

Affine connection

Last lecture we introduced a Riemannian metric on a manifold. Then we defined the length of a curve. Given two points on the manifold, it is desired to find a shortest curve connecting these two points. This would be easy if we are able to find a chart containing these two points in which the Riemannian metric looks as the standard Euclidean metric on \mathbb{R}^m , since we know that the shortest paths for the Euclidean metric are line segments. Hence, our next goal is to show that this is not always possible and that the existence of the chart with the desired properties is obstructed by the curvature of the metric.

In order to build this curvature tensor, we will use that there is a canonical affine connection for the given Riemannian metric, called the Levi-Civita connection. Using this connection, we will use get a differential equation that any distance minimizing curve must satisfy, this is so called geodesic equation.

Definition An affine connection on a smooth manifold M is a map

$$\nabla: \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M),$$

where we will write $\nabla_X Y := \nabla(X, Y)$, such that

$$(AC1) \quad \nabla_{f_1 X_1 + f_2 X_2} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} Y,$$

$$(AC2) \quad \nabla_X (Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2,$$

$$(AC3) \quad \nabla_X fY = (Xf)Y + f \nabla_X Y \quad (\text{Leibniz rule})$$

for every $X_1, X_2, X, Y_1, Y_2, Y \in \mathcal{X}(M)$ and $f_1, f_2, f \in C^\infty(M)$.

Local description of affine connection

Let $\varphi: U \rightarrow \mathbb{R}^m$ be a local chart on M and assume that

$$X = \sum_{i=1}^m a_i(x) \frac{\partial}{\partial x_i} \quad \text{and} \quad Y = \sum_{j=1}^m b_j(x) \frac{\partial}{\partial x_j} \quad \text{on } U.$$

Then

$$\begin{aligned} \nabla_X Y &= \nabla \left(\sum_{i=1}^m a_i(x) \frac{\partial}{\partial x_i}, \sum_{j=1}^m b_j(x) \frac{\partial}{\partial x_j} \right) \\ &= \sum_{i=1}^m a_i(x) \nabla_{\frac{\partial}{\partial x_i}} \left(\sum_{j=1}^m b_j(x) \frac{\partial}{\partial x_j} \right) \\ &= \sum_{i,j=1}^m a_i(x) \left(\frac{\partial}{\partial x_i} b_j \right)(x) \frac{\partial}{\partial x_j} + \sum_{i,j=1}^m a_i(x) b_j(x) \nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} \\ &= \sum_{i,j=1}^m a_i(x) \left(\frac{\partial}{\partial x_i} b_j \right)(x) \frac{\partial}{\partial x_j} + \sum_{i,j,k=1}^m a_i(x) b_j(x) \Gamma_{ij}^k(x) \frac{\partial}{\partial x_k} \end{aligned}$$

where $\Gamma_{ij}^k(x)$ are defined by $\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = \sum_{k=1}^m \Gamma_{ij}^k(x) \frac{\partial}{\partial x_k}$.

Note that $\nabla_X Y$ on U depends only on $X|_U$ and $Y|_U$.

The functions $\Gamma_{ij}^k \equiv \Gamma_{ij}^k(x)$ are so called Christoffel symbols of ∇ in the chart φ . If $\psi: V \rightarrow \mathbb{R}^m$ is another chart in M with coordinate functions x_1', \dots, x_m' with $U \cap V = \emptyset$, then the Christoffel symbols $\Gamma_{i'j'}^{k'}$ of ∇ in the chart ψ can be computed from Γ_{ij}^k on $U \cap V$ in the following way. We will index the coordinate functions for ψ by primed integers $i', j', k' = 1, \dots, m$. Then we know that

$$\frac{\partial}{\partial x_{j'}'} = \sum_{i=1}^m \frac{\partial}{\partial x_i} \frac{\partial(\psi \circ \varphi^{-1})^i}{\partial x_{j'}'}(x) = \sum_{i=1}^m \frac{\partial}{\partial x_i} \Lambda_{j'}^i(x)$$

