Math 6455 — Nov 1, 2006[1]
Differential Geometry I
Fall 2006, Georgia Tech

Lecture Notes 14

Connections

Suppose that we have a vector field X on a Riemannian manifold M. How can we measure how much X is changing at a point $p \in M$ in the direction $Y_p \in T_pM$? The main problem here is that there exists no canonical way to compare a vector in some tangent space of a manifold to a vector in another tangent space. Hence we need to impose a new kind of structure on a manifold. To gain some insight, we first study the case where $M = \mathbf{R}^n$.

0.1 Differentiation of vector fields in $\mathbf{R}^n$

Since each tangent space $T_p\mathbf{R}^n$ is canonically isomorphic to $\mathbf{R}^n$, any vector field on $\mathbf{R}^n$ may be identified as a mapping $X\colon \mathbf{R}^n \to \mathbf{R}^n$. Then for any $Y_p \in T_p\mathbf{R}^n$ we define the *covariant derivative* of X with respect to Y_p as

$$\nabla_{Y_p}X := \big(Y_p(X^1), \dots, Y_p(X^n)\big).$$

Recall that $Y_p(X^i)$ is the directional derivative of X^i at p in the direction of Y, i.e., if $\gamma\colon(-\epsilon,\epsilon) \to M$ is any smooth curve with $\gamma(0) = p$ and $\gamma'(0) = Y$, then

$$Y_p(X^i) = (X^i \circ \gamma)'(0) = \big\langle \operatorname{grad} X^i(p), Y \big\rangle.$$

The last equality is an easy consequence of the chain rule. Now suppose that $Y\colon \mathbf{R}^n \to \mathbf{R}^n$ is a vector field on $\mathbf{R}^n$, $p \overset{Y}{\longmapsto} Y_p$, then we may define a new vector field on $\mathbf{R}^n$ by

$$(\nabla_Y X)_p := \nabla_{Y_p}X.$$

Then the operation $(X,Y) \overset{\nabla}{\longmapsto} \nabla_X Y$ may be thought of as a mapping $\nabla\colon \mathcal{X}(\mathbf{R}^n) \times \mathcal{X}(\mathbf{R}^n) \to \mathcal{X}(\mathbf{R}^n)$, where $\mathcal{X}$ denotes the space of vector fields on $\mathbf{R}^n$.

Next note that if $X \in \mathcal{X}(\mathbf{R}^n)$ is any vector field and $f\colon M \to \mathbf{R}$ is a function, then we may define a new vector field $fX \in (\mathbf{R}^n)$ by setting $(fX)_p := f(p)X_p$ (do not confuse fX, which is a *vector field*, with Xf which is a *function* defined by $Xf(p) := X_p(f)$). Now we observe that the covariant differentiation of vector fields on $\mathbf{R}^n$ satisfies the following properties:

1. $\nabla_Y(X_1 + X_2) = \nabla_Y X_1 + \nabla_Y X_2$

[1] Last revised: November 14, 2006

2. $\nabla_Y(fX) = (Yf)\nabla_Y X + f\nabla_Y X$

3. $\nabla_{Y_1+Y_2} X = \nabla_{Y_1} X + \nabla_{Y_2} X$

4. $\nabla_{fY} X = f\nabla_Y X$

It is an easy exercise to check the above properties. Another good exercise to write down the pointwise versions of the above expressions. For instance note that item (2) implies that

$$\nabla_{Y_p}(fX) = (Y_p f)\nabla_{Y_p} X + f(p)\nabla_{Y_p} X,$$

for all $p \in M$.

0.2 Definition of connection and Christoffel symbols

Motivated by the Euclidean case, we define a connection ∇ on a manifold M as any mapping

$$\nabla \colon \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$$

which satisfies the four properties mentioned above. We say that ∇ is smooth if whenever X and Y are smooth vector fields on M, then $\nabla_Y X$ is a smooth vector field as well. Note that any manifold admits the trivial connection $\nabla \equiv 0$. In the next sections we study some nontrivial examples.

Here we describe how to express a connection in local charts. Let E_i be a basis for the tangent space of M in a neighborhood of a point p. For instance, choose a local chart (U, ϕ) centered at p and set $E_i(q) := d\phi^{-1}_{\phi(q)}(e_i)$ for all $q \in U$. Then if X and Y are any vector fields on M, we may write $X = \sum_i X^i E_i$, and $Y = \sum_i Y^i E_i$ on U. Consequently, if ∇ is a connection on M we have

$$\nabla_Y X = \nabla_Y\Big(\sum_i X^i E_i\Big) = \sum_i \big(Y(X^i)E_i + X^i \nabla_Y E_i\big).$$

Now note that since $(\nabla_{E_j} E_i)_p \in T_pM$, for all $p \in U$, then it is a linear combination of the basis elements of T_pM. So we may write

$$\nabla_{E_j} E_i = \sum_k \Gamma^k_{ji} E_k$$

for some functions Γ^k_{ji} on U which are known as the *Christoffel symbols*. Thus

$$\begin{aligned}
\nabla_Y X &= \sum_i \Big(Y(X^i)E_i + X^i \sum_j \Big(Y^j \sum_k \Gamma^k_{ji} E_k\Big)\Big) \\
&= \sum_k \Big(Y(X^k) + \sum_{ij} Y^i X^j \Gamma^k_{ij}\Big) E_k
\end{aligned}$$

Conversely note that, a choice of the functions Γ_{ij}^k on any local neighborhood of M defines a connection on that neighborhood by the above expression. Thus we may define a connection on any manifold, by an arbitrary choice of Christoffel symbols in each local chart of some atlas of M and then using a partition of unity.

