PRINCIPAL BUNDLES AND CURVATURE PART II; PRINCIPAL BUNDLES

1. INTRODUCTION

The[1] plan for this section is as follows. First we will present principal bundles through the example of frame bundles of vector bundles. This has the advantage of a canonical system of coordinates. Then we will look at the more general form of the theory.

There are several good reasons for moving from vector bundles to principal bundles.

1. Change of the local basis of sections is a group action on the vector bundle but in the vector bundle situation it is difficult to see what the group action is really doing. In principal bundles this knowledge becomes more systematic and we understand it better.

2. Historically connections origonated as Christoffel symbols in the middle of the 1800s. In metric differential geometry they were derived from the metric coefficients in a particular coordinate system and formulas were derived for their changes when the coordinate system was changed. Deeper understanding of these changes had to await the invention of Differential Forms by Elie Cartan and of vector bundles. However, it is not easy to give conditions under which a set of differential forms on a manifold gives a local connection in a coordinate patch, nor is the formula for change of basis very informative. All this changes when we move to principal bundles. Here connections can be characterized in a satisfying way by giving forms with certain invariance properties and the formulas for change of basis become almost obvious as due to the action of the Group of the principal bundle.

3. Once a connection is put on a principal bundle it metastasizes through all the associated bundles in a natural way. This replaces what look like haphazard and ad hoc methods for constructing connections on bundles related to a vector bundle.

4. The role of "preserving a structure" in relation to connections is clearly reflected in the group of the principal bundle. Various properties like skew symmetry of the connection for the tangent bundle using orthonormal bases in Riemannian Geometry become understandable as properties of the Lie Algebra of the group $O(n, \mathbb{R})$.

5. It is a convenient mathematical setting in which to study Gauge Theory.

These are fairly compelling reasons for introducing principal bundles. For amusement I will offer an analogy. The role of principal bundles in differential

[1] December 5, 2019

geometry might be compared to the role of evolution in biology. It is perfectly possible to work in a specific area of biology, for example turtles, without evolution. But the big scale structure of turtle theory will never be clear without evolution. Each new turtle requires a specific investigation.

For a specific example in Differential Geometry, consider the following.

It is natural to require in the Riemannian case that the covariant derivative satisfy

$$d < \sigma, \rho > = < D\sigma, \rho > + < \sigma, D\rho >$$

If we come at this from the principal bundle approach, we don't have to worry about whether it's natural; it will *automatically* fall out of the theory when the group is $O(n, \mathbb{R})$, which is the natural choice of group in the Riemannian context.

2. REVIEW OF VECTOR BUNDLES

In this section we will briefly review what we know about Vector Bundles and digest the formulas for future use. All these were derived in Part I.

Let $\pi : \mathcal{E} \to M$ be a vector bundle. We will denote by $\sigma = \{\sigma_1, \ldots, \sigma_n\}$ a local basis of sections over $U \subseteq M$ and by $\{\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n\}$ second such. We will denote individual sections over U by τ, ρ. Thus $\sigma_\alpha, \tau, \rho \in \Gamma(U, \mathcal{E})$.

Let $\{\sigma_\alpha\}$ be a local basis of sections over U. For any $x \in U$ there is a vector space $\pi^{-1}[x]$ isomorphic to E above x in $\mathcal{E}$, and $\sigma(x) = (\sigma_1(x), \ldots, \sigma_n(x))$ will be a basis of this vector space. (The basis may be arbitrary or it may be a special kind of basis, for example orthonormal, depending on additional structures on the manifold, for example a Riemannian structure.) In general, if $\{\tilde{\sigma}_\alpha\}$ is a second such local basis of sections, then they will be connected by a group element $\big(g^\alpha_\beta(x)\big) \in \mathrm{GL}(n, \mathbb{R})$ so that

$$\tilde{\sigma}_\beta = \sigma_\alpha g^\alpha_\beta(x)$$

We assume that everything is smooth (C^∞) so that we may regard $(g^\alpha_\beta(x))$ as a local section of $\mathrm{GL}(n, \mathbb{R})$ over U. (Here we are thinking of $\mathrm{GL}(n, \mathbb{R})$ as being the trivial bundle $U \times \mathrm{GL}(n, \mathbb{R})$ over U.)

Now we wish to digest the material on Vector Bundles that we derived in Part I so it will be handy for reference.

Vector Bundle $(\mathcal{E}, M, \pi)$ where $\pi : \mathcal{E} \to M$

Fibre is E

U coordinate patch of M

$\{\sigma_1, \ldots, \sigma_n\}$ fixed local basis of sections over U

$u^1, \ldots, u^d$ local coordinates on U

ρ is a local section of $\mathcal{E}$ and thus

$$\rho = \sigma_\alpha \rho^\alpha(u^1, \ldots, u^d)$$

$$\rho \text{ has local coordinates } u^1, \ldots, u^d, \rho^1, \ldots, \rho^n \qquad \text{(bundle coordinates in } \mathbb{R}\text{)}$$

$$T_\rho(\mathcal{E}) \text{ has a basis } \frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^d}, \frac{\partial}{\partial \rho^1}, \ldots, \frac{\partial}{\partial \rho^n}$$

$$v \in T_\rho(\mathcal{E}) \text{ has a representation } v = v^i \frac{\partial}{\partial u^i} + V^\alpha \frac{\partial}{\partial \rho^\alpha}$$

$$w \text{ is in the VERTICAL space } \Leftrightarrow w = W^\alpha \frac{\partial}{\partial \rho^\alpha} \Leftrightarrow \text{ all } w^i = 0$$

The connection is characterized by its HORIZONTAL SPACE. We can define this by a projection Π at each point in $\mathcal{E}$ which projects the tangent space at that point of the bundle to the vertical space which is the tangent space of the fibre and a subset of the full tangent space. Given a connection, the projection Π is defined by

$$\Pi\Big(v^i \frac{\partial}{\partial u^i} + V^\alpha \frac{\partial}{\partial \rho^\alpha}\Big) = \Big(V^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta v^i\Big) \frac{\partial}{\partial \rho^\alpha}$$

Note how Π is a projection onto the Vertical Space. The Horizontal Space is defined as the kernel of Π;

$$v \text{ is in the HORIZONTAL space } \Leftrightarrow 0 = \Pi v \Leftrightarrow$$

$$0 = \begin{pmatrix} W^1 \\ W^2 \\ \vdots \\ W^n \end{pmatrix} = \begin{pmatrix} \Gamma^1_{\beta\,1}\rho^\beta, & \cdots & \Gamma^1_{\beta\,d}\rho^\beta, & 1, & 0, & \cdots & 0 \\ \Gamma^2_{\beta\,,1}\rho^\beta, & \cdots & \Gamma^2_{\beta\,d}\rho^\beta, & 0, & 1, & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \Gamma^n_{\beta\,,1}\rho^\beta, & \cdots & \Gamma^n_{\beta\,d}\rho^\beta, & 0, & 0, & \cdots & 1 \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^d \\ V^1 \\ \vdots \\ V^n \end{pmatrix}$$

$$\Leftrightarrow 0 = \Gamma^\alpha_{\beta i} \rho^\beta v^i + V^\beta$$

A vector field X on M has the form

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

The LIFT of X to $H\mathcal{E}$ is

$$\tilde{X} = X^i \frac{\partial}{\partial u^i} + \tilde{X}^\alpha \frac{\partial}{\partial \rho^\alpha}$$

where $\tilde{X}$ is in the Horizontal Space and thus

$$\begin{aligned} 0 &= \Pi \tilde{X} \\ 0 &= \tilde{X}^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta X^i \end{aligned}$$

so

$$\tilde{X} = X^i \frac{\partial}{\partial u^i} - \Gamma^\alpha_{\beta i} \rho^\beta X^i \frac{\partial}{\partial \rho^\alpha}$$

We wish to lift the curve $x(t)$ from M to $\mathcal{E}$. We do this by lifting it's tangent vector $\dot{x}(t)$ to $H\mathcal{E} \subseteq \mathcal{TE}$ by using the lift equations and then integrating along the lift to get the curve in $\mathcal{E}$. (Technically, it would be best to embed $\dot{x}(t)$ in a vector field in a neighborhood of the curve $x(t)$, lift the whole field, and then pursue the integral curve in $\mathcal{E}$). Let's look at these equations explicitly.

Curve $x(t)$ with coordinates $u^i(t)$ and

$$\text{tangent vector } X(t) = \frac{du^i}{dt} \frac{\partial}{\partial u^i}$$

$$\text{Lift:} \qquad \tilde{X}(t) = \frac{du^i}{dt} \frac{\partial}{\partial u^i} - \Gamma^\alpha_{\beta i} \rho^\beta \frac{du^i}{dt} \frac{\partial}{\partial \rho^\alpha}$$

Equations of the lift $\tilde{x}(t)$ of $x(t)$ which is the integral curve of $\tilde{X}(t)$ and has coordinates $\tilde{u}^i(t), \rho^\alpha(t)$:

$$\begin{aligned} \frac{d\tilde{u}^i}{dt} &= \frac{du^i}{dt} \\ \frac{d\rho^\alpha}{dt} &= \tilde{X}^\alpha = -\Gamma^\alpha_{\beta i} \rho^\beta \frac{du^i}{dt} \end{aligned}$$

where $\tilde{u}^i(0) = u^i(0)$ are the coordinates of $x(0)$ so that $\pi \tilde{x}(t) = x(t)$ and where $\rho^\alpha(0)$ are the coordinates of an arbitrary vector lying over $x(0)$ in $\mathcal{E}$.

To further understand what the lift of the curve $x(t)$ means, recall that

$$\begin{aligned} D\rho &= \sigma_\alpha \mathcal{D} \rho^\alpha \\ &= \sigma_\alpha (d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta) \\ &= \sigma_\alpha (d\rho^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta du^i) \end{aligned}$$

Thus for a tangent vector v in $T(\mathcal{E})$

$$\begin{aligned} D\rho(v) &= \sigma_\alpha (d\rho^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta du^i) \Big(v^j \frac{\partial}{\partial u^j} + V^\beta \frac{\partial}{\partial \rho^\beta} \Big) \\ &= \sigma_\alpha \big(V^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta v^i \big) \end{aligned}$$

Since the tangent vector to the lift is

$$\dot{\tilde{x}} = \frac{du^i}{dt} \frac{\partial}{\partial u^i} - \Gamma^\alpha_{\beta i} \rho^\beta \frac{du^i}{dt} \frac{\partial}{\partial \rho^\alpha}$$

we see

$$D\rho(\dot{\tilde{x}}) = 0$$

Just to be completely clear here, ρ is an element of the vector bundle $\mathcal{E}$. It is determined by coordinates u^i which determine the point of U over which ρ

lies, and by the coordinates ρ^α. The u^i determine $\Gamma^\alpha_{\beta i}$, and the $\Gamma^\alpha_{\beta i}$ and the ρ^α determine $D\rho$ which is closely related to the projection Π from the *Tangent Space* of $\mathcal{E}$ to the *Vertical Space* of $\mathcal{E}$. Note that ρ lives in $\mathcal{E}$, *not* in the tangent space $T(\mathcal{E})$.

Now let's review the equations for parallel transport of a vector ρ along a curve $x(t)$ in U. We recall from part I that the equations for parallel transport

$$\begin{aligned}
0 &= D\rho(\dot{x}) \;=\; \sigma_\alpha \mathcal{D}\rho^\alpha(\dot{x}) \\
&= \sigma_\alpha(d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta)\Big(\frac{du^j}{dt}\frac{\partial}{\partial u^j}\Big) \\
&= \sigma_\alpha(\frac{\partial \rho^\alpha}{\partial u^i}du^i + \Gamma^\alpha_{\beta i}\rho^\beta du^i)\Big(\frac{du^j}{dt}\frac{\partial}{\partial u^j}\Big) \\
&= \sigma_\alpha\Big(\frac{\partial \rho^\alpha}{\partial u^i}\frac{du^i}{dt} + \Gamma^\alpha_{\beta i}\rho^\beta \frac{du^i}{dt}\Big) \\
&= \sigma_\alpha\Big(\frac{d\rho^\alpha}{dt} + \Gamma^\alpha_{\beta i}\rho^\beta \frac{du^i}{dt}\Big)
\end{aligned}$$

Thus we see that lifting $x(t)$ to $\tilde{x}(t)$ is just another way of describing parallel transport of vectors.

Finally, we recall that

$$\begin{aligned}
\widetilde{[X,Y]} - [\tilde{X},\tilde{Y}] &= \Omega_\beta{}^\alpha(X,Y)\rho^\beta \frac{\partial}{\partial \rho^\alpha} \\
&= R_\beta{}^\alpha{}_{ij} X^i Y^j \rho^\beta \frac{\partial}{\partial \rho^\alpha}
\end{aligned}$$

Thus if one knows the horizontal space, one knows the lifts of vector fields and thus one knows the curvature.

3. FROM VECTOR BUNDLES TO FRAME BUNDLES

Now we are ready to introduce the Principal Bundle. In the present context we have two pictures of the principle bundle, the first of which is called the frame bundle.

1. In the first picture we get the principal bundle by replacing the fibre $\pi^{-1}[x]$ of $\mathcal{E}$ by a fibre consisting of all the different bases of the fibre E of $\mathcal{E}$. An element of the new fibre over x will be denoted by $\sigma(x) = \{\sigma_1(x), \ldots, \sigma_n(x)\}$. A section over $U \subseteq M$ will be a smooth choice of a basis in the fibre of each $x \in U$.

2. In the second picture we get the principal bundle by replacing fibre $\pi^{-1}[x]$ of $\mathcal{E}$ by a fibre consisting of the group G of the vector bundle $\mathcal{E}$. In general

this group is $\mathrm{GL}(n,\mathbb{R})$ and this is what we are using in our examples but it could be, for example, $\mathrm{O}(n,\mathbb{R})$ or a number of other groups.

The two pictures are equivalent but not canonically so. To transition between the two pictures we must select a fixed local section σ_0 over U. Then for any $x \in U$ and any element $\sigma(x)$ in P, we can find a group element $k(x) \in G$ for which $\sigma(x) = \sigma_0(x)k(x)$. Thus the local section $k(x)$ of $U \times \mathrm{GL}(n,\mathbb{R})$ forms a local section in the second picture. We symbolize this as

$$\sigma \overset{\sigma_0}{\longleftrightarrow} k$$

Naturally if we select a different local basis σ_1 the correspondence will change. If $\sigma_1 = \sigma_0 h$ then

$$\sigma_0 k = \sigma = \sigma_1 \tilde{k} = \sigma_0 h \tilde{k}$$

and thus

$$\begin{aligned} k &= h\tilde{k} \\ \tilde{k} &= h^{-1}k \end{aligned}$$

so

$$\sigma \overset{\sigma_1}{\longleftrightarrow} h^{-1}k$$

Since all these objects are local sections I have quietly dropped out the x.

We are now going to perform some of the same activities in the Principal Bundle that we did in vector bundles. The new feature is that instead of the parallel translation of a vector field along a curve we are going to transform an entire frame. This will then produce a group element and eventually an element of the Lie Algebra of the group.

Since the frame is made up of vectors, initially the change will require little more than bookkeeping. As before, we let $x(t)$ be a curve in $U \subseteq M$ where U is a coordinate patch with coordinates $u^1, \ldots, u^d$ and $\{\sigma^\alpha\}$ is a basis made from sections σ^α over U where $1 \leq \alpha \leq n$. An element $p \in P$ will be represented by a frame, and we will assume that the frame coincides in the fibre over $x(0)$ with the base frame $\sigma_0 = \{\sigma^\alpha(x(0)\}$. That is, we impose the initial condition that the transported frame $\tilde{\sigma}(t)$ coincides with the base frame σ_0 over $x(0)$ when $t = 0$

Now we transport each vector $\sigma_\alpha(0)$ of the base frame along the curve $x(t)$ to get a frame field $\tilde{\sigma}_\alpha(t)$. For small t (at least) the vectors $\tilde{\sigma}_\alpha(t)$ are linearly independent since $\tilde{\sigma}_\alpha(0)$ are. Hence there are coefficients $k^\alpha_\beta(t)$ so that

$$\tilde{\sigma}_\beta(t) = \sigma_\alpha(t)k^\alpha_\beta(t)$$

where $k^\alpha_\beta(t)$ are the coefficients of the basis vector $\tilde{\sigma}_\beta(t)$ in the basis $\{\sigma_\alpha(t)\}$. Here and in the future we are using $\sigma(t)$ for the value of the base frame σ^α over $x(t)$.

