General Relativity: Curvature and torsion of a connection

Gonzalo E. Reyes
Université de Montréal

10 June 2006

The fundamental concepts of Lie bracket, connection, covariant derivative and curvature tensor have been defined in geometric terms, making these notions intuitive and easy to understand. In this section we describe an equivalent, algebraic approach which, although not so perspicuous as the geometric one, is easier to manipulate. This approach is based on the notion of affine space over an R-module. This whole pamphlet, with a few exceptions, is taken almost verbatim from Lavendhomme's book [3].

0.1 Affine spaces and strong difference

Let V be an R-module . An *affine space over* V is a set E together with two operations

$$\begin{cases} \dot{+} : V \times E \longrightarrow E \\ \dot{-} : E \times E \longrightarrow V \end{cases}$$

satisfying the following conditions for $e, e_1, e_2 \in E$ and $v, v_1, v_2 \in V$

$$\begin{cases} (a) & (e_2 \dot{-} e_1) \dot{+} e_1 = e_2 \\ (b) & (v \dot{+} e) \dot{-} e = v \\ (c) & 0 \dot{+} e = e \\ (d) & v_1 \dot{+} (v_2 \dot{+} e) = (v_1 + v_2) \dot{+} e \end{cases}$$

In other words, an affine space E over V is a V-set whose action $V \times E \longrightarrow E$ is faithful and transitive. (V is considered as an abelian group).

Proposition 0.1 (Some elementary consequences)

(1) $e \dot{-} e = 0$

(2) $v_1 \dot{+} e = v_2 \dot{+} e$ *iff* $v_1 = v_2$

(3) $e_3 \dot{-} e_1 = (e_3 \dot{-} e_2) + (e_2 \dot{-} e_1)$

(4) $(e_1 \dot{-} e_2) + (e_2 \dot{-} e_1) = 0$

(5) $e_2 \dot{-} e_1 = 0$ *iff* $e_1 = e_2$

Proof:

(1) $e \dot{-} e =_{(c)} (0 \dot{+} e) \dot{-} e =_{(b)} 0$

(2) Assume $v_1 \dot{+} e = v_2 \dot{+} e$. Then $v_1 =_{(b)} (v_1 \dot{+} e) \dot{-} e = (v_2 \dot{+} e) \dot{-} e =_{(b)} v_2$.

(3) $((e_3 \dot{-} e_2) + (e_2 \dot{-} e_1)) \dot{+} e_1 =_{(d)} (e_3 \dot{-} e_2) \dot{+} ((e_2 \dot{-} e_1) \dot{+} e_1) =_{(a)} (e_3 \dot{-} e_2) \dot{+} e_2$ $=_{(a)} e_3 =_{(a)} (e_3 \dot{-} e_1) \dot{+} e_1$. Now, apply (2).

(4) Put $e_3 = e_1$ in (3) and use (1).

(5) In one direction, use (1). In the other, use conditions (a) and (c) of the definition of affine space.

We shall often use these consequences tacitly.

Example: The euclidean 3-space E is an affine space over R^3 with the operations

$$\begin{cases} \dot{+} : R^3 \times E \longrightarrow E \\ \dot{-} : E \times E \longrightarrow R^3 \end{cases}$$

defined by $v \dot{+} e$ = the end point of the vector obtained by transporting v parallel to itself to the point e and $e_2 \dot{-} e_1$ = the vector which transported parallel to itself to e_1 has end-point e_2.

The affine space E is just like the vector space R^3, but with no element distinguished as the origin. This remark can be made precise as follows: fix $e_0 \in E$ and define $\phi : E \longrightarrow R^3$ by $\phi(e) = e \dot{-} e_0$. Then $\psi : R^3 \longrightarrow E$, defined by $\psi(v) = v \dot{+} e_0$ and ϕ are inverse of each other, as can be checked immediately from properties (a) and (b).

For other examples and general theory see volume 1 of M. Berger [1].

For our purposes, the next example is crucial.

Let $i : D(2) \subset D \times D$. Then $M^i : M^{D\times D} \longrightarrow M^{D(2)}$ is just the restriction map. Consider $\tau \in M^{D(2)}$ and let $x = \tau(0,0)$.

Let M^i_τ be the fiber of M^i over τ and M_x the fiber of π_M over x. Aiming at making M^i_τ into an affine space over M_x, we define two operations

$$\begin{cases} \dot{-} : M^i_\tau \times M^i_\tau \longrightarrow M_x \\ \dot{+} : M_x \times M^i_\tau \longrightarrow M^i_\tau \end{cases}$$

To define $\dot{-} : M^i_\tau \times M^i_\tau \longrightarrow M_x$ we need the following

Lemma 0.2 *The diagram*

$$\begin{array}{ccc} D(2) & \xrightarrow{\quad i \quad} & D \times D \\ \Big\downarrow{\scriptstyle i} & & \Big\downarrow{\scriptstyle \psi} \\ D \times D & \xrightarrow[\quad \phi \quad]{} & (D \times D) \bigvee D \end{array}$$

is an R-colimit, where $(D \times D) \bigvee D = \{(d_1, d_2, e) \in D^3 | d_1 e = d_2 e = 0\}$, $\phi(d_1, d_2) = (d_1, d_2, 0)$ *and* $\psi(d_1, d_2) = (d_1, d_2, d_1 d_2)$.

Let $\gamma_1, \gamma_2 \in M^i_\tau$. Define $(\gamma_2 \dot{-} \gamma_1)(d) = f(0,0,d)$ where $f : (D \times D) \bigvee D \longrightarrow M$ is the unique function satisfying

$$\begin{cases} f \circ \phi = \gamma_1 \\ f \circ \psi = \gamma_2 \end{cases}$$

Then $(\gamma_2 \dot{-} \gamma_1)(0) = f(0,0,0) = (f \circ \phi)(0,0) = \gamma_1(0,0) = x$, i.e. $\gamma_2 \dot{-} \gamma_1 \in M_x$.

