(III) Connections, Parallelism, and Covariant Derivatives.

Consider the geodesic equation again. In (U, x),

$$\ddot{x}^i(t) + \Gamma^i_{jk}(x(t))\dot{x}^j(t)\dot{x}^k(t) = 0, \quad i = 1, \dots, n \qquad (*)$$

Recall under coordinate change $(x^i) \to (y^\alpha)$, the Christoffel symbols behave as

$$\Gamma^i_{jk}(x) = \tilde{\Gamma}^\alpha_{\eta\gamma}(y(x)) \frac{\partial y^\eta}{\partial x^j} \frac{\partial y^\gamma}{\partial x^k} \frac{\partial x^i}{\partial y^\alpha} + \frac{\partial^2 y^\alpha}{\partial x^j \partial x^k} \frac{\partial x^i}{\partial y^\alpha}$$

Therefore Γ^i_{jk} is not coefficients of a tensor!! This fact suggest in particular that we should pay more attention to taking derivative in local coordinates. It would be nice if we have "the derivative of a tensor is again a tensor". This will be solved by ~~introducing~~ so-called "covariant derivatives".

On the other hand, the LHS of $(*)$ behaves under coordinate change $(x^i) \mapsto (y^\alpha)$ as

$$\left(\ddot{x}^i + \Gamma^i_{jk}(x)\dot{x}^j\dot{x}^k\right) = \left(\ddot{y}^\alpha + \tilde{\Gamma}^\alpha_{\eta\gamma}(y(x))\dot{y}^\eta\dot{y}^\gamma\right)\frac{\partial x^i}{\partial y^\alpha}.$$

That is, it behaves like a $(1,0)$-tensor. (i.e. vector field)

Recall, in local coordinates, if $X = X^i \frac{\partial}{\partial x^i} = Y^\alpha \frac{\partial}{\partial y^\alpha}$.

Then $X^i = Y^\alpha \frac{\partial x^i}{\partial y^\alpha}$.

This suggest that $\ddot{x}^i + \Gamma^i_{jk}(x)\dot{x}^j\dot{x}^k$ is coefficients of a

(1,0)-tensor. This ~~will~~ leads to the concepts of connections, and parallelism.

§1. Affine Connections. [WSY Chap.1] [doC, 2.2]

"Differentiate a vector field on a manifold".

On $\mathbb{R}^n$, let v be a vector at $p \in \mathbb{R}^n$, $f \in C^\infty(U)$, $p \in U \subseteq \mathbb{R}^n$, we have the following ~~derivat~~ "directional derivative" ~~of~~ of f (at p) along v:

$$D_v f = \lim_{t \to 0} \frac{f(p+tv) - f(p)}{t}.$$

Let X be a C^∞ vector field. In local coordinates (U, x), we have

$$X = (X^1, \cdots, X^n), \quad X^i \in C^\infty(U).$$

where $\quad X = X^i \frac{\partial}{\partial x^i}$.

Then the directional derivative of X along v ~~at~~ is defined as

$$D_v X = (D_v X^1, \cdots, D_v X^n).$$

That is $\quad D_v X = \sum_i (D_v X^i) \frac{\partial}{\partial x^i}$.

It is direct to check the following properties:

(a) $D_{\alpha v} X = \alpha D_v X, \quad \forall \alpha \in \mathbb{R}$

(b) $D_v (fX) = (D_v f) X + f D_v X, \quad \forall f$

(c) $D_v (X_1 + X_2) = D_v X_1 + D_v X_2, \quad \forall X_1, X_2$

(d) $D_{v_1 + v_2} X = D_{v_1} X + D_{v_2} X, \quad \forall v_1, v_2$.

In fact, we also have

$$D_v \frac{\partial}{\partial x^i} = 0$$

But this property can not be extended to manifold case.

~~This~~ In general, we can define the following concept:

<u>Definition 1</u> (Affine connection). An affine connection ∇ on a ~~d~~ smooth manifold M is a mapping

$$\nabla : \Gamma(TM) \times \Gamma(TM) \longrightarrow \Gamma(TM).$$

($\Gamma(TM)$ is the set of all smooth sections of TM, i.e., the set of all smooth vector fields on M).

which is denoted by $(X, Y) \xrightarrow{\nabla} \nabla_X Y$ and which satisfies the following properties:

(i) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$ (linear over the C^∞ fcts in the argument X).