Since the $\tilde{\sigma}_\beta(t)$ are linearly independent (for small t at least) the elements $k^\alpha_\beta(t)$ may be put into a matrix $(k^\alpha_\beta(t))$ and regarded as an element of the group GL$(n, \mathbb{R})$. For each β, the elements $(k^\alpha_\beta(t))$ play the same role in the principal bundle that the coefficients $\rho^\alpha(t)$ played in the vector bundle with parallel transfer. Thus the equations for the transfer of the frame are the same;

$$\begin{aligned} \mathcal{D}k^\alpha_\beta(\dot{x}) &= 0 \qquad \text{for } \beta = 1, \ldots, n \\ (dk^\alpha_\beta + \omega_\gamma{}^\alpha k^\gamma_\beta)\dot{x} &= 0 \\ \frac{dk^\alpha_\beta}{dt} + \Gamma^\alpha_{\gamma i} k^\gamma_\beta \frac{du^i}{dt} &= 0 \end{aligned}$$

The initial conditions for the $k^\alpha_\beta(t)$ are

$$k^\alpha_\beta(0) = \delta^\alpha_\beta$$

reflecting that $\tilde{\sigma}(0) = \sigma(0)$. Thus the connection coefficients $\Gamma^\alpha_{\beta i}$ for the principal bundle are identical to those for the vector bundle $\mathcal{E}$ from which it arose.

Note that because parallel transfer from 0 to $s+t$ is the same as from 0 to s and from s to $s+t$, we are dealing with a one parameter subgroup of GL$(n, \mathbb{R})$. This then introduces a new feature, since $\frac{dk}{dt}|_{t=0}$ will then be an element of the Lie Algebra $\mathfrak{g}l(n, \mathbb{R})$ of the group GL$(n, \mathbb{R})$. Let's look at what it will be. We have

$$\frac{dk^\alpha_\beta}{dt} = -\Gamma^\alpha_{\gamma i} k^\gamma_\beta \frac{du^i}{dt}$$

so, letting $t \to 0$, we have

$$\begin{aligned} \frac{dk^\alpha_\beta}{dt}|_{t=0} &= [-\Gamma^\alpha_{\gamma i} k^\gamma_\beta \frac{du^i}{dt}]_{t=0} \\ &= -\Gamma^\alpha_{\gamma i} \delta^\gamma_\beta \frac{du^i}{dt}|_{t=0} \\ &= -\Gamma^\alpha_{\beta i} \frac{du^i}{dt}|_{t=0} \\ &= -\omega_\beta{}^\alpha(\dot{x}(0)) \end{aligned}$$

Since $\dot{x}(0)$ can be any element of $T_{\pi(p)}(M)$, we can now interpret $\omega_\beta{}^\alpha$ in a new way; $\omega_\beta{}^\alpha$ is a Lie Algebra valued one form. We input a tangent vector of M and it outputs a element in the Lie Algebra $\mathfrak{g}l(n, \mathbb{R})$ of G. We have looked at this for the example $G =$GL$(n, \mathbb{R})$ but it would work exactly the same way if G were some other matrix group. This is our first new insight; the connection on the Vector Bundle $\mathcal{E}$ functions as a connection ω in the Principal Bundle and there forms a Lie Algebra valued one-form.

This is one of the two properties we wish to have for connections on principal bundles. The second property is an invariance principle so that the covariant derivative D will not depend on the choice of basis. We will investigate this in a subsequent section.

Next we need to discuss the analog of the Horizontal Space for vector bundles. We also want to see that the effect of parallel translation can be looked at as an activity in GL$(n, \mathbb{R})$ and we can use this to connect it with the Lie Algebra $\mathfrak{g}l(n, \mathbb{R})$.

Since the coordinates in the principal bundle can be taken as u^i, k^α_β, the vectors in the tangent space T_pP look like

$$v = v^i \frac{\partial}{\partial u^i} + V^\alpha_\beta \frac{\partial}{\partial k^\alpha_\beta}$$

Vectors in the vertical space V_p will look like

$$w = W^\alpha_\beta \frac{\partial}{\partial k^\alpha_\beta}$$

The connection ${\omega_\beta}^\alpha$ gives rise, just as in the vector bundle, to a projection Π_p defined by

$$\begin{aligned} W^\alpha_\beta \frac{\partial}{\partial k^\alpha_\beta} \; = \; \Pi_p\, v &= \Pi_p\Big(v^i \frac{\partial}{\partial u^i} + V^\alpha_\beta \frac{\partial}{\partial k^\alpha_\beta}\Big) \\ &= \Big(V^\alpha_\beta + \Gamma^\alpha_{\gamma i} k^\gamma_\beta v^i\Big)\frac{\partial}{\partial k^\alpha_\beta} \end{aligned}$$

It is obvious that $\Pi_p[T_p(P)] = V_p(P)$ and that if $w \in V_p(P)$ then $\Pi_p(w) = w$, since all $w^i = 0$. Hence Π_p is a genuine projection. Since dim $T_p(P) = n^2 + d$ and dim $V_p = n^2$ we have dim ker $\Pi_p = d =$ dim M. We define

$$H_p(P) = \text{ker } \Pi_p$$

and thus dim $H_p(P) = d$. Thus $H_p(P)$ is isomorphic to $T_{\pi(p)}(M)$. We will illustrate this later. We note that if the structure group G is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ instead of the whole group, the details would be only slightly different. We will consider this later.

Next we note that the elements of V_p are essentially $n \times n$ matrices. Indeed

$$W^\alpha_\beta \frac{\partial}{\partial k^\alpha_\beta} \longleftrightarrow W^\alpha_\beta$$

and these matrices may be regarded as elements of the Lie Algebra $\mathfrak{g}l(n, \mathbb{R})$.

4. COORDINATES ON THE PRINCIPAL BUNDLE AND INVARIANCE OF THE CONNECTION

It is critical that the Horizontal Space $H(P)$ be invariant under the right $G-$action. This means we must have

$$H_{\sigma g}(P) = R_{g*} H_\sigma(P)$$

This will be a corollary of the formula

$$\Pi_{\sigma g} R_{g*}(\vec{v}) = R_{g*}\Pi_\sigma(\vec{v})$$

which we are going to derive in this section. This formula can be written in another way, as

$$\mathcal{D}_{\sigma g} = \mathrm{Ad}(g^{-1})\,\mathcal{D}_{\sigma}$$

Formulas like these were derived for Vector Bundles and we need "only" replicate the process. However, there are some tricky aspects.

As has often been the case in these notes I wish to do this proof in coordinates so we have the techniques and coordinate system available should we need them later. Having a "good" coordinate system will perhaps facilitate this and later adventures. What to choose for a good coordinate system is often a matter of experience, which one accumulates by trying coordinate systems which are less good. Recall that we are working over a single coordinate patch and that we have chosen a local basis of sections $\sigma_{01}, \ldots, \sigma_{0n}$ which we will keep in the shadows but is necessary for coordinates. All sections of P can then be written as $\sigma = \sigma_0 k$ where $k = \big(k^\alpha_\beta(u^1, \ldots, u^d)\big)$.

We will work with two sections $\sigma = \sigma_0 k$ and $\tilde\sigma = \sigma_0 \tilde k$ related by $\tilde\sigma = \sigma g$ where $g = \big(g^\alpha_\beta(u^1, \ldots, u^d)\big)$. We then have

$$\sigma = \sigma_0 k \qquad\qquad \sigma_0 \tilde k = \tilde\sigma = \sigma g = \sigma_0 k g$$
$$\tilde k \;=\; k g$$

Since $\sigma = \sigma_0 k$ it would be possible to take (k^α_β, u^i) as coordinates for P. Possible but not such a good idea. The (k^α_β, u^i) transform in an inconvenient manner under the $G-$action because of the equation $\tilde k = kg$. We are free to use any coordinates we like, so we will use a more convenient set.

I now suggest how we might find a more convenient set. Recall that the coordinates $\{\rho^1, \ldots, \rho^n\}$ for a section ρ of a vector bundle transformed under change of local basis of sections as $\tilde\rho = h\rho$ (where $h = g^{-1}$). We will be better able to utilize the insights of the vector bundles if we have coordinates the transform the same way. Set $h = g^{-1}, \quad \ell = k^{-1}, \quad \tilde\ell = \tilde k^{-1}$. Then since $\tilde k = kg$, we have $\tilde k^{-1} = g^{-1}k^{-1}$ and thus $\tilde\ell = h\ell$. Thus I propose to take $\{\ell^\alpha_\beta\}$ as the coordinates of $\sigma = \sigma_0 k$ (and thus $\{\tilde\ell^\gamma_\delta\}$ become the coordinates of $\tilde\sigma = \sigma_0 \tilde k$). The equation $\tilde\ell = h\ell$ now mirrors $\tilde\rho = h\rho$, as desired.

Explicitly, if $p \in P$ then p is represented by some σ on the fibre over some $x \in U$ with coordinates $(u^1, \ldots, u^d)$ and $\sigma = \sigma_0 k$ for some $k \in \mathrm{GL}(n, \mathbb{R})$ and, setting $\ell = k^{-1}$ we have the coordinates of $p \in P$ are

$$p \longleftrightarrow (\ell^\alpha_\beta, u^i)$$

Since $\tilde\ell = h\ell$

$$\tilde\ell^\alpha_\beta = h^\alpha_\gamma \ell^\gamma_\beta$$

Now we recall that for a vector bundle if we have

$$\vec v = V^\alpha \frac{\partial}{\partial \rho^\alpha} + v^i \frac{\partial}{\partial u^i}$$

then

$$\Pi_\sigma \vec{v} = \Big(V^\alpha_\beta + \Gamma^\alpha_{\gamma i} v^i \rho^\gamma\Big) \frac{\partial}{\partial \rho^\alpha}$$

where we use coordinates (ρ^α, u^i). We define Π_σ by a similar formula where we replace ρ^α by ℓ^α_β. Then we have

$$\begin{aligned} \vec{v} &= V^\alpha_\beta \frac{\partial}{\partial \ell^\alpha_\beta} + v^i \frac{\partial}{\partial u^i} \\ \Pi_\sigma \vec{v} &= \Big(V^\alpha_\beta + \Gamma^\alpha_{\gamma i} v^i \ell^\gamma_\beta\Big) \frac{\partial}{\partial \ell^\alpha_\beta} \end{aligned}$$

Ten seconds thought shows that this is a projection onto the vertical space; if $\vec{v} = V^\alpha_\beta \frac{\partial}{\partial \ell^\alpha_\beta}$ (all $v^i = 0$ so $\vec{v} \in V_\sigma(P)$)) then $\Pi_\sigma \vec{v} = \vec{v}$. Thus the crucial question is whether it is invariant under the $G-$action. This is where the choice of coordinates is important; with these coordinates the calculation greatly resembles the calculation showing the covariant derivative in a vector bundle does not depend on the choice of basis. Formally one just replace ρ^α by ℓ^α_β.

Our goal is to prove the following

Theorem

$$\begin{aligned} \Pi_{\sigma g}\Big(R_{g*}\vec{v}\Big) &= R_{g*}\Big(\Pi_\sigma \vec{v}\Big) \\ \Pi_{R_g \sigma} \circ R_{g*} &= R_{g*} \circ \Pi_\sigma \end{aligned}$$

The proof consists of two parts. First we find a formula for R_{g*} and second we verify the formula; neither is difficult.

Since $R_g \sigma = \sigma g = \tilde{\sigma}$,

$$R_{g*} : T_\sigma(P) \to T_{\tilde{\sigma}}(P)$$

set $\vec{v} \in T_\sigma(P)$, $\quad \vec{w} \in T_{\tilde{\sigma}}(P)$

$$\begin{aligned} \vec{v} &= V^\alpha_\beta \frac{\partial}{\partial \ell^\alpha_\beta} + v^i \frac{\partial}{\partial u^i} \\ \vec{w} &= w^\gamma_\delta \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} + w^i \frac{\partial}{\partial u^i} \end{aligned}$$

Then we have only to construct the Jacobian of R_g to connect V^α_β to W^γ_δ. A tricky point is that $\tilde{\ell}^\gamma_\delta$ depends on both the ℓ^α_β and the u^i. Indeed

$$\tilde{\ell}^\gamma_\delta = h^\gamma_\alpha \ell^\alpha_\delta = h^\gamma_\alpha \delta^\beta_\delta \ell^\alpha_\beta$$

so

$$\frac{\partial \tilde{\ell}^\gamma_\delta}{\partial \ell^\alpha_\beta} = h^\gamma_\alpha \delta^\beta_\delta \qquad \text{and} \qquad \frac{\partial \tilde{\ell}^\gamma_\delta}{\partial u^i} = \frac{\partial h^\gamma_\alpha}{\partial u^i} \ell^\alpha_\delta$$

Then

$$\begin{aligned} W_\delta^\gamma &= V_\beta^\alpha \frac{\partial \tilde{\ell}_\delta^\gamma}{\partial \ell_\beta^\alpha} + v^i \frac{\partial \tilde{\ell}_\alpha^\gamma}{\partial u^i} \\ &= V_\beta^\alpha h_\alpha^\gamma \delta_\delta^\beta + v^i \, \frac{\partial h_\alpha^\gamma}{\partial u^i} \, \ell_\delta^\alpha \\ &= h_\alpha^\gamma V_\delta^\alpha + \frac{\partial h_\alpha^\gamma}{\partial u^i} \, v^i \ell_\delta^\alpha \end{aligned}$$

Comparison with the Vector Bundle calculation shows the results are formally identical (with ℓ_β^α replacing ρ^α) although the method of derivation was (superficially) different.

Part two is now a little calculation. We have as usual

$$\begin{aligned} \vec{v} &= V_\beta^\alpha \frac{\partial}{\partial \ell_\beta^\alpha} + v^i \frac{\partial}{\partial u^i} \\ \Pi_\sigma \vec{v} &= \big(V_\beta^\alpha + \Gamma_{\mu i}^\alpha v^i \ell_\beta^\mu \big) \frac{\partial}{\partial \ell_\beta^\alpha} \end{aligned}$$

Note that in $\Pi_\sigma \vec{v}$ there are no terms with $\frac{\partial}{\partial u^i}$. Hence

$$\begin{aligned} R_{g*} \Pi_\sigma \vec{v} &= R_{g*} \Big[\big(V_\delta^\alpha + \Gamma_{\mu i}^\alpha v^i \ell_\delta^\mu \big) \frac{\partial}{\partial \ell_\delta^\alpha} \Big] \\ &= h_\alpha^\gamma \big(V_\delta^\alpha + \Gamma_{\mu i}^\alpha v^i \ell_\delta^\mu \big) \frac{\partial}{\partial \tilde{\ell}_\delta^\gamma} \end{aligned}$$

using the formula previously derived.