Next note that for every $p \in U$ we have:

$$(\nabla_Y X)_p = \sum_k \Big(Y_p(X^k) + \sum_{ij} Y^i(p) X^j(p) \Gamma_{ij}^k(p) \Big) E_k(p). \tag{1}$$

This immediately shows that

Theorem 0.1. *For any point $p \in M$, $(\nabla_Y X)_p$ depends only on the value of X at p and the restriction of Y to any curve $\gamma\colon (-\epsilon, \epsilon) \to M$ which belongs to the equivalence class of curves determined by X_p.* □

Thus if $p \in M$, $Y_p \in T_pM$ and X is any vector field which is defined on an open neighborhood of p, then we may define

$$\nabla_{Y_p} X := (\nabla_Y X)_p$$

where Y is any extension of Y_p to a vector field in a neighborhood of p. Note that such an extension may always be found: for instance, if $Y_p = \sum Y_p^i E_i(p)$, where E_i are some local basis for tangent spaces in a neighborhood U of p, then we may set $Y_q := \sum Y_p^i E_i(q)$ for all $q \in U$. By the previous proposition, $(\nabla_Y X)_p$ does not depend on the choice of the local extension Y, so $\nabla_{Y_p} X$ is well defined.

0.3 Induced connection on submanifolds

As we have already seen M admits a standard connection when $M = \mathbf{R}^n$. To give other examples of manifolds with a distinguished connection, we use the following observation.

Lemma 0.2. *Let $\overline{M}$ be a manifold, M be an embedded submanifold of M, and X be a vector field of M. Then for every point $p \in M$ there exists an open neighborhood $\overline{U}$ of p in $\overline{M}$ and a vector filed $\overline{X}$ defined on $\overline{U}$ such that $\overline{X}_p = X_p$ for all $p \in M$.*

Proof. Recall that, by the rank theorem, there exists a local chart $(\overline{U}, \overline{\phi})$ of $\overline{M}$ centered at p such that $\overline{\phi}(U \cap M) = \mathbf{R}^{n-k}$ where $k = \dim(\overline{M}) - \dim(M)$. Now, note that $d\overline{\phi}(X)$ is a vector field on $\mathbf{R}^{n-k}$ and let Y be an extension of $d\overline{\phi}(X)$ to $\mathbf{R}^n$ (any vector field on a subspace of $\mathbf{R}^n$ may be extended to all of $\mathbf{R}^n$). Then set $\overline{X} := d\overline{\phi}^{-1}(Y)$. □

Now if $\overline{M}$ is a Riemannian manifold with connection $\overline{\nabla}$, and M is any submanifold of $\overline{M}$, we may define a connection on M as follows. First note that for any $p \in M$,

$$T_p\overline{M} = T_pM \oplus (T_pM)^\perp,$$

that is any vector $X \in T_p\overline{M}$ may written as sum of a vector $X^\top \in T_pM$ (which is tangent to M and vector $X^\perp := X - X^\top$ (which is normal to M). So for any vector fields X and Y on M we define a new vector field on M by setting, for each $p \in M$,

$$(\nabla_Y X)_p := (\overline{\nabla}_{\overline{Y}}\overline{X})_p^\top$$

where $\overline{Y}$ and $\overline{X}$ are local extensions of X and Y to vector fields on a neighborhood of p in M. Note $(\nabla_Y X)_p$ is well-defined, because it is independent of the choice of local extensions $\overline{X}$ and $\overline{Y}$ by Theorem 0.1.

0.4 Covariant derivative

We now describe how to differentiate a vector field along a curve in a manifold M with a connection ∇. Let $\gamma\colon I \to M$ be a smooth immersion, i.e., $d\gamma_t \neq 0$ for all $t \in I$, where $I \subset \mathbf{R}$ is an open interval. By a vector filed along γ we mean a mapping $X\colon I \to TM$ such that $X(t) \in T_{\gamma(t)}M$ for all $t \in I$. Let $\mathcal{X}(\gamma)$ denote the space of vector fields along γ.