The operation $\dot{-}$ is called *strong difference.*

To define the action $\dot{+} : M_x \times M^i_\tau \longrightarrow M^i_\tau$ we need the following

Lemma 0.3 *The diagram*

$$\begin{array}{ccc} 1 & \xrightarrow{\;0\;} & D \\ \Big\downarrow{\scriptstyle (0,0)} & & \Big\downarrow{\scriptstyle \epsilon} \\ D \times D & \xrightarrow[\;\phi\;]{} & (D \times D) \bigvee D \end{array}$$

with $\phi(d_1, d_2) = (d_1, d_2, 0)$ *and* $\epsilon(d) = (0, 0, d)$ *is an R-colimit.*

Let $\gamma : D \times D \longrightarrow M$ with $\gamma|_{D(2)} = \tau$ and $v : D \longrightarrow M$ with $v(0) = x$. Define $(v \dot{+} \gamma)(d_1, d_2) = g(d_1, d_2, d_1 d_2)$, where $g : (D \times D) \bigvee D \longrightarrow M$ is the unique function such that

$$\begin{cases} g \circ \phi = \gamma \\ g \circ \epsilon = v \end{cases}$$

Then, whenever $(d_1, d_2) \in D(2)$. $(v \dot{+} \gamma)(d_1, d_2) = g(d_1, d_2, 0) = (g \circ \phi)(d_1, d_2) = \gamma(d_1, d_2)$, i.e., $(v \dot{+} \gamma)|_{D(2)} = \gamma|_{D(2} = \tau$.

Proposition 0.4 *With the operations thus defined,* M^i_τ *is an affine space over* M_x*.*

Proof:

(a) By definition of $\dot{+}$, $((\gamma_2 \dot{-} \gamma_1) \dot{+} \gamma_1)(d_1, d_2) = g(d_1, d_2, d_1 d_2)$ where g is characterized by $g(d_1, d_2, 0) = \gamma_1(d_1, d_2)$ and $g(0, 0, d) = (\gamma_2 \dot{-} \gamma_1)(d)$. But $(\gamma_2 \dot{-} \gamma_1)(d) = l(0, 0, d)$ where l is characterized by $l(d_1, d_2, 0) = \gamma_1(d_1, d_2)$ and $l(d_1, d_2, d_1 d_2) = \gamma_2(d_1, d_2)$. Hence $l(0.0.d) = g(0.0.d)$ and $l(d_1, d_2, 0) = g(d_1, d_2, 0)$. Thus, $g = l$ and so, $((\gamma_2 \dot{-} \gamma_1) \dot{+} \gamma_1)(d_1, d_2) = g(d_1, d_2, d_1 d_2) = l(d_1, d_2, d_1 d_2) = \gamma_2(d_1, d_2)$.

(b) and (c): Similar and left to the reader

(d) The proof is a little more involved and is based on the easily checked statement that the diagram

$$\begin{array}{ccc} D \times D & \xrightarrow{\;\phi\;} & (D \times D) \bigvee D \\ \Big\downarrow{\scriptstyle \psi} & & \Big\downarrow{\scriptstyle \zeta_2} \\ (D \times D) \bigvee D & \xrightarrow[\;\zeta_1\;]{} & (D \times D) \bigvee D(2) \end{array}$$

where

$$\begin{array}{c}(D\times D)\bigvee D(2)=\{(d_1,d_2,e_1,e_2)\in D^4|d_ie_j=e_1e_2=0\}\\ \zeta_1(d_1,d_2,e)=(d_1,d_2,e,0)\\ \zeta_2(d_1,d_2,e)=(d_1,d_2,d_1d_2,e)\end{array}$$

is an R-colimit.

Consider $\alpha,\beta:(D\times D)\bigvee D\longrightarrow M$ characterized by $\alpha(0.0.e)=v_2(e)$, $\alpha(d_1,d_2,0)=\gamma(d_1,d_2)$ and $\beta(0.0.e)=v_1(e)$, $\beta(d_1,d_2,0)=\alpha(d_1,d_2,d_1d_2)$, respectively. Thus, $\alpha\circ\psi=\beta\circ\phi$ and so, there exists a unique $\delta:(D\times D)\bigvee D(2)\longrightarrow M$ with $\delta\circ\zeta_1=\alpha$ and $\delta\circ\zeta_2=\beta$

Clearly

$$[v_1\dot{+}(v_2\dot{+}\gamma)](d_1,d_2)=\beta(d_1,d_2,d_1d_2)$$

We claim that

$$[(v_1+v_2)\dot{+}\gamma](d_1,d_2)=\beta(d_1,d_2,d_1d_2)$$

Let $\zeta(d_1,d_2,e)=\delta(d_1,d_2,e,e)$. Then $\zeta(d_1,d_2,0)=\delta(d_1,d_2,0,0)=\alpha(d_1,d_2,0)=\gamma(d_1,d_2)$ and $\zeta(0.0.e)=\delta(0,0,e,e)=(v_1+v_2)(e)$ since $\delta(0,0,e,0)=v_2(e)$ and $\delta(0,0,0,e)=v_2(e)$

Finally,

$$\begin{array}{rcl}[(v_1+v_2)\dot{+}\gamma)(d_1,d_2) & = & \zeta(d_1,d_2,d_1d_2)\\ & = & \delta(d_1,d_2,d_1d_2,d_1d_2)\\ & = & \beta(d_1,d_2,d_1d_2)\end{array}$$

This finishes the proof.