For $\Pi_{\sigma g} R_{g*} \vec{v}$ we must use the formula $\tilde{\omega} = g^{-1}\, dg + \mathrm{Ad}(g^{-1})\omega$ for transforming the connection one-form from T_σ to $T_{\sigma g}$. This is automatic for the frame bundle, but for further development we should keep in mind that this would work in any Principal Bundle with a connection, (that is, a $\mathfrak{g}-$valued one form), satisfying this formula. The calculations would be identical. Hence the proof proves something more general than our original goal. Unpacking the formula we have

$$\begin{aligned} {\omega_\beta}^\alpha(\vec{v}) &= \Gamma_{\mu i}^\alpha v^i \\ {\omega_\nu}^\gamma(\vec{w}) &= \tilde{\Gamma}_{\nu i}^\gamma v^i \\ &= h_\alpha^\gamma \, \frac{\partial g_\nu^\alpha}{\partial u^i} \, v^i + h_\alpha^\gamma \Gamma_{\rho i}^\alpha \, g_\nu^\rho v^i \end{aligned}$$

Next we note $h_\alpha^\gamma g_\nu^\alpha = \delta_\nu^\gamma$ and $\frac{\partial}{\partial u^i}(h_\alpha^\gamma g_\nu^\alpha) = 0$ so we have

$$\frac{\partial h_\alpha^\gamma}{\partial u^i} \, g_\nu^\alpha + h_\alpha^\gamma \frac{\partial g_\nu^\alpha}{\partial u^i} = 0$$

Using this we finally have

$$
\begin{aligned}
\Pi_{\sigma g} R_{g*}\vec{v} &= \Pi_{\sigma g}\vec{w} \\
&= \Pi_{\sigma g}(W^\gamma_\delta \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} + v^i \frac{\partial}{\partial u^i}) \\
&= (W^\gamma_\delta + \tilde{\Gamma}^\gamma_{\nu i} v^i \tilde{\ell}^\nu_\delta) \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} \\
&= \Big(h^\gamma_\alpha V^\alpha_\delta + \frac{\partial h^\gamma_\alpha}{\partial u^i} v^i \ell^\alpha_\delta + (h^\gamma_\alpha \frac{\partial g^\alpha_\nu}{\partial u^i} v^i + h^\gamma_\alpha \Gamma^\alpha_{\rho i} g^\rho_\nu v^i) h^\nu_\sigma \ell^\sigma_\delta \Big) \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} \\
&= \Big(h^\gamma_\alpha V^\alpha_\delta + \frac{\partial h^\gamma_\alpha}{\partial u^i} v^i \ell^\alpha_\delta - \frac{\partial h^\gamma_\alpha}{\partial u^i} g^\alpha_\nu v^i h^\nu_\sigma \ell^\sigma_\delta + h^\gamma_\alpha \Gamma^\alpha_{\rho i} g^\rho_\nu v^i h^\nu_\sigma \ell^\sigma_\delta \Big) \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} \\
&= \Big(h^\gamma_\alpha V^\alpha_\delta + \frac{\partial h^\gamma_\alpha}{\partial u^i} v^i \ell^\alpha_\delta - \frac{\partial h^\gamma_\alpha}{\partial u^i} \delta^\alpha_\sigma v^i \ell^\sigma_\delta + h^\gamma_\alpha \Gamma^\alpha_{\rho i} \delta^\rho_\sigma v^i \ell^\sigma_\delta \Big) \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} \\
&= \Big(h^\gamma_\alpha V^\alpha_\delta + \frac{\partial h^\gamma_\alpha}{\partial u^i} v^i \ell^\alpha_\delta - \frac{\partial h^\gamma_\alpha}{\partial u^i} v^i \ell^\alpha_\delta + h^\gamma_\alpha \Gamma^\alpha_{\rho i} v^i \ell^\rho_\delta \Big) \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} \\
&= h^\gamma_\alpha \Big(V^\alpha_\delta + \Gamma^\alpha_{\rho i} v^i \ell^\rho_\delta \Big) \frac{\partial}{\partial \tilde{\ell}^\gamma_\delta} \\
&= R_{g*}\Pi_\sigma \vec{v}
\end{aligned}
$$

and the theorem is proved. Note carefully how the non-tensorial term $g^{-1}dg$ in the transformation formula for ω cancels out the "extra" term $\frac{\partial h^\gamma_\alpha}{\partial u^i} v^i \ell^\alpha_\delta$ in the transformation of V^α_δ in the seventh step.

5. ALTERNATE WAYS OF LOOKING AT CONNECTIONS ON THE PRINCIPAL BUNDLE

We now wish to examine Π_p from another point of view. Since the fibres are isomorphic to the group G, the vertical space $V_p(P)$ is isomorphic to the Lie Algebra $\mathfrak{g}$ of G. We wish to develop this isomorphism explicitly, which is easy.

Recall that an element $X \in \mathfrak{g}$ generates a local one parameter subgroup $k(t)$ of G which, for matrix groups can be written down explicitly as

$$k(t) = e^{Xt} = I + Xt + \frac{1}{2!}X^2t^2 + \frac{1}{3!}X^3t^3 + \dots$$

We can use the right action of G on p to form a path in the fibre

$$pk(t)$$

This is a path in the manifold P, so its derivative at $t = 0$ will yield a tangent vector to the fibre and thus be in $V_p(P)$

$$\frac{d}{dt}pk(t)|_{t=0} \in V_p(P)$$

More abstractly, we may obtain the vector field on P which corresponds to $X \in \mathfrak{g}$ by setting

$$\tilde{X}_p(f) = \frac{d}{dt} f(p\ \exp(Xt))|_{t=0} \qquad X \in \mathfrak{g}$$

where f is any smooth function on a neighborhood of p.

Since we are working with $\mathrm{GL}(n,\mathbb{R})$ all this can be carried out extremely explicitly. A one parameter group $k(t) = (k^\alpha_\beta(t))$ can be correlated with a Lie Algebra element

$$W^\alpha_\beta = \frac{d}{dt} k^\alpha_\beta(t)|_{t=0}$$

and $W = W^\alpha_\beta$ produces $k(t)$ by

$$k(t) = \exp(Wt)$$

The only limitation on $k(t)$ is that $\det k(t) \neq 0$ and $\mathfrak{g} = \mathfrak{g}l(n,\mathbb{R})$ is the full matrix algebra over $\mathbb{R}$. the action on P is given by

$$\sigma k(t) = (\sigma_\alpha k^\alpha_\beta(t))$$

and the element of $V_p(P)$ will be

$$\begin{aligned} \tilde{W}_p &= \frac{d}{dt}(\sigma_\alpha k^\alpha_\beta(t))|_{t=0} \\ &= W^\alpha_\beta \frac{\partial}{\partial k^\alpha_\beta} \end{aligned}$$

(For the last equation, recall that if a path on a manifold has coordinates $u^i(t)$ then the tangent vector is $\frac{du^i}{dt}\frac{\partial}{\partial u^i}$)

Using these methods we can now rewrite the projection Π_p in a way which is abstractly preferable. We have

a) the projection $\Pi_p : T_p(P) \to V_p(P)$

b) the isomorphism $W \mapsto \tilde{W}_p : \mathfrak{g} \to V_p$

These combine, using the first and the inverse of the second, to form a map

$$\mathcal{Z}_p : T_p(P) \to \mathfrak{g}$$

Notice that these constructions are not dependent on the particular choice of $G = \mathrm{GL}(n,\mathbb{R})$; they are available as long as Π_p is. The converse is also true; given a $\mathcal{Z}_p$ it is easy to construct a $\Pi_p : T_p(P) \to V_p(P)$ by composing $\mathcal{Z}_p$ with the map $W \to \tilde{W}_p$.

Thus we see that a connection can be specified in four formally different ways:

1. Through the differential forms ${\omega_\beta}^\alpha = \Gamma^\alpha_{\beta i} du^i$

2. Through the Horizontal Space $H_p(P)$

3. Through the projection $\Pi_p : T_p(P) \to V_p(P)$

4. Through the $\mathcal{Z}_p : T_p(P) \to \mathfrak{g} = \mathrm{Lie}(G)$

The specification of any of these four will give a connection on a Principal Bundle in an abstract sense, but we do not wish to study the situation in this generality. We wish to restrict our consideration to connections on Principal Bundles that arise from connections on Vector Bundles. For this to be true one needs a property we call invariance which is related to base change in the Vector Bundle, which we will study in a subsequent section.

In the more general situation we can even make a general definition for connection on a general fibre bundle, although we will not follow this trail. Given any fibre bundle $(\mathcal{E}, M, \pi, G)$ where G is the structure group, we set V_p equal to the tangent space to the *fibre* at p. A connection form on $\mathcal{E}$ s a vector valued differential form Π_p from $T_p(P)$ (one for each p) onto $V_p(P)$ and which is the identity on $V_p(P)$. Then for $t \in T_p(P)$

$$t \in H_p(P) \Longleftrightarrow \Pi_p(t) = 0$$

Such a differential form is called HORIZONTAL.

I do not wish to pursue the subject at this level of generality. For Principal Bundles and Vector Bundles we wish to restrict our attention to connections which are well behaved under the right action of G. (There need be no such action in a general fibre bundle.) This will be treated in a subsequent section. Briefly, we will study connections that satisfy the additional condition that

$$R_{g*}\Pi_p = \Pi_{pg} R_{g*}$$

which is equivalent to

$$\omega_{pg} = g^{-1}dg + \mathrm{Ad}(g^{-1})\omega_p$$

6. COVARIANT DERIVATIVE OF HIGHER ORDER VECTOR VALUED FORMS

Let $\pi : \mathcal{E} \to M$ be a vector bundle and $\pi : P \to M$ the corresponding frame bundle, both with group G. The covariant derivative D acts on objects in the bundles

$$\begin{aligned} D\rho &= \sigma_\alpha \mathcal{D}\rho^\alpha = \sigma_\alpha(d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta) \\ D\sigma &= \sigma\omega \end{aligned}$$

In each case the output is a vector valued one form or frame valued 1-form. For the moment, let us concentrate on $\mathcal{E}$. We will denote the algebra of differential forms on the tangent bundle of M by

$$\mathcal{A}(M) = \bigoplus_{i=0}^{\infty} \mathcal{A}_i(M)$$

where $\mathcal{A}_i(M)$ is the differential forms of degree i on M. $\mathcal{A}_0(M)$ is just the functions on M. In a coordinate patch elements of $\mathcal{A}_p(M)$ look locally like

$$\sum_{0\le j_1,<\ldots<j_p\le n} f_{j_1,\ldots,j_p} du^{i_1}\wedge\ldots\wedge du^{i_p}$$

Vector valued differential forms look locally like

$$\sigma_\alpha \sum_{0\le j_1,<\ldots<j_p\le n} f^\alpha_{j_1,\ldots,j_p} du^{i_1}\wedge\ldots\wedge du^{i_p}$$

We will denote these by

$$\mathcal{A}(M,\mathcal{E}) = \mathcal{E}\otimes\mathcal{A}(M)$$

Thus the operator D is

$$\begin{aligned} D:\mathcal{A}_0(M,\mathcal{E}) &\to \mathcal{A}_1(M,\mathcal{E}) \\ \sigma_\alpha\otimes\rho^\alpha &\mapsto \sigma_\alpha(d\rho^\alpha + {\omega_\beta}^\alpha\rho^\beta) \end{aligned}$$

This is fine as far as it goes, but we want a D that works on $\mathcal{A}_p(M,\mathcal{E})$. We can define this inductively so that

$$D:\mathcal{A}_p(M,\mathcal{E})\to\mathcal{A}_{p+1}(M,\mathcal{E})$$

by means of Leibniz' rule:

$$\text{for } \eta\in\mathcal{A}_p(M) \text{ and } \rho \text{ a section of } \mathcal{E}$$
$$D(\rho\otimes\eta) = \rho\otimes d\eta \quad + \quad D\rho\wedge\eta$$

This works because $\mathcal{A}_p(M)$ is a free algebra.

We will do a couple of examples. Let $f\in\mathcal{A}_0(M)$ (a function). Then the rule says

$$D(\rho f) = \rho\, df + (D\rho)\, f$$

However, we could also do this the old way:

$$\begin{aligned} D(\rho f) &= \sigma_\alpha\mathcal{D}(\rho^\alpha f) \\ &= \sigma_\alpha\big(d(\rho^\alpha f) + {\omega_\beta}^\alpha\rho^\beta f\big) \\ &= \sigma_\alpha(\rho^\alpha\, df) + \sigma_\alpha\big(d\rho^\alpha + {\omega_\beta}^\alpha\rho^\beta\big)\, f \\ &= \rho\, df + (D\rho) f \end{aligned}$$

so we have consistency.

For our second example we compute $D^2\rho$:

$$\begin{aligned} D^2\rho &= D\,D(\sigma_\alpha\rho^\alpha) \\ &= D\big(\sigma_\alpha(d\rho^\alpha + {\omega_\beta}^\alpha\rho^\beta)\big) \\ &= \sigma_\alpha\big[d(d\rho^\alpha + {\omega_\gamma}^\alpha\rho^\gamma)\big] + (D\sigma_\alpha)\wedge(d\rho^\alpha + {\omega_\gamma}^\alpha\rho^\gamma) \\ &= \sigma_\alpha\big[d{\omega_\gamma}^\alpha\,\rho^\gamma - {\omega_\gamma}^\alpha\, d\rho^\gamma\big] + \sigma_\beta{\omega_\alpha}^\beta\wedge(d\rho^\alpha + {\omega_\gamma}^\alpha\rho^\gamma) \\ &= \sigma_\beta\big[d{\omega_\gamma}^\beta\,\rho^\gamma - {\omega_\gamma}^\beta\, d\rho^\gamma + {\omega_\alpha}^\beta d\rho^\alpha + {\omega_\alpha}^\beta\wedge{\omega_\gamma}^\alpha\,\rho^\gamma\big] \\ &= \sigma_\beta\big[d{\omega_\gamma}^\beta + {\omega_\alpha}^\beta\wedge{\omega_\gamma}^\alpha\big]\rho^\gamma \\ &= \sigma_\beta\,{\Omega_\gamma}^\beta\,\rho^\gamma \end{aligned}$$

Thus the second covariant derivative is intimately related to curvature, as second derivatives always are. Note the lack of any terms involving derivatives of ρ. This means that $D^2\rho$ is tensorial. We will discuss this later.

7. ASSOCIATED BUNDLES, Part I

In this section we show the algebraic construction of associated bundles and give some important examples. All our constructions here can be regarded as taking place on a single fibre, so the constructions will actually work for any vector space, although we have set up the notation to be suitable for bundles.

Let (P, M, π, G) be a Principal Bundle and let $\rho : G \to \mathrm{Aut}(V)$ be a representation of G on a vector space V. We will create a vector bundle with structure group G from ρ and V. To do this we we first form the Cartesian product $P \times V$. There is a natural right action of G on $P \times V$ that comes from the right action of G on P, namely

$$(\sigma, v)g = (\sigma g, \rho(g^{-1})v)$$

This is indeed a right action:

$$\begin{aligned}((\sigma, v)g)h &= (\sigma g, \rho(g^{-1})v)h \;=\; \big((\sigma g)h, \rho(h^{-1})\rho(g^{-1})v\big) \\ &= (\sigma(gh), \rho(h^{-1}g^{-1})v) \;=\; (\sigma(gh), \rho((gh)^{-1})v) \\ &= (\sigma, v)(gh)\end{aligned}$$

We can now take the quotient of $P \times V$ by the G-action, which is to say we make an equivalence class out of each orbit of G, and these equivalence classes form the desired associated vector space to P. We denote the equivalence class by $[\sigma, v]$. Thus $(\sigma, v) \sim (\sigma g, \rho(g^{-1})v)$ and we have

Def Let $g \in G$, $\sigma \in P$ and $v \in V$. Then

$$[\sigma, v] = [\sigma g, \rho(g^{-1})v]$$

This is indeed an equivalence relation; transitivity is checked in the previous calculation. Of course we could take a less sophisticated view and simply define the equivalence relation and form the classes without mentioning the G-action, but this way is more elegant.