For any vector field $X \in \mathcal{X}(\gamma)$, we define another vector field $D_\gamma X \in \mathcal{X}(\gamma)$, called the covariant derivative of X along γ, as follows. First recall that γ is locally one-to-one by the inverse function theorem. Thus, by the previous lemma on the existence of local extensions of vector fields on embedded submanifolds, there exists an open neighborhood U of $\gamma(t_0)$ and a vector field $\overline{X}$ defined on U such that $\overline{X}_{\gamma(t)} = X(t)$ for all $t \in (t_0 - \epsilon, t_0 + \epsilon)$. Set

$$D_\gamma X(t_0) := \nabla_{\gamma'(t_0)}\overline{X}.$$

Recall that $\gamma'(t_0) := d\gamma_{t_0}(1) \in T_{\gamma(t_0)}M$. By Theorem 0.1, $D_\gamma X(t_0)$ is well defined, i.e., it does not depend on the choice of the local extension $\overline{X}$. Thus we obtain a mapping $D_\gamma\colon \mathcal{X}(\gamma) \to \mathcal{X}(\gamma)$. Note that if X, $Y \in \mathcal{X}(\gamma)$, then $(X+Y)(t) := X(t) + Y(t) \in \mathcal{X}(\gamma)$. Further, if $f\colon I \to \mathbf{R}$ is any function then $(fX)(t) := f(t)X(t) \in \mathcal{X}(\gamma)$. It is easy to check that

$$D_\gamma(X+Y) = D_\gamma(X) + D_\gamma(Y) \quad \text{and} \quad D_\gamma(fX) = fD_\gamma(X).$$

Proposition 0.3. *If $\gamma\colon I \to \mathbf{R}^n$, and $X \in \mathcal{X}(\gamma)$, then $D_\gamma X = X'$. In particular, $D_\gamma\gamma' = \gamma''$.*

Proof. Let $\overline{X}$ be a vector field on an open neighborhood of $\gamma(t_0)$ such that

$$\overline{X}(\gamma(t)) = X(t),$$

for all $t \in (t_0 - \epsilon, t_0 + \epsilon)$. Then

$$D_\gamma X(t_0) = \nabla_{\gamma'(t_0)}\overline{X} = (\overline{X} \circ \gamma)'(t_0) = X'(t_0).$$

□

Corollary 0.4. *Let M be an immersed submanifold of $\mathbf{R}^n$ with the induced connection $\overline{\nabla}$, and corresponding covariant derivative $\overline{D}$. Suppose $\gamma\colon I \to M$ is an immersed curve, and $X \in \mathcal{X}_M(\gamma)$ is a vector field along γ in M. Then $\overline{D}_\gamma X = (X')^\top$.* □

0.5 Geodesics

Note that, by the last exercise, the only curves $\gamma\colon I \to \mathbf{R}^n$ with the property that

$$D_\gamma \gamma' \equiv 0$$

are given by $\gamma(t) = at + b$, which trace straight lines. With this motivation, we define a *geodesic* (which is meant to be a generalization of the concept of lines) as an immersed curve $\gamma\colon I \to M$ which satisfies the above equality for all $t \in I$. A nice supply of examples of geodesics are provided by the following observation:

Proposition 0.5. *Let $M \subset \mathbf{R}^n$ be an immersed submanifold, and $\gamma\colon I \to M$ an immersed curve. Then γ is a geodesic of M (with respect to the induced connection from $\mathbf{R}^n$) if and only if $\gamma''^\top \equiv 0$. In particular, if $\gamma\colon I \to M$ is a geodesic, then $\|\gamma'\| = const$.* □

Proof. The first claim is an immediate consequence of the last two results. The last sentence follows from the leibnitz rule for differentiating inner products in Euclidean space: $\langle \gamma', \gamma' \rangle' = 2\langle \gamma'', \gamma' \rangle$. Thus if $\gamma''^\top \equiv 0$, then $\|\gamma'\|^2 = const$. □

As an application of the last result, we can show that the geodesics on the sphere $\mathbf{S}^2$ are those curves which trace a great circle with constant speed:

Example 0.6 (Geodesics on $\mathbf{S}^2$)**.** A C^2 immersion $\gamma\colon I \to \mathbf{S}^2$ is a geodesic if and only if γ has constant speed and lies on a plane which passes through the center of the sphere, i.e., it traces a segment of a great circle.

First suppose that $\gamma\colon I \to \mathbf{S}^2$ has constant speed, i.e. $\|\gamma'\| = const.$, and that γ traces a part of a great circe, i.e., $\langle \gamma, u \rangle = 0$ for some fixed vector $u \in \mathbf{S}^2$ (which is the vector orthogonal to the plane in which γ lies). Since $\langle \gamma', \gamma' \rangle = \|\gamma'\|^2$ is constant, it follows from the Leibnitz rule for differentiating the innerproduct that $\langle \gamma'', \gamma' \rangle = 0$. Furthermore, differentiating $\langle \gamma, u \rangle = 0$ yields that $\langle \gamma'', u \rangle = 0$. So, γ'' lies in the plane of γ, and is orthogonal to γ. So, since γ traces a circle, γ'' must be parallel to γ. This in turn implies that γ'' must be orthogonal to $T_\gamma \mathbf{S}^2$, since γ is orthogonal to $T_\gamma \mathbf{S}^2$. So we conclude that $(\gamma'')^\top = 0$.