(ii) $\nabla_X (Y+Z) = \nabla_X Y + \nabla_X Z$

(iii) $\nabla_X (fY) = f\nabla_X Y + X(f)Y$.

in which $X, Y, Z \in \Gamma(TM)$ and f, g are any real-valued C^∞ functions on M. The vector field $\nabla_X Y$ is called the <u>covariant derivative</u> of Y along X (with respect to the connection ∇).

<u>Digest</u>: ①. on $\mathbb{R}^n$, the "directional derivative" provides an affine connection.

For $X, Y \in \Gamma(T\mathbb{R}^n)$, define

$$(\nabla_X Y)(p) = D_{X(p)} Y.$$
$$p \in \mathbb{R}^n$$

Then one can check ~~∇~~ ∇ satisfies (i) – (iii).

② Let $X, Y \in \Gamma(TM)$. In a local coordinates $(U, x^1, \ldots, x^n)$, X, Y can be considered as vector fields on $x(U) \subset \mathbb{R}^n$. ~~A natural question is~~ In ~~this~~ (U, x), we can define $\nabla_X Y$ as the directional derivative $D_X Y$. A natural question is: Can we obtain an affine connection by defining it as directional derivatives in every local coordinates?

The Answer is No! Suppose we have two coordinates $(U, x^1, \ldots, x^n)$, and $(V, y^1, \ldots, y^n)$. When $U \cap V \neq \phi$, we have

$$D_X Y = \sum_i (D_X f^i) \frac{\partial}{\partial x^i}, \quad \text{where } Y = f^i \frac{\partial}{\partial x^i} \text{ in } U.$$
$$= \sum_i (D_X g^i) \frac{\partial}{\partial y^i}, \quad \text{where } Y = g^i \frac{\partial}{\partial y^i} \text{ in } V.$$
$$= \sum_i (D_X g^i) \frac{\partial x^k}{\partial y^i} \frac{\partial}{\partial x^k}.$$

Need ~~$D_X g^i = D$~~ $D_X f^i = D_X g^j \frac{\partial x^i}{\partial y^j}$

$$= D_X\left(f^k \frac{\partial y^j}{\partial x^k}\right) \frac{\partial x^i}{\partial y^j}$$
$$= D_X f^k \cdot \delta^i_k + f^k D_X\left(\frac{\partial y^j}{\partial x^k}\right) \frac{\partial x^i}{\partial y^j}$$
$$= D_X f^i + f^k \left[D_X(\delta^i_k) - \frac{\partial y^j}{\partial x^k} D_X \frac{\partial x^i}{\partial y^j} \right]$$

$$= D_x f^i - f^k \frac{\partial y^j}{\partial x^k} D_x \frac{\partial x^i}{\partial y^j}$$

$$= D_x f^i - g^j D_x \frac{\partial x^i}{\partial y^j} .$$

That is we need

$$g^j D_x \frac{\partial x^i}{\partial y^j} = 0, \ \forall i \qquad (*)$$

We can find examples that (*) does not hold.

③ Existence: Many "trivial" connections: Fix a coordinate neighborhood U, define a "local" connection ∇^U on U via directional derivatives on $\mathbb{R}^n$. This can be extended "trivially" to a connection on M.

Lemma: The set of all affine connections on M form a convex set. Namely, if $\nabla^{(1)}, \cdots, \nabla^{(k)}$ are affine connections on M, and $f_1, \cdots, f_k \in C^\infty(M)$, s.t. $\sum_i f_i = 1$. Then $\sum_i f_i \nabla^{(i)}$ is also an affine connection on M.

Proof: Properties (i) (ii) of an affine connection can be checked directly.

For (iii), we check for $X, Y \in \Gamma(TM)$, $g \in C^\infty(M)$

$$\left(\sum_i f_i \nabla^{(i)}\right)_X (gY) = \sum_i f_i \left(\nabla_X^{(i)}(gY)\right)$$

$$= \sum_i f_i \left(X(g)Y + g\nabla_X^{(i)} Y\right)$$

$$= \left(\sum_i f_i\right) X(g) Y + g\left(\sum_i f_i \nabla^{(i)}\right)_X Y.$$

Here we need the property that $\sum_i f_i = 1$. $\square$

Exercise 1. Find a nontrivial connection on M via "partition of unity".

④. Locality : " $\nabla_X Y$ depends only on local information of X and Y ".

