General Relativity: Affine connections, parallel transport and sprays

Gonzalo E. Reyes
Université de Montréal

21 January 2009

In the following, we will use tacitly two results which in many cases reduce linearity of an application between R-modules to just homogeneity.

Proposition 0.1 *If M is a microlinear space, then all the tangent space of M satisfy the Kock-Lawvere axiom in the sense that the canonical map $\alpha : T_m(M) \times T_m(M) \longrightarrow T_m^D$ given by $\alpha(u,v) = u + dv$ is a bijection for each $m \in M$.*

Proof: (cf. [3] page 186) Let $\phi : D \longrightarrow T_x(M)$ be given and let $u = \phi(0)$. We claim the existence of a unique $v \in T_x(M)$ such that $\phi(d) = u + dv$. By replacing $\phi(d)$ by $\phi(d) - u$ in $T_x(M)$ we may assume that $\phi(0)$ is the null vector in $T_x(M)$. Define $\tau : D \times D \longrightarrow M$ by $\tau(d_1, d_2) = \phi(d_1)(d_2)$. Since M is microlinear and

$$D \times D \underset{p_2}{\overset{p_1}{\rightrightarrows}} D \times D \xrightarrow{\ \cdot\ } D$$

is an R-coequalizer, (here $p_1(d_1, d_2) = (d_1, 0), p_2(d_1.d_2) = (0, d_2)$) and $\tau \circ p_1 = \tau \circ p_2$, there is a unique $v \in T_x(M)$ such that $\tau(d_1, d_2) = v(d_1 \cdot v_2) = d_1 \cdot v(d_2)$. Thus, $\phi(d_1) = d_1 \cdot v_1$, completing the proof.

Proposition 0.2 *Let $H : V \longrightarrow W$ be an arbitrary (not necessarily linear) map of R-modules. Assume that W is a microlinear space which satisfies the Kock-Lawvere axiom in the sense that the canonical map $\alpha : W \times W \longrightarrow W^D$ defined by $\alpha(w_1, w_2)(d) = w_1 + dw_2$ is a bijection. Then H is linear iff H is homogeneous.*

Proof: Let $H : V \longrightarrow W$ be homogeneous, i.e., $H(\alpha v) = \alpha H(v)$ for every $\alpha \in R$ and $v \in V$. Take $v_1, v_2 \in V$ and consider the maps

$$D(2) \overset{\phi}{\underset{\psi}{\rightrightarrows}} W$$

defined by

$$\begin{cases} \phi(d_1, d_2) = H(d_1 v_1 + d_2 v_2) \\ \psi(d_1, d_2) = H(d_1 v_1) + H(d_2 v_2) \end{cases}$$

Since $\phi \circ i_k = \psi \circ i_k$ for $k = 1, 2$ and

$$\begin{array}{ccc} 1 & \xrightarrow{\;0\;} & D \\ {\scriptstyle 0}\downarrow & & \downarrow{\scriptstyle i_2} \\ D & \xrightarrow[i_1]{} & D(2) \end{array}$$

is an R-pushout, it follows that $\phi = \psi$. In particular $\phi(d, d) = \psi(d, d)$, i.e. $H(d(v_1 + v_2)) = H(dv_1) + H(dv_2)$. Using homogeneity and the fact that α is an isomorphism, $H(v_1) + H(v_2) = H(v_1 + v_2)$.

1 Connections and Sprays on a Microlinear Space

Let M be a microlinear space and let $K : M^{D \times D} \longrightarrow M^D \times_M M^D$ be the canonical map defined by $K(t)(d) = (t(d, 0), t(0, d))$

An *affine connection on* M is section of K

$$\nabla : M^D \times_M M^D \longrightarrow M^{D \times D}$$

satisfying some homogeneity conditions. More precisely, we require ∇ to have the following properties

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

A connection is *symmetric or torsion-free* if $\nabla(t_1, t_2)(d_1, d_2) = \nabla(t_2, t_1)(d_2, d_1)$.

There is a simple geometric interpretation of a connection in terms of parallel transport. A *parallel transport on* M is a function which associates with each $(t, h) \in M^D \times D$ a bijection

$$\tau_h(t, -) : \pi^{-1}(t(0)) \longrightarrow \pi^{-1}(t(h))$$

such that

$$\begin{cases} \tau_0(t_1, t_2) = t_2 \\ \tau_h(t_1, \lambda t_2) = \lambda \tau_h(t_1, t_2) \\ \tau_h(\lambda t_1, t_2) = \lambda \tau_h(t_1, t_2) \end{cases}$$

for all $\lambda \in R$.

We say that that $\tau_h(t_1, t_2)$ is the result of transporting t_2 parallel to itself along t_1 during h seconds.

Proposition 1.1 *Assume that* M *is a microlinear space. If* τ *is a parallel transport on* M*, then the map* $\nabla : M^D \times_M M^D \longrightarrow M^{D \times_D}$ *defined by*

$$\nabla(t_1, t_2)(h_1, h_2) = \tau_{h_1}(t_1, t_2)(h_2)$$

is an affine connection on M*. Conversely, any affine connection on* M *determines a parallel transport* τ *defined by the same formula.*

Proof: See [3] page 190.

In terms of parallel transport, the notion of a symmetric connection also has a clear geometrical interpretation that the reader is encouraged to find.