where $\Lambda_{j'}^i(x) = \frac{\partial(\psi \circ \varphi^{-1})^i}{\partial x_{j'}'}$. Let $\Lambda \equiv \Lambda(x) = (\Lambda_{j'}^i)_{\substack{i=1, \dots, m \\ j'=1, \dots, m}} \in M(m, \mathbb{R})$

and $\Lambda^{-1} = (\Lambda^{-1})_{j'}^{i'} = (\Lambda^{-1})_{i'}^{j'}$ be the inverse matrix so that

$$\frac{\partial}{\partial x_{j'}} = \sum_{i=1}^m \frac{\partial}{\partial x_{i'}} (\Lambda^{-1})_{j'}^{i'}$$

Then we have

$$\begin{aligned} \sum_{k'=1}^m \Gamma_{j'i'}^{k'} \frac{\partial}{\partial x_{k'}'} &= \nabla_{\frac{\partial}{\partial x_{j'}}} \frac{\partial}{\partial x_{i'}} = \sum_{\ell, k=1}^m \nabla_{\frac{\partial}{\partial x_\ell}} \Lambda_{j'}^\ell(x) \frac{\partial}{\partial x_k} \Lambda_{i'}^k(x) \frac{\partial}{\partial x_k} \\ &= \sum_{\ell, k=1}^m \Lambda_{j'}^\ell(x) \nabla_{\frac{\partial}{\partial x_\ell}} \left(\Lambda_{i'}^k(x) \frac{\partial}{\partial x_k} \right) \\ &= \sum_{\ell, k=1}^m \Lambda_{j'}^\ell(x) \left(\frac{\partial}{\partial x_\ell} \Lambda_{i'}^k(x) \right) \frac{\partial}{\partial x_k} \\ &\quad + \sum_{\ell, k=1}^m \Lambda_{j'}^\ell(x) \Lambda_{i'}^k(x) \Gamma_{\ell k}^i(x) \frac{\partial}{\partial x_i} \\ &= \sum_{\ell, k=1}^m \Lambda_{j'}^\ell(x) \left(\frac{\partial}{\partial x_\ell} \Lambda_{i'}^k(x) \right) \sum_{k'=1}^m \frac{\partial}{\partial x_{k'}'} (\Lambda^{-1})_k^{k'} \\ &\quad + \sum_{\ell, k=1}^m \Lambda_{j'}^\ell(x) \Lambda_{i'}^k(x) \Gamma_{\ell k}^i(x) \sum_{k'=1}^m \frac{\partial}{\partial x_{k'}'} (\Lambda^{-1})_i^{k'} \end{aligned}$$

Hence we have:

$$(CCS) \quad \Gamma_{i'j'}^{k'}(x) = \sum_{\ell, k=1}^m \left[\Lambda_{j'}^\ell(x) \left(\frac{\partial}{\partial x_\ell} \Lambda_{i'}^k(x) \right) (\Lambda^{-1})_k^{k'} + \Lambda_{j'}^\ell(x) \Lambda_{i'}^k(x) \Gamma_{\ell k}^i(x) (\Lambda^{-1})_i^{k'} \right]$$

Note that the first term on the right side of (CCS) involves partial derivatives of $\Lambda_{i'}^k(x)$ and so the Christoffel symbols do NOT transform as a tensor field. This is caused by (AC3). However, it can be shown that a difference of two connections (defined in the obvious way) is a tensor field. Hence the space of all affine connections on M is an affine space modelled on the vector space $\mathcal{T}^{2,1}(M)$.

Example (Flat connection on \mathbb{R}^m)

Let us work in the standard chart on \mathbb{R}^2 with coordinates x_1, x_2 and let us consider the associated vector fields

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}.$$

Let ∇ be the affine connection on \mathbb{R}^2 determined by $\Gamma_{ij}^k(x) = 0$ for every $i, j, k = 2$. This is the flat connection on \mathbb{R}^2 .