Proposition 1. For any open subset $U \subset M$, if

$$X|_U = \tilde{X}|_U$$

and

$$Y|_U = \tilde{Y}|_U,$$

then

$$\nabla_X Y|_U = \nabla_{\tilde{X}} \tilde{Y}|_U.$$

Proof : We will show $\nabla_X Y|_U \underset{(1)}{=} \nabla_{\tilde{X}} Y|_U \underset{(2)}{=} \nabla_{\tilde{X}} \tilde{Y}|_U$.

For (1), it's enough to show $X|_U = 0 \Rightarrow \nabla_X Y|_U = 0$ (a)

For (2), it's enough to show $Y|_U = 0 \Rightarrow \nabla_X Y|_U = 0$ (b)

Proof of (a) : $\forall\, p \in U$, $\exists$ open $V \subset U$ and a function $f \in C_0^\infty(U)$ s.t.

$$f = 1 \text{ on } V$$

We check $(1-f)X = X$ since $X|_U = 0$.

Then $\nabla_X Y = \nabla_{(1-f)X} Y \overset{\text{Connection (i)}}{=} (1-f)\nabla_X Y$.

In particular, $\nabla_X Y(p) = (1-f(p))\nabla_X Y = 0$.

Therefore $\nabla_X Y|_U = 0$.

Exercise 2 : Show that $Y|_U = 0$ implies $\nabla_X Y|_U = 0$.

Proposition 2. If $X(p) = \tilde{X}(p)$, then $\nabla_X Y(p) = \nabla_{\tilde{X}} Y(p)$.

Proof: Again, it's enough to show $X(p) = 0 \Rightarrow \nabla_X Y(p) = 0$. (*)

By Prop 1, we only need to show (*) for X supported in an coordinate neighborhood (U, x), with $x(p) = 0$, the origin of $\mathbb{R}^n$.

Now we can write

$$X = X^i \frac{\partial}{\partial x^i}$$

with $X^i(0) = 0$. By Taylor's theorem, $\exists$ functions X^i_k s.t.

$$X^i(x^1, \dots, x^n) = X^i(0) + x^k X^i_k = x^k X^i_k.$$

So

$$\nabla_X Y = \nabla_{x^k X^i_k \frac{\partial}{\partial x^i}} Y = x^k \nabla_{X^i_k \frac{\partial}{\partial x^i}} Y.$$

In particular at p,

$$\nabla_X Y(p) = x^k(p) \nabla_{X^i_k \frac{\partial}{\partial x^i}} Y(p) = 0. \qquad \square$$

Consequently, for $v \in T_pM$, and $Y \in \Gamma(TM)$, we can define $\nabla_v Y(p) := \nabla_X Y(p)$, where X is any vectorfield with $X(p) = v$.

(This is like a "directional derivative" of Y at p along v).

But, it is not true that $Y(p) = \tilde{Y}(p) \Rightarrow \nabla_X Y(p) = \nabla_X \tilde{Y}(p)$.

It is not hard to construct counterexamples.

Proposition 3: Let $\gamma : (-\varepsilon, \varepsilon) \to M$ be a smooth curve on M. with $\gamma(0) = p$ and $\dot{\gamma}(0) = v$. Suppose $X, Y, \tilde{Y}$ are vector fields on M s.t.

$$X(p) = v, \quad Y(\gamma(t)) = \tilde{Y}(\gamma(t)), \quad -\varepsilon < t < \varepsilon.$$

then: $\nabla_X Y(p) = \nabla_X \tilde{Y}(p)$.

Proof: It's enough to show

$$Y=0 \text{ along } \gamma \Rightarrow \nabla_v Y(p) = 0.$$

Let $(U, x^1, \dots, x^n)$, $p \in U$ be a coordinate neighborhood around p with $x(p)=0$. Let

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

Then $$\nabla_v Y(p) = \nabla_v \left(f^i \frac{\partial}{\partial x^i}\right)_{(p)} = \left(v(f^i)\frac{\partial}{\partial x^i} + f^i \nabla_v \frac{\partial}{\partial x^i}\right)(p)$$

$$= \frac{d}{dt}\Big|_{t=0} f^i \circ \gamma(t) \frac{\partial}{\partial x^i} + f^i(p) \nabla_v \frac{\partial}{\partial x^i}(p).$$

Since $f^i \circ \gamma(t) = 0$, $t \in (-\varepsilon, \varepsilon)$, we have

$$\nabla_v Y(p) = 0. \qquad \square$$

§2. Parallelism.