The equivalence classes now form a vector space, for let $\xi_1 = [\sigma_1, v_1]$ and $\xi_2 = [\sigma_2, v_2]$. Since G is transitive on the fibres of P there is a $g \in G$ for which $\sigma_1 g = \sigma_2$. We now define

$$\begin{aligned}\xi_1 + \xi_2 &= [\sigma_1, v_1] + [\sigma_2, v_2] \\ &= [\sigma_1 g, \rho(g^{-1})v_1] + [\sigma_2, v_2] \\ &= [\sigma_2, \rho(g^{-1})v_1] + [\sigma_2, v_2] \\ &= [\sigma_2, \rho(g^{-1})v_1 + v_2]\end{aligned}$$

It is easy to show that with this definition that addition is well defined and with the obvious definition $\alpha[\sigma, v] = [\sigma, \alpha v]$ for scalar multiplication that the equivalence classes form a vector space.

It would be nice to have a notation for this new construction. The equivalence class lives in the space $(P \times V)$ so one notation would be $(P \times V)/\sim$. We will use the standard notation which is

$$P \times_\rho V$$

There are some standard isomorphisms for sections of these bundles which we will discuss after a few examples. The point of the construction is that a connection on a Vector Bundle $\mathcal{E}$ uploads to a connection on its principal bundle P which then metastasizes through all of P's associated bundles.

We should also pause a minute and look at the situation from a higher point of view. We are interested in associated *vector* bundles, but the construction above will work in more general circumstances. Suppose that V, instead of being a vector space, is any object on which G acts from the LEFT. This means that for $h, g \in G$ and $v \in V$ we have, denoting the action by $g.v$,

$$h.(g.v) = (hg).v$$

In our case we have the action given by

$$g.v = \rho(g)v$$

and this is a left action since $h.(g.v) = \rho(h)\rho(g)v = \rho(hg)v = (hg).v$

There is confusion lurking here because in the equivalence relation

$$[\sigma, v] = [\sigma g, \rho(g^{-1})v]$$

the action $g,\ v \to \rho(g^{-1})v$ is a *right* action. This is *irrelevant*. To see things in the proper light, we rewrite the equivalence relations as follows:

$$[\sigma g, v] = [\sigma g, \rho(g^{-1})\rho(g)v] = [\sigma, \rho(g)v]$$

Then the right action of G on P and the left action of G on V are more clearly visible. Unfortunately the equation $[\sigma g, v] = [\sigma, \rho(g)v]$ does not correlate too well with the forms we usually use in writing down expressions in vector spaces. However, it can be very useful for calculation; for example to decode $[\sigma(gh), \rho((gh)^{-1})v]$, we can use it as follows:

$$\begin{aligned}
[\sigma(gh), \rho((gh)^{-1})v] &= [(\sigma g)h, \rho(h^{-1})\rho(g^{-1})v] \\
&= [(\sigma g), \rho(h)\rho(h^{-1})\rho(g^{-1})v] \\
&= [\sigma g, \rho(g^{-1})v] \\
&= [\sigma, \rho(g)\rho(g^{-1})v] \\
&= [\sigma, v]
\end{aligned}$$

This should also convince you that it is correct, even though for some people it intuitively feels wrong.

Now suppose that P is the bundle of frames of a vector bundle $\mathcal{W}$ of dimension n and fibre W and that $G =$GL$(n, \mathbb{R})$. We then recover the bundle $\mathcal{W}$ from the construction in the following obvious way. We take V to be $\mathbb{R}^n$ $\rho(g) = g \in \mathrm{GL}(\mathbb{R}, n)$ and correlate $w \in W$ with $[\sigma, v]$ as follows. Let σ be a frame and $w \in W$. Then $w = \sigma_i w^i$ and we correlate

$$w \leftrightarrow [\sigma, \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}]$$

If we write $w = \sigma_i w^i$ as

$$w = \sigma \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix} = (\sigma_1, \ldots, \sigma_n) \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}$$

Then the familiar relations

$$w = (\sigma_1, \ldots, \sigma_n) \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix} w = (\sigma_1, \ldots, \sigma_n) g g^{-1} \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix} = (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n) \begin{pmatrix} \tilde{w}^1 \\ \vdots \\ \tilde{w}^n \end{pmatrix}$$

correlate perfectly with

$$[\sigma, \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}] = [\sigma g, \rho(g^{-1}) \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}] = [(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n), \begin{pmatrix} \tilde{w}^1 \\ \vdots \\ \tilde{w}^n \end{pmatrix}]$$

so that we see we merely have two different notations for the same idea. Note that each equivalence class correlates to a unique vector and that different members of the same equivalence class correlate with representations of the same vector in different bases. This is very nice way to look at basis change.

It is now easy to show that W^* and $\mathrm{Hom}(W, W)$ are examples of the Associated Bundle construction. First we review how the dual basis works and construct a matrix notation for it. Recall that we may form a basis σ^* for W^* by taking the dual basis of σ. We define σ^{*i} by

$$\sigma^{*i}(\sigma_j) = \delta^i_j$$

It is convenient, as we will see, to represent the dual basis as a *column*

$$\sigma^* = \begin{pmatrix} \sigma^{*1} \\ \vdots \\ \sigma^{*n} \end{pmatrix}$$

We shall refer to the elements λ of W^* as *covectors*. Then a covector λ may be expanded in terms of the basis σ^* as $\lambda = \lambda_i \sigma^{*i}$, and in matrix notation for this is

$$\lambda = \lambda_i \sigma^{*i} = (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} \sigma^{*1} \\ \vdots \\ \sigma^{*n} \end{pmatrix}$$

This is convenient because we can now find $\lambda(v)$ where $v \in W$ by matrix multiplication

$$\begin{aligned} \lambda(v) &= (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} \sigma^{*1} \\ \vdots \\ \sigma^{*n} \end{pmatrix} (\sigma_1, \ldots, \sigma_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\ &= (\lambda_1, \ldots, \lambda_n) \big(\sigma^{*i}(\sigma_j) \big) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\ &= (\lambda_1, \ldots, \lambda_n) \big(\delta^i_j \big) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\ &= \lambda_i v^i \end{aligned}$$

as expected. This looks even better (and more symmetric) if we replace the notation $\lambda(v)$ by the notation $\langle \lambda, v \rangle$, which emphasizes the duality between W^* and W.

Since $\lambda(v)$ does not depend upon the basis, we should be able to incorporate the G action into the above equation. As before we write $\tilde{\sigma} = \sigma g$ to indicate the G action on a basis of W. The corresponding dual basis of W^* will be $\tilde{\sigma}^*$ and since both σ^* and $\tilde{\sigma}^*$ are bases there will be matrix h for which $h\sigma^* = \tilde{\sigma}^*$ The h is written to the left of σ^* because of the way we write σ as a *column*. We now want to find h in terms of g, which is easy with our efficient notation. Using I for the identity matrix we have

$$\mathrm{I} \; = \tilde{\sigma}^* \tilde{\sigma} = h \sigma^* \sigma g = h \; \mathrm{I} \; g = hg$$

so that

$$\begin{aligned} h &= g^{-1} \\ \tilde{\sigma}^* &= g^{-1} \sigma^* \end{aligned}$$

Because of the inverse, this is a *right* action of G on the set of bases of W^*.

In order to proceed further without typographical stress we introduce some additional notation. We will denote an element of $\mathbb{R}^n$ by $\vec{v}$ and an element of $\mathbb{R}^{n*}$ by $\underline{\lambda}$;

$$\vec{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \qquad\qquad \underline{\lambda} = (\lambda_1, \ldots, \lambda_n)$$

and thus with $w \in W$ and $\lambda \in W^*$ we have

$$\begin{aligned} v &= \sigma\vec{v} = (\sigma_1, \ldots, \sigma_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\ \lambda &= \underline{\lambda}\sigma^* = (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} \sigma^{*1} \\ \vdots \\ \sigma^{*n} \end{pmatrix} \end{aligned}$$

This will simplify writing things down. Let's practice with the new notation by checking that change of basis works as it should:

$$\begin{aligned} v &= \sigma\,\vec{v} \;=\; \sigma g g^{-1}\vec{v} \;=\; \tilde{\sigma}\,\vec{\tilde{v}} \\ \lambda &= \underline{\lambda}\,\sigma^* = \underline{\lambda} g g^{-1}\sigma^* = \underline{\lambda}\,g\,\tilde{\sigma}^* \end{aligned}$$

Note that we forgot to find out how the coordinates of λ change when we change the basis, but from the above we can read off immediately that the formula must be $\tilde{\underline{\lambda}} = \underline{\lambda} g$. We already knew that $\vec{\tilde{v}} = g^{-1}\vec{v}$ which we can read off from the first line. Thus we have

$$\begin{aligned} \vec{\tilde{v}} &= g^{-1}\vec{v} \\ \tilde{\underline{\lambda}} &= \underline{\lambda} g \end{aligned}$$

Now let's check that we can compute the value of a linear functional on a vector in any coordinate system. Recall that we saw $\lambda(v) = \lambda_\alpha v^\alpha$.

$$\lambda(v) = \lambda_\alpha v^\alpha = \underline{\lambda}\vec{v} = \tilde{\underline{\lambda}} g^{-1} g \vec{\tilde{v}} = \tilde{\underline{\lambda}}\vec{\tilde{v}} = \tilde{\lambda}_\beta \tilde{v}^\beta$$

Next we want to describe W^* as an associated bundle to the bundle P of frames of W. To do this, we must find a representation ρ of G on some vector space V so that $P \times_\rho V$ will be isomorphic to W^*. This is only a matter of of getting the transformation under basis change correct.

For V we take $\mathbb{R}^{n*}$ and write elements of V as rows $\underline{\lambda} = (\lambda_1, \ldots, \lambda_n)$. Define ρ by

$$g.\underline{\lambda} = \rho(g)\underline{\lambda} = \underline{\lambda} g^{-1}$$

If $\lambda \in W$ and $\lambda = \underline{\lambda}\,\sigma^*$ is the representation of λ in the dual basis then the isomorphism is

$$\lambda \leftrightarrow [\sigma, \underline{\lambda}]$$

Now we must check that the isomorphism correctly survives basis change:

$$[\sigma, \underline{\lambda}] = [\sigma g, \rho(g^{-1})\underline{\lambda}] = [\tilde{\sigma}, \underline{\lambda} g] = [\tilde{\sigma}, \tilde{\underline{\lambda}}]$$

using $\tilde{\underline{\lambda}} = \underline{\lambda} g$ and other formulas derived above. It is trivial to verify the homomorphism properties so

$$W^* \leftrightarrow P \times_\rho \mathbb{R}^{n*}$$

is an isomorphism and we see that W^* is an associated vector bundle of P.

Without going into the trivial details, it should now be clear that with $\rho_1(g)\vec{v} = g\vec{v}$ on $\mathbb{R}^n$ and $\rho_2(g)\underline{\lambda} = \underline{\lambda}g^{-1}$ on $\mathbb{R}^{n*}$ we have a Tensor Product representation

$$\rho = \rho_1 \otimes \cdots \otimes \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_2$$

of G on

$$\mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n \otimes \mathbb{R}^{n*} \otimes \cdots \otimes \mathbb{R}^{n*}$$

(any finite number of $\mathbb{R}^n$'s and $\mathbb{R}^{n*}$'s) and this shows that

$$W \otimes \cdots \otimes W \otimes W^* \otimes \cdots \otimes W^*$$

is an associated bundle of P.

It will be instructive to examine $W \otimes W^*$, which is generated by elements of the form $w \otimes \lambda$, in more detail. We have the isomorphism

$$w \otimes \lambda \leftrightarrow [\sigma,\ \vec{w} \otimes \underline{\lambda}]$$

(where $\vec{w}$ and $\underline{\lambda}$ are the coordinates of w and λ in the basis σ). Let's look at the equivalence classes

$$\begin{aligned}
[\sigma,\ \vec{w} \otimes \underline{\lambda}] &\sim [\sigma g,\ \rho(g^{-1})(\vec{w} \otimes \underline{\lambda})] \\
&\sim [\tilde{\sigma},\ \rho_1(g^{-1})\vec{w} \otimes \rho_2(g^{-1})\underline{\lambda}] \\
&\sim [\tilde{\sigma},\ g^{-1}\vec{w} \otimes \underline{\lambda}g] \\
&\sim [\tilde{\sigma},\ \vec{\tilde{w}} \otimes \underline{\tilde{\lambda}}]
\end{aligned}$$

just as we would expect.

We can extract a little more out of $W \otimes W^*$. We define an action of $W \otimes W^*$ on W by

$$(w \otimes \lambda)(v) = w\lambda(v)$$

Moving this over to the matrix representation, any matrix may be counterfeited in $\mathbb{R}^n \otimes \mathbb{R}^{n*}$ by using its rows as in the following example.

$$\begin{aligned}
\begin{pmatrix} 1 & -3 \\ 2 & 5 \end{pmatrix} &\leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (1, -3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes (2, 5) \\
\begin{pmatrix} 1 & -3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix} &\leftrightarrow \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (1, -3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes (2, 5)\right]\begin{pmatrix} 3 \\ 2 \end{pmatrix} \\
&\leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}(-3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(16) \ = \ \begin{pmatrix} -3 \\ 16 \end{pmatrix}
\end{aligned}$$

This immediately generalizes; $\mathrm{Hom}(W, W)$ is isomorphic to $W \otimes W^*$ (with the action described above). If T has matrix (t^α_β) in the basis σ then we have an isomorphism

$$T \leftrightarrow [\sigma,\ t] \in P \times_{\rho_1 \otimes \rho_2} (\mathbb{R}^n \otimes \mathbb{R}^{n*})$$

where t is formed from the matrix (t^α_β) as shown above.

Then

$$T \leftrightarrow [\sigma g,\ \rho(g^{-1})t] \sim [\tilde{\sigma},\ g^{-1}tg]$$

using the action of ρ defined above. From this we read off that if T has matrix (t^α_β) in the σ basis then it will have matrix $g^{-1}(t^\alpha_\beta)g$ in the $\tilde{\sigma} = \sigma g$ basis. This indicates how our formalism can be a handy tool in linear algebra.

8. ASSOCIATED BUNDLES, Part II

Having defined Associated Bundles in the previous section, our next task is to use the connection on the principal bundle to generate a connection on each of the associated bundles. It is easy enough to give a formula for this, but to motivate the formula will take some time.

While there are no doubt many ways to do this, we will come at the problem by means of parallel displacement. One reason for this approach is that parallel displacement is a rather easy to visualize geometric process, which we hope will make the construction seem more natural and geometric.

Let $\mathcal{V} = (P, V, \rho)$ be an associated vector bundle to the principal bundle P. We visualize P as a bundle of frames of a vector bundle $\mathcal{E}$ and also as the bundle with fibre G as discussed above; there is a fixed local basis of sections σ which in this section will remain fixed.

The vector bundle structure for $\mathcal{V}$ is defined as follows: take $v = [\sigma,\ \vec{v}] \in \mathcal{V}$ and define $\pi(v)$ to be $\pi(\sigma)$. Since σ and σg are in the same fibre over M, $\pi(v)$ is independent of the choice of equivalence class for v.