Proposition 0.5 *Let* $\gamma_1,\gamma_2\in M^i_\tau$ *and* $a\in R$. *Then*

$$\left\{\begin{array}{rcl}a(\gamma_2\dot{-}\gamma_1) & = & (a._1\gamma_2)\dot{-}(a._1\gamma_1)\\ & = & (a._2\gamma_2)\dot{-}(a._2\gamma_1)\end{array}\right.$$

Proof: We prove the first, the other is similar. By definition, $(\gamma_1\dot{-}\gamma_2)(d)=l(0,0,d)$ where $l:(D\times D)\bigvee D\longrightarrow M$ is characterized by

$$\left\{\begin{array}{rcl}l\circ\phi & = & \gamma_1\\ l\circ\psi & = & \gamma_2\end{array}\right.$$

Let $l' = a._1 l$. Then

$$\begin{array}{rcl}(a._1 l)(d_1, d_2) & = & (l' \circ \phi)(d_1, d_2) \\ & = & l'(\phi(d_1, d_2)) \\ & = & l'(d_1, d_2, 0) \\ & = & a._1 l(d_1, d_2, 0) \\ & = & l(ad_1, d_2, 0) \\ & = & (l \circ \phi)(ad_1, d_2, 0) \\ & = & \gamma_1(ad_1, d_2.0) \\ & = & a._1 \gamma(d_1, d_2)\end{array}$$

It follows that $(a._1\gamma_2 \dot{-} a._1\gamma_1)(d) = a._1 l(0,0,d) = a_1(\gamma_2 \dot{-} \gamma_1)(d)$, finishing the proof.

Proposition 0.6 *Let $\gamma_1, \gamma_2 \in M^i_\tau$. Then*

$$\sum(\gamma_2) \dot{-} \sum(\gamma_1) = \gamma_2 \dot{-} \gamma_1$$

Proof: Recall that $(\gamma_2 \dot{-} \gamma_1)(d) = l(0,0,d)$ where $l : (D \times D) \bigvee D \longrightarrow M$ is the unique map that satisfying $l \circ \phi = \gamma_1$ and $l \circ \psi = \gamma_2$. But then $\sum l \circ \phi = \sum \gamma_1$ and $\sum l \circ \psi = \sum \gamma_2$ and this shows that $(\sum \gamma_2 \dot{-} \sum \gamma_1)(d) = \sum l(0,0,d) = l(0,0,d) = (\gamma_2 \dot{-} \gamma_1)(d)$.

Proposition 0.7 *If $\gamma \in M^i_\tau$ and $v \in M_a$ where $\tau(0,0) = a$, Then*

$$\sum(v \dot{+} \gamma) = v \dot{+} \sum(\gamma)$$

Proof: Similar to the preceding ones.

Example: $M = R$

$\dot{+}$: Let $v \in R_a$, $\tau \in M_a^{D(2)}$ and $\gamma \in M^i_\tau$. By Kock-Lawvere,$v(d) = a + bd$, $\gamma(d_1, d_2) = a + b_1 d_1 + b_2 d_2 + c d_1 d_2$ and hence $\tau : D(2) \longrightarrow R$ is given by $\tau(d_1, d_2) = a + b_1 d_1 + b_2 d_2$. Then

$$(v \dot{+} \gamma)(d_1, d_2) = g(d_1, d_2, d_1 d_2)$$

where $g : (D \times D) \bigvee D \longrightarrow R$ is the unique map such that

$$\left\{\begin{array}{c} g \circ \phi(d_1, d_2) = g(d_1, d_2, 0) = \gamma(d_1, d_2) \\ g \circ \epsilon(d) = g(0,0,d) = a + bd \end{array}\right.$$

Thus, $g(d_1, d_2, d) = a + b_1 d_1 + b_2 d_2 + bd + c d_1 d_2$ and hence

$$(v \dot{+} \gamma)(d_1, d_2) = a + b_1 d_1 + b_2 d_2 + (b + c) d_1 d_2.$$

$\dot{-}$: Let $\gamma_1.\gamma_2 \in M^i_\tau$ with τ as above. Then

$$\begin{cases} \gamma_1(d_1,d_2) = a + b_1d_1 + b_2d_2 + c_1d_1d_2 \\ \gamma_2(d_1,d_2) = a + b_1d_1 + b_2d_2 + c_2d_1d_2 \end{cases}$$

Then $(\gamma_2 \dot{-} \gamma_1)(d) = l(0,0,d)$ where $l : (D \times D) \bigvee D \longrightarrow R$ is the unique map such that

$$\begin{cases} (l \circ \phi)(d_1,d_2) = l(d_1,d_2,d_1d_2) = \gamma_2(d_1,d_2) \\ (l \circ \psi)(d_1,d_2 = l(d_1,d_2,0)) = \gamma_1(d_1,d_2) \end{cases}$$

Thus $l(d_1,d_2,d) = a + b_1d_1 + b_2d_2 + (c_2 - c_1)d + c_1d_1d_2$ and hence

$$(\gamma_2 \dot{-} \gamma_1)(d) = l(0.0.d) = a + (c_2 - c_1)d$$

Connecting mapping: Let ∇ be a connection on M. We define the *connecting mapping* of ∇ to be the map $C : M^{D\times D} \longrightarrow M^D$ by the formula

$$C(\gamma) = \gamma \dot{-} \nabla K\gamma$$

Proposition 0.8 *The connecting mapping is R-linear for the two structures of fiber bundles of $M^{D\times D}$ on M^D.*

Proof: We do this only for the first structure, the other being similar. If $\gamma \in M^{D\times D}$, then $K(\gamma) = (\gamma_1, \gamma_2)$ where $\gamma_1 = [d \mapsto \gamma(d,0)]$ and $\gamma_2 = [d \mapsto \gamma(0,d)]$. Thus,

$$\begin{array}{rcl} C(a._1\gamma) & = & a._1\gamma \dot{-} \nabla K(a._1\gamma) \\ & = & a._1\gamma \dot{-} \nabla(a._1\gamma_1, a._1\gamma_2) \\ & = & a._1\gamma \dot{-} \nabla(a._1\gamma_1, \gamma_2) \\ & = & a._1\gamma \dot{-} a._1\nabla(\gamma_1, \gamma_2) \\ & = & a(\gamma \dot{-} \nabla K\gamma) \\ & = & aC(\gamma) \end{array}$$