What is going on geometrically? [Spivak II, Chap 6,]

Consider a curve

$$c: [a,b] \longrightarrow M.$$

By a vector field V along c, we mean

$$t \in [a,b] \longmapsto V(t) \in T_{c(t)}M.$$

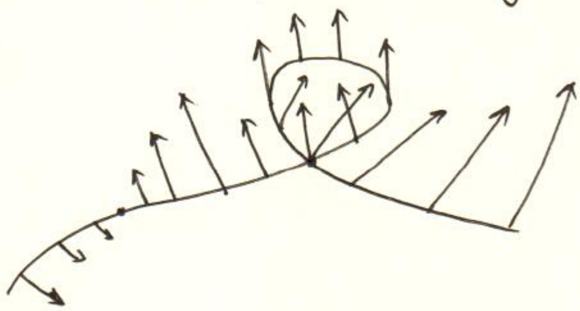
In a coordinate neighborhood $(U, x^1, \dots, x^n)$, we can write

$$V(t) = \sum_{i=1}^{n} v_i(t) \cdot \frac{\partial}{\partial x^i}\Big|_{c(t)}$$

We call V a C^∞ vector field along c if the functions v_i are C^∞ on $[a,b]$; This is equivalent to saying that

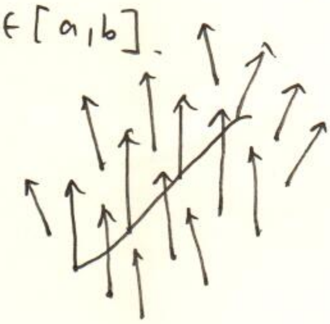
$t \longmapsto V(t)(f)$ is C^∞ for every C^∞ function f on M.

Notice that a vector field V along c may not be extended to a vector field on M.



When c is an embedding, $V(t)$ can be extended to a vector field $\tilde{V}$ on M. We have

$$\tilde{V}(c(t)) = V(t), \quad \forall\, t \in [a,b].$$



~~By locality, we~~ Then

$$(*) \qquad \nabla_{\underbrace{\frac{dc}{dt}}_{dc(\frac{\partial}{\partial t})}} \tilde{V}$$

is a C^∞ vector field along c. By locality, we know $(*)$ ~~only~~ does not depend on the extension of $\tilde{V}$. We call $(*)$ the <u>covariant derivative of V along c</u>. We denote it by the convenient symbolism

$$\frac{DV}{dt}$$

We would like to generalize this covariant derivative along c to any curve c. (This is actually the concept of "<u>induced connection</u>", for which we will discuss later).

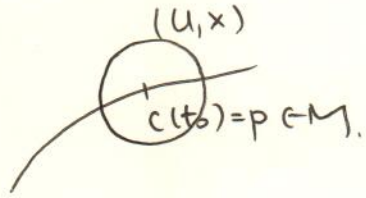
[doC, Prop 2.2]

<u>Proposition 4</u>: Let M be a differentiable manifld with an affine connection ∇. There exists a unique correspondence from C^∞ vector fields V along the smooth curve $c: [a,b] \to M$ to C^∞ vector fields along c: $V \longmapsto \frac{DV}{dt}$, called the <u>covariant derivative</u> of V along c, such that

(a) $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$.

(b). $\frac{D}{dt}(fV) = \frac{df}{dt} V + f \frac{DV}{dt}$, for $f \in C^\infty([a,b])$

(c) If $V(t) = Y(c(t))$ for some ~~$Y \in \Gamma(TM)$~~ C^∞ vector field Y defined in a neighborhood of $c(t)$, then
$$\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} Y.$$



<u>Proof</u>: Let us suppose initially that there exists a correspondence satisfying (a), (b), and (c).

Let $p = c(t_0) \in M$ and $(U, x^1, \ldots, x^n)$ is a coordinate neighborhood of p. For t sufficiently close to t_0, we can ~~write~~ express V locally as
$$V(t) = \sum_{j=1}^{n} v^j(t) \cdot \frac{\partial}{\partial x^j}\Big|_{c(t)}$$