Next we show how to describe a section. Let $(u^1, \ldots, u^n)$ be coordinates on some open set $U \subset M$. Then we can define a section by giving

$$[\sigma(u^1, \ldots, u^d),\ \vec{v}(u^1, \ldots, u^d)]$$

in which it is required that

$$\begin{aligned} \pi\big(\sigma(u^1, \ldots, u^d) &= x(u^1, \ldots, u^d) \in M \\ \pi[\sigma(u^1, \ldots, u^d),\ \vec{v}(u^1, \ldots, u^d)] &= \pi\big(\sigma(u^1, \ldots, u^d) = x(u^1, \ldots, u^d) \in M \end{aligned}$$

To see how to define parallel transport for $\mathcal{V}$ we will first review previous work in $\mathcal{E}$ and P. Let $x(t)$ be a curve in M with 0 in its domain. We will concentrate our attention at $t = 0$. Recall parallel transport in $\mathcal{E}$. Let $\rho_0 \in \pi^{-1}[x(0)]$. Then the parallel transport $\rho(t)$ of ρ_0 along $x(t)$ is governed by the equations

$$\begin{aligned} \rho(0) &= \rho_0 \\ D_{\dot{x}(t)}\rho &= \sigma_\alpha \left(\frac{d\rho^\alpha}{dt} + \Gamma^\alpha_{\beta i}\, \rho^\beta \dot{u}^i \right) \ = \ 0 \end{aligned}$$

For notational simplicity we will denote the frame σ along $x(t)$ by $\sigma(t) = \big(\sigma(u^1(t), \ldots, u^n(t)\big)$. Next we remind ourselves that the transport can be done for each of the particular vectors $\sigma_\alpha(0)$ of the basis $\sigma(0)$. The equations will then be, with $\tilde{\sigma}_\gamma(t) = \sigma_\alpha(t) g^\alpha_\gamma(t)$

$$\begin{aligned} \tilde{\sigma}_\alpha(0) &= \sigma_\alpha(0) \\ D_{\dot{x}(t)}\tilde{\sigma}_\gamma(t) &= \sigma_\alpha \left(\frac{dg^\alpha_\gamma}{dt} + \Gamma^\alpha_{\beta i}\, g^\beta_\gamma(t) \dot{u}^i(t) \right) = 0 \end{aligned}$$

where the $\Gamma^\alpha_{\beta i}$ are the connection coefficients for the frame σ. The first condition shows that $g^\alpha_\gamma(t_0) = \delta^\alpha_\gamma$. Thus we can replace these equations by

$$\begin{aligned} g^\alpha_\gamma(0) &= \delta^\alpha_\gamma \\ \frac{dg^\alpha_\gamma}{dt} + \Gamma^\alpha_{\beta i}\, g^\beta_\gamma(t) \dot{u}^i(t) &= 0 \end{aligned}$$

We collect the transported vectors $\tilde{\sigma}_\alpha(t)$ into a new frame $\tilde{\sigma} = (\sigma_1(t), \ldots, \sigma_n(t))$. Now suppose we let $\tilde{\rho}(t) = \tilde{\sigma}_\gamma(t)\rho^\gamma_0$ where the ρ^γ_0 are *constant*. The we have $\tilde{\rho}(0) = \tilde{\sigma}_\gamma(0)\rho^\gamma_0 = \sigma_\gamma(0)\rho^\gamma_0 = \rho_0$ and we also have

$$D_{\dot{x}(0)}\tilde{\rho}(t) = D_{\dot{x}(0)}\left[\tilde{\sigma}_\gamma(t)\rho^\gamma_0\right] = D_{\dot{x}(0)}\left[\tilde{\sigma}_\gamma(t)\right]\rho^\gamma_0 = 0$$

Thus $\rho(t)$ and $\tilde{\rho}(t)$ are both solutions of the same linear system of differential equations and have he same initial conditions. Hence by uniqueness of solutions $\tilde{\rho}(t) = \rho(t)$. Putting this into words, we can solve the parallel transport problem for each ρ_0 or we can solve it by translating the whole frame and using the coordinates of ρ at t_0 as the coefficients of the translated frame. Thus solving the problem in P automatically solves it in $\mathcal{E}$.

Next we note that the uniqueness of solution for $\left(g^\alpha_\beta(t)\right)$ implies that $\left(g^\alpha_\beta(t)\right)$ forms a one parameter group in G =GL$(n, \mathbb{R})$. From this we see that

$$\begin{aligned} \frac{d}{dt}\left(g^\alpha_\beta(t)\right)|_{t=0} &= -\left(\Gamma^\alpha_{\gamma i}\, g^\gamma_\beta(t)\dot{u}^i(t)\right)|_{t=0} \\ &= -\left(\Gamma^\alpha_{\gamma i}\, \delta^\gamma_\beta(t)\dot{u}^i(0)\right) \\ &= -\left(\Gamma^\alpha_{\beta i}\, \dot{u}^i(0)\right) := -{\omega_\beta}^\alpha(\dot{x}_0) \end{aligned}$$

But the derivative of a one-parameter group at $t = 0$ in a Lie Group is an element of the Lie Algebra of the Lie Group. Thus ${\omega_\beta}^\alpha(\dot{x}_0)$ is an element of the Lie Algebra. For example, if $G = \mathrm{O}(n, \mathbb{R})$ then ${\omega_\beta}^\alpha(\dot{x}_0)$ is in $\mathfrak{o}(n, \mathbb{R})$ which implies that ${\omega_\beta}^\alpha(\dot{x}_0)$ is skew symmetric. Since ${\omega_\beta}^\alpha(\dot{x}_0)$ depends linearly on the tangent vector $\dot{x}_0 \in T_{\dot{x}_0}(M)$ we see that ${\omega_\beta}^\alpha$ is a Lie algebra valued one form. It is important that the connection ${\omega_\beta}^\alpha$ can be found from

$${\omega_\beta}^\alpha(v) = \left(\Gamma^\alpha_{\beta i}\, \dot{u}^i(0)\right) = -\frac{d}{dt}\left(g^\alpha_\beta(t)\right)_{t=0}$$

where the $\left(g^\alpha_\beta\right)$ are found by parallel transfer of $\sigma(0)$ along any curve for which $\dot{x}(0) = v$.

9. ASSOCIATED BUNDLES, Part III

Before we can introduce connections into all the associated bundles it will be convenient to introduce a concept we have neglected up till now. Another way to think of parallel transport in a bundle is that it sets up an isomorphism between $\pi^{-1}[x(0)]$ and $\pi^{-1}[x(t)]$. Specifically, for the principal bundle P we have the isomorphism

$$\begin{aligned} \theta_t : \pi^{-1}[x(0)] &\rightarrow \pi^{-1}[x(t)] \\ \theta_t : \sigma_0 &\mapsto \tilde{\sigma}(t) \end{aligned}$$

If we write this in terms of the local basis of sections $\sigma(t)$ (the frame) we have

$$\theta_t : \sigma_0 \mapsto \sigma(t)g(t)$$

where we are using the same notation as in Part II, where $g(t) = (g^\alpha_\beta(t))$ and $\tilde{\sigma}(t) = \sigma(t)g(t)$ means $\tilde{\sigma}_\beta(t) = \sigma_\alpha(t)g^\alpha_\beta(t)$. We are really most interested in θ_t^{-1} which moves the basis $\sigma(t)$ in $\pi^{-1}[x(t)]$ to the basis in $\pi^{-1}[x(0)]$ which parallel translates into the basis $\sigma(t)$. For convenience let $h(t) = g(t)^{-1}$. Then

$$\theta_t(\sigma_0\, h(t)) = \theta_t(\sigma_0)h(t) = \tilde{\sigma}(t)h(t) = \sigma(t)g(t)h(t) = \sigma(t)$$

so we have

$$\theta_t^{-1}\big(\sigma(t)\big) = \sigma_0\, h(t)$$

All of this works because parallel transport preserves linear combinations of vectors.

In a similar way we can define θ_t on $\mathcal{E}$. Let ρ be a section and let $\rho_0 = \rho(x(0))$ and $\rho(t) = \rho(x(t))$. Let $\tilde{\rho}(t)$ be the parallel transport of ρ_0 along $x(t)$. Define

$$\begin{aligned} \theta_t : \pi_{\mathcal{E}}^{-1}[x(0)] &\rightarrow \pi_{\mathcal{E}}^{-1}[x(t)] \\ \theta_t : \rho_0 &\mapsto \tilde{\rho}(t) = \tilde{\sigma}(t)\rho_0 = \sigma(t)g(t)\rho_0 \end{aligned}$$

as we discussed in Part II.

Again as in Part I it is extremely critical to understand that the isomorphism θ_t ***depends on the path*** $x(t)$ in M. A different path would give a different isomorphism θ_t. So θ_t is path dependent.

Now we want to consider taking the derivative of ρ. This is a special case of a result we will derive later so discussing it now is inefficient, but perhaps illuminating psychologically. The naive formula for the derivative is of course

$$\rho'(0) = \lim_{t\to 0} \frac{\rho(t) - \rho(0)}{t} \qquad\qquad \textit{WRONG}!$$

However, this is nonsense since $\rho(t) \in \pi_{\mathcal{E}}^{-1}[x(t)]$ and $\rho(0) \in \pi_{\mathcal{E}}^{-1}[x(0)]$ and these are different spaces and so the subtraction is impossible. We want to subtract things in the same space, preferably $\pi_{\mathcal{E}}^{-1}[x(0)]$, and one way to do this is to replace $\rho(t)$ by $\theta_t^{-1}(\rho(t))$. So we are going to try this and compute

$$\lim_{t\to 0} \frac{\theta_t^{-1}\rho(t) - \rho(0)}{t}$$

Our first job is to find the derivative of $h(t) = g(t)^{-1}$ we have

$$\begin{aligned} h(t)g(t) &= \mathrm{Id} \\ \frac{dh}{dt}g(t) + h(t)\frac{dg}{dt} &= 0 \\ \frac{dh}{dt} &= -h(t)\frac{dg}{dt}g^{-1}(t) \;=\; -h(t)\frac{dg}{dt}h(t) \end{aligned}$$

We will have need of this below.

Our second job is to compute $\theta_t^{-1}\rho(t)$. Let $\theta_t\rho = \rho(t)$ and $\rho = \sigma_\alpha(0)\rho^\alpha$ and $\rho(t) = \sigma_\alpha(t)\rho^\alpha(t)$. If we set

$$\vec{\rho} = \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} \qquad \text{and} \qquad \vec{\rho}(t) = \begin{pmatrix} \rho^1(t) \\ \vdots \\ \rho^n(t) \end{pmatrix}$$

then we can write more briefly

$$\begin{aligned} \theta_t(\rho) &= \theta_t(\sigma(0)\vec{\rho}) \;=\; \theta_t(\sigma(0))\vec{\rho} \\ &= \tilde{\sigma}(t)\vec{\rho} \;=\; \sigma(t)g(t)\vec{\rho} \end{aligned}$$

using the techniques we developed in Associated Bundles Part II and the fact that θ_t is a linear map and $\vec{\rho}$ is constant. But since

$$\rho(t) = \sigma(t)\vec{\rho}(t)$$

we have

$$\begin{aligned} \sigma(t)g(t)\vec{\rho} \;=\; \rho(t) &= \theta_t(\rho) \;=\; \sigma(t)\vec{\rho}(t) \\ g(t)\vec{\rho} &= \vec{\rho}(t) \\ \vec{\rho} &= h(t)\vec{\rho}(t) \end{aligned}$$

where $h(t)$ is the matrix inverse to $g(t)$. Thus

$$\rho = \sigma(0)\vec{\rho} = \sigma(0)h(t)\vec{\rho}(t)$$

Now we have everything necessary to compute the derivative.

$$\begin{aligned} \lim_{t\to 0} \frac{\theta_t^{-1}\rho(t) - \rho(0)}{t} &= \lim_{t\to 0} \frac{\sigma(0)h(t)\vec{\rho}(t) - \sigma(0)\vec{\rho}(0)}{t} \\ &= \sigma(0)\lim_{t\to 0} \frac{h(t)\vec{\rho}(t) - h(0)\vec{\rho}(0)}{t} \end{aligned}$$

$$
\begin{aligned}
&= \sigma(0)\big[\frac{d}{dt}h(t)\vec{\rho}(t)\big]_{t=0} \\
&= \sigma(0)\big[h(t)\frac{d\vec{\rho}(t)}{dt} + \frac{dh}{dt}\vec{\rho}(t)\big]_{t=0} \\
&= \sigma(0)\big[h(t)\frac{d\vec{\rho}(t)}{dt} - h(t)\frac{dg}{dt}h(t)\vec{\rho}(t)\big]_{t=0} \\
&= \sigma(0)\big[h(t)\frac{d\vec{\rho}(t)}{dt} + h(t)\omega(\dot{x}(t))h(t)\vec{\rho}(t)\big]_{t=0} \\
&= \sigma(0)\big[\frac{d\vec{\rho}(t)}{dt}\Big|_{t=0} + \omega(\dot{x}(0))\vec{\rho}_0\big] \\
&= \sigma(0)\mathcal{D}_{\dot{x}(0)}\rho \\
&= D_{\dot{x}(0)}\rho
\end{aligned}
$$

Thus our new methods have reinvented the covariant derivative, which is very reassuring. Also notice that the critical technical pieces were the knowledge of the connection on P and the use of the map θ_t^{-1}.

With the above as a model we can now guess the proper formula for a connection on any associated bundle. First the setup. Let the bundle be $\mathcal{V}$ with representation space V and representation ρ of G on V. As before we have a curve $x(t)$ and a section

$$v(u^1, \ldots, u^n) = [\sigma(u^1, \ldots, u^d), \vec{v}(u^1, \ldots, u^d)]$$

with values along the curve

$$v(t) = v(u^1(t), \ldots, u^d(t)) = [\sigma(u^1(t), \ldots, u^d(t)), \vec{v}(u^1(t), \ldots, u^d(t))]$$

where $\vec{v} \in V$.

We define parallel transport of $v(0)$ along $x(t)$ in the obvious way. Recall that $\sigma(t)$ is the "original basis of sections" over U and $\tilde{\sigma}(t)$ is the result of parallel transport of $\sigma(0)$ along the path $u(t)$.

Def The parallel transport of $v(0) \in \pi_{\mathcal{V}}^{-1}[x(0]$ along $x(t)$ is

$$v(t) = [\tilde{\sigma}(t), \vec{v}(0)] = [\sigma(t)g(t), \vec{v}(0)] = [\sigma(t), \rho\big(g(t)\big)(\vec{v}(0)]$$

Then we have the usual parallel transport functions

$$
\begin{aligned}
\theta_t[\sigma(0), \vec{v}(0)] &\mapsto [\sigma(t), \rho\big(g(t)\big)(\vec{v}(0)] \\
\theta_t^{-1}[\sigma(t), \vec{v}(t)] &\mapsto [\sigma(0), \rho\big(h(t)\big)(\vec{v}(t)]
\end{aligned}
$$

where $h(t) = g(t)^{-1}$. Now we define

Def The *covariant derivative* of a section $v(u^1, \ldots, u^d) = [\sigma(u^1, \ldots, u^d), \vec{v}(u^1, \ldots, u^d)$ of the vector bundle $\mathcal{V}$ at $v(0)$ is

$$D_{\dot{x}(0)}v = \lim_{t\to 0}\frac{\theta_t^{-1}(v(t)) - v(0)}{t} = \Big[\sigma(0), \frac{d}{dt}\big[\rho\big(h(t)\big)\vec{v}(t)\big]_{t=0}\Big]$$

In detail,

$$\begin{aligned}
\lim_{t\to 0}\frac{\theta_t^{-1}(v(t))-v(0)}{t} &= \lim_{t\to 0}\frac{[\sigma(0),\rho\big(h(t)\big)(\vec{v}(t))]-[\sigma(0),\rho\big(h(0)\big)(\vec{v}(0))]}{t}\\
&= \Big[\sigma(0),\lim_{t\to 0},\frac{\rho\big(h(t)\big)(\vec{v}(t))-\rho\big(h(0)\big)(\vec{v}(0))]}{t}\Big]\\
&= \Big[\sigma(0),\frac{d}{dt}\big[\rho(h(t))\vec{v}(t)\big]_{t=0}\Big]
\end{aligned}$$

We are now in the enviable position of having defined parallel transport in any associated bundle, and we have a formula for the covariant derivative of sections of any associated bundle. The formula will be effective if we have sufficient control of the representation ρ, as we will illustrate in the following section, where we find the covariant derivatives in various popular bundles. We will also investigate some theoretical considerations involving tensor products which are important in the sequel.