Conversely, suppose that $(\gamma'')^\top = 0$. Then γ'' is parallel to γ. So if $u := \gamma \times \gamma'$, then $u' = \gamma' \times \gamma' + \gamma \times \gamma'' = 0 + 0 = 0$. So u is constant. But γ is orthogonal to u, so γ lies in the plane which passes through the origin and is orthogonal to u. Finally, γ has constant speed by the last proposition.

0.6 Ordinary differential equations

In order to prove an existence and uniqueness result for geodesic in the next section we need to develop first a basic result about differential equations:

Theorem 0.7. *Let $U \subset \mathbf{R}^n$ be an open set and $F\colon U \to \mathbf{R}^n$ be C^1, then for every $x_0 \in U$, there exists an $\overline{\epsilon} > 0$ such that for every $0 < \epsilon < \overline{\epsilon}$ there exists a* unique *curve $x\colon (-\epsilon, \epsilon) \to U$ with $x(0) = x_0$ and $x'(t) = F(x(t))$.*

Note that, from the geometric point of view the above theorem states that there passes an integral curve through every point of a vector field. To prove this result we need a number of preliminary results. Let $I \subset \mathbf{R}$ be an interval, (X, d) be a compact metric space, and $\Gamma(I, X)$ be the space of maps $\gamma\colon I \to X$. For every pair of curves γ_1, $\gamma \in \Gamma(I, X)$ set

$$\delta(\gamma_1, \gamma_2) := \sup_{t \in I} d\big(\gamma_1(t), \gamma_2(t)\big).$$

It is easy to check that (Γ, δ) is a metric space. Now let $C(I, X) \subset \Gamma(I, X)$ be the subspace of consisting of *continuous* curves.

Lemma 0.8. *(C, δ) is a complete metric space.*

Proof. Let $\gamma_i \in C$ be a Cauchy sequence. Then, for every $t \in I$, $\gamma_i(t)$ is a Cauchy sequence in X. So $\gamma_i(t)$ converges to a point $\overline{\gamma}(t) \in X$ (since every compact metric space is complete). Thus we obtain a mapping $\overline{\gamma}\colon I \to X$. We claim that $\overline{\gamma}$ is continuous which would complete the proof. By the triangular inequality,

$$\begin{aligned} d(\overline{\gamma}(s), \overline{\gamma}(t)) &\leq d(\overline{\gamma}(s), \gamma_i(s)) + d(\gamma_i(s), \gamma_i(t)) + d(\gamma_i(t), \overline{\gamma}(t)) \\ &\leq 2\delta(\overline{\gamma}, \gamma_i) + d(\gamma_i(s), \gamma_i(t)). \end{aligned}$$

So, since γ_i is continuous,

$$\lim_{t \to s} d(\overline{\gamma}(s), \overline{\gamma}(t)) \leq 2\delta(\overline{\gamma}, \gamma_i).$$

All we need then is to check that $\lim_{i \to \infty} \delta(\overline{\gamma}, \gamma_i) = 0$: Given $\epsilon > 0$, choose i sufficiently large so that $\delta(\gamma_i, \gamma_j) < \epsilon$ for all $j \geq i$. Then, for all $t \in I$, $d(\gamma_i(t), \gamma_j(t)) \leq \epsilon$, which in turn yields that $d(\gamma_i(t), \overline{\gamma}(t)) \leq \epsilon$. So $\delta(\gamma_i, \overline{\gamma}) \leq \epsilon$. □

Now we are ready to prove the main result of this section:

Proof of Theorem 0.7. Let $B = B_r^n(x_0)$ denote a ball of radius r centered at x_0. Choose $r > 0$ so small that that $\overline{B} \subset U$. For any continuous curve $\alpha \in C((-\epsilon, \epsilon), \overline{B})$ we may define another continuous curve $s(\alpha) \in ((-\epsilon, \epsilon), \mathbf{R}^n)$ by

$$s(\alpha)(t) := x_0 + \int_0^t F(a(u))du.$$

We claim that if ϵ is small enough, then $s(\alpha) \in C((-\epsilon, \epsilon), \overline{B})$. To see this note that

$$\|s(\alpha)(t) - x_0\| = \left\| \int_0^t F(\alpha(u))du \right\| \leq \int_0^t \|F(\alpha(u))\| du \leq \epsilon \sup_{\overline{B}} \|F\|.$$

So setting $\epsilon \leq r/\sup_{\overline{B}} \|F\|$, we may then assume that

$$s \colon C((-\epsilon, \epsilon), \overline{B}) \to C((-\epsilon, \epsilon), \overline{B}).$$