By (a), (b), (c), we have
$$\frac{DV}{dt} \overset{(a)}{=} \sum_{j=1}^{n} \frac{D}{dt}\left(v^j(t) \cdot \frac{\partial}{\partial x^j}\Big|_{c(t)}\right)$$

$$\overset{(b)}{=} \sum_{j=1}^{n} \left[\frac{dv^j(t)}{dt} \cdot \frac{\partial}{\partial x^j}\bigg|_{c(t)} + v^j(t) \cdot \frac{D}{dt}\left(\frac{\partial}{\partial x^j}\bigg|_{c(t)}\right)\right]$$ (73)

$$\overset{(c)}{=} \sum_{j=1}^{n} \left[\frac{dv^j}{dt} \cdot \frac{\partial}{\partial x^j}\bigg|_{c(t)} + v^j(t) \; \nabla_{\frac{dc}{dt}} \frac{\partial}{\partial x^j}\right]$$

$$\frac{dc}{dt} = dc\left(\frac{\partial}{\partial t}\right) = \frac{dc^i}{dt} \frac{\partial}{\partial x^i}\bigg|_{c(t)}$$

$$= \sum_{j=1}^{n} \left[\frac{dv^j}{dt} \cdot \frac{\partial}{\partial x^j}\bigg|_{c(t)} + v^j(t) \frac{dc^i}{dt} \nabla_{\frac{\partial}{\partial x^i}\big|_{c(t)}} \frac{\partial}{\partial x^j}\right]$$

Note $\nabla_{\frac{\partial}{\partial x^i}\big|_{c(t)}} \frac{\partial}{\partial x^j}$ is a C^∞ vector field along c

Hence $\exists \left\{ {k \atop ij} \right\}$ s.t.

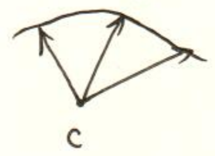
$$\nabla_{\frac{\partial}{\partial x^i}\big|_{c(t)}} \frac{\partial}{\partial x^j} = \left\{ {k \atop ij} \right\}(c(t)) \frac{\partial}{\partial x^k}\bigg|_{c(t)}$$

$$\Rightarrow \quad \frac{DV}{dt} = \sum_{k=1}^{n} \left(\frac{dv^k}{dt} + \sum_{i,j} \left\{ {k \atop ij} \right\}(c(t)) \frac{dc^i}{dt} v^j(t)\right) \frac{\partial}{\partial x^k}\bigg|_{c(t)} \qquad (*)$$

The expression (*) shows us that if there is a correspondence satisfying (a), (b), and, (c), then such a correspondence is unique.

To show existence, define $\frac{DV}{dt}$ in (U, x) by (*). We can ~~show~~ verify that (*) possesses the desired properties. If (V, y) is another coordinate neighborhood with $U \cap V \neq \phi$, Then we define $\frac{DV}{dt}$ in V by (*), the definitions agree in $U \cap V$ by the uniqueness of $\frac{DV}{dt}$ in U. Therefore, the definition can be extended over all of M. $\square$

<u>Remark</u>: Even at points where $\frac{dc}{dt} = 0$, $\frac{DV}{dt}$ is not necessarily 0. !! If c is a constant curve, $c(t) = p \in M, \forall t$.



Then a vector field V along c is just a curve in T_pM, and $\frac{DV}{dt}$ is just the ordinary derivative of this curve.

<u>Definition 2</u> (Parallelism) Let M be a differentiable manifold with an affine connection ∇. A vector field V along a curve $c : [a,b] \to M$ is called <u>parallel</u> when $\frac{DV}{dt} = 0, \forall t \in [a,b]$

When $M = \mathbb{R}^n$, ∇ be the directional derivative, we obtain the standard picture of a parallel vector field.

<u>Proposition 5</u> [doC Prop 2.6]. Let M be a differentiable manifold with an affine connection ∇. Let $c : I \to M$ be a smooth curve in M, and let $V_0 \in T_{c(t_0)}M$, $t_0 \in I$. Then there exists a unique parallel vector field V along c, such that $V(t_0) = V_0$.

<u>Rmk</u>: $V(t)$ is called the <u>parallel transport</u> of $V(t_0)$ along c.

<u>Proof</u>: First consider the case when $c(I)$ is contained in a coordinate neighborhood $(U, x^1, \dots, x^n)$. Then V_0 can be expressed

as: $V_0 = \sum_j v_0^j \frac{\partial}{\partial x_j}\Big|_{c(t_0)}$

Suppose there exists a vector field V in U which is parallel along c, with $V(t_0) = V_0$. Then $V = \sum v^j(t) \frac{\partial}{\partial x^j}\Big|_{c(t)}$

satisfies

$$0 = \frac{DV}{dt} = \sum_k \left(\frac{dv^k}{dt} + \sum_{i,j} \left\{ {k \atop ij} \right\} (c(t)) \frac{dc^i}{dt} v^j(t) \right) \frac{\partial}{\partial x^k}\Big|_{c(t)}$$