Recall from "affineconnections.pdf" 0.2 that if $p : E \longrightarrow M$ is a vector bundle and ∇ a connection on M, the connection map $C : E^D \longrightarrow E$ associated to ∇ is defined in terms of the splitting of the map $H : E \times_M E \longrightarrow E^D$ induced by ∇. In more details: ∇ induces a split-exact sequence of vector bundles over E (i.e. with respect to $\oplus$):

$$0 \hookrightarrow E \times_M E \xrightarrow{H} E^D \underset{\nabla}{\overset{K}{\rightleftarrows}} M^D \times_M E \twoheadrightarrow 0$$

Hence H splits as well. Indeed, there is a map $C_1 : E^D \longrightarrow E \times_M E$ such that

$$(*) \begin{cases} C_1 \circ H = id_{E \times_M E} \\ H \circ C_1 = id_{E^D} - \nabla K \end{cases}$$

The connection map associated to ∇ was defined to be the map

$$C : E^D \longrightarrow E$$

given by $C = p_2 \circ C_1$.

Proposition 0.9 *The connecting map* $M^{D \times D} \longrightarrow M^D$ *and the connection map* $E^D \longrightarrow E$ *coincide for* $E = M^D$.

Proof: By definition of the connection map, $C_1(\gamma) = (\pi_{M^D}(\gamma), C(\gamma))$. Therefore, $HC_1(\gamma)(d) = \pi_{M^D}(\gamma) + dC(\gamma)$ (definition of H). Letting $K(\gamma) = (\gamma_1, \gamma_2)$ where $\gamma_1(d) = \gamma(d, 0)$ and $\gamma_2(d) = \gamma(0, d)$, we conclude from (*) that

$$(**) \quad (\gamma -_1 \nabla(\gamma_1, \gamma_2))(d_1, d_2) = \gamma_2(d_2) + d_1 C(\gamma)(d_2)$$

On the other hand, by definition of strong difference, $(\gamma \dot{-} \nabla(\gamma_1, \gamma_2))(e) = l(0, 0, e)$ where $l : (D \times D) \longrightarrow M$ is the unique map satisfying

$$\begin{cases} l(d_1, d_2, 0) = \nabla(t_1, t_2)(d_1, d_2) \\ l(d_1, d_2, d_1 d_2) = \gamma(d_1, d_2) \end{cases}$$

Let us fix d_1 and compute $\gamma_2(-) + d_1(\gamma \dot{-} \nabla(\gamma_1, \gamma_2))(-)$. Since

$$l(0, d_2, d_1 d_2') = \begin{cases} = t_d(d_2) & (d_2' = 0) \\ = d_1(\gamma \dot{-} \nabla(\gamma_1, \gamma_2))(d_2') & (d_2 = 0) \end{cases}$$

then $\gamma_2(d_2) + d_1(\gamma \dot{-} \nabla(\gamma_1, \gamma_2)(d_2) = l(0, d_2, d_1 d_2)$ by putting $d_2' = d_2$, according to the definition of addition. Proceeding in a similar way, considering this time $l(d_1, d_2, d_1' d_2)$

$$l(d_1, d_2, d_1 d_2) = (\nabla(\gamma_1, \gamma_2)(d_1, d_2) +_1 (\gamma_2 + d_1(\gamma \dot{-} \nabla(\gamma_1, \gamma_2)))(d_2).$$

But the left hand side is $\gamma(d_1, d_2)$. Therefore, $(\gamma_2 + d_1 C(\gamma))(d_2) = (\gamma -_1 \nabla(\gamma_1, \gamma_2))(d_1, d_2) = (\gamma_2 + d_1(\gamma \dot{-} \nabla(\gamma_1, \gamma_2))(d_2)$. Hence $C(\gamma) = \gamma \dot{-} \nabla(\gamma_1, \gamma_2)$.

Lie bracket: For X and Y vector fields over M, let $\phi : D \times D \longrightarrow M^M$ be defined by $\phi(d_1, d_2) = Y_{-d_2} \circ X_{-d_1} \circ Y_{d_2} \circ X_{d_1}$. Since $\phi(d.0) = \phi(0, d) = Id$, there is, by microlinearity, a unique function $[X, Y] : D \longrightarrow M$ such that $\phi(d_1, d_2) = [X, Y](d_1 d_2)$.

Proposition 0.10 *The vector fields on M equipped with the Lie bracket form a Lie algebra*

Proof: This means that $[-,-]$ satisfies the following equations

(a) $[Y, X] = -[X, Y]$ (antisymmetry)

(b) $[\lambda X, Y] = [X, \lambda Y] = \lambda[X, Y]$ (bilinearity)

(c) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (Jacobi identity)

This proposition is a particular case of the more general

Proposition 0.11 *If G is a microlinear space which is a monoid, then $G_e = \{t \in G^D | t(0) = e\}$ has a natural Lie algebra structure, i.e, a natural bilinear, antisymmetric operation*

$$[-,-] : G_e \times G_e \longrightarrow G_e$$

satisfying Jacobi identity.

