$.

Proposition 2.17. *The map $I\!I$ defined above is symmetric and tensorial.*

Proof. We first argue that $I\!I$ is symmetric.

$$\begin{aligned} I\!I(X,Y) - I\!I(Y,X) &= (\nabla_{\tilde{X}}\tilde{Y} - \bar{\nabla}_X Y) - (\nabla_{\tilde{Y}}\tilde{X} - \bar{\nabla}_Y X) \\ &= (\nabla_{\tilde{X}}\tilde{Y} - \nabla_{\tilde{Y}}\tilde{X}) - (\bar{\nabla}_X Y - \bar{\nabla}_Y X) \\ &= [\tilde{X}, \tilde{Y}] - [X, Y] = 0. \end{aligned}$$

Next we argue that $I\!I$ is tensorial in the first slot. Let $\tilde{f}$ be a local extension of a function f on N. Then at the points of N,

$$I\!I(fX, Y) = \nabla_{\tilde{f}\tilde{X}}\tilde{Y} - \bar{\nabla}_{fX}Y = \tilde{f}\nabla_{\tilde{X}}\tilde{Y} - f\bar{\nabla}_X Y = f\, I\!I(X, Y).$$

□

It follows that for all points $x \in N$ there is a symmetric bilinear map

$$II_x : T_xN \times T_xN \to (T_xN)^\perp.$$

Remark 2.18. In classical terminology the *first fundamental form* of an embedding is the induced metric.

Next suppose that the embedded submanifold N is a *hypersurface*, that is, that $\dim M - \dim N = 1$. Then the normal bundle $TN^\perp$ has 1-dimensional fibers hence, locally, a frame on $TN^\perp$ is defined by one nowhere zero vector field. By rescaling, if necessary, we may assume that this vector n field has length 1 everywhere:

$$g_x(n_x, n_x) = 1$$

for all points $x \in N$. We furthermore make an extra assumption that unit vector field n normal to N is defined on *all* of N. That is, N is orientable inside M. This is true for the sphere embedded in $\mathbb{R}^3$ but false for the central circle of the Möbius band inside the band. If $N \subset M$ has a globally defined unit normal n, we can write

$$II_x(v, w) = h_x(v, w)n_x$$

for a symmetric bilinear map $h_x : T_xN \times T_xN \to \mathbb{R}$. Unwinding the definitions we see that for any vector fields X, Y on N

$$h(X, Y) = g(\nabla_{\tilde{X}}\tilde{Y}, n).$$

We will refer to $h \in \Gamma(TN^* \otimes TN^*)$ also as the second fundamental form. The second fundamental form h allows us to relate the curvature tensor R of the Levi-Civita connection on M, the *Riemann curvature* of M, and the curvature $\bar{R}$ of the induced connection on N:

Theorem 2.19. *Let $N \hookrightarrow M$ be an embedded orientable hypersurface of a Riemannian manifold (M, g). Let $h \in \Gamma(T^*N^{\otimes 2})$ be the second fundamental form of the embedding. Then for any vector fields $X, Y, Z, W \in TN$*

$$g(R(X, Y)Z, W) = g^N(\bar{R}(X, Y)Z, W) - h(Y, Z)h(X, W) + h(X, Z)h(Y, W), \tag{2.4}$$

where R is the Riemann curvature tensor of M and $\bar{R}$ is the induced Riemann curvature tensor of N.

We prove an easy lemma before tackling the computations involved in the proof of the theorem.

Lemma 2.20. *Let (M, g), ∇, N, n and h be as above. Then*

$$h(X, W) = -g(\nabla_X n, W).$$

for any vector fields $X, W \in \Gamma(TN)$ (here we didn't bother with putting tildes on the extensions).