The equations

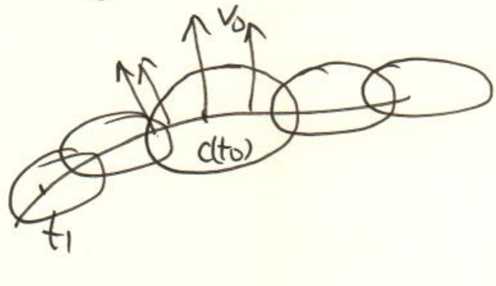
$$\frac{dv^k}{dt} + \sum_{i,j} \left\{ {k \atop ij} \right\} (c(t)) \frac{dc^i}{dt} v^j(t) = 0, \quad k = 1, \dots, n$$

are linear ~~differentiable~~ differential equations. So there is a unique solution satisfying the initial condition

$$v^k(t_0) = v_0^k \, . \quad k = 1, \dots, n.$$

Due it's linearity, the solution is defined for all $t \in I$. This proves the existence and uniqueness of V in this case.

In general, for any $t_1 \in I$ there is a finite cover of $c([t_0, t_1])$ by coordinate neighborhood.

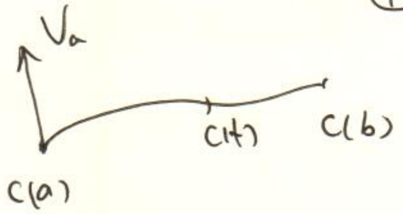


In each of those coordinate neighborhood, V is defined. By uniqueness, the definitions coincide when the intersections are not empty, thus allowing the definition of V along all of $[t_0, t_1]$

$\square$

Now consider $c : [a,b] \to M$

$V_a \in T_{c(a)} M$.



Then there is a unique $V_t \in T_{c(t)} M$

s.t. V_t is the parallel transport of V_a along c.

It is clear from the definition that

$$(V+W)_t = V_t + W_t, \quad (\lambda \cdot V)_t = \lambda V_t.$$

That is, we have a linear transformation

$$P_{c,a,t} = P_t : T_{c(a)} M \longrightarrow T_{c(t)} M.$$
$$V_a \longmapsto V_t$$

Moreover, P_t is one-to-one. Its inverse is given by the parallel transport along the reversed portion of c from t to a.

$$\xi : [a,b] \to M$$
$$t \mapsto c(a+b-t).$$

$$\xi(a) = c(b), \quad \xi(b) = c(a)$$

$$\frac{d\xi}{dt}(t) = -\frac{dc}{dt}(a+b-t)$$

~~Moreover, we check~~ Therefore. $V_b \in T_{c(b)} M$ is the parallel transportation of $V_a \in T_{c(a)} M$ along c iff V_a is the parallel transportation of V_b along ξ.

(When c is embedding, this is seen from

$$\nabla_{\frac{d\xi}{dt}(t)} \tilde{V} = \nabla_{-\frac{dc}{dt}(a+b-t)} \tilde{V} = -\nabla_{\frac{dc}{dt}(a+b-t)} \tilde{V}.)$$

Hence $\boxed{P_t \text{ is an } \boxed{\text{isomorphism}} \text{ between two vector space } T_{c(a)} M \text{ and } T_{c(t)} M.}$

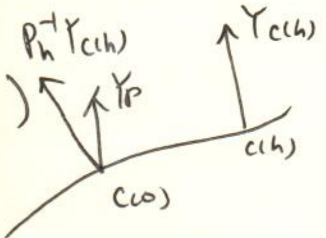
Remark: (justification of the term "connection"). A connection ∇ gives the possibility of comparing, or, "connecting", tangent spaces at different points.

Note the isomorphism between two tangent spaces given by the parallel transport depends on the choice of curves connecting the two points. $\square$

The parallel transport P_t is defined in terms of ∇, but we can also reverse the process.

Proposition 6. [Spivak II, Chapt. 6. Prop 3] Let c be a curve with $c(0) = p$ and $\dot{c}(0) = X_p$. Let $Y \in \Gamma(TM)$. Then

$$\nabla_{X_p} Y = \lim_{h \to 0} \frac{1}{h} \left(P_h^{-1} Y_{c(h)} - Y_p \right)$$



Rmk: Parallel transport enables us to use the idea of "directional derivative" to define $\nabla_{X_p} Y$.

Proof: Let $V_1, \dots, V_n$ be parallel vector fields along c which are linearly independent at $c(0)$, and (since parallel transports are isomorphisms), hence at all points of c.

Set $\quad Y(c(t)) = \sum_{i=1}^{n} f^i(t) V_i(t)$.