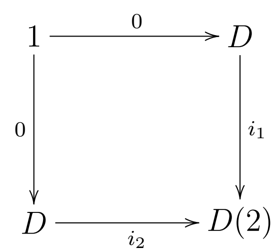
Proof: (Taken almost verbatim from Moerdijk/Reyes [4]). Let $t_1, t_2 \in G_e$ and define $h(t_1, t_2)(d_1, d_2) = t_2(-d_2) \circ t_1(-d_1) \circ t_2(d_2) \circ t_1(d_1)$. Since $h(t_1, t_2)(d_1, 0) = h(t_1, t_2)(0, d_2) = e$, there is a unique function $[-,-] : D \longrightarrow G$ such that $h(t_1, t_2)(d_1, d_2) = [t_1, t_2](d_1 d_2)$. This results from the following R-coequalizer

$$D \overset{i_1}{\underset{i_2}{\underset{(0,0)}{\rightrightarrows}}} D \times D \xrightarrow{\;\cdot\;} D$$

We first notice that $t_1(d) \circ t_2(d) = (t_1 + t_2)(d)$ and $t(-d) = (-t)(d) = t(d)^{-1}$ are immediate consequences of the definition of $+$ in G_e. In fact, define $l : D(2) \longrightarrow G$ by $l(d_1, d_2) = t_1(d_1) \circ t_2(d_2)$. Then $l(d_1, 0) = t_1(d_1)$ and $l(0, d_2) = t_2(d_2)$. Thus $(t_1 + t_2)(d) = l(d, d) = t_1(d) \circ t_2(d)$.

Antisymmetry follows from $h(t_1, t_2)(d_1, d_2) \circ h(t_2, t_1)(d_2, d_1) = e$ and $h(t_1, t_2)(d_1, d_2) = h(t_1, t_2)(d_2, d_1)$. Since homogeneity (which implies linearity) is clear, we are left with the Jacobi identity.

Since



is an R-pushout, it follows that whenever $t, t' \in G_e$ and $(d, d') \in D(2)$, then

$$(1) \qquad t(d) \circ t'(d') = t'(d') \circ t(d).$$

Writing $a = t_1(d_1), b = t_2(d_2), c = t_3(d_3)$, we may rewrite the Jacobi identity as

$$(2) \quad (a, (b, c))(b, (c, a))(c, (a, b)) = e$$

where (x, y) is the commutator $y^{-1}x^{-1}yx$ (the operation $\circ$ suppresed). Therefore, we have to prove

$$(3) \quad (b, c)^{-1}a^{-1}(b, c)a(c, a)^{-1}b^{-1}(c, a)b(a, b)^{-1}c^{-1}(a, b)c = e$$

We notice the following consequence of (1): we can interchange (x, y) with x or y or (u, v) provided that $\{x, y\} \cap \{u, v\} \neq \emptyset$.

Using this remark, we may move $(c, a)^{-1}$ to second place and $(a, b)^{-1}$ just right of a in the left hand side of (2) to get

$$(4) \quad (b, c)^{-1}(c, a)^{-1}a^{-1}(b, c)a(a, b)^{-1}b^{-1}(c, a)bc^{-1}(a, b)c$$

Rewriting bc^{-1} as $c^{-1}b(c^{-1}, b)$ and interchanging (c^{-1}, b) and (a, b) we get

$$(5) \quad (b, c)^{-1}(c, a)^{-1}a^{-1}(b, c)a(a, b)^{-1}b^{-1}(c, a)c^{-1}b(a, b)(c^{-1}, b)c$$

Spelling out the commutators, we obtain e.

Lemma 0.12 $[X,Y](d_1d_2) = X_{-d_1} \circ Y_{d_2} \circ X_{d_1} \circ Y_{-d_2}$

Proof: This is a simple consequence of $X_{-d} = (X_d)^{-1}$. In fact,

$$\begin{aligned}[X,Y](d_1d_2) &= ([X,Y]_{d_1(-d_2)})^{-1} = (Y_{d_2} \circ X_{-d_1} \circ Y_{-d_2} \circ X_{d_1})^{-1} \\ &= X_{-d_1} \circ Y_{d_2} \circ X_{d_1} \circ Y_{-d_2}\end{aligned}$$

As a consequence, we obtain the following characterization of the Lie bracket in terms of strong difference

Corollary 0.13 $[X,Y] = Y \star X \dot{-} \sum(X \star Y)$

Proof: By definition,

$$\begin{cases} (Y \star X)(d_1, d_2) = Y_{d_2} \circ X_{d_1} \\ \sum(X \star Y)(d_1, d_2) = X_{d_1} \circ Y_{d_2} \end{cases}$$

Thus, these functions coincide on $D(2)$ and their strong difference can be performed. Let $l : (D \times D) \bigvee D \longrightarrow M^M$ be given by

$$l(d_1, d_2, h) = X_{d_1} \circ [X,Y]_h \circ Y_{d_2}$$

Then, a trivial computation gives

$$\begin{cases} l(d_1, d_2, 0) = \sum(X \star Y)(d_1, d_2) \\ l(d_1, d_2, d_1d_2) = (Y \star X)(d_1, d_2) \end{cases}$$

Thus the strong difference is

$$(Y \star X \dot{-} \sum(X \star Y))(h) = l(0,0,h) = [X,Y]_h$$

Covariant derivative: If X, Y are vector fields on M we define the *covariant derivative* of Y along X to be the new vector field $\nabla_X Y = C(Y \star X)$. Since the connection and the connecting map coincide, a geometrical interpretation has already been provided in "affineconnections.pdf" 0.2 in the more general context of X a vector field on M and Y a vector field on a vector bundle $E \longrightarrow M$ having an affine connection ∇. We recall that in this case, $\nabla_X Y$ is the E-vector field $\nabla_X Y : M \longrightarrow E$ characterized by the equation

$$\nabla([d \mapsto \tilde{X}_m(h+d)], Y_{X_m(h)})(-h) - Y_m = h.(\nabla_X Y)_m$$

for every $h \in D$. Here $\tilde{X} : M \longrightarrow M^{D(2)}$ is the (unique) extension of X.

Grosso modo, this equation says that $(\nabla_X Y)_m$ is computed by transporting $Y_{X_m(h)}$ back along X and substracting Y_m from the result.