Next note that for every α, $\beta \in C((-\epsilon, \epsilon), \overline{B})$, we have

$$\delta(s(\alpha), s(\beta)) = \sup_t \left\| \int_0^t F(\alpha(u)) - F(\beta(u))du \right\| \leq \sup_t \int_0^t \|F(\alpha(u)) - F(\beta(u))\| du.$$

Further recall that, since F is C^1, by the mean value theorem there is a constant K such that

$$\|F(x) - F(y)\| \leq K\|x - y\|,$$

for all x, $y \in \overline{B}$ (in particular recall that we may set $K := \sqrt{n} \sup_{\overline{B}} |D_j F^i|$). Thus

$$\int_0^t \|F(\alpha(u)) - F(\beta(u))\| du \leq K \int_0^t \|\alpha(u) - \beta(u)\| du \leq K\epsilon\delta(\alpha, \beta).$$

So we conclude that

$$\delta(s(\alpha), s(\beta)) \leq K\epsilon\delta(\alpha, \beta).$$

Now assume that $\epsilon < 1/K$ (in addition to the earlier assumption that $\epsilon \leq r/\sup_{\overline{B}} \|F\|$), then, s must have a unique fixed point since it is a contraction mapping. So for every $0 < \epsilon < \overline{\epsilon}$ where

$$\overline{\epsilon} := \min \left\{ \frac{r}{\sup_{\overline{B}} \|F\|}, \frac{1}{\sqrt{n} \sup_{\overline{B}} |D_j F^i|} \right\}$$

there exists a unique curve $x \colon (-\epsilon, \epsilon) \to \overline{B}$ such that $x(0) = s(x)(0) = x_0$, and $x'(t) = s(x)'(t) = F(x(t))$.

It only remains to show that $x \colon (-\epsilon, \epsilon) \to U$ is also the unique curve with $x(0) = x_0$ and $x'(t) = F(x(t))$, i.e., we have to show that if $y \colon (-\epsilon, \epsilon) \to U$ is any curve with $y(0) = x_0$ and $y'(t) = F(y(t))$, then $y = x$ (so far we have proved this only for $y \colon (-\epsilon, \epsilon) \to \overline{B}$). To see this recall that $\epsilon \leq r/\sup_{\overline{B}} \|F\|$ where r is the radius of $\overline{B}$. Thus

$$\|y(t) - x_0\| \leq \int_0^t \|y'(u)\| du = \int_0^t \|F(y(u))\| du \leq \epsilon \sup_{\overline{B}} \|F\| \leq r.$$

So the image of y lies in $\overline{B}$, and therefore we must have $y = x$. □

0.7 Existence and uniqueness of geodesics

Note that for every point $p \in \mathbf{R}^n$ and and vector $X \in T_p\mathbf{R}^n \simeq \mathbf{R}^n$, we may find a geodesic through p and with velocity vector X at p, which is given simply by $\gamma(t) = p + Xt$. Here we show that all manifolds with a connection share this property:

Theorem 0.9. *Let M be a manifold with a connection. Then for every $p \in M$ and $X \in T_pM$ there exists an $\overline{\epsilon} > 0$ such that for every $0 < \epsilon < \overline{\epsilon}$ there is a* unique *geodesic $\gamma\colon (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ and $\gamma'(0) = X$.*

To prove this theorem, we need to record some preliminary observations. Let M and $\widetilde{M}$ be manifolds with connections ∇ and $\widetilde{\nabla}$ respectively. We say that a diffeomorphism $f\colon M \to \widetilde{M}$ is connection preserving provided that

$$(\nabla_Y X)_p = \big(\widetilde{\nabla}_{df(Y)} df(X)\big)_{f(p)}$$

for all $p \in M$ and all vector fields X, $Y \in \mathcal{X}(M)$. It is an immediate consequence of the definitions that

Lemma 0.10. *Let $f\colon M \to \widetilde{M}$ be a connection preserving diffeomorphism. Then $\gamma\colon I \to M$ is a geodesic if and only of $f \circ \gamma$ is a geodesic.* $\square$

Note that if $f\colon M \to \widetilde{M}$ is a diffeomorphism, and M has a connection ∇, then f induces a connection $\widetilde{\nabla}$ on $\widetilde{M}$ by

$$(\widetilde{\nabla}_{\widetilde{Y}} \widetilde{X})_{\widetilde{p}} := \big(\nabla_{df^{-1}(\widetilde{X})} df^{-1}(\widetilde{Y})\big)_{f^{-1}(\widetilde{p})}.$$

It is clear that then $f\colon M \to \widetilde{M}$ will be connection preserving. So we may conclude that

Lemma 0.11. *Let (U, ϕ) be a local chart of M, then $\gamma\colon I \to U$ is a geodesic if and only of $\phi \circ \gamma$ is a geodesic with respect to the connection induced on $\mathbf{R}^n$ by ϕ.* $\square$

Now we are ready to prove the main result of this section:

Proof of Theorem 0.9. Let (U, ϕ) be a local chart of M centered at p and let ∇ be the connection which is induced on $\phi(U) = \mathbf{R}^n$ by ϕ. We will show that there exists an $\overline{\epsilon} > 0$ such that for every $0 < \epsilon < \overline{\epsilon}$ there is a unique geodesic $c\colon (-\epsilon, \epsilon) \to \mathbf{R}^n$, with respect to the induced connection, which satisfies the initial conditions

$$c(0) = \phi(p) \quad \text{and} \quad c'(0) = d\phi_p(X).$$

Then, by a previous lemma, $\gamma := \phi^{-1} \circ c\colon (-\epsilon, \epsilon) \to M$ will be a geodesic on M with $\gamma(0) = p$ and $\gamma'(0) = X$. Furthermore, γ will be unique. To see this suppose that $\overline{\gamma}\colon (-\epsilon, \epsilon) \to M$ is another geodesic with $\overline{\gamma}(0) = p$ and $\overline{\gamma}'(0) = X$. Let ϵ' be the supremum of $t \in [0, \epsilon]$ such that $\overline{\gamma}(-t, t) \subset U$, and set $\overline{c} := \phi \circ \overline{\gamma}|_{(-\epsilon', \epsilon')}$.

Then, by Theorem 0.7, $c = \overline{c}$ on $(-\epsilon', \epsilon')$, because $\epsilon' < \overline{\epsilon}$. So it follows that $\gamma = \overline{\gamma}$ on $(-\epsilon', \epsilon')$, and we are done if $(-\epsilon', \epsilon') = (-\epsilon, \epsilon)$. This is indeed the case, for otherwise, $(-\epsilon' - \delta, \epsilon' + \delta) \subset (-\epsilon, \epsilon)$, for some $\delta > 0$. Further $\overline{\gamma}(\pm\epsilon') = \gamma(\pm\epsilon') \in U$. So if δ is sufficiently small, then $\overline{\gamma}(-\epsilon' - \delta, \epsilon' + \delta) \subset U$, which contradicts the definition of ϵ'.

So all we need is to establish the existence and uniqueness of the geodesic $c\colon (-\epsilon, \epsilon) \to \mathbf{R}^n$ mentioned above. For c to be a geodesic we must have

$$D_c c' \equiv 0.$$

We will show that this may be written as a system of ordinary differential equations. To see this first recall that

$$D_c \dot{c}(t) = \nabla_{\dot{c}(t)} \overline{\dot{c}}$$

where $\overline{\dot{c}}$ is a vector filed in a neighborhood of $c(t)$ which is a local extension of $\dot{c}$, i.e.,

$$\overline{\dot{c}}(c(t)) = \dot{c}(t).$$

By (1) we have

$$\nabla_{\dot{c}(t)} \overline{\dot{c}} = \sum_k \Big(\dot{c}(t)(\overline{\dot{c}}^k) + \sum_{ij} \dot{c}^i(t) \dot{c}^j(t) \Gamma_{ij}^k(c(t)) \Big) e_k,$$

where e_i are the standard basis of $\mathbf{R}^n$ and $\Gamma_{ij}^k(p) = \langle (\nabla_{e_i} e_j)_p, e_k \rangle$. But

$$\dot{c}(t)(\overline{\dot{c}}^k) = (\overline{\dot{c}}^k \circ c)'(t) = (\dot{c}^k)'(t) = \ddot{c}^k(t).$$

So $D_c c' \equiv 0$ if and only if

$$\ddot{c}^k(t) + \sum_{ij} \dot{c}^i(t) \dot{c}^j(t) \Gamma_{ij}^k(c(t)) = 0$$

for all $t \in I$ and all k. This is a system of n second order ordinary differential equations (ODEs), which we may rewrite as a system of $2n$ first order ODEs, via substitution $\dot{c} = v$. Then we have

$$\begin{aligned} \dot{c}^k(t) &= v^k(t) \\ \dot{v}^k(t) &= -\sum_{ij} v^i(t) v^j(t) \Gamma_{ij}^k(c(t)). \end{aligned}$$

Now let $\alpha(t) := (c(t), v(t))$, and define $F\colon \mathbf{R}^{2n} \to \mathbf{R}^{2n}$, $F = (F^1, \ldots, F^{2n})$ by

$$F^\ell(x, y) = y_\ell, \qquad \text{and} \qquad F^{\ell+n}(x, y) = -\sum_{ij} y^i y^j \Gamma_{ij}^\ell(x)$$

for $\ell = 1, \ldots, n$. Then the system of $2n$ ODEs mentioned above may be rewritten as

$$\alpha'(t) = F(\alpha(t)),$$

which has a unique solution with initial conditions $\alpha(0) = (\phi(p), d\phi(X))$.

□

0.8 Parallel translation

Let M be a manifold with a connection, and $\gamma\colon I \to M$ be an immersed curve. Then we say that a vector field $X \in \mathcal{X}(\gamma)$ is *parallel* along γ if

$$D_\gamma X' \equiv 0.$$

Thus, in this terminology, γ is a geodesic if its velocity vector field is parallel. Further note that if M is a submanifold of $\mathbf{R}^n$, the, by the earlier results in this section, X is parallel along γ if and only $(X')^\top \equiv 0$.