10. EXAMPLES WITH TENSOR PRODUCTS

In this section we are going to calculate covariant derivatives using the technology of the previous section for some of out favourite bundles. Let ρ be a representation of the group G of P on some vector space V and $\mathcal{V}$ the corresponding associated bundle, and let $x(t)$ be a curve on the base space M which we may take as being in the coordinate neighbourhood U. The notations and formulas we will use are as follows:

$$\begin{aligned}
g(t) &= g(u(t))\\
h(t) &= g(t)^{-1}\\
\frac{dh}{dt} &= -h(t)\frac{dg}{dt}\,h(t)\\
\frac{dg}{dt} &= -\omega_\beta{}^\alpha(\dot{x}(t))\\
D_{\dot{x}(0)}v &= \big[\sigma(0),\frac{d}{dt}\left(\rho(h(t))\vec{v}(t)|_{t=0}\right)\big]
\end{aligned}$$

where $v=[\sigma,\vec{v}]\in P\times_\rho V$. These formulas are intended to be generic for $\mathcal{V}$ the associated bundle for V, the representation space; the specific shapes of the formulas will vary slightly depending on the situation. Also, $v\in\Gamma(\mathcal{V})$ and $\vec{v}$ the representation of v in a basis. It is annoying that the notational details vary slightly from example to example, but the cure for this would be worse than the disease.

The Bundle $\mathcal{E}$

We begin with the simplest associated bundle. If P is the bundle of frames of a bundle $\mathcal{E}$ the we know $\mathcal{E}$ from P using $V=\mathbb{R}^n$ and $\rho(g)v=g\vec{v}$. We will now show how the coveriant derivative and the connection on $\mathcal{E}$ conforms to our previous formulas. Let $({\omega_\beta}^\alpha) = (\Gamma^\alpha_{\beta i} du^i)$ be the connections one forms on P for the local basis of sections $\sigma = (\sigma_1, \ldots, \sigma_n)$. We regard $\mathcal{E}$ as

$$\mathcal{E} = \mathcal{P} \times_\rho \mathbb{R}^\backslash$$

with typical element

$$v = [\sigma, \vec{v}] \qquad \text{where} \qquad \vec{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

Now let $x(t)$ be a curve in $U \subseteq M$ and $\dot{x}_0 = \dot{x}(0)$. Then

$$\begin{aligned}
D_{\dot{x}_0} v &= [\sigma(0), \frac{d}{dt}\big(\rho(h(t))(\vec{v}(t))\big)\big|_{t=0}] \\
&= [\sigma(0), \frac{d}{dt}\big(h(t)\vec{v}(t)\big)\big|_{t=0}] \\
&= [\sigma(0), \big(h(t)\frac{d\vec{v}(t)}{dt} + \frac{dh}{dt}\,\vec{v}(t)\big)\big|_{t=0}] \\
&= [\sigma(0), \big(h(t)\frac{d\vec{v}(t)}{dt} - h(t)\frac{dg}{dt}\,h(t)\vec{v}(t)\big)\big|_{t=0}] \\
&= [\sigma(0), \big(h(t)\frac{d\vec{v}(t)}{dt} + h(t)\omega(x(t))h(t)\vec{v}(t)\big)\big|_{t=0}] \\
&= [\sigma(0), \frac{d\vec{v}(t)}{dt}\big|_{t=0} + \omega(x_0)\vec{v}(0)]
\end{aligned}$$

Since there is nothing special about $t = 0$ we can translate this into more normal notation for any T as

$$D_{\dot{x}(t)} v = \sigma_\alpha(t)\Big(\frac{dv^\alpha}{dt} + \Gamma^\alpha_{\beta j}\frac{du^j}{dt}\,v^\beta(t)\Big)$$

Since for any $w \in T_x(M)$ we can find a curve with $\dot{x}(0) = w$ we can write the above formula (again with $t = 0$) as

$$\begin{aligned}
D_w v &= \sigma_\alpha\Big(\frac{\partial v^\alpha}{\partial u^j}\frac{du^j}{dt} + \Gamma^\alpha_{\beta j} w^j v^\beta(0)\Big) \\
&= \sigma_\alpha\Big(\frac{\partial v^\alpha}{\partial u^j} + \Gamma^\alpha_{\beta j} v^\beta(0)\Big) w^j
\end{aligned}$$

and thus the covariant differnetial is

$$\begin{aligned}
Dv &= \sigma_\alpha\Big(\frac{\partial v^\alpha}{\partial u^j} du^j + \Gamma^\alpha_{\beta j} v^\beta du^j\Big) \\
&= \sigma_\alpha\Big(dv^\alpha + {\omega_\beta}^\alpha v^\beta\Big) \\
&= \sigma_\alpha \mathcal{D} v^\alpha
\end{aligned}$$

which we recognize as our old formula for the covariant derivative. We also want to make contiact with the ancient notation. We set, as usual

$$v^\alpha_{|j} = \frac{\partial v^\alpha}{\partial u^j} + \Gamma^\alpha_{\beta j} b^\beta$$

and then

$$\begin{aligned} Dv &= \sigma_\alpha v^\alpha_{|j} du^j \\ D_{\dot{x}(t)} v &= \sigma_\alpha v^\alpha_{|j} \frac{du^j}{dt} \\ D_w v &= \sigma_\alpha v^\alpha_{|j} w^j \end{aligned}$$

The Bundle $\mathcal{E}^*$

We obtain the bundle $\mathcal{E}^*$ by choosing $V = \mathbb{R}^{n*}$ and $\rho(g)(\lambda_1, \ldots, \lambda_n) = (\lambda_1, \ldots, \lambda_n) g^{-1}$. (Note this is a left action on V.) As before we will use the notation $\underline{\lambda} = (\lambda_1, \ldots, \lambda_n)$ so the above equation becomes $\rho(g)\underline{\lambda} = \underline{\lambda} g^{-1}$. Then, using our usual formulas,

$$\begin{aligned} D_{\dot{x}_0} \lambda &= [\sigma(0), \frac{d}{dt}(\rho(h(t)))\underline{\lambda}(t)\big|_{t=0}] \\ &= [\sigma(0), \frac{d}{dt}(\underline{\lambda}(t) h(t)^{-1}\big|_{t=0}] \\ &= [\sigma(0), \frac{d}{dt}(\underline{\lambda}(t) g(t)\big|_{t=0}] \\ &= [\sigma(0), \frac{d\underline{\lambda}}{dt} g(t)\big|_{t=0} + \underline{\lambda}(t) \frac{dg}{dt}\big|_{t=0}] \\ &= [\sigma(0), \frac{d\underline{\lambda}}{dt} g(t)\big|_{t=0} - \underline{\lambda}(t)({\omega_\beta}^\alpha(\dot{x}(t))\big|_{t=0}] \\ &= [\sigma(0), \frac{d\underline{\lambda}}{dt}\big|_{t=0} - \underline{\lambda}(0)({\omega_\beta}^\alpha(\dot{x}(0))] \end{aligned}$$

Since there is nothing special about $t = 0$ we may replace the 0 by Y and then switch to a more normal notation. We have then,

$$\begin{aligned} D_{\dot{x}_0} \lambda &= \sigma^\beta \Big(\frac{\partial \lambda_\beta}{\partial u^j} - \lambda_\alpha \Gamma^\alpha_{\beta j} \Big) \frac{du^j}{dt} \\ D_w \lambda &= \sigma^\beta \Big(\frac{\partial \lambda_\beta}{\partial u^j} - \lambda_\alpha \Gamma^\alpha_{\beta j} \Big) w^j \\ D\lambda &= \sigma^\beta \Big(\frac{\partial \lambda_\beta}{\partial u^j} - \lambda_\alpha \Gamma^\alpha_{\beta j} \Big) du^j \end{aligned}$$

and setting

$$\lambda_{\beta|j} = \frac{\partial \lambda_\beta}{\partial u^j} - \lambda_\alpha \Gamma^\alpha_{\beta j}$$

we have

$$D\lambda = \sigma^\beta \lambda_{\beta|j} du^j$$

which is the covariant differential.

Linear Transformations: the bundle $\mathcal{E}\otimes\mathcal{E}^*$

For the space V we choose $M(n,\mathbb{R})$ and, for the representation for $a\in M(n,\mathbb{R})$ we choose $\rho(g)a = gag^{-1}$. (For reasons connected with Lie Groups this is often referred to as the *adjoint representation.*) Let the Linear Transformation T be represented by the matrix a in the basis σ so that

$$T = [\sigma, a]$$

then in the basis $\tilde{\sigma} = \sigma g$ we will have the representation matrix $b = g^{-1}ag$. In the associated bundle notation we have

$$[\sigma, a] = [\sigma g g^{-1}, a] = [\sigma g, \rho(g^{-1})a] = [\sigma, g^{-1}ag] = [\tilde{\sigma}, b]$$

as we expected.

By our general forula, (with $h = g^{-1}$)

$$\begin{aligned}
D_{\dot{x}(0)} &= [\sigma, \frac{d}{dt}\big(\rho(h(t))a(t)\big|_{t=0}\big)] \\
&= [\sigma, \frac{d}{dt}\big(h(t)a(t)h^{-1}(t)\big|_{t=0}\big)] \\
&= [\sigma, \frac{d}{dt}\big(h(t)a(t)g(t)\big|_{t=0}\big)] \\
&= [\sigma, \big(h(t)\frac{da}{dt}\,g(t) + \frac{dh}{dt}a(t)g(t) + h(t)a(t)\frac{dg}{dt}\big)\big|_{t=0}] \\
&= [\sigma, \big(h(t)\frac{da}{dt}\,g(t) - h(t)\frac{dg}{dt}h(t)a(t)g(t) + h(t)a(t)\frac{dg}{dt}\big)\big|_{t=0}] \\
&= [\sigma, \big(h(t)\frac{da}{dt}\,g(t) + h(t)\omega(x(t))h(t)a(t)g(t) - h(t)a(t))\omega(x(t))\big)\big|_{t=0}] \\
&= [\sigma, \frac{da}{dt}\big|_{t=0} + \omega(x(0))a(t) - a(t)\omega(x(0))]
\end{aligned}$$

Decoding and replacing $\dot{x}(0)$ by $\dot{x}(t)$ since there is nothing special about $t=0$, we find the representation for $D_{\dot{x}(t)}$ to be

$$\begin{aligned}
D_{\dot{x}(t)} &\stackrel{\sigma}{\longleftrightarrow} \frac{da}{dt} + \omega(x(t))a(t) - a(t)\omega(x(t)) \\
&\stackrel{\sigma}{\longleftrightarrow} \frac{\partial a^\alpha_\beta}{\partial u^i}\frac{du^i}{dt} + \Gamma^\alpha_{\gamma i}a^\gamma_\beta\frac{du^i}{dt} - \Gamma^\gamma_{\beta i}a^\alpha_\gamma\frac{du^i}{dt} \\
&\stackrel{\sigma}{\longleftrightarrow} \Big(\frac{\partial a^\alpha_\beta}{\partial u^i} + \Gamma^\alpha_{\gamma i}a^\gamma_\beta - \Gamma^\gamma_{\beta i}a^\alpha_\gamma\Big)\frac{du^i}{dt}
\end{aligned}$$

Then, setting as the ancients did,

$$a^\alpha_{\beta|i} = \frac{\partial a^\alpha_\beta}{\partial u^i} + \Gamma^\alpha_{\gamma i}a^\gamma_\beta - \Gamma^\gamma_{\beta i}a^\alpha_\gamma$$

we have

$$
\begin{aligned}
DT &\overset{\sigma}{\longleftrightarrow} a^{\alpha}_{\beta|i} du^i \\
DT &= [\sigma,\ a^{\alpha}_{\beta|i} du^i]
\end{aligned}
$$

The equation above can be written in another way but since it uses 1-forms ω we must multiply by du^i:

$$
\begin{aligned}
a^{\alpha}_{\beta|i} du^i &= \frac{\partial a^{\alpha}_{\beta}}{\partial u^i} du^i + \Gamma^{\alpha}_{\gamma i} du^i a^{\gamma}_{\beta} - a^{\alpha}_{\gamma} \Gamma^{\gamma}_{\beta i} du^i \\
&= da^{\alpha}_{\beta} + \omega_{\gamma}{}^{\alpha} \wedge a^{\gamma}_{\beta} - a^{\alpha}_{\gamma} \wedge \omega_{\beta}{}^{\gamma} \\
\mathcal{D}a &= da + \omega \wedge a - a \wedge \omega \\
&= da + [\omega, a]
\end{aligned}
$$

Note this is $\mathcal{D}$ not D. The wedges are strictly speaking not necessary since a is a matrix of 0-forms (functions on U) but are included to make this formula look like others of its type.

Bilinear Forms

This example is more difficult than the preceding ones because it is not convenient here to use matrix technology. I could force the use of such technology for the specific case of bilinear forms (using the transpose of a matrix) but I want to claim that the method used here will work for any tensor, and for the more general case the transpose will not work, so best to do without it. Keep in mind this claim to generality as we march through the somewhat unpleasant proof. This is a proof by example of the general case.

We will use the notation $B[u, v] \in \mathcal{E}^* \otimes \mathcal{E}^*$ for this treatment and we can take $V = \mathbb{R}^{n*} \otimes \mathbb{R}^{n*} \simeq M(n, \mathcal{R})$ but here it is natural to write the matrix with low indices. The representation will be the tensor product of the previously introduced ρ for $\mathcal{E}^*$ with itself, although we will allow ourselves some flexibility in the notation. We could follow the same trail we trod with the linear transformations but I want to use a related technique because it is more readily generalizable to more complicated tensor situations.