Riemann-Christoffel tensor: We define *the curvature of the connection* ∇ to be the map

$$R_{XY}Z = (C \circ C^D - C \circ C^D \circ \Sigma) \circ (Z \star Y \star X)$$

where

$$\begin{cases} (Z \star Y \star X)(d_1, d_2, d_3) = Z_{d_3} \circ Y_{d_2} \circ X_{d_1} \\ \Sigma(Z \star Y \star X)(d_1, d_2, d_3) = Z_{d_3} \circ Y_{d_1} \circ X_{d_2} \end{cases}$$

Notice that $Z \star Y \star X$ is a transpose of $(Z^D)^D \circ Y^D \circ X : M \longrightarrow ((M^D)^D)^D$.

Whether this algebraic notion is equivalent to the corresponding geometric one (as defined in [4]) seems to be the main open problem in this area. For finite dimensional manifolds, these notions are equivalent (cf. "Riemann-Christoffel.pdf" page 4.)

The following identity will be used later on

Proposition 0.14

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

Proof: The following is taken almost verbatim from the original (French) version of [3] page 175.

A simple computation gives

$$\begin{aligned} \nabla_X(\nabla_Y Z) &= C \circ (C \circ Z^D \circ Y)^D \circ X \\ &= C \circ C^D \circ (Z^D)^D \circ Y^D \circ X \\ &= C \circ C^D(Z \star Y \star X) \end{aligned}$$

since $Z \star Y \star X$ is a transpose of $(Z^D)^D \circ Y^D \circ X$.

In the following, it is convenient to use variables and the λ notation for functions (e.g., $f = \lambda x f(x)$). To fix the ideas, we shall use d_1 as a variable ranging in the outermost D of the expression $((M^D)^D)^D$, d_2 as a variable ranging in the next, etc. Thus, $C \circ C^D : ((M^D)^D)^D \longrightarrow M^D$ sends ϕ into

$C(\lambda_{d_1} C(\lambda_{d_2}(\lambda_{d_3}\phi(d_1,d_2,d_3))))$. To simplify the notation, we write $\lambda_{d_1}\lambda_{d_2}...$ instead of $\lambda_{d_1}(\lambda_{d_2}...)$. Thus,

$$\begin{cases} C \circ C^D(\phi) &= C(\lambda_{d_1} C(\lambda_{d_2}\lambda_{d_3}\phi(d_1,d_2,d_3))) \\ C \circ C^D(\Sigma\phi) &= C(\lambda_{d_1} C(\lambda_{d_2}\lambda_{d_3}\phi(d_2,d_1,d_3))) \end{cases}$$

Notice that since d_1 and d_2 are bound variables, we may interchange them in the last equation to obtain, equivalently,

$$C \circ C^D(\Sigma\phi) = C(\lambda_{d_2} C(\lambda_{d_1}\lambda_{d_3}\phi(d_1,d_2,d_3))).$$

NB: In [2], this last expression is written as $C[C\phi(\underline{d_1}, \underline{\underline{d_2}}, \underline{d_3})))]$, clearly showing the scope of the bound variables d_1, d_2, d_3. The English translation [3] obliterates the distinction between $\underline{d_2}$ and $\underline{\underline{d_2}}$, making this proof nonsensical. For instance (1) and (2) become literally identical, making $R = 0$!

We shall use both notations. Thus,

$$\begin{array}{rcll} (C \circ C^D)(Z \star Y \star X) &=& C(\lambda d_1\ C(\lambda d_2 \lambda d_3\ Z_{d_3} \circ Y_{d_2} \circ X_{d_1})) & (1) \\ (C \circ C^D \circ \Sigma)(Z \star Y \star X) &=& C(\lambda d_2\ C(\lambda d_1 \lambda d_3\ Z_{d_3} \circ Y_{d_2} \circ X_{d_1})) & (2) \\ \nabla_X(\nabla_Y Z) &=& C(\lambda d_1\ C(\lambda d_2 \lambda d_3\ Z_{d_3} \circ Y_{d_2} \circ X_{d_1})) & (3) \\ \nabla_Y(\nabla_X Z) &=& C(\lambda d_1\ C(\lambda d_2 \lambda d_3\ Z_{d_3} \circ X_{d_2} \circ Y_{d_1})) & (4) \\ \nabla_{[X,Y]} Z &=& C \circ Z^D \circ [X,Y] & (5) \end{array}$$

We must show that $(1) - (2) = (3) - (4) - (5)$. Since $(1) = (3)$, it is enough to show that $(2) - (4) = (5)$. Notice, furthermore, that by interchanging d_1 and d_3, we can re-write (4) as $C(\lambda d_2\ C(\lambda d_1 \lambda d_3\ Z_{d_3} \circ X_{d_1} \circ Y_{d_2}))$.

Let us compute

$$(2) - (4) = C[C(Z_{\underline{d_3}} \circ Y_{\underline{\underline{d_2}}} \circ X_{\underline{d_1}})] - C[C(Z_{\underline{d_3}} \circ X_{\underline{d_1}} \circ Y_{\underline{\underline{d_2}}})]$$

at m.

Define

$$\begin{cases} \delta_1(d_2, d) = C((Z_{\underline{d}_3} \circ Y_{d_2} \circ X_{\underline{d}_1})(m)(d) \\ \delta_2(d_2, d) = C((Z_{\underline{d}_3} \circ X_{\underline{d}_1} \circ Y_{d_2})(m)(d) \end{cases}$$

Then,

$$\begin{cases} \delta^1(d_2, 0) = C((Z_{\underline{d}_3} \circ Y_{d_2} \circ X_{\underline{d}_1})(m))(0) = Y_{d_2}(m) \\ \delta^2(d_2, 0) = C((Z_{\underline{d}_3} \circ X_{\underline{d}_1} \circ Y_{d_2})(m))(0) = Y_{d_2}(m) \end{cases}$$

These follow from the definition of C as the strong difference $C(\psi) = \psi \dot{-} \nabla K \psi$. In fact, recall that by definition of strong difference, $C(\psi)(d) = l(0,0,d)$ where $l : D \times D \vee D \longrightarrow M$ is the unique map such that $l(d_1, d_2, d_1 d_2) = \psi(d_1, d_2)$ and $\psi(d_1, d_2, 0) = \nabla K \psi(d_1, d_2)$. In particular, $C(\psi)(0) = l(0,0,0) = \psi(0,0) = x$, the base point.