Let us now consider the polar coordinates on \mathbb{R}^2 which can be viewed as a map

$$\Phi: (0, +\infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2$$

$$\Phi(r, \varphi) = (\underbrace{r \cos \varphi}_{x_1}, \underbrace{r \sin \varphi}_{x_2}).$$

By the chain rule:

$$\frac{\partial}{\partial r} = \frac{\partial}{\partial x_1} \frac{\partial x_1}{\partial r} + \frac{\partial}{\partial x_2} \frac{\partial x_2}{\partial r} = \cos \varphi \frac{\partial}{\partial x_1} + \sin \varphi \frac{\partial}{\partial x_2}$$

$$\frac{\partial}{\partial \varphi} = \frac{\partial}{\partial x_1} \frac{\partial x_1}{\partial \varphi} + \frac{\partial}{\partial x_2} \frac{\partial x_2}{\partial \varphi} = -r \sin \varphi \frac{\partial}{\partial x_1} + r \cos \varphi \frac{\partial}{\partial x_2}$$

and so

$$\begin{aligned} \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} &= \left(\cos \varphi \nabla_{\frac{\partial}{\partial x_1}} + \sin \varphi \nabla_{\frac{\partial}{\partial x_2}} \right) \left(\cos \varphi \frac{\partial}{\partial x_1} + \sin \varphi \frac{\partial}{\partial x_2} \right) \\ &= \left[\cos \varphi \left(\frac{\partial}{\partial x_1} \cos \varphi \right) + \sin \varphi \left(\frac{\partial}{\partial x_2} \cos \varphi \right) \right] \frac{\partial}{\partial x_1} \\ &\quad + \left[\cos \varphi \left(\frac{\partial}{\partial x_1} \sin \varphi \right) + \sin \varphi \left(\frac{\partial}{\partial x_2} \sin \varphi \right) \right] \frac{\partial}{\partial x_2} = (*) \end{aligned}$$

Now it remains to express φ as a function of x_1, x_2 . In the first quadrant this leads to $\tan \varphi = \frac{x_1}{x_2}$, $\varphi = \arctan \frac{x_1}{x_2}$. Hence we get

$$\begin{aligned} (*) &= \left[-\cos \varphi \sin \varphi \frac{\frac{1}{x_2}}{1 + \left(\frac{x_1}{x_2}\right)^2} - \sin \varphi \sin \varphi \frac{\frac{-x_1}{x_2^2}}{1 + \left(\frac{x_1}{x_2}\right)^2} \right] \frac{\partial}{\partial x_1} \\ &\quad + \left[\cos^2 \varphi \frac{\frac{1}{x_2}}{1 + \left(\frac{x_1}{x_2}\right)^2} + \sin^2 \varphi \frac{\frac{-x_1}{x_2^2}}{1 + \left(\frac{x_1}{x_2}\right)^2} \right] \frac{\partial}{\partial x_2} \\ &= \left[-\cos \varphi \sin \varphi \frac{r \sin \varphi}{r^2} + \sin^2 \varphi \frac{r \cos \varphi}{r^2} \right] \frac{\partial}{\partial x_1} \\ &\quad + \left[\cos^2 \varphi \frac{r \sin \varphi}{r^2} + \sin^2 \varphi \frac{-r \sin \varphi}{r^2} \right] \frac{\partial}{\partial x_2} \\ &= 0. \end{aligned}$$

We see that

$$0 = \Gamma_{rr}^r \frac{\partial}{\partial r} + \Gamma_{rr}^\varphi \frac{\partial}{\partial \varphi} = \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r}$$

and so $\Gamma_{rr}^n = \Gamma_{rr}^\varphi$. Here we use the coordinate functions r, φ to index the Christoffel symbols.