When $M = R^n$, a vector t at $\underline{a} \in R^n$ may be written uniquely as $t(d) = \underline{a} + d\underline{b}$, by using Kock-Lawvere axiom. The same axiom also implies that a connection can be written as

$$\nabla(t_1, t_2)(d_1, d_2) = \underline{a} + \underline{b}_1 d_1 + \underline{b_2} d_2 + \tilde{\nabla}_{\underline{a}}(b_1, b_2) d_1 d_2$$

with $t_i(d) = \underline{a} + \underline{b}_i d$ $(i = 1, 2)$

Thus, ∇ is completely determined by by

$$\tilde{\nabla}_{\underline{a}}(-, -) : R^n \times R^n \longrightarrow R^n$$

The defining properties of ∇ are translated into properties of $\tilde{\nabla}_{\underline{a}}(-, -)$ and express simply that this last map is bilinear. If ∇ is symmetric, this map is also symmetric. Notice, however, that in this case such a map is completely determined by its values at the diagonal. Indeed,

Proposition 1.2 *Let* $\sigma : R^n \times R^n \longrightarrow R^n$ *be a map that is bilinear and symmetric and let* $\delta : R^n \longrightarrow R^n$ *be defined by* $\delta(t) = \sigma(t,t)$*. Then*

$$\sigma(t_1, t_2) = (1/2)[\delta(t_1 + t_2) - \delta(t_1) - \delta(t_2)]$$

Proof: We have $\delta(t_1+t_2) = \sigma(t_1+t_2, t_1+t_2) = \sigma(t_1, t_1)+2\sigma(t_1, t_2)+\sigma(t_2, t_2)$, i.e., $\sigma(t_1, t_2) = (1/2)[\delta(t_1 + t_2) - \delta(t_1) - \delta(t_2)]$.

NB For $n = 1$, i.e., for $\sigma : R \times R \longrightarrow R$, homogeneity implies symmetry. In fact, by homogeneity

$$(1) \qquad \sigma(ax, y) = a\sigma(x, y) = \sigma(x, ay)$$

Taking partial derivatives with respect to a we get

$$(2) \qquad x(\partial\sigma/\partial x)(ax, y) = \sigma(x, y) = y(\partial\sigma/\partial y)(x, ay)$$

Taking partial derivatives with respect to x we get

$$(3) \qquad (\partial\sigma/\partial x)(ax, y) + ax(\partial^2\sigma/\partial^2 x)(ax, y) = (\partial\sigma/\partial x)(x, y).$$

Letting $a = 1$,

$$(4) \qquad x(\partial^2\sigma/\partial^2 x)(x, y) = 0$$

Similarly for y. From (1) and (2),

$$\begin{cases} \sigma(x, 0) = \sigma(0, y) = 0 \\ x(\partial\sigma/\partial x)(x, 0) = 0 \\ y(\partial\sigma/\partial y)(0, y) = 0 \end{cases}$$

Lavendhomme principle implies

$$(5) \qquad \begin{cases} (\partial^2\sigma/\partial x^2)(x, y) = 0 \\ (\partial\sigma/\partial x)(x, 0) = 0 \\ (\partial\sigma/\partial y)(0, y) = 0 \end{cases}$$

By Taylor developing around $(0, 0)$,

$$\sigma(x, y) = cxy$$

where $c = (\partial^2\sigma/\partial x\partial y)(0, 0)$.

Thus, σ is symmetric, $\delta(x) = cx^2$ is a quadratic form and $\sigma(x,y) = 1/2[\delta(x+y) - \delta(x) - \delta(y)]$.

For $n > 1$ homogeneity does not imply symmetry. In fact, already for $n = 2$ the map $f : R^2 \times R^2 \longrightarrow R^2$ defined by $f((x_1, x_2), (y_1, y_2)) = (x_1 y_2, x_2 y_1)$ is homogeneous, but not symmetric.

In the following we prove a generalization of the last proposition.

Let M be a microlinear space. A *spray* on M is a map $\sigma : M^D \longrightarrow M^{D_2}$ satisfying the conditions

$$\begin{array}{lll} (1) & \sigma(t)(d) = t(d) & \text{for all } d \in D \\ (2) & \sigma(\lambda t)(\delta) = \sigma(t)(\lambda\delta) & \text{for all } \delta \in D_2, \lambda \in R \end{array}$$

For $M = R^n$, we easily see that a spray has the following description

$$\sigma(t)(\delta) = \underline{a} + \underline{b}\delta + \tilde{\sigma_a}(\underline{b})\delta^2$$

with the homogeneity condition (2)

$$\tilde{\sigma_a}(\lambda\underline{b}) = \lambda^2 \tilde{\sigma_a}(\underline{b})$$

Proposition 1.3 *Given any affine connection (not necessarily symmetric) ∇ on M there is a unique spray σ on M such that*

$$\sigma(t)(d_1 + d_2) = \nabla(t,t)(d_1, d_2)$$

Proof: This follows from the fact that

$$D \underset{j_2}{\overset{j_1}{\rightrightarrows}} D \times D \xrightarrow{\;+\;} D_2$$

is an R-coequalizer, where $j_1(d) = (d, 0)$ and $j_2(d) = (0, d)$. Indeed, for $t \in M^D$, $\nabla(t,t)(d_1, d_2)$ satisfies $\nabla(t,t)(d,0) = t(d) = \nabla(t,t)(0,d)$. Using the previous R-coequalizer, we conclude that there is a unique map $\sigma(t) : D_2 \longrightarrow M$ satisfying

$$\sigma(t)(d_1 + d_2) = \nabla(t,t)(d_1, d_2)$$

It is easy to check that σ satisfies the conditions to be a spray.

The main result of this section is the following version of the Ambrose-Palais-Singer theorem:

Theorem 1.4 *Let M be a microlinear space. Then there is a natural one-to-one correspondence between symmetric affine connections ∇ on M and sprays on M given by*

$$\sigma(t)(d_1 + d_2) = \nabla(t,t)(d_1, d_2)$$

Proof: cf [3] page 192.