To show the first equality, for example, put $\psi(d_1, d_3) = (Z_{d_3} \circ Y_{d_2} \circ X_{d_1})(m)$. with d_2 fixed.

Since C is R-linear for the structure $(+_1, \cdot_1)$ of the vector bundle $M^{D \times D}$ over M^D,

$$\begin{array}{rcl} (2) - (4) & = & C(\delta^1) - C(\delta^2) \\ & = & C(\delta^1 -_1 \delta^2) \end{array}$$

To compute $\delta = (\delta^1 -_1 \delta^2)$, define

$$\phi(d_2, d_2', d) = C((Z_{\underline{d}_3} \circ Y_{d_2} \circ X_{\underline{d}_1} \circ Y_{-d_2'})(m))(d)$$

whenever $(d_2, d_2') \in D(2)$.

We have

$$\phi(d_2, d_2', d) = \begin{cases} \delta^1(d_2, d) & (d_2' = 0) \\ \delta^2(-d_2', d) & (d_2 = 0) \end{cases}$$

Therefore,

$$\begin{array}{rcl} \delta(d_2, d) & = & (\delta^1 -_1 \delta^2)(d_2, d) \\ & = & \phi(d_2, d_2, d) \\ & = & C((Z_{\underline{d}_3} \circ Y_{d_2} \circ X_{\underline{d}_1} \circ Y_{-d_2})(m))(d) \end{array}$$

Notice that

$$\begin{array}{rcl} \delta(d_2, 0) & = & C((Z_{\underline{d}_3} \circ Y_{d_2} \circ X_{\underline{d}_1} \circ Y_{-d_2})(m))(0) \\ & = & (Y_{d_2} \circ Y_{-d_2})(m) \\ & = & m \end{array}$$

On the other hand,

$$\delta(0, d) = C((Z_{\underline{d}_3} \circ X_{\underline{d}_1}(m))(d)$$

Therefore

$$\delta(d_2, d) = C((Z_{\underline{d}_3} \circ X_{\underline{d}_1}(m))(d) + d_2.C(\delta))(d)$$

since the derivative of δ is $C(\delta) = (2) - (4)$.

In fact, letting as usual $K(\delta) = (\delta_1, \delta_2)$, we have that $\delta_1(d_2) = \delta(d_2, 0) = 0$ for every $d_2 \in D$. Therefore, $\nabla(\delta_1, \delta_2) = 0$, by bilinearity and this implies that

$$\begin{aligned} \delta(d_2, d) &= (\delta -_1 \nabla(\delta_1, \delta_2))(d_2, d) \\ &= \delta(0, d) + d_2 C(\delta)(d) \quad \text{as shown in proposition } 0.9 \ (**) \end{aligned}$$

Isolating the last term,

$$\begin{aligned} d_2 C(\delta) &= \delta(d_2, \underline{d}) - C((Z_{\underline{d}_3} \circ X_{\underline{d}_1})(m))(\underline{d}) \\ &= C((Z_{\underline{d}_3} \circ Y_{d_2} \circ X_{\underline{d}_1} \circ Y_{-d_2})(m)) - C((Z_{\underline{d}_3} \circ X_{\underline{d}_1})(m)) \\ &= C(\epsilon^1) - C(\epsilon^2) \end{aligned}$$

with

$$\begin{cases} \epsilon^1(d_1, d_3) = (Z_{d_3} \circ Y_{d_2} \circ X_{d_1} \circ Y_{-d_2})(m) \\ \epsilon^2(d_1, d_3) = (Z_{d_3} \circ X_{d_1})(m) \end{cases}$$

By using R-linearity of C for the $(+_1, \cdot_1)$ structure,

$$C(\epsilon^1) - C(\epsilon^2) = C(\epsilon^1 -_1 \epsilon^2)$$

To compute $\epsilon^1 -_1 \epsilon^2$, fix $d_2 \in D$ and let

$$\psi(d_1, d_1', d_3) = (Z_{d_3} \circ X_{-d_1'} \circ Y_{d_2} \circ X_{d_1} \circ Y_{-d_2})(m)$$

for $(d_1, d_1') \in D(2)$. Since

$$\psi(d_1, d_1', d_3) = \begin{cases} (Z_{d_3} \circ Y_{d_2} \circ X_{d_1} \circ Y_{-d_2})(m) = \epsilon^1(d_1, d_3) & (d_1' = 0) \\ (Z_{d_3} \circ X_{d_1'})(m) = \epsilon^2(-d_1', d_3) & (d_1 = 0) \end{cases}$$

then

$$\begin{aligned} (\epsilon^1 -_1 \epsilon^2)(d_1, d_3) &= \psi(d_1, d_1, d_3) \\ &= (Z_{d_3} \circ X_{-d_1} \circ Y_{d_2} \circ X_{d_1} \circ Y_{-d_2})(m) \\ &= (Z_{d_3} \circ [X, Y](d_1 d_2))(m) \end{aligned}$$

Therefore,

$$\begin{aligned} d_2 C(\delta) &= C(\epsilon^1 -_1 \epsilon^2) \\ &= C((Z_{\underline{d}_3} \circ [X, Y](\underline{d}_1 d_2))(m) \\ &= d_2 C((Z_{\underline{d}_3} \circ [X, Y](\underline{d}_1))(m) \end{aligned}$$