Proof. The function $g(n, W)$ is identically 0 on N. Hence

$$0 = X(g(n, W)) = g(\nabla_X n, W) + g(n, \nabla_X W)$$

since ∇ is a metric connection. □

Proof of Theorem 2.19. Recall that

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

$$\begin{aligned} \nabla_X(\nabla_Y Z) &= \nabla_X(\bar{\nabla}_Y Z + h(Y, Z)n) \\ &= \bar{\nabla}_X(\bar{\nabla}_Y Z) + h(X, \bar{\nabla}_Y Z)n + (Xh(Y, Z))n + h(Y, Z)\bar{\nabla}_X n. \end{aligned}$$

Hence

$$g(\nabla_X(\nabla_Y Z), W) = g(\bar{\nabla}_X(\bar{\nabla}_Y Z), W) + h(Y, Z)g(\bar{\nabla}_X n, W) = g(\bar{\nabla}_X(\bar{\nabla}_Y Z), W) - h(Y, Z)h(X, W). \tag{2.5}$$

Similarly,

$$g(\nabla_Y(\nabla_X Z), W) = g(\bar{\nabla}_Y(\bar{\nabla}_X Z), W) - h(X, Z)h(Y, W), \tag{2.6}$$

while

$$g(\nabla_{[X,Y]}Z, W) = g(\bar{\nabla}_{[X,Y]}Z, W). \tag{2.7}$$

Subtracting (2.6) and (2.7) from (2.5) we get (2.4). □

Let us see what the theorem tells us about the curvature of oriented surfaces in $\mathbb{R}^3$. If $N \subset \mathbb{R}^3$ is an oriented embedded manifold, then the unit normal field n assigns to every point in N a unit vector in $\mathbb{R}^3$. Hence we can think of n as a map to the unit sphere,

$$n : N \to S^2.$$

This is the *Gauss map*. Since T_xN and $T_{n_x}S^2$ are two planes perpendicular to the same vector n_x, they are the same two plane in $\mathbb{R}^3$. Therefore we may think of the differential dn_x of the Gauss map as a map

$$dn_x : T_xN \to T_xN.$$

Definition 2.21. The *Gauss curvature* κ of an oriented surface N in $\mathbb{R}^3$ is the determinant of the differential of the Gauss map:

$$\kappa(x) = \det dn_x.$$

We compute a few examples of Gauss curvature by brute force.

Example 2.22. Consider

$$N = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = 0\},$$

a plane. The normal vector field $n(x)$ is constant, and so the Gauss curvature $\kappa(x)$ is 0.

Example 2.23. Now let N be a round cylinder:

$$N = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2^2 + x_3^2 = R^2\},$$

Here the unit normal $n(x)$ is constant in the x_1 direction. Hence, $dn_x(e_1) = 0$, and so the Gauss curvature is again zero.

Example 2.24. Let N be the standard round sphere of radius R:

$$N = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = R^2\}.$$

Then the normal vector field n is given by $n(x) = \frac{1}{R}x$, hence

$$dn = \frac{1}{R} \cdot id.$$

Therefore

$$\kappa(x) = \frac{1}{R^2}.$$

Note that the Gauss curvature is constant and positive. Also, the bigger the radius of the sphere the smaller the Gauss curvature. This makes sense since the sphere gets flatter as its radius increases.

In general one computes the Gauss curvature from the first and second fundamental form.Once again we denote the Levi-Civita connection on $\mathbb{R}^3$ by D. Then for any vector v and vector field $Y : \mathbb{R}^3 \to \mathbb{R}^3$

$$D_vY = dY(v).$$

Hence for any two vector fields X, Y on a surface N,

$$h_x(X_x, Y_x) = -g_x((D_Xn)(x), Y_x) = -g_x(dn_x(X_x), Y_x). \tag{2.8}$$

In particular the differential of the Gauss map is completely determined by the induced metric and the second fundamental form. We will see shortly that the Gauss curvature depends only on the metric g and its first and second partials. But first we extract Gauss curvature from the above equation.