Parallel vector fields and geodesics

Proposition Let $X, Y \in \mathcal{X}(M)$ and ∇ be an affine connection on M .
 If $x \in M$, then

$$(\nabla_X Y)(x) \in T_x M$$

depends only on $X(x)$ the values of Y on any curve $\gamma: (-\epsilon, \epsilon) \rightarrow M$ such that $\gamma(0) = x, [\gamma] = X(x) \in T_x M$.

Proof: From Section Local description of affine connection we know that

$$\nabla_X Y(x) = \sum_{k=1}^m \left(\sum_{i,j=1}^m a_i(x) \frac{\partial}{\partial x_i} b_j(x) + a_i(x) b_j(x) \Gamma_{ij}^k(x) \right) \frac{\partial}{\partial x_k}$$

We see that $\nabla_X Y(x)$ depends on $a_i(x), b_j(x)$ and on X by where $i, j = 1, \dots, m$ and $X = X(x) \in T_x M = D_x M$ is the derivation at x determined by $X(x)$. But we know from week 2 that

$$X b_j = \left. \frac{d}{dt} b_j(\gamma(t)) \right|_{t=0}$$

where γ is any curve with $\gamma(0) = x$ and $[\gamma] = X$. This completes the proof. \square

The previous Proposition shows that the following definition makes sense.

Definition Let $\gamma: [a, b] \rightarrow M$ be a curve on M and X be a vector field defined on the image of γ , that is $X: [a, b] \rightarrow TM$ is smooth with $X(t) \in T_{\gamma(t)} M$. The covariant derivative of X along γ is defined as

$$\frac{dX}{dt}(t) = (\nabla_{\gamma'(t)} X)(t).$$

The vector field X is called parallel along γ if $\frac{dX}{dt} = 0$ on $[a, b]$. The curve γ is a geodesic for ∇ if γ' is parallel along γ .

If the image of γ is contained in U and Γ_{ij}^k are the Christoffel symbols for ∇ in the chart $\varphi: U \rightarrow \mathbb{R}^m$ and

$$X(t) = \sum_{i=1}^m a_i(t) \frac{\partial}{\partial x_i}, \text{ then } X \text{ is parallel iff}$$

it satisfies the system of first order ODE's

$$(PT) \quad a'_k(t) + \sum_{i,j=1}^m \Gamma_{ji}^k(\gamma(t)) \gamma'_j(t) a_i(t) = 0, \quad k=1, \dots, m.$$

The curve γ is a geodesic for ∇ iff it satisfies the system of second order ODE's

$$(GE) \quad \gamma''_k(t) + \sum_{i,j=1}^m \Gamma_{ij}^k(\gamma(t)) \gamma'_j(t) \gamma'_i(t) = 0, \quad k=1, \dots, m.$$

Here $\gamma(t) = (\gamma_1(t), \dots, \gamma_m(t))$ in the coordinates on U and so

$$\gamma'(t) = \sum_{j=1}^m \gamma'_j(t) \frac{\partial}{\partial x_j}, \quad t \in [a, b].$$

By Picard-Lindelöf theorem, given $t_0 \in [a, b]$ and a vector $(a_1, \dots, a_m) \in \mathbb{R}^m$, there is a unique parallel vector field X along γ such that $X(t_0) = \sum_{i=1}^m a_i \frac{\partial}{\partial x_i}$.

Moreover, given $x \in X$ and $v \in T_x M$, there is a unique (maximal) geodesic γ s.t. $\gamma(0) = x$ and $\gamma'(0) = v$.

Example: If $M = \mathbb{R}^m$ and ∇ is the flat connection on \mathbb{R}^m with Christoffel symbols $\Gamma_{ij}^k = 0$. Then a vector field

$$X(t) = \sum_{i=1}^m a_i(t) \frac{\partial}{\partial x^i}$$

is parallel along a curve γ iff the coefficient functions a_1, \dots, a_m are constant on the domain of γ . The curve γ is a geodesic iff γ' is constant, hence γ is a parametrization of line segment with constant speed.