Before starting the proof, we observe that the following are R-coequalizers

$$(i) \qquad D \underset{j_2}{\overset{j_1}{\rightrightarrows}} D \times D \xrightarrow{+} D_2$$

$$(ii) \qquad D(2) \times D(2) \underset{(p_2,p_1)}{\overset{id}{\rightrightarrows}} D(2) \times D(2) \xrightarrow{+} D_2(2)$$

$$(iii) \qquad D \times D \times D \underset{(p_2, m \circ p_{13})}{\overset{(p_1, m \circ p_{23})}{\rightrightarrows}} D \times D \xrightarrow{m} D$$

$$(iv) \qquad D \times D \times D(2) \underset{(p_2, m \circ p_{13})}{\overset{(p_1, m \circ p_{23})}{\rightrightarrows}} D \times D(2) \xrightarrow{m} D(2)$$

$$(v) \qquad D_2 \times D_2 \times D_2(2) \underset{(p_2, m \circ p_{13})}{\overset{(p_1, m \circ p_{23})}{\rightrightarrows}} D_2 \times D_2(2) \xrightarrow{m} D_2(2)$$

where m stands for "multiplication."

From connections to sprays: already done using the R-coequalizer (i) (cf. last proposition)

From sprays to connection: we use the R-coequalizer (v). Let σ be a spray, $t_1, t_2 \in M^D$ such that $t_1(0) = t_2(0) = x$. The map $\tau : D_2 \times D_2(2) \longrightarrow M$ defined by $\tau(\delta, \lambda_1, \lambda_2) = \sigma(\lambda_1 t_1 + \lambda_2 t_2)(\delta)$ equalizes the parallel arrows in (v), by homogeneity of σ. Therefore, there is a unique map $\nabla'(t_1, t_2) : D_2(2) \longrightarrow M$ such that $\nabla'(t_1, t_2)(\delta\lambda_1, \delta\lambda_2) = \tau(\delta, \lambda_1, \lambda_2) = \sigma(\lambda_1 t_1 + \lambda_2 t_2)(\delta)$. Then $\nabla = \nabla'|_{D \times D} : D \times D \longrightarrow M$ is an affine symmetric connection, as can be easily checked. It remains to prove that these operations are inverse of each other: see [3].

2 Vector bundles and vector fields

Let M be a microlinear space. A *vector bundle over* M is a map $p : E \longrightarrow M$ with E microlinear such that every fiber $E_m = p^{-1}(m)$ is equipped with an R-module structure satisfying the Kock-Lawvere axiom, i.e., the canonical map

$$H_m : E_m \times E_m \longrightarrow (E_m)^D$$

defined by $H_m(u,v)(d) = u + dv$ is a bijection.

Examples:

(1) the canonical map $\pi_M : M^D \longrightarrow M$ sending a vector v into $v(0)$ is a vector bundle over M. (This was already proved. Cf. Proposition 0.1)

(2) the canonical map $\pi_E : E^D \longrightarrow E$. We denote the R-module structure by $\oplus$ and $\odot$ (same as (1))

(3) the map $p^D : E^D \longrightarrow M^D$. This is not so obvious. We check that its fibers satisfy the Kock-Lawvere axiom. Indeed, given $\phi : D \longrightarrow (E^D)_t = (p^D)^{-1}(t)$ with $t \in M^D$, consider $\phi_d : D \longrightarrow E_{t(d)} = p^{-1}(t(d))$ defined by $\phi_d(d') = \phi(d')(d)$. Since E is a vector bundle, there is a unique couple $(u_d, v_d) \in E^2_{t(d)}$ such that $\phi(d') = u_d + d'v_d$ $(\forall d' \in D)$. where $u = [d \mapsto u_d]$ and $v = [d \mapsto v_d]$ are in $(E^D)_t$. The R-module structure is defined pointwise and denoted by $+$ and $.$

(4) the first projection $p_1 : E \times_M E \longrightarrow E$. The R-module structure, (again denoted by $\oplus$ and $\odot$) is defined by

$$(u,v) \oplus (u,v') = (u, v + v') \qquad \alpha \odot (u,v) = (u, \alpha v)$$

(5) the canonical map $E \times_M E \longrightarrow M$. The R-module structure is defined by

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2) \qquad \alpha(u,v) = (\alpha u, \alpha v)$$

(If all E_m's satisfy the Kock-Lawvere axiom, so do all products $E_m \times E_m$.)

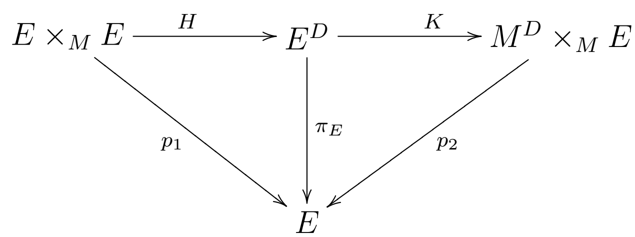
(6) the second projection $p_2 : M^D \times_M E \longrightarrow E$. The corresponding R-structure is given by

$$(t,v) \oplus (t',v) = (t \oplus t', v) \qquad \alpha \odot (t,v) = (\alpha t, v)$$

(7) the first projection $p_1 : M^D \times_M E \longrightarrow M^D$. The corresponding structure is defined by

$$(t, v) + (t, v') = (t, v + v') \qquad \alpha(t, v) = (t, \alpha v)$$

Whenever $p : E \longrightarrow M$ is a vector bundle, we have the following diagram



where $H(u, v)(d) = u + dv$ and $K(t) = (p \circ t, t(0))$.

Proposition 2.1 *1. H is linear as a map of vector bundles (4) and (2)*

2. K is linear as a map of vector bundles (2) and (6)

3. The sequence of linear maps over E

$$0 \longrightarrow E \times_M E \xrightarrow{\ H\ } E^D \xrightarrow{\ K\ } M^D \times_M E$$

is exact, i.e., H is 1-1 and $Im(H) = Ker(K)$.

Proof: (1) and (2). Linearity follows from homogeneity, but this is obvious in each case. (3) H is monic is a consequence of the fact that the fibers of $p : E \longrightarrow M$ satisfy the Kock-Lawvere axiom. It remains to show that $Im(H) = Ker(K)$. Let $m \in M$ and $E_m = p^{-1}(m)$ the fiber of p over m. Then

$$\begin{aligned} t \in E_m^D &\leftrightarrow \forall d \in D\ t(d) \in E_m \\ &\leftrightarrow \forall d \in D\ p(t(d)) = m \\ &\leftrightarrow p \circ t = 0_m \\ &\leftrightarrow t \in Ker(K) \end{aligned}$$

But, by definition of vector bundle, $t \in E_m^D$ iff there is a unique couple $(u, v) \in E_m^2$ such that $\forall d \in D\ t(d) = u + dv$, i.e., if $t = H(u, v)$. This finishes the proof.