This shows that

$$
\begin{array}{rcl}
(2)-(4) & = & C(\delta) \\
 & = & C((Z_{\underline{d}_3} \circ [X,Y]_{\underline{d}_1})(m)) \\
 & = & (5)
\end{array}
$$

The last line follows from our identification $X_d(m) = X(m)(d)$ and conventions about the order of variables in the exponent of $(M^D)^D$:

$$
\begin{array}{rcl}
(Z^D \circ [X,Y])(m)(d_1)(d_3) & = & (Z \circ [X,Y])(d_1, m)(d_3) \\
 & = & Z([X,Y]_{d_1}(m)(d_3) \\
 & = & Z_{d_3}([X,Y]_{d_1}(m)) \\
 & = & (Z_{d_3} \circ [X,Y]_{d_1})(m)
\end{array}
$$

This concludes the proof.

Torsion of a connection: We define the *torsion of a connection* ∇ by $T(X,Y) = C(Y \star X) - C(\sum(Y \star X))$

Proposition 0.15 *If X and Y are vector fields on M, $T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$*

Proof: From the definition of the torsion of a connection,

$$
T(X,Y) = (Y \star X \dot{-} \nabla(X,Y)) - (\sum(Y \star X) \dot{-} \nabla(Y,X))
$$

On the other hand,

$$
\begin{array}{rcl}
\nabla_X Y - \nabla_Y X & = & C(Y \star X) - C(X \star Y) \\
 & = & (Y \star X \dot{-} \nabla(X,Y)) - (X \star Y \dot{-} \nabla(Y,X))
\end{array}
$$

Therefore,

$$
\begin{array}{rcl}
\nabla_X Y - \nabla_Y X - T(X,Y) & = & (\sum(Y \star X) \dot{-} \nabla(Y,X)) \\
 & & -(X \star Y \dot{-} \nabla(Y,X)) \\
 & = & \sum(Y \star X) \dot{-} X \star Y \\
 & = & [X,Y]
\end{array}
$$

The torsion T measures the lack of symmetry of the connection, as the following proposition makes clear

Proposition 0.16

$$T(X,Y) = \sum \nabla(Y,X) \dot{-} \nabla(X,Y)$$

Proof: The proof hinges on the following facts

1. $\sum(\gamma_1 \dot{-} \gamma_2) = \gamma_1 \dot{-} \gamma_2$
2. $\nabla K(Y \star X) = \nabla(X,Y)$

The second is immediate, recalling that $K(\phi) = (\phi_1, \phi_2)$ with $\phi_1(d) = \phi(d,0)$ and $\phi_2(d) = \phi(0,d)$. The first has been proved already. (Cf. Proposition 06)

Return to the proof of the proposition

$$\begin{array}{rcl} T(X,Y) & = & C(Y \star X) - C(\sum(Y \star X)) \\ & = & (Y \star X \dot{-} \nabla K(Y \star X)) - (\sum(Y \star X) \dot{-} \nabla K(\sum(Y \star X))) \\ & = & (Y \star X \dot{-} \nabla K(Y \star X)) - ((Y \star X) \dot{-} \sum \nabla K(\sum(Y \star X)) \\ & = & \sum \nabla K(\sum(Y \star X)) \dot{-} \nabla K(Y \star X) \\ & = & \sum \nabla(Y,X) \dot{-} \nabla(X,Y) \end{array}$$

Since $\sum \nabla(Y,X)(d_1,d_2) = \nabla(Y,X)(d_2,d_1)$, then ∇ is symmetric iff $T = 0$. (Cf. Proposition 0.1 (5)). In particular, we have

Corollary 0.17 *If the connection ∇ is symmetric, then*

$$\nabla_X Y = \nabla_Y X + [X,Y]$$

Corollary 0.18 *Let ∇ be connection. Then*

$$R_{YX}Z = -R_{XY}Z$$

Furthermore, if ∇ is symmetric, then

$$R_{XY}Z + R_{YZ}X + R_{ZX}Y = 0$$

Proof: The first is immediate from proposition 0.14. The second uses repeatedly $\nabla_X Y - \nabla_Y X = [X,Y]$ to reduce the equation to Jacobi's identity. In more details, we have (from the same proposition)

$$\left\{ \begin{array}{rcl} R_{XY}Z & = & \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \\ R_{YZ}X & = & \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X - \nabla_{[Y,Z]} X \\ R_{ZX}Y & = & \nabla_Z \nabla_X Y - \nabla_X \nabla_Z Y - \nabla_{[Z,X]} Y \end{array} \right.$$

Then the cyclic sum $\sum_0 R_{XY}Z$ may be computed by pairing e.g. the first term of the RHS of the first equation with the second term of the RHS of the last formula to obtain $\nabla_X\nabla_Y Z - \nabla_X\nabla_Z Y = \nabla_X[Y,Z]$. We then pair the result with the last term of the RHS of the second formula to obtain $\nabla_X[Y,Z] - \nabla_{[Y,Z]}X = [X,[Y,Z]]$ and then we add all of these results. Then, $\sum_0 R_{XY}Z = \sum_0[X,[Y,Z]] = 0$. (By the Jacobi's identity)

References

[1] Berger M., *Géométrie*, 5 volumes, CEDIC/Fernand Nathan, 2ème édition revue et corrigée, Paris 1979

[2] Lavendhomme R., *Leçons de Géométrie Différentielle Synthétique Naïve*, Institut de Mathématique, Ciaco, Louvain-la-Neuve, 1987

[3] Lavendhomme R., *Basic Concepts of Synthetic Differential Geometry*, Kluwer Academic Publishers, Dordrecht /Boston/London, 1996

[4] Moerdijk I., G.E.Reyes, *Models for Smooth Infinitesimal Analysis*, Springer-Verlag