Lemma 2.25. *Let g be a positive definite inner product on a vector space V, $h : V \times V \to \mathbb{R}$ a symmetric bilinear map and $S : V \to V$ the linear map uniquely defined by*

$$h(v, w) = g(Sv, w).$$

Let $\{e_i\}$ be a basis of V. Then

$$\det(h(e_i, e_j)) = \det(g(e_i, e_j)) \det S.$$

Proof. The matrix (s_{ki}) of S with respect to the basis $\{e_i\}$ is defined by

$$Se_i = \sum_k s_{ki} e_k.$$

Therefore

$$h(e_i, e_j) = g(Se_i, e_j) = g(\sum_k s_{ki} e_k, e_j) = \sum_k s_{ki} g(e_k, e_j).$$

Therefore the matrix $(h(e_i, e_j))$ is the product of matrices (s_{ki}) and $(g(e_j, e_k)) = (g(e_k, e_j))$. Thus

$$\det(h(e_i, e_j)) = \det(g(e_j, e_k)) \det(s_{ki}).$$

$\square$

Together Lemma 2.25 above and (2.8) tell us how to compute the Gauss curvature: pick a basis $\{e_1, e_2\}$ of the tangent space $T_x N$. Then

$$\kappa(x) = \frac{\det(h(e_i, e_j))}{\det(g(e_i, e_j))}.$$

In particular, if the basis $\{e_1, e_2\}$ is orthonormal with respect to the induced metric g,

$$\kappa(x) = \det(h(e_i, e_j)).$$

We are now ready to prove Gauss' *theorema egregium* ("remarkable theorem") from 1828!

Theorem 2.26. *Let $N \hookrightarrow \mathbb{R}^3$ be an oriented embedded surface. Let $\bar{R}$ denote the Riemann curvature on N. Then the Gauss curvature κ is given by*

$$\kappa(x) = -g^N_x(\bar{R}_x(e_1, e_2)e_1, e_2)$$

where $\{e_1, e_2\}$ is a basis of $T_x N$ orthonormal with respect to the induced metric g^N.

Hence the Gauss metric depends only on the induced metric and its first and second partials and not on the embedding.

Proof. The Riemann curvature of the standard Levi-Civita connection D on $\mathbb{R}^3$ is 0. Hence, by Theorem 2.19

$$g^N_x(\bar{R}_x(e_1, e_2)e_1, e_2) = h_x(e_2, e_1)h_x(e_1, e_2) - h_x(e_1, e_1)h_x(e_1, e_1) = -\det(h_x(e_i, e_j)) = -\kappa(x).$$

The curvature of a connection depends on the Christoffel symbols and their first partials. The Christoffel symbols of a Levi-Civita connection are functions of the metric and its first partials. $\square$

Exercise 2.1. Let $f(x, y)$ be a smooth function on $\mathbb{R}^2$ and N its graph in $\mathbb{R}^3$:

$$N = \{(x, y, f(x, y)) \mid (x, y) \in \mathbb{R}^2\}$$

Show that the Gauss curvature κ is given by

$$\kappa = \frac{f_{xx} f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2}$$

where $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$ and so on.

3. Geodesics as critical points of the energy functional

This section is a brief excursion into the calculus of variations. The basic setup is this. Let M be a manifold. Consider the set of all maps $\mathcal{P}$ from a fixed interval $[a, b]$ to M with fixed end points:

$$\mathcal{P} = \mathcal{P}([a, b], q_1, q_2) = \{\gamma : [a, b] \to M \mid \gamma(a) = q_1, \gamma(b) = q_2\},$$

where $q_1, q_2 \in M$ are two points. Every path $\gamma \in \mathcal{P}$ gives rise to a path $\dot{\gamma} : [a, b] \to TM$. Therefore, a smooth function $L : TM \to \mathbb{R}$ on the tangent bundle of M (a "Lagrangian") defines a map ("action")

$$\mathcal{A} : \mathcal{P} \to \mathbb{R}, \quad \mathcal{A}(\gamma) = \int_a^b L(\dot{\gamma}(t))\, dt.$$

For example, if g is a Riemannian metric on a manifold M then

$$L(x, v) = \frac{1}{2} g_x(v, v) \quad x \in M, v \in T_x M$$

is a Lagrangian and the corresponding action

$$\mathcal{A}_L(\gamma) = \int_a^b \frac{1}{2} g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))\, dt$$

is the "energy" of the path. The term "energy" comes from the fact that for a particle of mass m moving in $\mathbb{R}^3$ the quantity $\frac{1}{2}m(v_1^2 + v_1^2 + v_3^2) = \frac{1}{2}m||v||^2$ is the kinetic energy.