A vector $t \in E^D$ such that $K(t) = 0$ is called a *vertical vector.*

Our aim is to generalize the definition of affine connections to vector bundles, and introduce the connection maps.

We recall some notions to be used in the sequel (cf. [3]and [2]).

Let M be a micro-linear space and let $p : E \longrightarrow M$ be a vector bundle. An *affine connection* on $p : E \longrightarrow M$ (briefly, on E) is a section of K

$$\nabla : M^D \times_M E \longrightarrow E^D$$

which is linear with respect to both vector bundles structures $\oplus$ over E and $+$ over M^D. Recall that

$$\begin{cases} (u,v) \oplus (u,v') = (u, v+v'), & \alpha \odot (u,v) = (u, \alpha v) \\ (u_1,v_1) + (u_2,v_2) = (u_1+u_2, v_1+v_2), & \alpha \cdot (u,v) = (\alpha u, \alpha v) \end{cases}$$

In other words, ∇ should satisfy the following identities

$$\begin{gathered} p \circ \nabla(t,v)(d) = t(d), \nabla(t,v)(0) = v \\ \nabla(\alpha t, v)(d) = (\alpha \odot \nabla(t,v))(d) = \nabla(t,v)(\alpha d) \\ \nabla(t, \alpha v)(d) = (\alpha \cdot \nabla(t,v))(d) = \alpha \cdot (\nabla(t,v)(d)) \end{gathered}$$

Let $\nabla : M^D \times_M E \longrightarrow E^D$ be an affine connection on a vector bundle $p : E \longrightarrow M$. The connection ∇ 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

$$\begin{cases} C_1 \circ H = id_{E \times_M E} \\ H \circ C_1 = id_{E^D} \ominus \nabla K \end{cases}$$

Indeed, $K(t \ominus \nabla K(t)) = K(t) - K(t) = 0$. Thus, $t \ominus \nabla K(t) = H(u,v)$ for a unique couple $(u,v) \in E^2$. Define $C_1(t) = (u,v)$. Then $C_1 \circ H = id_{E \times_M E}$ and $H \circ C_1 = id_{E^d} \ominus \nabla K$.

We define *the connection map associated to* ∇ to be

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

given by $C = p_2 \circ C_1$.

The connection map $C : E^D \longrightarrow E$ is bilinear in the sense that for every $(t, t') \in E^D \times_{M^D} E^D$, we have

$$\begin{cases} C(t + t') = C(t) + C(t') \\ C(t \oplus t') = C(t) + C(t') \end{cases}$$

(The two +'s on the right denote addition in $E_{pt(0)} = E_{pt'(0)}$).

Intuitively, C measures the extent to which a tangent vector $t \in E^D$ coincides with the parallel transport of $t(0)$ along $p \circ t$ given by ∇. (Cf. [3] V3.9-11 and V4.7)

We shall apply these notions to the vector bundle $E = M^D$ (and $p = \pi_0$). So, we assume that $M = E^D$. If X, Y are vector fields, we define the new vector field

$$\nabla_X Y = C \circ (X \star Y)$$

with $(X \star Y)(d_1, d_2) = Y_{d_2} \circ X_{d_1}$. We notice that $X \star Y$ is a transpose of the map $Y^D \circ X : M \longrightarrow (M^D)^D$. Finally, 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 $(Z^D)^D \circ Y^D \circ X : M \longrightarrow ((M^D)^D)^D$.

There is a geometrical interpretation of $\nabla_X Y$:

Proposition 2.2 *Let $E \longrightarrow M$ be a vector bundle, ∇ an affine connection on E, X a vector field on M and Y an E-vector field on M. Then $\nabla_X Y$ is uniquely determined by the following identity*

$$\forall h \in D \; h(\nabla_X Y)_m = \nabla(d \mapsto \tilde{X}_m(h + d), Y^D \circ X(h))(-h) - Y_m$$

where $\tilde{X} : M \longrightarrow M^{D_2}$ is the (unique) extension of X.

Proof: (Almost verbatim from [3]). We first notice that H is a bijection (since E_m satisfies the Kock-Lawvere axiom) and this implies that

$$\begin{aligned} H \circ C_1(t)(h) &= p_1 \circ H^{-1} H \circ C_1(t) + h p_2 H^{-1} H C_1(t) \\ &= H \circ C_1(t)(0) + hC(t) \end{aligned}$$

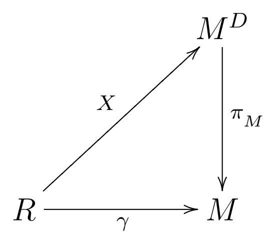
In particular, for $t = Y^D \circ X - \nabla(X, Y)$ we have

$$\forall h \in D \; [(Y^D \circ X)_m - \nabla(X_m, Y_m)](h) = Y_m + h(\nabla_X Y)_m]$$

The rest of the proof is straightforward and may be found in [3].

2.1 Vector fields over curves

Let M be a microlinear space and $\gamma : R \longrightarrow M$. A *vector space over* γ is a map $X : R \longrightarrow M^D$ such that the diagram



is commutative.

Example: the "velocity field" of γ, $\gamma^\bullet : R \longrightarrow M^D$ defined by $\gamma^\bullet(t)(d) = \gamma(t + d)$.

If ∇ is an affine connection on M and X a vector field over γ, we define $\nabla_{\gamma^\bullet} X$, a new vector field over γ, by the equation, for all $h \in D$,

$$h \nabla_{\gamma^\bullet} X(t) = \nabla(\gamma^\bullet(t), X(\gamma(t + h))(-h) - X(\gamma(t))$$

(Notice that existence and uniqueness of $\nabla_{\gamma^\bullet} X(t)$ follow from the validity of Kock-Lawvere axiom for the fiber $M_{\gamma(t)}$.) For the properties of this operation, see [5].