Example 0.12. Let M be a two dimensional manifold immersed in $\mathbf{R}^n$, $\gamma\colon I \to M$ be a geodesic of M, and $X \in \mathcal{X}_M(\gamma)$ be a vector field along γ in M. Then X is parallel along γ if and only if X has constant length and the angle between $X(t)$ and $\gamma'(t)$ is constant as well. To see this note that $(\gamma'')^\top \equiv 0$ since γ is a geodesic; therefore,

$$\langle X, \gamma'\rangle' = \langle X', \gamma'\rangle + \langle X, \gamma''\rangle = \langle X', \gamma'\rangle.$$

So, if $(X')^\top = 0$, then it follows that $\langle X, \gamma'\rangle$ is constant which since γ' and X have both constant lengths, implies that the angle between X and γ' is constant. Conversely, suppose that X has constant length and makes a constant angle with γ'. Then $\langle X, \gamma'\rangle$ is constant, and the displayed expression above implies that $\langle X, \gamma'\rangle = 0$ is constant. Furthermore, $0 = \langle X, X\rangle' = 2\langle X, X'\rangle$. So $X'(t)$ is orthogonal to both $X(t)$ and $\gamma'(t)$. If $X(t)$ and $\gamma'(t)$ are linearly dependent, then this implies that $X'(t)$ is orthogonal to $T_{\gamma(t)}M$, i.e., $(X')^\top \equiv 0$. If $X(t)$ and $\gamma'(t)$ are linearly dependent, then $(X')^\top = D_\gamma(X) = D_\gamma(f\gamma') = fD_\gamma(\gamma') \equiv 0$.

Lemma 0.13. *Let $I \subset \mathbf{R}$ and $U \subset \mathbf{R}^n$ be open subsets and $F\colon I \times U \to \mathbf{R}^n$, be C^1. Then for every $t_0 \in I$ and $x_0 \in U$ there exists an $\bar\epsilon >$ such that for every $0 < \epsilon < \bar\epsilon$ there is a unique curve $x\colon (t_0 - \epsilon, t_0 + \epsilon) \to \mathbf{R}^n$ with $x(t_0) = x_0$ and $x'(t) = F(t, x(t))$.*

Proof. Define $\overline{F}\colon I \times U \to \mathbf{R}^{n+1}$ by $\overline{F}(t, x) := (1, F(t, x))$. Then, by Theorem 0.7, there exists an $\bar\epsilon > 0$ and a unique curve $\overline{x}\colon (t_0-\epsilon, t_0+\epsilon) \to \mathbf{R}^{n+1}$, for every $0 < \epsilon < \bar\epsilon$, such that $\overline{x}(t_0) = (1, x_0)$ and $\overline{x}'(t) = \overline{F}(\overline{x}(t))$. It follows then that $\overline{x}(t) = (t, x(t))$, for some unique curve $x\colon (t_0 - \epsilon, t_0 + \epsilon) \to \mathbf{R}^n$. Thus $\overline{F}(\overline{x}(t)) = (1, F(t, x(t)))$, and it follows that $x'(t) = F(t, x(t))$. □

Lemma 0.14. *Let $A(t)$, $t \in I$, be a C^1 one-parameter family of matrices. Then for every $x_0 \in \mathbf{R}^n$ and $t_0 \in I$, there exists a unique curve $x\colon I \to \mathbf{R}^n$ with $x(t_0) = x_0$ such that $x'(t) = A(t) \cdot x(t)$.*

Proof. Define $F\colon I \times \mathbf{R}^n \to \mathbf{R}^n$ by $F_t(x) = A(t) \cdot x$. By the previous lemma, there exists a unique curve $x\colon (t_0 - \epsilon, t_0 + \epsilon) \to \mathbf{R}^n$ with $x(t_0) = x_0$ such that $F_t(x(t)) = x'(t)$ for all $t \in (t_0 - \epsilon, t_0 + \epsilon)$.

Now let $J \subset I$ be the union of all open intervals in I which contains t_0 and such that $x'(t) = F(x(t))$ for all t in those intervals. Then J is open in I and nonempty. All we need then is to show that J is closed, for then it would follow that $J = I$. Suppose that $\bar{t}$ is a limit point of J in I. Just as we argued in the first paragraph, there exists a curve $y\colon (\bar{t} - \bar{\epsilon}, \bar{t} + \bar{\epsilon}) \to \mathbf{R}^n$ such that $y'(t) = F(y(t))$ and $y'(\bar{t}) \neq 0$. Thus we may assume that $y' \neq 0$ on $(\bar{t} - \bar{\epsilon}, \bar{t} + \bar{\epsilon})$, after replacing $\bar{\epsilon}$ by a smaller number. In particular $y'(\widetilde{t}) \neq 0$ for some $\widetilde{t} \in (\bar{t} - \bar{\epsilon}, \bar{t} + \bar{\epsilon}) \cap J$, and there exists a matrix B such that $B \cdot y'(\widetilde{t}) = x'(\widetilde{t})$.