The space of bilinear forms can be visualized as $\mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$. Here is how we do this. First we define an action of $\mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$ on $\mathbb{R}^n \otimes \mathbb{R}^n$ by

$$
(\underline{\lambda} \otimes \underline{\mu})(\vec{v} \otimes \vec{w} = \underline{\lambda}(\vec{v})\underline{\mu}(\vec{w})
$$

Next we define, with local basis of sections $\sigma = (\sigma_1, \ldots, \sigma_n)$,

$$
\begin{aligned}
b_{\alpha\beta} &= B[\sigma_\alpha, \sigma_\beta] \\
\underline{b}_\alpha &= (b_{\alpha 1}, \ldots, b_{\alpha n}) \\
\underline{\sigma}_\alpha &= (0, \ldots, 0, 1, 0, \ldots, 0) \qquad 1 \text{ in } \alpha^{\text{th}} \text{ place}
\end{aligned}
$$

Note that with $u = [\sigma, \vec{u}$ and other previously established conventions we have

$$
\begin{aligned}
B[u, v] &= B[\sigma_\alpha u^\alpha, \sigma_\beta v^\beta] \\
&= B[\sigma_\alpha, \sigma_\beta] u^\alpha v^\beta \\
&= b_{\alpha\beta} u^\alpha v^\beta
\end{aligned}
$$

Then we can represent $B[u,v]$ in the form

$$b = \sum_\alpha \underline{\sigma}_\alpha \otimes \underline{\beta}_\alpha \in \mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$$

Using the action of $\mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$ on $\mathbb{R}^n \otimes \mathbb{R}^n$ defined above we have

$$\begin{aligned} b(\vec{u},\vec{v}) &= \sum_\alpha \underline{\sigma}_\alpha \otimes \underline{\beta}_\alpha(\vec{u},\vec{v}) \\ &= \sum_\alpha \underline{\sigma}_\alpha(\vec{u})\underline{\beta}_\alpha(\vec{v}) \\ &= \sum_\alpha u^\alpha \big(\sum_\beta b_{\alpha\beta} v^\beta\big) \\ &= \sum_{\alpha\beta} b_{\alpha\beta} u^\alpha v^\beta \\ &= B[u,v] \end{aligned}$$

Now we set things up for application of the general formula for covariant derivative. Recall that $\rho_2 : \mathrm{GL}(n,\mathbb{R}) \to \mathrm{Aut}(\mathbb{R}^{n*})$ is given by

$$\rho_2(g)(\lambda_1,\ldots,\lambda_n) = (\lambda_1,\ldots,\lambda_n)g^{-1} = \underline{\lambda} g^{-1}$$

and we can define $\rho : \mathrm{GL}(n,\mathbb{R}) \to \mathrm{Aut}(\mathbb{R}^{n*} \otimes \mathbb{R}^{n*})$ by

$$\begin{aligned} \rho &= \rho_2 \otimes \rho_2 \\ \rho(g)(\underline{\lambda} \otimes \underline{\mu}) &= \rho_2(g)(\underline{\lambda}) \otimes \rho_2(g)(\underline{\mu}) \\ &= \underline{\lambda} g^{-1} \otimes \underline{\mu} g^{-1} \end{aligned}$$

Now using the formula

$$\begin{aligned} D_{\dot{x}_0} &= [\sigma(0), \frac{d}{dt}\,(\rho(h(t))\Big(\sum_\alpha \underline{\sigma}_\alpha \otimes \underline{\beta}_\alpha(t))\Big)\big|_{t=0}] \\ &= [\sigma(0), \frac{d}{dt}\,(\Big(\sum_\alpha \underline{\sigma}_\alpha h(t)^{-1} \otimes \underline{\beta}_\alpha(t))h(t)^{-1}\Big)\big|_{t=0}] \\ &= [\sigma(0), \frac{d}{dt}\,(\Big(\sum_\alpha \underline{\sigma}_\alpha g(t) \otimes \underline{\beta}_\alpha(t))g(t)\Big)\big|_{t=0}] \\ &= [\sigma(0), \sum_\alpha \Big((\frac{d\underline{\sigma}_\alpha}{dt}\,g(t) + \underline{\sigma}_\alpha \frac{dg(t)}{dt}) \otimes \underline{\beta}_\alpha(t))g(t) \\ &+ \underline{\sigma}_\alpha g(t) \otimes (\frac{d\underline{\beta}_\alpha}{dt}\,g(t) + \underline{\beta}_\alpha \frac{dg(t)}{dt})\Big)\big|_{t=0}] \end{aligned}$$

Using the fact that $\underline{\sigma}_\alpha$ is a constant vector so $\frac{d}{dt}\underline{\sigma}_\alpha = 0$, we can eliminate one term. We also have $\frac{dg}{dt} = -\omega(x(t))$. Keeping in mind the $\underline{\beta}_\alpha$ is a row vector and ω is a matrix we have

$$\begin{aligned}
D_{\dot{x}_0} &= [\sigma(0), \sum_\alpha \Big(\underline{\sigma}_\alpha \frac{dg(t)}{dt} \otimes \underline{\beta}_\alpha(t) g(t) + \underline{\sigma}_\alpha g(t) \otimes (\frac{d\underline{\beta}_\alpha}{dt}\, g(t) + \underline{\beta}_\alpha \frac{dg(t)}{dt})\Big)\big|_{t=0}] \\
&= [\sigma(0), \sum_\alpha \Big(- \underline{\sigma}_\alpha \omega(x(t)) \otimes \underline{\beta}_\alpha(t) g(t) + \underline{\sigma}_\alpha g(t) \otimes (\frac{d\underline{\beta}_\alpha}{dt}\, g(t) - \underline{\beta}_\alpha \omega(x(t))\Big)\big|_{t=0}] \\
&= [\sigma(0), \sum_\alpha \Big(- \underline{\sigma}_\alpha \omega(x(0)) \otimes \underline{\beta}_\alpha(0) + \underline{\sigma}_\alpha \otimes (\frac{d\underline{\beta}_\alpha}{dt} - \underline{\beta}_\alpha \omega(x(0))\Big)] \\
&= [\sigma(0), \sum_\alpha \Big(\underline{\sigma}_\alpha \otimes \frac{d\underline{\beta}_\alpha}{dt} - \underline{\sigma}_\alpha \omega(x_0) \otimes \underline{\beta}_\alpha - \underline{\sigma}_\alpha \otimes \underline{\beta}_\alpha \omega(x_0)\Big)]
\end{aligned}$$

the first and third terms are no problem, but the second term is, because it does not match up well with matrix calculations. Hence we must do some unpleasant footwork. First we transform the term a bit. We will use ω for $\omega(x_0)$

$$\begin{aligned}
\sum_\gamma \underline{\sigma}_\gamma \omega \otimes \underline{\beta}_\gamma &= \sum_\gamma (0,\dots,0,1,0,\dots,0)\omega \otimes (b_{\gamma 1},\dots,b_{\gamma n}) \\
&= \sum_\gamma {\omega_\beta}^\gamma \underline{\sigma}_\beta \otimes (b_{\gamma 1},\dots,b_{\gamma n}) \\
&= \sum_\gamma \underline{\sigma}_\beta \otimes ({\omega_\beta}^\gamma b_{\gamma 1},\dots,\omega_\beta^\gamma b_{\gamma n})
\end{aligned}$$

In the same way that

$$\begin{aligned}
(b_{\alpha\beta}) &\stackrel{\sigma}{\longleftrightarrow} \sum_\alpha \underline{\sigma}_\alpha \otimes \underline{\beta}_\alpha \\
&\stackrel{\sigma}{\longleftrightarrow} \sum_\alpha \underline{\sigma}_\alpha \otimes (b_{\alpha 1},\dots,b_{\alpha n})
\end{aligned}$$

we have

$$({\omega_\alpha}^\gamma)(b_{\gamma\beta}) \stackrel{\sigma}{\longleftrightarrow} \sum_\alpha \underline{\sigma}_\alpha \otimes ({\omega_\alpha}^\gamma b_{\gamma 1},\dots,{\omega_\alpha}^\gamma b_{\gamma n})$$

Thus the second term referree to above is

$$[\sigma(0), \sum_\alpha \underline{\sigma}_\alpha \otimes ({\omega_\alpha}^\gamma b_{\gamma 1}),\dots,{\omega_\alpha}^\gamma b_{\gamma n})] = [\sigma(0), -({\omega_\alpha}^\gamma b_{\gamma n})]$$

Thus if we write out $D_{\dot{x}_0}$ in matrix form and give ${\omega_\alpha}^\gamma$ its argument back, we get

$$\begin{aligned}
D_{\dot{x}_0} b &= [\sigma_0, \Big(\frac{db_{\alpha\beta}}{dt} - {\omega_\alpha}^\gamma(x_0) b_{\gamma\beta} - b_{\alpha\gamma}{\omega_\beta}^\gamma(x_0)\Big)] \\
&= [\sigma_0, \Big(\frac{\partial b_{\alpha\beta}}{\partial u^j}\frac{du^j}{dt} - \Gamma^\gamma_{\alpha j} b_{\gamma\beta}\frac{du^j}{dt} - b_{\alpha\gamma}\Gamma^\gamma_{\beta j}\frac{du^j}{dt}\Big)] \\
&= [\sigma_0, (b_{\alpha\beta|j}\frac{du^j}{dt})]
\end{aligned}$$

where

$$b_{\alpha\beta|j} = \frac{\partial b_{\alpha\beta}}{\partial u^j} - \Gamma^\gamma_{\alpha j} b_{\gamma\beta} - \Gamma^\gamma_{\beta j} b_{\alpha\gamma} \qquad \text{classical formula}$$

and thus

$$Db = b_{\alpha\beta|j} du^j$$

I hope it is now clear that the preceding caluclations are general enough so that we can see the general pattern. That is

$$\begin{aligned} \text{if} \quad T &= [\sigma_0,\, T^{\alpha\beta\ldots\delta}_{\rho\sigma\ldots\upsilon} \sigma_\alpha \otimes \sigma_\beta \otimes \ldots \otimes \delta \otimes \sigma^\rho \otimes \sigma^\sigma \otimes \ldots \otimes \sigma^\upsilon] \\ \text{then} \quad DT &= [\sigma_0,\, T^{\alpha\beta\ldots\delta}_{\rho\sigma\ldots\upsilon|j} \sigma_\alpha \otimes \sigma_\beta \otimes \ldots \otimes \delta \otimes \sigma^\rho \otimes \sigma^\sigma \otimes \ldots \otimes \sigma^\upsilon \otimes du^j] \end{aligned}$$

where

$$T^{\alpha\beta\ldots\delta}_{\rho\sigma\ldots\upsilon|j} = \frac{\partial T^{\alpha\beta\ldots\delta}_{\rho\sigma\ldots\upsilon}}{\partial u^j} + \Gamma^\alpha_{\theta j} T^{\theta\beta\ldots\delta}_{\rho\sigma\ldots\upsilon} + \ldots + \Gamma^\delta_{\theta j} T^{\alpha\beta\ldots\theta}_{\rho\sigma\ldots\upsilon} - \Gamma^\theta_{\rho j} T^{\alpha\beta\ldots\delta}_{\theta\sigma\ldots\upsilon} - \ldots - \Gamma^\theta_{\upsilon j} T^{\alpha\beta\ldots\delta}_{\rho\sigma\ldots\theta}$$

Thus we have derived the formulas for classical tensor calculus.

The "bundle" of functions

It is interesting to see how our technology recaptures the ordinary directional derivative of a function on a manifold. Nobody would doubt that in the case of functions $Df = df$ but we shall look at this a little more closely than necessary just for fun. If you don't think it's fun, skip it.

We don't usually think of the functions on a manifold as being an associated bundle but they can be looked at that way, and we need to think of them that way if we want to use our general formula. Thus let $f : U \to \mathbb{R}$. We need a vector space V and a representation ρ :$\mathrm{GL}(n, \mathbb{R}) \to \mathrm{Aut}(V)$ to set up the associated bundle. Thus we take $V = \mathbb{R}$ and ρ to have the output 1 for any input. We then identify the function f with $[\sigma, f]$ for any function f over $U \subseteq M$ and any local basis of sections σ over U. Thus

$$f = [\sigma,\, f] \qquad \text{and} f(x) = [\sigma(x), f(x)]$$

Now we change the basis: $\tilde{\sigma} = \sigma g$. Then

$$[\tilde{\sigma},\, f] = [\sigma g,\, f] = [\sigma,\, \rho(g) f] = [\sigma,\, f]$$

so all is consistent. Note the change of local basis is without effect as it should be. Let $x(t)$ be a curve in U. Setting, for convenience, $f(t) = f(x(t))$, we have

$$\begin{aligned} \theta_t[\sigma(0), f(0)] &= [\tilde{\sigma}(t), f(0) \\ &= [\sigma(t) g,\, f(0)] \\ &= [\sigma(t),\, \rho(g) f] \\ &= [\sigma(t),\, f] \end{aligned}$$

So

$$\theta_t[\sigma(0),\ f(t] = [\sigma(t),\ f(t)$$

and we have for the curve $x(t)$

$$\begin{aligned} D_{\dot{x}_0} f &= \lim_{t\to 0} \frac{\theta_t^{-1}[\sigma(t),\ f(t)] - [\sigma(0), f(0)]}{t} \\ &= \lim_{t\to 0} \frac{[\sigma(0),\ f(t)] - [\sigma(0), f(0)]}{t} \\ &= [\sigma(0),\ \frac{d}{dt} f(x(t))\big|_{t=0}] \\ &= [\sigma(0),\ \frac{\partial f}{\partial u^i} \frac{du^i}{dt} \big|_{t=0}] \\ &= [\sigma(0),\ \frac{\partial f}{\partial u^i} du^i(\dot{x}_0)] \\ &= [\sigma(0),\ df(\dot{x}_0)] \end{aligned}$$

So, unwinding the identification, we have

$$\begin{aligned} D_{\dot{x}_0} f &= df(\dot{x}_0) \\ Df &= df \end{aligned}$$

which is the result we wished to establish.

It is important to realize here that we are taking covariant derivatives of real valued functions on M. Later we will consider functions on P and then formula $Df = df$ will not hold for this more general case. If the function f is constant on the fibres we will of course be fack to the original case and the formula $Df = df$ will again hold.

11 TENSORS OF TYPE Ad AND THE CURVATURE TENSOR

As we saw, linear transformations live in $\mathcal{E} \otimes \mathcal{E}^*$ and are tensors that use the representation $\rho(g)a = gag^{-1}$ for $g \in \mathrm{GL}(n, \mathbb{R})$. Because this is the formula for the Adjoint (left) Action of a Matrix Lie Group on its Lie Algebra, tensors associated with this action are called *tensors of type* Ad. Recall that tensors of type Ad form an associated tensor bundle of the Principal Bundle P with the formula

$$[\sigma, a] = [\sigma g g^{-1}, a] = [\sigma g, \rho(g^{-1})a] = [\sigma g, g^{-1} a g]$$

The most important of these tensors is the curvature tensor Ω. (the curvature tensor Ω is acutally an element of $\mathcal{E} \otimes \mathcal{E}^* \otimes \mathcal{T}(\mathcal{M}) \otimes \mathcal{T}(\mathcal{M})^*$ but we are interested at this point only in the $\mathcal{E} \otimes \mathcal{E}^*$ part of it. Recall that

$$\Omega = \frac{1}{2} R_{\alpha\ \ kl}^{\ \beta} \, e_\beta \otimes e^\alpha \otimes du^k \wedge du^l$$

$$
\begin{aligned}
&= \sum_{\alpha,\beta,(k<l)} R_{\alpha\ kl}^{\ \beta}\, e_\beta \otimes e^\alpha \otimes du^k \wedge du^l \\
&= d\omega + \omega \wedge \omega
\end{aligned}
$$

where

$$\omega = (\omega_\alpha^{\ \beta}) = (\Gamma^\beta_{\alpha k} du^k)$$

and

$$R_{\alpha\ kl}^{\ \beta} = \frac{\partial \Gamma^\beta_{\alpha l}}{\partial u^k} - \frac{\partial \Gamma^\beta_{\alpha k}}{\partial u^l} + \Gamma^\beta_{\gamma k}\Gamma^\gamma_{\alpha l} - \Gamma^\beta_{\gamma l}\Gamma^\gamma_{\alpha k}$$

We will not use the last two formulas in our development; they are only provided to remind the user of the background.

We wish to show that Ω is a tensor of type Ad. We could certaihnly do this using the previous material, but this is a convenient moment to do a little review and recapitulation and compress a lot of the previous theory insto a small space. We also want to show how the desire to have the curvature *be* a tensor gives a big hint for what its formula should be.

As part of our review we will again derive the change of basis formula for ω. For this we will use only

$$D\sigma = \sigma\omega \qquad \text{and} \qquad D(fv) = df\, v + f Dv$$

for $\sigma = (\sigma_1, \ldots, \sigma_n) \in P$, (a local basis of sections), v a section over $U \subseteq M$ and f a function on M. Then we have, for $\sigma' = \sigma g$,

$$
\begin{aligned}
\sigma'\omega' &= D\sigma' \;=\; D(\sigma g) \;=\; \sigma dg + (D\sigma)g \\
\sigma g\omega' &= \sigma dg + (\sigma\omega)g \;=\; \sigma(dg + \omega g) \\
g\omega' &= dg + \omega g \\
\omega' &= g^{-1}dg + g^{-1}\omega g \;=\; \omega_G + Ad(g^{-1})\omega
\end{aligned}
$$

where ω_G is the fundamental left invariant 1-form on G and Ad is the adjoint action of G on it's Lie Algebra $\mathfrak{g}$ introduced above. For the remainder of this section we will use only the formula

$$\omega' = g^{-1}dg + g^{-1}\omega g$$

relating the ω' for σ' to the ω for $sigma$. It is interesting how much can be pulled out of this formula and illustrates how important the formulas for change of basis can be. Let's rewrite the formula in the forms

$$g\omega' = dg + \omega g \qquad \text{and} \qquad dg = g\omega' - \omega g$$

Remember this is matrix multiplication! Recall that for a wedge product of the forms $\omega \in \Lambda^r$ and $\eta \in \Lambda^s$ we have

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^r \omega \wedge d\eta$$

Recall that if f is a function ($f \in \Lambda^0$) then $f \wedge \omega = f\omega$ by definition. Functions commute with forms, $f\omega = \omega f$ but this is *not* true for matrices of functions and matrices of forms, where order must be strictly preserved. On the other hand if we have $(\omega g) \wedge \omega' = (\omega \wedge g) \wedge \omega'$ then since both matrix multiplication and wedge product are associative, this will be equal to $\omega \wedge (g \wedge \omega') = \omega \wedge (g\omega')$. This is important below.