2.2 Computations for $M = R^n$

Proposition 2.3 (e.g. Fock page 128) *If* $X = \sum_i a^i \partial/\partial x^i$ *and* $Y = \sum_i b^i \partial/\partial x^i$ *are two vector fields on* $M = R^n$*, then*

$$\nabla_X Y = \sum_{ij} a^i (\partial b^j/\partial x^i) \partial/\partial x^j + \sum_{ijk} \Gamma^k_{ij} a^i b^j \partial/\partial x^k$$

In particular the k^{th}*-component of* $L_X Y$ *is*

$$(\nabla_X Y)^k = \sum_i a^i (\partial b^k/\partial x^i) + \sum_{ij} \Gamma^k_{ij} a^i b^j$$

Proof: We first prove the particular case $X = \partial/\partial x^i$:

$$\begin{aligned} \nabla_{\partial/\partial x^i} Y &= \nabla_{\partial/\partial x^i} \textstyle\sum_j b^j \partial/\partial x^j \\ &= \textstyle\sum_j \nabla_{\partial/\partial x^i} b^j \partial/\partial x^j \\ &= \textstyle\sum_j ((\partial b^j/\partial x^i) \partial/\partial x^j + b^j \nabla_{\partial/\partial x^i} \partial/\partial x^j) \\ &= \textstyle\sum_j ((\partial b^j/\partial x^i) \partial/\partial x^j + b^j \sum_k \Gamma^k_{ij} \partial/\partial x^k) \end{aligned}$$

The general case follows from the properties of $\nabla_X Y$ (cf. Lavendhomme [2] page 147). In fact,

$$\begin{aligned} \nabla_X Y &= \nabla_{\sum_i a^i \partial/\partial x^i} Y \\ &= \textstyle\sum_i \nabla_{a^i \partial/\partial x^i} Y \\ &= \textstyle\sum_i a^i \nabla_{\partial/\partial x^i} Y \\ &= \textstyle\sum_i a^i [\sum_j ((\partial b^j/\partial x^i) \partial/\partial x^j + b^j \sum_k \Gamma^k_{ij} \partial/\partial x^k)] \\ &= \textstyle\sum_{ij} a^i (\partial b^j/\partial x^i) \partial/\partial x^j + \sum_{ijk} a^i b^j \Gamma^k_{ij} \partial/\partial x^k \end{aligned}$$

3 Geodesics

Proposition 3.1 *Let* $\gamma : R \longrightarrow M$ *and* ∇ *a connection on* M*. Then the following are equivalent:*

1. $\nabla(\gamma^\bullet, \gamma^\bullet) = \gamma^{\bullet\bullet}$
2. $\nabla_{\gamma^\bullet} \gamma^\bullet = 0$

Proof: Unwinding definitions and writing everything explicitly, these expressions are shorthands for

1. $\gamma^\bullet(t+h_1) = \tau_{h_1}(\gamma^\bullet(t), \gamma^\bullet(t))$
2. $\tau_{-h}(\gamma^\bullet(t+h), \gamma^\bullet(t+h)) = \gamma^\bullet(t)$

We obtain the second from the first by substituting

$$\begin{cases} h_1 \mapsto -h \\ t \mapsto t+h \end{cases}$$

To obtain the other direction, we make the substitution

$$\begin{cases} h \mapsto -h_1 \\ t \mapsto t+h_1 \end{cases}$$

We say that $\gamma : R \longrightarrow M$ is a *geodesic* iff γ satisfies any of the conditions of the proposition.

Thus, geodesic curves have a clear geometrical content: "by parallel transporting an infinitesimal portion of the curve (seen as a tangent vector) along itself, the result is again an infinitesimal portion of the curve, i. e, a new tangent vector to the curve." No metrics are needed to define this notion.

Let M is an n-dimensional manifold. By introducing x-coordinates, we may assume that $M = R^n$, and use Tensor Calculus ("les débauches d'indices" according to Élie Cartan). In this case,

$$\nabla : M \times R^n \longrightarrow M \times R^n \times R^n$$

is completely determined by by its last component ∇_4, since K is a projection and we define the *Chritoffel symbols of the second kind* by

$$(\Gamma^l_{ji})_x(p) = -\nabla_4(p, e_i, e_j)(e_l)$$

We shall usually omit 'x' from the notation, if the coordinate frame is clear from the context.

Lemma 3.2 *If ∇ is a connection, ∇^4 is bilinear.*

Proof: Immediate from the properties $\nabla(\lambda b_1, b_2)(d_1, d_2) = \nabla(b_1, b_2)(\lambda d_1, d_2)$ and $\nabla(b_1, \lambda b_2)(d_1, d_2) = \nabla(b_1, b_2)(d_1, \lambda d_2)$.

Proposition 3.3 *A geodesics $\gamma = (x^1, \ldots, x^n)$ satisfies the second order differential equation*

$$d^2x^k/dt^2 + \sum_{i,j=1}^{n} \Gamma^k_{\alpha\beta}(dx^\alpha/dt)(dx^\beta/dt) = 0$$