We want to make sense of a path $\gamma \in \mathcal{P}$ being *critical* for a an action $\mathcal{A}_L : \mathcal{P} \to \mathbb{R}$. This is a bit delicate since we have been careless with the topology on $\mathcal{P}$ and since $\mathcal{P}$ is infinite dimensional. The cheapest way to do it is by analogy with a finite dimensional case: a point is critical for a function f if and only if for every path $\sigma(s)$ through the point, we have $\frac{d}{ds}\Big|_{s=0} f(\sigma(s)) = 0$. Now, a path in the space $\mathcal{P}$ through $\gamma^0 \in \mathcal{P}$ is a family of curves γ_s with $\gamma_s|_{s=0} = \gamma^0$, where s varies in some open interval $(-\epsilon, \epsilon)$. We say that γ_s depends smoothly on s if the map

$$(-\epsilon, \epsilon) \times [a, b] \to M, \quad (s, t) \mapsto \gamma_s(t)$$

is smooth.

Definition 3.1. Let $\mathcal{P} = \mathcal{P}([a, b], q_1, q_2)$ be a space of paths in a manifold M and $L : TM \to \mathbb{R}$ a Lagrangian. A path $\gamma^0 \in \mathcal{P}$ is *L-critical* if for any family γ_s of paths through γ^0 we have

$$\frac{d}{ds}\Big|_{s=0} (\mathcal{A}_L(\gamma_s)) = 0,$$

where $\mathcal{A}_L$ is the associated action.

A connection between variational problems and Riemannian geometry is provided by the following theorem.

Theorem 3.2. *Let (M, g) be a Riemannian manifold and $L(x, v) = \frac{1}{2} g_x(v, v)$ the associated Lagrangian. A path γ is L-critical if and only if γ is a geodesic of the Levi-Civita connection.*

We will first prove the theorem above locally, when the image of the path is contained in a coordinate chart. We will then show that any L-critical path is a geodesic. We will not have time to prove the converse. We start by examining what critical paths for an arbitrary Lagrangian look like locally.

Theorem 3.3. *Let $L : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $(x, v) \mapsto L(x, v)$ be a Lagrangian. A path $\gamma^0(t) = (\gamma_1^0(t), \dots, \gamma_m^0(t)) : [a, b] \to \mathbb{R}^m$ is L-critical if and only if it satisfies the Euler-Lagrange equations:*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\gamma(t), \dot\gamma(t))\right) - \frac{\partial L}{\partial x_i}(\gamma(t), \dot\gamma(t)) = 0, \quad 1 \le i \le m. \tag{3.1}$$

Proof. Let $\gamma_s(t) = \gamma(s, t) = (\gamma_1(s, t), \dots, \gamma_m(s, t))$ be a variation of γ^0. Then $\gamma(0, t) = \gamma^0(t)$ for all t, and $\gamma(s, a) = \gamma^0(a)$, $\gamma(s, b) = \gamma^0(b)$ for all s. Hence

$$h(t) := \frac{\partial}{\partial s}\Big|_{s=0} \gamma(s, t) : [a, b] \to \mathbb{R}^m$$

has to vanish at $t = a$ and at $t = b$. It's important that there are no other restrictions on h: given an arbitrary curve $h : [a, b] \to \mathbb{R}^m$ which vanishes at the endpoints,

$$\gamma(s, t) := \gamma^0(t) + s h(t)$$

is a variation of γ^0. Note further that $\dot\gamma_s(t) = \frac{\partial}{\partial t}\Big|_t \gamma(s, t)$ and consequently

$$\frac{\partial}{\partial s}\Big|_{s=0} \dot\gamma_s(t) = \frac{\partial^2 \gamma}{\partial s \partial t}\Big|_{(0,t)} = \frac{d}{dt}\Big|_t (\frac{\partial}{\partial s}\Big|_{s=0} \gamma(s, t) = \dot h(t).$$