Covariant derivative of tensor fields

Theorem Let $\omega \in \Omega^1(M)$ and ∇ be an affine connection on M .

Then $\nabla \omega : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{C}^\infty(M)$ defined by

$$\nabla \omega(X, Y) = X(\omega(Y)) - \omega(\nabla_X Y)$$

is tensor field of type $(2,0)$ on M .

Proof: It is clear that $\nabla \omega$ is multilinear. Using Definition A' of tensor fields, it is enough to show that (LOF) holds for $\nabla \omega$. Let $f, g \in \mathcal{C}^\infty(M)$, then we have

$$\nabla \omega(fX, Y) = fX(\omega(Y)) - \omega(\nabla_{fX} Y)$$

$$= f(X(\omega(Y)) - \omega(\nabla_X Y)) = f \nabla \omega(X, Y),$$

$$\nabla \omega(X, fY) = X(\omega(fY)) - \omega(\nabla_X (fY)) =$$

$$= X(f\omega(Y)) - \omega((Xf)Y + f\nabla_X Y)$$

$$= (Xf)\omega(Y) + fX(\omega(Y)) - (Xf)\omega(Y) - f\omega(\nabla_X Y)$$

$$= f \nabla \omega(X, Y). \quad \square$$

Definition The tensor field $\nabla \omega \in \mathcal{T}^{2,0}(M)$ is called the covariant derivative of $\omega \in \Omega^1(M)$ and we also write

$$(\nabla_X \omega)(Y) := \nabla \omega(X, Y).$$

Theorem Let $T \in \mathcal{T}^{k,l}(M)$ and ∇ be an affine connection on M .

Then $\nabla T : \underbrace{\mathcal{X}(M) \times \dots \times \mathcal{X}(M)}_{k+1 \text{ copies}} \times \underbrace{\Omega^1(M) \times \dots \times \Omega^1(M)}_{l \text{ copies}} \rightarrow \mathcal{C}^\infty(M)$

$$\nabla T(X_0, \dots, X_k, \theta_1, \dots, \theta_l) =$$

$$= X_0 T(X_1, \dots, X_k, \theta_1, \dots, \theta_l) - T(\nabla_{X_0} X_1, X_2, \dots)$$

$$- T(X_1, \nabla_{X_0} X_2, \dots) - \dots - T(X_1, \dots, \theta_{l-1}, \nabla_{X_0} \theta_l)$$

is a tensor field of type $(k+1, l)$ on M .

Proof is just a repetition of that for differential forms. \square

Definition Tensor field $T \in \mathcal{T}^{k,l}(M)$ is called parallel for ∇ if $\nabla T = 0$.

Let $T \in \mathcal{T}^{k,1}(M)$ and $x \in M$, $v_1, \dots, v_k \in T_x M$. Then the map

$$(E) \quad T_x^* M \ni \alpha \mapsto T_x(v_1, \dots, v_k, \alpha) \in \mathbb{R}$$

is linear and so there is a unique $w \in T_x M$ so that (E) is equal to $\alpha \mapsto \alpha(w)$. Moreover, it is easy to see that if $X_1, \dots, X_k \in \mathcal{X}(M)$, then there is a unique smooth vector field $\Psi \in \mathcal{X}(M)$ such that for every $x \in M$ and $\alpha \in T_x M$ we have that

$$T_x^* M \ni \alpha \mapsto T_x(X_1(x), \dots, X_k(x), \alpha) = \alpha(\Psi(x)).$$

Hence we can view T as a map

$$T: \mathcal{X}(M) \times \dots \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M), (X_1, \dots, X_k) \mapsto \Psi$$

which is linear in every argument and if $f \in \mathcal{C}^\infty(M)$, then

$$(LOF)' \quad T(X_1, \dots, f X_i, \dots, X_k) = f T(X_1, \dots, X_k), \quad i=1, \dots, k.$$

Torsion of affine connection

Theorem Let ∇ be an affine connection on M , then

$$T^\nabla \equiv T: \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M)$$

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y], \quad X, Y \in \mathcal{X}(M),$$

is a tensor field of type (2,1) on M .