Proof: Recall that the vector field along γ, $\gamma^\bullet : R \longrightarrow M^D$ is defined by $\gamma^\bullet(t) = [d \mapsto \gamma(t+d)]$, i.e., $\gamma^\bullet(t)(d) = \gamma(t+d)$. It follows that $\gamma^{\bullet\bullet} : R \longrightarrow (M^D)^D$ is given by $\gamma^{\bullet\bullet}(t)(d_1, d_2) = \gamma(t + d_1 + d_2)$. Hence

$$\begin{aligned} \gamma^\bullet(t)(d) &= (x^1(t+d), \ldots, x^n(t+d)) \\ &= (x^1(t) + d(dx^1/dt)(t)), \ldots, x^n(t) + d(dx^n/dt(t))) \\ &= \underline{x}(t) + d(d\underline{x}/dt)(t) \end{aligned}$$

with $\underline{x} = (x^1(t), \ldots, x^n(t))$ and $d\underline{x}/d(t) = (dx^1/dt(t), \ldots, dx^n/dt(t))$. Leaving the "$t$" out,

$$\begin{aligned} \gamma^{\bullet\bullet}(d_1, d_2) &= (x^1 + d_1(dx^1/dt) + d_2(dx^1/dt)) + d_1d_2(d^2x^1/dt^2), \ldots, \\ &\quad x^n + d_1(dx^n/dt) + d_2(dx^n/dt) + d_1d_2(d^2x^n/dt^2)) \\ &= \underline{x} + d_1(d\underline{x}/dt) + d_2(d\underline{x}/dt) + d_1d_2(d^2\underline{x}/dt^2) \end{aligned}$$

Similarly,

$$\nabla(\gamma^\bullet, \gamma^\bullet)(d_1, d_2) = \underline{x} + d_1(d\underline{x}/dt) + d_2(d\underline{x}/dt) + d_1d_2\nabla^4_{\underline{x}}(d\underline{x}/dt, d\underline{x}/dt)$$

The equation of the geodesic becomes then

$$(*) \quad d^2\underline{x}/dt^2 = \nabla^4_{\underline{x}}(d\underline{x}/dt, d\underline{x}/dt)$$

Writing $d\underline{x}/dt = \sum_\alpha (dx^\alpha/dt)\partial/\partial x^\alpha$ and using the bi-linearity of ∇^4 we have

$$\nabla^4(d\underline{x}/dt, d\underline{x}/dt) = \sum_{\alpha\beta}(dx^\alpha/dt)(dx^\beta/dt)\nabla^4(\partial/\partial x^\alpha, \partial/\partial x^\beta)$$

Taking the k^{th} component in the equation (*), we finally obtain

$$d^2x^k/dt^2 + \sum_{\alpha\beta} \Gamma^k_{\alpha\beta}(dx^\alpha/dt)(dx^\beta/dt) = 0$$

This is the expression found in texts of Differential Geometry (See e,g, [6]).

Proposition 3.4 *Given an affine connection ∇ on M, there is a symmetric affine connection $\tilde{\nabla}$ having the same geodesics.*

Proof: Let ∇ be an affine connection and let $\sigma : M^D \longrightarrow M^{D_2}$ be the unique spray associated to ∇, i.e., defined by the formula $\sigma(t)(d_1+d_2)) = \nabla(t,t)(d_1,d_2)$. Define $\tilde{\nabla}$ to be the symmetric affine connection associated with σ (given by the Ambrose-Palais-Singer Theorem). We claim that ∇ and $\tilde{\nabla}$ have the same geodesics. In fact, if $\gamma : R \longrightarrow M$ is a curve,

$$\begin{array}{rcl} \tilde{\nabla}(\gamma^\bullet(t),\gamma^\bullet(t))(d_1,d_2) & = & \sigma(\gamma^\bullet(t)(d_1+d_2) \\ & = & \nabla(\gamma^\bullet(t),\gamma^\bullet(t))(d_1,d_2) \end{array}$$

(the first by definition of $\tilde{\nabla}$ (cf. Ambrose-Palais-Singer theorem already mentioned); the second by definition of σ). But γ is a geodesic for a connection ∇ iff $\nabla(\gamma^\bullet(t),\gamma^\bullet(t)) = \gamma^{\bullet\bullet}(t)$. Thus, ∇ and $\tilde{\nabla}$ have the same geodesics.

This shows that, so far as geodesics of a connection are concerned, we may assume, without loss of generality, that the connection is symmetric.

4 Transformation law for vectors and vector fields

In this section we study how vectors transform from x-coordinates to x'-coordinates.

Let $\phi = (\phi^1,\ldots,\phi^n) : R^n \longrightarrow R^n$ be a diffeomorphism and $v = (v^1,\ldots,v^n) : D \longrightarrow R^n$ a tangent vector. We want to find the the law of transformation of the components of v under ϕ. Assume that $v^i(d) = a^i + db^i$ $(i = 1,\ldots,n)$. We have

$$D \xrightarrow{\;v\;} R^n \xrightarrow{\;\phi\;} R^n$$

$$(\phi\circ v)(d) = (\phi^1(a) + d\sum_j b_j(\partial\phi^1/\partial x^j)(a),\ldots,\phi^n(a) + d\sum_j b_j(\partial\phi^n/\partial x^j)(a))$$

This means that the i^{th}-component of $\phi\circ v$ is

$$(\phi\circ v)^i = \sum_j b^j(\partial\phi^i/\partial x^j).$$

Define $v'^i = \pi_i(\phi \circ v)$. By identifying v^i with b^i and letting $x'^i(x) = \phi^i(x)$, as is common in texts of GR, our previous expression can be written

$$v'^i = \sum_j v^j(\partial x'^i/\partial x^j)$$

This is the transformation rule "from the x variables to the x' variables". Using a terminology used extensively, vectors transform "contravariantly." Clearly this argument gives the law of transformation for vector fields by just interpreting v'^i and v^j as components of a vector field.

There is a corresponding transformation law for co-vectors which can be deduced from this one, namely

$$\partial/\partial x^\mu = \sum_\alpha (\partial \xi^\alpha/\partial x^\mu)\partial/\partial \xi^\alpha$$

4.1 Transformation law for Γ^k_{ij}

In this section, we present the proof of Synge/Schild [7] of the transformation rule for Γ^k_{ij} when we pass from x-coordinates to x'-coordinates:

$$\begin{array}{rcl}(\Gamma^\gamma_{\alpha\beta})_{x'} & = & \sum_{ijk}(\Gamma^k_{ij}) + x(\partial x'^\gamma/\partial x^k)(\partial x^i/\partial x'^\alpha)(\partial x^j/\partial x'^\beta) \\ & + & \sum_k(\partial^2 x^k/\partial x'^\alpha \partial x'^\beta)(\partial x'^\gamma/\partial x^k)\end{array}$$

(cf. e.g. [6] II page 232 or [7] page 48).