Since γ^0 is L-critical,

$$\begin{aligned}
0 &= \frac{d}{ds}\Big|_{s=0} \int_a^b L(\gamma(t,s), \dot\gamma(t,s))\, dt \\
&= \int_a^b \frac{\partial}{\partial s}\Big|_{s=0} L(\gamma(t,s), \dot\gamma(t,s))\, dt \\
&= \int_a^b \sum_i \left(\frac{\partial L}{\partial x_i}(\gamma^0, \dot\gamma^0) \frac{\partial \gamma_i}{\partial s}\Big|_{s=0} + \frac{\partial L}{\partial v_i}(\gamma^0, \dot\gamma^0) \frac{\partial \dot\gamma_i}{\partial s}\Big|_{s=0} \right) dt \\
&= \sum_i \int_a^b \Big(\frac{\partial L}{\partial x_i}(\gamma^0, \dot\gamma^0) h_i + \frac{\partial L}{\partial v_i}(\gamma^0, \dot\gamma^0) \dot h_i \Big)\, dt.
\end{aligned}$$

Integration by parts gives

$$\int_a^b \frac{\partial L}{\partial v_i}(\gamma^0, \dot\gamma^0)\dot h_i\, dt = \frac{\partial L}{\partial v_i}(\gamma^0, \dot\gamma^0) h_i \Big|_a^b - \int_a^b \frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\gamma^0(t), \dot\gamma^0(t)) \right) h_i(t)\, dt.$$

Therefore

$$0 = \sum_i \int_a^b \left(\frac{\partial L}{\partial x_i}(\gamma^0, \dot\gamma^0) - \frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\gamma^0(t), \dot\gamma^0(t)) \right) \right) h_i(t)\, dt.$$

Since $h_i(t)$ are arbitrary, the equation above forces (3.1): see Lemma 3.4 below.

Running the computations backwards we see that if γ^0 satisfies the Euler-Lagrange equations then γ^0 is L-critical. $\square$

Lemma 3.4. *If $f \in C^\infty([a,b])$ is a smooth function and if for any $h \in C^\infty([a,b])$ with $h(a) = h(b) = 0$ we have $\int_a^b f(t)h(t)\, dt = 0$, then $f(t) \equiv 0$.*

Proof. Exercise. $\square$

Proposition 3.5. *Let g be a metric on $\mathbb{R}^m$ and $L(x,v) = \frac{1}{2} g_x(v,v)$ the associated Lagrangian. Then γ is L-critical if and only if it is a geodesic for the Levi-Civita connection defined by the metric g.*

Proof. We have

$$2L(x,v) = \sum_{k,l} g_{kl}(x)\, v_k v_l.$$

Therefore, for each index i,

$$2\frac{\partial L}{\partial x_i} = \sum_{k,l} \frac{\partial g_{kl}}{\partial x_i} v_k v_l$$

and

$$2\frac{\partial L}{\partial v_i} = \sum_{k,l} (g_{il} v_l + g_{ki} v_k).$$

The Euler-Lagrange equations in this case then are

$$\sum_{k,l} \frac{\partial g_{kl}}{\partial x_i} \dot\gamma_k \dot\gamma_l = \frac{d}{dt}\left(\sum_{k,l} (g_{il}\dot\gamma_l + g_{ki}\dot\gamma_k) \right).$$

Differentiating and gathering $\ddot\gamma_s$ terms on one side, we get:

$$\sum_s g_{is}\ddot\gamma_s = -\frac{1}{2}\sum_{k,l} \Big(\frac{\partial g_{ki}}{\partial x_l} + \frac{\partial g_{il}}{\partial x_k} - \frac{\partial g_{kl}}{\partial x_i} \Big) \dot\gamma_l \dot\gamma_k. \tag{3.2}$$

Here we used the fact that $\gamma_{is} = \gamma_{si}$; this is where the $\frac{1}{2}$ comes from. As before we denote the entries of the inverse of the matrix $(g_{\alpha\beta})$ by $g^{\alpha\beta}$ so that $\sum_\beta g^{\alpha\beta} g_{\beta\gamma} = \delta_{\alpha\gamma}$. Therefore if we multiply both sides of (3.2) by g^{ji} and sum on i we get

$$\ddot\gamma_j = -\frac{1}{2}\sum_{i,k,l} g^{ji} \Big(\frac{\partial g_{ki}}{\partial x_l} + \frac{\partial g_{il}}{\partial x_k} - \frac{\partial g_{kl}}{\partial x_i} \Big) \dot\gamma_l \dot\gamma_k = -\sum_{k,l} \Gamma^j_{kl} \dot\gamma_k \dot\gamma_l,$$

where Γ^i_{kl} are the Christoffel symbols for the Levi-Civita connection (cf. (2.3)). We now see that this is the geodesic equation. Thus, L-critical curves are geodesics and vice versa. $\square$