Then $\displaystyle \lim_{h \to 0} \frac{1}{h} \left(P_h^{-1} Y_{c(h)} - Y_p \right) = \lim_{h \to 0} \frac{1}{h} \left(\sum_i f^i(h) P_h^{-1} V_i(h) - \sum_i f^i(0) V_i(0) \right)$

$$= \lim_{h\to 0}\frac{1}{h}\sum_i \Big(f^i(t)\, V_i(0) - f^i(0) V_i(0)\Big)$$

$$= \sum_i \lim_{h\to 0}\frac{1}{h}(f^i(t) - f^i(0))\, V_i(0) = \sum_i \frac{df^i}{dt}\Big|_{t=0} V_i(0)$$

$$= \frac{D}{dt}\Big|_{t=0} \underbrace{\sum_i f^i(t) V_i(t)}_{\circledast}$$

$$= \nabla_{X_p} Y. \qquad \square$$

Remark: Recall (*) on page 73. and geodesic equation in last Chapter. If $\gamma : [a,b] \to M$ is a geodesic, then we have

$$\frac{D\dot\gamma(t)}{dt} = 0.$$

where $\frac{D}{dt}$ is determined by a connection ∇ on M, for which

in (U,x), $\quad \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}$.

§3. Covariant derivatives of a tensor field.

In this section, we extend the covariant derivative of a vector field Y along X to that of a tensor field along X. Similar as previous cases, we can do this via pure algebraic discussions, or via parallel transport.

For $(0,0)$-tensor ($=$ functions), we have a nice derivative:

$$\nabla_X : C^\infty(M) \to C^\infty(M) \;:\; f \mapsto \nabla_X f = X(f) = df(X).$$

We can check that this derivative satisfies (i)–(iii) in Def 1 (P. 65).

The following property enables us to define the covariant derivative

$\nabla_X A$ of (r,s)-tensor A. (via an algebraic discussion).

In fact we can define a connection on (r,s)-tensor fields

$$\nabla : \Gamma(TM) \times \Gamma(\otimes^{r,s} TM) \to \Gamma(\otimes^{r,s} TM)$$
$$(X, A) \longmapsto \nabla_X A .$$

Proposition 7: Let M be a differentiable manifold with an affine connection ∇. (on vector fields) There is a unique connection on all tensor fields $\nabla : \Gamma(TM) \times \Gamma(\otimes^{r,s} TM) \to \Gamma(\otimes^{r,s} TM)$ that satisfies.

(i) $\nabla_{fX+gY} A = f \nabla_X A + g \nabla_Y A$.

(ii) $\nabla_X (A_1 + A_2) = \nabla_X A_1 + \nabla_X A_2$

(iii) $\nabla_X (fA) = X(f) A + f \nabla_X A$.

and

(iv) ∇ coincide with the given connections on $\Gamma(TM)$ and $C^\infty(M)$.

(v) $\nabla_X (T_1 \otimes T_2) = (\nabla_X T_1) \otimes T_2 + T_1 \otimes (\nabla_X T_2)$

(vi) $C(\nabla_X T) = \nabla_X C(T)$, where

$$C : \Gamma(\otimes^{r,s} TM) \to \Gamma(\otimes^{r-1, s-1} TM)$$

is the contraction map that pairs the first vector with the first covector.

Rmk: (i)–(iii) is the properties for a connection. (iv)–(vi) provides a unique extension to all tensor fields.

<u>Proof</u>: First, we derive the formula of ∇ on 1-forms.

Let $\omega \in \Omega^1(M) = \Gamma(T^*M)$ be any 1-form, then

$$X(\omega(Y)) \underset{(iv)}{=} \nabla_X(\omega(Y)) = \nabla_X(C(\omega \otimes Y))$$

$$\underset{(vi)}{=} C\,\nabla_X(\omega \otimes Y) \underset{(v)}{=} C(\nabla_X\omega \otimes Y + \omega \otimes \nabla_X Y)$$

$$= (\nabla_X\omega)(Y) + \omega(\nabla_X Y)$$

So we conclude

① $$\boxed{(\nabla_X\omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y)}$$

Next, we can use (v) iteratively to show that for any (r,s)-tensor field A,

$$(\nabla_X A)(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s) = X(A(\omega_1, \dots, \omega_r, Y_1, \dots, Y_s))$$

② $$- \sum_i A(\omega_1, \dots, \nabla_X\omega_i, \dots, \omega_r, Y_1, \dots, Y_s)$$

$$- \sum_j A(\omega_1, \dots, \omega_r, Y_1, \dots, \nabla_X Y_j, \dots, Y_s)$$

This shows the uniqueness.