Proof is left as an exercise. One has to verify that T is linear in every argument and that (LOF)' holds. \square

Definition The tensor field $T \in \mathcal{T}^{2,0}(M)$ defined above is called the torsion tensor (or simply torsion) of ∇ .

The affine connection is called symmetric or torsion-free if $T=0$.

Local formula for torsion tensor

Let $\varphi: U \rightarrow \mathbb{R}^m$ be a chart on M and Γ_{ij}^k be the Christoffel symbols of ∇ in the chart φ . If $X = \sum_{i=1}^m a_i(x) \frac{\partial}{\partial x_i}$ and $Y = \sum_{j=1}^m b_j(x) \frac{\partial}{\partial x_j}$ on U , then

$$\begin{aligned} T(X, Y) &= \sum_{i,j=1}^m \left(a_i(x) \nabla_{\frac{\partial}{\partial x_i}} \left(b_j(x) \frac{\partial}{\partial x_j} \right) - b_j(x) \nabla_{\frac{\partial}{\partial x_j}} \left(a_i(x) \frac{\partial}{\partial x_i} \right) \right) \\ &\quad - \sum_{i,j=1}^m \left(a_i(x) \left(\frac{\partial}{\partial x_i} b_j \right)(x) - b_j(x) \left(\frac{\partial}{\partial x_i} a_i \right)(x) \right) \frac{\partial}{\partial x_j} \\ &= \sum_{i,j,k=1}^m \left(a_i(x) b_j(x) \Gamma_{ij}^k(x) - b_j(x) a_i(x) \Gamma_{ji}^k(x) \right) \frac{\partial}{\partial x_k}. \end{aligned}$$

This shows that

$$T(x) = \sum_{i,j=1}^m T_{ij}^k(x) dx_i \otimes dx_j \otimes \frac{\partial}{\partial x_k}$$