Start with a curve $\gamma : R \longrightarrow M$. Then we know (previous section) that $V = \gamma^{\bullet\bullet} + \nabla(\gamma^\bullet, \gamma^\bullet)$ is a contravariant vector field. In x coordinates, the γ-th component of V is

$$v^\gamma = d^2x^\gamma/dt^2 + \sum_{\alpha\beta} \Gamma^\gamma_{\alpha\beta}(dx^\alpha/du)(dx^\beta/du)$$

Thus, by the transformation law of contravariant vector fields,

$$(2.502) \quad v'^\gamma = \sum_k v^k(\partial x'^\gamma/\partial x^k)$$

Lemma 4.1 (2.504)

$$(d^2x^k/dt^2)(\partial x'^\gamma/\partial x^k) = (d^2x'^\gamma/dt^2) + \sum_{\alpha\beta}(\partial x'^\gamma/\partial x^k)(\partial^2 x^k/\partial x'^\alpha \partial x'^\beta)(dx'^\alpha/dt)(dx'^\beta/dt)$$

Proof: By Leibniz's rule,

$$\begin{array}{rcl} d^2x^k/dt^2 & = & d/dt[(\sum_\alpha \partial x^k/\partial x'^\alpha)(dx'^\alpha/dt)] \\ & = & \sum_\alpha(\partial x^k/\partial x'^m)(d^2x'^m/dt^2) + \sum_{\alpha\beta}(\partial^2 x^k/\partial x'^\alpha \partial x'^\beta)(dx'^\alpha/dt)(dx'^\beta)dt \end{array}$$

Multiplying both sides by $\partial x'^\gamma/\partial x^k$ we obtain the result sought.

Corollary 4.2 (Transformation law for Γ)

$$\begin{array}{rcl} (\Gamma^\gamma_{\alpha\beta})_{x'} & = & \sum_{ijk}(\Gamma^k_{ij})_x(\partial x'^\gamma/\partial x^k)(\partial x^i/\partial x'^\alpha)(\partial x^j/\partial x'^\beta) \\ & + & \sum_k(\partial^2 x^k/\partial x'^\alpha \partial x'^\beta)(\partial x'^\gamma/\partial x^k) \end{array}$$

Proof: From (2.502)

$$\begin{array}{l} d^2x'^\gamma/dt^2 + (\Gamma^\gamma_{\alpha\beta})_{x'}(dx'^\alpha/dt)(dx'^\beta/dt) = \\ \sum_k(d^2x^k/dt^2 + (\Gamma^k_{\alpha\beta})_x(dx^\alpha/dt)(dx^\beta/dt))(\partial x'^\gamma/\partial x^k)) \end{array}$$

Using 2.504, changing bound variables α, β by i, j and noticing that

$$\left\{ \begin{array}{rcl} dx^i/dt & = & \sum_\alpha(\partial x^i/\partial x'^\alpha)(dx'^\alpha/dt) \\ dx^j/dt & = & \sum_\alpha(\partial x^j/\partial x'^\alpha)(dx'^\alpha/dt) \end{array} \right.$$

we obtain

$$A^\gamma_{\alpha\beta}(dx'^\alpha/dt)(dx'^\beta/dt) = 0$$

with

$$\begin{array}{rcl} A^\gamma_{\alpha\beta} & = & (\Gamma'^\gamma_{\alpha\beta})_x - \sum_{ijk}(\Gamma^k_{ij})_x(\partial x'^\gamma/\partial x^k)(\partial x^i/\partial x'^\alpha)(\partial x^j/\partial x'^\beta) \\ & - & \sum_k(\partial x'^\gamma/\partial x^k)(\partial^2 x^k/\partial x'^\alpha \partial x'^\beta) \end{array}$$

Now, $A^\gamma_{\alpha\beta}$ is a function on M, independent of both dx'^α/dt and dx'^β/dt. Therefore, $A^\gamma_{\alpha\beta} = 0$. In more details, let $p \in M$. Let γ be a geodesic starting at p with initial velocity $(1, \ldots, 1)$ in the x'-coordinates. Then $A^\gamma_{\alpha\beta}(p) = 0$. We have obtained the transformation law.

4.2 Transformation law for Γ (bis)

Recall from [5] (page 4) that

$$(*)\ (\nabla_{\partial/\partial x^i}\partial/\partial x^j)_x = \sum_{k=1}^n \Gamma^k_{ij}(x)\partial/\partial x^k|_x$$

By the properties of ∇, if $X = \sum_i a^i \partial/\partial x^j$ and $Y = \sum_j b^j \partial/\partial x^i$, then ([5] page 5)

$$(**)\ \nabla_X Y = \sum_{ij} a^i(\partial b^j/\partial x^i)\partial/\partial x^j + \sum_{ijk} \Gamma^k_{ij} a^i b^j \partial/\partial x^k$$

In x' coordinates, $(*)$ becomes

$$(*')\ (\nabla_{\partial/\partial x'^\alpha} \partial/\partial x'^\beta)_{x'} = \sum_\gamma \Gamma^\gamma_{\alpha\beta}(x')\partial/\partial x'^\gamma|_{x'}$$

and the $\Gamma' s$ are determined by this equation, since $\{\partial/\partial x'^\gamma|_{x'}\}_\gamma$ is a basis.