Now let $\overline{y}(t) := B \cdot y(t)$. Since $F(y(t)) = y'(t)$, we have $F(\overline{y}(t)) = \overline{y}'(t)$. Further, by construction $\overline{y}(\widetilde{t}) = x(\widetilde{t})$, so by uniqueness part of the previous result we must have $\overline{y} = x$ on $(\bar{t} - \bar{\epsilon}, \bar{t} + \bar{\epsilon}) \cap J$. Thus x is defined on $J \cup (\bar{t} - \bar{\epsilon}, \bar{t} + \bar{\epsilon})$. But J was assumed to be maximal. So $(\bar{t} - \bar{\epsilon}, \bar{t} + \bar{\epsilon}) \subset J$. In particular $\bar{t} \in J$, which completes the proof that J is closed in I. □

Theorem 0.15. *Let $X\colon I \to M$ be a C^1 immersion. For every $t_0 \in I$ and $X_0 \in T_{\gamma(t_0)}M$, there exists a* unique *parallel vector field $X \in \mathcal{X}(\gamma)$ such that $X(t_0) = X_0$.*

Proof. First suppose that there exists a local chart (U, ϕ) such that $\gamma\colon I \to U$ is an embedding. Let $\overline{X}$ be a vector field on U and set $X(t) := \overline{X}(\gamma(t))$. By (1),

$$D_\gamma(X)(t) = \nabla_{\gamma'(t)}\overline{X} = \sum_k \Big(\gamma'(t)(\overline{X}^k) + \sum_{ij} \gamma^i(t) X^j(t) \Gamma_{ij}^k(\gamma(t))\Big) E_k(\gamma(t)).$$

Further note that

$$\gamma'(t)\overline{X} = (\overline{X} \circ \gamma)'(t) = X'(t).$$

So, in order for X to be parallel along γ we need to have

$$\dot{X}^k + \sum_{ij} \gamma^i(t) \Gamma_{ij}^k(\gamma(t)) X^j(t) = 0,$$

for $k = 1, \dots, n$. This is a linear system of ODE's in terms of X^i, and therefore by the previous lemma it has a unique solution on I satisfying the initial conditions $X^i(t_0) = X_0^i$.

Now let $J \subset I$ be a compact interval which contains t_0. There exists a finite number of local charts of M which cover $\gamma(J)$. Consequently there exist subintervals $J_1, \dots, J_n$ of J such that γ embeds each J_i into a local chart of M. Suppose that $t_0 \in J_\ell$, then, by the previous paragraph, we may extend X_0 to a parallel vector field defined on J_ℓ. Take an element of this extension which lies in a subinterval $J_{\ell'}$ intersecting J_ℓ and apply the previous paragraph to $J_{\ell'}$. Repeating this procedure, we obtain a parallel vector field on each J_i. By the uniqueness of each local extension mentioned above, these vector fields coincide on the overlaps of J_i. Thus we obtain a well-defined vector filed X on J which is a parallel extension of X_0. Note that if $\overline{J}$ is any other compact subinterval of I which contains t_0, and $\overline{X}$ is the parallel extension of X_0 on $\overline{J}$, then X and $\overline{X}$ coincide on $J \cap \overline{J}$, by the uniqueness of local parallel extensions. Thus, since each point of I is contained in a compact subinterval containing t_0, we may consistently define X on all of I.

Finally let $\overline{X}$ be another parallel extension of X_0 defined on I. Let $A \subset I$ be the set of points where $\overline{X} = X$. Then A is closed, by continuity of $\overline{X}$ and X. Further A is open by the uniqueness of local extensions. Furthermore, A is nonempty since $t_0 \in A$. So $A = I$ and we conclude that X is unique. □

Using the previous result we now define, for every $X_0 \in T_{\gamma(t_0)}M$,

$$P_{\gamma,t_0,t}(X_0) := X(t)$$

as the *parallel transport* of X_0 along γ to $T_{\gamma(t)}M$. Thus we obtain a mapping from $T_{\gamma(t_0)}M$ to $T_{\gamma(t)}M$.

Exercise 0.16. Show that $P_{\gamma,t_0,t}\colon T_{\gamma(t_0)}M \to T_{\gamma(t)}M$ is an isomorphism (*Hint:* Use the fact hat $D_\gamma\colon \mathcal{X}(\gamma) \to \mathcal{X}(\gamma)$ is linear). Also show that $P_{\gamma,t_0,t}$ depends on the choice of γ.

Exercise 0.17. Show that

$$\nabla_{\gamma'(t_0)}X = \lim_{t\to t_0} \frac{X_{\gamma(t_0)} - P^{-1}_{\gamma,t_0,t}(X_{\gamma(t)})}{t}.$$

(*Hint:* Use a parallel frame along γ.)