The result for Lagrangians on $\mathbb{R}^n$, Theorem 3.3, and the corresponding result for geodesics, Proposition 3.5, generalize to the manifold setting. To be precise, recall that if $(x_1, \ldots, x_n) : U \to \mathbb{R}^m$ is a coordinate chart on a manifold M, then it defines an associated coordinate chart $(x_1, \ldots, x_m, v_1, \ldots, v_m) : TU \to \mathbb{R}^m \times \mathbb{R}^m$ on the tangent bundle of M. Namely, if $q \in U$ is a point and $w \in T_qU = T_qM$ is a vector, then there are unique numbers $v_1 = v_1(w), \ldots, v_m = v_m(w)$ so that

$$w = \sum_i v_i(w) \frac{\partial}{\partial x_i}\Big|_q,$$

since $\left\{ \frac{\partial}{\partial x_i}\Big|_q \right\}$ is a basis of T_qM. Of course, $v_i(w) = (dx_i)_q(w)$.

Proposition 3.6. *Let M be a manifold and $L : TM \to \mathbb{R}$ a Lagrangian. If a path $\gamma^0 : [a, b] \to M$ lies entirely inside a coordinate chart $(x_1, \ldots, x_n) : U \to \mathbb{R}^m$ (i.e., $\gamma([a, b]) \subset U$), then*

$$(\gamma^0_1(t), \ldots, \gamma^0_m(t), \dot\gamma^0_1(t), \ldots, \dot\gamma^0_m(t)) := (x_1 \circ \gamma^0(t), \ldots, x_m \circ \gamma^0(t), v_1 \circ \dot\gamma^0(t), \ldots, v_m \circ \dot\gamma^0(t))$$

satisfies the Euler-Lagrange equations. Here, as above, $(x_1, \ldots, x_m, v_1, \ldots, v_m) : TU \to \mathbb{R}^m \times \mathbb{R}^m$ is the coordinate chart on the tangent bundle TM associated with the chart $(x_1, \ldots, x_n) : U \to \mathbb{R}^m$ on the manifold M.

Proof. The only possible concern is that the image of a variation γ_s of our curve γ^0 lies outside the domain U of our coordinate chart. But we only care about γ_s for s small, and for small values of the parameter s the variation $\gamma_s(t)$ is close to $\gamma^0(t)$, hence lies in U. $\square$

From Propositions 3.5 and 3.6 we deduce:

Corollary 3.6.1. *Let M be a manifold with a Lagrangian L. A path $\gamma^0 : [a, b] \to M$ lying inside a coordinate chart on M is a geodesic for a Riemannian metric g if and only if γ^0 is critical for the energy Lagrangian $L(x, v) = \frac{1}{2} g_x(v, v)$.*

What about L-critical paths whose images cannot be covered by a single coordinate chart? Suppose $\gamma : [a, b] \to M$ is L-critical and for some time t_0 the point $\gamma(t_0)$ lies in a coordinate chart $(x_1, \ldots, x_m) : U \to \mathbb{R}^m$. Then $\gamma([a', b']) \subset U$ for some subinterval $[a', b'] \subset [a, b]$ containing t_0. Any variation of $\gamma|_{[a',b']}$ is a variation of γ. Hence $\gamma|_{[a',b']}$ is also L-critical. Therefore it satisfies Euler-Lagrange equations in the chart U. In particular, if γ is critical for the energy Lagrangian, then γ is a geodesic in every coordinate chart, hence a geodesic. This proves one global direction of Theorem 3.2, as promised.

The converse is true as well, but this requires a coordinate-free description of L-critical curves which we don't have time for.