We are now going to show that the curvature tensory $\Omega = d\omega + \omega \wedge \omega$ is a tensor of type Ad using only the transformation formula for ω. We have first

$$\begin{aligned} d(g\omega') &= dg \wedge \omega' + (-1)^0 g \wedge d\omega' \\ &= dg \wedge \omega' + g d\omega' \end{aligned}$$

and second

$$\begin{aligned} g\omega' &= dg + \omega \wedge g \\ d(g\omega') &= ddg + d\omega \wedge g + (-1)^1 \omega \wedge dg \\ &= 0 + d\omega \wedge g - \omega \wedge dg \end{aligned}$$

Note the two equations have different signs on the second term which is critical for the derivation. Now we can equate the two formulas for $d(g\omega')$

$$dg \wedge \omega' + g \wedge d\omega' = d\omega \wedge g - \omega \wedge dg$$

Into this we insert the formula $dg = g \wedge \omega' - \omega \wedge g$ (where we have used $\omega g = \omega \wedge g$) and we get

$$\begin{aligned} (g \wedge \omega' - \omega \wedge g) \wedge \omega' + g \wedge d\omega' &= d\omega \wedge g - \omega \wedge (g \wedge \omega' - \omega \wedge g) \\ g \wedge \omega' \wedge \omega' - \omega \wedge g \wedge \omega' + g \wedge d\omega' &= d\omega \wedge g - \omega \wedge g \wedge \omega' + \omega \wedge \omega \wedge g \end{aligned}$$

We drop the term $-\omega \wedge g \wedge \omega'$ which is common to both sides and rearrange to get

$$\begin{aligned} g \wedge (d\omega' + \omega' \wedge \omega') &= (d\omega + \omega \wedge \omega) \wedge g \\ d\omega' + \omega' \wedge \omega' &= g^{-1}(d\omega + \omega \wedge \omega) g \end{aligned}$$

so that $d\omega + \omega \wedge \omega$ is a tensor of type Ad. This suggests that $d\omega + \omega \wedge \omega$ might be something important, and of course we know it is the curvature tensor Ω. The point is that one of the ways of knowing *what* is important is keeping an eye on the transformation properties. Notice that we derived the transformation properties of $d\omega + \omega \wedge \omega$ solely from the formula $\omega' = g^{-1}dg + g^{-1}\omega g$ (which is *not* tensorial).

So we have finally

$$\Omega' = g^{-1}\Omega g = \mathrm{Ad}(g^{-1})\Omega$$

which shows that Ω is a tensor of type Ad, as promised.

We will now develop some theory for the curvature tensor Ω. We will begin with $D\sigma = \sigma\omega$. Then by the definition of $\mathcal{D}$ we have

$$D(D\sigma) = D(\sigma\omega) = \sigma\mathcal{D}(\omega)$$

But we have

$$\begin{aligned} D(D\sigma) &= D(\sigma\omega) \;=\; \sigma d\omega + D\sigma \wedge \omega \\ &= \sigma d\omega + (\sigma\omega) \wedge \omega \;=\; \sigma(d\omega + \omega \wedge \omega) \end{aligned}$$

Comparing the two expressions for $D(D\sigma)$ we see

$$\sigma\mathcal{D}(\omega) = D^2\sigma = \sigma(d\omega + \omega \wedge \omega) = \sigma\Omega$$

amd thus

$$\mathcal{D}(\omega) = d\omega + \omega \wedge \omega = \Omega$$

None of these formulas are new for us, but we have derived them very efficiently here. We also have

$$\begin{aligned} d\Omega &= d(d\omega + \omega \wedge \omega) \;=\; dd\omega + d(\omega \wedge \omega) \\ &= 0 + (d\omega) \wedge \omega - \omega \wedge d\omega \\ &= (\Omega - \omega \wedge \omega) \wedge \omega - \omega \wedge (\Omega - \omega \wedge \omega) \\ &= \Omega \wedge \omega - (\omega \wedge \omega) \wedge \omega - \omega \wedge \Omega + \omega \wedge (\omega \wedge \omega) \\ &= \Omega \wedge \omega - \omega \wedge \Omega \\ &= [\Omega, \omega] \end{aligned}$$

which is an important formula. Notice that there are no exterior derivatives in this formula; they have all cancelled out. To indicate what I mean more explicitly, ω is expressed in terms of $\Gamma^{\alpha}_{\beta i}$, Ω in terms of $R_{\alpha\;kl}^{\;\beta}$ and $d\Omega$ is a complicated linear combination of products of these terms with no additional derivatives.

Note $(\omega \wedge \omega) \wedge \omega = \omega \wedge (\omega \wedge \omega)$ by the associativity of matrix multiplication even though the terms are being multiplied by wedge.

Now let us generalize to a tensor T of type Ad whose expression in matrix form is an $n \times n$ matrix of k-forms ${t_{\sigma\beta}}^{\alpha}$, where the σ subscript indicates the basis for which ${t_{\sigma\beta}}^{\alpha}$ is the matrix expression. Thus $t_\sigma \in M(n, \Lambda^k(T^*M))$.

We now assume that T is a tensor of type Ad. This means that for $\sigma' = \sigma g$

$$t_{\sigma'} = \text{Ad}(g^{-1})t_\sigma = g^{-1}t_\sigma g$$

We are now going to duplicate the calculation we made with Ω which will work since both ω and T are tensor of type Ad. We begin as before:

$$g t_{\sigma'} = t_\sigma g$$

and take the exteror derivative of both sides to get

$$\begin{aligned} d(g t_{\sigma'}) &= d(t_\sigma g) \\ dg \wedge t_{\sigma'} + g\, dt_{\sigma'} &= (dt_\sigma)\, g + (-1)^k t_\sigma \wedge dg \end{aligned}$$

using the usual rule $d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^{\deg\,\omega} \omega \wedge \eta$. Recalling that $\omega' = g^{-1}\,dg + g^{-1}\omega g$ we see that $dg = g\omega' - \omega\,dg$ and then we have from the last equation

$$\begin{aligned} (g\omega' - \omega\, g) \wedge t_{\sigma'} + g\, dt_{\sigma'} &= (dt_\sigma)\, g + (-1)^k t_\sigma \wedge (g\omega' - \omega\, dg) \\ g(dt_{\sigma'} + \omega' \wedge t_{\sigma'}) - \omega g \wedge t_{\sigma'} &= (dt_\sigma + (-1)^{k+1} t_\sigma \wedge \omega) g + (-1)^k t_\sigma \wedge g\omega' \end{aligned}$$

Now, using the fact that T is of type Ad, so $g t_{\sigma'} = t_\sigma g$,

$$g(dt_{\sigma'} + \omega' \wedge t_{\sigma'}) - \omega g \wedge g^{-1} t_\sigma g = (dt_\sigma + (-1)^{k+1} t_\sigma \wedge \omega) g + (-1)^k g t_{\sigma'} g^{-1} \wedge g\omega'$$

There is cancellation of g and g^{-1} in the second term on each side. We now move each of the simplified second terms to the opposite side of the equation to get

$$\begin{aligned} g(dt_{\sigma'} + \omega' \wedge t_{\sigma'}) - (-1)^k g t_{\sigma'} \wedge \omega &= (dt_\sigma + (-1)^{k+1} t_\sigma \wedge \omega) g + \omega \wedge t_\sigma g \\ g(dt_{\sigma'} + \omega' \wedge t_{\sigma'} - (-1)^k t_{\sigma'} \wedge \omega) &= (dt_\sigma + (-1)^{k+1} t_\sigma \wedge \omega + \omega \wedge t_\sigma) g \\ dt_{\sigma'} + \omega' \wedge t_{\sigma'} - (-1)^k t_{\sigma'} \wedge \omega &= g^{-1}(dt_\sigma - (-1)^k t_\sigma \wedge \omega + \omega \wedge t_\sigma) g \end{aligned}$$

which shows that

$$dt_\sigma - (-1)^k t_\sigma \wedge \omega + \omega \wedge t_\sigma \quad \text{is a tensor of type Ad}$$

In fact it is $\mathcal{D} t_\sigma$ as can easily be shown.

As a first application, Ω_σ is a tensor of type Ad. Hence

$$\begin{aligned} \mathcal{D}\Omega_\sigma &= d\Omega_\sigma + \omega_\sigma \wedge \Omega_\sigma - (-1)^2 \Omega_\sigma \wedge \omega_\sigma \\ &= d\Omega_\sigma + [\omega_\sigma,\, \Omega_\sigma] \end{aligned}$$

But earlier we showed that

$$d\Omega_\sigma = [\Omega_\sigma, \omega_\sigma] = -[\omega_\sigma,\, \Omega_\sigma]$$

so we must have

$$\mathcal{D}\Omega_\sigma = 0$$

and hence

$$D\Omega = \sigma \mathcal{D}\Omega_\sigma = 0$$

Now we want a general formula for $D^2 T$ where T is a tensor of type Ad represented by a matrix of k-forms as before. Since the formulas involve tensors we may as well use an invariant notation. We have

$$\begin{aligned} DT &= dt + \omega \wedge T - (-1)^k T \wedge \omega \\ D(DT) &= d(DT) + \omega \wedge DT - (-1)^{k-1} DT \wedge \omega \\ &= d(dt + \omega \wedge T - (-1)^k T \wedge \omega) + \omega \wedge (dt + \omega \wedge T - (-1)^k T \wedge \omega) \\ &\quad -(-1)^{k-1}(dt + \omega \wedge T - (-1)^k T \wedge \omega) \wedge \omega \\ &= 0 + d\omega \wedge T - \omega \wedge dT - (-1)^k dT \wedge \omega - (-1)^{2k} T \wedge d\omega \\ &\quad +\omega \wedge dT + \omega \wedge \omega \wedge T - (-1)^k \omega \wedge T \wedge \omega \\ &\quad -(-1)^{k+1} dT \wedge \omega - (-1)^{k+1} \omega \wedge T \wedge \omega + (-1^{k+1+k} T \wedge \omega \wedge \omega \end{aligned}$$

This gives us 10 terms (including the 0). Moving terms 8 and 9 into positions after term 3 and 7 and simplifying the (-1) factors gives us

$$\begin{aligned} D^2T &= d\omega \wedge T - \omega \wedge dT - (-1)^k dT \wedge \omega + (-1)^k dT \wedge \omega \\ &\quad +\omega \wedge dT + \omega \wedge \omega \wedge T - (-1)^k \omega \wedge T \wedge \omega + (-1)^k \omega \wedge T \wedge \omega \\ &\quad -T \wedge \omega \wedge \omega \end{aligned}$$

Removing the cancelling terms we have

$$\begin{aligned} D^2T &= (d\omega + \omega \wedge \omega) \wedge T - T \wedge (d\omega + \omega \wedge \omega) \\ &= \Omega \wedge T - T \wedge \Omega \\ &= [\Omega, T] \end{aligned}$$

This shows how intricately the curvature Ω is intertwined with the second covariant derivative of tensors of type Ad.

Appendix A notationial variant.

The formula

$$\Omega = d\omega + \omega \wedge \omega$$

has a variant often found in more advanced books. The idea is that since ω is a Lie Algebra valued one form, it would make sense to express it in a form using Lie Algebra symbols, namely the bracket $[\ \ ,\ \ ]$ which is the basic Lie Algebra notation for product. Not all Lie Algebras consist of matrices, but if one does then the Lie Algebra product $[A, B]$ where $A, B \in \mathfrak{g}$ is

$$[A, B] = AB - BA$$

In matrix land this is called the anticommutator, but if course the above formula will not work for Lie Algebras that do not consist of matrices. Fortunately for our purposes the Lie Algebras *do* consist of matrices.

Now if ω and η are two Lie Algebra valued one forms, it is natural to define

$$[\omega, \eta] = \omega \wedge \eta - \eta \wedge \omega$$

as we did for matrices. (Note ω, η become Lie Algebra elements, matrices, when vectors are inserted as their arguments and the wedges become matrix multiplication.) In our case ω, η are Lie Algebra valued one forms and thus we have

$$[\omega, \eta] = \omega \wedge \eta - (-1)^{1\cdot 1} \omega \wedge \eta$$

so

$$[\omega, \eta] = \omega \wedge \eta + \omega \wedge \eta = 2\, \omega \wedge \eta$$

Hence, letting $\eta = \omega$ we have

$$[\omega,\, \omega] = 2\, \omega \wedge \omega$$

Hence we see

$$\omega \wedge \omega = \frac{1}{2}\,[\omega,\,\omega]$$

and finally the sought for revision

$$\Omega = d\omega + \frac{1}{2}\,[\omega,\,\omega]$$

12 ABSTRACT THEORY OF CONNECTIONS

Up to this point, we have mainly concentrated on connections on a principal bundle that derived from connections on a vector bundle. It would be more aesthetically pleasing to have a definition that could be applied when the given data is a principal bundle, without any mention of a vector bundle from which it comes. So the question is, what are the conditions on on a Lie Algebra valued differential form that would allow us to use it as a connection on the given principal bundle?

The key idea here is visible right from the beginning of our treatment in vector bundles. A connection there is, in a basis, related to a system of coefficients, the Christoffel symbols, that we called $\Gamma^{\alpha}_{\beta i}$ *that were functions on* U*,* the open set of the coordinate neighborhood. These coefficients did not depend on the section. For example the directional derivative of the section ρ in the direction $X \in T_p(U)$ is given by

$$D_X(\rho) = \sigma_\beta \rho^{\beta}{}_{|j} X^j$$

where

$$\rho^{\beta}{}_{|j} = \frac{\partial \rho^\beta}{\partial u^j} + \Gamma^{\beta}_{\alpha j}\rho^\alpha$$

and $\sigma_1, \ldots, \sigma_n$ is a local basis of sections. (Naturally the coefficients $\Gamma^{\beta}_{\alpha j}$ depend on the choice of local basis.) The important thing to get here is that the $\Gamma^{\beta}_{\alpha j}$ depend only on the coordinate $u^1, \ldots, u^d$ on U. *They do not depend on the section* ρ*.* The same thing is true for a connection in a principal bundle.

What we need to do, then, is to find a way to say this in the abstract case of a given principal bundle and a given Lie Algebra valued differential form, and to say it in a way that is consistent with the given data.

We recall that a connection can be specified in four formally different ways:

1. Through the differential forms ${\omega_\beta}^\alpha = \Gamma^{\alpha}_{\beta i} du^i$

2. Through the Horizontal Space $H_p(P)$

3. Through the projection $\Pi_p : T_p(P) \to V_p(P)$

4. Through the $\mathcal{Z}_p : T_p(P) \to \mathfrak{g} = \mathrm{Lie}(G)$

Recall also $H_p(P) = \ker(\Pi_p)$ and $V_p(P) = \mathrm{im}(\Pi_p)$. Ordinary linear algebra then tells us that

$$T_p(P) = H_p(P) \oplus V_p(P)$$

where the sum is direct but not orthogonal. In general there is no inner product given so orthogonal makes no sense. Thus we see

$$\begin{aligned} u \in H_p(P) \Longleftrightarrow \Pi_p(u) &= 0 \\ u \in V_p(P) \Longleftrightarrow \Pi_p(u) &= u \end{aligned}$$