For the particular case

$$\begin{cases} X = \partial/\partial x'^\alpha = \sum_i (\partial x^i/\partial x'^\alpha)\partial/\partial x^i \\ Y = \partial/\partial x'^\beta = \sum_j (\partial x^j/\partial x'^\beta)\partial/\partial x^j \end{cases}$$

$a^i = \partial x^i/\partial x'^\alpha$ and $b^j = \partial x^j/\partial x'^\beta$ and

$$(***)\ \nabla_{\partial/\partial x'^\alpha} \partial/\partial x'^\beta = \sum_{ij} (\partial x^i/\partial x'^\alpha)\partial/\partial x^j + \sum_{ijk} \Gamma^k_{ij}(\partial x^i/\partial x'^\alpha)(\partial x^j/\partial x'^\beta)\partial/\partial x^k$$

From

$$\partial/\partial x^k = \sum_\sigma (\partial x'^\sigma/\partial x^k)\partial/\partial x'^\sigma$$

we conclude that the γ^{th} term of the second summand of $(***)$ is

$$\sum_{ijk} \Gamma^k_{ij}(\partial x^i/\partial x'^\alpha)(\partial x^j/\partial x'^\beta)(\partial x'^\gamma/\partial x^k)$$

Similarly, from

$$\partial/\partial x^j = \sum_\sigma (\partial x^\sigma/\partial x^j)\partial/\partial x'^\sigma$$

we conclude that the γ^{th} term of the first summand is

$$\begin{aligned} &\textstyle\sum_{ij} (\partial x^i/\partial x'^\alpha)\partial/\partial x^i(\partial x^j/\partial x'^\beta)(\partial x'^\gamma/\partial x^j) \\ &= \textstyle\sum_{ij} \partial x^i/\partial x'^\alpha \sum_\sigma (\partial x'^\sigma/\partial x^i)\partial/\partial x'^\sigma(\partial x^j/\partial x'^\beta)(\partial x'^\gamma/\partial x^j) \\ &= \textstyle\sum_{ij\sigma} (\partial x^i/\partial x'^\alpha)(\partial x'^\sigma/\partial x^i)(\partial^2 x^j/\partial x'^\sigma \partial x'^\beta)(\partial x'^\gamma/\partial x^j) \\ &= \textstyle\sum_{ij\sigma} \delta^\sigma_\alpha(\partial^2 x^j/\partial x'^\sigma \partial x'^\beta)(\partial x'^\gamma/\partial x^j) \\ &= \textstyle\sum_{ij} (\partial^2 x^j/\partial x'^\alpha \partial x'^\beta)(\partial x'^\gamma/\partial x^j) \\ &= \textstyle\sum_i (\partial^2 x^i/\partial x'^\alpha \partial x'^\beta)(\partial x'^\gamma/\partial x^i) \end{aligned}$$

This gives the law of transformation

$$(\Gamma^\gamma_{\alpha\beta})_{x'} = \sum_{ijk} (\Gamma^k_{ij})_x(\partial x^i/\partial x'^\alpha)(\partial x^j/\partial x'^\beta)(\partial x'^\gamma/\partial x^k) + \sum_i (\partial^2 x^i/\partial x'^\alpha \partial x'^\beta)(\partial x'^\gamma/\partial x^i)$$

4.3 A note on the equivalence between definitions of tangent vectors in classical DF and SDG

Recall that classically, a tangent vector at p is defined as a linear derivation at p

$$V_p : C^\infty(U) \longrightarrow R$$

where $p \in U$ and U is an open subset of M. (NB: this definition is independent of U CHECK!)

On the other hand, a tangent vector at p in SDG is simply a map $v_p : D \longrightarrow M$ with $v_p(0) = p$.

Proposition 4.3 *If M is a smooth n-dimensional manifold, there is a one-to-one correspondence between*

$$\frac{\frac{R\text{-linear derivations at } p}{}}{\text{Maps } v : D \longrightarrow M \text{ with } v(0) = p}$$

Proof: We may assume as well that $M = R^n$ and $p = 0$. This proposition is a consequence of Hadamard's. Let V_p be a R-linear derivation at p and $f : U \longrightarrow R$ with $0 \in U$ and U convex open in M. By Hadamard's

$$f(x) = f(0) + \sum_i x^i g_i(x)$$

Since $V_p(c) = 0$ (as easily checked), $V_0(f) = V_0(\sum_i x^i g_i(x))$. Hence,

$$\begin{aligned} V_0(f) &= V_0(\textstyle\sum_i x^i g_i(x)) \\ &= \textstyle\sum_i V_0(x^i g_i(x)) \\ &= \textstyle\sum_i g_i(0) V_0(x^i) + 0 \\ &= \textstyle\sum_i \partial f / \partial x^i(0) V_0(x_i) \\ &= (\textstyle\sum_i a_i \partial / \partial x^i|_0)(f) \end{aligned}$$

Thus the operator V_0 is completely determined by the n-tuple $(a_1, \ldots, a_n)$ (since the operators $\partial/\partial x^i|_0$'s are clearly linearly independent)

By Kock-Lawvere, a map $v_0 : D \longrightarrow R^n$ is of the form $v_0(d) = (a_1 d, \ldots, a_n d)$ for unique n-tuple $(a_1, \ldots, a_n)$. Thus the correspondence can be described as follows: if V_0 is given, we let $v_0(d) = \sum_i V_0(x^i)d$. Conversely, if v_0 is given, say, $v_0(d) = (a_1 d, \ldots, a_n d)$, we let $V_0 = \sum_i a_i \partial/\partial x^i|_0$.

References

[1] V.Fock, *The theory of space time and gravitation*. Pergamon Press, New York, London, Paris, Los Angeles, 1959

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

[3] I.Moerdijk and G.E.Reyes, *Models for Smooth Infinitesimal Analysis*, Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo, Hong Kong, Barcelona 1991

[4] G.E.Reyes, *Pamphlets on SDG and GR*, in reyes-reyes.com

[5] G.E.Reyes, covariantderivation.pdf in *Pamphlets on SDG and GR*

[6] M.Spivak, *A comprehensive Introduction to Differential Geometry*, volumes I-V, (second edition) Publish or Perish Inc., 1979

[7] J.L.Synge and A.Schild,*Tensor Calculus*, Dover Publications, Inc. New York, 1978