For the existence, one need to check that the connection defined by ① and ② satisfies all conditions. (i) – (vi). □

<u>Remark</u>: $\nabla_X A$ is called the <u>covariant derivative</u> of the (r,s)-tensor fields A along X.

The properties (iv) – (vi) are very natural. To elaborate this point ~~further~~, We briefly discuss another way ~~to~~ of defining $\nabla_X A$, via parallel transport.

Recall for an isomorphism $\varphi : V \longrightarrow W$ between two vector spaces V and W, there is an induced isomorphism

$$\varphi^* : W^* \longrightarrow V^*$$

between their dual spaces W^*, V^* defined by

for $\alpha \in W^*$: $\quad \varphi^*(\alpha)(v) := \alpha(\varphi(v)), \quad \forall v \in V.$

Then for any $v_i \in V$, $\alpha^j \in V^*$, define

$$\tilde{\varphi}(v_1 \otimes \cdots \otimes v_r \otimes \alpha^1 \otimes \cdots \otimes \alpha^s)$$

$$:= (\varphi(v_1) \otimes \cdots \otimes \varphi(v_r) \otimes (\varphi^*)^{-1}(\alpha^1) \otimes \cdots \otimes (\varphi^*)^{-1}(\alpha^s))$$

Using linearity, we can extend $\tilde{\varphi}$ to $\otimes^{r,s} V$ all (r,s)-tensor over V. This defines an isomorphism between

$$\otimes^{r,s} V \longrightarrow \otimes^{r,s} W.$$

Recall the parallel transport along c. $P_{c,t} : T_{c(0)}M \longrightarrow T_{c(t)}M$ is an isomorphism. We can extend it to be an isomorphism

$$\tilde{P}_{c,t} : \otimes^{r,s} T_{c(0)}M \longrightarrow \otimes^{r,s} T_{c(t)}M.$$

As in Proposition 6, we define

$$\nabla_{X_p} A := \lim_{h \to 0} \frac{1}{h} \left(\tilde{P}_{c,h}^{-1} A_{c(h)} - A_p \right). \quad (**)$$

where c is a curve with $c(0) = p$, $\dot{c}(0) = X_p$.

Clearly if $A \in \Gamma(\otimes^{r,s} TM)$, then $\nabla_{X_p} A \in \Gamma(\otimes^{r,s} TM)$.

We also check that $\nabla_{X_p} A$ given in $(**)$ satisfies prop. (iv) – (vi).

Excercise 3 Let $Y \in \Gamma(TM)$, $\omega, \eta \in \Gamma(T^*M)$.

Consider the tensor field $K = Y \otimes \omega \otimes \eta$.

Let $X_p \in T_pM$, and $\nabla_{X_p} K$ be defined in (**).

(i) Show $\nabla_{X_p} K = \nabla_{X_p} Y \otimes \omega \otimes \eta + Y \otimes \nabla_{X_p}\omega \otimes \eta + Y \otimes \omega \otimes \nabla_{X_p}\eta$.

(ii) Let $CK = \omega(Y)\eta$.

Show $\nabla_{X_p}(CK) = C(\nabla_{X_p} K)$.

Remark: The definition (**) looks to be dependent on X_p and the curve c. However, (**) does not depend on choices of c.

Recall $\nabla_{X_p} Y$ depends only on X_p. We only need to show for any $\eta \in \Gamma(T^*M)$, $\nabla_{X_p}\eta$ also depends only on X_p.

We need show $(\nabla_{X_p}\eta)(Y)$, $\forall\, Y \in \Gamma(TM)$, depends only on X_p, not on c.

Consider $Y \otimes \eta$, we have

$$\nabla_{X_p}(Y \otimes \eta) = (\nabla_{X_p} Y) \otimes \eta + Y \otimes \nabla_{X_p}\eta.$$

exchange with contraction.

$$\Rightarrow \quad X_p(\eta(Y)) = \eta(\nabla_{X_p} Y) + (\nabla_{X_p}\eta)(Y)$$

$$\Leftrightarrow \quad (\nabla_{X_p}\eta)(Y) = X_p(\eta(Y)) - \eta(\nabla_{X_p} Y).$$

RHS only depends on X_p, not on c. $\square$

Now, for any tensor field A, and a vector field X, we can define $(\nabla_X A)(p) = \nabla_{X_p} A$, $\forall\, p \in M$.