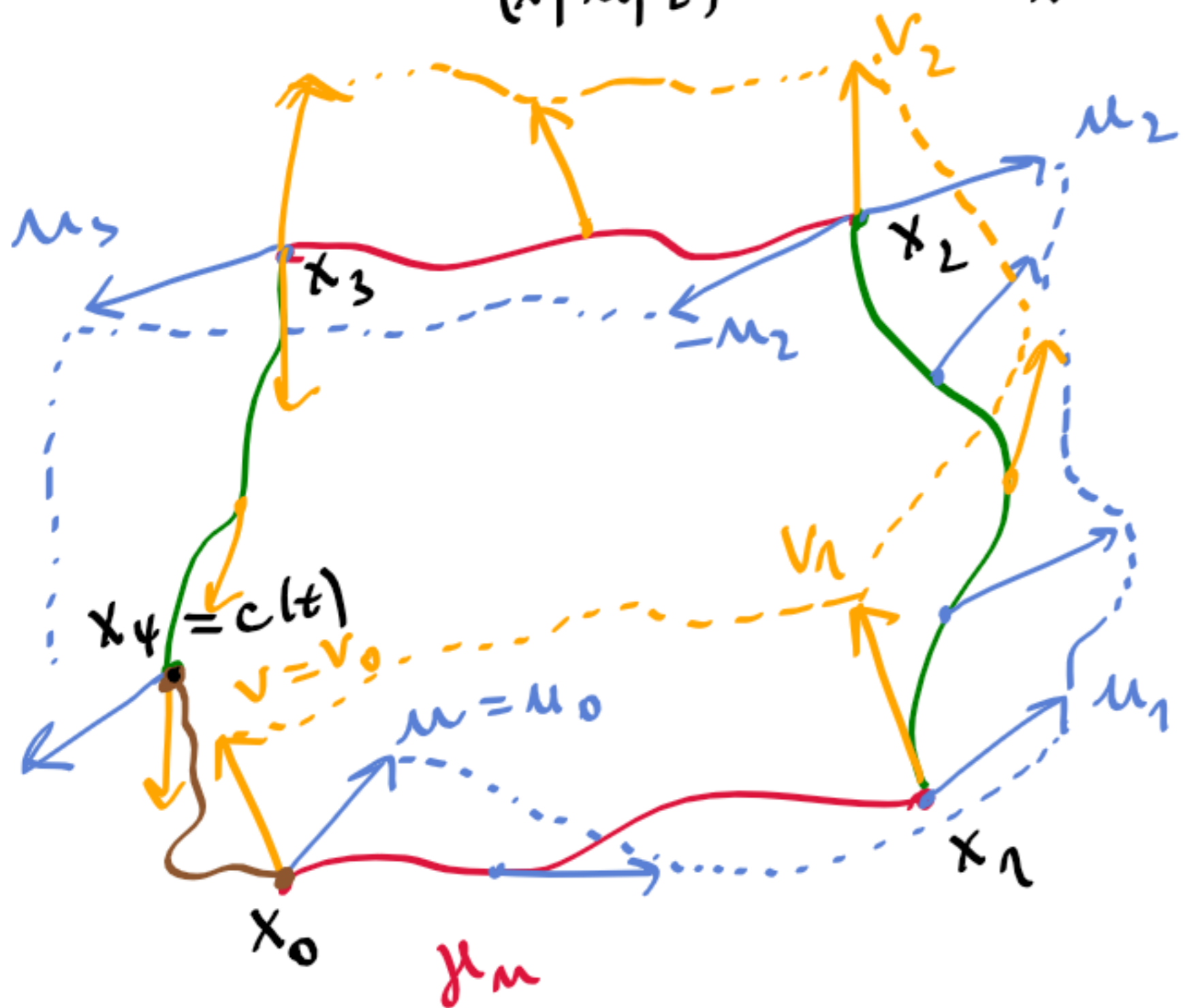
where

$$T_{ij}^k(x) = \Gamma_{ij}^k(x) - \Gamma_{ji}^k(x).$$

Geometric interpretation of torsion

Let x and $u, v \in T_x M$. Let $\gamma_{u,x}$ be the geodesic for \mathcal{D} with $\gamma_{u,x}(0) = x$ and $\dot{\gamma}_{u,x}(0) = u$. Let V be the parallel vector field along $\gamma_{u,x}$ with $V(0) = v$. This defines a linear map (since the parallel transport is governed by linear ODE's) for time t close to 0

$$\Gamma_{(x, u, t)}^V : T_x M \rightarrow T_{\gamma_{u,x}(t)} M, \quad v \mapsto V(t).$$



1) Put $x_0 = x, u_0 = u, v_0 = v$.

2) Put $x_1 = \gamma_{u_0, x_0}(t), u_1 = \Gamma_{(x_1, u_1, t)}(u_0),$

$$v_1 = \Gamma_{(x_0, u_0, t)}(v_0).$$

3) Put $x_2 = \gamma_{v_1, x_1}(t), u_2 = \Gamma_{(x_2, v_2, t)}(u_1)$ and

$$v_2 = \Gamma_{(x_1, v_1, t)}(v_1)$$

4) Put $x_3 = \gamma_{-u_2, x_2}(t), u_3 = \Gamma_{(x_3, -u_3, t)}(u_2)$ and

$$v_3 = \Gamma_{(x_2, -u_2, t)}(v_2).$$

5) Put $x_4(t) = \gamma_{-v_3, x_3}(t).$

Now the map $t \mapsto c(t) = x_4(t)$ is a smooth curve such that $c(0) = x$ and $c'(0) = 0$. If we reparametrize and set $\tilde{c}(s^2) = c(s)$, then \tilde{c} is smooth and

$$\left. \frac{d\tilde{c}}{ds}(s) \right|_{s=0^+} = T_x(u, v).$$